N^o d'ordre : 7453

Université Paris XI
UFR Scientifique d'Orsay

Thèse

présentée
pour obtenir

le grade de DOCTEUR EN SCIENCES
de l'université Paris XI Orsay

par

Emmanuel Breuillard

Sujet : Marches aléatoires, equirépartition, et sous-groupes denses dans les groupes de Lie.

Soutenue le 19 décembre 2003 devant la commission d'examen :

M. Benoist Yves, Université Paris VII (rapporteur)
M. Bougerol Philippe, Université Paris VI (rapporteur)
M. Guivarc'h Yves, Université de Rennes 1 (rapporteur)
M. Ledrappier François, University of Notre Dame (président du jury)
M. Margulis Gregory, Yale University (directeur de thèse)
M. Paulin Frédéric, Université Paris XI (directeur de thèse)

À la mémoire de Martine Babillot,

Remerciements

Je tiens d'abord à rendre hommage à Martine Babillot dont nous avons appris avec douleur la mort prématurée l'été dernier. Martine Babillot a été, à maints égards, l'initiatrice de mon travail de doctorat. C'est elle qui, à travers son cours de troisième cycle donné à Paris VI en 1999, m'a introduit à la théorie ergodique et ses diverses ramifications, puis a stimulé mon intérêt pour la théorie des probabilités sur les groupes à laquelle est consacrée la première partie de cette thèse. C'est elle aussi qui m'a suggéré la possibilité d'aller étudier à l'étranger et m'a encouragé dans ce choix qui n'allait pas de soi, mais qui s'est révélé être une grande chance. C'est elle enfin qui, par le soutien constant qu'elle m'a apporté tout au long de ces dernières années, par la patience qu'elle a eue pour relire mes brouillons et l'enthousiasme inaltérable dont elle faisait preuve à l'égard de tant de sujets variés en mathématiques, a fait en sorte que mes années de thèse auront été l'expérience la plus enrichissante que j'ai vécue durant toutes mes études. C'est avec une profonde tristesse que je lui adresse ici ces modestes remerciements posthumes.

Mon directeur de thèse à Yale, Gregory A. Margulis, a suivi mon travail de près pendant ces trois années en m'apportant régulièrement des remarques précises et perspicaces. Il a grandement contribué a m'apporter la motivation et l'optimisme nécessaires pour mener à bien cette recherche. Grâce à ses "graduate courses" que j'ai suivis à Yale et de par son attitude générale à l'égard des mathématiques, il m'a apporté une expérience qui a joué et jouera certainement aussi dans l'avenir un très grand rôle pour moi.

Je remercie mon directeur de thèse à Paris, Frédéric Paulin, dont l'aide m'a été très précieuse et grâce à qui j'ai été chaleureusement accueilli dans le laboratoire du DMA de l'Ecole Normale Supérieure et à l'université Paris-Sud d'Orsay où j'ai pu passer le deuxième semestre 2003 dans le cadre de la cotutelle de thèse entre Yale et Orsay, ainsi que de nombreux autres séjours de plus courte durée. Je lui suis aussi reconnaissant pour tout le temps qu'il m'a consacré et pour sa relecture attentive de mon manuscrit.

Tsachik Gelander est le co-auteur de la seconde partie de cette thèse et je tiens à lui exprimer ici ma profonde reconnaissance pour cette collaboration qui a été si fructueuse pour moi. Nos échanges constants depuis plus de deux ans ont été extrêmement stimulants intellectuellement. Cela a été une grande chance, et toujours une grande joie pour moi, de pouvoir réfléchir et échanger des idées avec Tsachik.

Je voudrais aussi exprimer ma profonde gratitude envers tous mes autres professeurs, en particulier J.F. Le Gall, mon tuteur à l'ENS pendant quatre ans, J. Bertoin pour ses cours de probabilités en maîtrise et en DEA à Paris et A. Lubotzky de qui j'ai beaucoup appris à Yale. Je suis aussi tout particulièrement reconnaissant à Hassan Tarfaoui, mon professeur de Terminale, qui a joué un grand rôle dans mon orientation et mon engagement dans les mathématiques.

Je remercie également tous les mathématiciens avec lesquels j'ai eu des conversations enrichissantes au sujet de mon travail, notamment Y. Guivarc'h qui m'a consacré beaucoup de temps, H. Abels, G. Alexopoulos, Y. Benoist, E. Ghys, L. Pyber et G. Soifer qui ont stimulé ma recherche de diverses façons.

Merci aussi à tous ceux qui m'ont tant appris au gré des conversations et des séminaires à Paris, Cambridge et Yale, Azer Akhmedov, Nalini Anantharaman, Uri Bader, David Bessis, Sacha Boufetov, Eliott Brenner, Gaëtan Chenevier, Yves Coudène, David Fisher, Peter Friz, Grigor Grigorov, Olivier Guichard, Bruno Klingler, Seonhee Lim, Keivan Mallahi-Karai, Roman Muchnik, Yann Ollivier, Jean-François Quint, Bertrand Rémy, Alireza Salehi-Golsefidy, Hadi Salmasian, Barbara Schapira, Pete Storm, Andrzej Zuk, et beaucoup d'autres encore que j'oublie sûrement. Un grand merci aussi à Paul Lukasiewicz, bibliothécaire du département de mathématiques de Yale, dont l'aide et le dévouement m'ont beaucoup touché.

Last but not least, je remercie toute ma famille et bien sûr Léa pour leur patience et leur soutien.

Avant-propos

Ce mémoire rassemble plusieurs travaux réalisés dans le cadre du doctorat de mathématiques sous la tutelle de deux universités : l'université de Yale à New Haven (USA) sous la direction de G.A. Margulis d'une part et l'université Paris XI à Orsay et le laboratoire de l'Ecole Normale Supérieure DMA à Paris (France) sous la direction de F. Paulin d'autre part. Il est écrit en partie en anglais et en partie en français conformément à la convention de cotutelle du 19 juin 2001 signée par les deux universités.

Il comporte deux parties relativement distinctes. Chaque partie débute par un chapitre introductif qui présente les problématiques développées dans les chapitres suivants, fait le point sur les résultats connus auparavant, et énonce les principaux théorèmes de la thèse en donnant parfois soit des indications sur les méthodes utilisées soit la démonstration d'un cas particulier. Les chapitres introductifs 1 et 5 exceptés, les chapitres peuvent être lus de manière indépendante. Les chapitres 3, 4 et 7 ont été chacun soumis à publication et le chapitre 6 a déjà fait l'objet d'une publication dans le *Journal of Algebra* de mars 2003.

La partie I traite des marches aléatoires et de la théorie des probabilités sur les groupes. Les résultats proposés généralisent certains théorèmes classiques du calcul des probabilités pour les sommes de variables aléatoires au cas non-commutatif des produits de matrices aléatoires, notamment le théorème limite local et les théorèmes d'équirépartition de marches aléatoires. Le chapitre 2 présente un travail en cours sur l'équirépartition des marches symétriques sur les groupes de Lie nilpotents et les marches unipotentes sur les espaces homogènes. Le chapitre 3 reprend un article où l'on démontre un théorème limite local sur le groupe de Heisenberg pour les distributions centrées à support compact sans hypothèse de densité, ainsi qu'un théorème de comparaison entre la marche aléatoire et le noyau de la chaleur associé. Cela permet en outre d'obtenir une version probabiliste du théorème d'équirépartition de Ratner pour les flots unipotents sur les espaces homogènes. Au chapitre 4, on introduit la notion de distribution diophantienne sur $\mathbb{R}^d$ et on montre que cette propriété gouverne la vitesse de convergence dans le théorème local classique. On montre aussi un théorème de comparaison pour les écarts modérés ainsi qu'un principe d'invariance pour les fonctions bornées (voir ci-dessous).

La partie II rassemble des articles qui sont le fruit d'une recherche en commun avec Tsachik Gelander (Université Hébraïque de Jérusalem). Le thème général de ce travail est la construction de groupes libres dont le plongement dans les groupes de Lie ou les groupes algébriques satisfont certaines contraintes. Ces problèmes et leurs méthodes se situent dans la lignée de la célèbre alternative de Tits sur les groupes linéaires. Notre théorème principal énonce que tout sous-groupe dense d'un groupe de Lie non résoluble contient un sous-groupe libre et dense. Au chapitre 6, on démontre ce fait pour les groupes de Lie réels connexes et on indique au passage une méthode simple pour construire des sous-groupes denses dans les groupes de Lie. Au chapitre 7, on généralise ce théorème au cas non connexe et p-adique et on présente certaines applications à des domaines variés comme la théorie des actions moyennables ou bien l'étude des groupes profinis. Enfin dans l'appendice à cette partie, on donne l'esquisse d'un travail en cours sur une question proche : la détermination d'une version effective de l'alternative de Tits.

Pour conclure cet avant-propos, nous résumons ci-dessous quelques-uns des énoncés, parmi les plus significatifs, qu'on démontrera dans cette thèse.

Partie I

Théorème (3.1.1 et 3.1.2 Théorème limite local sur le groupe de Heisenberg) : *Soit N le groupe de Heisenberg des matrices triangulaires supérieures 3×3. Soit μ une probabilité centrée, apériodique et à support compact sur N. Soit $(\nu_t)_t$ un semi-groupe gaussien associé à μ. Alors pour toute fonction f continue à support compact sur N, on a*

$$\lim_{n \to +\infty} n^2 \int_N f(x) d\mu^n(x) = c(\mu) \int_N f(x) dx$$

où $c(\mu) > 0$ est la valeur en e de la densité p_1 correspondant à ν_1 (i.e. le noyau de la chaleur associé à μ). De plus pour tout borélien borné B de N dont la frontière est négligeable par rapport à la mesure de Lebesgue (i.e. $|\partial B| = 0$), on a

$$\lim_{n \to +\infty} n^2 \sup_{x \in N} |\mu^n(xB) - \nu_n(xB)| = 0$$

Ce théorème généralise le théorème limite local classique sur $\mathbb{R}^d$ et sa version uniforme due à Stone [161] au groupe de Heisenberg. La méthode utilise les représentations unitaires de N ainsi que le théorème limite central sur N, ce qui permet de s'affranchir de l'hypothèse d'existence d'une densité continue pour la probabilité μ, hypothèse sous-laquelle ce type de théorème avait été démontré par le passé (voir Alexopoulos [5] [6] [7]).

Théorème (2.0.4 Version probabiliste du théorème d'équirépartition de Ratner) *: Soit G un groupe de Lie réel connexe et Γ un réseau de G. Soit H un sous-groupe simplement connexe unipotent de G. Soit μ une mesure de probabilité symétrique et à support fini sur H, et $(S_n)_{n \geq 0}$ la marche aléatoire sur H associée. Alors pour tout x* dans G/Γ *et toute fonction f continue et bornée sur G/Γ*

$$\lim_{n \to +\infty} \mathbb{E}(f(S_n \cdot x)) = \int_{G/\Gamma} f dm_x$$

où, pour tout point x de G/Γ, *on note* m_x *la probabilité ergodique* H*-invariante dont le support est* $\overline{H \cdot x}$ *fournie par le théorème de Ratner.*

Ce théorème répond à une question de G.A. Margulis et complète l'étude des marches aléatoires sur G/Γ développée dans [54].

Théorème (4.3.2 Théorème local pour les écarts modérés) : *Soit* μ *une probabilité centrée et apériodique sur* $\mathbb{R}$ *admettant un moment fini d'ordre* $r > 2$ *et* ν *la loi gaussienne associée. Soit* σ^2 *la variance de* μ, $I = [a, b]$ *un intervalle fermé de* $\mathbb{R}$ *et* c *un réel dans* $]0, r-2[$. *Alors on a*

$$\lim_{n \to +\infty} \frac{\mu^n(I + x)}{\nu^n(I + x)} = 1$$

uniformément quand $|x| \leq \sqrt{c\sigma^2 n \log n}$.

Théorème (4.4.5 Principe d'invariance pour les fonctions bornées) : *Soit* μ *une probabilité centrée sur* $\mathbb{R}$ *admettant un moment d'ordre 4 et soit* ν *la loi gaussienne associée. On suppose que* μ *est diophantienne. Alors il existe* $k_0 \geq 0$ *tel que*

$$\lim_{n \to +\infty} \frac{\int f d\mu^n}{\int f d\nu^n} = 1$$

pour toute fonction C^k , *avec* $k \geq k_0$, *bornée non nulle* $f \geq 0$ *sur* $\mathbb{R}$ *dont toutes les dérivées jusqu'à l'ordre* k *sont bornées.*

Au chapitres 1 et 4, on montre comment ce théorème peut être appliqué à l'équidistribution des marches aléatoires sur les espaces homogènes.

Partie II

Théorème (7.1.2 Version topologique de l'alternative de Tits) *Soit* Γ *un groupe de type fini et* $\sigma : \Gamma \hookrightarrow GL_n(k)$ *un homomorphisme injectif de* Γ *dans* $GL_n(k)$ *où* k *est un corps local. On munit* Γ *de la topologie la moins fine pour que* σ *soit continue. Alors soit* Γ *contient un sous-groupe ouvert résoluble, soit* Γ *contient un sous-groupe dense qui est un groupe libre de type fini.*

En particulier, ce théorème montre que tout sous-groupe dense d'un groupe de Lie connexe non résoluble contient un sous-groupe libre et dense.

Terminons par trois corollaires de ce théorème :

Théorème (7.8.3 Conjecture de Connes-Sullivan généralisée) *Soit* G *un groupe localement compact et* Γ *un sous-groupe dénombrable de* G. *Alors* Γ *agit moyennablement sur* G *par translations à gauche si et seulement s'il existe un sous-groupe* Γ_0 *de* Γ *qui* *est moyennable en tant que groupe abstrait et qui est relativement ouvert dans* Γ *pour la topologie de* Γ *induite par celle de* G.

Théorème (7.1.6) *Dans un feuilletage Riemannien sur une variété compacte, la croissance des feuilles est bornée uniformément par une fonction polynômiale, ou bien le revêtement d'holonomie de chaque feuille est à croissance exponentielle.*

Théorème (7.7.1 Conjecture de Dixon-Pyber-Seress-Shalev) *Soit* Γ *un groupe linéaire de type fini et* $\widehat{\Gamma}$ *sa complétion profinie. Alors soit* Γ *contient un sous-groupe résoluble d'indice fini, soit* $\widehat{\Gamma}$ *contient un sous-groupe libre et dense de type fini.*

Paris, le 6 octobre 2003

Table des matières

Première partie

Marches aléatoires sur les groupes de Lie

Chapitre 1

Introduction

Dans ce premier chapitre introductif, nous allons présenter les problématiques qui font l'objet de ce travail en commençant par les resituer dans le cadre de la théorie des probabilités sur les groupes. Les premières sections sont consacrées à la description succincte de résultats connus concernant les marches aléatoires. Par marche aléatoire sur un groupe G, nous entendrons le processus aléatoire $S_n = X_1 \cdot \ldots \cdot X_n$ donné par le produit de n variables aléatoires indépendantes et de même loi à valeurs dans un groupe G que l'on s'est fixé au départ. De façon similaire, on parle de marche aléatoire sur un espace X, muni d'un groupe de transformations G, si à chaque pas on applique une transformation aléatoire choisie dans G selon une même loi de probabilité et de façon indépendante. Dans les premières sections, la plupart des énoncés généralisent les théorèmes classiques du calcul des probabilités (la convergence en loi, la loi des grands nombres, le théorème limite central[1], son équivalent local, les principes de grandes déviations, etc) à un contexte non-commutatif. Ensuite, on annonce les principaux résultats démontrés aux chapitres suivants en en traitant parfois certains cas particuliers significatifs. Ce chapitre a notamment pour objectif de préparer la suite en indiquant les préliminaires nécessaires.

L'accent sera mis sur les problèmes d'équirépartition (on dit aussi équidistribution) des marches aléatoires dans les groupes de Lie (théorèmes locaux et théorèmes quotient en particulier). Il s'agit de comprendre comment une marche aléatoire, après un temps très grand, finit par occuper tout l'espace de façon plus ou moins homogène, et en particulier d'estimer la vitesse à laquelle ce phénomène se produit. Ces problèmes ont déjà fait l'objet de très nombreuses études par le passé. Nous renvoyons le lecteur aux exposés [78] et [77] pour une introduction historique et bibliographique ainsi qu'à la monographie [68], à ma connaissance le premier ouvrage entièrement consacré à la théorie des probabilités sur les groupes.

Comme nous le verrons, les propriétés arithmétiques des distributions considérées jouent un rôle important. Dans la grande majorité des situations (le cas abélien ex-

[1]Nous adopterons le point de vue de G. Pólya, père de la terminologie “théorème limite central”, selon lequel c'est le théorème qui est central (i.e. fondamental dans la théorie des probabilités) et non la limite.

cepté), les théorèmes d'équirépartition connus ont été obtenus sous l'hypothèse que la distribution initiale possède une densité par rapport à la mesure de Lebesgue ou la mesure de Haar du groupe de Lie ambiant. En l'absence de cette hypothèse, par exemple lorsque le support de la distribution est donné par un nombre fini d'éléments possibles qui engendrent ensemble un sous-groupe dense, les questions arithmétiques entrent en jeu de façon cruciale. C'est le cas par exemple pour les marches aléatoires sur la sphère où, à chaque pas, on applique une isométrie aléatoire : comprendre la vitesse d'équidistribution des groupes libres et denses dans $SO(3)$ reste un important problème ouvert (voir plus bas §1.2.2). Il en est de même pour le théorème limite local sur le groupe des déplacements de $\mathbb{R}^d$: au paragraphe 1.2.3 on en donne une preuve pour les déplacements du plan inspirée du théorème quotient de Kazhdan ([96], [74]), le cas $d \geq 3$ restant largement ouvert. Au chapitre 4 aussi est introduite la notion de distribution diophantienne sur la droite réelle : on montre que cette propriété gouverne la vitesse d'équidistribution des marches centrées sur $\mathbb{R}$.

Les études faites aux chapitres 2 et 3 concernent le cas particulier du théorème limite local pour les groupes de Lie nilpotents. Il existe un vaste ensemble de travaux de nombreux mathématiciens à propos des marches aléatoires sur les groupes de Lie nilpotents, mais le théorème local en lui-même a été peu étudié (voir cependant [105]) jusqu'à l'imposant récent travail d'Alexopoulos ([5], [6] et [7]) qui traite le cas des mesures centrées possédant une densité. Plus bas dans ce chapitre, on décrira quelques-uns des résultats qui nous serons utiles par la suite, et on introduira quelques notions de base sur les groupes nilpotents et les marches aléatoires sur ceux-ci. Nous présenterons deux approches différentes pour traiter le problème du théorème local. La première, développée au chapitre suivant, consiste à tirer parti des résultats déjà connus dans le cadre des marches aléatoires symétriques sur les groupes nilpotents de type fini (voir [171], [175], [85], [6]) et de les combiner à un théorème d'équidistribution à la Weyl sur les groupes de Lie nilpotents. La seconde, expliquée dans le chapitre 3, met à profit la théorie des représentations unitaires irréductibles des groupes de Lie nilpotents (voir [99]) et suit un schéma semblable à la preuve du théorème local sur la droite réelle due à Stone (voir [30]). Cependant, pour mener à bien cette stratégie simple, un important travail analytique est nécessaire et nous a forcé à nous restreindre au cas du groupe de Heisenberg[2]. Néanmoins, cette méthode présente l'avantage de fournir des résultats très précis sur la distribution des marches aléatoires (voir le théorème 3.23) ainsi que de permettre d'obtenir le théorème local et sa version uniforme pour toutes les distributions apériodiques, centrées et à support compact sans hypothèse de densité, ce qui constitue quasiment autant d'information que ce dont on dispose dans le cas classique où $G = \mathbb{R}^d$.

Ces études ont été motivées en partie par un travail récent d'Eskin et Margulis [54] qui met en lumière une propriété de récurrence des marches aléatoires dans les espaces homogènes de volume fini G/Γ. En guise d'application, ils obtiennent en [113]

[2]La démonstration du chapitre 3 est écrite pour le groupe de Heisenberg, mais elle est valide telle quelle pour tout groupe nilpotent de rang 2.

une nouvelle preuve du théorème de Borel et Harish-Chandra sur la finitude du volume des quotients arithmétiques des groupes algébriques semi-simples et ils complètent la preuve de la conjecture de Raghunathan sur la classification des mesures ergodiques sous l'action d'un sous-groupe quelconque engendré par des unipotents (voir [54] et [160] Theorem 19.14). Les marches aléatoires considérées dans [54] sont algébriquement denses dans un groupe semi-simple. A l'inverse, nous nous intéresserons aux marches aléatoires qui évoluent dans un sous-groupe unipotent de G. Nous disposons alors de la théorie très riche de la dynamique des sous-groupes unipotents dans les espaces homogènes G/Γ et en particulier du théorème de Ratner (voir §1.5.1). Ce que nous proposons au chapitre 3 est un analogue probabiliste du théorème d'équirépartition de Ratner-Shah (voir aussi le paragraphe §1.5.2 de ce chapitre pour le cas des flots unipotents) pour les marches aléatoires symétriques et à support fini dans un sous-groupe unipotent quelconque de G. Nous démontrons qu'il y a récurrence comme dans [54] et même équirépartition. Un phénomène intéressant apparaît lorsque l'on considère une marche décentrée le long d'un flot unipotent. Lorsque l'espace G/Γ n'est pas compact alors une marche décentrée peut se retrouver dans un voisinage de l'infini avec grande probabilité et ce à des temps arbitrairement grands. Ce phénomène dépend des propriétés diophantiennes du point de départ de la marche aléatoire (voir ci-dessous le §1.5.4). Les théorèmes locaux établis aux chapitres 2 et 3, ou plus exactement leur version fine qui donnent un contrôle uniforme sur les translatés d'un compact dans une boule de rayon $\sqrt{n}$, permettent de passer de l'équirépartition déterministe à l'équirépartition probabiliste.

Au chapitre 4, on effectue une étude fine de l'équirépartition des marches aléatoires centrées sur $\mathbb{R}^d$. Dans ce cas, déjà très étudié par les probabilistes dans le cadre de la théorie du renouvellement, on étudie la vitesse de convergence dans le théorème local et on montre qu'elle dépend des propriétés diophantiennes de la distribution qui gouverne la marche. On obtient aussi un théorème local pour les écarts modérés sous une condition faible de moment : il stipule que le théorème local habituel qui fournit l'estimation de $\mathbb{P}(S_n \in I + x)$, où I est un intervalle de $\mathbb{R}$ et S_n une marche centrée, est uniforme en x sur toute la plage $[-c\sqrt{n \log n}, c\sqrt{n \log n}]$ dès que la distribution possède un moment d'ordre $> c^2 + 2$. Enfin, on s'intéressera dans ce chapitre (voir aussi plus bas le §1.4.3) à l'asymptotique des quantités $\mathbb{E}(f(S_n))$ lorsque f est une fonction continue bornée sur $\mathbb{R}$ ou $\mathbb{R}^d$ sans hypothèse sur son comportement à l'infini. Il s'avère que si f est assez régulière ou bien si μ a de bonnes propriétés arithmétiques, alors cet asymptotique est indépendant de la marche centrée de variance 1 choisie. Un tel principe d'invariance est la clé derrière les théorèmes d'équirépartition déjà mentionnés pour les flots sur les espaces homogènes.

1.1 Définitions et notions de base

Soit G un groupe localement compact à base dénombrable engendré par un voisinage compact de l'identité e. On considère G en tant qu'espace mesurable pour la tribu des boréliens. On munit G d'une mesure de Haar invariante à gauche dg. On notera μ une mesure de probabilité sur G, c'est-à-dire une mesure positive σ-additive sur les boréliens de G telle que $\mu(G) = 1$. Le support de μ est le plus petit fermé de G de μ-mesure 1. Commençons par définir quelques termes de base :

Définition 1.1 *On dit que la mesure de probabilité μ sur G est **adaptée** si son support engendre un sous-groupe dense dans G.*

Définition 1.2 *On dit que μ est **apériodique** si son support n'est pas contenu dans une classe (à gauche ou à droite) d'un sous-groupe fermé propre de G.*

Soit U un voisinage compact de l'identité engendrant G. On pose

$$\delta_U(g) = \inf\{n \geq 0, g \in U^n\}$$

La fonction δ_U est sous-additive et constitue une mesure de l'éloignement par rapport à e dans G.

Définition 1.3 *On dit que μ possède un **moment d'ordre** $\alpha > 0$ si*

$$\int \delta_U(g)^\alpha d\mu(g) < +\infty$$

On vérifie aisément que cette définition ne dépend pas du choix de U. L'existence d'un moment d'ordre 1 implique que les homomorphismes continus $G \to (\mathbb{R}, +)$ sont dans $\mathbb{L}^1(\mu)$.

Définition 1.4 *On dit que μ est **centrée** si μ admet un moment d'ordre 1 et si pour tout homomorphisme continu $\chi : G \to (\mathbb{R}, +)$ on a*

$$\int \chi(g) d\mu(g) = 0$$

Définition 1.5 *On note μ^{-1} la symétrique de μ, c'est-à-dire la probabilité définie par $\mu^{-1}(E) = \mu(E^{-1})$ pour tout borélien E de G. On dit que μ est **symétrique** si $\mu = \mu^{-1}$.*

Remarquons que les mesures symétriques possédant un moment d'ordre 1 sont centrées et que si μ est symétrique alors μ est apériodique si et seulement si $\mu * \mu$ est adaptée.

1.1.1 Marches aléatoires et récurrence

De façon analogue à la théorie classique des sommes de variables aléatoires réelles, on va s'intéresser au produit de variables aléatoires à valeurs dans le groupe G, en particulier au produit de variables aléatoires indépendantes et identiquement distribuées. On note $S_n = X_n \cdot ... \cdot X_1$ le n-ième terme de la marche, les variables X_i étant indépendantes et distribuées selon une unique mesure de probabilité μ sur G. On suppose en général que la mesure μ est *adaptée*. Dans le cas contraire, c'est que la marche vit en fait sur un sous-groupe fermé propre de G, et on peut se restreindre à celui-ci. La loi de S_n est tout simplement la n-ième puissance de convolution de μ, c'est-à-dire μ^{*n} (plus loin on notera par abus μ^n au lieu de μ^{*n}). Rappelons que la convolution de deux mesures sera notée $\mu * \nu$ et définie par la formule

$$\int f d(\mu * \nu) = \int \int f(gh) d\mu(g) d\nu(h)$$

où f est une fonction quelconque continue et bornée sur G. Soit X une variable aléatoire à valeurs dans G dont la loi de probabilité est μ. Si Y est une autre variable aléatoire de loi ν cette fois et indépendante de X, alors la loi du produit XY est précisément la mesure $\mu * \nu$.

Une des premières constations simples (voir [75] pour la preuve) que l'on peut faire sur le comportement de S_n est l'existence de la dichotomie suivante :

• ou bien, la suite S_n part à l'infini presque sûrement (c'est-à-dire pour tout compact K, $S_n \notin K$ après un certain temps), on parle alors de *transience*.

• ou bien, presque sûrement, la suite passe un temps infini dans chaque ouvert de G, on parle alors de *récurrence*.

On remarque aussi que la marche est transiente si et seulement si le potentiel $\sum_{n \geq 0} \mu^{*n}$ est une mesure de Radon sur G, i.e. finie sur les compacts. Dans le cas contraire, le potentiel d'un ouvert quelconque est infini.

1.1.2 Convergence en loi et représentations unitaires

Le théorème classique de Lévy qui affirme l'équivalence entre la convergence en loi d'une suite de probabilités et la convergence point par point des transformées de Fourier, ou fonctions caractéristiques, s'étend naturellement dans le cadre de l'analyse harmonique non-commutative, avec essentiellement le même énoncé, comme nous allons le voir plus bas.

Si G est un groupe localement compact, l'ensemble des classes d'isomorphismes de représentations unitaires irréductibles de G s'appelle le dual unitaire de G et est noté $\widehat{G}$. Rappelons qu'une représentation unitaire $\pi \in \widehat{G}$ est une action continue de G sur un espace de Hilbert $\mathcal{H}$ par automorphismes linéaires qui préservent le produit scalaire. Elle est dite irréductible s'il n'y a pas de sous-espace fermé invariant.

Si μ est une mesure de probabilité définie sur les boréliens de G, et π une représentation unitaire de G, on leur associe naturellement l'opérateur $\pi(\mu)$ défini pour $\xi \in \mathcal{H}$ par

$$\pi(\mu)\xi = \int_G \pi(g)\xi\mu(dg)$$

Par exemple, si $G = \mathbb{R}^d$, les représentations unitaires irréductibles sont de dimension 1 et sont données par les caractères $\pi_t(x) = e^{it \cdot x}$, où $t \in \widehat{G} \cong \mathbb{R}^d$ et $x \in G = \mathbb{R}^d$. Dans ce cas, $\pi_t(\mu)$ est simplement un nombre complexe de module inférieur ou égal à un : il s'agit de la transformée de Fourier classique, ou fonction caractéristique $\widehat{\mu}(t)$, de μ. De plus, chaque représentation unitaire π de G établit un morphisme d'algèbres de Banach entre les mesures complexes sur G et les opérateurs bornés de $\mathcal{H}$, l'espace de la représentation π. En d'autres termes,

$$\pi(\mu * \nu) = \pi(\mu)\pi(\nu)$$

Pour étudier le comportement de la marche aléatoire S_n, on est donc amené à étudier les puissances des opérateurs $\pi(\mu)$ pour $\pi \in \widehat{G}$.

Avec ces notations, la généralisation du théorème de Lévy s'énonce ainsi :

Théorème 1.6 *Soit G un groupe localement compact et $\widehat{G}$ son dual unitaire.*

• *Si μ et ν sont deux probabilités sur G telles que $\pi(\mu) = \pi(\nu)$ pour tout $\pi \in \widehat{G}$, alors $\mu = \nu$ (unicité de la transformée de Fourier).*

• *Si $(\mu_n)_n$ est une suite de probabilités sur G qui converge étroitement vers une autre probabilité ν (c'est-à-dire $\int f d\mu_n \to \int f d\nu$ pour toute fonction f continue et bornée sur G), alors $\pi(\mu_n)$ converge fortement vers $\pi(\nu)$ (c'est-à-dire $\pi(\mu_n)\xi \to \pi(\nu)\xi$ pour tout $\xi \in \mathcal{H}$).*

• *Si $\pi(\mu_n)$ converge faiblement vers $\pi(\nu)$ (c'est-à-dire $\langle \pi(\mu_n)\xi, \eta \rangle \to \langle \pi(\nu)\xi, \eta \rangle$ pour tout $\xi, \eta \in \mathcal{H}$), alors $(\mu_n)_n$ converge étroitement vers ν.*

La preuve de cet énoncé (voir [68]) est une généralisation naturelle de la preuve classique du théorème de Lévy, par approximation des fonctions continues sur les compacts par des polynômes trigonométriques (i.e. des combinaisons linéaires de coefficients matriciels $g \mapsto \langle \pi(g)\xi, \eta \rangle$).

1.1.3 Equirépartition

Nous définissons ici ce que nous entendons par équirépartition.

Définition 1.7 *Nous dirons que la marche aléatoire associée à la probabilité μ sur le groupe G est **équirépartie** s'il existe une mesure de Radon m sur G telle que*

$$\lim_{n \to +\infty} \frac{\int f d\mu^n}{\int g d\mu^n} = \frac{\int f dm}{\int g dm} \tag{1.1}$$

pour toutes fonctions continues et à support compact f et g sur G avec $g \geq 0$ non nulle.

Si la marche aléatoire est équirépartie, on dit aussi, de façon équivalente, que la mesure μ satisfait un *théorème quotient*. Lorsque l'on dispose d'une suite *explicite* $(a_n)_n$ de nombres réels positifs telle que

$$\lim_{n \to +\infty} a_n \int f d\mu^n = \int f dm$$

pour toute fonction f continue à support compact sur G, alors on dit que μ satisfait un *théorème limite local*[3].

C'est cette propriété particulière des marches aléatoires sur les groupes que nous allons étudier dans la suite dans certains cas particuliers.

1.1.4 Processus de diffusion et formule de Lévy-Khinchine-Hunt

Dans son article de 1956 [90], Hunt caractérise les semi-groupes continus de mesures de probabilité sur les groupes de Lie connexe. Comme dans le cas classique où $G = \mathbb{R}^d$ ces semi-groupes correspondent aux processus à accroissements indépendants et stationnaires (PAIS). Quand $G = \mathbb{R}^d$ la célèbre formule de Lévy-Khinchine caractérise leur loi de probabilité en donnant une formule explicite pour la fonction caractéristique (voir [30], [65]). Sur un groupe de Lie, la généralisation naturelle consiste à caractériser les PAIS par le générateur infinitésimal du semi-groupe de mesures associé agissant sur les fonctions C^2 sur G.

Soit G un groupe de Lie connexe. Par **semi-groupe continu** de mesures de probabilité sur G nous entendons une famille $(\mu_t)_{t>0}$ de probabilités sur G telle que :

(i) $\mu_t * \mu_s = \mu_{t+s}$ pour tous $s, t > 0$.

(ii) $\mu_t \Rightarrow \delta_e$ quand $t \to 0$ (i.e. $\int f d\mu_t \to f(e)$ pour toute fonction $f \in C_b(G)$).

On pose $T_t f(g) = \int f(gh) d\mu_t(h)$. Alors les (T_t) forment un semi-groupe d'opérateurs linéaires bornés sur l'espace $C_b(G)$ des fonctions continues bornées de G tel que $T_t f$ converge uniformément vers f pour tout $f \in C_b(G)$ quand $t \to 0$. De plus on définit le générateur infinitésimal L de T_t par la formule

$$Lf(g) = \frac{d}{dt}_{|t=0} T_t f(g)$$

Le théorème ci-dessous montre que le domaine de L contient $C_c^2(G)$, (l'espace des fonctions à support compact et deux fois différentiables).

Un **processus à accroissements indépendants et stationnaires** (PAIS) sur G est un processus stochastique $(X_t, \mathbb{P}, \Omega)$ tel que :

(i) $X_0 = e$.

(ii) pour tous temps s, t $0 < s < t < +\infty$ la loi de $X_s^{-1} X_t$ ne dépend que de $t - s$.

(iii) pour tous $t_1, ..., t_k$ tels que $0 < t_1 < ... < t_k$ les variables aléatoires $X_{t_i}^{-1} X_{t_{i+1}}$ sont indépendantes.

(iv) quand $t \to 0$, X_t converge en loi vers e.

Les PAIS sont en correspondance biunivoque avec les semi-groupes continus de mesures de probabilité sur G. Plus précisément, si $(X_t, \mathbb{P}, \Omega)$ est un PAIS, alors si l'on note μ_t la loi de X_t, les $(\mu_t)_t$ forment un semi-groupe continu de mesures de probabilité sur

[3] Par théorème limite central en revanche, on entend en général un théorème qui précise le comportement limite en loi d'une suite Y_n de variables aléatoires *renormalisées* par un coefficient de contraction qui tend vers l'infini avec n (ou une dilatation du groupe si Y_n prend ses valeurs dans G et si G possède de telles dilatations, voir les paragraphes suivants)

G, et réciproquement, si $(\mu_t)_t$ est un tel semi-groupe, alors on peut construire un espace probabilisé $(\mathbb{P}, \Omega)$ et un PAIS (X_t) défini sur $(\mathbb{P}, \Omega)$ tel que μ_t soit la loi de X_t. De plus, on peut toujours trouver une version càdlàg (continue à droite avec limite à gauche en tout point) de (X_t).

On fixe une base $X_1, ..., X_d$ de l'espace vectoriel de l'algèbre de Lie $\mathfrak{g}$ de G et un voisinage U_0 de l'identité dans G sur lequel le logarithme est un difféomorphisme bien défini. Ceci permet d'associer à tout élément $g \in G$ appartenant à U_0 des coordonnées $(x_i(g))_{i=1,\ldots,d}$ déterminées par l'équation

$$g = \exp(\sum x_i(g) X_i)$$

On fixe une fois pour toute une fonction ϕ sur G qui est identiquement égale à 1 en dehors d'un voisinage compact de l'identité et qui vaut $\phi(g) = \sum x_i(g)^2$ sur U_0. Les éléments de l'algèbre de Lie $\mathfrak{g}$ de G peuvent être vus comme appartenant à l'algèbre universelle enveloppante de $\mathfrak{g}$ qui s'identifie à l'algèbre des opérateurs différentiels invariants à gauche sur G. En particulier, pour une fonction différentiable f sur G et pour $X \in \mathfrak{g}$, on note

$$Xf(g) = \frac{d}{dt}_{|t=0} f(ge^{tX})$$

On a alors :

Théorème 1.8 *(Hunt) Soit $(\mu_t)_{t>0}$ un semi-groupe continu de mesures de probabilité sur G et (T_t) le semi-groupe d'opérateurs associé sur $C_b(G)$. Alors le générateur infinitésimal L de (T_t) est bien défini pour toute fonction f de classe C^2 et à support compact sur G et admet la forme suivante :*

$$Lf(g) = \sum_i b_i X_i f(g) + \frac{1}{2} \sum_{i,j} a_{ij} X_i X_j f(g) + \int_G \left\{ f(gh) - f(g) - \sum_i X_i f(g) x_i(h) \right\} \frac{d\eta(h)}{\phi(h)} \tag{1.2}$$

où les (b_i) (a_{ij}) sont des nombres réels, la matrice (a_{ij}) est symétrique semi-definie positive et $d\eta$ est une mesure positive finie sur G telle que $\eta(\{e\}) = 0$. De plus, l'opérateur du second ordre $\sum_{i,j} a_{ij} X_i X_j$ et la mesure $\frac{d\eta(g)}{\phi(g)}$ sont indépendants du choix de la base (X_i) et de la fonction $\phi(g)$. Enfin, le semi-groupe $(\mu_t)_t$ est déterminé de façon unique par l'opérateur L restreint à $C_c^2(G)$.

Réciproquement, tout opérateur L défini sur $C_c^2(G)$ par la formule (1.2) est le générateur infinitésimal d'un semi-groupe continu de mesures de probabilité.

Lorsque la mesure $\frac{\eta}{\phi}$ (appelée mesure de Lévy) est nulle, on dit que le semi-groupe $(\mu_t)_t$ est **gaussien** (de façon équivalente $(\mu_t)_t$ est gaussien si et seulement si pour tout voisinage U de l'identité $\mathbb{P}(X_t \notin U) = o(t)$). Dans ce cas et seulement dans ce cas, le processus associé (X_t) est à trajectoires continues presque sûrement. De plus le semi-groupe gaussien $(\mu_t)_t$ est symétrique (i.e. on a l'égalité en loi $X_t \stackrel{d}{=} X_t^{-1}$) si et seulement si les b_i sont tous nuls.

Dans le cas gaussien, l'opérateur du second ordre $L = \sum_i b_i X_i + \frac{1}{2} \sum_{i=1}^d a_{ij} X_i X_j$ peut être mis sous la forme $E_0 + \sum_{i=1}^p E_i^2$, $1 \leq p \leq d$, pour certains vecteurs E_i dans $\mathfrak{g}$. Soit $\mathfrak{h}$ l'algèbre de Lie engendrée par les vecteurs $E_1, ..., E_p$ ainsi que tous les crochets de tout ordre entre les E_i, $0 \leq i \leq p$, faisant intervenir au moins une fois E_0. On dit que l'opérateur $L = L_\mu$ est un **sous-laplacien**, si l'opérateur différentiel $\frac{\partial}{\partial t} - L_\mu$ sur $\mathbb{R}_+^* \times G$ est hypoelliptique, ce qui revient à dire, d'après le théorème de Hörmander, que $\mathfrak{h} = \mathfrak{g}$. Dans ce cas et dans ce cas seulement, le semi-groupe gaussien $(\mu_t)_{t>0}$ possède une densité $p_t(x)$ par rapport à la mesure de Haar de G : c'est le noyau de la chaleur associé à L_μ (voir par exemple [174] dans le cas où μ_t est symétrique, [156] et [157] dans le cas général). Les $(p_t(x))_{t>0}$ sont de classe C^∞, forment une famille de Dirac, et vérifient l'équation de la chaleur $(\frac{\partial}{\partial t} - L_\mu)p_t = 0$. On montre que $p_t(x)$ décroît plus vite, à t fixé, que $e^{-cd(e,x)^2}$ pour une certaine constante $c = c(t) > 0$ (voir [171] pour une étude précise du noyau de la chaleur p_t en fonction de la géométrie du groupe G). Si la matrice (a_{ij}) est définie positive (i.e. $p = d$) alors les densités $p_t(x)$ sont des fonctions analytiques sur G ([87] Theorem 6.3.1).

Evidemment, le support de μ_t (μ_t est toujours supposé gaussien) est inclus dans l'adhérence du sous-groupe analytique G_0 de G dont l'algèbre de Lie est l'algèbre de Lie $\mathfrak{g}_0$ engendrée par les vecteurs $E_0, ..., E_p$. La sous-algèbre de Lie $\mathfrak{h}$ est un idéal de $\mathfrak{g}_0$ et $\mathfrak{h} = \mathfrak{g}_0$ si et seulement si les μ_t sont absolument continues par rapport à la mesure de Haar de G_0. Si $\mathfrak{h}$ est un idéal strict de $\mathfrak{g}_0$ alors le support de μ_t est concentré sur la classe He^{tE_0} où H est le sous-groupe de Lie de G d'algèbre de Lie $\mathfrak{h}$. D'autre part le support de μ_t contient la classe Me^{tE_0} où M est le sous-groupe de G correspondant à l'algèbre de Lie $\mathfrak{m}$ engendrée en tant qu'algèbre de Lie par les vecteurs $E_1, ..., E_p$ ([156] Theorem 4). En particulier si $E_1, ..., E_p$ engendrent tout $\mathfrak{g}$ alors $supp(\mu_t) = G$. Lorsque G est simplement connexe nilpotent $\mathfrak{m}$ est stricte si et seulement si $\mathfrak{h}$ est stricte. Donc dans ce cas $supp(\mu_t) = G$ si et seulement si μ_t possède une densité par rapport à la mesure de Haar. Pour un groupe de Lie G quelconque, il se peut que le support de μ_t soit strict bien que μ_t soit absolument continue (voir [156] example 3.4b, en particulier $p_t(x)$ n'est pas analytique en général).

1.1.5 Loi des grands nombres, théorème limite central

Il y a plusieurs façons de généraliser la loi des grands nombres d'une part et le théorème limite central d'autre part aux marches aléatoires sur les groupes de Lie. Une approche consiste à étudier les variables aléatoires $d(e, S_n)$, où d est une métrique Riemannienne invariante à gauche sur G, et obtenir une loi des grands nombres pour celles-ci, c'est-à-dire une convergence du type

$$\frac{d(e, S_n)}{n} \to \gamma$$

où γ est un réel ≥ 0. Dans [76], Guivarc'h montre qu'une telle convergence a toujours lieu presque sûrement (le théorème de [76] traite le cas général d'un groupe localement compact) et que l'on peut préciser le nombre γ dans de nombreux cas : par exemple si G n'est pas moyennable alors $\gamma > 0$. Plus précisément, on peut se demander si un théorème limite central est valide, c'est-à-dire si on a une convergence du type

$$\frac{d(e, S_n) - \gamma n}{\sqrt{n}} \to X$$

où X est une variable gaussienne centrée non-dégénérée. Nous ne nous occuperons pas de cette question ici (cf. [76] et [79] et le livre [28]).

Une seconde approche consiste à considérer des produits $P_n = X_{1,n} \cdot ... \cdot X_{n,n}$ de variables aléatoires indépendantes telles que la taille typique de $X_{k,n}$ est très petite, de l'ordre de $\frac{1}{n}$ ou $1/\sqrt{n}$. Lorsque G est $\mathbb{R}^d$ ou un groupe de Lie nilpotent homogène (voir plus bas 1.3.1), G possède un semi-groupe à un paramètre de dilatations (δ_t) qui sont des automorphismes ayant la propriété que $\delta_t(K) \to e$ quand $t \to 0$ pour tout compact K de G. On peut alors poser $X_{k,n} = \delta_{\frac{1}{\sqrt{n}}}(X_k)$ et étudier $P_n = \delta_{\frac{1}{\sqrt{n}}}(S_n)$. Quand $G = \mathbb{R}^d$ la convergence en loi, sous certaines hypothèses, des variables P_n vers une loi gaussienne constitue le théorème limite central classique. Dans le cas d'un groupe de Lie connexe général, cette approche infinitésimale permet d'obtenir le théorème limite fondamental énoncé ci-dessous.

On garde les notations du théorème de Hunt énoncé plus haut.

On se donne à présent une famille $(\mu_{k,n})_{1 \le k \le r_n < +\infty}$ de mesures de probabilité sur G et $X_{k,n}$ des variables aléatoires indépendantes à valeurs dans G de loi $\mu_{k,n}$. On pose $\mu_n = \mu_{1,n} * ... * \mu_{r_n,n}$. On fait de plus l'hypothèse que, au sein d'une même ligne à n fixé, les $\mu_{k,n}$ $1 \le k \le r_n$ commutent, i.e. $\mu_{i,n} * \mu_{j,n} = \mu_{j,n} * \mu_{i,n}$. Alors on a le théorème suivant :

Théorème 1.9 *(Wehn [172] [173] [68]) Sous les hypothèses suivantes :*

(i) $\sum_{k=1}^{r_n} \int x_i(g) x_j(g) d\mu_{k,n}(g)$ *converge vers un réel* a_{ij} *quand* $n \to +\infty$ *quels que soient* i *et* j *entre* 1 *et* d*,*

(ii) $\sum_{k=1}^{r_n} \int x_i(g) d\mu_{k,n}(g)$ *converge vers un réel* b_i *quand* $n \to +\infty$ *et la suite* $\sum_{k=1}^{r_n} \left| \int x_i(g) d\mu_{k,n}(g) \right|$ *reste uniformément bornée quel que soit* i *entre* 1 *et* d*,*

(iii) $\sum_{k=1}^{r_n} \int_{G \setminus U} \phi(g) d\mu_{k,n}(g)$ *converge vers* 0 *quand* $n \to +\infty$ *quel que soit le voisinage* U *de l'identité dans* G*,*

Alors, la suite de mesures $(\mu_n)_n$ *converge vers la mesure* ν_1 *qui appartient à un semi-groupe gaussien* $(\nu_t)_{t>0}$ *de mesures de probabilité sur* G *dont le générateur infinitésimal est donné par*

$$L_\nu = \sum_{i=1}^{d} b_i X_i + \frac{1}{2} \sum_{i=1}^{d} a_{ij} X_i X_j$$

Dans la formule ci-dessus, les éléments X_i de l'algèbre de Lie sont considérés en tant qu'opérateurs différentiels agissant sur $C^2(G)$. La matrice (a_{ij}) est par définition semi-définie positive et le semi-groupe (ν_t) associé à L_ν par la formule de Lévy-Khinchine-Hunt est gaussien.

La condition (iii) est l'analogue de la condition de Lindeberg dans le cas classique. La preuve de Wehn repose sur la théorie des semi-groupes et certains théorèmes de convergence pour les suites de semi-groupes établis par Hunt. Stroock et Varadhan [165] ont généralisé ce théorème en donnant une condition nécessaire et suffisante semblable aux conditions du théorème 1.9 (et sans l'hypotèse de commutativité au sein d'une même ligne du système triangulaire) pour que le processus par morceaux $X_t(n)$ égal à $X_{1,n} \cdot ... \cdot X_{[r_n t],n}$ sur l'intervalle $[\frac{[r_n t]}{r_n}, \frac{[r_n t]+1}{r_n}[$ converge en loi vers le processus gaussien $(\nu_t)_{t\in[0,1]}$. Leur preuve est purement probabiliste et, comme la preuve de Wehn, ne fait intervenir aucune propriété de structure spécifique au groupe de Lie G.

1.2 Le théorème local sur les groupes compacts ou abéliens et leurs extensions

1.2.1 Théorèmes locaux sur les groupes abéliens

Sur $\mathbb{R}^d$ on dispose du théorème suivant, classiquement appelé théorème limite local, dont la preuve repose sur l'analyse réelle et la transformation de Fourier.

Théorème 2.1 *(voir [30] Theorem 10.17) Soit μ une probabilité centrée et apériodique sur $\mathbb{R}^d$ admettant un moment d'ordre 2 fini, et K sa matrice de covariance. Alors pour toute fonction continue f à support compact sur $\mathbb{R}^d$, on a*

$$\lim_{n\to+\infty} n^{d/2} \int f d\mu^n = \frac{1}{\sqrt{(2\pi)^d \det K}} \int_{\mathbb{R}^d} f(x)dx$$

Remarquons que la condition d'apériodicité est nécessaire pour obtenir la convergence ci-dessus. Il est naturel qu'on la retrouve dans tous les autres énoncés d'équirépartition. D'une certaine façon, le cas $\mathbb{R}^d$ est le cas idéal, le cas le mieux connu, et on cherche à obtenir des théorèmes semblables sur d'autres groupes. Dans cette thèse, il sera question presque exclusivement de groupes moyennables ou à croissance polynômiale. Dans les groupes non moyennables le comportement des μ^n est très différent mais les questions d'équirépartition s'y posent de la même façon (voir §1.6).

Sur les groupes abéliens plus généraux on dispose toujours d'un théorème quotient. On a

Théorème 2.2 *(Stone [162]) Soit G un groupe abélien localement compact à base dénombrable et engendré par un voisinage compact de l'identité. Soit μ une probabilité apériodique et centrée sur G. Alors on a*

$$\lim_{n\to+\infty} \frac{\mu^n(A)}{\mu^n(B)} = \frac{|A|}{|B|}$$

pour tous boréliens relativement compacts A et B de G de mesure de Haar > 0 et dont la frontière est négligeable.

Remarquons que dans ce dernier théorème, on suppose seulement l'existence d'un moment d'ordre 1.

1.2.2 Groupes compacts

Si le groupe est compact, on a toujours équirépartition. C'est un vieux théorème de Itô et Kawada (cf. [95]).

Théorème 2.3 *(Itô-Kawada) Soit μ une probabilité apériodique sur G compact. Alors la suite μ^n converge en loi vers la mesure de Haar normalisée de G.*

Une preuve possible consiste à appliquer le critère de Lévy pour la convergence des mesures (i.e. le théorème 1.6). En effet, les représentations unitaires irréductibles de G sont de dimension finie et donc les opérateurs $\pi(\mu)$ sont des matrices. Il suffit de vérifier que si π n'est pas la représentation triviale, alors les puissances $\pi(\mu)^n$ tendent vers 0. Pour cela, il suffit de voir que la norme d'opérateur de $\pi(\mu)$ est < 1 ou encore que toutes les valeurs propres de $\pi(\mu)$ sont de module < 1. On vérifie aisément que l'existence d'une valeur propre de module 1 contredit l'apériodicité de μ. Cette méthode est en substance celle qu'avait déjà utilisée Poincaré dans son traité sur les Probabilités de 1912 [134] pour démontrer ce théorème dans le cas particulier où le groupe G est le groupe fini des permutations de n éléments.

Ce théorème implique en particulier l'équirépartition de toute marche aléatoire évoluant sur un groupe compact. Par exemple, le problème de l'équirépartition sur la sphère considéré par Arnol'd et Krylov (cf. [13]). Toute marche aléatoire sur la sphère S^2 par isométries engendrant une orbite dense dans S^2 est équirépartie.

Le problème de la vitesse de convergence dans le théorème précédent est un problème très délicat. Si f est une fonction C^∞ sur G, peut-on estimer la vitesse de convergence de $\int f d\mu^n$ vers $\int_G f(g)dg$? En décomposant f en une somme d'harmoniques correspondant à la décomposition de la représentation régulière de G en somme directe de représentations irréductibles $(\pi_n)_n$, on peut ramener en partie ce problème à étudier la suite des normes $\|\pi_n(\mu)\|$. Nous ne nous intéresserons pas ici à cette question difficile. Bornons-nous à faire quelques remarques. Si μ n'est pas étrangère à la mesure de Haar alors il existe un "trou spectral" c'est-à-dire que les $\|\pi_n(\mu)\|$ sont $< \alpha < 1$ pour une certaine borne α indépendante de n. En revanche, si μ est singulière, et en particulier atomique, le problème reste largement ouvert lorsque G n'est pas commutatif. Quand G est un groupe de Lie compact semisimple, Dolgopyat [51] a récemment obtenu une borne polynômiale du type $\|\pi_n(\mu)\| \leq 1-\frac{c}{n^\beta}$ ce qui permet d'obtenir une vitesse de convergence polynômiale pour les fonctions C^∞ dans le théorème local. Cependant il est conjecturé, par exemple pour $SO(3,\mathbb{R})$, que l'on a toujours un trou spectral (voir le livre de Sarnak [147]). Pour ce groupe, seuls des exemples très spécifiques de mesures μ possédant un trou spectral ont été exhibés (voir [147] [60] et [107]), tous provenant de considérations arithmétiques dans le cadre de la solution au problème de Ruziewicz (voir [107]).

1.2.3 Equirépartition dans le plan

Dans ce paragraphe, nous allons considérer le cas d'une marche aléatoire dans l'espace euclidien $\mathbb{R}^d$, vu comme espace homogène du groupe des déplacements de $\mathbb{R}^d$. On pose donc $G = SO(d) \cdot \mathbb{R}^d$, $X = \mathbb{R}^d$ et une marche aléatoire $S_n \cdot x$. Partant de x, à chaque étape on applique un déplacement aléatoire en suivant une probabilité μ fixée sur G. Un élément de G s'écrit $g = (\tau, \rho)$ où $\tau \in \mathbb{R}^d$ est une translation et $\rho \in SO(d)$ est une rotation. Le premier résultat important concernant cette marche aléatoire est qu'elle satisfait un théorème limite central, à savoir

Théorème 2.4 ***(TLC pour*** $SO(d)\cdot\mathbb{R}^d$***)*** *Soit μ une probabilité adaptée sur $G = SO(d)\cdot \mathbb{R}^d$ telle que l'image de μ sur $\mathbb{R}^d$ par la projection $G \to \mathbb{R}^d$, $g \mapsto \tau$ possède un moment d'ordre 2 fini*[4]*, que l'on note $\sigma^2 = \int_G |\tau|^2 \, d\mu(g)$. Soit $S_n = (\tau_n, \rho_n)$ le produit de n variables aléatoires indépendantes de même loi μ. Alors la variable aléatoire $\frac{1}{\sqrt{n}}\tau_n$ converge en loi vers une gaussienne $\mathcal{N}(0, \sigma^2)$, centrée et de covariance diagonale $\sigma^2 Id$.*

Il en résulte que $\frac{1}{\sqrt{n}}S_n \cdot x$ converge aussi en loi vers la même limite. Ce théorème est une conséquence du théorème d'Ibragimov-Billingsley sur les différences de martingales (cf. [91] et [23] Theorem 35.12). De tels processus satisfont toujours un théorème limite central sous la condition de Lindeberg. La suite S_n écrite sous la forme $X_1 \cdot ... \cdot X_n$ (dans cet ordre) et projetée sur $\mathbb{R}^d$ est un tel processus pourvu que $\mathbb{E}(T_i) = 0$, car elle s'écrit $T_1 + \rho_1(T_2) + ... + \rho_{n-1}(T_n)$ où T_i est la projection de X_i sur $\mathbb{R}^d$. Voir aussi [75] pour une autre preuve par Roynette et aussi l'article de Gorostiza [67].

Passons maintenant au problème de l'équirépartition. En 1965, Kazhdan (cf. [96]) a démontré le théorème suivant (la preuve fut plus tard corrigée et complétée par Guivarc'h dans [74]),

Théorème 2.5 *([96] et [74]). Soit $G = SO(2) \cdot \mathbb{R}^2$ le groupe des déplacements du plan. Soit μ une mesure sur G adaptée, symétrique et à support fini. Alors pour $x \in \mathbb{R}^2$ et pour toutes les fonctions continues positives à support compact ϕ et ψ sur $\mathbb{R}^2$, on a le théorème quotient suivant :*

$$\lim_{n\to+\infty} \frac{\int_G \phi(g\cdot x)d\mu^{2n}(g)}{\int_G \psi(g\cdot x)d\mu^{2n}(g)} = \frac{\int_{\mathbb{R}^2} \phi(y)dy}{\int_{\mathbb{R}^2} \psi(y)dy}$$

Dans [105], Le Page a donné une autre preuve (qui traite le cas un peu plus général d'une mesure symétrique à support compact) de ce résultat reposant sur l'absence de solution autre que la mesure de Lebesgue à un scalaire près à l'équation de Choquet-Deny $\mu * \sigma = \sigma$.

Ci-dessous nous généralisons ce résultat en prouvant le théorème local correspondant en toute généralité :

[4]Remarquons que cette condition d'existence d'un moment d'ordre 2 fini est indépendante de la projection $G \to \mathbb{R}^2$, $g \mapsto g \cdot x$ choisie.

Théorème 2.6 ***(TLL dans le plan)*** *Soit μ une probabilité apériodique sur $G = SO(2)\cdot \mathbb{R}^2$ telle que l'image de μ sur $\mathbb{R}^2$ par la projection $G \to \mathbb{R}^2$, $g \mapsto g\cdot 0$, possède un moment d'ordre 2 fini σ^2. Alors quel que soit $x \in \mathbb{R}^2$, et f une fonction continue à support compact sur $\mathbb{R}^2$, on a*

$$\lim_{n\to+\infty} n \int_G f(g\cdot x)d\mu^n(g) = \frac{1}{2\pi\sigma^2}\int_{\mathbb{R}^2} f(y)dy \tag{1.3}$$

Preuve. L'idée consiste à généraliser la preuve classique du théorème limite local sur $\mathbb{R}$ en considérant la transformée de Fourier. On fait tout d'abord la réduction classique suivante (voir [30]) :

Affirmation 2.7 *Il suffit de démontrer la convergence* (1.3) *pour les fonctions f sur $\mathbb{R}^2$ dont la transformée de Fourier est continue et à support compact (ces fonctions sont intégrables mais ne sont plus à support compact).*

Preuve de l'affirmation. Soit h une fonction continue, intégrable et strictement positive sur $\mathbb{R}^2$ telle que sa transformée de Fourier $\widehat{h}$ soit à support compact (une telle fonction existe ! voir [30] 10.2). D'après le critère de Lévy sur la convergence étroite des mesures, la suite de mesures finies $\nu_n = nh(y)\mu^n * \delta_x(dy)$ converge vers la mesure $\frac{1}{2\pi\sigma^2}h(y)dy$ car la masse totale $\nu_n(\mathbb{R}^2)$ converge vers $\frac{1}{2\pi\sigma^2}\int h$ et les fonctions caractéristiques $\widehat{\nu_n}(t)$ convergent point par point vers la fonction caractéristique de la mesure limite. En effet $\widehat{\nu_n}(t)$ n'est autre que $n\int_G h_t(g\cdot x)d\mu^n(g)$ où $h_t(y) = e^{it\cdot y}h(y)$ et $\widehat{h_t}$ est la translatée de $\widehat{h}$ par t, donc est à support compact. Maintenant si f est continue à support compact sur $\mathbb{R}^2$, f/h l'est aussi et la quantité $\int f/h d\nu_n$ égale $n\int fd(\mu^n * \delta_x)$ et converge vers la limite souhaitée. $\square$

Reprenons maintenant la preuve du théorème. Clairement, il suffit de démontrer l'équidistribution pour $x = 0$. On a alors (θ est un vecteur unitaire dans $\mathbb{R}^2$)

$$\begin{aligned}\int_G f(\tau)d\mu^n(g) &= \int_G\int_{\mathbb{R}^2} \widehat{f}(x)e^{i\tau\cdot x}dxd\mu^n(g) \qquad (1.4)\\ &= \int_0^{+\infty}\int_G\int_0^{2\pi} \widehat{f}(r\theta)e^{ir\tau\cdot\theta}d\theta d\mu^n(g)rdr\\ &= \int_0^{+\infty}\left\langle \pi_r(\mu)^n\mathbf{1}, \widehat{f_r}\right\rangle_{\mathbb{L}^2(S^1)} rdr\end{aligned}$$

où $\mathbf{1}$ est la fonction constante égale à 1 sur le cercle S^1, $\widehat{f_r}$ est la fonction définie S^1 par $\widehat{f}(r\theta)$ et π_r est la représentation unitaire irréductible de G définie sur l'espace $\mathbb{L}^2(S^1)$ par

$$\pi_r(g)\phi(\theta) = e^{ir\tau\cdot\theta}\phi(\rho^{-1}\theta) \tag{1.5}$$

où $r > 0$ et $g = (\tau, \rho) \in G$. Ensuite, on effectue le changement de variable $r \to r\sqrt{n}$ dans (1.4) et on constate que d'après le théorème limite central 2.4, pour tout $r > 0$ et $\theta \in S^1$, on a

$$\lim_{n\to+\infty} \widehat{f}_{r/\sqrt{n}}(\theta) = \widehat{f}(0)$$

et

$$
\begin{aligned}
\lim_{n\to+\infty} \pi_{r/\sqrt{n}}(\mu)^n \mathbf{1}(\theta) &= \lim_{n\to+\infty} \mathbb{E}\left(e^{ir\frac{S_n}{\sqrt{n}}\cdot\theta}\right) \\
&= \mathbb{E}(e^{irN\cdot\theta}) \\
&= e^{-r^2\sigma^2/2}
\end{aligned}
$$

où N est la loi normale $\mathcal{N}(0,\sigma^2)$. Par conséquent $\left\langle \pi_{r/\sqrt{n}}(\mu)^n\mathbf{1}, \widehat{f}_{r/\sqrt{n}}\right\rangle$ converge pour chaque $r>0$ vers $2\pi\widehat{f}(0)e^{-r^2\sigma^2/2}$. Si l'on peut justifier le passage à la limite sous le signe intégral, on aura donc la conclusion recherchée, à savoir

$$
\lim_{n\to+\infty} n\int_G f(\tau)d\mu^n(g) = \widehat{f}(0)\int_{\mathbb{R}} 2\pi e^{-r^2\sigma^2/2} r dr = \frac{1}{2\pi\sigma^2}\int_{\mathbb{R}^2} f(y)dy
$$

car $\widehat{f}(0) = \frac{1}{(2\pi)^2}\int f$.

Puisqu'on a choisi $\widehat{f}$ à support compact, on peut limiter l'intégration dans (1.4) à un compact $[0,M]$. On a alors le lemme suivant :

Lemme 2.8 *Il existe une constante $c>0$ telle que $\|\pi_r(\mu)\| \le 1 - cr^2$ lorsque r est dans un voisinage de 0, soit $[0,\varepsilon]$, et de plus $s = \sup_{r\in[\varepsilon,M]} \|\pi_r(\mu)\| < 1$.*

Avant de démontrer ce lemme on termine la preuve du théorème. D'après le lemme, on obtient

$$
\left|\left\langle \pi_{r/\sqrt{n}}(\mu)^n\mathbf{1}, \widehat{f}_{r/\sqrt{n}}\right\rangle\right| \le s^n \left\|\widehat{f}\right\|_\infty \longrightarrow 0
$$

dès que $r\in[\varepsilon\sqrt{n}, M\sqrt{n}]$, ce qui permet de négliger ce terme dans l'intégrale. Et

$$
\left|\left\langle \pi_{r/\sqrt{n}}(\mu)^n\mathbf{1}, \widehat{f}_{r/\sqrt{n}}\right\rangle\right| \le (1-\frac{cr^2}{n})^n \left\|\widehat{f}\right\|_\infty \le e^{-cr^2}\left\|\widehat{f}\right\|_\infty
$$

pour l'intervalle $r\in[0,\varepsilon\sqrt{n}]$, ce qui permet finalement d'appliquer le théorème de convergence dominée de Lebesgue pour justifier la convergence. $\square$

Preuve du lemme. La preuve de ces inégalités repose sur le fait essentiel que G est résoluble parce que $d=2$. Remarquons d'abord que puisque $\|\pi_r(\mu)\|^2 = \|\pi_r(\mu*\mu^{-1})\|$, on peut supposer que μ est symétrique et adaptée. Démontrons la première inégalité. Pour cela, fixons deux éléments x_0 et y_0 qui ne commutent pas et appartiennent au support μ. Le commutateur (x_0,y_0) est une translation pure non triviale de G. De même on peut se donner deux autres éléments w_0 et z_0 ayant les mêmes propriétés et donnant lieu à une translation pure (w_0,z_0) qui soit non colinéaire à (x_0,y_0). Ceci résulte du fait que μ est adaptée. On fixe alors un voisinage de l'identité U dans G tel que la norme du vecteur de translation (x,y) où $x\in x_0U$ et $y\in y_0U$ et celle de (w,z), $w\in w_0U$ et $z\in z_0U$ soient minorées par un certain réel positif, disons $\alpha>0$ ainsi que l'angle entre ces deux translations. On a

$$
\mu = \int \nu_{x,y,w,z}\mu(dx)\mu(dy)\mu(dw)\mu(dz)
$$

où
$$\nu_{x,y,w,z} = \frac{\delta_x + \delta_{x^{-1}} + \delta_y + \delta_{y^{-1}} + \delta_w + \delta_{w^{-1}} + \delta_z + \delta_{z^{-1}}}{8}$$
Soient μ_{x_0U} et μ_{y_0U} les restrictions normalisées de μ à x_0U et y_0U. Alors il existe $c_0 > 0$ et deux mesures de probabilité ν_1 et ν_2 telles que

$$\begin{aligned}
\mu &= c_0\nu_1 + (1-c_0)\nu_2 \\
\nu_1 &= \int \nu_{x,y,w,z}\mu_{x_0U}(dx)\mu_{y_0U}(dy)\mu_{w_0U}(dw)\mu_{z_0U}(dz)
\end{aligned}$$

On a donc

$$\begin{aligned}
1 - \|\pi_r(\mu)\| &\geq c_0(1 - \|\pi_r(\nu_1)\|) \\
&\geq c_0(1 - \sup_{x\in x_0U,\dots,z\in z_0U} \|\pi_r(\nu_{x,y,w,z})\|)
\end{aligned}$$

Il suffit donc de montrer l'inégalité voulue pour $\nu_{x,y,w,z}$ et ce uniformément quand $x \in x_0U, ..., z \in z_0U$. Mais il est facile d'obtenir cette borne car il suffit de l'obtenir pour $\nu^4_{x,y,w,z}$ à la place de $\nu_{x,y,w,z}$ et $\nu^4_{x,y,w,z}$ est une probabilité symétrique dont le support fini contient les translations pures (x, y), $(x, y)^{-1}$, (w, z) et $(w, z)^{-1}$. La longueur de ces translations étant minorée par $\alpha > 0$ ainsi que l'angle entre elles, on a immédiatement que pour r petit, et uniformément en $x, ..., z$

$$1 - \|\pi_r(\nu_{x,y,w,z})\| \geq C \cdot r^2$$

pour une constante $C > 0$ qui ne dépend que de μ. Cela résulte du fait que puisque l'angle entre $t_1 = (x, y)$ et $t_2 = (w, z)$ est borné inférieurement, la quantité $(t_1 \cdot \theta)^2 + (t_2 \cdot \theta)^2$ est minorée par une constante strictement positive quand θ varie dans S^1. Cela termine la preuve de la première estimation.

Pour la seconde, plus facile, on renvoie le lecteur à la proposition 3.9 plus bas, ou bien au chapitre 3 où on retrouvera le même argument (déjà présent pour l'essentiel dans [73]). $\square$

Notons que le théorème est énoncé pour la dimension 2 seulement. Sa validité en dimension supérieure est une question ouverte à ce jour. Même l'extension du théorème quotient 2.5 à la dimension supérieure reste inconnue.

A la lumière de cette preuve, il apparaît qu'un ingrédient essentiel est l'estimation de la norme $\|\pi(\mu)\|$ pour une représentation unitaire π donnée et en particulier la preuve d'un "trou spectral" $\|\pi(\mu)\| < 1$.

1.3 Le cas des groupes nilpotents

Dans cette section, N désignera toujours un groupe de Lie réel nilpotent simplement connexe. On notera $(C^i(N))_{i=1,\dots,r}$ la suite centrale descendante qui lui correspond, où

$C^1(N) = N$, $C^{i+1}(N) = [N, C^i(N)]$ et r est le plus grand indice tel que $C^r(N)$ est non trivial. De plus, on note

$$d(N) = \sum_{i \geq 1} i \cdot \dim(C^i(N)/C^{i+1}(N)). \tag{1.6}$$

nombre que l'on appelle *exposant de croissance* du groupe N, ou encore *dimension homogène* de N.

1.3.1 Normes homogènes et jauges

Pour ce paragraphe, on renvoie le lecteur aux deux articles de Y. Guivarc'h [76] et [73] ainsi qu'au livre [66].

Algèbre graduée associée, dilatations

Soit $\mathcal{N}$ l'algèbre de Lie de N. L'application exponentielle établit un difféomorphisme $exp : \mathcal{N} \to N$. On note $gr(\mathcal{N})$ l'algèbre graduée canoniquement associée à $\mathcal{N}$. Par définition $gr(\mathcal{N}) = \bigoplus_{i \geq 1} C^i(\mathcal{N})/C^{i+1}(\mathcal{N})$ est munie du crochet de Lie induit par celui de $\mathcal{N}$. Soient $D_k = \bigoplus_{i \geq k} C^i(\mathcal{N})/C^{i+1}(\mathcal{N})$. Ils forment une filtration de $gr(\mathcal{N})$, c'est-à-dire :

$$gr(\mathcal{N}) = D_1 \supsetneq D_2 \supsetneq ... \supsetneq D_{r+1} = \{0\} \text{ et } [D_k, D_l] \subset D_{k+l}$$

de la même façon que les $C^i(\mathcal{N})$ forment une filtration de $\mathcal{N}$.

On se donne un ensemble de sous-espaces vectoriels $(m_i)_{i \geq 1}$ de $\mathcal{N}$ tels que m_i est un supplémentaire de $C^{i+1}(\mathcal{N})$ dans $C^i(\mathcal{N})$. Alors $\mathcal{N} = \oplus_{i \geq 1} m_i$ et, dans cette décomposition, on notera un élément quelconque x de $\mathcal{N}$ (ou N par abus de notation) sous la forme

$$x = \sum_{i \geq 1} x_i$$

Cela permet aussi de définir une application linéaire

$$\phi = (\phi_1, ..., \phi_r) : \mathcal{N} \to gr(\mathcal{N})$$

par la propriété que si $x \in m_i$, $\phi(x) = \phi_i(x) = x \; mod C^{i+1}(\mathcal{N})$. Ainsi ϕ est un isomorphisme d'espaces vectoriels qui préserve les filtrations respectives, i.e. $\phi(C^i(\mathcal{N})) \subset D_i$. De plus, ϕ induit sur m_i un isomorphisme avec $C^i(\mathcal{N})/C^{i+1}(\mathcal{N})$ qui coincide avec l'application quotient canonique.

Ceci permet de définir sur $\mathcal{N}$ une autre structure d'algèbre de Lie en prenant l'image réciproque par ϕ de la structure d'algèbre de Lie de $gr(\mathcal{N})$. Soit $\mathcal{N}'$ l'espace $\mathcal{N}$ muni de cette nouvelle structure : $\mathcal{N}' \simeq gr(\mathcal{N})$. Le nouveau crochet de Lie vérifie alors $[x_i, y_j]' = [x_i, y_j]_{i+j}$.

Le choix des supplémentaires $(m_i)_{i\geq 1}$ permet de définir un semi-groupe $(\delta_t)_{t\geq 0}$ d'applications linéaires de $\mathcal{N}$ appelées **dilatations** par la formule

$$\delta_t(x_i) = t^i x_i$$

Ces dilatations préservent la nouvelle structure d'algèbre de Lie $\mathcal{N}'$ mais ne préservent pas a priori la structure initiale de $\mathcal{N}$. D'autre part on voit que $\mathcal{N}$ et $\mathcal{N}'$ (ou $gr(\mathcal{N})$) sont isomorphes en tant qu'algèbres de Lie si et seulement si les δ_t sont des automorphismes de $\mathcal{N}$. Nous dirons que N (ou $\mathcal{N}$) est **homogène** (on dit aussi gradué) s'il existe un bon choix de supplémentaires (m_i) telle que ce soit le cas.

Normes

Le but ici est d'introduire une classe naturelle de normes sur le groupe N. Soit G un groupe de Lie connexe et U un voisinage compact de l'identité. On définit alors comme plus haut pour tout $g \in G$

$$\delta_U(g) = \inf\{n \geq 0, g \in U^n\}$$

On vérifie que la fonction $\delta_U(\cdot)$ est sous-additive (i.e. $\delta_U(gh) \leq \delta_U(g) + \delta_U(h)$) et que pour différents choix de U, les fonctions obtenues sont *équivalentes* (i.e. il existe des constantes positives A et B telles que $\delta_U(g) \leq A\delta_V(g) + B$ et vice-versa).

Définition 3.1 *On dit qu'une fonction mesurable positive $|\cdot|$ sur G est une **jauge** s'il existe une constante $c_0 \geq 0$ telle que pour tous g et h dans G, $|gh| \leq |g| + |h| + c_0$. C'est une **jauge principale** si elle est équivalente à δ_U.*

Il est facile de voir que toute jauge est localement bornée (i.e. $\sup_{g\in K} |g| < +\infty$ pour tout compact K). En particulier la jauge $|\cdot|$ est principale si et seulement s'il existe un voisinage compact U de l'identité dans G tel que pour tout entier n la boule $\{g \in G, |g| \leq n\}$ est contenue dans U^n.

Supposons maintenant que $G = N$ est nilpotent simplement connexe. Soit $(m_i)_i$ une suite de supplémentaires de $C^{i+1}(\mathcal{N})$ dans $C^i(\mathcal{N})$ comme au paragraphe précédent. Soit $\|\cdot\|_i$ une norme quelconque sur m_i.

Proposition 3.2 *([73] lemme II.1) Quitte à changer chaque $\|\cdot\|_i$ en une norme proportionnelle $\lambda_i \|\cdot\|_i$ ($\lambda_i > 0$), la fonction $|x|_N = \max_i \|x_i\|_i^{1/i}$ est une jauge principale sur N.*

Les définitions suivantes nous seront utiles dans la suite :

Définition 3.3 *Etant donné un choix de supplémentaires m_i comme plus haut, et une base $(X_1, ..., X_n)$ de $\mathcal{N}$ associée à ce choix, on appelle **rectangle** de N tout ensemble constitué des éléments $x = t_1X_1 + ... + t_nX_n$ tels que $t_i \in [a_i, b_i]$ pour des intervalles fermés quelconques $[a_i, b_i]$.*

Définition 3.4 *On dit qu'une fonction continue positive* $|\cdot|$ *est une* ***norme homogène*** *sur* N *pour le semi-groupe* $(\delta_t)_{t\geq 0}$ *de dilatations lorsque*

(i) $|x| = 0 \Leftrightarrow x = 0$

(ii) $|\delta_t(x)| = t|x|$ *pour tout* $t > 0$

Clairement, deux normes homogènes quelconques sont équivalentes (plus précisément, il y a une constante $c > 0$ telle que $\frac{1}{c}|\cdot|_1 \leq |\cdot|_2 \leq c|\cdot|_1$). La fonction introduite à la proposition 3.2 présente l'avantage[5] d'être à la fois une norme homogène et une jauge sur N.

On en déduit que pour toute norme homogène $|\cdot|$ sur N il existe $C > 0$ tel que :

(i) $|x^{-1}| \leq C \cdot |x|$

(ii) $\|x_i\|_i \leq C \cdot |x|^i$

Remarque 3.5 *On peut aussi définir une jauge en considérant des coordonnées de deuxième espèce au lieu de la paramétrisation utilisée ici d'un élément de* N *par les coordonnées de son logarithme dans l'algèbre de Lie. Si* $(X_1, ..., X_n)$ *est une base adaptée à la somme directe* $\mathcal{N} = \oplus_{i\geq 1} m_i$ *telle que* $vect(X_k, ..., X_n)$ *soit un idéal de* $\mathcal{N}$ *pour chaque* k*, on peut considérer l'application*

$$\phi : \mathcal{N} \to N$$
$$\sum t_i X_i \mapsto \prod \exp(t_i X_i)$$

Alors ϕ *est un difféomorphisme polynômial et son inverse aussi est polynômial (voir [43]). De plus on peut poser* $\delta(x) = |\phi^{-1}(x)|_N$ *où* $|\cdot|_N$ *est la jauge définie à la proposition 3.2. On obtient alors une jauge sur* N*.*

1.3.2 Croissance des groupes nilpotents

Dans toute la suite, on notera $|\cdot|_N$ ou simplement $|\cdot|$ la norme introduite à la proposition 3.2. L'existence d'une telle norme sur N permet de démontrer le fait important suivant :

Théorème 3.6 *(Guivarc'h) Soit* U *un voisinage compact de l'identité dans* N*. Il existe deux constantes* $C_1 > 0$ *et* $C_2 > 0$ *telles que pour tout entier* $n \geq 1$

$$C_1 \cdot n^{d(N)} \leq |U^n| \leq C_2 \cdot n^{d(N)} \tag{1.7}$$

où $|X|$ *désigne la mesure de Lebesgue de l'ensemble mesurable* X *et* $d(N)$ *est l'exposant de croissance défini en* (1.6).

[5] Dans [86] on montre que si N est homogène (i.e. δ_t est un automorphisme pour tout $t > 0$) il existe une norme homogène $|\cdot|$ qui est lisse sur $N\backslash\{e\}$ et telle que $|xy| \leq |x| + |y|$ et $|x^{-1}| = |x|$.

Comme la norme $|\cdot|$ est une jauge principale, il existe une constante $c > 0$ telle que, en notant $B_n = \{x \in N, |x| \leq n\}$ on a

$$U^{n/c} \subset B_n \subset U^{cn}$$

Ainsi, on peut remplacer U^n par la boule B_n dans l'estimation (1.7).

Si maintenant Γ est un groupe nilpotent de type fini, disons sans torsion, alors d'après un théorème de Malcev (voir [137] Theorem 2.18), il existe un groupe de Lie nilpotent simplement connexe $\widetilde{\Gamma}$ tel que Γ est un sous-groupe discret cocompact de $\widetilde{\Gamma}$. On en déduit aussitôt le

Théorème 3.7 *(Guivarc'h) Soit Γ un groupe nilpotent de type fini et S un système de générateurs symétrique. Alors il existe deux constantes $C_1 > 0$ et $C_2 > 0$ telles que pour tout entier $n \geq 1$*

$$C_1 \cdot n^{d(\Gamma)} \leq \#B_S(n) \leq C_2 \cdot n^{d(\Gamma)}$$

où $d(\Gamma) = d(\widetilde{\Gamma})$ est l'exposant $\sum i \cdot rk(C^i(\Gamma)/C^{i+1}(\Gamma))$.

Ce théorème a aussi été démontré ultérieurement mais indépendamment par H. Bass par une méthode plus combinatoire (voir [84] pour une discussion de ce résultat et des références antérieures, voir aussi [92]).

1.3.3 Représentations unitaires irréductibles des groupes nilpotents

Fonctions harmoniques

On a déjà souligné l'importance des fonctions μ-harmoniques dans l'étude des problèmes d'équirépartition. D'après un résultat récent de Raugi [143], si G est un groupe nilpotent localement compact à base dénombrable (l.c.b.d.), alors pour une mesure de probabilité μ adaptée quelconque sur G, toute fonction μ-harmonique continue et bornée sur G est constante. Ce résultat, reposant sur le théorème de convergence des martingales, prouve en toute généralité ce qui avait été obtenu auparavant par d'autres méthodes sous des hypothèses restrictives sur la mesure μ. Le cas où G est de type fini et le support de μ engendre G en tant que semi-groupe a été démontré par Dynkin et Malioutov dans [52]. Dans [73], Guivarc'h traite le cas général d'un groupe nilpotent l.c.b.d. pour une mesure μ possédant un moment d'ordre > 0. En fait Guivarc'h démontre un résultat plus précis qui s'apparente au théorème local :

Théorème 3.8 *(Guivarc'h) Soit G un groupe nilpotent l.c.b.d. et μ une mesure de probabilité apériodique sur G. Alors pour toute fonction mesurable f intégrable sur G et d'intégrale totale nulle $\int_G f = 0$, on a*

$$\lim_{n \to +\infty} \|\mu^n * f\|_{\mathbb{L}^1(G)} = 0.$$

Ce théorème nous sera utile au chapitre 2 pour démontrer certains théorèmes d'équirépartition sur les espaces homogènes. On en déduit l'absence de fonctions harmoniques non-constantes de la façon suivante : si ϕ est continue bornée et μ-harmonique, alors $\phi * \mu^n = \phi$ pour tout entier n. Ainsi

$$\|\phi * f\|_\infty = \|\phi * \mu^n * f\|_\infty \leq \|\phi\|_\infty \|\mu^n * f\|_1$$

et d'après la théorème ci-dessus, $\phi * f$ est nulle pour toute fonction intégrable f d'intégrale nulle. Cela signifie que ϕ est constante.

La méthode utilisée repose essentiellement sur la théorie des représentations et sur la propriété de trou spectral suivante que vérifient les opérateurs associés à une probabilité apériodique :

Proposition 3.9 *Soit G un groupe l.c.b.d. nilpotent et μ une mesure de probabilité sur G. Alors les conditions suivantes sont équivalentes :*
(i) μ est apériodique
(ii) pour toute représentation unitaire irréductible non triviale π de G, $\|\pi(\mu)\| < 1$.

Des arguments semblables pour figurent déjà dans [68] et aussi dans [73] (Prop. V.4).

Preuve. Supposons que μ est apériodique et π est une représentation unitaire irréductible non triviale de G et démontrons $(i) \Rightarrow (ii)$. Tout d'abord, si G est abélien, alors π est un caractère non trivial χ de G et si $\|\pi(\mu)\| = |\chi(\mu)| = 1$, alors il existe un nombre complexe z de module 1 tel que pour μ-presque tout $g \in G$, $\chi(g) = z$. Ceci contredit l'apériodicité de μ car $\ker \chi$ est un sous-groupe fermé propre de G.

Supposons maintenant G quelconque et $\|\pi(\mu)\| = 1$. D'après le lemme de Shur, la restriction de π à $C^r(G)$ (qui est inclus dans le centre de G) coïncide avec un caractère de $C^r(G)$, soit $\chi : C^r(G) \to \mathbb{C}^\times$. On choisit (ξ_n) telle que $\|\pi(\mu)\xi_n\| \to 1$, ce qui s'exprime aussi

$$\int_G \langle \pi(g)\xi_n, \xi_n \rangle \, \mu * \mu^{-1}(dg) \to 1$$

Quitte à extraire une sous-suite, on peut alors supposer que $\mu * \mu^{-1}$-presque partout $\langle \pi(g)\xi_n, \xi_n \rangle \to 1$ ce qui équivaut à $\pi(g)\xi_n - \xi_n \to 0$. Mais $\Gamma = \{g \in G, \pi(g)\xi_n - \xi_n \to 0\}$ est un sous-groupe de G, tel que $\mu * \mu^{-1}(\Gamma) = 1$. La condition d'apériodicité entraîne alors que Γ est dense dans G. Il en résulte que $C^r(\Gamma)$ est dense dans $C^r(G)$. De plus, si $\gamma \in C^r(\Gamma)$, $(\chi(\gamma) - 1)\xi_n \to 0$, et cela entraîne que $\chi(\gamma) = 1$. Par conséquent, $\chi(g) = 1$ pour tout $g \in C^r(G)$ et χ est le caractère trivial. Ainsi la représentation π passe au quotient $G/C^r(G)$ et $\pi(\mu) = \overline{\pi}(\overline{\mu})$ où $\overline{\mu}$ est l'image de μ dans la projection canonique sur $G/C^r(G)$ et $\overline{\pi}$ est une représentation unitaire irréductible non triviale de $G/C^r(G)$. Mais puisque μ est apériodique, $\overline{\mu}$ l'est aussi et par récurrence sur le rang de la suite centrale descendante de G, il résulte que $\|\overline{\pi}(\overline{\mu})\| < 1$ ce qui fournit la contradiction recherchée.

Passons à la réciproque : si μ n'est pas apériodique, alors il existe un sous-groupe fermé propre H de G et $g \in G$ tel que μ soit supportée sur gH. La conclusion résulte alors du lemme suivant qui présente un intérêt en soi[6] :

Lemme 3.10 *([72]) Soit H un sous groupe fermé propre d'un groupe nilpotent localement compact. G. Alors il existe un caractère continu non trivial χ de G tel que $H \subset \ker \chi$.*

Ainsi il existe $\theta \in \mathbb{R}$ tel que $\chi(x) = e^{i\theta}$ pour μ-presque tout x dans G. Donc $|\chi(\mu)| = 1$. Pour démontrer ce lemme, on remarque d'abord que si H est distingué alors on peut prendre n'importe quel caractère non trivial du quotient G/H. Sinon, on procède par récurrence sur le plus grand entier p tel que $\overline{H}$ contienne $C^{p+1}(G)$. Si $p = 1$ alors l'image de $\overline{H}$ dans la projection canonique sur $G/\overline{[G, G]}$ est un sous-groupe fermé propre de $G/\overline{[G, G]}$ qui est abélien : on est ramené au cas distingué. Supposons la propriété vraie pour $k \leq p - 1$. Si H n'est pas distingué, son normalisateur L est un sous-groupe fermé propre contenant $C^p(G)$, donc par hypothèse de récurrence L (et donc H) est contenu dans $\ker \chi$ pour un certain caractère non trivial de G. $\square$

Estimation quantitative du trou spectral

La preuve classique du théorème local sur $\mathbb{R}$ est repose sur une estimation au voisinage de l'identité de la norme de la fonction caractéristique $\widehat{\mu}(t)$ de la mesure μ. Il en va de même pour la preuve que l'on donne plus haut du théorème local sur le groupe des déplacements de $\mathbb{R}^2$ (lemme 2.8). Au chapitre 3, on démontre le théorème local pour les mesures centrées sur le groupe de Heisenberg en suivant une stratégie semblable. Il nous faut estimer la norme de $\|\pi(\mu)\|$ lorsque π est dans un voisinage de la représentation triviale pour la topologie de Fell sur le dual unitaire du groupe. Pour la définition de la topologie de Fell, on renvoie à [99] et à [84].

Pour un groupe de Lie nilpotent, le dual unitaire est bien compris d'après la théorie de Dixmier-Kirillov (voir [100]). Pour toute forme linéaire ℓ sur l'algèbre de Lie $Lie(N)$ de N, on peut trouver une sous-algèbre $\mathfrak{m}$ telle que $[\mathfrak{m}, \mathfrak{m}] \subset \ker \ell$ et qui soit maximale pour cette propriété. Cela permet de définir un caractère du groupe $M = \exp(\mathfrak{m})$ en posant $\chi_\ell(\exp(v)) = e^{i\ell(v)}$, puis de définir une représentation $\pi_{\ell,\mathfrak{m}}$ de N en induisant ce caractère à N tout entier $\pi_{\ell,\mathfrak{m}} = Ind_M^N \chi_\ell$. On a alors (voir [100]) :

Théorème 3.11 *(Kirillov) La représentation $\pi_{\ell,\mathfrak{m}}$ est irréductible et deux choix distincts pour l'algèbre maximale $\mathfrak{m}$ conduisent à des représentations équivalentes. Deux formes ℓ et ℓ' sont conjuguées sous l'action de N si et seulement si elles conduisent à des représentations équivalentes. De plus toute représentation unitaire irréductible de N est équivalente à une représentation de cette forme.*

Il en résulte que le dual unitaire de N s'identifie à l'espace des orbites de l'action de N sur l'espace vectoriel dual de $Lie(N)$ (action co-adjointe). Il se trouve que cette

[6] En particulier n éléments de N engendrent un groupe dense dans N si et seulement si leur projection dans le quotient abélien $N/[N, N]$ engendre un sous-groupe dense.

identification est de plus un homéomorphisme entre la topologie de Fell d'un côté et la topologie quotient de l'autre ([34]).

La preuve du théorème 3.11 passe par l'étude préliminaire d'un cas particulier important : le groupe de Heisenberg. Pour ce groupe, ces résultats sont la conséquence du théorème suivant de Stone et Von Neumann.

Théorème 3.12 *(Stone-Von Neumann) Soit $\mathcal{H}$ un espace de Hilbert séparable et A et B deux opérateurs auto-adjoints sur $\mathcal{H}$ tels que $AB - BA = iId$. Alors $\mathcal{H}$ se décompose en une somme directe dénombrable $\mathcal{H} = \oplus_{i\geq 0}\mathcal{H}_i$ de sous-espaces fermés invariants $\mathcal{H}_i$ isomorphes à $\mathbb{L}^2(\mathbb{R})$ et sur lesquels A et B sont simultanément équivalents (i.e. conjugués par une même transformation unitaire) aux transformations $T_1 = i\frac{d}{dt}$ et $T_2 = t$, la multiplication par t dans $\mathbb{L}^2(\mathbb{R})$.*

Pour les sous-groupes à un paramètre d'opérateurs unitaires, ce théorème admet la formulation équivalente suivante : si ρ_1 et ρ_2 sont deux représentations unitaires du groupe $(\mathbb{R}, +)$ dans $\mathcal{H}$ telles que $\rho_1(u)\rho_2(v)\rho_1^{-1}(u)\rho_2^{-1}(v) = e^{icuv}$ $(c > 0)$ alors $\mathcal{H}$ se décompose en $\mathcal{H} = \oplus_{i\geq 0}\mathcal{H}_i$ où $\mathcal{H}_i \simeq \mathbb{L}^2(\mathbb{R})$ et $\rho_1(u)$ et $\rho_2(v)$ sont conjugués à la translation par cu et la multiplication par e^{itv}.

Ce théorème permet de classer les représentations irréductibles du groupe de Heisenberg comme suit[7]. Rappelons d'abord que le groupe de Heisenberg H est par définition le groupe des matrices triangulaires supérieures unipotentes dans $GL_3(\mathbb{R})$. Dans ces coordonnées matricielles, la multiplication s'écrit :

$$(x, y, z) \cdot (x', y', z') = (x + x', y + y', z + z' + xy')$$

Les représentations irréductibles sont de deux sortes : il y a d'une part les caractères ($\chi(x, y, z) = e^{i(ax+by)}$ pour tous réels a et b) et d'autre part une famille à un paramètre de représentations de dimension infinie $(\pi_\lambda)_\lambda$ paramétrée par les caractères non-nuls λ du centre de H. Une réalisation de π_λ est donnée de la façon suivante dans $\mathbb{L}^2(\mathbb{R})$:

$$\pi_\lambda(x, y, z) \cdot f(t) = e^{i\lambda(z+yt)} f(t + x)$$

Au chapitre 3, nous utilisons le théorème de Stone-Von Neumann pour obtenir l'estimation suivante :

Proposition 3.13 *Soit μ une probabilité sur le groupe de Heisenberg H dont le support n'est pas contenu dans une classe d'un sous-groupe abélien propre de H. Alors il existe une constante $c > 0$ telle que pour tout $\lambda \neq 0$ assez petit :*

$$\|\pi_\lambda(\mu)\| < 1 - c|\lambda| \tag{1.8}$$

[7] A la place du théorème de Stone-Von Neumann, on peut aussi utiliser un théorème de Mackey sur les représentations induites : voir par exemple [179] Theorem 7.3.1 et Example 7.3.2.

En particulier, ce résultat s'applique lorsque μ est la mesure symétrique $\mu_0 = \frac{1}{4}(\delta_a + \delta_{a^{-1}} + \delta_b + \delta_{b^{-1}})$ où $a = (1,0,0)$ et $b = (0,1,0)$. On peut voir alors l'opérateur $\pi_\lambda(\mu_0)$ comme agissant sur $\ell^2(\mathbb{Z})$. On retrouve alors l'opérateur de Schrödinger discret bien connu dit de Harper dont le spectre a été très étudié : voir notamment la figure 1 de [19] et le théorème 2.1. de [24] qui justifient l'estimation (1.8).

1.3.4 Théorèmes limites pour les groupes discrets nilpotents

Dans ce paragraphe, nous allons citer quelques-uns des nombreux résultats connus sur le comportement asymptotique des puissances de convolution d'une mesure de probabilité sur un groupe de type fini. Nous nous limiterons aux groupes à croissance polynômiale, c'est-à-dire virtuellement nilpotents d'après le théorème de Gromov [71]. Si le groupe G est discret, toute mesure de probabilité sur G est absolument continue par rapport à la mesure de Haar de G, i.e. la mesure de comptage. Dans ce cas, les méthodes analytiques initiées par N. Varopoulos et développées aussi par de nombreux autres mathématiciens s'avèrent remarquablement efficaces (voir les livres [171] et [175]). Cependant elles ne permettent pas à ma connaissance de traiter le cas des probabilités singulières (par exemple atomiques) sur les groupes de Lie, cas qui comporte d'une certaine façon un aspect arithmétique.

Commençons par un théorème très général dû à Avez [14] :

Théorème 3.14 *(Avez) Soit Γ un groupe moyennable de type fini et μ une mesure de probabilité symétrique à support fini sur Γ. Alors pour tous x et y dans Γ,*

$$\lim_{n\to+\infty} \frac{\mu^{2n}(x)}{\mu^{2n}(y)} = 1$$

Ce résultat est une conséquence du critère de Kesten sur la probabilité de retour (cf. [98]). Dans [171], Varopoulos obtient un théorème local faible pour les mesures symétriques à support fini sur un groupe de type fini à croissance polynômiale. Plus exactement :

Théorème 3.15 *(Varopoulos) Soit Γ un groupe de type fini à croissance polynômiale et μ une mesure de probabilité symétrique, adaptée et à support fini sur Γ. Alors il existe une constante $c > 1$ telle que pour tout $n \geq 1$*

$$\frac{1}{c}\frac{1}{n^{d(\Gamma)/2}} \leq \mu^{2n}(e) \leq c\frac{1}{n^{d(\Gamma)/2}}$$

L'exposant $d(\Gamma)$ est l'exposant de croissance du groupe Γ. D'après le théorème de Gromov, le groupe à croissance polynômiale Γ contient un sous-groupe nilpotent Γ_n d'indice fini. Alors $d(\Gamma)$ est donné par la formule (1.6) avec Γ_n en place de N.

Comme μ est symétrique, $\mu^{2n}(x) \leq \mu^{2n}(e)$ pour tout $x \in \Gamma$ (comme il résulte immédiatement de l'inégalité de Cauchy-Schwartz : voir [14]). De façon plus générale, sans l'hypothèse de symétrie, on a toujours la borne supérieure suivante :

Théorème 3.16 *(Varopoulos, [170]) Soit Γ un groupe de type fini à croissance polynômiale et μ une mesure de probabilité adaptée et à support fini sur Γ. Alors il existe une constante $C > 0$ telle que pour tout $n \geq 1$*

$$\sup_{x \in \Gamma} \mu^n(x) \leq \frac{C}{n^{d(\Gamma)/2}}$$

Ces résultats ont été améliorés notamment par Hebisch et Saloff-Coste dans [85] (cas symétrique) et par Alexopoulos dans [6] (cas centré) qui obtiennent une estimation gaussienne valable sur une large portion du support de μ^n.

Théorème 3.17 *([85] et [6]) Supposons que Γ est à croissance polynômiale et μ est une mesure de probabilité centrée sur Γ dont le support est fini et contient un système symétrique S ($S \ni e$) de générateurs de Γ. Alors il existe des constantes $c > 1$ et $\theta \in (0,1)$ telles que*

(i) pour tout entier $n \geq 1$ et tout $x \in \Gamma$

$$\mu^n(x) \leq c\frac{1}{n^{d(\Gamma)/2}} \exp(-\frac{d(e,x)^2}{cn})$$

(ii) pour tout entier $n \geq 1$ et tout $x \in \Gamma$ tel que $d(e,x) \leq \theta n$

$$\mu^n(x) \geq \frac{1}{cn^{d(\Gamma)/2}} \exp(-c\frac{d(e,x)^2}{n})$$

où d est la distance invariante à gauche induite par le système de générateurs S.

Dans [6], Alexopoulos améliore ces estimations gaussiennes et obtient le théorème limite local avec un contrôle de la vitesse de convergence dans sa version uniforme la plus précise comme suit. On suppose toujours que Γ est à croissance polynômiale et que μ est une mesure de probabilité centrée sur Γ dont le support est fini et contient un système symétrique $S \ni e$ de générateurs de Γ. D'après le théorème de Gromov, Γ contient un sous-groupe nilpotent d'indice fini. Il est facile de voir que, quitte à prendre un sous-groupe d'indice fini plus petit, on peut toujours supposer que Γ contient un sous-groupe nilpotent d'indice fini Γ_N qui soit distingué et sans torsion ([137] Lemma 4.6). D'après le théorème de Malcev (voir [137]), il existe à isomorphisme près un unique groupe de Lie nilpotent simplement connexe N tel que Γ_N s'identifie à un sous-groupe discret cocompact de N.

Soit $g_1, ..., g_k$ des représentants dans Γ des classes de Γ/Γ_N. D'après le théorème de rigidité de Malcev ([137] 2.11), les automorphismes $h \mapsto g_i h g_i^{-1}$ se prolongent à tout N. Cela permet de définir le groupe de Lie $G = \{hg_i, h \in N, 1 \leq i \leq k\}$, dans lequel Γ se plonge comme sous-groupe discret cocompact.

Alexopoulos définit alors sur N un sous-laplacien invariant à gauche L_μ associé à la mesure μ. Lorsque Γ lui-même est nilpotent ce sous-laplacien a la forme simple donnée par la formule (1.9) intervenant dans le théorème central. Le noyau de la chaleur $\widetilde{p}_t(x,y)$

$(x, y \in N)$ associé à L_μ est lisse (voir ci-dessus §1.1.4) et son comportement asymptotique est bien connu (voir [171]). On l'étend à G en posant

$$p_t(g_i x, g_j y) = \frac{1}{k}\widetilde{p}_t(x, y).$$

Théorème 3.18 *([6]) On a*

(i) Il existe un réel $C(\mu) > 0$ tel que pour tout $x \in \Gamma$

$$\lim_{n\to\infty} n^{d(\Gamma)/2} \cdot \mu^n(x) = C(\mu)$$

(ii) Il existe une constante $c > 0$ telle que pour tout entier $n \geq 1$ et tout $x \in \Gamma$ on a

$$|\mu^n(x) - p_n(e, x)| \leq \frac{c}{n^{(d(\Gamma)+1)/2}} \exp\left(-\frac{d(e, x)^2}{cn}\right)$$

Ce théorème est dans une large mesure optimal (il généralise évidemment les théorèmes de [65] obtenus dans le cas abélien par l'analyse de Fourier) et permet de répondre à de nombreuses questions sur le comportement des marches aléatoires centrées sur un groupe de type fini à croissance polynômiale (récurrence, équirépartition, fonctions harmoniques, etc). Les preuves reposent sur une inégalité de Harnack ad hoc pour les solutions de l'équation de la chaleur associée à μ. Remarquons qu'on a ici une vitesse de convergence en $1/\sqrt{n}$. Au chapitre 4, on fera l'observation que dans le cas non-discret, et déjà sur $\mathbb{R}$, la vitesse de convergence dans le théorème local peut être arbitrairement lente pour une mesure μ arbitraire et dépend en fait des propriétés arithmétiques de μ.

1.3.5 Théorèmes limites pour les groupes de Lie nilpotents

Théorèmes centraux

Nous avons déjà mentionné plus haut le théorème limite central de Wehn (§ 1.1.5). Dans le cas des groupes de Lie nilpotents simplement connexes, ce théorème prend la forme simple énoncée ci-dessous. Fixons une suite de supplémentaires m_i de $C^{i+1}(\mathcal{N})$ dans $C^i(\mathcal{N})$ (cf. §1.3.1) et notons δ_t le semi-groupe de dilatations associé à ce choix. Notons aussi N' le groupe de Lie gradué correspondant et $\mathcal{N}'$ son algèbre de Lie. Fixons $X_1, ..., X_d$ des champs de vecteurs invariants à gauche sur N tels que $X_1(e), ..., X_d(e)$ soit une base de $\mathcal{N}$ adaptée à la décomposition en somme directe $\mathcal{N} = \oplus_{i\geq 1} m_i$. Pour $g \in N$, on note $x_i(g)$ les coordonnées de $\log(g)$ dans cette base, i.e. $\log(g) = \sum x_i(g)X_i(e)$. On note aussi $n_i = \dim(m_1 \oplus ... \oplus m_i)$ et X_i' les champs de vecteurs invariants à gauche pour le produit dans N' tels que $X_i'(e) = X_i(e)$.

Théorème 3.19 ***(TLC cas centré)*** *Soit μ une mesure de probabilité centrée sur N ayant un moment d'ordre 2. Soit S_n la marche aléatoire associée à μ et partant de l'identité. Alors on a la convergence en distribution suivante :*

$$\delta_{\frac{1}{\sqrt{n}}}(S_n) \xrightarrow{d} X$$

où $X = X_1$ est la valeur au temps 1 du processus gaussien $(X_t)_{t\geq 0}$ sur $\mathcal{N}'$ dont le générateur infinitésimal (voir §1.1.5) est le sous-laplacien N'-invariant à gauche donné sur $C^2(N)$ par

$$L_\mu = \sum_{i=n_1+1}^{n_2} b_i X'_i + \frac{1}{2} \sum_{1\leq i,j\leq n_1} a_{ij} X'_i X'_j \tag{1.9}$$

avec

$$a_{ij} = \int x_i(x)x_j(x)d\mu(x) \text{ si } 1 \leq i,j \leq n_1 \tag{1.10}$$
$$b_i = \int x_i(x)d\mu(x) \text{ si } n_1 < i \leq n_2$$

La loi limite est non-dégénérée si et seulement si la matrice (a_{ij}) est définie positive, ou bien ce qui revient au même, si le support de la mesure μ n'est pas contenu dans un sous-groupe fermé connexe propre de N. Dans ce cas, la loi limite possède une densité C^∞ par rapport à la mesure de Haar sur N (voir plus haut 1.1.4).

Le processus $(X_t)_t$ est stable pour (δ_t) c'est-à-dire qu'on a l'identité en loi $X_t \overset{d}{=} \delta_{\sqrt{t}}(X_1)$.

Dans l'écriture de L_μ les éléments de $\mathcal{N}$ sont considérés en tant qu'opérateurs différentiels agissant sur $C^2(N)$. Notons que la limite (X et L_μ) dépend du choix du semi-groupe de dilatations, i.e. du choix des supplémentaires (m_i). si l'on considère deux choix possibles, disons (a) et (b), tels que $\mathcal{N} = \oplus_{i\geq 1} m_i^{(a)} = \oplus_{i\geq 1} m_i^{(b)}$ alors les processus de diffusion limites $X_t^{(a)}(\mu)$ et $X_t^{(b)}(\mu)$ vérifient l'identité en loi

$$\phi_{ab}(X_t^{(a)}(\mu)) \overset{d}{=} X_t^{(b)}(\mu) \tag{1.11}$$

où ϕ_{ab} est l'endomorphisme de l'espace vectoriel $\mathcal{N}$ sur lui-même qui à tout élément de $m_i^{(a)}$ associe sa projection sur $m_i^{(b)}$. L'application ϕ_{ab} établit un isomorphisme entre les algèbres de Lie induites $\mathcal{N}^{(a)}$ et $\mathcal{N}^{(b)}$ sur $\mathcal{N}$ par le choix (a) ou (b) de supplémentaires (cf. §1.3.1) et $\phi_{ba} = \phi_{ab}^{-1}$. La relation entre les semi-groupes de dilatations est donnée par

$$\phi_{ab} \circ \delta_t^{(a)} = \delta_t^{(b)} \circ \phi_{ab}$$

La relation (1.11) résulte aisément du théorème et de l'observation suivante : on a l'identité en loi $X_t^{(a)}(\phi_{ba}(\mu)) \overset{d}{=} X_t^{(a)}(\mu)$ car $a_{ij}(\phi_{ba}(\mu)) = a_{ij}(\mu)$ si $1 \leq i,j \leq n_1$ et $b_i(\phi_{ba}(\mu)) = b_i(\mu)$ si $n_1 < i \leq n_2$. En effet, on vérifie que $x_i^{(a)} \circ \phi_{ba} = x_i^{(a)}$ si $1 \leq i \leq n_1$ et $x_i^{(a)} \circ \phi_{ba} = x_i^{(a)} + \ell_i(x_1^{(a)}, ..., x_{n_1}^{(a)})$ où ℓ_i est une certaine forme linéaire si $n_1 < i \leq n_2$; en prenant les moyennes par rapport à μ et en utilisant le fait que μ est centrée, on obtient les identités voulues.

Lorsque N est homogène (cf. §1.3.1), le théorème est un cas particulier du théorème de Wehn 1.9. Pour une preuve élémentaire, en supposant N homogène et sous la condition que μ possède un moment d'ordre $2r$, voir [79]. Sans hypothèse sur N, le théorème est dû à Crépel et Raugi [46]. On peut pousser plus loin l'approximation de $\delta_{\frac{1}{\sqrt{n}}}(S_n)$ par la loi gaussienne et obtenir les termes suivant dans ce développement asymptotique (développement de Edgeworth). Pour cela, on renvoie à [22].

Dans le cas non centré la situation est radicalement différente (voir [142] et §1.6 ci-dessous).

Théorèmes locaux et équirépartition

Passons maintenant aux théorèmes d'équirépartition. Le premier théorème de ce type est dû à Le Page et concerne les mesures μ symétriques à support fini telles que $\mu(e) > 0$:

Théorème 3.20 *(Le Page [105]) Soit G un groupe nilpotent l.c.b.d. et μ une probabilité symétrique à support fini sur G telle que $\mu(e) > 0$. Alors pour toutes fonctions continues ϕ et ψ positives à support compact*

$$\lim_{n \to +\infty} \frac{\int \phi d\mu^n}{\int \psi d\mu^n} = \frac{\int \phi dg}{\int \psi dg}$$

La preuve de ce théorème repose sur le fait que $(\int \phi d\mu^n)^{1/n}$ tend vers 1 car G est moyennable (voir [98] et [76]) et sur l'absence de fonctions harmoniques positives démontrée dans ce cas par Margulis dans [112].

Un autre résultat ancien de ce type concerne l'équirépartition dans les espaces homogènes et est un corollaire du théorème 3.8 :

Théorème 3.21 *(Guivarc'h [73] [74]) Soit μ une probabilité apériodique sur N. Soit X un espace métrique compact muni d'une mesure de probabilité m. Supposons que N agisse continûment sur X en préservant m et de sorte que m soit la seule mesure de probabilité invariante sous l'action de N (unique ergodicité). Alors pour tout $x \in X$ et toute fonction continue f sur X, on a*

$$\lim_{n \to +\infty} \int f(g \cdot x) d\mu^n(g) = \int_X f(y) dm(y) \tag{1.12}$$

Par exemple, X peut être un espace homogène compact de N, ou bien X peut être le fibré unitaire tangent d'une surface de Riemann compacte avec l'action de $N = \mathbb{R}$ par le flot horocyclique, dont l'unique ergodicité a été démontrée en premier par Furstenberg (cf. [58]). Remarquons qu'on ne suppose pas μ centrée dans ce théorème. On peut se passer d'une telle hypothèse car on a supposé l'unique ergodicité. Plus bas on donne l'exemple du flot horocyclique sur une surface de Riemann non compacte X pour laquelle la convergence (4.4) n'a pas lieu pour presque tout point de départ $x \in X$ dès que μ n'est pas centrée sur $\mathbb{R}$.

Lorsque la mesure μ possède une densité à support compact, les travaux récents d'Alexopoulos permettent d'obtenir l'analogue du théorème 3.18. Le théorème suivant est obtenu selon les mêmes méthodes que 3.18. On fixe une norme homogène $|\cdot|$ sur N.

Théorème 3.22 *(Alexopoulos [7]) Supposons que* μ *est centrée et possède une densité continue à support compact* ϕ *sur le groupe de Lie nilpotent* N. *On suppose aussi que* $\phi(e) > 0$. *Soit* L_μ *le sous-laplacien invariant à gauche sur* N *donné par*

$$L_\mu = \sum_{i=n_1+1}^{n_2} b_i X_i + \frac{1}{2} \sum_{1 \leq i,j \leq n_1} a_{ij} X_i X_j$$

où les coefficients a_{ij} *et* b_i *sont définis par les relations* (1.10). *Soit* $p_t(x)$ *le noyau de la chaleur associé à* L_μ *sur* N *(voir §1.1.4). Alors il existe une constante* $c > 0$ *telle que pour tout* $x \in N$

$$|\phi^{*n}(x) - p_n(x)| \leq \frac{c}{n^{(d(N)+1)/2}} \exp\left(-\frac{|x|^2}{cn}\right) \tag{1.13}$$

De plus, il existe une constante $c(\phi) > 0$ *telle que, uniformément quand* x *reste dans un compact de* N,

$$\lim_{n \to +\infty} n^{d(N)/2} \phi^{*n}(x) = c(\phi) \tag{1.14}$$

Alexopoulos obtient aussi une estimation de ce type dans le cas général des groupes de Lie connexes à croissance polynômiale, mais dans ce cas, la définition de L_μ est plus délicate.

Lorsque μ n'admet pas de densité par rapport à la mesure de Haar, les analogues de (1.14) ou (1.13), c'est-à-dire le théorème limite local et sa version uniforme, restent des problèmes ouverts. Au chapitre 3 cependant on démontre ces résultats pour le groupe de Heisenberg. La preuve reste valide pour les groupes nilpotents de rang deux et il est possible que les méthodes du chapitre 3 puissent apporter un jour une preuve du cas général. Nous avons :

Théorème 3.23 *(voir Théorème 3.1.1 et 3.1.2) Soit* N *le groupe de Heisenberg des matrices triangulaires supérieures* 3×3. *Soit* μ *une probabilité centrée, apériodique et à support compact sur* N. *Soit* $(\nu_t)_t$ *un semi-groupe gaussien associé à* μ *et défini par son générateur infinitésimal* L_μ *donné en* (1.9). *Alors pour toute fonction* f *continue à support compact sur* N, *on a*

$$\lim_{n \to +\infty} n^2 \int f(x) d\mu^n(x) = c(\mu) \int_N f(x) dx$$

où $c(\mu) > 0$ *est la valeur en* e *de la densité* p_1 *correspondant à* ν_1 *(i.e. le noyau de la chaleur pour le sous-laplacien* L_μ*). De plus pour tout borélien borné* B *de* N *dont la frontière est négligeable par rapport à la mesure de Lebesgue (i.e.* $|\partial B| = 0$*) on a*

$$\lim_{n \to +\infty} n^2 \sup_{x \in N} |\mu^n(xB) - \nu_n(xB)| = 0 \tag{1.15}$$

Le théorème est valable pour un choix quelconque de ν. Notons que N est homogène et le choix de ν ne dépend que du choix d'un supplémentaire du centre de $\mathcal{N}$. A chaque choix correspond un unique semi-groupe de dilatations (δ_t) de N, qui sont ici des automorphismes de N. Le semi-groupe (ν_t) est stable par rapport au semi-groupe de dilatations correspondant (δ_t), i.e. $\nu_t = \delta_{\sqrt{t}}(\nu_1)$. Si X, Y, Z est une base de $\mathcal{N}$ telle que $[X, Y] = Z$, alors le groupe des automorphismes de $\mathcal{N}$ est

$$\left\{ \begin{pmatrix} A & 0 \\ C & b \end{pmatrix},\ A \in GL_2(\mathbb{R}),\ \det A = b \right\}$$

Soit δ_t l'automorphisme tel que $A = tId$, $b = t^2$, $C = 0$ et (ν_t) le semi-groupe gaussien correspondant. Alors les autres groupes de dilatations sont de la forme $\phi \circ \delta_t \circ \phi^{-1}$ pour un certain automorphisme ϕ de la forme $A = Id$, $b = 1$. Le semi-groupe gaussien correspondant à $(\phi \circ \delta_t \circ \phi^{-1})_t$ est $(\phi(\nu_t))_t$. Les densités correspondantes ont la même valeur en e et donc $c(\mu)$ est bien indépendant du choix du semi-groupe gaussien.

Notons que dans le théorème 3.22, l'estimation (1.13) montre qu'on a une vitesse de convergence au moins en $\frac{1}{\sqrt{n}}$ dans le théorème local (1.14) (car $n^{d/2}p_n(x)$ converge aussi avec une vitesse en $1/\sqrt{n}$, par [7] 1.9.2). Comme on l'a déjà mis en évidence dans le cas commutatif ($N = \mathbb{R}^d$), une telle vitesse dépend fortement de la régularité de la mesure μ et est liée aux propriétés diophantiennes de μ (voir §1.4.2). Dans le cas où $N = \mathbb{R}^d$ et μ a une densité comme dans l'énoncé de 3.22, on a en fait une convergence plus forte en $1/n$ comme il résulte par exemple de [57] ch. XVI, Théorème 1. Si en revanche le support de μ est fini et très bien approchable par des éléments d'un sous-groupe discret de N, on ne peut espérer obtenir aucune vitesse de convergence dans (1.15).

Au chapitre 2, on obtient des résultats plus faibles que ceux du théorème 3.23, mais pour un groupe nilpotent N quelconque en suivant une autre méthode. Nous avons :

Théorème 3.24 *(voir Théorème 2.0.4) Soit N un groupe de Lie nilpotent simplement connexe et μ une probabilité à support fini symétrique et adaptée sur N. Alors il existe une constante $C > 0$ telle que pour tout borélien B de mesure de Haar positive dont la frontière est négligeable, on a*

$$\frac{1}{C}\frac{|B|}{n^{d(N)/2}} \leq \mu^n(B) \leq C\frac{|B|}{n^{d(N)/2}}$$

dès que $n \geq n_1 = n_1(B)$. De plus la borne supérieure dans cet encadrement est uniforme quand B est remplacé par un quelconque de ses translatés xB, $x \in N$.

La méthode de preuve ici consiste à utiliser les résultats connus pour les marches aléatoires sur les groupes discrets, plus précisément le théorème 3.17 dû à Hebisch et Saloff-Coste dans le cas symétrique, et de démontrer un théorème d'équidistribution "déterministe" à la Weyl pour les plongements denses de groupes discrets nilpotents dans les groupes de Lie.

D'autres théorèmes limites sont connus sur les groupes nilpotents (on peut voir [127]), et notablement un principe de grandes déviations ([17]). Nous n'en aurons pas besoin dans la suite.

1.4 Equirépartition et version fine du théorème local

1.4.1 Equirépartition le long d'une orbite

Soit μ une probabilité apériodique et centrée sur $\mathbb{Z}$. Soit D une partie de $\mathbb{Z}$ telle que la densité $d(D)$ de D sur $\mathbb{Z}$ existe, à savoir

$$\lim_{n\to+\infty} \frac{\#\{D \cap [-n,n]\}}{2n+1} = d(D) \geq 0$$

Une question naturelle se pose : a-t-on

$$\lim_{n\to+\infty} \mu^n(D) = d(D) \text{ ?} \tag{1.16}$$

Comme nous allons le démontrer plus bas, la réponse est oui dès que, par exemple, μ possède un moment d'ordre 2 fini σ^2. Ce type de question apparaît lorsque l'on cherche à démontrer l'équirépartition d'une marche aléatoire. Supposons que (X,T,m) est un système dynamique ergodique et inversible qui conserve la mesure de probabilité m sur un espace topologique X et supposons que la trajectoire $\{T^n x\}_{n\in\mathbb{Z}}$ du point x est équidistribuée, c'est-à-dire que

$$\lim_{N\to+\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} f(T^n x) = \int_X f dm$$

pour toute fonction f continue à support compact sur X (d'après le théorème ergodique de Birkhoff, ceci a lieu pour m-presque tout point x). Alors la marche aléatoire de loi μ évoluant le long de la trajectoire est aussi équidistribuée, c'est-à-dire que

$$\lim_{n\to+\infty} \int f(T^k x) d\mu^n(k) = \int_X f dm \tag{1.17}$$

En effet, il suffit de montrer cette convergence pour les fonctions indicatrices $f = \mathbf{1}_B$ où B est disons un ouvert de X avec $m(\partial B) = 0$. Alors on prend $D = \{k \in \mathbb{Z}, T^k x \in B\}$ et on est ramené à la question envisagée ci-dessus. Il est important de noter ici que la convergence (1.17) a lieu pour le *même* point de départ x. si l'on recherche (1.17) seulement pour m -presque tout point, alors la conclusion est plus facile et résulte d'un théorème ergodique aléatoire (par exemple [130]).

La preuve de (1.16), simple, nécessite la version uniforme du théorème limite local sur $\mathbb{Z}$:

$$\lim_{n\to+\infty} \sqrt{n} \sup_{x\in\mathbb{Z}} \left|\mu^n(x) - p(x/\sqrt{n})\sqrt{n}\right| = 0 \tag{1.18}$$

où p est la gaussienne centrée de variance σ^2. Ce théorème classique pour $\mathbb{Z}$ est démontré dans [65] par l'analyse de Fourier. Si μ est à support fini, c'est évidemment un cas particulier du théorème 3.18 (ii). On a aussi d'après le théorème limite central : pour tout $\varepsilon > 0$ il existe $C > 0$ tel que

$$\sum_{|x| \geq C\sqrt{n}} \mu^n(x) \leq \varepsilon \tag{1.19}$$

En combinant (1.18) et (1.19) on voit qu'il suffit de montrer la convergence pour $p(x/\sqrt{n})/\sqrt{n}$, c'est-à-dire

$$\lim_{n \to +\infty} \sum_{x \in D} p(x/\sqrt{n})/\sqrt{n} = d(D)$$

En approximant p par une fonction par morceaux, on se ramène à démontrer que pour tout $I = \{x \in \mathbb{R}, |x| \in [a, b]\}$, $b > a \geq 0$ on a

$$\lim_{n \to +\infty} \frac{1}{\sqrt{n}} \#\{I\sqrt{n} \cap D\} = d(D)(b - a)$$

Mais ceci résulte directement de l'hypothèse faite que D admet pour densité $d(D)$.

De façon semblable, on montre que si cette fois μ n'est pas centrée alors on a toujours une convergence au sens de Cesaro. Plus précisément, si la moyenne de μ est > 0 et si

$$\lim_{n \to +\infty} \frac{\#\{D \cap [1, n]\}}{n} = d(D) \geq 0$$

alors

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu^k(D) = d(D)$$

Toute cette discussion faite ci-dessus sur $\mathbb{Z}$ s'étend à $\mathbb{R}$ et $\mathbb{R}^d$. En particulier l'outil principal, le théorème limite local uniforme (1.18) admet la forme suivante sur $\mathbb{R}$. Sa preuve, assez délicate, est due à Stone [161].

Théorème 4.1 *(Stone) Soit μ une probabilité centrée et apériodique sur $\mathbb{R}$ admettant un moment d'ordre 2 fini σ^2 et ν la loi gaussienne centrée de variance σ^2. Alors il existe une suite ε_n décroissante vers 0 et ne dépendant que de μ telle que pour tout intervalle fermé I de $\mathbb{R}$ on ait*

$$\lim_{n \to +\infty} \sqrt{n} \sup_{x \in \mathbb{R}} |\mu^n(I + x) - \nu^n(I + x)| \leq \varepsilon_n(1 + |I|)$$

où $|I|$ est la mesure de Lebesgue de I.

Le théorème entraîne comme plus haut un corollaire analogue. Les détails de la preuve sont donnés au chapitre 4.

Théorème 4.2 *(voir Corollaire 4.5.4) Soit μ une mesure de probabilité sur $\mathbb{R}$ centrée, apériodique et ayant un moment d'ordre 2 fini. On suppose que f est une fonction uniformément continue et bornée sur $\mathbb{R}$ telle que la limite suivante existe*

$$\lim \frac{1}{T} \int_0^T f(t)dt = \ell \tag{1.20}$$

quand $|T| \to +\infty$. Alors

$$\lim_{n\to+\infty} \int f d\mu^n = \ell$$

Remarque 4.3 *Le théorème 4.2 est énoncé pour les marches centrées et la même conclusion est évidemment fausse pour les marches non centrées. Cependant, dans ce cas, on a quand même, comme dans le cas évoqué plus haut d'une marche sur $\mathbb{Z}$, une convergence au sens de Cesaro. A savoir, si μ est apériodique et décentrée et f vérifie les hypothèses du théorème, alors*

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{k=0}^{n-1} \int f d\mu^k = \ell$$

La preuve de cette assertion est semblable à celle du théorème.

De même, le théorème 4.2 permet d'obtenir un résultat d'équidistribution probabiliste dès que l'on dispose du résultat déterministe correspondant. Plus précisément, soit X un espace localement compact et $(\phi_t)_t$ un flot agissant continûment sur X en y préservant une mesure de borélienne finie m. Le corollaire qui suit montre que si la trajectoire d'un point $x \in X$ par le flot est équidistribuée par rapport à m, alors toute marche aléatoire centrée le long de cette trajectoire s'équidistribue de la même façon. On fixe une mesure de probabilité μ centrée et apériodique sur $\mathbb{R}$ avec un moment d'ordre 2 fini et S_n la marche aléatoire de loi μ^n associée.

Corollaire 4.4 *Supposons que la trajectoire $(\phi_t)_{t\in\mathbb{R}} \cdot x$ soit équidistribuée par rapport à m, c'est-à-dire*

$$\lim_{|T|\to+\infty} \frac{1}{T} \int_0^T f(\phi_t \cdot x)dt = \int_X f(y)dm(y)$$

pour toute fonction continue à support compact f sur X. Alors on a aussi

$$\lim_{n\to+\infty} \mathbb{E}(f(\phi_{S_n} \cdot x)) = \int_X f(y)dm(y)$$

La preuve découle aussitôt du théorème 4.2 appliquée à la fonction $t \mapsto f(\phi_t \cdot x)$ qui est bornée et uniformément continue sur $\mathbb{R}$, puisque f est à support compact. Dans cet énoncé, le fait que μ est centrée est essentiel. Si μ est décentrée, on a convergence au sens de Cesaro d'après la remarque 4.3.

1.4.2 Théorème local et vitesse de convergence

Au chapitre 4, on étudie le problème de la vitesse de convergence dans le théorème limite local sur $\mathbb{R}$ (i.e. théorème 2.1). Pour cela, on introduit la notion de "probabilité diophantienne" sur $\mathbb{R}$. On dit que μ est *diophantienne* si elle ne peut pas être très bien approximée par une mesure à valeurs dans une progression arithmétique de $\mathbb{R}$. Plus précisément on a :

Définition 4.5 *(voir 4.4.2.1) Soit $l \geq 0$. On dit qu'une probabilité μ sur $\mathbb{R}$ est **l-diophantienne** si l'une des propositions équivalentes suivantes est vérifiée :*

(i) $\exists C > 0$ telle que si $|x|$ est assez grand,

$$\inf_{y \in \mathbb{R}} \int \{xa + y\}^2 d\mu(a) \geq \frac{C}{|x|^l}$$

où $\{t\}$ désigne la partie fractionnaire de t.

(ii) $\exists C > 0$ telle que si $|x|$ est assez grand,

$$|\widehat{\mu}(x)| \leq 1 - \frac{C}{|x|^l}$$

où $\widehat{\mu}$ est la fonction caractéristique de μ.

On obtient alors les résultats suivants qui précisent la vitesse de convergence dans le théorème local et montrent le rôle joué par les propriétés diophantiennes de la mesure μ.

Théorème 4.6 *Soit μ une probabilité centrée sur $\mathbb{R}$ qui possède un moment fini d'ordre 4.*

(i) On suppose que μ est l-diophantienne. Alors pour toute fonction f à support compact et de classe C^k avec $k > 3l/2 + 1$, il existe une constante $C = C(f) > 0$ telle que

$$\left| \int f d\mu^n - \frac{1}{\sqrt{2\pi\sigma^2 n}} \int f(x)dx \right| \leq \frac{C(f)}{n\sqrt{n}}$$

(ii) Si μ est symétrique de la forme $\mu = \nu * \nu^{-1}$ et si μ n'est l-diophantienne pour aucun $l \geq 0$ alors pour tout $\varepsilon > 0$ et tout entier $p > 0$ il existe une fonction C^p à support compact f telle que

$$\overline{\lim_{n \to +\infty}} n^{\frac{1}{2}+\varepsilon} \left| \int f d\mu^n - \frac{1}{\sqrt{2\pi\sigma^2 n}} \int f(x)dx \right| = +\infty$$

(iii) La mesure μ est l-diophantienne pour un certain $l \geq 0$ si et seulement s'il existe $k \geq 0$ tel que, pour toute fonction C^k à support compact sur $\mathbb{R}$, on ait

$$\sup_{t \in \mathbb{R}} \left| \int f(t + \cdot)d\mu^n - \int f(t + \cdot)d\nu^n \right| = O(\frac{1}{n})$$

où ν est la loi gaussienne centrée de même variance que μ.

1.4.3 Théorèmes locaux et grandes déviations

Un autre phénomène lié aux deux paragraphes précédents apparaît lorsque l'on cherche à déterminer le comportement à l'infini de l'expression

$$\int f d\mu^n$$

quand f est une fonction continue bornée sur $\mathbb{R}$ qui ne tend pas nécessairement vers une limite en l'infini. Un premier résultat dans ce sens est le théorème 4.2 cité plus haut. Mais que se passe-t-il si f ne satisfait pas l'hypothèse (1.20) ? Il est alors naturel d'essayer de comparer le comportement de $\int f d\mu^n$ à celui de $\int f d\nu^n$ où ν est la loi gaussienne associée à μ. En effet la quantité $\int f d\nu^n$ est plus facile à comprendre car ν est connue explicitement. Le théorème ci-dessous montre que sous certaines hypothèses de régularité, le comportement asymptotique de $\int f d\mu^n$ est le même pour toutes les lois μ centrées de même variance. Plus précisément :

Théorème 4.7 *(voir Théorème 4.4.5) Soit μ une probabilité centrée sur $\mathbb{R}$ admettant un moment d'ordre 4 et soit ν la loi gaussienne associée. On suppose que μ est l-diophantienne. Alors pour tout $k > 3l/2 + 1$ on a*

$$\lim_{n\to+\infty} \frac{\int f d\mu^n}{\int f d\nu^n} = 1 \tag{1.21}$$

pour toute fonction C^k bornée non nulle $f \geq 0$ sur $\mathbb{R}$ dont toutes les dérivées jusqu'à l'ordre k sont bornées.

Au chapitre 4, on donne, pour chaque $k \geq 0$, un exemple d'une fonction f_k de classe C^k, dont toutes les dérivées jusqu'à l'ordre k tendent vers 0 à l'infini, et d'une loi μ_k diophantienne centrée et à support fini, tel que la limite (1.21) n'ait pas lieu.

La démonstration du théorème 4.2 requiert un contrôle uniforme de la quantité $\mu^n(x+I)$ sur un large intervalle de valeurs de x (I est un intervalle fermé fixé). Cette information est donnée par le théorème local uniforme de Stone (théorème 4.1). Remarquons que ce théorème ne donne une information significative que dans la plage $|x| \leq C\sqrt{n}$ où C est une constante positive. Comme conséquence immédiate du théorème 4.1, on a

$$\lim_{n\to+\infty} \frac{\mu^n(I+x)}{\nu^n(I+x)} = 1$$

uniformément en x quand $|x|/\sqrt{n}$ reste borné. Lorsque l'on cherche à contrôler $\mu^n(I+x)$ pour des valeurs de x plus grandes que $C\sqrt{n}$, on est confronté à un problème de grandes déviations et une hypothèse de moment supplémentaire sur μ est nécessaire. Pour les déviations modérées, on obtient le théorème local suivant :

Théorème 4.8 *(voir Théorème 4.3.2) Soit μ une probabilité centrée et apériodique sur $\mathbb{R}$ admettant un moment fini d'ordre $r > 2$ et ν la loi gaussienne associée. Soit σ^2 la variance de μ, $I = [a, b]$ un intervalle fermé de $\mathbb{R}$ et c un réel $\in]0, r-2[$. Alors on a*

$$\lim_{n\to+\infty} \frac{\mu^n(I+x)}{\nu^n(I+x)} = 1$$

uniformément quand $|x| \leq \sqrt{c\sigma^2 n \log n}$.

Ce théorème étend donc le théorème local de Stone aux déviations modérées. Dans le cas particulier où μ possède une densité par rapport à la mesure de Lebesgue, ce théorème était déjà connu (voir [11]).

1.5 Marches aléatoires unipotentes et théorème de Ratner

1.5.1 Equirépartition des orbites unipotentes dans G/Γ

Dans ce paragraphe, nous présentons brièvement le théorème de Ratner et certains résultats qui s'y rattachent. Soit G un groupe de Lie connexe et Γ un sous-groupe discret de G de covolume fini. On dit qu'un élément $g \in G$ est *unipotent* si l'automorphisme $Ad(g)$ de l'algèbre de Lie de G est unipotent, c'est-à-dire que toutes ses valeurs propres sont égales à 1. Un sous-groupe U de G est dit unipotent si tous ses éléments sont unipotents. De façon assez surprenante, et en contraste fort avec les actions de sous-groupes diagonaux par exemple, les orbites des points de G/Γ sous l'action d'un sous-groupe unipotent ou plus généralement d'un sous-groupe engendré par des unipotents, ont un comportement très régulier.

Ce comportement a d'abord été conjecturé par Raghunathan, Dani et Margulis au cours des années 70 et 80, puis prouvé partiellement par de nombreux auteurs (Dani, Margulis, Shah, etc.) et finalement prouvé en toute généralité par M. Ratner au début des années 90 (cf. [160], [138]).

Théorème 5.1 *(Ratner) Soit H un sous-groupe connexe de G engendré par des éléments unipotents. Soit $x \in G/\Gamma$. Alors il existe un sous-groupe fermé F de G contenant H tel que $\overline{H \cdot x} = F \cdot x$ et tel que l'orbite fermée $F \cdot x$ possède une mesure de probabilité invariante à gauche par F, mesure que l'on note m_x. De plus la mesure m_x est ergodique pour l'action de H. Enfin toute mesure de probabilité sur G/Γ qui est ergodique sous l'action de H est de cette forme.*

Un ingrédient essentiel de la preuve est le théorème de récurrence suivant dû à Dani :

Théorème 5.2 *(Dani) Soit $U = (u_t)_t$ un sous-groupe unipotent à un paramètre de G. Fixons $x \in G/\Gamma$ et $\varepsilon > 0$. Alors, il existe un sous-ensemble compact K de G/Γ tel que*

$$\overline{\lim_{T\to+\infty}} \frac{1}{T} |\{t \in [0,T], u_t \cdot x \notin K\}| \leq \varepsilon$$

Ce théorème implique en particulier le fait non trivial suivant : aucune orbite d'un flot unipotent $(u_t)_t$ ne s'en va à l'infini dans G/Γ, elles reviennent toutes dans un compact. Ce phénomène a été pour la première fois mis en évidence dans [111] par Margulis qui s'en était servi dans la preuve originale de l'arithméticité des réseaux non cocompacts en rang supérieur. La preuve du théorème 5.2 repose sur la nature polynômiale des flots unipotents et utilise les méthodes introduites dans [111].

Afin de démontrer le théorème 5.1, Ratner généralise le théorème de Dani et montre qu'en fait les orbites des sous-groupes unipotents à un paramètre sont équiréparties dans leur adhérence.

Théorème 5.3 *(Ratner) Si $U = \{u_t, t \in \mathbb{R}\}$ est un sous-groupe à un paramètre unipotent de G, alors pour tout $x \in G/\Gamma$, il existe un sous-groupe fermé F de G contenant U tel que l'orbite $F \cdot x$ soit fermée et porte une mesure de probabilité F-invariante m_x telle que*

$$\lim_{T\to+\infty} \frac{1}{T} \int_0^T f(u_t \cdot x)dt = \int_{G/\Gamma} f dm_x$$

pour toute fonction f continue et bornée sur G/Γ.

Plus tard, répondant à une question de Ratner, N. Shah a généralisé ce résultat au cas d'un groupe unipotent simplement connexe quelconque (voir [150]). Soit U un sous-groupe unipotent simplement connexe de G et $X_1, ..., X_n$ une base de son algèbre de Lie, telle que pour chaque $k \in [1,n]$ les vecteurs $X_k, ..., X_n$ engendrent un idéal de $Lie(U)$. On introduit les sous-ensembles de U

$$S(s_1, ..., s_n) = \left\{ \prod \exp(t_i X_i) \in U, 0 \leq t_i \leq s_i \right\}$$

Théorème 5.4 *(Shah) Soit U un sous-groupe unipotent simplement connexe de G et $x \in G/\Gamma$. Alors on a*

$$\lim \frac{1}{\lambda(S(s_1, ..., s_n))} \int_{\lambda(S(s_1,...,s_n))} f(u \cdot x)\lambda(du) = \int_{G/\Gamma} f dm_x$$

où λ représente une mesure de Haar sur U et f est une fonction continue bornée sur G/Γ et tous les s_i tendent vers $+\infty$.

Dans un [151], Shah généralise les théorèmes de Ratner pour l'action d'un groupe H tel que $Ad(H)$ est contenu dans l'adhérence de Zariski du sous-groupe engendré par (i.e. le plus petit sous-groupe algébrique contenant) les éléments unipotents de $Ad(H)$.

On ne suppose plus H connexe. Cela permet de traiter les cas des sous-groupes discrets engendrés par des unipotents, par exemple les réseaux non cocompacts des groupes semisimples. Il obtient

Théorème 5.5 *(Shah) Pour tout $x \in G/\Gamma$, il existe un sous-groupe fermé $F \supset H$ de G tel que $\overline{H \cdot x} = F \cdot x$ et tel que l'orbite fermée $F \cdot x$ possède une mesure de Radon invariante à gauche par F, mesure que l'on note m_x. De plus la mesure m_x est ergodique pour l'action de H. Enfin toute mesure de Radon sur G/Γ qui est ergodique sous l'action de H est de cette forme.*

On observe que dans cet énoncé la mesure m_x est une mesure de Radon (i.e. finie sur les compacts). Shah conjecture dans [151] que les m_x sont en fait finies et montre que cette conjecture se ramène au cas où G est semisimple de rang supérieur et Γ est un réseau irréductible. Typiquement, le fait que m_x est fini entraîne que si H est discret alors toute orbite discrète de H dans G/Γ est en fait finie. Dans le même article, il conjecture que les seules mesures de Radon H-ergodiques sont des mesures finies. Récemment Eskin et Margulis [54] ont répondu affirmativement à cette question et montrent que si H est Zariski dense dans G semisimple, alors toute mesure de Radon H-invariante sur G/Γ est en fait finie. Leur preuve remarquable passe par l'étude d'une propriété de récurrence des marches aléatoires sur G/Γ évoluant sur une orbite du groupe H. Nous décrivons leur résultat au paragraphe suivant.

1.5.2 Une version probabiliste du théorème d'équirépartition de Ratner

On se donne comme précédemment un groupe de Lie connexe G et un sous-groupe discret Γ de G tel que G/Γ soit de volume fini. On garde les notations du paragaphe précédent, en particulier $m_{\mathbf{x}}$ désigne la mesure invariante associée au point $\mathbf{x}$ dans le théorème de Ratner (théorème 5.1).

Eskin et Margulis obtiennent dans [54] le théorème suivant :

Théorème 5.6 *(Eskin-Margulis) Supposons G semisimple de centre fini et Γ irréductible dans G. Soit μ une mesure de probabilité sur G tel que son support soit Zariski-dense dans G. Alors pour tout $\varepsilon > 0$ il existe un compact K dans G/Γ tel que, quelque soit $\mathbf{x} \in G/\Gamma$, on ait*

$$\mu^n * \delta_{\mathbf{x}}(K) \geq 1 - \varepsilon$$

dès que $n \geq n_0 = n_0(\mathbf{x})$. De plus $\mathbf{x} \mapsto n_0(\mathbf{x})$ est localement bornée.

En particulier, la suite de mesures $(\mu^n * \delta_{\mathbf{x}})_n$ est relativement compacte dans l'espace des probabilités sur G/Γ. La preuve utilise les propriétés de l'opérateur associé à μ sur les fonctions sur G et notamment la positivité du premier exposant de Lyapounov (voir par exemple [79] pour une preuve rapide de ce théorème dû à Furstenberg), et s'inspire des méthodes introduites dans [55] pour démontrer la version quantitative de la conjecture d'Oppenheim (voir le survol [15]). Comme corollaire immédiat, ils obtiennent que toute mesure de Radon μ-stationnaire sur G/Γ est en fait finie, répondant ainsi à la question de N. Shah.

Dans ce théorème, le support de μ est Zariski dense dans G supposé semisimple. On peut se demander ce qu'il en est si le support de μ est par exemple contenu dans un sous-groupe unipotent de G. Ce cas est d'une certaine façon plus simple car on dispose déjà du théorème de Ratner et on sait que les orbites des sous-groupes unipotents sont équiréparties (théorèmes 5.3 et 5.4). Plus bas et aux chapitres 2 et 3 on étudie ce problème et on met en évidence le fait que les marches aléatoires centrées évoluant le long d'une orbite d'un sous-groupe unipotent de G (G est supposé de Lie connexe quelconque) sont équiréparties. On peut passer de l'équirépartition de l'orbite (théorèmes 5.3 et 5.4) à l'équirépartition de la marche aléatoire en utilisant une version fine du théorème limite local sur le groupe unipotent considéré (voir le corollaire 4.4).

Lorsqu'un tel théorème est disponible, on obtient aussitôt comme corollaire l'équirépartition de la marche aléatoire correspondante dans l'espace homogène G/Γ. Pour les sous-groupes unipotents à un paramètre, ou plus généralement pour les sous-groupes unipotents abéliens, nous avons :

Théorème 5.7 *Soit* $U = \{u_t, t \in \mathbb{R}\}$ *un sous-groupe à un paramètre unipotent de* G *et* μ *une probabilité apériodique et centrée sur* U *possédant un moment d'ordre* 2 *fini. Si* $\mathbf{x} \in G/\Gamma$*, alors*

$$\lim_{n\to+\infty} \mu^n * \delta_{\mathbf{x}} = m_{\mathbf{x}} \tag{1.22}$$

au sens de la convergence faible des mesures.

Ce théorème est la conséquence directe de la combinaison du théorème 5.3 et du corollaire 4.4 que nous avons déduit du théorème limite local de Stone à la section précédente (voir chapitre 4).

Lorsque U n'est pas abélien, le théorème limite local est encore à l'état de conjecture. Au chapitre 3, nous démontrons le théorème limite local et sa version fine (i.e. l'énoncé analogue au théorème de Stone 4.1) pour le groupe de Heisenberg (voir théorème 3.23 cidessus) et on en déduit comme corollaire l'équirépartition des marches centrées évoluant le long d'une orbite d'un sous-groupe unipotent de G isomorphe au groupe de Heisenberg. C'est le cas par exemple pour les marches aléatoires centrées sur les horosphères des variétés hyperboliques complexes de volume fini.

Notons aussi que si la mesure μ possède une densité continue à support compact alors le théorème d'Alexopoulos permet de la même manière de démontrer la convergence (1.22).

Au chapitre 2, on démontre un théorème limite local faible pour les mesures symétriques à support fini (voir théorème 3.24 plus haut) sur un groupe de Lie nilpotent simplement connexe quelconque. En combinant ce résultat au théorème 3.8 de Guivarc'h, nous parvenons à montrer l'équirépartition de la marche associée dans G/Γ. Nous obtenons :

Théorème 5.8 *Soit U un sous-groupe simplement connexe unipotent de G et μ une mesure de probabilité symétrique et à support fini sur U, alors pour tout $\mathbf{x} \in G/\Gamma$*

$$\lim_{n\to+\infty} \mu^n * \delta_{\mathbf{x}} = m_{\mathbf{x}}$$

1.5.3 Convergence au sens de Cesaro

Dans ce paragraphe, nous allons montrer qu'on a toujours équirépartition au sens de Cesaro, que la marche soit centrée ou non.

Proposition 5.9 *Soit U un sous-groupe à un paramètre de G et μ une probabilité apériodique sur U de variance finie. Alors pour tout $\mathbf{x} \in G/\Gamma$, on a la convergence des mesures*

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu^k * \delta_{\mathbf{x}} = m_{\mathbf{x}} \tag{1.23}$$

Pour démontrer cette proposition, nous pourrions faire comme dans le cas centré et utiliser la remarque 4.3. Cependant, nous allons utiliser une méthode un peu différente qui a l'avantage de demander moins de précisions sur la marche aléatoire que n'en donne la version uniforme du théorème local. Au chapitre 2, on applique cette idée au cas des groupes nilpotents.

Nous allons démontrer que la suite des moyennes de Cesaro est relativement compacte dans l'espace des probabilités sur G/Γ puis que toute mesure limite est absolument continue par rapport à $m_{\mathbf{x}}$. Cela résultera du lemme ci-dessous comme on l'explique plus bas. Or toute mesure limite de la suite de Cesaro est μ-stationnaire. D'après le théorème de Choquet-Deny [42] une mesure μ-stationnaire est nécessairement invariante par U tout entier. Comme $m_{\mathbf{x}}$ est U-ergodique, il en résulte alors que toute mesure limite est précisément égale à $m_{\mathbf{x}}$, ce qui achève la preuve de la proposition. Il nous faut montrer :

Lemme 5.10 *Soit μ une probabilité apériodique sur $\mathbb{R}$ de variance $\sigma^2 < +\infty$ et de moyenne $d = 1$. On note $S_n + n$ la marche aléatoire associée (S_n est la marche centrée associée). Alors pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que si f est une fonction bornée $0 \le f \le 1$ et uniformément continue sur $\mathbb{R}$, telle que*

$$\overline{\lim_{T\to+\infty}} \frac{1}{T} \int_0^T f(t)dt \le \delta \tag{1.24}$$

alors

$$\overline{\lim_{N\to+\infty}} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}(f(S_n + n)) \le \varepsilon \tag{1.25}$$

Soit ϕ une fonction continue à valeurs dans $[0,1]$ sur G/Γ telle que $\phi \equiv 0$ sur un compact K et $\phi \equiv 1$ en dehors d'un compact plus grand. En prenant $f(t) = \phi(u_t \cdot \mathbf{x})$ dans le lemme ci-dessus, on obtient que la suite de mesures $(\frac{1}{n}\sum_{k=0}^{n-1}\mu^k * \delta_{\mathbf{x}})_n$ est relativement compacte. Il suffit de voir que (1.24) a bien lieu si K est assez grand d'après le théorème de récurrence de Dani.

Maintenant si ϕ est continue à support compact sur G/Γ et satisfait $\int \phi(y)dm_{\mathbf{x}}(y) \leq \delta$ alors $f(t) = \phi(u_t \cdot \mathbf{x})$ satisfait (1.24) d'après le théorème d'équidistribution de Ratner, donc pour toute mesure limite ν de la suite de Cesaro, $\int \phi(y)d\nu(y) \leq \varepsilon$ d'après le lemme. Cela montre que ν est absolument continue par rapport à $m_{\mathbf{x}}$ et on conclut par ergodicité.

Preuve du lemme : Tout d'abord on remarque qu'il suffit de démontrer

$$\overline{\lim_{N\to+\infty}}\frac{1}{N}\sum_{N/2\leq n\leq N}\mathbb{E}(f(S_n+n)) \leq \delta \tag{1.26}$$

au lieu de (1.25) comme on le voit en découpant la somme dans (1.25) en sous-sommes de $N/2^{i+1}$ à $N/2^i$. De plus, comme f est uniformément continue, il suffit de montrer (1.26) en remplaçant f par une fonction g en escalier de la forme

$$g(t) = \sum_{i\in\mathbb{Z}} f_i \mathbf{1}_{I_i}$$

où I_i est l'intervalle $[ih, (i+1)h]$ et $h > 0$ est fixé et $f_i \in [0,1]$ (pour obtenir (1.26) pour f on prendra $h < \omega(\delta/2)$ où ω est un module de continuité uniforme pour f) . La seule information probabiliste que l'on utilise est la borne supérieure suivante qui résulte du théorème local de Stone : il existe une constante $C > 0$ telle que, pour tout intervalle fermé I de $\mathbb{R}$, on a, à partir d'un rang n_0 qui dépend de I

$$\sup_{x\in\mathbb{R}}\mathbb{P}(S_n \in I + x) \leq \frac{C\cdot|I|}{\sqrt{n}} \tag{1.27}$$

Ceci étant donné, on fixe une grande constante $D > 0$ et on écrit :

$$\frac{1}{N}\sum_{N/2\leq n\leq N}\mathbb{E}(g(S_n+n)) \leq \frac{1}{N}\sum_{N/2\leq n\leq N}\mathbb{E}(g(S_n+n)1_{|S_n|\leq D\sqrt{N}}) + \max_{N/2\leq n\leq N}\mathbb{P}(|S_n| \geq D\sqrt{N})$$

On peut choisir D assez grand de sorte que le reste dans le terme de droite soit plus petit que $\varepsilon/2$ dès que N est assez grand. Pour le premier terme, on écrit pour N assez grand en utilisant (1.27) :

$$\begin{aligned}
\frac{1}{N}\sum_{N/2\le n\le N}\mathbb{E}(g(S_n+n)1_{|S_n|\le D\sqrt{N}}) &\le \frac{1}{N}\sum_{N/2\le n\le N}\ \sum_{|ih-n|\le D\sqrt{N}} f_i\mathbb{P}(S_n\in I_i-n)\\
&\le \frac{1}{N}\sum_{N/2\le n\le N}\frac{C}{\sqrt{N/2}}\int_{n-D\sqrt{N}}^{n+D\sqrt{N}} g(t)dt\\
&\le \frac{1}{N}\sum_{N/2\le n\le N}\frac{C}{\sqrt{N/2}}\int_{-D\sqrt{N}}^{D\sqrt{N}} g(t+n)dt\\
&\le \frac{1}{N}\sum_{N/2\le n\le N}\frac{C}{\sqrt{N/2}}\sum_{|k|\le D\sqrt{N}}\int_0^1 g(t+n+k)dt\\
&\le \frac{C}{\sqrt{N/2}}\sum_{|k|\le D\sqrt{N}}\frac{1}{N}\int_{N/2+k}^{N+k} g(t)dt
\end{aligned}$$

mais grâce à l'hypothèse faite sur g, à savoir (1.24), chaque terme dans la somme ci-dessus est, disons, $\le 2\delta$ pour N assez grand. Ainsi on obtient

$$\overline{\lim_{N\to+\infty}}\frac{1}{N}\sum_{N/2\le n\le N}\mathbb{E}(g(S_n+n)1_{|S_n|\le D\sqrt{N}})\le 8CD\delta$$

Il suffit alors de choisir δ tel que $8CD\delta\le\varepsilon/2$ et cela termine la preuve. □

Remarque 5.11 *Notons que la proposition (5.9) n'est pas un corollaire du théorème ergodique aléatoire de Kakutani. D'après ce théorème, quelle que soit la probabilité m sur G/Γ, si m est U-ergodique (et ici U peut-être n'importe quel sous-groupe de G), alors la convergence (1.23) a bien lieu pour m-presque tout point de départ $\mathbf{x}$ dans G/Γ (sous l'hypothèse que μ est adaptée à U). Ici au contraire, on cherche à contrôler le comportement de la marche aléatoire quelle que soit l'orbite de U sur laquelle elle se promène.*

1.5.4 Un contre-exemple dans le cas non-centré

Nous donnons ici un exemple qui montre que si le flot U n'est pas uniquement ergodique, alors on ne peut se passer de l'hypothèse que μ est centrée dans le théorème 5.7.

Proposition 5.12 *Soit $G=SL_2(\mathbb{R})$ et $\Gamma=SL_2(Z)$ et U un sous-groupe unipotent à un paramètre de G. Soit μ une mesure de probabilité non centrée sur U de variance $\sigma^2<+\infty$ et de moyenne $d\neq 0$. Alors pour tout compact K de G/Γ et pour presque tout point $\mathbf{x}\in G/\Gamma$, on a*

$$\lim\inf_{n\to+\infty}\mu^n*\delta_{\mathbf{x}}(K)=0 \tag{1.28}$$

Preuve : Nous dirons que le réel θ est *bien approximable des deux côtés* si pour tout $\varepsilon > 0$ et $\sigma \in \{-1, 1\}$ on peut trouver deux entiers x et y dans $\mathbb{Z}^2\backslash\{(0,0)\}$ tels que

$$\begin{aligned} |x(y-\theta x)| &< \varepsilon \\ \sigma x(y-\theta x) &> 0 \end{aligned}$$

Il est facile de vérifier que θ est bien approximable des deux côtés si et seulement si son développement en fractions continues $[a_0, a_1, ..., a_n, ...]$ est tel que les deux sous-suites $(a_{2n})_n$ et (a_{2n+1}) soient non bornées. De plus, au sens de la mesure de Lebesgue sur $\mathbb{R}$, presque tout réel θ est bien approximable des deux côtés.

On identifie $SL_2(\mathbb{R})/SL_2(\mathbb{Z})$ à l'espace des réseaux unimodulaires de $\mathbb{R}^2$ et on pose comme à l'habitude $SL_2(\mathbb{R}) = G$ et $SL_2(\mathbb{Z}) = \Gamma$. Soit $\|\cdot\|$ la norme euclidienne canonique de $\mathbb{R}^2$. Rappelons que d'après le *critère de Mahler* (cf. [25] 1.9), un sous-ensemble $K \subset G/\Gamma$ est relativement compact si et seulement s'il existe $\delta > 0$ tel que $\|v\| > \delta$ pour tout réseau $\mathbf{x} \in K$ et tout vecteur non nul $v \in \mathbf{x}$. On pose aussi $\mathbf{x}_0 = \mathbb{Z}^2$

Soit $(u_t)_t$ un sous-groupe unipotent de $SL_2(\mathbb{R})$, D la droite de $\mathbb{R}^2$ qu'il laisse stable et $\theta \in \mathbb{R}$ la pente (supposée finie) de cette droite. Dans la base canonique de $\mathbb{R}^2$, l'action de u_t s'écrit :

$$u_t \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + \alpha t(y-\theta x) \\ y + \alpha\theta t(y-\theta x) \end{pmatrix}$$

pour un certain $\alpha \in \mathbb{R}\backslash\{0\}$. On fixe un compact quelconque K de G/Γ et on se donne un $\delta > 0$ qui lui correspond par le critère de Mahler (on peut choisir δ petit de sorte que $2|\theta|\delta^3 < 1$). On note Ω l'ensemble des réels t tels que $u_t \cdot \mathbf{x}_0 \notin K$. Pour $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{Z}^2\backslash\{(0,0)\}$ on note $A_{x,y} = \left\{t \in \mathbb{R}, \left| u_t \begin{pmatrix} x \\ y \end{pmatrix} \right| < \delta \right\}$. Alors

$$\Omega = \{t, u_t \cdot \mathbf{x}_0 \notin K\} = \bigcup_{x,y \in \mathbb{Z}^2\backslash\{(0,0)\}} A_{x,y}$$

On fait maintenant l'hypothèse que $\theta \notin \mathbb{Q}$ et que θ est bien approximable des deux côtés. Fixons $\kappa = |2\alpha| \max\{|\theta|, 1\}$ et notons $t_{x,y} = \frac{x}{\alpha(\theta x - y)}$ et $I_{x,y} = \{t, |t| < \frac{\delta}{\kappa|\theta x - y|}\}$. On peut alors trouver des entiers x et y arbitrairement grands tels que

$$|x(y-\theta x)| < \delta^4 \tag{1.29}$$

et $t_{x,y}$ a le signe que l'on veut. On vérifie que, puisqu'on a choisi δ assez petit,

$$t_{x,y} + I_{x,y} \subset A_{x,y}.$$

Soit maintenant $S_n = X_1 + ... + X_n$ une marche aléatoire centrée de variance $\sigma^2 > 0$ et $d \in \mathbb{R}\backslash\{0\}$. Alors $S_n + nd$ est une marche aléatoire décentrée avec un terme de dérive (drift) égal à d, donc S_n est de loi μ^n. Comme x peut être pris aussi grand que l'on veut, on peut supposer que $|I_{x,y}| > 2|d|$. Ainsi on pourra trouver un entier positif n tel que

$$nd \in t_{x,y} + \frac{1}{2} I_{x,y}. \tag{1.30}$$

Alors, si S_n n'est pas trop grand, i.e. si $|S_n| < \frac{1}{4}|I_{x,y}|$, on aura $S_n + nd \in \Omega$. Mais si n vérifie (1.30) alors

$$nd \leq \frac{2|x|}{|\alpha||\theta x - y|}.$$

Grâce à (1.29), il en résulte que si $|S_n| < C_\delta \sqrt{n}$, où $C_\delta = \frac{1}{\delta}\sqrt{|\alpha d|/2\kappa^2}$ alors $|S_n| < \frac{1}{4}|I_{x,y}|$.

Ainsi nous avons trouvé un entier positif n arbitrairement grand (car $|t_{x,y}|$ est arbitrairement grand) tel que l'on a

$$S_n + nd \in \Omega$$

dès que $|S_n| < C_\delta \sqrt{n}$. Comme δ peut être pris aussi petit que l'on veut, C_δ est arbitrairement grand et pour tout $\varepsilon > 0$, il résulte du théorème central limite classique que pour tout n assez grand :

$$\mathbb{P}(|S_n| < C_\delta \sqrt{n}) \geq 1 - \varepsilon.$$

Ainsi on a, pour tout compact K de G/Γ,

$$\limsup \mathbb{P}(u_{S_n + nd} \cdot \mathbf{x}_0 \notin K) = 1. \tag{1.31}$$

Nous avons donc établi (1.28) pour $\mathbf{x} = \mathbf{x}_0$ et dès que la pente θ de la droite fixée par U vérifie la condition diophantienne énoncée au début de ce paragraphe. Revenons à la situation de l'énoncé de la proposition. Soit E l'ensemble des $g \in G$ tels que (1.28) a lieu pour $\mathbf{x} = g^{-1} \cdot \mathbf{x}_0$. Alors E est précisément l'ensemble des $g \in G$ tels que (1.31) a lieu pour tout K lorsque (u_t) est remplacé par son conjugué $(g u_t g^{-1})_t$ dont la droite fixe est $g^{-1}D$. Donc E contient l'ensemble des $g \in G$ tels que la pente de gD est irrationnelle et bien approximable des deux côtés. L'application de G dans $\mathbb{R}$ qui associe à g la pente de gD est différentiable et sans points critiques (elle s'identifie à la translation à gauche dans G/P où P est le stabilisateur de D). Il en résulte que, pour la mesure de Haar de G, presque tout g appartient à E. Cela termine la preuve de la proposition. $\square$

Remarque 5.13 *Dans cet exemple, on a montré que la marche décentrée peut rester très probablement très loin dans la pointe à des temps très grands, mais la même idée montre que la marche peut aussi rester très probablement très proche d'une orbite fermée de U et ce à des temps arbitrairement grands. Dans tous les cas, cela empêche qu'il y ait équirépartition de la marche. L'exemple est décrit dans $SL_2(\mathbb{R})/SL_2(\mathbb{Z})$, mais un phénomène semblable se produit dès que U n'est pas uniquement ergodique, sur G/Γ, i.e. s'il existe des sous-espaces homogènes propres de G/Γ stables par U.*

1.6 Questions et remarques

(i) Revenons au théorème 3.23 et remarquons que l'estimation uniforme (1.15) implique en particulier que

$$\lim_{n\to+\infty} \frac{\mu^n(xB)}{\nu_n(xB)} = 1$$

uniformément quand $|x|/\sqrt{n}$ reste borné. Cela résulte en effet de la borne inférieure gaussienne satisfaite par le noyau de la chaleur $p_t(x)$ associé à μ (i.e. la densité de ν_t). Cependant, cela ne dit rien sur le comportement de ce quotient sur une plage plus large de valeurs de x. Il serait intéressant d'obtenir un énoncé semblable au théorème 4.8 dans le cas des groupes nilpotents, sous une condition de moment supplémentaire à préciser. Le même problème se pose déjà quand μ a une densité, i.e. dans le théorème 3.22.

(ii) Nous n'avons pas parlé de ce qui se passe sur un groupe de Lie nilpotent lorsque la marche aléatoire n'est pas centrée. Dans ce cas, on dispose du théorème limite central de Raugi [142], qui montre que sous une certaine renormalisation $d_{1/\sqrt{n}}$ de $\mathcal{N}$ par des dilatations linéaires (pas des automorphismes de l'algèbre de Lie) déterminées par la loi μ, la loi de $d_{1/\sqrt{n}}(\mu^n \circ \delta_{e^{-nX}})$, où $X \in \mathcal{N}$ représente la moyenne de μ dans $\mathcal{N}/[\mathcal{N},\mathcal{N}]$, converge vers une loi non-dégénérée qui est la loi au temps 1 d'un certain processus de diffusion sur l'espace vectoriel $\mathcal{N}$ dont on connaît le générateur infinitésimal. Le trait remarquable qui distingue ce cas du cas centré, est que la renormalisation $d_{1/\sqrt{n}}$ peut être plus prononcée dans le cas non-centré. Plus précisément, soit $D(\mu)$ l'exposant de dilatation des d_t, i.e. l'entier tel que si B est un voisinage compact de l'identité $|d_t(B)| = t^{D(\mu)}|B|$ pour tout $t \geq 1$, alors on a toujours $d(N) \leq D(\mu) \leq 2d(N) - dim(N)$. De plus $D(\mu) = d(N)$ (resp. $D(\mu) = 2d(N) - dim(N)$) si et seulement si l'application $ad(X) : \mathcal{N}^i/\mathcal{N}^{i+1} \to \mathcal{N}^{i+1}/\mathcal{N}^{i+2}$ est nulle (resp. surjective). Donc en général, on peut avoir $D(\mu) > d(N)$; c'est que les automorphismes intérieurs "étalent" la marche sur un domaine plus vaste. Par exemple, pour le groupe de Heisenberg, $d(N) = 4$ et $D(\mu) = 4$ (cas centré) ou 5 (cas non centré). On peut alors se demander ce qu'il en est pour les théorèmes locaux. En particulier, dans le cas d'un groupe nilpotent de type fini Γ vu comme plongé dans un groupe de Lie nilpotent en tant que réseau cocompact, si μ n'est pas centrée, a-t-on en fait (voir le théorème de Varopoulos 3.16)

$$\sup_{x\in\Gamma} \mu^n(x) \leq \frac{C}{n^{D(\mu)/2}} \ ?$$

(iii) Revenons maintenant au problème de l'équidistribution dans l'espace euclidien (§ 1.2.3). La preuve que l'on a donnée du théorème local pour les groupes des déplacements du plan passe par une estimation de la norme de certains opérateurs associés à μ par les représentations irréductibles de G (lemme 2.8). Lorsque $d = 2$, cette estimation est possible essentiellement parce que G est alors résoluble. Si $d \geq 3$, G possède un facteur de Levi $SO(d)$ qui est un groupe compact semisimple non trivial et la preuve du lemme 2.8 est caduque. Il est cependant légitime de se demander si on a toujours un trou spectral dans ce cas, c'est-à-dire si

$$\|\pi_r(\mu)\| < 1$$

dès que $r > 0$. La représentation unitaire π_r de G est irréductible si $r > 0$ et est définie sur $\mathbb{L}^2(S^{d-1})$ par l'équation (1.5). Lorsque $r = 0$, on retrouve la représentation régulière de $SO(d)$. L'existence d'un trou spectral quand $r = 0$ est précisément la question posée par Sarnak dans [147], et semble très délicate.

(iv) Considérons le théorème 5.6. Il implique en particulier que la suite des moyennes $(\frac{1}{n}\sum_{k=0}^{n-1}\mu^k * \delta_{\mathbf{x}})_n$ est relativement compacte dans l'espace des probabilités sur G/Γ. On peut se demander si cette suite converge vers la probabilité invariante m_x du théorème 5.5. Clairement toute valeur d'adhérence de cette suite est μ-stationnaire. Il suffirait alors de démontrer que toutes les mesures μ-stationnaires sont invariantes par les éléments du support de μ, c'est-à-dire en quelque sorte, une propriété à la Choquet-Deny. Dans [59], Furstenberg appelle cette propriété "stiffness" et montre qu'elle a bien lieu si μ n'est pas étrangère à la mesure de Haar de G. Il pose aussi la question de savoir ce qu'il en est si l'on suppose seulement que le support de μ engendre un groupe dense, ou Zariski-dense dans G.

(v) Nous n'avons pas abordé ci-dessus le problème de l'équidistribution dans les groupes non-moyennables. Dans les groupes de Lie semisimples, on dispose du théorème limite local de Bougerol [27] pour les mesures de probabilité μ dont une certaine puissance n'est pas étrangère à la mesure de Haar. Cependant le problème reste ouvert à ma connaissance pour les mesures singulières. En particulier, soit G un groupe semi-simple et a et b deux éléments pris au hasard (pour la mesure de Haar) dans un petit voisinage de l'identité. Alors comme on l'indiquera dans la Partie B, le groupe engendré par a et b est libre et dense dans G. Soit μ la mesure donnant un poids $1/4$ à a et b et leurs inverses. A-t-on un théorème quotient pour μ ? De façon semblable, on peut se poser la question analogue pour les moyennes faites sur les boules du groupe libre au lieu de la moyenne effectuée selon μ^n.

Chapitre 2

Equidistribution of symmetric random walks on nilpotent Lie groups

In this chapter, we present one approach to tackle the local limit and equidistribution problem for random walks on nilpotent Lie groups. We will work under the hypothesis that the probability measure is finitely supported and symmetric. This allows to make use of the known results for random walks on discrete nilpotent groups. They need to be combined with "purely deterministic" information about how a finitely generated dense subgroup of a nilpotent Lie group is embedded in it. More precisely, we need to show that elements of the dense subgroup densify uniformly in the Lie group. In this respect, Theorem 1.1 is a (weak) analogue of the classical Weyl's equidistribution theorem for dense $\mathbb{Z}$-orbits in the torus.

In the problem of the local limit theorem on a nilpotent group, the upper bound (see estimate (2.1)) is usually the more difficult part of the question. It is obtained here via the equidistribution of dense subgroups just mentioned and a gaussian upper bound for probability measures on discrete nilpotent groups. The lower bound follows from a standard argument involving the central limit theorem.

As an application of these methods we obtain the "probabilistic Ratner theorem" for symmetric finitely supported random walks on unipotent subgroups (see Theorem 0.4 below and Chapter 1 for background on this question and Ratner's theorem). We show that symmetric unipotent random walks on a finite volume homogeneous space do equidistribute to the invariant ergodic measure supported on the closed unipotent orbit on which they live. Three main ingredients are used in the proof : first the "deterministic" Ratner-Shah equidistribution theorem for actions of unipotent subgroups (see Chapter 1), second the uniform upper bound for translates of a bounded set (Theorem 0.1 below), third a theorem of Y. Guivarc'h which allows to show that every limit measure is invariant.

We stress the importance of the uniformity in x in the upper bound (2.1). Already in the local limit theorem on $\mathbb{R}^d$ this uniformity does not follow from the classical local theorem, but requires Stone's uniform estimate. Here, it is the key to show the recurrence of the random walks, when the homogeneous space is not compact (see Proposition 3.3).

Let us state first the main theorems of this chapter. Throughout the rest of the chapter, N will denote a simply connected nilpotent Lie group, and

$$d(N) = \sum k \dim(C^k(N)/C^{k+1}(N))$$

the exponent of polynomial growth of N. We will freely make references to Chapter 1 when is appropriate.

Theorem 0.1 *(Uniform local upper bound) Let μ be an aperiodic symmetric probability measure on N with finite support. Then there exists a constant $C = C(\mu) > 0$ such that for any bounded borelian subset $B \subset N$ with negligeable boundary (i.e. $|\partial B| = 0$) and positive Lebesgue measure, there exists an integer $n_0 = n_0(B) \geq 0$ such that whenever $n \geq n_0$*

$$\mu^n(xB) \leq C\frac{|B|}{n^{d(N)/2}} \tag{2.1}$$

uniformly in $x \in N$.

Theorem 0.2 *(Weak local limit theorem). Let μ be an aperiodic symmetric probability measure on N with finite support. Then one can find positive constants C_1 and C_2 depending only on μ such that for any bounded borelian subset $B \subset N$ with negligeable boundary and positive Lebesgue measure, there exists an integer $n_1 = n_1(B) \geq 0$ such that whenever $n \geq n_1$*

$$C_1\frac{|B|}{n^{d(N)/2}} \leq \mu^n(B) \leq C_2\frac{|B|}{n^{d(N)/2}}$$

Remark 0.3 *As will follow from the proof, in this theorem, the lower bound is uniform with r when B is the ball B_r of radius r centered at identity and $r \in [0,1]$, but is not uniform for, say, all translates of a fixed B. Whereas the upper bound is by essence not uniform for $B = B_r$, $r \leq 1$, but is uniform for translates of a given B as Theorem 0.1 shows.*

Theorem 0.4 *(Equidistribution of unipotent random walks) Let G be a connected Lie group and Γ a lattice in G. Let N be a connected unipotent subgroup in G. Let μ be an aperiodic symmetric probability measure on N with finite support. Then for every point $x \in G/\Gamma$, the sequence of measures $\mu^n * \delta_x$ converges to the unique N-invariant ergodic probability measure m_x whose support is $\overline{Nx}$.*

2.1 Equidistribution of dense subgroups

In order to prove Theorem 0.1, we will use two main ingredients : one is a similar upper bound for random walks on discrete nilpotent groups and the other is a result of equidistribution of dense subgroups in nilpotent Lie groups.

Theorem 1.1 *Let be given a finitely generated nilpotent group Γ together with an homomorphism $\sigma : \Gamma \to N$ into a simply connected Lie group N. Let S be a finite symmetric set of generators of Γ and $B_S(n)$ the ball of radius n in the word metric on Γ defined by S. Let $d(N)$ (resp. $d(\Gamma)$) the exponent of growth of N (resp. Γ). Assume that σ has dense image. Then there is a constant $C = C(S, \sigma) > 0$ such that for every Borel measurable subset $B \subset N$ with negligeable boundary (i.e. $|\partial B| = 0$) and positive Lebesgue measure (i.e. $|B| > 0$), one can find an integer $n_0 \geq 0$ such that whenever $n \geq n_0$*

$$\#\{\gamma \in B_S(n), \sigma(\gamma) \in xB\} \leq C \cdot |B| \cdot n^{d(\Gamma)-d(N)} \tag{2.2}$$

uniformly in $x \in N$.

The general strategy of the proof is to proceed by induction on the rank of Γ and thus reduce to the case when Γ is abelian. The abelian case follow from an application of Weyl's equidistribution theorem.

Proof: Since the elements of torsion of Γ form a finite normal subgroup of Γ and N is simply connected, they belong to the kernel of σ. Hence we can assume that Γ is torsion free. We start by pointing out the fact that changing the generating set S only affects the constant C, so it is enough to prove the theorem for one choice of a generating set. More precisely, by Malcev's theorem there exists a simply connected nilpotent Lie group $\widetilde{\Gamma}$ such that Γ is a lattice in $\widetilde{\Gamma}$ (see [137]). Let us choose a homogeneous norm $|\cdot|_\Gamma$ on $\widetilde{\Gamma}$ as in Proposition 1.3.2. Since $|\cdot|_\Gamma$ is "almost" subadditive (see Definition 1.3.1), there exists a constant $C_1 > 0$, depending on S, such that for all positive integers n, we have $B_S(n) \subset \{\gamma \in \Gamma, |\gamma|_\Gamma \leq C_1 n\}$. Consequently, in the proof of the theorem, we can replace the left hand side of (2.2) by $\#\{\gamma \in \Gamma, |\gamma|_\Gamma \leq n$ and $\sigma(\gamma) \in xB\}$.

Now observe that every borelian subset $B \subset N$ with negligeable boundary can be approximated by a finite union of rectangles in the following way. For every $\varepsilon > 0$ there are finitely many pairwise disjoint rectangles $R_1, ..., R_m$ in N (see the definition of a rectangle above in 1.3.3) such that $B \subset \bigcup R_i$ and $\sum |R_i| \leq |B| + \varepsilon$. Since $|B| > 0$, we can choose $\varepsilon < |B|$, so that $\sum |R_i| \leq 2|B|$. Therefore it suffices to prove the theorem when B is a rectangle.

First, assume that Γ is abelian. Then N is abelian too and we will proceed by induction on $\dim(N)$. First assume $\dim(N) = 1$ and identify N with the additive group $\mathbb{R}$. Then $d(N) = 1$ and $d(\Gamma)$ is equal to the rank of Γ as a finitely generated abelian group. In this case, the result is a consequence of the classical Weyl equidistribution theorem as follows. Let $k := d(\Gamma)$ and $\xi_1, ..., \xi_k$ be free generators of the free abelian group Γ. Since $\sigma(\Gamma)$ is dense in N, there must be two generators, say ξ_1, ξ_2 such that $\sigma(\xi_1)$ and

$\sigma(\xi_2)$ are not commensurable. Rescaling N if necessary, we can assume that $\sigma(\xi_2) = 1$ and $\sigma(\xi_1) = \alpha \notin \mathbb{Q}$. Splitting B in smaller intervals if necessary, we can assume that B is an interval $[a, b]$ lying in $[0, 1]$. Then

$$\begin{aligned}
\#\{\gamma \in \Gamma \cap xB, |\gamma|_\Gamma \leq n\} &= \#\{(n_1, ..., n_k) \in \mathbb{Z}^k, |n_i| \leq n \text{ and } \sum n_i \xi_i \in B + x\} \\
&\leq (2n+1)^{k-2} \cdot \sup_{x \in \mathbb{R}} \#\{(n_1, n_2) \in \mathbb{Z}^2, |n_i| \leq n \text{ and } n_1\alpha + n_2 \in B + x\} \\
&\leq (2n+1)^{k-2} \cdot \sup_{x \in \mathbb{R}} \#\{n_1 \in \mathbb{Z}, |n_1| \leq n \text{ and } n_1\alpha \in B + x \bmod(1)\} \qquad (2.3)
\end{aligned}$$

Weyl's equidistribution theorem says that as n tends to infinity

$$\frac{1}{2n+1} \sum_{k=-n}^{n} \chi_{B+x}(\{k\alpha\}) \to |B|$$

where χ_{B+x} is the indicator function of the set $B + x$ considered projected in $\mathbb{R}/\mathbb{Z}$, and $\{k\alpha\}$ is the fractional part of $k\alpha$. From the proof of Weyl's theorem via Fourier analysis, it is easy to check that this convergence is uniform in $x \in \mathbb{R}$. Combining this with equation 2.3, we obtain the existence of an integer $n_0 = n_0(B) \geq 0$ such that whenever $n \geq n_0$

$$\#\{\gamma \in \Gamma \cap xB, |\gamma|_\Gamma \leq n\} \leq 2(2n+1)^{k-1} \cdot |B|$$

This shows (2.2) with $C := 2^{k+1}$ for instance.

Let us now pass to the case when N is abelian and $\dim(N) = d$. Since $\sigma(\Gamma)$ is dense, there are d generators among $\xi_1, ..., \xi_k$ of Γ, say $\xi_1, ..., \xi_d$, such that $\sigma(\xi_1), ..., \sigma(\xi_d)$ are linearly independent over $\mathbb{R}$. We identify N with $\mathbb{R}^d$ according to this basis and let T be the torus $\mathbb{R}^d/\mathbb{Z}^d$ in these coordinates. Suppose now that for some $\ell \in [d+1, k]$, say $\ell = d+1$, the cyclic group generated by $\sigma(\xi_\ell)$ is dense in T. Then we can conclude as above by the multi-dimensional version of Weyl's equidistribution theorem. More precisely if B is a given rectangle in $\mathbb{R}^d$ which we assume contained in $[0, 1]^d$ and not reduced to a point,

$$\begin{aligned}
\#\{\gamma &\in \Gamma \cap xB, |\gamma|_\Gamma \leq n\} = \#\{(n_1, ..., n_k) \in \mathbb{Z}^k, |n_i| \leq n \text{ and } \sum n_i \sigma(\xi_i) \in B + x\} \\
&\leq (2n+1)^{k-(d+1)} \cdot \sup_{x \in \mathbb{R}^d} \#\{(n_1, ..., n_{d+1}) \in \mathbb{Z}^{d+1}, |n_i| \leq n \text{ and } \sum_{i=1}^{d+1} n_i \sigma(\xi_i) \in B + x\} \\
&\leq (2n+1)^{k-(d+1)} \cdot \sup_{x \in \mathbb{R}^d} \#\{n_{d+1} \in \mathbb{Z}, |n_{d+1}| \leq n \text{ and } n_{d+1}\sigma(\xi_{d+1}) \in B + x \; mod(\mathbb{Z}^d)\} \\
&\leq 2(2n+1)^{k-d} \cdot |B|
\end{aligned}$$

where the last inequality holds as soon as n is large enough.

We can thus assume that the subgroup generated by $\sigma(\xi_\ell)$, $d+1 \leq \ell \leq k$ is never dense in T. However since $\sigma(\Gamma)$ is dense in N, this subgroup cannot always be finite. It follows that for some $\ell \in [d+1, k]$, the group generated by $\xi_1, ..., \xi_d, \xi_\ell$ is a proper closed subgroup H of $\mathbb{R}^d$ whose connected component of identify H° is non trivial. Hence Γ contains a subgroup Γ_0, the inverse image of H° under σ, such that $\overline{\sigma(\Gamma_0)}$ is a connected subgroup of $\mathbb{R}^d$ of strictly lower dimension.

Now observe that the short exact sequence

$$1 \to \Gamma_0 \to \Gamma \to \Gamma/\Gamma_0 \to 1$$

splits, because Γ/Γ_0 has no torsion : indeed it embeds in the quotient vector space N/H°. We have $rk(\Gamma) = rk(\Gamma_0) + rk(\Gamma/\Gamma_0)$ and we can then find free abelian generators of Γ, which we again denote by $\xi_1, ..., \xi_k$ such that $\xi_1, ..., \xi_l$ are generators of Γ_0 and $\xi_{l+1}, ..., \xi_k$ are representants in Γ of the generators of Γ/Γ_0. We take $|\cdot|_\Gamma$ to be the homogeneous norm on Γ defined by this set of generators. We also identify N with $\mathbb{R}^d$ so that $\mathbb{R}^{d'} \times \{0\}$ coincide with $\overline{\Gamma_0}$, where $d' = \dim \overline{\Gamma_0}$. As above we can assume that B is a rectangle, then $B = R_1 \times R_2$ where $R_1 \subset \mathbb{R}^{d'}$ and $R_2 \subset \mathbb{R}^{d-d'}$. Then, if $x = (x_1, x_2)$ and n large enough

$$\begin{aligned}
\#\{\gamma &\in \Gamma \cap xB, |\gamma|_\Gamma \leq n\} = \#\{(n_1, ..., n_k) \in \mathbb{Z}^k, |n_i| \leq n \\
\text{and } \sum_{i=1}^{l} n_i\xi_i &\in R_1 + x_1, \sum_{i=l+1}^{k} n_i\xi_i \in R_2 + x_2\} \\
&= \#\{(n_1, ..., n_l) \in \mathbb{Z}^l, |n_i| \leq n \text{ and } \sum_{i=1}^{l} n_i\xi_i \in R_1 + x_1\} \times \\
\#\{(n_{l+1}, ..., n_k) &\in \mathbb{Z}^{k-l}, |n_i| \leq n \text{ and } \sum_{i=l+1}^{k} n_i\xi_i \in R_2 + x_2\} \\
&\leq C_1 \cdot |R_1| \cdot n^{l-d'} \times C_2 \cdot |R_2| \cdot n^{k-l-(d-d')}
\end{aligned}$$

where in the last inequality we have used the induction hypothesis on the dimension of N. This concludes the proof when Γ is assumed abelian.

Now suppose that Γ is a general finitely generated torsion free nilpotent group. We will proceed by induction on the rank of Γ. Let $C^i(\Gamma)$ be the descending central series, with $C^0(\Gamma) = \Gamma$ and $C^1(\Gamma) = [\Gamma, \Gamma]$. Let $r \geq 0$ be the first integer such that $C^{r+1}(\Gamma)$ is trivial. Then $\Gamma_0 := C^r(\Gamma)$ is a non trivial subgroup in the center of Γ. Similarly we define $C^i(N)$. Since $\sigma(\Gamma)$ is dense in N, every $\sigma(C^i(\Gamma))$ is dense in $C^i(N)$ and in particular $C^{r+1}(N) = \{1\}$ and Γ_0 is dense in $N_0 := C^r(N)$. Then the map σ induces a map $\overline{\sigma} : \Gamma/\Gamma_0 \to N/N_0$ with dense image. We identify N with its Lie algebra and write elements of N in the form $x = (\pi_1(x), \pi_2(x))$ where $\pi_1(x) \in Lie(N)$ and $\pi_2(x) \in Lie(N)$ are the components of x in the decomposition the Lie algebra of N as a direct sum of vector subspaces $Lie(N) = V_1 \oplus V_2$ where $V_2 = Lie(N_0)$ and V_1 is just a supplementary subspace which is identified with $Lie(N/N_0)$. So we have $x = e^{\pi_1(x)}e^{\pi_2(x)} = e^{\pi_1(x)+\pi_2(x)}$ since N_0 belongs to the center of N. Note that in these coordinates, elements of N_0 are written $(0, y)$ for any $y \in Lie(N_0)$. Let B be a rectangle under this decomposition, i.e. $B = R_1 \cdot R_2$ where $R_i = e^{L_i}$ for some rectangle $L_i \subset V_i$ for $i = 1, 2$.

Let $A_\sigma(x,B,n)$ be the set $\{\gamma \in \Gamma,\ |\gamma|_\Gamma \leq n$ and $\sigma(\gamma) \in xB\}$. Given $c \in \Gamma/\Gamma_0$ we define in a similar way the set $A_\sigma(x,B,n,c)$ to be the subset of $A_\sigma(x,B,n)$ consisting of elements lying in the coset c. Now note that if $\sigma(\gamma) \in xB$ then $\overline{\sigma}(\overline{\gamma}) = \overline{\sigma(\gamma)} \in \overline{xB} = e^{\pi_1(x)}R_1$. Moreover, observe (from the choice of $|\cdot|_\Gamma$ in Chapter 1 Proposition 1.3.2) that $|\overline{\gamma}|_{\Gamma/\Gamma_0} \leq |\gamma|_\Gamma$. It follows that

$$\{c \in \Gamma/\Gamma_0, A_\sigma(x,B,n,c) \neq \emptyset\} \subset A_{\overline{\sigma}}(e^{\pi_1(x)}, R_1, n)$$

If $N = N_0$ then $\overline{\sigma}$ is a constant map and $\pi_1(x) = 0$ and $R_1 = \{1\}$. Then the cardinality of this set is bounded by the growth of Γ/Γ_0, which is well known ([73]) :

$$\#\{c \in \Gamma/\Gamma_0, A_\sigma(x,B,n,c) \neq \emptyset\} \leq \left\{c \in \Gamma/\Gamma_0, |c|_{\Gamma/\Gamma_0} \leq n\right\} \leq C \cdot n^{d(\Gamma/\Gamma_0)}$$

Otherwise, if N_0 is proper in N, since B has positive measure, R_1 also. And since Γ_0 is non trivial, we can apply the induction hypothesis and we get that for some constant $C_1 > 0$ independent of x_1 and R_1 the following inequality holds as soon as n is large enough $(n \geq n_0(R_1))$ uniformly in x_1 :

$$\#A_{\overline{\sigma}}(x_1, R_1, n) \leq C_1 \cdot |R_1| \cdot n^{d(\Gamma/\Gamma_0) - d(N/N_0)} \tag{2.4}$$

Finally, pick an element $\gamma_1 \in A_\sigma(x,B,n,c)$. If γ also belongs to $A_\sigma(x,B,n,c)$, then $\gamma = \gamma_1\gamma_0$ for some $\gamma_0 \in \Gamma_0$. But from the definition of the homogeneous norm on Γ,

$$|\gamma|_\Gamma = \max\{|\overline{\gamma_1}|_{\Gamma/\Gamma_0}, \|\pi_2(\gamma_1) + \pi_2(\gamma_0)\|^{1/r}\}.$$

It follows that $\|\pi_2(\gamma_0)\| \leq 2n^r$, or equivalently, $|\gamma_0|_{\Gamma_0} \leq 2n^r$. Now let $\sigma(\gamma_1) = xb$ and $\sigma(\gamma) = xb'$ where $b, b' \in B$. Then $xb\sigma(\gamma_0) = xb'$, so $b\sigma(\gamma_0) \in B$, or equivalently $\pi_2(b) + \pi_2(\sigma(\gamma_0)) \in L_2 = \log(R_2)$. Therefore

$$\begin{aligned} A_\sigma(x,B,n,c) &\subset \gamma_1 \cdot \{\gamma_0 \in \Gamma_0, |\gamma_0|_{\Gamma_0} \leq 2n^r, \sigma(\gamma_0) \in e^{-\pi_2(b)}R_2\} \\ &\subset \gamma_1 \cdot A_{\sigma_0}(e^{-\pi_2(b)}, R_2, 2n^r) \end{aligned} \tag{2.5}$$

where $\sigma_0 : \Gamma_0 \longrightarrow N_0$ is the restriction of σ to Γ_0 and N_0. It has dense image. If N_0 is trivial then the map is constant $R_2 = \{1\}$ and $\pi_2(b) = 0$, hence a bound on the cardinality of this set is obtained by the growth of Γ_0

$$\#A_\sigma(x,B,n,c) \leq \{\gamma_0 \in \Gamma_0, |\gamma_0|_{\Gamma_0} \leq 2n^r\} \leq C \cdot (n^r)^{rk(\Gamma_0)}$$

Combining this with (2.4) we obtain the desired result, because $d(\Gamma) = d(\Gamma/\Gamma_0) + r \cdot rk(\Gamma_0)$.

Now assume N_0 is not trivial then R_2 has positive measure, because B has positive measure too. And since Γ_0 is abelian, we can apply the first part of the proof to conclude that there is a constant $C_2 > 0$ independent of x_2 and R_2 such that for any large enough n $(n \geq n_0(R_2))$ uniformly in x_2,

$$\#A_{\sigma_0}(x_2, R_2, n) \leq C_2 \cdot |R_2| \cdot n^{rk(\Gamma_0) - rk(N_0)} \tag{2.6}$$

Combining (2.4), (4.4) and (2.6) we obtain for large enough n,

$$A_\sigma(x, B, n) \leq 2^r C_1 C_2 \cdot |R_1||R_2| \cdot n^{d(\Gamma/\Gamma_0)+r\cdot rk(\Gamma_0)-d(N/N_0)-r\cdot rk(N_0)}$$

This concludes the proof of the theorem if we observe that $|B| = |R_1||R_2|$ and $d(\Gamma) = d(\Gamma/\Gamma_0) + r \cdot rk(\Gamma_0)$ and $d(N) = d(N/N_0) + r \cdot rk(N_0)$. $\square$

2.2 Uniform local limit theorem on nilpotent Lie groups

Here we complete the proof of the "weak local limit theorem" for finitely supported aperiodic symmetric probability measures on a simply connected nilpotent Lie group and its uniform version for translates (Theorem 0.1).

2.2.1 Proof of the uniform upper bound for translates

Proof of Theorem 0.1. We are going to make use of the following well known theorem (see [HS], [Woe] Theorem 14.12 and [Ale] Theorem 1.8.) and combine it with theorem 1.1.

Theorem 2.1 *Let Γ be a finitely generated nilpotent group and μ a finitely supported symmetric probability measure on Γ. Given a homogeneous norm $|\cdot|_\Gamma$ on Γ (see above Chapter 1, Definition 1.3.4) there exists a constant $c > 0$ such that for all $\gamma \in \Gamma$ and all $n \in \mathbb{N}$,*

$$\mu^n(\gamma) \leq \frac{c}{n^{d(\Gamma)/2}} \exp\left(-\frac{|\gamma|_\Gamma^2}{cn}\right) \tag{2.7}$$

Remark 2.2 *The corresponding lower estimate has also been established (see [HS], [Ale]). It holds for $|\gamma|_\Gamma \leq \theta n$ for some $\theta \in (0, 1)$. The proofs of these estimates are based on a Harnack principle for harmonic functions whose proof makes use of on an earlier uniform upper bound due to Varopoulos (see [Var]), which asserts that if μ is an adapted probability measure on the discrete nilpotent group Γ, then*

$$\sup_{\gamma\in\Gamma} \mu^n(\gamma) \leq \frac{c}{n^{d(\Gamma)/2}} \tag{2.8}$$

We keep the notations of Theorem 0.1. We consider the measure μ as a symmetric probability measure on the group Γ generated by the support of μ. Let $(S_n)_n$ be the random walk on N corresponding to μ and starting from the identity element in N. We let $|\cdot|$ be a homogeneous norm on Γ. We write :

$$\begin{aligned}\mathbb{P}(S_n \in xB) &= \sum_{\gamma\in\Gamma} \mu^n(\gamma)\chi_{xB}(\gamma) \\ &= \sum_{|\gamma|\le\sqrt{n}} \mu^n(\gamma)\chi_{xB}(\gamma) + \sum_{\sqrt{n}<|\gamma|\le n} \mu^n(\gamma)\chi_{xB}(\gamma)\end{aligned} \tag{2.9}$$

The first term satisfies the following upper estimate :

$$\begin{aligned}\sum_{|\gamma|\le\sqrt{n}} \mu^n(\gamma)\chi_{xB}(\gamma) &\le \frac{c}{n^{d(\Gamma)/2}} \cdot \#\{\gamma\in\Gamma, |\gamma|\le\sqrt{n} \text{ and } \gamma\in xB\} \\ &\le \frac{c}{n^{d(\Gamma)/2}} \cdot C\cdot|B|\cdot(\sqrt{n})^{d(\Gamma)-d(N)} \\ &\le \frac{C_1}{n^{d(N)/2}}|B|\end{aligned}$$

where the first inequality follows from (2.7) and the second inequality follows from Theorem 1.1 and holds uniformly in $x \in N$ as soon as n is larger that some integer $n_0(B)$.

The second term of (2.9) can be written

$$\begin{aligned}\sum_{\sqrt{n}<|\gamma|\le n} \mu^n(\gamma)\chi_{xB}(\gamma) &= \sum_{i\ge0}\ \sum_{2^i\sqrt{n}<|\gamma|\le 2^{i+1}\sqrt{n}} \mu^n(\gamma)\chi_{xB}(\gamma) \\ &\le \sum_{i\ge0} \frac{c}{n^{d(\Gamma)/2}} e^{-4^i/c}\cdot\#\{\gamma\in\Gamma, |\gamma|\le 2^{i+1}\sqrt{n} \text{ and } \gamma\in xB\} \\ &\le \frac{c}{n^{d(\Gamma)/2}} \sum_{i\ge0} e^{-4^i/c}\cdot C\cdot|B|\cdot(2^{i+1}\sqrt{n})^{d(\Gamma)-d(N)} \\ &\le \frac{c\cdot C}{n^{d(N)/2}}\cdot|B|\cdot\sum_{i\ge0} e^{-4^i/c}(2^{i+1})^{d(\Gamma)-d(N)} \\ &\le \frac{C_1}{n^{d(N)/2}}\cdot|B|\end{aligned} \tag{2.10}$$

where $C_1 > 0$ is the (finite!) constant $c\cdot C\cdot\sum_{i\ge0} e^{-4^i/c}(2^{i+1})^{d(\Gamma)-d(N)}$. The inequality holds (thanks to Theorem 1.1) uniformly in $x \in N$ and for all n larger than some integer n_0 depending on the choice of B.

Combining the two terms of (2.9) we obtain the desired result. □

Proof of Theorem 0.2. We have to prove the lower bound. First we are going to show that the lower bound holds for $B = B_r$ being any ball centered at the identity element in N and of radius $r \le 1$ (for some left-invariant metric on N).

2.2.2 Lower bound for centered balls

This follows from a standard Cauchy-Schwarz argument and from the central limit theorem on N.

Lemma 2.3 *Let ϕ be a compactly supported bounded measurable function on N and $x \in N$. Then for $n \in \mathbb{N}$*

$$\left\langle \mu^{2n} * \phi, \phi * \delta_x \right\rangle_{\mathbb{L}^2(N)} \leq \left\langle \mu^{2n} * \phi, \phi \right\rangle_{\mathbb{L}^2(N)}$$

Proof: Since μ is symmetric, $\langle \mu^{2n} * \phi, \phi * \delta_x \rangle = \langle \mu^n * \phi, \mu^n * \phi * \delta_x \rangle$. By the Cauchy-Schwarz inequality, this is $\leq \|\mu^n * \phi\| \, \|\mu^n * \phi * \delta_x\| = \|\mu^n * \phi\|^2$. Again, since μ is symmetric,$\|\mu^n * \phi\|^2 = \langle \mu^{2n} * \phi, \phi \rangle$. $\square$

Let $\phi = \chi_{B_r}$ is the above lemma. Then $\langle \mu^{2n} * \phi, \phi \rangle = \int_{B_r} \mathbb{P}(S_{2n} x \in B_r) dx \leq |B_r| \cdot \mathbb{P}(S_{2n} \in B_r B_r^{-1})$. Hence

$$\left\langle \mu^{2n} * \phi, \phi \right\rangle \leq |B_r| \cdot \mathbb{P}(S_{2n} \in B_{2r})$$

Besides, $\phi * \delta_x = \chi_{B_r x^{-1}}$. From lemma 3.4 below, we can find $m_0 > 0$ depending only on N and, for any radius $r > 0$, countably many points $x_i \in N$ such that for any $t > 0$, $\chi_{|x| \leq t} \leq \sum_{|x_i| \leq t+r} \phi * \delta_{x_i} \leq m_0$. Then, there is $C_0 > 0$ and $n_0 \geq 0$, such that as soon as $n > n_0$

$$\begin{aligned}
\left\langle \mu^{2n} * \phi, \sum_{|x_i|_N \leq 2\sqrt{n}} \phi * \delta_{x_i} \right\rangle &\geq \left\langle \mu^{2n} * \phi, \chi_{|x| \leq \sqrt{n}} \right\rangle \\
&\geq \int_{|x| \leq \sqrt{n}} \mathbb{E}(\phi(S_{2n} x)) dx \\
&\geq \mathbb{E}(\int_{|S_{2n} x| \leq \sqrt{n}} \phi(x) dx) \\
&\geq \mathbb{P}(|S_{2n}| \leq \frac{1}{2}\sqrt{n}) \cdot |B_r| \\
&\geq C_0 \cdot |B_r|
\end{aligned}$$

where the last inequality follows from the central limit theorem on N. On the other hand, applying lemma 2.3

$$\begin{aligned}
\left\langle \mu^{2n} * \phi, \sum_{|x_i|_N \leq 2\sqrt{n}} \phi * \delta_{x_i} \right\rangle &\leq \sum_{|x_i|_N \leq 2\sqrt{n}} \left\langle \mu^{2n} * \phi, \phi * \delta_{x_i} \right\rangle \\
&\leq \#\{i, |x_i|_N \leq 2\sqrt{n}\} \left\langle \mu^{2n} * \phi, \phi \right\rangle \\
&\leq m_0 \cdot \frac{(3\sqrt{n})^d}{|B_r|} |B_r| \cdot \mathbb{P}(S_{2n} \in B_{2r})
\end{aligned}$$

Hence

$$\mathbb{P}(S_{2n} \in B_{2r}) \geq \frac{C_0 \cdot |B_r|}{m_0 (3\sqrt{n})^d} \geq \widetilde{C} \cdot \frac{|B_{2r}|}{n^{d/2}} \tag{2.11}$$

for some absolute constant $\widetilde{C} > 0$ (since $|B_r|/|B_{2r}|$ is bounded below by a positive constant uniformly in $r \in (0,1)$, see lemma 3.4 below).

2.2.3 Lower bound for translates

We now prove the lower bound for translates of the ball B_r. This can be done directly using a ratio limit theorem proved by Le Page (see [LeP0]). However we will give another argument following the spirit of the proof of Theorem 0.1.

Let $B = B_r$. From the upper gaussian estimate (2.7), there is $n_0 = n_0(B) \geq 0$ such that if $n \geq n_0$ and $x \in N$,

$$\begin{aligned} \sum_{n^{2/3}<|\gamma|\leq n} \mu^n(\gamma)\chi_{xB}(\gamma) &\leq \sum_{n^{2/3}<|\gamma|\leq n} \mu^n(\gamma) \\ &\leq c \cdot n^{d(\Gamma)} \exp(-n^{1/3}/c) \\ &\leq \frac{\widetilde{C}}{100} \cdot \frac{|B|}{n^{d/2}} \end{aligned} \tag{2.12}$$

Combining the upper and lower gaussian estimates, we find positive constants c_0 and c_1 such that

$$\frac{\mu^n(\gamma_1)}{\mu^n(\gamma_2)} \geq c_0 \exp(-\frac{c_1}{n} \left||\gamma_1|_\Gamma^2 - |\gamma_2|_\Gamma^2\right|)$$

for any γ_1 and γ_2 such that $|\gamma_i|_\Gamma \leq \theta n$. Now let γ_0 be some fixed element in Γ. Then, for n large enough (i.e. $n \geq n_1 = n_1(\gamma_0) \geq 0$)

$$\begin{aligned} \mathbb{P}(S_{2n} &\in \gamma_0 B) = \sum \mu^{2n}(\gamma)\chi_{\gamma_0 B}(\gamma) = \sum_{\gamma\in B} \mu^{2n}(\gamma_0\gamma) \\ &\geq \sum_{\gamma\in B, |\gamma|_\Gamma \leq n^{2/3}} \mu^{2n}(\gamma_0\gamma) \\ &\geq c_0 \cdot \sum_{\gamma\in B, |\gamma|_\Gamma \leq n^{2/3}} \mu^{2n}(\gamma) \exp(-\frac{c_1}{2n} \left||\gamma|_\Gamma^2 - |\gamma\gamma_0|_\Gamma^2\right|) \\ &\geq \frac{c_0}{2} \sum_{\gamma\in B, |\gamma|_\Gamma \leq n^{2/3}} \mu^{2n}(\gamma) \end{aligned} \tag{2.13}$$

The last inequality above follows from the fact that if $|\gamma|_\Gamma \leq n^{2/3}$

$$\begin{aligned} \left||\gamma|_\Gamma^2 - |\gamma\gamma_0|_\Gamma^2\right| &= \left||\gamma|_\Gamma - |\gamma\gamma_0|_\Gamma\right| \left||\gamma|_\Gamma + |\gamma\gamma_0|_\Gamma\right| \\ &\leq C(\gamma_0) \cdot n^{2/3} \end{aligned}$$

Then combining 2.13 and 2.12 we get, as soon as n is larger than some integer n_1 depending on γ_0 and B

$$\begin{aligned} \mathbb{P}(S_{2n} &\in \gamma_0 B) \geq \frac{c_0}{2}(\mathbb{P}(S_{2n} \in B) - \frac{\widetilde{C}}{100} \cdot \frac{|B|}{n^{d/2}}) \\ &\geq \frac{c_0\widetilde{C}}{4} \cdot \frac{|B|}{n^{d/2}}. \end{aligned} \tag{2.14}$$

Now let $x \in N$ be arbitrary. Since Γ is dense in N we can pick $\gamma_0 \in \Gamma$ inside the ball xB_r so that $\gamma_0 B_{r/2}$ is contained in xB_r. The lower bound (2.14) shows that for n large enough

$$\mathbb{P}(S_{2n} \in xB_r) \geq 2C_1 \cdot \frac{|B_r|}{n^{d/2}}.$$

where $C_1 > 0$ is an absolute constant such that $2C_1 \leq (c_0\widetilde{C}|B_r|)/(4|B_{r/2}|)$.

To pass from S_{2n} to S_n is now easily done.

$$\mathbb{P}(S_{2n+1} \in xB_r) = \int \mathbb{P}(S_{2n} \in y^{-1}xB_r)d\mu(y)$$

Since μ is finitely supported, for n large enough, the right hand side is also larger than $2C_1|B_r|/n^{d/2}$. This establishes the lower bound for translates.

2.2.4 General case

The general case follows immediately from the lower bounds for translates, because given a bounded open set O, one can always find finitely many disjoint balls B_i contained in O such that $|O| \leq 2(\sum |B_i|)$. Hence, for all large enough n

$$\mathbb{P}(S_n \in O) \geq C_1 \cdot \frac{|O|}{n^{d/2}}.$$

□

Remark 2.4 *This upper and lower estimate still do not prove the expected local limit theorem on N, i.e. the conjecture that in fact $n^{d(N)/2}\mathbb{P}(S_n \in B)$ converges to $c(\mu) \cdot |B|$ where $c(\mu) > 0$ is the value at identity of the heat kernel associated with μ by the central limit theorem. The above estimates show that every limit measure is an absolutely continuous Radon measure on N. Making use of an argument in the proof of a theorem of Guivarc'h (Theorem 3.6 below), and of the uniform upper bound above, it is possible to show that every limit measure is invariant by some abelian subgroup of N. It is reasonable to believe that a modification of Guivarc'h's argument would yield the invariance by the whole of N.*

2.3 Completion of the proof of Theorem 0.4

The proof will follow several steps. First we show, by use of the above uniform upper bound for translates (Theorem 0.1), that the sequence of measures $(\mu^n * \delta_x)_n$ is relatively compact in the space of probability measures on G/Γ. Then we show that any limit measure must be absolutely continuous with respect to the ergodic measure m_x. Making use of a theorem of Guivarc'h (Theorem 1.3.8 in Chapter 1 and 3.6 below) we then prove that any limit measure is N-invariant. These last two properties imply that the limit measure is unique and equal to m_x thus establishing the convergence of $(\mu^n * \delta_x)_n$.

Proposition 3.1 *Under the assumptions of Theorem 0.4, the sequence of measures $\mu^n * \delta_x$ is relatively compact in the space of probability measures on G/Γ. In other words, for every $x \in G/\Gamma$, and every $\varepsilon > 0$, there exists a compact $K \subset G/\Gamma$ such that $\mu^n * \delta_x(K) \geq 1 - \varepsilon$ as soon as n is large enough.*

Proposition 3.2 *Under the assumptions of Theorem 0.4, suppose that a subsequence of measures $\{\mu^{n_k} * \delta_x\}_k$ converges to a probability measure ν on G/Γ. Then ν is absolutely continuous with respect to m_x.*

The two last propositions will follow from the following result about random walks on N :

Proposition 3.3 *Under the assumptions of Theorem 0.1, for every $\varepsilon > 0$ there exists $\delta > 0$ such that if h denotes any right-uniformly continuous function on N such that $0 \leq h \leq 1$ and such that the following bound holds :*

$$\overline{\lim_{t \to +\infty}} \frac{\int_{|x| \leq t} h(x)dx}{vol(\{x \in N, |x| \leq t\})} \leq \delta \tag{2.15}$$

then we also have

$$\overline{\lim_{n \to +\infty}} \int h(x) d\mu^n(x) \leq \varepsilon$$

Proof: Let $\varepsilon > 0$ be given. Let S_n be the random walk in N associated with μ and starting at the identity. The proof makes use of the uniform upper bound for translates given in Theorem 0.1 and of the central limit theorem on N (see above Theorem 1.3.19 in Chapter 1).

Let h be any right-uniformly continuous function on N such that $0 \leq h \leq 1$. Let B_r be the ball of radius r in N centered at identity for some left-invariant metric on N. Let $\omega_h(r) = \sup_{u \in B_r} \|h(xu) - h(x)\|_\infty$. From the definition of h we have that $\lim_{r \to 0} \omega_h(r) = 0$. We need the following lemma :

Lemma 3.4 *Let G be a connected Lie group endowed with a left invariant metric. Let B_r be the ball of radius r centered at the identity element e in G. Then there exists an integer $m_0 = m_0(G) \geq 1$ such that, for every real number $r \in (0, 1)$ there exist a countable set of points $\{x_i\}_{i \in I}$ in G which gives rise to a covering*

$$G = \bigcup_{i \in I} x_i B_r$$

with multiplicity at most m_0.

Proof: The proof is essentially a Vitali covering argument. We look at the collection $\mathcal{C}$ of all translates of B_r. We take a maximal subset $\mathcal{S}$ of $\mathcal{C}$ consisting of pairwise disjoint balls. It must be a countable subset ; we denote by I the set of all centers of these balls :

$\mathcal{S} = \{x_i B_r, i \in I\}$. Since the metric is left invariant, xB_r is the ball of radius r centered at x. Now any $y \in G$ lies at distance less or equal to $2r$ from one of the x_i's. Indeed, if not then the ball yB_r would be disjoint from all $x_i B_r$'s, contradicting the maximality of $\mathcal{S}$. Hence

$$G = \bigcup_{i \in I} x_i B_{2r}$$

Finally, let us show that the multiplicity od the covering in bounded. Given $y \in G$, suppose that $y \in x_i B_{2r}$ for some $i \in I$. Then for any $z \in x_i B_r$, $d(y, z) \leq 3r$ hence $z \in yB_{3r}$, i.e. $x_i B_r \subset yB_{3r}$ or $y^{-1} x_i B_r \subset B_{3r}$. But all $y^{-1} x_i B_r$ are pairwise disjoint. If there are at least N different such i's then $N \cdot vol(B_r) \leq vol(B_{3r})$. But one can find positive constants C_1 and C_2 such that for all $r \leq 3/2$,

$$C_1 \cdot r^d \leq vol(B_r) \leq C_2 \cdot r^d$$

where $d = \dim(G)$. We conclude that $N \leq 3^d C_2/C_1$. Take m_0 to be the largest integer smaller or equal to $3^d C_2/C_1$. We are done. $\square$

Let $p(x)$ be the heat kernel associated to μ by the central limit theorem. This theorem shows in particular that for any positive $A > 0$, the following limit holds :

$$\lim_{n \to +\infty} \mathbb{P}(|S_n| \geq A\sqrt{n}) = \int_{|x| \geq A} p(x)dx \tag{2.16}$$

for any homogeneous norm $|\cdot|$ on N.

We fix once and for all $A > 0$ such that $\int_{|x| \geq A} p(x)dx \leq \varepsilon/(4m_0)$ where $m_0 = m_0(N)$ is the multiplicity in the above lemma for $G = N$.

Now, with the notations of lemma 3.4, write $\widetilde{h}(x) = \sum_{i \in I} h(x_i) \chi_{x_i B_r}(x)$. Then we have for all $x \in N$

$$h(x) \leq \widetilde{h}(x) + \omega_h(r)$$

Hence $\mathbb{E}(h(S_n)) \leq \mathbb{E}(\widetilde{h}(S_n)) + \omega_h(r)$. Then for any $A > 1$

$$\begin{aligned} \mathbb{E}(\widetilde{h}(S_n)) &= \sum h(x_i) \mathbb{P}(S_n \in x_i B_r) \\ &\leq \sum_{|x_i| \geq (A-1)\sqrt{n}} \mathbb{P}(S_n \in x_i B_r) + \sum_{|x_i| \leq (A-1)\sqrt{n}} h(x_i) \mathbb{P}(S_n \in x_i B_r) \end{aligned} \tag{2.17}$$

If $x \in x_i B_r$ and $|x_i| \geq (A-1)\sqrt{n}$ then $|x| \geq |x_i| - diam(B_r) - c_0$ for some constant $c_0 > 0$ (see above the almost subadditivity of the homogeneous norm in Chapter 1, 1.3.1) hence $|x| \geq A\sqrt{n}$ as soon as $r \leq 1$ and n is large enough. Therefore the first term in (2.17) satisfies :

$$\sum_{|x_i| \geq (A-1)\sqrt{n}} \mathbb{P}(S_n \in x_i B_r) \leq m_0 \cdot \mathbb{P}(|S_n| \geq A\sqrt{n}) \leq \varepsilon/3 \tag{2.18}$$

as soon as n is larger than an integer $n_1 \geq 0$ depending only on μ and determined by the convergence (2.16) in the central limit theorem. Note that the multiplicity m_0 is independent of the radius $r > 0$.

Let us now focus on the second term in (2.17). We can apply Theorem 0.1 to the random walk S_n and we obtain the existence of a constant $C > 0$ such that for every $r > 0$, there exists an integer $n(r) > 0$ such that whenever $n \geq n(r)$, we have for any $i \in I$

$$\mathbb{P}(S_n \in x_i B_r) \leq \frac{C \cdot |B_r|}{n^{d(N)/2}} \tag{2.19}$$

Hence

$$\sum_{|x_i| \leq (A-1)\sqrt{n}} h(x_i) \mathbb{P}(S_n \in x_i B_r) \leq \frac{C \cdot |B_r|}{n^{d(N)/2}} \cdot \sum_{|x_i| \leq (A-1)\sqrt{n}} h(x_i)$$

But for any $i \in I$ and $r > 0$,

$$h(x_i)|B_r| \leq \int h(x) \chi_{x_i B_r}(x) dx + |B_r| \omega_r(h)$$

Hence for all $n \geq n(r)$,

$$\begin{aligned} |B_r| \cdot \sum_{|x_i| \leq (A-1)\sqrt{n}} h(x_i) &\leq \sum_{|x_i| \leq (A-1)\sqrt{n}} \int h(x) \chi_{x_i B_r}(x) dx + |B_r| \omega_r(h) \cdot \#\{x_i, |x_i| \leq (A-1)\sqrt{n}\} \\ &\leq m_0 \int_{|x| \leq A\sqrt{n}} h(x) dx + m_0 \cdot \omega_r(h) \cdot vol(\{x \in N, |x| \leq A\sqrt{n}\}) \end{aligned}$$

But, from Guivarc'h's theorem on the growth of nilpotent Lie groups (Theorem 1.1.7 in Chapter 1), there is a constant $C_2 > 0$ such that whenever $A\sqrt{n} > 1$

$$vol(\{x \in N, |x| \leq A\sqrt{n}\}) \leq C_2 \cdot (A\sqrt{n})^{d(N)}$$

Hence

$$\sum_{|x_i| \leq (A-1)\sqrt{n}} h(x_i) \mathbb{P}(S_n \in x_i B_r) \leq m_0 C C_2 \cdot A^{d(N)} \cdot \left(\frac{\int_{|x| \leq A\sqrt{n}} h(x) dx}{vol(\{x \in N, |x| \leq A\sqrt{n}\})} + \omega_r(h) \right)$$

Now let $\delta := \varepsilon/(12 \times m_0 C C_2 \times A^{d(N)}) > 0$ and choose once and for all $r > 0$ so that $\omega_r(h) \times m_0 C C_2 \times A^{d(N)} \leq \varepsilon/6$ and $\omega_r(h) \leq \varepsilon/2$.

Suppose (2.15) holds for h and this δ. Then for n larger than some $n_2 = n_2(h, A)$ and larger than $n_1(r)$ we have

$$\sum_{|x_i| \leq (A-1)\sqrt{n}} h(x_i) \mathbb{P}(S_n \in x_i B_r) \leq \frac{\varepsilon}{6} + \frac{\varepsilon}{6} = \frac{\varepsilon}{3} \tag{2.20}$$

Putting together (2.20), (2.18) and (2.17) we obtain for all $n \geq \max\{n_1, n_2\}$

$$\mathbb{E}(h(S_n)) \leq \varepsilon$$

We are done. $\square$

Remark 3.5 *Obviously the proposition also holds for left-uniformly continuous functions on N.*

We now turn to the proof of the two propositions 3.1 and 3.2 that we had left behind : they are corollaries of Proposition 3.3.

Proof of Proposition 3.1. Let $x \in G/\Gamma$ and $\varepsilon > 0$ be given. From the Margulis-Dani-Shah recurrence theorem (see above Theorems 1.5.2, 1.5.3 and 1.5.4 in Chapter 1), there exists a compact subset K in G/Γ such that

$$\overline{\lim_{t \to +\infty}} \frac{vol\{g \in N, |g| \le t, g \cdot x \in K^c\}}{vol\{g \in N, |g| \le t\}} \le \delta \tag{2.21}$$

where K^c is the complement of K in G/Γ and $\delta > 0$ is the number obtained in Proposition 3.3. Let K' be another compact subset of G/Γ containing the interior of K, i.e. such that K and $\overline{(K')^c}$ are disjoint (this is always possible since G/Γ can be exhausted by compact subsets). By Uryshon's lemma we can construct a function ϕ on G/Γ which is continuous and such that

$$\chi_{(K')^c} \le \phi \le \chi_{K^c}$$

Now define $h(g) := \phi(g \cdot x)$, which turns out to be a left-uniformly continuous function on G since ϕ may vary only within the compact set K'. Clearly the estimate (2.21) yields the bound (2.15) for h. It follows from Proposition 3.3 that all large enough integer n

$$\mathbb{P}(S_n \cdot x \notin K') \le \mathbb{E}(\phi(S_n \cdot x)) \le \varepsilon$$

This is what we needed. $\square$

Proof of Proposition 3.2. What we have to do is to find, given an arbitrary $\varepsilon > 0$, a real number $\delta > 0$ such that whenever ϕ is a non-negative compactly supported function on G/Γ such that $\int \phi(y) dm_x(y) \le \delta$ then $\int \phi(y) d\nu(y) \le \varepsilon$. Let $h(g) := \phi(g \cdot x)$. By Shah's equidistribution theorem (see above Theorem 1.5.4 above in Chapter 1), we have

$$\lim_{t \to +\infty} \frac{\int_{|g| \le t} h(g) dg}{vol\{g \in N, |g| \le t\}} = \int \phi(y) dm_x(y)$$

Besides, h is a left-uniformly continuous function on N and we can apply Proposition 3.3. We obtain that

$$\mathbb{E}(\phi(S_n \cdot x)) \le \varepsilon$$

for all n large enough. Hence if ν is a limit probability measure, then $\int \phi(y) d\nu(y) \le \varepsilon$. $\square$

We now finish up the proof of Theorem 0.4. By Proposition 3.1, it suffices to show that every limiting probability measure ν, arising from a subsequence $\mu^{n_k} * \delta_x$ is equal to m_x. We will first show that ν is N-invariant. This is a consequence of the following theorem of Y. Guivarc'h. The same idea was used by Guivarc'h in his proof of equidistribution of random walks on nil-manifolds in [73]. Recall (Theorem 1.3.8 in Chapter 1) :

Theorem 3.6 ***(Guivarc'h)*** *Let μ be an aperiodic compactly supported probability measure on a nilpotent Lie group N. Let $f \in \mathbb{L}^1_0(N)$ an integrable function with zero integral on N. Then*
$$\lim_{n\to+\infty} \|\mu^n * f\|_{\mathbb{L}^1(N)} = 0$$

Let ϕ be a bounded continuous function on G/Γ. Let $\tau_g(\phi)$ be its translation by an element of $G : \tau_g(\phi)(x) = \phi(g \cdot x)$. This is again a bounded continuous function on G/Γ and we have the following convergence for every $g_0 \in G$:

$$\lim_{k\to+\infty} \int \tau_{g_0}(\phi)(g \cdot x)d\mu^{n_k}(g) = \int \tau_{g_0}(\phi)(y)d\nu(y)$$

Let $\psi(g_0)$ denote the limit on the right hand side. It is a continuous and bounded function on G. Its restriction to N is also continuous and bounded, we call it ψ again. Let f be any integrable function on N whose total integral vanishes, i.e.

$$\int_N f(h)dh = 0$$

By Lebesgue dominated convergence theorem, we obtain the following convergence :

$$\lim_{k\to+\infty} \int_N f(h) \int \tau_h(\phi)(g \cdot x)d\mu^{n_k}(g)dh = \int_N \psi(h)f(h)dh$$

On the other hand :

$$\begin{aligned}
\int_N f(h) \int \tau_h(\phi)(g \cdot x)d\mu^{n_k}(g)dh &= \int_N f(h) \int \phi(hg \cdot x)d\mu^{n_k}(g)dh \\
&= \int_N \int f(hg^{-1})\phi(h \cdot x)d\mu^{n_k}(g)dh \\
&= \int_N \phi(h \cdot x) \int f(hg^{-1})d\mu^{n_k}(g)dh
\end{aligned}$$

and

$$\begin{aligned}
\left| \int_N f(h) \int \tau_h(\phi)(g \cdot x)d\mu^{n_k}(g)dh \right| &\leq \|\phi\|_\infty \int_N \left| \int f(hg^{-1})d\mu^{n_k}(g) \right| dh \\
&\leq \|\phi\|_\infty \|f * \mu^{n_k}\|_{L^1(N)}
\end{aligned}$$

From Guivarc'h's Theorem 3.6 we have $\|f * \mu^n\|_{L^1(N)} \to 0$ as $n \to +\infty$. It follows that

$$\int_N f(h)\psi(h)dh = 0$$

for every choice of an integrable function $f \in \mathbb{L}^1(N)$ with $\int f = 0$. Since ψ is continuous, this implies that ψ is in fact a constant function on N. In particular

$$\int \tau_h(\phi)(y)d\nu(y) = \int \phi(y)d\nu(y)$$

for all $h \in N$ and all bounded continuous functions ϕ on G/Γ. This means that the measure ν is invariant under the action of N. Yet we know, from Proposition 3.2, that every limiting measure ν is absolutely continuous with respect to m_x, hence is of the form $d\nu = f \cdot dm_x$ for some non-negative function $f \in \mathbb{L}^1(m_x)$. Hence f is essentially N-invariant. The ergodicity of m_x under the action of N shows that f is essentially equal to the constant 1 and $\nu = m_x$. This concludes the proof. $\square$

Remark 3.7 *In the above proof of Proposition 3.3, we made use of the local uniform upper bound (Theorem 0.1), whose proof was based on the upper gaussian estimate for discrete nilpotent groups. However, Varopoulos' uniform upper bound* (2.8)*, combined with the central limit theorem for* Γ *(viewed as a lattice in a simply connected nilpotent group* $\widetilde{\Gamma}$*), yields in the same way a proof of this proposition.*

Chapitre 3

Local limit theorems and equidistribution of random walks on the Heisenberg group

The goal of this chapter is to show the local limit theorem for centered distributions on the Heisenberg group along with some of its refinements and applications. The local limit problem on non-commutative Lie groups has been studied by many authors in the last thirty or forty years (Ito-Kawada, Arnold-Krylov, Kazhdan, Bougerol, Le Page, Guivarc'h, Varopoulos, etc.). In the classical commutative case or in the compact group case, the local limit theorem is available under the weakest assumptions on the probability distribution (see [161] and [95]). Random walks on semisimple Lie groups as well as on nilpotent Lie groups have been extensively studied and in particular the problem of the local limit theorem. In a recent work, Alexopoulos [7] obtains a very precise local limit theorem and estimates à la Berry-Essen for distributions with a continuous density of compact support on an arbitrary connected Lie group of polynomial growth.

However, in this problem, an assumption of absolute continuity of the distribution with respect to the Haar measure is often made, while the case of a possibly singular (e.g. finitely supported) distribution remains generally open. Such cases include finitely supported distributions on the group of isometries of the Euclidean 3-space (see [96], [74]) or the case of semisimple groups [27]. Similarly the speed of convergence to equidistribution is not well understood and seems to depend on difficult arithmetic questions (cf. the spectral gap conjecture [147] for equidistribution on the sphere). In this paper however, we shall treat the case of an arbitrary, possibly singular, measure. We shall focus on the simplest nilpotent Lie group : the first Heisenberg group.

We shall also obtain a uniform version of the local limit theorem yielding a very precise estimate on the asymptotic behavior of centered random walks on the Heisenberg group. This generalizes a result of Stone [161] in the commutative case. This estimate allows to show further equidistribution results for random walks on homogeneous spaces. Following this strategy, we show at the end of the paper that centered Heisenberg-unipotent random walks on homogeneous spaces G/Γ, where Γ is discrete of finite co-volume in a Lie group G, are equidistributed in the closure of the orbit on which they live. This yields a "probabilistic equivalent" of M. Ratner's equidistribution theorem for orbits of unipotent subgroups in homogeneous spaces (see [138], [150], [160]). In the non-centered case, an interesting phenomenon can occur : non-centered unipotent random walks may not converge to any distribution on G/Γ. For instance if Γ is not co-compact, they may wander outside an arbitrary compact set with high probability at arbitrary large times. Hence the hypothesis that the walk should be centered is crucial (unless of course the Haar measure is uniquely ergodic for the unipotent subgroup).

Finally let us remark that the results of this paper should extend to the case of an arbitrary simply connected nilpotent Lie group, but the technical difficulty of the forthcoming proofs forced us to restrain our attention to the Heisenberg group.

3.1 Statement of the results

Let G be the group of 3×3 upper-triangular unipotent matrices and e the identity in G. Let us fix the Haar measure on G, $dg = dxdydz$ where

$$g = \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}$$

is simply denoted by $g = (x, y, z)$. We shall also denote by $|A|$ the Haar measure of a borelian set A and we fix a homogeneous norm $\|g\| = \max\{|x|, |y|, |z|^{1/2}\}$ on G.

We shall consider a probability measure μ on G with the following properties :

- compactly supported
- centered : $\int p(x)d\mu(x) = 0$ where $p : G \to G/[G,G]$ is the canonical map.
- aperiodic : for any proper closed subgroup $H \subsetneq G$ and any $x \in G$, $\mu(xH) < 1$

In particular, we make no assumption of smoothness for μ, which can be for instance finitely supported.

The convolution product of measures is denoted by $\mu * \nu$ and convolution powers are simply denoted by μ^n.

The central limit theorem for G is well known (see the work of Wehn [172], as well as Tutubalin [169] and Crépel-Raugi [142]). It states that if $(d_t)_t$ is the semigroup of dilations given by $d_t(x, y, z) = (tx, ty, t^2z)$ then the sequence $d_{\frac{1}{\sqrt{n}}}(\mu^n)$ converges to some gaussian measure ν on G (in the sense of probability measures, i.e. $\int f \circ d_{\frac{1}{\sqrt{n}}} d\mu^n \to \int f d\nu$ for every bounded continuous function on G). The measure ν lies inside a gaussian semigroup of probability measures $(\nu_t)_{t>0}$ defined by its generating distribution

$$\mathcal{A}f = \overline{z}\partial_z f(e) + \overline{xy}\partial^2_{xy} f(e) + \frac{1}{2}\overline{x^2}\partial^2_x f(e) + \frac{1}{2}\overline{y^2}\partial^2_y f(e) \tag{3.1}$$

where $\overline{z} = \int z d\mu(x,y,z)$ and $\overline{xy} = \int xy d\mu(x,y,z)$. Then $\nu = \nu_1$ and $\nu_t = p_t(g)dg$ where $p_t(g)$ is the heat kernel associated to the operator defined by (3.1) hence is a strictly positive fastly decreasing smooth function on G. It is the density of the Brownian Motion corresponding to $\mathcal{A}$ on G.

Let $c(\mu) = p_1(e) > 0$. For general references about gaussian semigroups and the Levy-Khintchin-Hunt formula, see [68], [127] and the original article of Hunt [90].

We shall say that μ satisfies *Cramer's condition* if

$$\sup_{t^2+s^2\geq 1} \left| \int e^{i(tx+sy)} d\mu(x,y,z) \right| < 1$$

We obtain the theorems below without this assumption, but at some point we get a slightly stronger result if this assumption is made (apparently for technical reasons, but we have no guess whether it is necessary).

Let us state the local limit theorem for G together with a uniform version for translates of a bounded set. These results generalize to the Heisenberg group the well known theorems of classical probability theory on $\mathbb{R}^n$ (see Bre and [161]). The method makes use of the unitary representations of G.

Theorem 1.1 *(Local limit theorem) Let μ be a compactly supported aperiodic centered probability measure on G. Let f be a continuous function on G which is compactly supported. Then the following convergence holds uniformly when z varies in compact subsets of G*

$$\lim_{n\to\infty} n^2 \int_G f(gz) d\mu^n(g) = c(\mu) \int_G f(g) dg$$

Theorem 1.2 *(Uniform local limit theorem) Let μ be a compactly supported aperiodic centered probability measure on G and $(\nu_t)_{t>0}$ the corresponding limit gaussian semigroup. Then for all bounded borelian $B \subset G$ with $|\partial B| = 0$,*

$$\lim_{n\to+\infty} \sup_{x\in G} n^2 \left|\mu^n(xB) - \nu_n(xB)\right| = 0$$

If we assume additionally Cramer's condition, then

$$\lim_{n\to+\infty} \sup_{x,y\in G} n^2 \left|\mu^n(xBy) - \nu_n(xBy)\right| = 0$$

Let us remark that the choice of $(\nu_t)_{t>0}$ depends on the choice we made of a semi-group of dilations $(d_t)_{t>0}$. Any other choice $(d'_t)_{t>0}$ for the semi-group of dilations is of the form $\phi \circ d_t \circ \phi^{-1}$ for some automorphism ϕ of G. The associated gaussian semi-group $(\nu'_t)_{t>0}$ is obtained from $(\nu_t)_{t>0}$ by composing by some automorphism of G and theorem 1.2 remains valid if we take $(\nu'_t)_{t>0}$ instead of $(\nu_t)_{t>0}$.

Theorem 1.3 *(Concentration function) Under the assumptions above for μ, for every bounded set $K \subset G$, there is a constant C_K such that for all integers n*

$$\sup_{x\in G} \mu^n(xK) \leq \frac{C_K}{n^2}$$

If we suppose additionally Cramer's condition, then

$$\sup_{x,y\in G} \mu^n(xKy) \leq \frac{C_K}{n^2}$$

This uniform version of the local limit theorem for translates enables to show the following corollary. The point of this result is that the function f is not assumed to tend to 0 at infinity, and in particular, can have a periodic type of behavior. For the classical commutative case see Chapter 4.

Corollary 1.4 *Let f be an arbitrary left-uniformly continuous function on G such that the following limit exists*

$$\lim_{T_x\to\pm\infty, T_y\to\pm\infty, T_z\to\pm\infty} \frac{1}{T_xT_yT_z} \int_0^{T_x} \int_0^{T_y} \int_0^{T_z} f(x,y,z)dxdydz = \ell \tag{3.2}$$

where $\ell \in \mathbb{C}$. Then

$$\lim_{n\to\infty} \int_G f(g)d\mu^n(g) = \ell$$

This corollary enables to prove further equidistribution results on homogeneous spaces. In the last section we show how to derive a "probabilistic Ratner's theorem" for Heisenberg-unipotent random walks on homogeneous spaces, such as horospheric random walks on complex hyperbolic manifolds, namely,

Theorem 1.5 *Let G be a connected Lie group and Γ a lattice in G. Let H be a closed subgroup of G consisting of unipotent elements and isomorphic to the Heisenberg group. Let μ be a centered compactly supported aperiodic probability measure on H. Then for an arbitrary $x \in G/\Gamma$ and for any bounded and continuous function f on G/Γ,*

$$\lim_{n\to+\infty} \int_H f(hx)d\mu^n(h) = \int_G f(g)dm_x(g)$$

where m_x is the unique H-invariant ergodic probability measure on G/Γ whose support is the closure of the orbit $\overline{Hx}$.

The existence of the measure m_x is given by Ratner's theorem (see [Rat1 to 3], [160]). The corresponding deterministic result was proved for one parameter subgroups by Ratner (cf. [138]) elaborating on a weaker qualitative recurrence result due to Margulis and generalized by Dani, and for general unipotent groups by Shah (cf. [150]).

Note that this also implies that the only μ-stationary probability measures on G/Γ (i.e. the measures ν such that $\mu * \nu = \nu$) are the H-invariant ones. This means, in the terminology of Furstenberg (cf. [59]), that the action of H on G/Γ is *stiff* with respect to μ. But this also follows from the fact that (H, μ) has the Choquet-Deny property (i.e. the absence of bounded μ-harmonic functions) as follows from the work of Guivarc'h [73].

Also note that according to a general random ergodic theorem of Oseledec (cf. [130]), which is proved in the case when μ is symmetric (i.e. $\mu(A) = \mu(A^{-1})$), the convergence of theorem 1.5 above holds for m-almost every x in G/Γ for any H-ergodic probability measure m on G/Γ.

In the above theorem, the assumption that μ is centered cannot be removed. A simple use of the central limit theorem shows that for certain lattices in $\mathbb{R}^2$, any non centered unipotent random walk starting at that point in the space of lattices $SL_2(\mathbb{R})/SL_2(\mathbb{Z})$ will diverge, i.e. may remain very far with high probability at some arbitrary large time.

When H is uniquely ergodic on G/Γ (e.g. the horocycle flow on a compact Riemann surface), then theorem 1.5 follows easily from Théorème V.5. in [73] and holds even when μ is not centered.

This application was originally motivated by the work of Eskin and Margulis [54], where they studied the case of random walks on G/Γ obtained by a measure μ whose support is Zariski dense in a semisimple group. Their main result is that the sequence $\mu^n * \delta_x$ is relatively compact in the space of probability measures on G/Γ.

3.2 Notations and outline of the paper

We keep the notations and terminology introduced in the last section. Let ρ be the regular representation of G on the functions on $G : \rho(g)f(x) = f(g^{-1}x)$.

G is identified with its Lie algebra $\mathfrak{g}$ by writing $g = (x, y, z)$ with the help of the diffeomorphism

$$\begin{array}{rcl} \mathfrak{g} & \to & G \\ xX + yY + zZ & \mapsto & e^{yY}e^{xX}e^{zZ} \end{array} \tag{3.3}$$

where X, Y and Z are the upper triangular elementary matrices, with $[X, Y] = Z$.

The product in G is given by

$$gg' = (x + x', y + y', z + z' + xy')$$

In the sequel, we will need to look at possible other parametrizations of G, in particular at those of the form

$$\begin{array}{rcl} \phi_\sigma & : & \mathbb{R}^3 \to G \\ (x, y, z) & \mapsto & (x, y, z + \sigma(x, y)) \end{array}$$

where $\sigma(x,y)$ is a quadratic form in x and y. The Fourier transform of a function on G can be defined in different ways depending on the choice of a parametrization. Given a function $f : G \to \mathbb{C}$, we shall denote by $F_\sigma(f) : (\mathbb{R}^3)^\vee \to \mathbb{C}$ the Fourier transform taken in the parametrization defined by ϕ_σ. When $\sigma = 0$, we simply write $F_0(f) = \widehat{f}$. The variable in the dual space $(\mathbb{R}^3)^\vee$ will be denoted by $\xi = (t, s, \lambda)$. The formula reads :

$$\begin{aligned} F_\sigma(f) &: (\mathbb{R}^3)^\vee \to \mathbb{C} \\ \xi &= (t,s,\lambda) \mapsto \frac{1}{(2\pi)^{3/2}} \int f(\phi_\sigma(x,y,z)) e^{i(tx+sy+\lambda z)} dxdydz \end{aligned} \tag{3.4}$$

Let $C_c(G)$ be the space of continuous and compactly supported functions on G. Let μ be a probability measure satisfying the properties listed in the introduction and (X, Y, Z) a random variable on G distributed according to μ. We set once and for all

$$\alpha = \frac{\mathbb{E}(XY)}{2\mathbb{E}(Y^2)} \tag{3.5}$$

Theorem 2.1 *Let $\sigma(x,y) = \alpha y^2$ and F_σ the Fourier transform just defined. Let f and g be two functions on G with $f \in C_c(G)$ and $F_\sigma(g) \in C_c(G)$. Then*

$$\lim_{n\to\infty} n^2 \langle \rho(\mu^n) f, g \rangle = c(\mu) \int_G f \int_G \overline{g}$$

and, if we suppose additionally that $F_\sigma(g)$ is absolutely continuous, then uniformly when z varies in compact subsets,

$$\lim_{n\to\infty} n^2 \int g(z^{-1}x) d\mu^n(x) = c(\mu) \int_G g$$

Below, we derive theorem 1.1 from theorem 2.1. The strategy for proving theorem 2.1 follows the general scheme provided by Stone's proof of the local limit theorem on $\mathbb{R}^d$ [161] and is as follows. Looking at the Fourier transform of the integral, we give an explicit decomposition of the regular representation of G into a continuous direct sum of primary representations and treat each part of the integral to show that only the part with small λ's and small s and t's gives a contribution. Then we gain control on this part by showing a domination condition on the integrand. This is achieved by performing a Taylor expansion in s, t, λ of the characteristic function of μ^n. Lebesgue's dominated convergence theorem combined with the point-wise convergence granted by the central limit theorem on G completes the proof. The proof of theorem (1.2) makes use of the estimates previously obtained and goes along similar lines.

3.3 Irreducible unitary representations of G

The irreducible unitary representations of G are well known. Apart from characters, there is a one-parameter family of irreducible unitary representations π_λ ($\lambda \in \mathbb{R}\backslash\{0\}$)

modeled on $\mathbb{L}^2(\mathbb{R})$ by

$$\pi_\lambda(g)f(t) = e^{i(tx+\lambda z)}f(t+\lambda y)$$

The following two propositions will be crucial in the proof. The first is quite standard (see [68], and [73]) :

Proposition 3.1 *Let μ be an aperiodic probability measure on G. Then for any closed interval $I \subset \mathbb{R}\backslash\{0\}$*

$$\sup_{\lambda\in I} \|\pi_\lambda(\mu)\| < 1 \tag{3.6}$$

Proof: Let $\lambda_n \to \lambda \in I$ such that

$$\left\langle \pi_{\lambda_n}(\mu * \mu^{-1})f_n(t), f_n(t)\right\rangle \to 1$$

for some sequence of vectors $f_n \in \mathbb{L}^2(\mathbb{R})$ of norm 1. Then, up to taking a subsequence, for $\mu * \mu^{-1}$-almost every x,

$$\langle \pi_{\lambda_n}(x)f_n(t), f_n(t)\rangle \to 1$$

But $\Gamma = \{x \in G, \langle \pi_{\lambda_n}(x)f_n(t), f_n(t)\rangle \to 1\}$ is clearly a subgroup of G, and $\mu*\mu^{-1}(\Gamma) = 1$. Since μ is aperiodic, Γ is dense in G, hence $[\Gamma,\Gamma]$ dense in the center of G. In particular, we can find $(0,0,z) \in \Gamma$, such that $\lambda z \notin 2\pi\mathbb{Z}$. Then

$$\left\|e^{i\lambda_n z}f_n - f_n\right\| \to 0$$

which implies $e^{i\lambda z} = 1$ and provides the desired contradiction. $\square$

The next proposition gives an estimate of the norm of the operators $\pi_\lambda(\mu)$. When μ is taken to be the symmetric Dirac measure on $(1,0,0)$ and $(0,1,0)$, this operator can be viewed as acting on $\ell^2(\mathbb{Z})$ and then coincide with the well-known Harper operator (see [19]) studied in mathematical physics.

Proposition 3.2 *Let μ be a probability measure on G whose support is not contained in a coset of an abelian subgroup of G. Then we have*

$$\liminf_{\lambda\to 0} \frac{1-\|\pi_\lambda(\mu)\|}{|\lambda|} > 0 \tag{3.7}$$

Proof: Note that if we define $\mu^{-1}(B) = \mu(B^{-1})$ for every borelian subset B of G, then $\pi_\lambda(\mu^{-1})$ is the adjoint of $\pi_\lambda(\mu)$ and $\pi_\lambda(\mu^{-1} * \mu)$ is self-adjoint and non-negative.

If (7.3.1) does not hold, then we can find unit vectors $f_\lambda \in \mathbb{L}^2(\mathbb{R})$ such that for arbitrarily small λ's

$$\left\langle \pi_\lambda(\mu^{-1} * \mu)f_\lambda, f_\lambda\right\rangle \geq 1 - |o(\lambda)|$$

From the assumption made on μ, we can find two non-commuting elements x_0 and x_1 lying in the support of $\mu^{-1} * \mu$. For each λ we can then find x_0^λ close to x_0 and x_1^λ close to x_1 (i.e. $\left\|x_1^\lambda - x_1\right\|$ and $\left\|x_0^\lambda - x_0\right\| < \|x_1 - x_0\|/3$) such that for $i = 0, 1$

$$Re\left\langle \pi_\lambda(x_i^\lambda)f_\lambda, f_\lambda\right\rangle \geq 1 - |o(\lambda)|$$

The commutator $(x_0^\lambda, x_1^\lambda)$ belongs to the center, hence is of the form $(0, 0, c_\lambda)$. From the choice of x_i^λ it follows that $c < c_\lambda < 1/c$ for some constant $c \in (0, 1)$. From the Stone-Von Neumann theorem (cf. [43]) we can then find an isometry I_λ of $\mathbb{L}^2(\mathbb{R})$ such that, conjugating by I_λ, $\pi_\lambda(x_0^\lambda)$ is turned into the translation by $c_\lambda\lambda$ and $\pi_\lambda(x_1^\lambda)$ is turned into the multiplication by e^{it}. Hence we can assume that for arbitrarily small λ's

$$\begin{aligned} Re\,\langle f_\lambda(t + c_\lambda\lambda), f_\lambda(t)\rangle &\geq 1 - |o(\lambda)| \\ Re\,\langle e^{it} f_\lambda(t), f_\lambda(t)\rangle &\geq 1 - |o(\lambda)| \end{aligned}$$

Or equivalently

$$\|f_\lambda(t + c_\lambda\lambda) - f_\lambda(t)\| = o(\sqrt{|\lambda|}) \tag{3.8}$$

$$\left\|e^{it} f_\lambda(t) - f_\lambda(t)\right\| = o(\sqrt{|\lambda|}) \tag{3.9}$$

Let $A_\varepsilon = \{t \in \mathbb{R}, d(t, 2\pi\mathbb{Z}) < \varepsilon\}$. We deduce from (3.9) that

$$\int_{A_\varepsilon^c} |f_\lambda(t)|^2 dt = o(|\lambda|/\varepsilon^2) \tag{3.10}$$

and from (3.8) that for any positive integer n

$$\|f_\lambda(t + nc_\lambda\lambda) - f_\lambda(t)\| = o(n\sqrt{|\lambda|})$$

or

$$Re\,\langle f_\lambda(t + nc_\lambda\lambda), f_\lambda(t)\rangle \geq 1 - o(n^2\lambda) \tag{3.11}$$

We now take $n = [\frac{1}{2\sqrt{|\lambda|}}] + 1$ and $\varepsilon = c_\lambda\sqrt{|\lambda|}/12$. Then for small enough λ we have $1 > |nc_\lambda\lambda| > 3\varepsilon$. So for small λ and as soon as $\varepsilon < (\pi - 1)/2$, making use of (3.10) and applying the Cauchy-Schwarz inequality we have

$$\begin{aligned} |\langle f_\lambda(t + nc_\lambda\lambda), f_\lambda(t)\rangle| &\leq \sqrt{\int_{A_\varepsilon} |f_\lambda(t + nc_\lambda\lambda)|^2 \int_{A_\varepsilon} |f_\lambda(t)|^2} + \\ &\quad \sqrt{\int_{A_\varepsilon^c} |f_\lambda(t + nc_\lambda\lambda)|^2 \int_{A_\varepsilon^c} |f_\lambda(t)|^2} \\ &\leq 2\sqrt{\int_{A_\varepsilon^c} |f_\lambda(t)|^2} = o(\sqrt{|\lambda|}/\varepsilon) = o(1) \end{aligned}$$

which yields the desired contradiction with (3.11). $\square$

3.4 Reducing to small values of λ

Here we will begin the proof of theorem (2.1). The center of G is the one parameter subgroup $H = e^{\mathbb{R}Z}$. We are going to decompose the regular representation of G into a continuous direct sum of other unitary representations. Every character of H is determined by a number $\lambda \in \widehat{H} \cong \mathbb{R}^\vee$. If $f \in \mathbb{L}^1(G)$, we define for $\mathbf{x} \in G$

$$f_\lambda(\mathbf{x}) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(\mathbf{x}e^{zZ}) e^{i\lambda z} dz$$

We check that $f_\lambda(\mathbf{x}e^{zZ}) = e^{-i\lambda z} f_\lambda(\mathbf{x})$. By the Fourier isometry, if $z \mapsto f(\mathbf{x}e^{zZ})$ is in $\mathbb{L}^2(\mathbb{R})$, then $\lambda \mapsto f_\lambda(\mathbf{x})$ is in $\mathbb{L}^2(\mathbb{R}^\vee)$, and

$$\int_{\mathbb{R}} |f(\mathbf{x}e^{zZ})|^2 dz = \int_{\mathbb{R}^\vee} |f_\lambda(\mathbf{x})|^2 d\lambda$$

By Fubini's theorem, it follows that, if f also belongs to $\mathbb{L}^2(G)$ then $|f_\lambda(\mathbf{x})|$ is in $\mathbb{L}^2(G/H)$ for almost every λ. Let us write $\mathcal{H}_\lambda$ the Hilbert space of measurable functions F on G such that $F(\mathbf{x}e^{zZ}) = e^{-i\lambda z} F(\mathbf{x})$ and $|F(\mathbf{x})|$ is square integrable on G/H. Then $\mathcal{H}_\lambda$ is a realization of the induced representation $\rho_\lambda = Ind_H^G \lambda$, where G acts by left translations. The above Plancherel formula for f_λ shows that we have the continuous sum decomposition

$$\rho = \int^\oplus \rho_\lambda d\lambda$$

and if f, g belong to $\mathbb{L}^2(G)$

$$\langle \rho(\mu^n) f, g \rangle = \int \langle \rho_\lambda(\mu^n) f_\lambda, g_\lambda \rangle_{\mathcal{H}_\lambda} d\lambda$$

It is easy to see that ρ_λ is a primary unitary representation of G and that the representation π_λ defined above is the only irreducible representation of G contained in ρ_λ. Moreover, its multiplicity is infinite :

$$\rho_\lambda = \int^\oplus \pi_\lambda(s) ds \tag{3.12}$$

with $\pi_\lambda(s) \simeq \pi_\lambda$ for all s.

Now from Proposition (3.1) and from (3.12) we obtain that for some $\varepsilon \in (0,1)$ there exists some $c > 0$ such that for all $\lambda \in \mathbb{R}$ with $|\lambda| \leq \varepsilon$ and all $n \in \mathbb{N}$

$$\|\rho_\lambda(\mu^n)\| \leq \|\rho_\lambda(\mu)\|^n \leq e^{-c|\lambda|n} \tag{3.13}$$

therefore for any integer $k_0 \geq 3$, whenever $D > k_0/c$ we have

$$\begin{aligned}
\left| n^2 \int_{D\frac{\log n}{n} \leq |\lambda| \leq \varepsilon} \langle \rho_\lambda(\mu^n) f_\lambda, g_\lambda \rangle_{\mathcal{H}_\lambda} \, d\lambda \right| &\leq n^2 \int_{D\frac{\log n}{n} \leq |\lambda| \leq \varepsilon} e^{-k_0 \log n} \, \|f_\lambda\|_{\mathcal{H}_\lambda} \, \|g_\lambda\|_{\mathcal{H}_\lambda} \, d\lambda \\
&\leq \frac{1}{n^{k_0-2}} \int_{\lambda \in \mathbb{R}} \|f_\lambda\|_{\mathcal{H}_\lambda} \, \|g_\lambda\|_{\mathcal{H}_\lambda} \, d\lambda \\
&\leq \frac{1}{n^{k_0-2}} \, \|f\|_{L^2(G)} \, \|g\|_{L^2(G)} \to 0 \text{ as } n \to \infty
\end{aligned} \tag{3.14}$$

The last step follows from the Cauchy-Schwarz inequality. Hence this part of the integral tends to 0. Similarly if I denotes any of the intervals $[-A, -\varepsilon]$ or $[\varepsilon, A]$ where A is some positive number such that g_λ is identically zero outside $[-A, A]$, then it follows from Proposition (3.1) and from (3.12) that there is some constant $\beta \in (0, 1)$ such that

$$\sup_{\lambda \in I} \|\rho_\lambda(\mu^n)\| \leq \beta^n$$

Hence

$$\begin{aligned}
\left| n^2 \int_{\lambda \in I} \langle \rho_\lambda(\mu^n) f_\lambda, g_\lambda \rangle \, d\lambda \right| &\leq n^2 \beta^n \int_{\lambda \in \mathbb{R}} \|f_\lambda\|_{\mathcal{H}_\lambda} \, \|g_\lambda\|_{\mathcal{H}_\lambda} \, d\lambda \\
&\leq n^2 \beta^n \, \|f\|_{L^2(G)} \, \|g\|_{L^2(G)} \\
&\leq \frac{C}{n^{k_0-2}} \, \|f\|_{L^2(G)} \, \|g\|_{L^2(G)}
\end{aligned} \tag{3.15}$$

for some constant $C > 0$ depending on β. The right hand side clearly tends to 0 as n tends to infinity.

Now observe that if $F_\sigma(g)$ (defined in theorem 2.1) has compact support, then g_λ is identically zero outside some bounded set of values of λ. More precisely, if for some fixed λ, $F_\sigma(g)(t, s, \lambda) = 0$ for all t and s, then g_λ vanishes identically. The last argument allows then to reduce to small values of λ (i.e. less that $D\frac{\log n}{n}$ for some fixed $D > 0$).

We are now going to perform Fourier integration one step further, i.e. on G/H. We fix an arbitrary Borel section $\overline{\sigma} : G/H \to G$ (given in the above coordinates by $(x, y) \mapsto (x, y, \sigma(x, y))$ where σ is some measurable function on $\mathbb{R}^2$). Then $f_\lambda(\overline{\sigma}(\overline{\mathbf{x}}))$ is in $\mathbb{L}^2(G/H)$ for almost every λ.

Let now f and g be two functions in $\mathbb{L}^1(G)\cap\mathbb{L}^2(G)$ and $\mathbf{y}$ be given in G. We have

$$\begin{aligned}
\langle\rho(\mathbf{y})f,g\rangle &= \int_{G/H}\int_H f(\mathbf{y}^{-1}\mathbf{x}e^{zZ})\overline{g}(\mathbf{x}e^{zZ})dzd\overline{\mathbf{x}}\\
&= \int_{G/H}\int_{\mathbb{R}^\vee} f_\lambda(\mathbf{y}^{-1}\mathbf{x})\overline{g}_\lambda(\mathbf{x})d\lambda d\overline{\mathbf{x}}\\
&= \int_{\mathbb{R}^\vee}\int_{G/H} f_\lambda(\mathbf{y}^{-1}\mathbf{x})\overline{g}_\lambda(\mathbf{x})d\overline{\mathbf{x}}d\lambda\\
&= \int_{\mathbb{R}^\vee}\int_{\mathbb{R}^2} f_\lambda(\mathbf{y}^{-1}\overline{\sigma}(x,y))\overline{g}_\lambda(\overline{\sigma}(x,y))dxdyd\lambda\\
&= \int_{\mathbb{R}^\vee}\left\langle\rho_\lambda(\mathbf{y}^{-1})f_\lambda,g_\lambda\right\rangle_{\mathcal{H}_\lambda}d\lambda
\end{aligned}$$

For notational convenience we set $\phi_{\mathbf{y},\sigma,\lambda}(x,y):=f_\lambda(\mathbf{y}^{-1}\overline{\sigma}(x,y))$ and $\psi_{\sigma,\lambda}(x,y):=\overline{g}_\lambda(\overline{\sigma}(x,y))$. For almost all $\lambda\in\mathbb{R}$, these functions belong to $\mathbb{L}^2(\mathbb{R}^2)$. Performing Fourier transform on $\mathbb{L}^2(\mathbb{R}^2)$ now we get :

$$\langle\rho(\mathbf{y})f,g\rangle=\int_{\mathbb{R}^\vee}\int_{\mathbb{R}^{2\vee}}\widehat{\phi_{\mathbf{y},\sigma,\lambda}}(t,s)\overline{\widehat{\psi_{\sigma,\lambda}}}(t,s)dtdsd\lambda$$

Now a straightforward computation yields that $\widehat{\psi_{\sigma,\lambda}}(t,s)$ is the Fourier transform at (t,s,λ) of g^σ defined by

$$g^\sigma:=(x,y,z)\mapsto g(x,y,\sigma(x,y)+z) \tag{3.16}$$

$$\widehat{\psi_{\sigma,\lambda}}(t,s)=\widehat{g^\sigma}(t,s,\lambda)$$

Similarly if $\mathbf{y}=(\mathbf{y}_x,\mathbf{y}_y,\mathbf{y}_z)$ we compute,

$$\widehat{\phi_{\mathbf{y},\sigma,\lambda}}(t,s)=\frac{1}{(2\pi)^{3/2}}\int e^{i(\mathbf{y}_x(t+\lambda y)+\mathbf{y}_ys+\lambda(\mathbf{y}_z-\sigma(x+\mathbf{y}_x,y+\mathbf{y}_y)}e^{i(tx+sy+\lambda z)}f(x,y,z)dxdydz$$

Hence, if we suppose additionally that f has compact support on G and $F_\sigma(g)=\widehat{g^\sigma}$ has compact support on $\mathbb{R}^3$, we obtain

$$\langle\rho(\mathbf{y})f,g\rangle=\frac{1}{(2\pi)^{3/2}}\int d\mathbf{x}f(\mathbf{x})\int_{\mathbb{R}^\vee}\int_{\mathbb{R}^{2\vee}}e^{i[\mathbf{y}_x(t+\lambda y)+\mathbf{y}_ys+\lambda(\mathbf{y}_z-\sigma(x+\mathbf{y}_x,y+\mathbf{y}_y))]}e^{i(tx+sy+\lambda z)}\overline{\widehat{g^\sigma}}(t,s,\lambda)d\xi$$

in other words,

$$\langle\rho(\mathbf{y})f,g\rangle=\int_G\frac{d\mathbf{x}}{(2\pi)^{3/2}}f(\mathbf{x})\left\langle e^{i\psi(\xi,\mathbf{y},\mathbf{x})},\widehat{g^\sigma}(t-\lambda y,s,\lambda)\right\rangle_{\mathbb{R}^3}$$

where $\mathbf{x}=(x,y,z)$, $\mathbf{y}=(\mathbf{y}_x,\mathbf{y}_y,\mathbf{y}_z)$, $\xi=(t,s,\lambda)$ and $\psi(\xi,\mathbf{y},\mathbf{x})=t(\mathbf{y}_x+x)+s(\mathbf{y}_y+y)+\lambda(z-xy+\mathbf{y}_z-\sigma(x+\mathbf{y}_x,y+\mathbf{y}_y))$.

Let S is a random variable in G with distribution μ^n. The quantity we are interested in is

$$n^2\mathbb{E}\left(\langle \rho(S)f, g\rangle\right) = n^2 \int_G \frac{d\mathbf{x}}{(2\pi)^{3/2}} f(\mathbf{x}) \left\langle \mathbb{E}(e^{i\psi(\xi,S,\mathbf{x})}), \widehat{g^\sigma}(t-\lambda y, s, \lambda)\right\rangle_{\mathbb{R}^3}$$

Let $D > 0$. We split it in two parts and write $n^2\mathbb{E}\left(\langle \rho(S)f, g\rangle\right) = A_n + B_n$ where

$$\begin{aligned} A_n &= \int_G \frac{d\mathbf{x}}{(2\pi)^{3/2}} f(\mathbf{x}) J_n(\mathbf{x}) \\ B_n &= \int_G \frac{d\mathbf{x}}{(2\pi)^{3/2}} f(\mathbf{x}) I_n(\mathbf{x}) \end{aligned}$$

and

$$J_n(\mathbf{x}) = n^2 \int_{|\lambda|\geq D\frac{\log n}{n}} \left\langle \mathbb{E}(e^{i\psi(\xi,S,\mathbf{x})}), \widehat{g^\sigma}(t-\lambda y, s, \lambda)\right\rangle_{\mathbb{L}^2(\mathbb{R}^2)} d\lambda$$

$$I_n(\mathbf{x}) = n^2 \int_{|\lambda|\leq D\frac{\log n}{n}} \left\langle \mathbb{E}(e^{i\psi(\xi,S,\mathbf{x})}), \widehat{g^\sigma}(t-\lambda y, s, \lambda)\right\rangle_{\mathbb{L}^2(\mathbb{R}^2)} d\lambda$$

The part A_n has already been dealt with, because the above computations show that

$$A_n = n^2 \int_{|\lambda|\geq D\frac{\log n}{n}} \langle \rho_\lambda(\mu^n) f_\lambda, g_\lambda\rangle_{\mathcal{H}_\lambda} d\lambda$$

and (applying (3.14) and (3.15)) there is $C \geq 0$ (depending on D, μ and the size of the set $\{\lambda, \exists (t,s), F_g(g)(s,t,\lambda) \neq 0\}$) such that if $n \geq 1$

$$|A_n| \leq \frac{C}{n^{k_0-2}} \|f\|_{L^2(G)} \|g\|_{L^2(G)} \tag{3.17}$$

which tends to zero as soon as k_0 is taken such that $D \geq k_0/c \geq 3/c$ (where c was the constant defined in (3.13)).

Hence in the sequel, fixing $\mathbf{x} = (x,y,z) \in G$, we shall focus on the term $I_n(\mathbf{x})$. Before going further, we shall fix once and for all the section σ. We take it of the form proposed in theorem 2.1, that is $\sigma(x,y) = \alpha y^2$ where α is defined in terms of the moments of μ in equation (3.5). In this case,

$$\begin{aligned} \mathbb{E}(e^{i\psi(\xi,S,\mathbf{x})}) &= \mathbb{E}(e^{i[t(S_x+x)+s(S_y+y)+\lambda(z+S_z-\sigma(x+S_x,y+S_y))]}) \\ &= e^{i(tx+sy+\lambda z)}\mathbb{E}(e^{i\left[tS_x+sS_y+\lambda S_z-\lambda\alpha(y+S_y)^2\right]}) \\ &= e^{i(tx+sy+\lambda(z-\alpha y^2))}\mathbb{E}(e^{i\left[tS_x+(s-2\alpha\lambda y)S_y+\lambda(S_z-\alpha S_y^2)\right]} \end{aligned}$$

Hence,

$$I_n(\mathbf{x}) = n^2 \int_{|\lambda|\leq D\frac{\log n}{n}} \left\langle \mathbb{E}(\theta_n(t,s,\lambda)), e^{-i\phi}\widehat{g^\sigma}(t-\lambda y, s+2\alpha\lambda y, \lambda)\right\rangle_{\mathbb{L}^2(\mathbb{R}^2)} d\lambda$$

where $\phi = tx + sy + \lambda(z - xy + \alpha y^2)$ and

$$\theta_n(t,s,\lambda) = e^{i\left[tS_x + sS_y + \lambda(S_z - \alpha S_y^2)\right]} \tag{3.18}$$

We shall next estimate $\mathbb{E}(\theta_n(t,s,\lambda))$ for small values of λ.

3.5 Evaluation of the integral for small λ

Let us fix $D > 0$.

The remainder of this section is devoted to finding a suitable bound for the expectation $\mathbb{E}(\theta_n(t,s,\lambda))$ when $|\lambda| \leq D\frac{\log n}{n}$ and s and t take values bounded away from 0 and infinity.

We shall need the following lemma about moderate deviations. Since we could not find a reference for this precise form of the estimate we want (about the maximum of the random walk up to time n), we include a proof. It follows the well known argument of Cramer via Laplace transforms.

Lemma 5.1 *Let $U_0 = 0$ and $U_n = Y_1 + ... + Y_n$ be a sum of independent identically distributed real random variables Y_n, which are assumed of compact support and centered. Let $\mathcal{A}_n$ be the event $\{\max_{0\leq k\leq n}|U_k| \leq \sqrt{n}\log n\}$ and let $\mathcal{A}_n^c$ be the complementary event. Then for every non-negative integer p, there are constants $c_p > 0$ and $C_p > 0$ such that for all integers $n \in \mathbb{N}$*

$$\mathbb{E}(\max_{0\leq k\leq n}|U_k|^p \mathbf{1}_{\mathcal{A}_n^c}) \leq C_p e^{-c_p \log^2 n}$$

Proof: Since Y has compact support, if $D > 0$ is a bound for the support, we obviously get

$$|U_n| \leq Dn$$

for all n. Hence for any fixed $p \geq 0$

$$\mathbb{E}(\max_{0\leq k\leq n}|U_k|^p \mathbf{1}_{\mathcal{A}_n^c}) \leq D^p n^p \mathbb{P}(\mathcal{A}_n^c)$$

Therefore it is enough to show the lemma when $p = 0$.

We can assume that Y_1 is not identically 0. Define the Laplace transform $\Lambda(\lambda)$ of Y_1 by $e^{\Lambda(\lambda)} = \mathbb{E}(e^{\lambda Y_1})$ for a positive real $\lambda > 0$. Then define the Fenchel transform $\Lambda^*(x) = \sup_{\lambda>0}(\lambda x - \Lambda(\lambda))$ for a given $x > 0$. Clearly, $\Lambda^*(x)$ is a non-decreasing function of $x > 0$. The function $\Lambda(\lambda)$ is strictly convex, since its second derivative $\Lambda''(\lambda) = e^{-2\Lambda(\lambda)}(\mathbb{E}(e^{\lambda Y})\mathbb{E}(Y^2 e^{\lambda Y}) - \mathbb{E}(Ye^{\lambda Y})^2)$ is > 0 from the Cauchy-Schwarz inequality. In particular the supremum $\Lambda^*(x)$ is attained for a unique value λ_x of λ given by the equation $\Lambda'(\lambda_x) = x$. And if $x > 0$ then $\lambda_x > 0$ because $\Lambda'(0) = 0$ since Y is centered. Differentiating the relation $\Lambda'(\lambda_x) = x$, we obtain

$$\frac{d\lambda_x}{dx} = \frac{1}{\Lambda''(\lambda_x)}$$

hence

$$\frac{d\Lambda^*}{dx} = \lambda_x$$

and

$$\frac{d^2\Lambda^*}{dx^2} = \frac{1}{\Lambda''(\lambda_x)} > 0$$

Therefore $\Lambda^*(x)$ is strictly convex for $x > 0$.

Since Y_1 has compact support, it has moments of any order and in particular we have the following Taylor expansion

$$\Lambda(\lambda) = \frac{\lambda^2}{2}\mathbb{E}(Y^2) + O(\lambda^3)$$

and

$$\Lambda'(\lambda) = \lambda\mathbb{E}(Y^2) + O(\lambda^2) \tag{3.19}$$

In particular, as x tends to 0 the value λ_x tends to 0 too. We set $\Lambda^*(0) = 0$. By definition of Λ^*, we have for any $x > x_0 > 0$

$$\Lambda^*(x) \geq \lambda_{x_0}x - \Lambda(\lambda_{x_0}) \tag{3.20}$$

Now let us study the function Λ^* near 0. From (3.19), we get

$$\lambda_x = \frac{x}{\mathbb{E}(Y^2) + O(\lambda_x)}$$

and

$$\lim_{x\to 0^+} \frac{\lambda_x\mathbb{E}(Y^2)}{x} = 1$$

Hence

$$\begin{aligned}
\Lambda^*(x) &= \lambda_x x - \Lambda(\lambda_x) \\
&= \lambda_x x - \lambda_x^2\frac{\mathbb{E}(Y^2)}{2} + O(\lambda_x^3) \\
&= \frac{x^2}{\mathbb{E}(Y^2)}(u - \frac{u^2}{2}) + O(ux)^3
\end{aligned}$$

where $u = \lambda_x\mathbb{E}(Y^2)/x$. As x tends to 0, u tends to 1, therefore

$$\Lambda^*(x) \geq \frac{x^2}{4\mathbb{E}(Y^2)} \tag{3.21}$$

for any sufficiently small x.

Similarly

$$\begin{aligned}\frac{\Lambda(\lambda_x)}{x} &= \frac{\lambda_x^2}{2x}\mathbb{E}(Y^2) + O(\frac{\lambda_x^3}{x}) \\ &= \lambda_x \frac{u}{2} + O(\lambda_x^2 u)\end{aligned}$$

Hence

$$\frac{\Lambda(\lambda_x)}{x} \leq \frac{3}{4}\lambda_x \tag{3.22}$$

for all sufficiently small x. Fix $x_0 > 0$ such that both (3.21) and (3.22) hold for all $x \in (0, x_0]$. From (3.20) we obtain immediately for all $x \geq x_0$

$$\Lambda^*(x) \geq \lambda_{x_0}(x - \frac{3}{4}x_0) \tag{3.23}$$

Now let us now apply these estimates obtained above for $\Lambda^*(x)$ to the probability $\mathbb{P}(U_k > \sqrt{n}\log(n))$. We write for $x > 0$ and $\lambda > 0$

$$\mathbb{E}(e^{\lambda U_n}) \geq e^{\lambda x}\mathbb{P}(U_n > x)$$

or

$$\mathbb{P}(U_n > x) \leq e^{-n(\lambda \frac{x}{n} - \Lambda(\lambda))}$$

Hence taking the supremum of $\lambda > 0$

$$\mathbb{P}(U_n > x) \leq \exp(-n\Lambda^*(\frac{x}{n})) \tag{3.24}$$

Take two integers k and n with $k \leq n$. We have

$$\mathbb{P}(U_k > \sqrt{n}\log(n)) \leq \exp(-k\Lambda^*(\frac{\sqrt{n}\log n}{k}))$$

Suppose first that $\frac{\sqrt{n}\log n}{k} \geq x_0$. Then it follows from (3.23) that

$$\begin{aligned}\mathbb{P}(U_k &> \sqrt{n}\log(n)) \leq \exp(-\lambda_{x_0}\sqrt{n}\log n + \frac{3}{4}kx_0\lambda_{x_0}) \\ &\leq \exp(-\frac{1}{4}\lambda_{x_0}\sqrt{n}\log n) \leq e^{-c(\log n)^2}\end{aligned}$$

where we can take $c = \lambda_{x_0}/4$.

Now assume that $\frac{\sqrt{n}\log n}{k} \leq x_0$. Then it follows from (3.21) that

$$\mathbb{P}(U_k > \sqrt{n}\log(n)) \leq e^{-c(\log n)^2}$$

where c can be taken to be $1/4\mathbb{E}(Y^2)$.

By taking the lesser of the two c above we can now write

$$\mathbb{P}(U_k > \sqrt{n}\log(n)) \le e^{-c(\log n)^2}$$

for all integers n and k with $n \ge k \ge 1$. Hence

$$\mathbb{P}(\mathcal{A}_n^c) \le 2ne^{-c(\log n)^2}$$

Therefore there is a constant $c_0 > 0$ smaller that c such that when n is larger than say n_0 we have

$$\mathbb{P}(\mathcal{A}_n^c) \le e^{-c_0(\log n)^2}$$

For C_0 we may take $e^{c_0(\log n_0)^2}$ and we have obtained the desired inequality. $\square$

Obviously, we may, and do, assume that $c_{p+1} < c_p$ for all p. This lemma will enable us to reduce to the case when U_k does not take very big values. Let us also remark that in the above proof, the constant c_p can be chosen to depend only on the size of the support of Y, i.e. on $M = \min\{t, P(|Y| < t) = 1\}$. In the sequel, we will use freely the result of lemma 5.1, in particular the fact that $\mathbb{E}(|U_n|^p) = O_p(n^{p/2}\log^p(n))$ for any $p \ge 0$.

Let $G_n = (X_n, Y_n, Z_n)$ be a sequence of independent random variables identically distributed according to the probability measure μ on G. We write $S_n = G_n \cdot ... \cdot G_1$ the product of these variables. The law of S_n is μ^n. Bearing in mind the form of the product on G, we get $S_n = (S_{n,x}, S_{n,y}, S_{n,z})$ where

$$\begin{aligned} S_{n,x} &= X_1 + ... + X_n \\ S_{n,y} &= Y_1 + ... + Y_n \\ S_{n,z} &= Z_1 + ... + Z_n + X_2Y_1 + X_3(Y_1 + Y_2) + ... + X_n(Y_1 + ... + Y_{n-1}) \end{aligned}$$

Let θ_n be the random variable defined in (3.18) as follows

$$\begin{aligned} \theta_n &= \theta_n(t, s, \lambda) = e^{i(tS_{n,x} + sS_{n,y} + \lambda(S_{n,z} - \alpha S_{n,y}^2))} \\ &= e^{-i\lambda\alpha U_n^2} \prod_{1 \le k \le n} e^{iX_k(t + \lambda U_{k-1})} e^{i(sY_k + \lambda Z_k)} \\ \text{and } \theta_0 &= 1 \end{aligned} \tag{3.25}$$

where $U_k = Y_1 + ... + Y_k$ and $U_0 = 0$ as above. In the sequel we fix the value of α (as in equation (3.5)) to be

$$\alpha = \mathbb{E}(XY)/2\mathbb{E}(Y^2).$$

We will also use the following notation :

$$\beta(t, s) = \mathbb{E}(e^{i(X_1 t + Y_1 s)})$$

3.5.1 Estimate for "large" values of s and t

Proposition 5.2 *Let ε, A, D be positive numbers with $A > \varepsilon > 0$. Let k be an arbitrary positive integer. Then the following estimate holds uniformly when $\lambda \in [-D\frac{\log n}{n}, D\frac{\log n}{n}]$ and $\varepsilon < s^2 + t^2 < A$:*

$$\mathbb{E}(\theta_n) = O(\frac{1}{n^k})$$

where the constant in O depends on ε, A, k, D and μ.

Proof: We fix ε, A and D and consider (t, s, λ) in the range defined above. The proof will proceed by induction. Recall that $\mathcal{A}_n$ is the event $\{\max_{0\leq k\leq n} |U_k| \leq \sqrt{n}\log n\}$. In fact we will take for induction hypothesis the following statement, which we denote by $H_{j,k}$ for $j \in \mathbb{N}$ and $k \in \mathbb{Z}$:

$$\mathbb{E}(\theta_n U_n^j) = O(\frac{\log^{2k+3j} n}{n^{k/2}})$$

where O may depend on j and on k.

Clearly $H_{0,k}$ for all $k \geq 0$ implies the proposition. Also note that $H_{j,k}$ implies $H_{j,k-1}$.

Let $k + j \geq 0$. Making use of lemma 5.1 and of the independence of the random variables G_i's, we have the following Taylor expansion : for all positive integers $p \leq n$,

$$\begin{aligned}
\mathbb{E}(\theta_p U_p^j) &= \mathbb{E}(\theta_p U_p^j \mathbf{1}_{\mathcal{A}_p}) + O(e^{-c_j \log^2 p}) \\
&= \mathbb{E}(\theta_{p-1} U_p^j e^{-i\lambda\alpha(2U_{p-1}Y_p + Y_p^2)} e^{iX_p(t+\lambda U_{p-1})} e^{i(sY_p+\lambda Z_p)} \mathbf{1}_{\mathcal{A}_p}) + O(e^{-c_j \log^2 p}) \\
&= \mathbb{E}[\theta_{p-1} U_p^j e^{i(X_p t + Y_p s)} \mathbf{1}_{\mathcal{A}_p} \sum_{l=0}^{k+j} \frac{(i\lambda)^l}{l!} \\
&\sum_{q=0}^{l} C_l^q (Z_p - \alpha Y_p^2)^{l-q}((X_p - 2\alpha Y_p)U_{p-1})^q] + \\
&+ \lambda^{k+j+1} O((\sqrt{p}\log p)^{k+2j+1}) + O(e^{-c_j \log^2 p})
\end{aligned}$$

Further expanding and using lemma 5.1 again :

$$\begin{aligned}
\mathbb{E}(\theta_p U_p^j) = \mathbb{E}[\theta_{p-1} & \left(\sum_{r=0}^{j} C_j^r U_{p-1}^r Y_p^{j-r}\right) e^{i(X_p t + Y_p s)} \sum_{l=0}^{k+j} \frac{(i\lambda)^l}{l!} \cdot \\
& \sum_{q=0}^{l} C_l^q (Z_p - \alpha Y_p^2)^{l-q}((X_p - 2\alpha Y_p)U_{p-1})^q] + O(\frac{\log^{2k+3j+2} p}{p^{(k+1)/2}}) \\
&= \sum_{l=0}^{k+j} \sum_{q=0}^{l} \sum_{r=0}^{j} C_j^r C_l^q \frac{(i\lambda)^l}{l!} \mathbb{E}(\theta_{p-1} U_{p-1}^{r+q}) \\
& \cdot \mathbb{E}(e^{i(X_1 t + Y_1 s)} Y_1^{j-r} (Z_1 - \alpha Y_1^2)^{l-q} (X_1 - 2\alpha Y_1)^q) + \\
& + O(\frac{\log^{2(k+1)+3j} p}{p^{(k+1)/2}})
\end{aligned}$$

The last expression can be written as a sum of $\beta(t,s)\mathbb{E}(\theta_{p-1}U_{p-1}^j)$ (where $\beta(t,s) = \mathbb{E}(e^{i(X_1 t+Y_1 s)})$) and a linear combination with bounded coefficients of a bounded number of terms of the form $\lambda^l\mathbb{E}(\theta_{p-1}U_{p-1}^m)$ with $0 \leq m \leq j+l$ and $0 \leq l \leq k+j$ and $(m,l) \neq (j,0)$.

Let $j \geq 0$ and $k+j \geq 0$. Now suppose $H_{m,k+1-2l}$ holds for all $0 \leq m \leq j+l$ and $0 \leq l \leq k+j$ except when $(m,l) = (j,0)$. We are going to show that it implies $H_{j,k+1}$. Since $|\lambda| \leq D\frac{\log n}{n} \leq D\frac{\log p}{p}$, we have for all these values of m and l :

$$\lambda^l\mathbb{E}(\theta_{p-1}U_{p-1}^m) = O(\frac{\log^{l+2(k+1-2l)+3m} p}{p^{(k+1)/2}}) = O(\frac{\log^{2(k+1)+3j} p}{p^{(k+1)/2}})$$

Hence,

$$\mathbb{E}(\theta_p U_p^j) = \beta(t,s)\mathbb{E}(\theta_{p-1}U_{p-1}^j) + O(\frac{\log^{2(k+1)+3j} p}{p^{(k+1)/2}})$$

Therefore, recursively on p, we obtain for all n

$$\begin{aligned} \mathbb{E}(\theta_n U_n^j) &= \sum_{p=1}^{n}\beta(s,t)^{n-p}O(\frac{\log^{2(k+1)+3j} p}{p^{(k+1)/2}}) \qquad (3.26)\\ &= n\beta(s,t)^{n/2}O(n^{|k|}) + (1+...+\beta(t,s)^{n/2+1})O(\frac{\log^{2(k+1)+3j} n}{n^{(k+1)/2}}) \end{aligned}$$

As above, since μ is aperiodic, it follows that the law of (X,Y) on $\mathbb{R}^2$ is aperiodic. Hence

$$\sup_{\varepsilon<s^2+t^2<A}|\beta(t,s)| < 1$$

therefore (3.26) yields

$$\mathbb{E}(\theta_n U_n^j) = O(\frac{\log^{2(k+1)+3j} n}{n^{(k+1)/2}})$$

Thus we obtain $H_{j,k+1}$.

Now, we note that from lemma 5.1, $H_{j,-j}$ holds for all $j \geq 0$. This also guarantees $H_{j,k}$ when $k+j \leq 0$. Then we proceed by induction on $h = j+k$. From the considerations above, we obtain $H_{j,k}$ from the knowledge of other $H_{j',k'}$ with $j'+k' < j+k$. So we are done. $\square$

Remark 5.3 *If we make the following additional assumption on* μ

$$\sup_{t^2+s^2\geq 1}|\beta(t,s)| < 1$$

then proposition (5.2) *holds uniformly in* A *and*

$$\mathbb{E}(\theta_n) = O(\frac{1}{n^k})$$

holds uniformly in (t,s,λ) *when* $\lambda \in [-D\frac{\log n}{n}, D\frac{\log n}{n}]$ *and* $t^2+s^2 \geq \varepsilon$ *and the constant in* O *depends only* ε, k, D *and* μ.

3.5.2 Estimate for intermediate values of s and t

Now we will treat the case when $|\lambda| \leq D\frac{\log n}{n}$ and $D\frac{\log^8 n}{\sqrt{n}} \leq |(s,t)| \leq \varepsilon$ for some fixed (large) $D > 0$.

Proposition 5.4 *With the notation above, there is some $\varepsilon > 0$, such that uniformly in λ, $|\lambda| \leq D\frac{\log n}{n}$*

$$\int_{D\frac{\log^8 n}{\sqrt{n}} \leq |(s,t)| \leq \varepsilon} |\mathbb{E}(\theta_n(t,s,\lambda))| dtds = O(\frac{1}{n\log^2 n})$$

where the constant in O may depend on μ, D and ε.

Proof: Let us write $r := |(t,s)| = \sqrt{t^2+s^2}$ and now choose $\varepsilon > 0$ so that for $|(t,s)| \leq \varepsilon$, $|\partial_1\beta(t,s)| = O(r)$ and $|\partial_2\beta(t,s)| = O(r)$. We suppose that $|\lambda| \leq D\frac{\log n}{n}$.

Choosing ε smaller if necessary, we can assume that when $|(t,s)| \leq \varepsilon$

$$\exists C > 0, |\beta(t,s)| \leq e^{-Cr^2} \Leftrightarrow \frac{1}{1-|\beta(t,s)|} = O(\frac{1}{r^2})$$

From the computation in the subsection above (with $j = 2$ and $k = -1$), we obtain that $\mathbb{E}(\theta_p U_p^2)$ is a sum of $\beta(t,s)\mathbb{E}(\theta_{p-1}U_{p-1}^2)$ and a linear combination with bounded coefficients of the terms $\lambda\mathbb{E}(\theta_{p-1}U_{p-1}^m)$ with $0 \leq m \leq 2$, $\mathbb{E}(\theta_{p-1})$, $\mathbb{E}(\theta_{p-1}U_{p-1})\partial_2\beta(t,s)$, and $\lambda\mathbb{E}(\theta_{p-1}U_{p-1}^3)(\partial_1\beta(t,s)+\partial_2\beta(t,s))$, with a rest of order $O(\log^6 n)$. Hence, making use of lemma 5.1, for all n and p, $p \leq n$, and if $D\frac{\log^2 n}{\sqrt{n}} \leq r \leq \varepsilon$, we have

$$\mathbb{E}(\theta_p U_p^2) = \beta(t,s)\mathbb{E}(\theta_{p-1}U_{p-1}^2) + rO(\sqrt{n}\log^4 n)$$

where O is independent of p and n. Iterating this equation, we then deduce for p, $\frac{n}{2} \leq p \leq n$ we have

$$\begin{aligned}
\mathbb{E}(\theta_p U_p^2) &= \beta^{n/4}(t,s)O(n\log^2 n) + (1+...+\beta(t,s)^{n/4})rO(\sqrt{n}\log^4 n) \\
&= \frac{1}{1-|\beta(t,s)|}rO(\sqrt{n}\log^4 n) \\
&= \frac{1}{r}O(\sqrt{n}\log^4 n) \qquad (3.27)
\end{aligned}$$

since $\beta^{n/4}(t,s)O(n\log^2 n) = O(e^{-Cnr^2})O(n\log^2 n) = \frac{1}{r}O(\sqrt{n}\log^4 n)$, because $D\frac{\log n}{\sqrt{n}} \leq r \leq \varepsilon$.

Similarly we can express $\mathbb{E}(\theta_p)$ as above (taking $j = 0$ and $k = 1$ in the proof of prop. 5.2) as a sum of $\beta(t,s)\mathbb{E}(\theta_{p-1})$ and a linear combination with bounded coefficients of the terms $\lambda\mathbb{E}(\theta_{p-1})$ and $\lambda\mathbb{E}(\theta_{p-1}U_{p-1})(\partial_1\beta+\partial_2\beta)$ with a rest of order $O(\frac{\log^4 n}{n})$. Since $D\frac{\log^2 n}{\sqrt{n}} \leq r \leq \varepsilon$ we have

$$\mathbb{E}(\theta_p) = \beta(t,s)\mathbb{E}(\theta_{p-1}) + rO(\frac{\log^2 n}{\sqrt{n}})$$

Iterating as above, we obtain for $\frac{n}{2} \leq p \leq n$

$$\mathbb{E}(\theta_p) = \frac{1}{r} O(\frac{\log^2 n}{\sqrt{n}}) \tag{3.28}$$

Now we look at $\mathbb{E}(\theta_p U_p)$. As above (take $j = 1$, $k = 0$) it is a sum of $\beta(t,s)\mathbb{E}(\theta_{p-1}U_{p-1})$ and a linear combination with bounded coefficients of the terms $\partial_2\beta(t,s)\mathbb{E}(\theta_{p-1})$, $\lambda\mathbb{E}(\theta_{p-1})$, $\lambda\mathbb{E}(\theta_{p-1}U_{p-1})$, and $\lambda(\partial_1\beta + \partial_2\beta)\mathbb{E}(\theta_{p-1}U_{p-1}^2)$, with a rest of order $O(\frac{\log^5 n}{\sqrt{n}})$. But thanks to (3.28) and (3.27) all these terms are $O(\frac{\log^5 n}{\sqrt{n}})$ when $\frac{n}{2} \leq p \leq n$. Hence,

$$\mathbb{E}(\theta_p U_p) = \beta(t,s)\mathbb{E}(\theta_{p-1}U_{p-1}) + O(\frac{\log^5 n}{\sqrt{n}})$$

And then, iterating this relation, for all p, $\frac{3n}{4} \leq p \leq n$, we get

$$\mathbb{E}(\theta_p U_p) = \frac{1}{r^2} O(\frac{\log^5 n}{\sqrt{n}}) \tag{3.29}$$

Finally we again decompose $\mathbb{E}(\theta_p)$ but pushing one step further the Taylor expansion and we see by the above calculation (take $j = 0$, $k = 2$) that it is a sum of $\beta(t,s)\mathbb{E}(\theta_{p-1})$ and a linear combination with bounded coefficients of the terms $\lambda\mathbb{E}(\theta_{p-1})$, $\lambda\mathbb{E}(\theta_{p-1}U_{p-1})\partial_1\beta(t,s)$, $\lambda^2\mathbb{E}(\theta_{p-1})$, $\lambda^2\mathbb{E}(\theta_{p-1}U_{p-1})$ and $\lambda^2\mathbb{E}(\theta_{p-1}U_{p-1}^2)$ with a rest of order $O(\frac{\log^6 n}{n\sqrt{n}})$. Thanks to (3.27), (3.28), (3.29) and lemma 5.1, they are all of order at most $\frac{1}{r}O(\frac{\log^6 n}{n\sqrt{n}})$ when $\frac{3n}{4} \leq p \leq n$. Thus

$$\mathbb{E}(\theta_p) = \beta(t,s)\mathbb{E}(\theta_{p-1}) + \frac{1}{r}O(\frac{\log^6 n}{n\sqrt{n}})$$

and consequently for large n

$$\mathbb{E}(\theta_n) = \frac{1}{r^3}O(\frac{\log^6 n}{n\sqrt{n}})$$

Then, integrating over r when $D\frac{\log^8 n}{\sqrt{n}} \leq r \leq \varepsilon$ we obtain :

$$\begin{aligned}
\int_{D\frac{\log^8 n}{\sqrt{n}} \leq |(s,t)| \leq \varepsilon} |\mathbb{E}(\theta_n)| dtds &\leq O(\frac{\log^6 n}{n\sqrt{n}}) \int_{D\frac{\log^8 n}{\sqrt{n}} \leq r \leq \varepsilon} \frac{1}{r^3} r dr \\
&\leq O(\frac{1}{n \log^2 n})
\end{aligned}$$

This concludes the proof of the proposition. $\square$

3.6 Study of a dynamical system

In this section we study a quadratic dynamical system in the complex plane and give precise estimates that will be crucial in the proof of the domination condition in the next section.

We first fix three real numbers x, y, z satisfying the following condition

$$\Delta := \det \begin{pmatrix} 1 & x & z \\ x & 1 & y \\ z & y & 1 \end{pmatrix} = 1 + 2xyz - x^2 - y^2 - z^2 \geq 0$$

And suppose additionally that $|x| \leq 1$, $|y| \leq 1$ and $|z| < 1$. Now define the following three sequences recursively :

$$\begin{aligned} a_{k+1} &= a_k + \frac{1}{2} - 2\lambda^2 c_k^2 + 2i\lambda c_k y \\ b_{k+1} &= b_k + \frac{1}{2} - 2\lambda^2 b_k^2 + 2i\lambda b_k z \\ c_{k+1} &= c_k + \frac{x}{2} - 2\lambda^2 b_k c_k + i\lambda b_k y + i\lambda c_k z \end{aligned} \tag{3.30}$$

where the initial values a_0, b_0 and c_0 are arbitrary and λ is a given real number. In the applications below, λ will be small (of order $\log(n)/n$) and the above dynamical system can be viewed as a perturbation of that given when $\lambda = 0$. In this section we will study the behavior of the above three sequences depending on initial values and also on the values of the parameters x, y, z and λ.

We first note that (b_k) is a quadratic dynamical system and is therefore conjugate to $P_\lambda : u \mapsto u^2 + c_\lambda$ for some complex number c_λ. A straightforward computation shows that if we set

$$x_k := \frac{1}{2} + i\lambda z - 2\lambda^2 b_k$$

then we have $x_{k+1} = P_\lambda(x_k)$ where $c_\lambda = \frac{1}{4} - \lambda^2(1 - z^2)$. In the limit when λ tends to 0, then c_λ tends to $\frac{1}{4}$, which is on the boundary of the Mandelbrot set. As long as λ is small enough and non-zero, then the Fatou set corresponding to P_λ has exactly one bounded connected component. Moreover, as soon as $|\lambda| < 1$, for every starting point x_0 lying inside this component, the resulting sequence of iterates (x_k) will converge to the attracting fixed point x_λ given by

$$x_\lambda = \frac{1}{2} - |\lambda|\sqrt{1 - z^2}$$

Thus for $\lambda \neq 0$ and $|\lambda| < 1$, the sequence (b_k) converges to

$$b_\lambda = \frac{iz + sgn(\lambda)\sqrt{1 - z^2}}{2\lambda}$$

Now let us define

$$v_k := c_k + \frac{yz - x}{1 - z^2} b_k$$

Then it is easy to check directly from the equations (3.30) that v_k satisfies

$$v_{k+1} - v = (x_k + \frac{1}{2})(v_k - v) \tag{3.31}$$

where

$$v = \frac{i}{2\lambda} \frac{y - xz}{1 - z^2} \tag{3.32}$$

Let us write $y_k = \frac{1}{2} + x_k$, then we obtain

$$v_k - v = (v_0 - v) y_0 \cdot \ldots \cdot y_{k-1} \tag{3.33}$$

Similarly, if we let

$$f_k := a_k + \frac{yz - x}{1 - z^2} c_k$$

then we find that

$$f_{k+1} - f_k = \frac{\Delta}{2(1 - z^2)} + (v_k - v)(1 - (\frac{1}{2} + x_k)) \frac{yz - x}{1 - z^2} - 2\lambda^2 (v_k - v)^2$$

Making use of (3.31), it follows that for $k \geq 1$

$$\begin{aligned} a_k &= f_0 + \left(\frac{yz - x}{1 - z^2} \right)^2 b_k + \frac{yz - x}{1 - z^2} (v_0 - 2v_k) + \\ & \quad k \frac{\Delta}{2(1 - z^2)} - 2\lambda^2 (v_0 - v)^2 \sum_{p=0}^{k-1} (y_0 \cdot \ldots \cdot y_{p-1})^2 \end{aligned} \tag{3.34}$$

We are now going to study the dynamical system (3.30) in the particular case when the initial values are defined by

$$\begin{aligned} a_0 &= 0 \\ b_0 &= \frac{i}{2\lambda} z \\ c_0 &= \frac{i}{2\lambda} w \end{aligned} \tag{3.35}$$

where w is a fixed real number. Then $x_0 = \frac{1}{2}$ belongs to the filled Julia set of P_λ, and the sequence (x_k) (hence (y_k) too) stays on the real line and satisfies

$$\frac{1}{2} - |\lambda| \sqrt{1 - z^2} = x_\lambda \leq x_k \leq \frac{1}{2}$$

Additionally,

$$v_0 = \frac{i}{2\lambda}(w + z\frac{yz-x}{1-z^2})$$

Together with (3.31) and (3.32) this shows that v_k belongs to $i\mathbb{R}$ for all k. With these initial values, (3.34) takes the form

$$\begin{aligned} a_k &= \frac{1}{2\lambda}\left(\frac{yz-x}{1-z^2}\right)^2 (iz + \frac{1}{\lambda}(1-y_k)) + \frac{i}{2\lambda}\left(\frac{yz-x}{1-z^2}\right)(2w - \frac{2y}{1-z^2} + z\frac{yz-x}{1-z^2}) - \\ &\quad -\frac{i}{\lambda}\left(\frac{yz-x}{1-z^2}\right)(w-y)y_0 \cdot \ldots \cdot y_{k-1} + \\ &\quad +k\frac{\Delta}{2(1-z^2)} + \frac{1}{2}(w-y)^2 \sum_{p=0}^{k-1}(y_0 \cdot \ldots \cdot y_{p-1})^2 \end{aligned}$$

taking the real part we get

$$\begin{aligned} Re(a_k) &= \frac{1}{2|\lambda|}\left(\frac{yz-x}{1-z^2}\right)^2 \frac{1-y_k}{|\lambda|} + k\frac{\Delta}{2(1-z^2)} + \\ &\quad \frac{1}{2}(w-y)^2 \sum_{p=0}^{k-1}(y_0 \cdot \ldots \cdot y_{p-1})^2 \end{aligned} \tag{3.36}$$

The following lemma summarizes the computations above and enclose the information that will be relevant to the sequel :

Lemma 6.1 *In the dynamical system defined by* (3.30)*, with initial values given by* (3.35)*, the following holds.*

Take $|\lambda| \leq 1/2$*. Then for all* $k \geq 0$

(i) $|\lambda k| \leq 1/\sqrt{1-z^2}$ *implies* $1 - y_k \geq k\frac{1}{2}\lambda^2(1-z^2)$*, and* $Re(b_k) \geq k\frac{1-z^2}{4}$*, and*

$$Re(a_k) \geq \frac{k}{4}[\frac{(yz-x)^2}{1-z^2} + 2e^{-4}(w-y)^2]$$

(ii) $|\lambda k| \geq 1/\sqrt{1-z^2}$ *implies* $|\lambda|\sqrt{1-z^2} \geq 1 - y_k \geq \frac{1}{2}|\lambda|\sqrt{1-z^2}$*, and* $Re(b_k) \geq \frac{1}{4|\lambda|}\sqrt{1-z^2}$ *and*

$$Re(a_k) \geq \frac{1}{|\lambda|}\frac{1}{4\sqrt{1-z^2}}[\frac{(yz-x)^2}{1-z^2} + \frac{1}{4}(w-y)^2]$$

(iii) $|\lambda b_k| \leq \frac{1}{2}$ *and* $Re(b_k) \geq 0$

(iv) $|\lambda v_k| \leq 2/(1-z^2) + |w| + 1$

(v) $|\lambda c_k| \leq 3/(1-z^2) + |w| + 1$

(vi) $Re(a_k)Re(b_k) \geq Re(c_k)^2$

(vii) $Re(a_k)$ *is a non-decreasing sequence.*

Proof: All these points are easy to check from what was done above. The proof of (i) follows by induction ; it is true when $k = 0$ and, assuming the inequality for k, we get

$$\begin{aligned} 1 - y_{k+1} &= (1 - y_k)y_k + \lambda^2(1 - z^2) \\ &\geq k\frac{1}{2}\lambda^2(1 - z^2)(1 - |\lambda|\sqrt{1 - z^2}) + \lambda^2(1 - z^2) \\ &\geq k\frac{1}{2}\lambda^2(1 - z^2) + \frac{1}{2}\lambda^2(1 - z^2)(2 - |k\lambda|\sqrt{1 - z^2}) \\ &\geq (k + 1)\frac{1}{2}\lambda^2(1 - z^2) \end{aligned} \tag{3.37}$$

Besides, for all $p \leq k$ we have $y_p \geq 1 - |\lambda|\sqrt{1 - z^2} \geq e^{-2|\lambda|\sqrt{1-z^2}}$ since $1 - t \geq e^{-2t}$ if $t \in [0, \frac{1}{2}]$. Hence

$$(y_0 \cdot \ldots \cdot y_{p-1})^2 \geq e^{-4|\lambda p|\sqrt{1-z^2}} \geq e^{-4}$$

The inequality for $Re(a_k)$ in (i) now follows instantly from (3.36) and the fact that $\Delta \geq 0$.

Point (ii) follows from the fact (granted by (3.37)) that $1-y_k \geq c|\lambda|$ implies $1-y_{k+1} \geq c|\lambda|$ for any real number c with $\sqrt{1 - z^2} \geq c \geq 0$.

We have $2\lambda b_k = iz+(1-y_k)/\lambda$ and $0 \leq 1-y_k \leq |\lambda|\sqrt{1 - z^2}$ so $|2\lambda b_k| \leq \sqrt{z^2 + (1 - z^2)} \leq 1$ yields (iii). Similarly $v_k - v = (v_0 - v)y_0 \cdot \ldots \cdot y_{k-1}$, hence

$$2\lambda v_k = i(y - xz)/(1 - z^2) + i(w - y)y_0 \cdot \ldots \cdot y_{k-1}$$

Since $|y_i| \leq 1$ for all i, we get $|2\lambda v_k| \leq |y - xz|/|1 - z^2| + |w - y|$ so $|2\lambda v_k| \leq 2/|1 - z^2| + |w| + 1$, and we have (iv) and also (v) because of (iii) and $v_k := c_k + \frac{yz-x}{1-z^2}b_k$.

For (vi), we have $Re(a_k) \geq \frac{1-y_k}{2\lambda^2}\left(\frac{yz-x}{1-z^2}\right)^2 = Re(b_k)\left(\frac{yz-x}{1-z^2}\right)^2$, so (recall that the v_k belong to $i\mathbb{R}$)

$$Re(a_k)Re(b_k) \geq \left(Re(b_k)\frac{yz - x}{1 - z^2}\right)^2 = Re(c_k)^2$$

Finally (vii) is easily checked from (3.36). $\square$

3.7 Proof of the domination condition for small values of the parameters s, t, λ

In this section, we shall give a domination estimate for a particular type of trigonometric sum that will arise in the proof of local limit theorem. We then apply these estimates to treat the part of the integral that yields a contribution to the limit, that is when s, t, λ are small.

3.7.1 Estimating a trigonometric sum

Let us consider throughout this section a sequence of independent and identically distributed random variables $(A_k, B_k, C_k, D_k)_{k\geq 1}$ in $\mathbb{R}^4$. We assume that the distribution has compact support in $\mathbb{R}^4$, and that $\mathbb{E}(A_k) = \mathbb{E}(B_k) = \mathbb{E}(C_k) = 0$. Let us use the shorthand $\overline{X}$ to denote the expectation $\mathbb{E}(X)$ of a random variable X. We fix the following notations for the covariances

$$x = \frac{\overline{A_1B_1}}{\sqrt{\overline{A_1^2}\,\overline{B_1^2}}},\ y = \frac{\overline{A_1C_1}}{\sqrt{\overline{A_1^2}\,\overline{C_1^2}}},\ z = \frac{\overline{B_1C_1}}{\sqrt{\overline{B_1^2}\,\overline{C_1^2}}}$$

We also assume that A_1 is not identically 0 and that the distribution of the marginal (B_k, C_k) is not degenerate, i.e. is not supported on a line. This is equivalent to the condition $|z| < 1$.

Fix $w \in \mathbb{R}$ and then consider the trigonometric product for r, λ in $\mathbb{R}$,

$$\theta_n = \left(\prod_{1\leq k\leq n} e^{ir(A_k/\sqrt{\overline{A_1^2}}-wC_k/\sqrt{\overline{C_1^2}})}e^{i\lambda D_k}\right)\left(\prod_{1\leq p<q\leq n} e^{i\lambda B_qC_p/\sqrt{\overline{B_1^2}\,\overline{C_1^2}}}\right)e^{-i\frac{\lambda}{2}z(C_1+...+C_n)^2/\overline{C_1^2}} \tag{3.38}$$

The following proposition yields the desired estimate for the trigonometric sum $\mathbb{E}(\theta_n)$.

Proposition 7.1 *Let us fix $D > 0$ a positive number and $m \geq 1$ an integer. For any integer $n \geq 1$ and any distribution (A_1, B_1, C_1, D_1), of compact support and as described above, the following estimate holds uniformly when r varies in $[-D\frac{\log^{2m} n}{\sqrt{n}}, D\frac{\log^{2m} n}{\sqrt{n}}]$ and λ varies in $[-D\frac{\log n}{n}, D\frac{\log n}{n}]\backslash[-\frac{2}{n\sqrt{1-z^2}}, \frac{2}{n\sqrt{1-z^2}}]$,*

$$|\mathbb{E}(\theta_n)| \leq \exp(-\left[n|\lambda|\frac{\sqrt{1-z^2}}{8} + \frac{r^2}{|\lambda|}\frac{C}{16\sqrt{1-z^2}}\right]) + \left[\frac{1+|w|}{1-z^2}\right]^4 O(\frac{\log^{6m} n}{\sqrt{n}})$$

where $C = \left((w-y)^2 + \frac{(yz-x)^2}{1-z^2}\right)$.

Similarly, if r varies in $[-D\frac{\log^{2m} n}{\sqrt{n}}, D\frac{\log^{2m} n}{\sqrt{n}}]$ and λ varies in $[-\frac{2}{n\sqrt{1-z^2}}, \frac{2}{n\sqrt{1-z^2}}]$, we obtain

$$|\mathbb{E}(\theta_n)| \leq \exp[-nr^2\frac{Ce^{-4}}{4}] + \left[\frac{1+|w|}{1-z^2}\right]^4 O(\frac{\log^{6m} n}{\sqrt{n}})$$

The constant in $O(\cdot)$ depends only on D and on the size of the distribution

$$M = \max\{|A_1|/\sqrt{\overline{A_1^2}}, |B_1|/\sqrt{\overline{B_1^2}}, |C_1|/\sqrt{\overline{C_1^2}}, |D_1|\}$$

Proof: Let $n \in \mathbb{N}$ and $r, \lambda \in \mathbb{R}$ be as in the statement of the proposition. Let $U_0 = 0$ and for $k \geq 1$,

$$U_k = (C_1 + ... + C_k)/\sqrt{\overline{C^2}}$$

Let q_k be the quadratic form

$$q_k(u,v) = a_k u^2 + b_k v^2 + 2c_k uv$$

and let $\pi_0 = 1$ and for $k \geq 1$

$$\pi_k = e^{ir(A_1+...+A_k)/\sqrt{\overline{A_1^2}}} \left(\prod_{1 \leq p < q \leq k} e^{i\lambda B_q C_p/\sqrt{\overline{B_1^2 C_1^2}}} \right) e^{i\lambda(D_1+...+D_k)}$$

and

$$P_k(r,\lambda) = \mathbb{E}(\pi_{n-k} e^{-q_k(r,\lambda U_{n-k})})$$

We shall define the coefficients a_k, b_k, c_k recursively as will be shown below.

Note that

$$\begin{aligned} q_k(r,\lambda U_{n-k}) &= q_k(r, \lambda U_{n-k-1} + \lambda C_{n-k}/\sqrt{\overline{C^2}}) \\ &= q_k(r,\lambda U_{n-k-1}) + b_k \lambda^2 C_{n-k}^2/\overline{C^2} + \\ & \quad 2b_k \lambda^2 U_{n-k-1} C_{n-k}/\sqrt{\overline{C^2}} + 2c_k \lambda r C_{n-k}/\sqrt{\overline{C^2}} \end{aligned}$$

We also recall that $\mathcal{A}_n$ is the event $\{\max_{0 \leq k \leq n} |U_k| \leq \sqrt{n} \log n\}$. We now let

$$P_k(r,\lambda) = \mathbb{E}(\pi_{n-k} e^{-q_k(r,\lambda U_{n-k})})$$

Suppose $|\lambda| \leq D\frac{\log n}{n}$ and $|r| \leq D\frac{\log^{2m} n}{n}$. Then for all k and n with $k \leq n-1$ we have

$$P_k(t,s) = e^{-b_k\lambda^2 + i\overline{D}\lambda} P_{k+1}(t,s) + \left[\frac{1+|w|}{1-z^2}\right]^4 O(\frac{\log^{6m} n}{n\sqrt{n}}) \tag{3.39}$$

where the constant in O depends only on D and on the size M of the support of the distribution $(A_1, ..., D_1)$. This crucial estimate follows from the computation below. Recall that near $t = 0$ we have the expansion $e^t = 1 + t + t^2/2 + O(t^3)$. Making use of lemma 5.1 and bearing in mind that μ is centered we can write

$$
\begin{aligned}
P_k &= \mathbb{E}\left(\pi_{n-k-1}e^{irA_{n-k}/\sqrt{\overline{A^2}}}e^{i\lambda B_{n-k}U_{n-k-1}/\sqrt{\overline{B^2}}}e^{i\lambda D_{n-k}}e^{-q_k(r,\lambda U_{n-k})}\right)\\
&= \mathbb{E}(\mathbf{1}_{\mathcal{A}_n}\pi_{n-k-1}e^{-q_k(r,\lambda U_{n-k-1})}e^{irA_{n-k}/\sqrt{\overline{A^2}}}e^{i\lambda B_{n-k}U_{n-k-1}/\sqrt{\overline{B^2}}}\\
&\quad e^{i\lambda D_{n-k}}e^{-b_k\lambda^2C_{n-k}^2/\overline{C^2}-2b_k\lambda^2U_{n-k-1}C_{n-k}/\sqrt{\overline{C^2}}-2c_k\lambda rC_{n-k}/\sqrt{\overline{C^2}}}) + e^{-c_0\log^2 n}\\
&= \mathbb{E}(\mathbf{1}_{\mathcal{A}_n}\pi_{n-k-1}e^{-q_k(r,\lambda U_{n-k-1})}(1 + irA_{n-k}/\sqrt{\overline{A^2}} + i\lambda B_{n-k}U_{n-k-1}/\sqrt{\overline{B^2}} +\\
&\quad i\lambda D_{n-k} - b_k\lambda^2C_{n-k}^2/\overline{C^2} - 2b_k\lambda^2U_{n-k-1}C_{n-k}/\sqrt{\overline{C^2}} - 2c_k\lambda rC_{n-k}/\sqrt{\overline{C^2}} +\\
&\quad \frac{1}{2}\begin{bmatrix} irA_{n-k}/\sqrt{\overline{A^2}} + i\lambda B_{n-k}U_{n-k-1}/\sqrt{\overline{B^2}} + i\lambda D_{n-k}\\ -b_k\lambda^2C_{n-k}^2/\overline{C^2} - 2b_k\lambda^2U_{n-k-1}C_{n-k}/\sqrt{\overline{C^2}} - 2c_k\lambda rC_{n-k}/\sqrt{\overline{C^2}}\end{bmatrix}^2\\
&\quad +O(\frac{\log^{6m} n}{n\sqrt{n}}))) + e^{-c_0\log^2 n}\\
&= \mathbb{E}(\pi_{n-k-1}e^{-q_k(r,\lambda U_{n-k-1})}[1 - b_k\lambda^2 + i\overline{D}\lambda - r^2(\frac{1}{2} - 2\lambda^2c_k^2 + 2i\lambda c_ky) -\\
&\quad \lambda^2U_{n-k-1}^2(\frac{1}{2} - 2\lambda^2b_k^2 + 2i\lambda b_kz) - 2\lambda rU_{n-k-1}(\frac{x}{2} - 2\lambda^2b_kc_k + i\lambda b_ky + i\lambda c_kz)]\\
&\quad +O(\frac{\log^{6m} n}{n\sqrt{n}}))\\
&= e^{-b_k\lambda^2+i\overline{D}\lambda}P_{k+1} + O(\frac{\log^{6m} n}{n\sqrt{n}})
\end{aligned}
$$

This computation makes sense as long as λb_k and λc_k remain uniformly bounded when n grows. With the help of lemma 5.1 and the remark following it, we verify that the constant involved in the O in the above calculations can be taken of the form $c{\cdot}D^4{\cdot}K^4$, where $c > 0$ depends only on the size of the support of the distribution of (A_1, B_1, C_1, D_1) and where K is that number that bounds λc_k and λb_k. We also need to insure that $Re(q_k) \geq 0$ everywhere, i.e. $Re(a_k)Re(b_k) \geq (Re(c_k))^2$ and $Re(a_k) \geq 0$.

For (3.39) to be valid, we must set

$$
\begin{aligned}
a_{k+1} &= a_k + \frac{1}{2} - 2\lambda^2c_k^2 + 2i\lambda c_ky\\
b_{k+1} &= b_k + \frac{1}{2} - 2\lambda^2b_k^2 + 2i\lambda b_kz\\
c_{k+1} &= c_k + \frac{x}{2} - 2\lambda^2b_kc_k + i\lambda b_ky + i\lambda c_kz
\end{aligned}
\tag{3.40}
$$

which are precisely the recurrence relations (3.30) defined in the last section. Finally, if we let $a_0 = 0$, $c_0 = \frac{i}{2\lambda}w$ for some $w \in \mathbb{R}$, and $b_0 = \frac{i}{2\lambda}z$ as in the last section, lemma 6.1 applies and the conditions λb_k and λc_k uniformly bounded (by $const/(1 - z^2) + |w| + 1$) and $Re(a_k)Re(b_k) \geq (Re(c_k))^2$ hold. Also note that $P_0 = \mathbb{E}(\pi_ne^{-q_0(r,\lambda U_n)}) =$

$\mathbb{E}(\pi_n e^{-\frac{i}{2}z\lambda U_n^2 - riwU_n})$, so that

$$P_0(r,\lambda) = \mathbb{E}(\theta_n).$$

And

$$P_n = \mathbb{E}(e^{-q_n(r,0)}) = e^{-a_n r^2}.$$

Iterating (3.39), we now deduce

$$\mathbb{E}(\theta_n) = e^{-\lambda^2 \sum_{k=0}^{n-1} b_k} e^{i\lambda n\overline{D}} e^{-a_n r^2} + O(\frac{\log^{6m} n}{\sqrt{n}})$$

The constant involved here in the O is bounded by some $cD^4(1+|w|)^4/(1-z^2)^4$ where c is a constant depending only on $M = \max\{|A_1|/\sqrt{\overline{A_1^2}}, |B_1|/\sqrt{\overline{B_1^2}}, |C_1|/\sqrt{\overline{C_1^2}}, |D_1|\}$.

$$|\mathbb{E}(\theta_n)| \leq e^{-\lambda^2 \sum_{k=0}^{n-1} Reb_k} e^{-r^2 Rea_n} + O(\frac{\log^{6m} n}{\sqrt{n}})$$

From lemma 6.1, if $|\lambda| \geq \frac{2}{\sqrt{1-z^2}}\frac{1}{n}$ then $Reb_k \geq \frac{1}{4|\lambda|}\sqrt{1-z^2}$ for all $k \geq n/2$ and $Re(b_k) \geq 0$ for all k. So the first factor in the above equation leads to the bound $e^{-n|\lambda|\sqrt{1-z^2}/8}$. While $Re(a_n) \geq \frac{1}{|\lambda|}\frac{1}{4(1-z^2)^{1/2}}\left(\frac{1}{4}(w-y)^2 + \frac{(yz-x)^2}{1-z^2}\right)$.

If $|\lambda| \leq \frac{2}{\sqrt{1-z^2}}\frac{1}{n}$, then $Re(a_n) \geq Re(a_{n/2}) \geq \frac{n}{4}e^{-4}\left((w-y)^2 + \frac{(yz-x)^2}{1-z^2}\right)$.

So we obtain the desired inequalities.

□

3.7.2 Domination condition

The purpose of this subsection is to give the precise estimate we wanted in the course of the proof of the local limit theorem (control of the part of the integral where all parameters s, t, λ are small). This is explained is the following proposition :

Proposition 7.2 *Let μ be a probability measure on the Heisenberg group with the properties described in the introduction. Let θ_n be the random variable defined in (3.25) at the beginning of section 3.5. There exist positive numbers $c_1 > 0$ and $c_2 > 0$ depending only on μ, such that, if $D > 0$ denotes some positive number, then the following estimates hold uniformly*

(i) when s and t vary in $[-D\frac{\log^8 n}{\sqrt{n}}, D\frac{\log^8 n}{\sqrt{n}}]$ and λ varies in $[-D\frac{\log n}{n}, D\frac{\log n}{n}]\backslash[-\frac{c_2}{n}, \frac{c_2}{n}]$,

$$|\mathbb{E}(\theta_n)| \leq \exp(-c_1\left[n|\lambda| + \frac{t^2+s^2}{|\lambda|}\right]) + O(\frac{\log^{24} n}{\sqrt{n}})$$

(ii) when s and t vary in $[-D\frac{\log^8 n}{\sqrt{n}}, D\frac{\log^8 n}{\sqrt{n}}]$ and λ varies in $[-\frac{c_2}{n}, \frac{c_2}{n}]$,

$$|\mathbb{E}(\theta_n)| \leq \exp(-c_1 n(t^2+s^2)) + O(\frac{\log^{24} n}{\sqrt{n}})$$

where the constant in O depends only on μ and on D.

Proof: We are going to apply the results of the last section. In order to do so, we need to choose carefully the variables A_1, B_1, C_1, D_1 as well as the coefficient w. Take $D_1 = Z/\sqrt{\overline{X^2} \cdot \overline{Y^2}}$, $C_1 = Y/\sqrt{\overline{Y^2}}$, $B_1 = X/\sqrt{\overline{X^2}}$ and $A_1 = \cos(\theta)\frac{X}{\sqrt{\overline{X^2}}} + 2\sin(\theta)\frac{Y}{\sqrt{\overline{Y^2}}}$ and $w = \frac{\sin(\theta)}{\alpha(\theta)}$, where $\alpha(\theta)^2 = \overline{A_1^2} = 1 + 3\sin^2(\theta) + 4z\sin(\theta)\cos(\theta)$ and $\alpha(\theta) \geq 0$. An easy computation shows that $8 \geq \alpha(\theta)^2 \geq (1 - z^2)/5$. Clearly the variables A_1, B_1 and C_1 are linearly dependent, hence $\Delta = 0$. But B_1 and C_1 are linearly independent since μ is aperiodic, hence $|z| < 1$.

Then the two expressions θ_n in (3.25) and (3.38) agree if we change λ into $\lambda/\sqrt{\overline{X^2 Y^2}}$ and let r and θ be determined by

$$
\begin{aligned}
t &= \frac{r\cos(\theta)}{\alpha(\theta)\sqrt{\overline{X^2}}} \\
s &= \frac{r\sin(\theta)}{\alpha(\theta)\sqrt{\overline{Y^2}}}
\end{aligned}
$$

From the above inequality on $\alpha(\theta)$, we conclude that if $M_{X,Y} = \max\{\frac{1}{\overline{X^2}}, \frac{1}{\overline{Y^2}}\}$ and $m_{X,Y} = \min\{\frac{1}{\overline{X^2}}, \frac{1}{\overline{Y^2}}\}$,

$$
m_{X,Y}\frac{1 - z^2}{5}(t^2 + s^2) \leq \frac{r^2}{\overline{X^2 Y^2}} \leq 8M_{X,Y}(t^2 + s^2)
$$

To compute the constant C appearing in proposition (7.1), let us first compute x, y and z.

$$
\begin{aligned}
z &= \overline{XY}/\sqrt{\overline{X^2} \cdot \overline{Y^2}} \\
y &= \frac{z\cos(\theta) + 2\sin(\theta)}{\alpha(\theta)} \\
x &= \frac{\cos(\theta) + 2z\sin(\theta)}{\alpha(\theta)}
\end{aligned}
$$

hence

$$
\begin{aligned}
C &= (w - y)^2 + \frac{(yz - x)^2}{1 - z^2} \\
&= \left(\frac{\sin(\theta) + z\cos(\theta)}{\alpha(\theta)}\right)^2 + \frac{(1 - z^2)\cos^2(\theta)}{\alpha(\theta)^2} \\
&= \frac{1}{\alpha(\theta)^2}\left[1 - 2z\cos(\theta)\sin(\theta)\right] \\
&\geq \frac{1}{\alpha(\theta)^2}(1 - |z|) \geq \frac{1 - z^2}{2 \cdot 8}
\end{aligned}
$$

Also note that $|w| \leq 1/\alpha(\theta) \leq \sqrt{5/(1-z^2)}$. The first estimate in (7.1) now yields

$$\begin{aligned}
|\mathbb{E}(\theta_n)| &\leq \exp(-\left[n|\lambda|\sqrt{\overline{X^2}\cdot\overline{Y^2}}\frac{\sqrt{1-z^2}}{8} + \frac{r^2}{|\lambda|}\frac{C}{4\sqrt{1-z^2}}\right]) + O(\frac{\log^6 n}{\sqrt{n}}) \\
&\leq \exp(-\sqrt{\overline{X^2}\cdot\overline{Y^2}}\sqrt{1-z^2}\left[\frac{n|\lambda|}{8} + \frac{r^2}{|\lambda|\overline{X^2}\cdot\overline{Y^2}}\frac{1}{64}\right]) + O(\frac{\log^6 n}{\sqrt{n}}) \\
&\leq \exp(-\sqrt{\overline{X^2}\cdot\overline{Y^2} - \overline{XY}^2}\left[\frac{n|\lambda|}{8} + \frac{M_{X,Y}(t^2+s^2)}{64\cdot|\lambda|}\right]) + O(\frac{\log^6 n}{\sqrt{n}}) \\
&\leq \exp(-c_1\left[n|\lambda| + \frac{(t^2+s^2)}{|\lambda|}\right]) + O(\frac{\log^6 n}{\sqrt{n}})
\end{aligned}$$

where we can take $c_1 \leq \frac{\sqrt{\overline{X^2}\cdot\overline{Y^2}-\overline{XY}^2}}{8}\min\{1, M_{X,Y}/8\}$, and the constant in O in the last line depends only on μ and D.

We thus have obtained (i), and (ii) follows similarly with $c_1 \leq \frac{e^{-4}}{80}(\overline{X^2}\cdot\overline{Y^2} - \overline{XY}^2)$.
□

3.8 Proofs of the main theorems

We can now finish the proof of theorem 2.1. Let f and g be like in the statement of theorem 2.1. Recall that the number α was defined in (3.5). As was remarked at the end of section 3, we can write

$$n^2\langle\rho(\mu^n)f, g\rangle = A_n + B_n$$

where A_n and B_n are defined as follows :

$$A_n = n^2\int_{|\lambda|\geq D\frac{\log n}{n}}\langle\rho_\lambda(\mu^n)f_\lambda, g_\lambda\rangle_{\mathcal{H}_\lambda}\,d\lambda$$

and there is a constant C depending on D, μ and the size of the set $\{\lambda, \exists(t,s), F_g(g)(s,t,\lambda) \neq 0\}$ such that

$$|A_n| \leq \frac{C}{n^{k_0-2}}\|f\|_{L^2(G)}\|g\|_{L^2(G)} \tag{3.41}$$

as soon as the integer k_0 satisfies $D > k_0/c$ (where $c > 0$ is a constant depending on μ only defined in (3.13)). And

$$B_n = \int_G \frac{d\mathbf{x}}{(2\pi)^{3/2}}f(\mathbf{x})I_n(\mathbf{x}) \tag{3.42}$$

with

$$I_n(\mathbf{x}) = n^2\int_{|\lambda|\leq D\frac{\log n}{n}}\left\langle\mathbb{E}(\theta_n(t,s,\lambda)), e^{-i\phi}F_\sigma(g)(t-\lambda y, s+2\alpha\lambda y, \lambda)\right\rangle_{\mathbb{L}^2(\mathbb{R}^2)}d\lambda$$

where, keeping the notations of theorem 2.1 $F_\sigma(g)$ is the Fourier transform defined in (3.4), $\theta_n(t,s,\lambda)$ is defined in (3.25) and $\mathbf{x} = (x,y,z) \in G$.

Splitting the integral on $\mathbb{R}^2$ in the expression of $I_n(\mathbf{x})$ above into the parts when $|(t,s)| \leq D\frac{\log^8 n}{\sqrt{n}}$ on the one hand and $|(t,s)| \geq D\frac{\log^8 n}{\sqrt{n}}$ on the other hand, we can write :

$$I_n(\mathbf{x}) = I_n^S(\mathbf{x}) + I_n^L(\mathbf{x}) \tag{3.43}$$

and assert that if $\|\mathbf{x}\| = \max\{|x|,|y|,|z|\}$ and $2|\alpha|$ are less than say $K > 1$ then $F_\sigma(g)(t-\lambda y, s+2\alpha\lambda y, \lambda) \neq 0$ implies that $\max\{|t|,|s|,|\lambda|\} \leq A(1+K^2)$ where $A > 0$ is a number such that the support of $F_\sigma(g)$ lies inside $[-A,A]^3$. Then we may write :

$$\begin{aligned}
|I_n^L(\mathbf{x})| &\leq n^2 \|F_\sigma(g)\|_\infty \int_{|\lambda|\leq D\frac{\log n}{n}} \int_{D\frac{\log^8 n}{\sqrt{n}} \leq |(t,s)| \leq A(1+K^2)} |\mathbb{E}(\theta_n(t,s,\lambda))| dt ds d\lambda \\
&\leq n^2 \|F_\sigma(g)\|_\infty \int_{|\lambda|\leq D\frac{\log n}{n}} O(\frac{1}{n\log^2(n)}) d\lambda \qquad (3.44)\\
&\leq \|F_\sigma(g)\|_\infty O(\frac{1}{\log(n)})
\end{aligned}$$

where line 3.44 is granted by the two propositions 5.2 and 5.4. The constant involved here in O depends only on μ, D and the size A of the support of $F_\sigma(g)$ and on the maximum of $\|\mathbf{x}\|$. In particular it is uniform when $\mathbf{x}$ varies in compact subsets of G.

We can now concentrate on the part $I_n^S(x)$ of the integral which actually gives a contribution to the limit. From proposition 7.2, we deduce that if $|(t,s)| \leq D\log^8 n$ and $c_2 \leq |\lambda| \leq D\log n$

$$|\mathbb{E}(\theta_n(\frac{t}{\sqrt{n}}, \frac{s}{\sqrt{n}}, \frac{\lambda}{n}))| \leq e^{-c_1(|\lambda| + t^2/|\lambda| + s^2/|\lambda|)} + O(\frac{\log^{24} n}{\sqrt{n}})$$

and if $|(t,s)| \leq D\log^8 n$ and $|\lambda| \leq c_2$

$$|\mathbb{E}(\theta_n(\frac{t}{\sqrt{n}}, \frac{s}{\sqrt{n}}, \frac{\lambda}{n}))| \leq e^{-c_1(t^2+s^2)} + O(\frac{\log^{24} n}{\sqrt{n}})$$

where the constant in O depends only on μ and D. But one can check that the function

$$(t,s,\lambda) \mapsto e^{-c(|\lambda| + t^2/|\lambda| + s^2/|\lambda|)}$$

is integrable over $\mathbb{R}^3$. And

$$(t,s,\lambda) \mapsto e^{-c(t^2+s^2)}$$

is integrable in $(t,s,\lambda) \in \mathbb{R}^2 \times [-c_2, c_2]$. Moreover

$$\int_{|\lambda|\leq D\log n} \int_{|(t,s)| \leq D\log^8 n} O(\frac{\log^{24} n}{\sqrt{n}}) = O(\frac{\log^{41} n}{\sqrt{n}}) \to 0$$

And finally, from the central limit theorem (see [169] or [142]) the following limit holds point-wise in t, s, λ :

$$\mathbb{E}(\theta_n(\frac{t}{\sqrt{n}}, \frac{s}{\sqrt{n}}, \frac{\lambda}{n})) \to \mathbb{E}(e^{i(tX+sY+\lambda(Z-\alpha Y^2))})$$

where (X, Y, Z) is the limit random variable with gaussian distribution ν_1 as mentioned in the introduction. Hence if we use the shorthand $t_n := t/\sqrt{n}$, $s_n = s/\sqrt{n}$ and $\lambda_n := \lambda/n$ we have (see the definition of $\phi_n = \phi(\frac{t}{\sqrt{n}}, \frac{s}{\sqrt{n}}, \frac{\lambda}{n})$ in (3.18) above), uniformly when $\mathbf{x}$ varies in compact subsets of G

$$\lim_{n\to+\infty} \mathbb{E}(\theta_n(t_n, s_n, \lambda_n))e^{i\phi_n}\overline{F_\sigma(g)}(t_n - \lambda_n y, s_n + 2\alpha\lambda_n y, \lambda_n) = \mathbb{E}(e^{i(tX+sY+\lambda(Z-\alpha Y^2))})\overline{F_\sigma(g)}(0)$$

Since the heat kernel corresponding to μ is a fastly decreasing smooth function $p(x, y, z)$ on $\mathbb{R}^3$ (see [171]) it follows that $\mathbb{E}(e^{i(tX+sY+\lambda(Z-\alpha Y^2))})$ is integrable in $(t, s, \lambda) \in \mathbb{R}^3$ and we compute by the Fourier inversion formula

$$\int_{\mathbb{R}^3} \mathbb{E}(e^{i(tX+sY+\lambda(Z-\alpha Y^2))})dtdsd\lambda = (2\pi)^3 p(0, 0, 0) = (2\pi)^3 c(\mu)$$

Therefore, by Lebesgue's dominated convergence theorem, we get uniformly when $\mathbf{x}$ varies in compact subsets

$$\begin{aligned}
\lim_{n\to+\infty} I_n(\mathbf{x}) &= \lim_{n\to+\infty} I_n^S(\mathbf{x}) \\
&= n^2 \int_{|\lambda|\le D\frac{\log n}{n}} \int_{|(t,s)|\le D\frac{\log^8 n}{\sqrt{n}}} \mathbb{E}(\theta_n(t, s, \lambda))e^{i\phi}\overline{F_\sigma(g)}(t - \lambda y, s + 2\alpha\lambda y, \lambda)dtdsd\lambda \\
&= (2\pi)^3 c(\mu)\overline{F_\sigma(g)}(0)
\end{aligned}$$

Integrating in $\mathbf{x}$ we finally obtain

$$\begin{aligned}
\lim_{n\to\infty} B_n &= \lim_{n\to+\infty} \int_G \frac{d\mathbf{x}}{(2\pi)^{3/2}} f(\mathbf{x}) I_n(\mathbf{x}) \\
&= (2\pi)^{3/2} c(\mu)\overline{F_\sigma(g)}(0) \int_G f \\
&= c(\mu) \int_G f \int_G \overline{g} \qquad (3.45)
\end{aligned}$$

Combining (3.45) and (3.41) we get

$$\lim_{n\to\infty} n^2 \langle \rho(\mu^n) f, g\rangle = c(\mu) \int_G f \int_G \overline{g} \qquad (3.46)$$

as desired.

We also remark that any translation of g (to the left ${}_zg(x) = g(z^{-1}x)$ or to the right $g_z(x) = g(xz)$) has again a compactly supported Fourier transform $F_\sigma({}_zg)$ and $F_\sigma(g_z)$. If z remains in a compact subset, then the supports of $F_\sigma({}_zg)$ and $F_\sigma(g_z)$ also remain within a prescribed bounded set. Moreover $\|F_\sigma(g_z)\|_\infty = \|F_\sigma({}_zg)\|_\infty$ and $\|g_z\|_2 = \|{}_zg\|_2$ are independent of z. This follows from the computation of of $F_\sigma({}_zg)$ which yields

$$F_\sigma({}_zg)(t,s,\lambda) = e^{i\tau} \cdot F_\sigma(g)(t, s - \lambda(2\alpha z_y - z_x), \lambda) \tag{3.47}$$

where $z = (z_x, z_y, z_z)$ and $\tau = z_x t + z_y s + \lambda(z_z - \alpha z_y^2)$. Consequently all the calculations above, and (3.46) in particular, hold uniformly for translates of g on compact subsets.

Now take f_n a Dirac sequence of positive functions supported on a neighborhood of 0 of diameter of order $1/n^3$. Then $\|f_n\|_2^2 = O(n^9)$. Choosing D large enough (so that A_n in (3.41) remains negligeable) we see that we can replace f by f_n in the above calculations, hence

$$\lim_{n\to\infty} n^2 \langle \rho(\mu^n) f_n, g\rangle = c(\mu) \int_G \overline{g} \tag{3.48}$$

And the same holds uniformly when z varies in compact subsets and g is replaced by g_z or ${}_zg$.

Suppose now that $F_\sigma(g)$ is absolutely continuous, then it follows[1] that $g \circ \phi_\sigma$ is real analytic and that $\|x\|\, d_x(g \circ \phi_\sigma)$ is a bounded function on $\mathbb{R}^3$ (where $\|x\|$ is the max of the coordinates of $x \in \mathbb{R}^3$). For all compact K we can then find a constant $C = C(g, K)$ such that whenever y is small enough

$$\sup_{z\in K} \|{}_{y^{-1}z}g -{}_z g\|_\infty \leq C\,\|y\|$$

and

$$\sup_{z\in G} \|{}_{y^{-1}}g_z - g_z\|_\infty \leq C\,\|y\|$$

Hence uniformly for z in compact subsets, $\|{}_zg * f_n -{}_z g\|_\infty = O(1/n^3)$ and $\|g_z * f_n - g_z\|_\infty = O(1/n^3)$. From (3.48) it now follows that uniformly for z in compact subsets :

$$\lim_{n\to\infty} n^2 \int g(z^{-1}x) d\mu^n(x) = c(\mu) \int_G g$$

and

$$\lim_{n\to\infty} n^2 \int g(xz) d\mu^n(x) = c(\mu) \int_G g \tag{3.49}$$

□

We now deduce Theorem 1.1 from Theorem 2.1.

[1]Remark that if a function $h(x) \in \mathbb{L}^1(\mathbb{R})$ is such that $\widehat{h}(t) \in C_c(\mathbb{R})$ and $\frac{d\widehat{h}}{dt} \in \mathbb{L}^1(\mathbb{R})$ (i.e. $\widehat{h}$ absolutely continuous) then $x\frac{dh}{dx}$ is the Fourier transform of $\frac{d}{dt}(t\widehat{h})$, hence is bounded.

Proof: For all $\varepsilon > 0$ one can find[2] a strictly positive function h in $\mathbb{L}^1(\mathbb{R}^3)$ such that $\widehat{h}$ is C^1 and compactly supported, and such that there exists $C > 0$, with $h(\mathbf{x}) \|\mathbf{x}\|^{6+\varepsilon} \geq C$ for $\mathbf{x} \in \mathbb{R}^3$ large enough. Let $g = h \circ \phi_\sigma^{-1}$. Now for all $\xi \in \mathbb{R}^3$ and $z \in G$, $e^{i\xi\cdot\phi_\sigma^{-1}(x)} g_z(x)$ also has C^1 Fourier transform F_σ of compact support, hence (3.49) holds for it uniformly when z varies in compact subsets. By Lévy's criterion for weak convergence of finite measures, this shows that the sequence of finite measures $d\nu_n^z = n^2 g_z d\mu^n$ converges weakly (in the space of finite measures on G) to $c(\mu) g_z(x) dx$ uniformly when z varies in compact subsets. Now let f be a function on G as in the statement of the theorem. Then f/g is a bounded continuous function, hence

$$n^2 \int f_z d\mu^n = \int f_z/g_z d\nu_n^z \to c(\mu) \int_G f$$

uniformly when z varies in compact subsets. $\square$

$\square$

3.9 Uniform local limit theorem for translates of a bounded set

We intend here to prove theorem 1.2. The probability measure μ is assumed to be aperiodic, centered and compactly supported. The letter ν denotes the associated gaussian probability distribution. Its density function, the heat kernel, is denoted by $p_1(x, y, z)$. It is a fastly decreasing smooth function on G.

We start by fixing a non-negative function K on G such that $F_\sigma(K)$ is smooth and has compact support and $F_\sigma(K)(0) = 1$. Then we form a Dirac family $(K_a)_{a>0}$ by letting $K_a(x) = a^4 K(d_a(x))$. Then $F_\sigma(K_a)(\xi) = F_\sigma(K)(d_{\frac{1}{a}}(\xi))$. Let us also write $K_a^\vee(z) = K_a(z^{-1})$. They also form a Dirac family when $a \to +\infty$.

Lemma 9.1 *There are two sequences of positive numbers $(\varepsilon_n)_n$ and $(a_n)_n$, depending only on μ, with $\varepsilon_n \to 0$ and $a_n \to +\infty$, such that for all bounded borelian sets $B \subset G$ for which $\max\{|y_1|, |y_2|\} \leq n/\log n$ whenever $\mathbf{y} = (y_1, y_2, y_3) \in B$, and all $\mathbf{x} \in G$, if we set $P_n^B(\mathbf{x}) = \mu^n(\mathbf{x}B)$ and $Q_n^B(\mathbf{x}) = \nu^n(\mathbf{x}B)$, the following inequality holds for all positive integers n,*

$$n^2 \left| P_n^B * K_{a_n}^\vee(\mathbf{x}) - Q_n^B * K_{a_n}^\vee(\mathbf{x}) \right| \leq \varepsilon_n \max\{1, |B|\} \tag{3.50}$$

where $|B|$ denotes the Haar measure of B.

[2] It is enough to find a function $f \in \mathbb{L}^1(\mathbb{R})$ such that $\widehat{f}$ is C^1 of compact support and $f(x) \geq \frac{C}{1+|x|^{2+\varepsilon}}$ for some $C > 0$ (e.g. see Chapter 4). Then take $h(\mathbf{x}) = f(x)f(y)f(z)$ if $\mathbf{x} = (x, y, z)$.

Proof: We may write

$$\begin{aligned}
P_n^B * K_a^\vee(\mathbf{x}) &= \int_G \mu^n(\mathbf{x}z^{-1}B)K_a^\vee(z)dz \\
&= \int_G \mathbb{E}(\mathbf{1}_{S_n^{-1}z\in B^{-1}})K_a(\mathbf{x}^{-1}z)dz \\
&= \langle \rho(\mu^n)f,_{\mathbf{x}}\, g\rangle
\end{aligned}$$

where ${}_{\mathbf{x}}g$ is the translate ${}_{\mathbf{x}}g(z) := g(\mathbf{x}^{-1}z)$ and $f := \mathbf{1}_{B^{-1}}$ and $g(z) := K_a(z)$ (note that K_a is real). Then we can make use of the calculations performed in the previous sections to estimate this scalar product. We use the notations introduced in section 3.8.

As noted at the beginning of section 3.8, we can write

$$n^2 \langle \rho(S_n)f,_{\mathbf{x}}\, g\rangle = A_n(\mu) + B_n(\mu)$$

where A_n is controlled by the estimation (3.41) and B_n given by (3.42). Since $\|f\|_{L^2(G)} = \sqrt{|B|}$ and $\|{}_x g\|_{L^2(G)} = a^2 \|K\|_{L^2(G)}$ we obtain (take $k_0 = 3$ and $D = 3/c$)

$$|A_n(\mu)| \le C(a) \cdot \frac{a^2}{n} \cdot \sqrt{|B|} \cdot \|K\|_{L^2(G)} \tag{3.51}$$

where $C(a)$ is a positive constant depending only on a. Integrating the decomposition $I_n(\mathbf{y}) = I_n^S(\mathbf{y}) + I_n^L(\mathbf{y})$ with respect to $f(\mathbf{y})d\mathbf{y}$ (see equations (3.42) and (3.43)), we obtain that $B_n(\mu)$ can be written as a sum $B_n^S(\mu) + B_n^L(\mu)$. Note that $\|F_\sigma({}_x K_a)\|_\infty = \|F_\sigma(K)\|_\infty$ as follows from (3.47). Spliting $B_n^L(\mu)$ into part when $\varepsilon \ge |(t,s)| \ge D\frac{\log^8 n}{n}$ and the part when $|(t,s)| \ge \varepsilon$ we have from (3.42) and (3.44)

$$B_n^L(\mu) = B_n^{L_0}(\mu) + B_n^{L_1}(\mu)$$

$$|B_n^{L_0}(\mu)| \le |B| \cdot \|F_\sigma(K)\|_\infty \cdot O_\mu(\frac{1}{\log n}) \tag{3.52}$$

as follows from proposition 5.4, and

$$|B_n^{L_1}(\mu)| \le \frac{|B| \cdot \|F_\sigma(K)\|_\infty}{(2\pi)^{3/2}} \cdot L \cdot n^2 \sup_{T \ge |(t,s)| \ge \varepsilon, |\lambda| \le D\frac{\log n}{n}} |\mathbb{E}(\theta_n^\mu(t,s,\lambda))| \tag{3.53}$$

where T is the size of the support of the functions $F_\sigma({}_{\mathbf{x}}g)$ and L is the Lebesgue measure of that support :

$$(t,s) \mapsto F_\sigma(K)(\frac{t-\lambda y_2}{a}, \frac{s + 2\lambda\alpha(y_2 - x_2) + \lambda x_1}{a}, \frac{\lambda}{a^2})$$

where $\mathbf{x} = (x_1, x_2, x_3)$ and $\mathbf{y} = (y_1, y_2, y_3)$. Note that L is a fixed multiple of a^2. Since $|\lambda| \le D\frac{\log n}{n}$ in (3.44) and by assumption $\max\{|y_1|, |y_2|\} \le n/\log n$, if we suppose additionally that $\max\{|x_1|, |x_2|\} \le 2n/\log n$ then the constant T in the above equation

(3.53) is bounded by some fixed function of a. In proposition 5.2 we showed that there is a constant C_T such that

$$\sup_{T \geq |(t,s)| \geq \varepsilon, |\lambda| \leq D\frac{\log n}{n}} |\mathbb{E}(\theta_n^\mu(t,s,\lambda))| \leq \frac{C_T}{n^3}$$

Hence for some number $\beta(a) < +\infty$

$$|B_n^L(\mu)| \leq |B| \cdot \frac{\beta(a)}{\log n} \tag{3.54}$$

The estimations above can be carried out in a similar way for ν instead of μ. In particular

$$\begin{aligned} Q_n^B * K_{a_n}^\vee(\mathbf{x}) &= \langle \rho(\nu^n) f,_\mathbf{x} g\rangle \\ &= A_n(\nu) + B_n(\nu) \end{aligned}$$

The term $A_n(\nu)$ is dealt with in exaclty the same way since estimate (3.41) is also valid for ν and we have

$$|A_n(\nu)| \leq C(a) \cdot \frac{a^2}{n} \cdot \sqrt{|B|} \cdot \|K\|_{L^2(G)} \tag{3.55}$$

In order to control $B_n(\mu)$ we made use of the compact support assumption on μ. For ν we can use the following direct argument because $F_\sigma(p_1)$ is integrable since $p_1(x,y,z)$ decays rapidly when (x,y,z) is large. Recall that $(\nu_t)_t$ is a stable semi-group, i.e. $\nu_n = \nu_1^n = d_{\sqrt{n}}(\nu_1)$. From (3.44) we have

$$\begin{aligned} |B_n^L(\nu)| &\leq \frac{1}{(2\pi)^{3/2}} \cdot |B| \cdot \|F_\sigma(K)\|_\infty \, n^2 \int_{|(t,s)| \geq D\frac{\log^8 n}{\sqrt{n}}} |\mathbb{E}(\theta_n^\nu(t,s,\lambda))| dt ds d\lambda \\ &\leq \frac{1}{(2\pi)^{3/2}} \cdot |B| \cdot \|F_\sigma(K)\|_\infty \cdot \int_{|(t,s)| \geq D \log^8 n} |F_\sigma(p_1)(t,s,\lambda)| dt ds d\lambda \\ &\leq \frac{1}{(2\pi)^{3/2}} \cdot |B| \cdot \|F_\sigma(K)\|_\infty \cdot o(1) \end{aligned} \tag{3.56}$$

Therefore it remains to treat $B_n^S(\mu) - B_n^S(\nu)$. From (3.42) and (3.43), we have

$$\left|B_n^S(\mu) - B_n^S(\nu)\right| \leq A \cdot n^2 \int_{|\lambda| \leq D\frac{\log n}{n}} \int_{|(t,s)| \leq D\frac{\log^8 n}{\sqrt{n}}} |\mathbb{E}(\theta_n^\mu(t,s,\lambda)) - \mathbb{E}(\theta_n^\nu(t,s,\lambda))| dt ds d\lambda \tag{3.57}$$

with

$$A = \frac{1}{(2\pi)^{3/2}} \cdot |B| \cdot \|F_\sigma(K)\|_\infty$$

Now the integral on the right hand side tends to 0 as it follows from Lebesgue's dominated converge theorem like we did in section 3.8. In section 7.4 we showed that there exists an integrable function $\psi(t,s,\lambda)$ such that for n sufficiently large and for all (t,s,λ) such that $|\lambda| \leq D\log n$ and $|(t,s)| \leq D\log^8 n$,

$$\left|\mathbb{E}(\theta_n^\mu(\frac{t}{\sqrt{n}},\frac{s}{\sqrt{n}},\frac{\lambda}{n}))\right| \leq \psi(t,s,\lambda) + O(\frac{\log^6 n}{\sqrt{n}})$$

On the other hand $F_\sigma(p_1)$ is integrable and

$$\mathbb{E}(\theta_n^\nu(\frac{t}{\sqrt{n}},\frac{s}{\sqrt{n}},\frac{\lambda}{n})) = F_\sigma(p_1)(t,s,\lambda)$$

And by the central limit theorem, we had point-wise

$$\lim_{n\to\infty} \mathbb{E}(\theta_n^\mu(\frac{t}{\sqrt{n}},\frac{s}{\sqrt{n}},\frac{\lambda}{n})) - \mathbb{E}(\theta_n^\nu(\frac{t}{\sqrt{n}},\frac{s}{\sqrt{n}},\frac{\lambda}{n})) = 0$$

Hence by Lebesgue's dominated convergence theorem,

$$\lim_{n\to+\infty} n^2 \int_{|\lambda|\leq D\frac{\log n}{n}} \int_{|(t,s)|\leq D\frac{\log^8 n}{\sqrt{n}}} |\mathbb{E}(\theta_n^\mu(t,s,\lambda)) - \mathbb{E}(\theta_n^\nu(t,s,\lambda))| dtdsd\lambda = 0$$

Therefore (3.57) reads :

$$\left|B_n^S(\mu) - B_n^S(\nu)\right| \leq |B| \cdot o(1) \tag{3.58}$$

And finally combining (3.51), (3.55), (3.56), (3.54) and (3.58), we have that for some sequence $\varepsilon_n \to 0$

$$n^2 \left|P_n^B * K_a^\vee(\mathbf{x}) - Q_n^B * K_a^\vee(\mathbf{x})\right| \leq \varepsilon_n|B| + \frac{\beta(a)}{\log n}|B| + C(a)\frac{a^2}{n} \cdot 2\left\|K\right\|_2 \sqrt{|B|}$$

whenever $\mathbf{x}$ satisfies $\max\{|x_1|,|x_2|\} \leq 2n/\log n$. But we can choose a sequence $a_n \to +\infty$ such that $\frac{\beta(a_n)}{\log n}$ and $C(a_n)\frac{a_n^2}{n}$ tend to 0. Changing ε_n if necessary, we obtain for these values of $\mathbf{x}$

$$n^2 \left|P_n^B * K_{a_n}^\vee(\mathbf{x}) - Q_n^B * K_{a_n}^\vee(\mathbf{x})\right| \leq \varepsilon_n \max\{|B|, \sqrt{|B|}\}$$

Now let us examine the case when $\mathbf{x}$ takes large values, i.e. when $\max\{|x_1|,|x_2|\} \geq 2n/\log n$. Then

$$P_n^B * K_{a_n}^\vee(\mathbf{x}) = \int_G \mu^n(xzB)K_{a_n}(z)dz \leq P(|T_n| \text{ or } |U_n| \geq n/2\log n) + \int_{|z|\geq\frac{n}{2\log n}} K_{a_n}(z)dz$$

where U_n is as before the sum $Y_1 + ... + Y_n$ and $T_n = X_1 + ... + X_n$. Now since $F_\sigma(K)$ is smooth, K decays rapidly and in particular

$$\int_{|z|\geq\frac{n}{2\log n}} K_{a_n}(z)dz = \int_{|z|\geq a_n\frac{n}{2\log n}} K(z)dz = O(\frac{1}{n^k})$$

for any $k \geq 0$. Hence we get from lemma 5.1

$$P_n^B * K_{a_n}^{\vee}(\mathbf{x}) = o(1/n^2)$$

Similarly the same holds for $Q_n^B * K_{a_n}^{\vee}(\mathbf{x})$ when $\max\{|x_1|, |x_2|\} \geq 2n/\log n$. Changing $(\varepsilon_n)_n$ if necessary, we obtain the desired conclusion, i.e. inequality (3.50) for all $\mathbf{x}$. □

The restrictions on the size of B in the above lemma disappear if we make the additional assumption (Cramer's condition) that

$$\sup_{t^2+s^2\geq 1} |\mathbb{E}(e^{i(tX+sY)})| < 1$$

Indeed in this case, we can control $\mathbb{E}(\theta_n^{\mu}(t,s,\lambda))$ uniformly for arbitrary large values of t and s (see remark (5.3)). Therefore we can take $T = +\infty$ in the estimation (3.53) above and the restriction on B is unecessary. We obtain

Lemma 9.2 *If we make the following additional assumption (Cramer's condition) on μ*

$$\sup_{t^2+s^2\geq 1} |\mathbb{E}(e^{i(tX+sY)})| < 1$$

then there are two sequences of positive numbers $(\varepsilon_n)_n$ and $(a_n)_n$, depending only on μ, with $\varepsilon_n \to 0$ and $a_n \to +\infty$, such that for all bounded borelian sets $B \subset G$ and all $x \in G$, if we set $P_n^B(x) = \mu^n(xB)$ and $Q_n^B(x) = \nu^n(xB)$, the following inequality holds for all positive integers n,

$$n^2 \left|P_n^B * K_{a_n}^{\vee}(x) - Q_n^B * K_{a_n}^{\vee}(x)\right| \leq \varepsilon_n \max\{1, |B|\}$$

where $|B|$ denotes the Haar measure of B.

Lemma 9.3 *Let $(a_n)_n$ be a sequence of positive numbers such that $a_n \to +\infty$. Then there exists another sequence $(\varepsilon_n)_n$ with $\varepsilon_n \to 0$ and a sequence of neighborhoods of identity (U_n) converging to identity, such that for all bounded borelian sets $B \subset G$, the following inequality holds for all positive integers n,*

$$n^2 \left|Q_n^B(x) - Q_n^B * K_{a_n}^{\vee}(x)\right| \leq p(|U_n B \backslash B| + \varepsilon_n |B|)$$

where $|B|$ denotes the Haar measure of B and where we have set $Q_n^B(x) = \nu^n(xB)$ and $p = p_1(0,0,0) > 0$.

Proof: The proof is straightforward. We first note that for any bounded borelian set B, $\nu^n(B) \leq p|B|/n^2$. Then we simply write :

$$\begin{aligned}
|Q_n^B * K_{a_n}^{\vee}(x) - Q_n^B(x)| &\leq \int |Q_n^B(xz) - Q_n^B(x)| K_{a_n}(z)dz \\
&\leq \int_{z\in U_n^{-1}} \nu^n(x(zB\Delta B)K_{a_n}(z)dz \\
&\quad +2p\frac{|B|}{n^2}\int_{z\notin U_n^{-1}} K_{a_n}(z)dz \\
&\leq p(|U_n B\backslash B| + \varepsilon_n |B|)/n^2
\end{aligned}$$

where U_n is a sequence of neighborhoods of identity tending to identity such that $\int_{z\notin U_n^{-1}} K_{a_n}(z)dz$ tends to 0 at infinity. $\square$

Now, let us complete the proof of theorem (1.2). Let B be an arbitrary bounded borelian set satisfying the condition of lemma (9.1), that is $\max\{|y_1|,|y_2|\} \leq n/\log n$ whenever $y=(y_1,y_2,y_3)\in B$ (resp. satisfying no additional condition if we assume Cramer's condition). In the sequel like above, the Landau notations o and O will correspond to functions depending only on μ. We keep notations of lemma (9.1),

$$\begin{aligned} P_n^{U_nB} * K_{a_n}^{\vee}(x) &= \int \mu^n(xzU_nB)K_{a_n}(z)dz \\ &\geq \mu^n(xB)\int_{U_n} K_{a_n}(z)dz \end{aligned}$$

where U_n is as in lemma 9.3. Now making use of lemma 9.1 we get uniformly in $x\in G$,

$$\begin{aligned} P_n^B(x) &\leq (1+o(1))P_n^{U_nB} * K_{a_n}^{\vee}(x) \\ &\leq (1+o(1))[Q_n^{U_nB} * K_{a_n}^{\vee}(x) + (1+|U_nB|)o(1/n^2)] \end{aligned} \tag{3.59}$$

And by lemma 9.3,

$$\begin{aligned} Q_n^{U_nB} * K_{a_n}^{\vee}(x) &\leq Q_n^{U_nB}(x) + \frac{p}{n^2}(|U_n^2B\backslash B| + \varepsilon_n|U_nB|) \\ &\leq Q_n^B(x) + \frac{p}{n^2}|U_nB\backslash B| + \frac{p}{n^2}(|U_n^2B\backslash B| + \varepsilon_n|U_nB|) \\ &\leq Q_n^B(x) + \frac{p}{n^2}(2|U_n^2B\backslash B| + \varepsilon_n|U_nB|) \end{aligned}$$

In particular we have

$$Q_n^{U_nB} * K_{a_n}^{\vee}(x) \leq \frac{4p}{n^2}|U_n^2B|$$

and, from (3.59),

$$P_n^B(x) \leq |U_n^2B|O(\frac{1}{n^2}) + o(\frac{1}{n^2}) \tag{3.60}$$

Additionally,

$$P_n^B(x) \leq Q_n^B(x) + \frac{2p}{n^2}|U_n^2B\backslash B| + (|U_n^2B|+1)o(\frac{1}{n^2}) \tag{3.61}$$

Now let us turn to the other direction of the inequality. We have, making use of (3.60)

$$\begin{aligned} P_n^B * K_{a_n}^{\vee}(x) &= \int \mu^n(xzB)K_{a_n}(z)dz \\ &\leq \int_{U_n} \mu^n(xzB)K_{a_n}(z)dz + \int_{U_n^c} \mu^n(xzB)K_{a_n}(z)dz \\ &\leq \mu^n(xU_nB)\int_{U_n} K_{a_n}(z)dz + (|U_n^2B|+1)o(\frac{1}{n^2}) \\ &\leq P_n^{U_nB}(x) + (|U_n^2B|+1)o(\frac{1}{n^2}) \end{aligned}$$

But from (3.60),

$$\begin{aligned} P_n^{U_nB}(x) - P_n^B(x) &= \mu^n(x(U_nB\backslash B)) \\ &\leq |U_n^2(U_nB\backslash B)|O(\frac{1}{n^2}) + o(\frac{1}{n^2}) \end{aligned}$$

Hence

$$P_n^B * K_{a_n}^{\vee}(x) \leq P_n^B(x) + |U_n^2(U_nB\backslash B)|O(\frac{1}{n^2}) + (|U_n^2B| + 1)o(\frac{1}{n^2})$$

Now it follows from lemma (9.1) that

$$Q_n^B * K_{a_n}^{\vee}(x) \leq P_n^B * K_{a_n}^{\vee}(x) + (1 + |U_nB|)o(1/n^2)$$

and from lemma (9.3)

$$Q_n^B(x) \leq Q_n^B * K_{a_n}^{\vee}(x) + \frac{p}{n^2}(|U_nB\backslash B| + \varepsilon_n|B|)$$

Combining the last three inequalities, we get

$$Q_n^B(x) \leq P_n^B(x) + |U_n^2(U_nB\backslash B)|O(\frac{1}{n^2}) + (|U_n^2B| + 1)o(\frac{1}{n^2}) \tag{3.62}$$

Equations (3.61) and (3.62) yield the desired result

$$|P_n^B(x) - Q_n^B(x)| \leq |U_n^2(U_nB\backslash B)|O(\frac{1}{n^2}) + (|U_n^2B| + 1)o(\frac{1}{n^2}) \tag{3.63}$$

But clearly $\bigcap_{n\geq 0} U_n^2(U_nB\backslash B)$ is contained in $\overline{B}\backslash \overset{\circ}{B}$. Hence for every bounded measurable set B such that $|\partial B| = 0$ we have

$$\lim_{n\to 0} \sup_{x\in G} n^2|\mu^n(xB) - \nu^n(xB)| = 0 \tag{3.64}$$

And if μ satisfies Cramer's condition the estimate (3.63) holds without the above restriction on B. Hence (3.63) holds uniformly on y for all By. We conclude

$$\lim_{n\to 0} \sup_{x,y\in G} n^2|\mu^n(xBy) - \nu^n(xBy)| = 0 \tag{3.65}$$

Remark 9.4 *It is easy to see from* (3.63) *that the limits in* (3.64) *and* (3.65) *are uniform in* B *when* B *ranges over the set of balls for a given norm on* G *lying in a given compact subset of* G.

Finally note that theorem 1.3 follows instantly from the inequality (3.60) above.

3.10 Applications

3.10.1 An equidistribution result for bounded uniformly continuous functions

In this section, we intend to give a proof of corollary (1.4). The proof splits into two steps.

Lemma 10.1 *Suppose f is a continuous and bounded function on G satisfying the condition* (3.2) *of corollary* (1.4) *then*

$$\lim_{n\to\infty}\int_G f(g)d\nu^n(g) = \ell$$

where $\nu = \nu_1$ is an arbitrary gaussian measure on G.

Proof: Let $(\nu_t)_t$ be the one-parameter semigroup of gaussian measures in which ν is embedded. By scaling invariance, ν_t coincide with the image of ν_1 under the automorphism $d_{\sqrt{t}}$ of G, where $d_t(x, y, z) = (tx, ty, t^2z)$. Hence

$$\int_G f(g)d\nu^n(g) = \int_G f\circ d_{\sqrt{n}}(g)p(g)dg$$

where $p(g)$ is the density of ν. It is known that p as well as its derivatives are smooth, fastly decreasing functions on G. Let $L > 0$ be a Lipschitz constant for p in the sense that $|p(g_1) - p(g_2)| \leq L\,\|g_1 - g_2\|$ for all $g_1, g_2 \in G$ where $\|g\| = \max\{|x|, |y|, |z|\} \geq C$ for $g = (x, y, z)$. Fix $\varepsilon > 0$ and let $C > 0$ be such that

$$\int_{\|g\|\geq C} p(g)dg \leq \varepsilon$$

We now have

$$\left|\int_{\|g\|\geq C} f(d_{\sqrt{n}}(g))p(g)dg\right| \leq \varepsilon\,\|f\|_\infty$$

Let us denote by $R(g, h)$ the rectangle $[x, x + h) \times [y, y + h) \times [z, z + h)$ where $g = (x, y, z) \in G$ and $h > 0$. We can find a subdivision of the hypercube $\{\|g\| \leq C\}$ by small cubes of the form $R(g_i, h)$. Hence

$$\int_{\|g\|\leq C} f(d_{\sqrt{n}}(g))p(g)dg = \sum\int_{R(g_i,h)} f(\sqrt{n}x, \sqrt{n}y, nz)p(g)dg$$

Also

$$\left|\sum\int_{R(g_i,h)} f(\sqrt{n}x, \sqrt{n}y, nz)p(g)dg - \sum p(g_i)\int_{R(g_i,h)} f(\sqrt{n}x, \sqrt{n}y, nz)dg\right| \leq Lh\,\|f\|_\infty\,(2C)^3$$

Now, note that, by viewing $R(g,h)$ as a difference of several rectangles in $\mathbb{R}^3$, the assumption (3.2) made on f easily implies that for all $g \in \mathbb{G}$ and $h > 0$

$$\lim_{T\to+\infty} \frac{1}{T^2} \int f(x)\chi_{R(g,h)}(d_{1/\sqrt{T}}(x))dx = \ell \cdot h^3 \tag{3.66}$$

since $\chi_{R(g,h)} \circ d_{1/\sqrt{T}}$ can be written as a finite sum of terms of the form $\pm\chi_{[0,T_1]\times\ldots\times[0,T_3]}$ for some positive or negative $T_{1,\ldots,}T_3$. Hence by (3.66)

$$\lim_{T\to+\infty} \sum p(g_i) \int_{R(g_i,h)} f(\sqrt{n}x, \sqrt{n}y, nz)dg = \ell \cdot \sum p(g_i)h^3$$

But

$$\begin{aligned} \left|\sum p(g_i)h^3 - 1\right| &\leq \int_{\|g\|\geq C} p(g)dg + \sum \int_{R(g_i,h)} |p(g) - p(g_i)|\, dg \\ &\leq \varepsilon + Lh(2C)^3 \end{aligned}$$

Therefore, combining the above inequalities, for n large enough

$$\left|\int f(d_{\sqrt{n}}(g))p(g)dg - \ell\right| \leq \ell(\varepsilon + Lh(2C)^3) + Lh\, \|f\|_\infty (2C)^3 + \varepsilon\, \|f\| + \varepsilon$$

We finally obtain the desired result since h can be taken arbitrarily small. □

The second step is about comparing the integrals with respect to the probability measure μ and its associated gaussian distribution ν. Here, we make use of the uniform version of the local limit theorem (theorem 1.2). Namely,

Lemma 10.2 *Let f be a bounded and uniformly continuous function (with respect to either right or left uniform structure on G). Let μ be a compactly supported aperiodic and centered probability measure on G. And let ν be its associated gaussian distribution. Then we have*

$$\lim_{n\to+\infty} \left|\int f(g)d\mu^n(g) - \int f(g)d\nu^n(g)\right| = 0$$

Proof: We may assume $\|f\|_\infty \leq 1$. Fix $\varepsilon > 0$ and let $\omega > 0$ be a modulus of continuity for f relatively to ε, i.e. $|f(ux) - f(x)| \leq \varepsilon$ if $\|u\| \leq \omega$ and $x \in G$, where $\|g\| = \max\{|x|, |y|, |z|\}$ for $g = (x,y,z) \in G$. As follows from the central limit theorem, we can find a number $C > 0$ such that if we let $A_n = \{g = (x,y,z) \in G, |x|, |y| \leq C\sqrt{n}$ and $|z| \leq Cn\}$, we have for large n

$$\mu^{*n}(A_n^c) \leq \varepsilon$$

and

$$\nu^{*n}(A_n^c) \leq \varepsilon$$

We can then find a cover B_n of the cube A_n by less than $O(n^2/\omega^4)$ disjoint translates $R_\omega h$ of a small cube of the form $R_\omega = d_\omega(R)$ where $h \in d_\omega(G(\mathbb{Z}))$, d_ω is the dilation on G with coefficient of contraction ω, and R is a fundamental domain for the co-compact lattice $G(\mathbb{Z})$ in G. Now we can write

$$\begin{aligned}
\left|\int f d\mu^{*n} - \int f d\nu^{*n}\right| &\leq \left|\int_{B_n} f d\mu^{*n} - \int_{B_n} f d\nu^{*n}\right| + 2\varepsilon \\
&\leq \sum_i f(h_i)\left|\mu^{*n}(Rh_i) - \nu^{*n}(Rh_i)\right| + 4\varepsilon \\
&\leq O(\frac{n^2}{\omega^4}) \sup_{h\in G} \left|\mu^{*n}(Rh) - \nu^{*n}(Rh)\right| + 4\varepsilon
\end{aligned}$$

Here we can apply the uniform local limit theorem (theorem 1.2) and get

$$\lim_{n\to\infty} n^2 \sup_{h\in G} \left|\mu^{*n}(Rh) - \nu^{*n}(Rh)\right| = 0$$

Thus, we obtain the desired result. □

The proof of corollary 1.4 now follows immediately from the combination of the last two lemmas.

3.10.2 Unipotent random walks and a probabilistic version of Ratner's theorem

Here we shall conclude this paper and give a proof of theorem 1.5.

Let G be a connected real Lie group and Γ a lattice in G, that is, a discrete subgroup of G such that the homogeneous space G/Γ bears a finite Borel measure invariant by the left action of G. An element $u \in G$ is called *Ad-unipotent*, when the automorphism $Ad(u) \in GL(\mathfrak{g})$ of the Lie algebra $\mathfrak{g}$ of G is unipotent, i.e. every eigenvalue of $Ad(u)$ equals 1. A subgroup $U \subset G$ is called *Ad-unipotent* or simply *unipotent* is every element $u \in U$ is *Ad*-unipotent. The action of U on G/Γ is called a *unipotent flow*.

In the early nineties, in a series of papers (see [Rat1-3]), M. Ratner proved the validity in full generality of the Raghunathan-Dani conjectures for the action of connected *Ad*-unipotent subgroups on G/Γ. These results have had a number of far reaching applications (some of which were found earlier by proving special cases of the conjecture, like in Margulis's proof of the Oppenheim conjecture (1986)), especially to number theory and lattice points counting problems (see the recent survey [15]). The results, can be summarized as follows. First, if U is a connected *Ad*-unipotent subgroup of G, for every $x \in G/\Gamma$, the orbit Ux has a "nice algebraic" closure, that is, there exists a closed subgroup $H \subset G$ such that $\overline{Ux} = Hx$ is closed and bears a unique H-invariant probability measure m_x. Secondly, every U-ergodic probability measure on G/Γ is of the form m_x for some $x \in G/\Gamma$. We refer the reader to the surveys [141] and [160] for a detailed exposition of these results and further references (see also [115] for an alternative proof).

One of the main steps in the proof of the latter conjecture is the following equidistribution theorem for the action of one-parameter unipotent flows :

Theorem 10.3 (M. Ratner) *Suppose G is a Lie group and Γ a lattice in G. Let $U = \{u(t), t \in \mathbb{R}\}$ be a one-parameter Ad-unipotent subgroup of G. Then for any $x \in G/\Gamma$, there is a closed subgroup H of G, such that Hx is closed and bears an H-invariant probability m_x, and the orbit Ux is equidistributed in Hx with respect to m_x. In other words, for all continuous and bounded functions f on G/Γ, we have*

$$\lim_{T \to +\infty} \frac{1}{T} \int_0^T f(u(t)x)dt = \int_{Hx} f dm_x$$

Let us emphasize the fact that this equidistribution holds for every point $x \in G/\Gamma$ and not only almost everywhere with respect to some U-ergodic measure. This shows that every point behaves 'generically'.

Making use of the equidistribution theorem above, N. Shah (cf. [150]) subsequently extended this result to the action of an arbitrary simply connected *Ad*-unipotent subgroup U of G. Let us introduce N. Shah's result.

Let U be any simply connected nilpotent Lie group and $(v_1, ..., v_k)$ be a basis of the Lie algebra $\mathfrak{u}$ of U. This basis is called a *triangular basis* (or strong Malcev basis) if the subspaces spanned by $(v_i, ..., v_k)$ for any i are ideals of $\mathfrak{u}$, that is $[v_i, v_j] \in span(v_l, ..., v_k)$ where $k = \max\{i, j\} + 1$. Such a basis gives rise to polynomial coordinates on U, i.e. the map

$$\begin{array}{rcl} \phi & : & \mathbb{R}^k \to U \\ (t_1, ..., t_k) & \longmapsto & \exp(t_k v_k) \cdot ... \cdot \exp(t_1 v_1) \end{array}$$

is polynomial diffeomorphism. It also sends the Lebesgue measure on $\mathbb{R}^k$ to the Haar measure on U. With this terminology, Shah proved (cf. [150] Cor. 1.3.)

Theorem 10.4 (N. Shah) *Suppose G is a Lie group and Γ a lattice in G. Let U be a simply connected Ad-unipotent subgroup of G. Let $(v_1, ..., v_k)$ be a triangular basis for U, and $x \in G/\Gamma$. Then for any continuous and bounded function f on G/Γ,*

$$\lim_{T_1 \to \infty, ..., T_k \to \infty} \frac{1}{T_1 ... T_k} \int_{[0,T_1] \times ... \times [0,T_k]} f(\phi(t_1, ..., t_k)x) dt_1 ... dt_k = \int_{Hx} f dm_x$$

where m_x is the H-invariant probability measure on $\overline{Ux} = Hx$.

This theorem is precisely what we need to apply corollary 1.4 to the situation of theorem 1.5. Keeping the notations of the statement of theorem 1.5, let f be a compactly supported function on G/Γ, and suppose that U is isomorphic to the Heisenberg group with triangular basis given by (3.3) in section 3.2. Then the function $F(u) = f(ux) = f(\phi(u_x, u_y, u_z)x)$ is a bounded uniformly continuous function for the left uniform structure on U satisfying the condition of corollary 1.4 with limit $\ell = \int_{Hx} f dm_x$. Therefore this is the end of the proof and of this paper.

Acknowledgements I sincerely thank G.A. Margulis (my dissertation adviser) for suggesting this problem and offering his help and guidance throughout the last couple of years. I am also very grateful to M. Babillot and Y. Guivarc'h for many enlightening discussions.

Chapitre 4

Distributions diophantiennes et théorème limite local sur $\mathbb{R}^d$

Soit $S_n = X_1 + ... + X_n$ la somme de n variables aléatoires centrées à valeurs dans $\mathbb{R}^d$ qui sont indépendantes et de même loi μ. Le problème du théorème limite local est de préciser le comportement asymptotique quand n tend vers l'infini de l'espérance

$$\mathbb{E}(f(S_n)) = \int f d\mu^n$$

quand f est une fonction définie sur $\mathbb{R}^d$. Lorsque f est la fonction indicatrice d'un intervalle borné I, cette espérance est la probabilité de retour $\mathbb{P}(S_n \in I)$ dans I au temps n. Le théorème limite local (voir [30]) affirme que si la loi μ possède un moment d'ordre 2 fini et si son support n'est pas contenu dans une classe d'un sous-groupe fermé propre de $\mathbb{R}^d$ ("non-lattice case"), alors $n^{d/2}\mathbb{E}(f(S_n))$ converge vers l'intégrale de f sur $\mathbb{R}^d$ par rapport à un multiple de la mesure de Lebesgue dès que f est continue et *à support compact*.

Il est naturel de se demander si l'on peut préciser le comportement asymptotique de $\mathbb{E}(f(S_n))$, et en particulier estimer la vitesse de convergence dans le théorème limite local. La question de la vitesse de convergence a été relativement peu abordée dans la littérature. Sous l'hypothèse que la loi μ est absolument continue par rapport à la mesure de Lebesgue (ou plus généralement sous la condition de Cramér, i.e. le module de la fonction caractéristique de μ reste éloigné de 1 de façon uniforme hors d'un voisinage de 0), on dispose du développement classique de Edgeworth (voir [57]), qui fournit tous les termes suivants de l'asymptotique de $\mathbb{P}(S_n \in I)$ (sous réserve de l'existence de moments). Dans le cadre de la théorie du renouvellement, les questions de vitesse ont été étudiées en partie, notamment dans [38] où Carlsson obtient, pour le renouvellement sur $\mathbb{R}$, une asymptotique très précise de $\nu(]-\infty, x])$ quand x tend vers $+\infty$, où ν est la mesure de renouvellement. Il montre aussi que si ν vérifie une condition diophantienne, la convergence en est d'autant plus rapide. Mais le cas multi-dimensionnel et en particulier le théorème local (qui n'est qu'un cas particulier du théorème de renouvellement pour la chaîne mixte (S_n, n) en dimension $d+1$) n'a pas fait l'objet d'une étude détaillée à notre connaissance. Il se trouve que la vitesse de convergence dépend de façon significative de la distribution des X_i, et non plus seulement de l'existence de moments. Nous donnons, dans la première partie de cet article, la condition exacte sur la loi μ pour que la vitesse de convergence soit optimale (i.e. en $\frac{1}{n}$), et en définitive pour que le développement de Edgeworth tout entier ait lieu lorsque f est suffisamment différentiable ainsi que sa transformée de Fourier. Cette condition est d'ordre diophantien : elle demande que la variable aléatoire X_1 ne puisse être très bien approximée par une variable à valeurs dans une réunion $H + \mathbb{Z}v$ d'hyperplans affines de $\mathbb{R}^d$. On dira que μ est *diophantienne*. On donne le développement de Edgeworth en en précisant les premiers termes au théorème 2.7. Notons que seules des puissances entières de $\frac{1}{n}$ interviennent dans ce développement.

L'autre volet de cet article est consacré à l'étude de l'asymptotique de la suite $\mathbb{E}(f(S_n))$ pour des fonctions f qui ne sont plus nécessairement intégrables sur $\mathbb{R}^d$. Cette étude est motivée par des problèmes d'équidistribution de marches aléatoires. Etant donnée la trajectoire d'un flot issu d'un point x dans un espace X, on montre que si cette trajectoire est équidistribuée pour une certaine mesure finie sur X, alors toute marche aléatoire centrée évoluant le long de cette trajectoire s'équidistribue de la même façon (voir le corollaire 5.4). Dans ce problème, nous avons besoin de pouvoir comparer la marche S_n à la marche gaussienne correspondante sur un domaine plus large que celui fourni par le théorème limite local. C'est l'objet du théorème local de Stone (théorème 3.3) qui donne une estimation uniforme sur une boule dont le rayon est de l'ordre de $\sqrt{n}$. On en déduit en premier lieu une forme très simple du reste dans l'approximation gaussienne en fonction de f et de sa variation totale (voir théorème 4.4). On étend ensuite le théorème de Stone aux écarts modérés, i.e. à un domaine de taille $\sqrt{cn \log n}$, sous l'hypothèse de l'existence d'un moment d'ordre $c+2$ (théorème 3.2). Ceci permet de montrer que lorsque μ est diophantienne et sous une condition de moment, on a bien l'équivalent attendu de $\mathbb{E}(f(S_n))$ pour toute fonction f bornée sur $\mathbb{R}^d$ et suffisamment différentiable. Plus précisément, si f prend des valeurs positives, quand $n \to +\infty$

$$\frac{\mathbb{E}(f(S_n))}{\int_{\mathbb{R}^d} f(x\sqrt{n})p(x)dx} \to 1 \tag{4.1}$$

où p est la densité de la gaussienne associée (voir théorème 4.5). Lorsque μ satisfait la condition de Cramér, l'équivalent (4.1) a lieu pour toutes les fonctions höldériennes bornées. Il y a une compétition entre la régularité de μ et celle de f. Sans hypothèse de régularité sur μ ni sur f, cet équivalent est faux. Ainsi pour tout entier $p \geq 0$ il existe des distributions (que l'on peut même choisir diophantiennes) et des fonctions C^p, intégrables, et dont toutes les dérivées d'ordre $\leq p$ tendent vers 0 à l'infini, pour lesquelles la limite (4.1) n'a pas lieu (voir l'exemple 4.4.4). En revanche, sans hypothèse sur μ autre que celles du théorème local cette fois, on montre que si le dénominateur dans l'équivalent (4.1) tend vers 0, alors le numérateur aussi. Plus généralement, si les moyennes de Cesaro de f sur de grands rectangles convergent, alors la suite $\mathbb{E}(f(S_n))$ tend vers la même limite, pour toute fonction f uniformément continue et bornée (voir théorème 5.1). Il est assez remarquable que l'on puisse déterminer dans ce cas le comportement de $\mathbb{E}(f(S_n))$ sous ces seules hypothèses. L'application à l'équidistribution des marches aléatoires en découle aussitôt.

4.1 Notations et préliminaires

Bien que la discussion se déroulera le plus souvent (pour alléger les notations) dans le cas plus simple des variables aléatoires réelles, on indiquera aussi dans la dernière section comment généraliser ces résultats au cas multi-dimensionnel.

Soit $(X_n)_n$ une suite de variables aléatoires réelles indépendantes et de même loi. Dans cet article, on désignera par μ la mesure de probabilité sur $\mathbb{R}$ correspondant à la loi commune des variables X_n. Tout le long de ce texte, on fera l'hypothèse que la mesure μ possède un moment d'ordre 2 fini, c'est-à-dire que $\int_{\mathbb{R}} x^2 d\mu(x) < \infty$.

On notera $S_n = X_1 + ... + X_n$ la somme partielle des variables X_i jusqu'au temps n et μ^{*n}, ou simplement μ^n, la loi de la variable aléatoire S_n, à savoir la convolée n fois de la mesure μ. La convolution des mesures μ et ν est notée $\mu * \nu$. De même μ^{-1} désigne la mesure adjointe à μ : $\mu^{-1}(A) := \mu(-A)$ pour tout borélien A.

On emploiera également dans la suite les notations de théorie de la mesure, comme $\mu(A)$ et $\int f(x)d\mu(x)$, et les notations probabilistes, comme $\mathbb{P}(A)$ et $\mathbb{E}(f(X_1))$, en passant de l'une à l'autre dès que cela semble mieux adapté.

On désignera par $\{x\} = d(x, \mathbb{Z})$ la distance du réel x à un entier le plus proche, entier que l'on notera $[x]$.

On notera $\widehat{f}$ la transformée de Fourier de la fonction f, à savoir

$$\widehat{f}(x) = \int e^{-itx} f(t)dt$$

Aussi pour une mesure de probabilité μ, loi de la variable aléatoire X, on note $\widehat{\mu}$ sa fonction caractéristique $\widehat{\mu}(x) = \mathbb{E}(e^{ixX})$ et on notera σ_p le moment d'ordre p de μ s'il existe, i.e.

$$\sigma_p = \int x^p d\mu(x)$$

La formule d'inversion de Fourier s'écrit alors $\widehat{\widehat{f}}(x) = 2\pi f(-x)$.

On note $\mathbb{L}^1(\mathbb{R})$ l'espace de Banach des classes de fonctions intégrables, muni de la norme usuelle $\|\cdot\|_1$. De même, $\|\cdot\|_\infty$ désigne la norme pour la convergence uniforme des fonctions. On notera $\langle f, g\rangle$ le produit scalaire usuel sur $\mathbb{L}^2(\mathbb{R})$.

On dira que μ est *apériodique* si son support n'est pas contenu dans une progression arithmétique. Cela revient à dire que

$$\forall x \in \mathbb{R}\backslash\{0\} \ \ |\widehat{\mu}(x)| < 1.$$

On dit que μ satisfait la *condition de Cramér* si elle vérifie l'une des trois conditions équivalentes suivantes

$$\sup_{|x|>1} |\widehat{\mu}(x)| < 1 \Leftrightarrow \lim \inf_{|x|\to+\infty} |1-\widehat{\mu}(x)| > 0 \Leftrightarrow \lim \inf_{|x|\to+\infty} \int \{xa\} d\mu(a) > 0$$

4.2 Distributions diophantiennes

Dans cette section, on introduit une certaine condition sur la mesure μ qui, lorsqu'elle est satisfaite (et seulement dans ce cas), permet d'obtenir une vitesse de convergence optimale dans le théorème limite local pour les fonctions à support compact suffisamment régulières. Il s'agit en réalité d'une condition diophantienne sur μ, stipulant qu'en moyenne la variable aléatoire X de loi μ ne peut être très bien approchée par les points d'une progression arithmétique réelle. Comme nous l'expliquerons ci-dessous, cette condition est vérifiée dans "la plupart" des cas.

4.2.1 Définitions

Un nombre réel a est habituellement appelé diophantien si $\{qa\} \geq C/|q|^l$ pour tout entier non nul q et pour un certain $C > 0$ et un certain entier $l \geq 0$ fixés, où, rappelons-le, $\{x\} = d(x, \mathbb{Z})$ désigne la distance du réel x à l'entier le plus proche. Cette définition dépend du choix du réseau $\mathbb{Z}$ des entiers dans $\mathbb{R}$. De façon similaire, sur la droite affine cette fois, on dira qu'un ensemble de points S est **diophantien** s'il est mal approximé par une progression arithmétique réelle, c'est-à-dire plus précisément s'il existe $C > 0$ et $l \geq 0$ tels que

$$\inf_{y\in\mathbb{R}} \sup_{a\in S} \{xa + y\} \geq C/|x|^l$$

pour tout réel x assez grand (on dit alors que S est $2l$-diophantien). Ainsi le réel a est diophantien si et seulement si le triplet $\{0, 1, a\}$ est diophantien, et tout ensemble de points dont une partie est diophantienne est lui-même diophantien. Par analogie, on a la définition suivante.

Définition 2.1 *Soit μ une mesure de probabilité borélienne sur $\mathbb{R}$ et l un réel ≥ 0. On dit que μ est l-**diophantienne** s'il existe $C > 0$ tel que pour tout $x \in \mathbb{R}$ assez grand en valeur absolue,*

$$\inf_{y\in\mathbb{R}} \int \{xa + y\}^2 d\mu(a) \geq \frac{C}{|x|^l}$$

*On dit que μ est **diophantienne** si elle est l-diophantienne pour un réel $l \geq 0$.*

Cette notion signifie donc que la moyenne du carré des écarts à toute progression arithmétique est minorée par une puissance fixée du pas de la progression. L'introduction du carré permet d'avoir la caractérisation simple ci-dessous, mais ne joue pas un rôle primordial. Notons aussi que la condition "0-diophantienne" est équivalente à la condition de Cramér énoncée dans la section précédente.

Avant d'aller plus loin, donnons quelques définitions équivalentes de la notion de mesure diophantienne :

Proposition 2.2 *Soit $l \geq 0$. Les assertions suivantes sont équivalentes :*
(i) la mesure μ est l-diophantienne
(ii) il existe un réel $C > 0$ tels que pour tout x assez grand (en valeur absolue), on a

$$|\widehat{\mu}(x)| \leq 1 - \frac{C}{|x|^l}$$

(iii) la mesure symétrisée $\mu * \mu^{-1}$ est l-diophantienne
(iv) il existe $C > 0$ tel que pour tout $x \in \mathbb{R}$ assez grand en valeur absolue,

$$\int \{x(a-b)\}^2 d\mu(a) d\mu(b) \geq \frac{C}{|x|^l}$$

Preuve: L'équivalence entre (i) et (iv) est immédiate. Clairement l'assertion (ii) est équivalente à la même assertion pour la mesure symétrisée $\mu * \mu^{-1}$. En tenant compte de cette remarque, on obtient aisément l'équivalence entre (ii) et (iv) si l'on note qu'il existe deux constantes positives c_1 et c_2 telles que

$$c_1\{x\}^2 \leq 1 - \cos(2\pi x) \leq c_2\{x\}^2$$

L'équivalence de (iii) avec les assertions précédentes est évidente au vu de (ii). $\square$

On voit aussi instantanément que μ est diophantienne si et seulement si μ^p est diophantienne pour un entier $p \geq 1$ quelconque. Observons que si μ est diophantienne, alors son support l'est aussi. On note aussi que si la mesure μ n'est pas étrangère à la mesure de Lebesgue, d'après le lemme de Riemann-Lebesgue, μ satisfait la condition de Cramér donc est 0-diophantienne.

Remarque 2.3 *Il faut remarquer que la condition : il existe $l > 0$ tel que*

$$\lim \inf_{|x| \to +\infty} \left| x^l (1 - \widehat{\mu}(x)) \right| > 0$$

est en général strictement plus faible que la condition "μ diophantienne", contrairement à ce qui se passe quand $l = 0$. Pour voir cela, il suffit de considérer une mesure supportée par deux points $\{1, a\}$ avec a diophantien.

Supposons maintenant que μ est étrangère à la mesure de Lebesgue. Soit $\Omega(\mu)$ l'ensemble des réels z tels que

$$\lim_{\varepsilon \to 0} \frac{\mu(I_\varepsilon(z))}{2\varepsilon} = +\infty$$

où $I_\varepsilon(z)$ représente l'intervalle ouvert de longueur 2ε centré au point z. On sait que $\mu(\Omega(\mu)) = 1$ (voir par exemple [146]). On peut alors énoncer la

Proposition 2.4 *Supposons que μ soit étrangère à la mesure de Lebesgue. S'il existe un ensemble fini S inclus dans $\Omega(\mu)$ qui soit diophantien, alors μ est diophantienne. Si μ est à support fini, alors μ est l-diophantienne si et seulement si support S de μ est l-diophantien.*

Preuve: Soit X une variable aléatoire de loi μ. Soit $l \geq 0$ et $C > 0$ tels que

$$\inf_{y \in \mathbb{R}} \max_{s \in S} \{sx + y\} \geq C / |x|^l$$

pour tout x assez grand. Comme $S \subset \Omega(\mu)$ est fini, pour tout x assez grand, $\mathbb{P}(|X - s| < C/2\,|x|^{l+1}) > 1/\,|x|^{l+1}$. Pour x grand et $y \in \mathbb{R}$, on peut trouver $s \in S$ tel que $\{sx + y\} \geq C/\,|x|^l$. Donc si $|X - s| < C/2\,|x|^{l+1}$, alors $\{xX + y\} \geq C/2\,|x|^l$. D'où

$$\mathbb{E}(\{xX + y\}^2) \geq \mathbb{P}(|X - s| < C/2\,|x|^{l+1}) \frac{C^2}{4\,|x|^{2l}} \geq \frac{C^2}{4\,|x|^{3l+1}}.$$

Le reste de la proposition est immédiat. $\square$

Rappelons que pour presque tout choix de réels a, b et c, par rapport à la mesure de Lebesgue, le triplet $\{a, b, c\}$ est diophantien. Cependant, il est facile de construire un exemple de mesure μ à support dans $[0, 1]$, qui soit à la fois apériodique et non diophantienne, et même dont le support soit diophantien. Dans le premier exemple, la mesure μ est atomique de support $[0, 1]$, dans le second, elle est diffuse et son support contient $[0, 1]$.

Exemple 2.5 *Soit E_n, n entier ≥ 1, l'ensemble des rationnels r dont l'écriture en fraction irréductible est $r = \frac{p}{n}$ pour un certain entier p premier à n et compris entre 1 et n. Soit μ une mesure de probabilité qui assigne à E_n le poids $\frac{1}{2^{n!}}$ si $n \geq 2$ et assigne un même poids aux points de E_n et X_1, X_2 des variables aléatoires indépendantes de loi μ. Clairement*

$$\begin{aligned} \mathbb{E}(\{n!(X_1 - X_2)\}^2) &\leq 2\mathbb{P}(n! X_1 \notin \mathbb{Z}) \\ &\leq 2 \sum_{p > n} 1/2^{p!} \leq 1/2^{n!} \end{aligned}$$

donc μ n'est pas diophantienne mais son support est l'intervalle $[0, 1]$ tout entier.

Exemple 2.6 *Soit $(n_i)_{i \geq 1}$ une suite d'entiers ≥ 1 telle que $n_{i+1} \geq 2^{n_i}$. Soit $E = \{n_i\}_i$ la partie de $\mathbb{N}$ correspondante et K l'ensemble des réels x qui s'écrivent*

$$x = \sum_{i \in I} \frac{1}{2^{n_i}}$$

pour un choix quelconque d'une partie I de l'ensemble E, la somme valant 0 si $I = \emptyset$. On voit que K est un compact inclus dans $[0, 1]$ sans point isolé, totalement discontinu et *de mesure de Lebesgue nulle (Cantor). Il est possible de construire une fonction continue f croissante de l'intervalle $[0,1]$ dans lui-même, avec $f(0)=0$ et $f(1)=1$, telle que la dérivée f' est nulle en tout point de $[0,1]\backslash K$ et $+\infty$ en tout point de K. La mesure de Stieljes associée μ est une mesure de probabilité diffuse et supportée par le compact K. Comme K n'est pas diophantien, on voit que μ n'est pas diophantienne. On peut aussi construire une mesure diffuse non-diophantienne dont le support contient tout l'intervalle $[0,1]$ (par exemple en prenant la somme $c\cdot\sum\mu*\delta_{x_n}/2^{k_n}$ où $(x_n)_n$ est une suite dense dans $[0,1]$ convenablement choisie et $k_n\uparrow+\infty$, $c>0$).*

4.2.2 Une classe de fonctions analytiques de type exponentiel

Dans ce paragraphe, on introduit une classe d'exemples de fonctions sur $\mathbb{R}$ qui nous seront utiles dans les paragraphes suivants, en particulier en tant que source de contre-exemples. Il s'agit de fonctions analytiques sur $\mathbb{R}$ dont la transformée de Fourier est à support compact.

Soit $(a_n)_{n\in\mathbb{Z}}$ une suite bornée quelconque de nombres complexes et p un entier ≥ 1. On lui associe la fonction d'une variable complexe $\beta(z)$ définie comme suit

$$\beta(z)=\frac{1}{2^{2p}}\sin^{2p}(\pi z)\sum_{n\in\mathbb{Z}}\frac{a_n}{(z-n)^{2p}}.$$

C'est une fonction holomorphe sur $\mathbb{C}$ qui vérifie $|\beta(z)|\leq Ce^{2p\pi|z|}$ où C est une constante dépendant de $\sup_{n\in\mathbb{Z}}|a_n|$. Donc β est entière et de type exponentiel. Restreinte à la droite réelle, β est analytique et bornée.

Remarquons d'abord que si la suite $(a_n)_n$ est dans $\ell^1(\mathbb{Z})$ alors $\beta\in\mathbb{L}^1(\mathbb{R})$. Il résulte alors du théorème de Paley-Wiener que la transformée de Fourier $\widehat{\beta}(t)$ sur $\mathbb{R}$ est une fonction continue à support compact inclus dans $[-2p\pi,2p\pi]$ et

$$\int\widehat{\beta}(t)dt=2\pi\beta(0)=2\pi(\frac{\pi}{2})^{2p}a_0.$$

Supposons maintenant que la suite $(a_n)_n$ soit un $O(1/n^{2p})$ au voisinage de l'infini. Alors, on voit aisément que $\beta(x)$ est elle-même un $O(1/|x|^{2p-1})$ au voisinage de l'infini, lorsque $x\in\mathbb{R}$. La transformée de Fourier $\widehat{\beta}$ est donc au moins de classe C^l dès que $2(p-1)>l$.

Revenons au cas où la suite $(a_n)_n$ est seulement bornée et supposons de plus que $a_n\geq 0$ pour tout n. Alors on a

$$\beta(x)\geq a_{[x]}$$

où $[x]$ désigne l'entier le plus proche du réel x. En effet, $\beta(x)\geq a_{[x]}\left(\frac{\sin\pi x}{2\{x\}}\right)^{2p}\geq a_{[x]}$.

4.2.3 Développements de Edgeworth locaux pour les mesures diophantiennes

Lorsque μ est diophantienne et admet un moment d'ordre $r+2$, alors $\sqrt{n}\mathbb{E}(f(S_n))$ admet un développement asymptotique avec un reste de l'ordre de $o(1/n^{r/2})$ dès que f est assez régulière. C'est l'objet des deux théorèmes qui suivent. Ceux-ci généralisent aux mesures diophantiennes le développement de Edgeworth classique pour les densités qui est valide sous l'hypothèse que $\widehat{\mu}$ soit intégrable (voir [57]).

On pose, pour une fonction numérique f

$$C_r(f) = \max_{0\leq j\leq r+1} \left\|x^j f\right\|_1, \; C^k(f) = \max_{0\leq j\leq k} \left\|f^{(j)}\right\|_1, \; C_r^k(f) = C^k(f) + C_r(f)$$

Théorème 2.7 *Soit r un entier ≥ 0. Supposons que la mesure de probabilité μ est centrée et possède un moment d'ordre $r+2$ fini. Si de plus μ est l-diophantienne pour $l \geq 0$, alors pour toute fonction f de classe C^k avec $k > l(r+1)/2+1$ et telle que $C_r^k(f) < +\infty$, on a le développement asymptotique (de Edgeworth) suivant*

$$\sqrt{n}\mathbb{E}(f(S_n)) = \sum_{p=0}^{[r/2]} \frac{1}{n^p} \left\langle f, Q_p \right\rangle_{L^2(\mathbb{R})} + C_r^k(f)\cdot o(\frac{1}{n^{r/2}})$$

où le $o()$ ne dépend que de r et de la loi μ et les Q_p sont des polynômes de degré $\leq 2p$ dont les coefficients dépendent des moments de μ d'ordre $\leq p+2$.

Notons que seules des puissances entières n'interviennent dans le développement (comparer avec Feller [57] chap. 16 (2.12) et (4.10), au (2.13) il faut en fait lire H_6 au lieu de H_3 dans l'expression de P_4). De plus on calcule aisément les premiers termes. Ainsi

$$\sqrt{2\pi\sigma_2}Q_0(X) = 1, \; \sqrt{2\pi\sigma_2}Q_1(X) = -\frac{1}{\sigma_2}X^2 - \frac{\sigma_3}{\sigma_2^2}X + \frac{1}{8}(\frac{\sigma_4}{\sigma_2^2} - 3) - \frac{5}{24}\frac{\sigma_3^2}{\sigma_2^3}$$

Preuve: On écrit

$$\sqrt{n}\sqrt{2\pi\sigma_2}\mathbb{E}(f(S_n)) = \sqrt{n}\sqrt{\frac{\sigma_2}{2\pi}} \int \widehat{f}(x)\widehat{\mu}(x)^n dx \tag{4.2}$$

La preuve reprend les techniques de Fourier classiques exposées dans [57]. On commence par traiter dans le lemme suivant la partie de l'intégrale correspondant aux grandes valeurs de x. Ce lemme nous sera aussi utile plus tard.

Lemme 2.8 *Soit μ une mesure de probabilité sur $\mathbb{R}$. On suppose que μ est l-diophantienne. Soit $r > 0$ un réel positif. Alors il existe un réel $D(r,\mu) > 0$ tel que pour tout $D > D(r,\mu)$, on a*

$$\int_{|t|\geq\sqrt{\frac{D\log n}{n}}} \widehat{f}(t)\widehat{\mu}^n(t)dt = C^k(f)\cdot o_\mu(\frac{1}{n^r})$$

uniformément pour toute fonction f de classe C^k telle que $C^k(f) < +\infty$, où $k > lr+1$.

Preuve: Si $l=0$ la loi μ satisfait la condition de Cramér, donc le lemme est clairement vrai. Supposons $l>0$. Puisque μ est apériodique (car diophantienne) et $\widehat{\mu}$ continue, pour tout $\varepsilon>0$ et $x_0>0$, il existe c_0, $0<c_0<1$, tel que $|\widehat{\mu}(x)|\leq c_0$ quel que soit x avec $\varepsilon<|x|<x_0$. D'où

$$\left|\int_{\varepsilon<|x|<x_0}\widehat{f}(x)\widehat{\mu}(x)^n dx\right|\leq c_0^n\left\|\widehat{f}\right\|_\infty x_0\leq c_0^n C^k(f)x_0.$$

Puisque μ est l-diophantienne, il existe un réel x_0 tel que si $|x|\geq x_0$ alors

$$|\widehat{\mu}(x)|\leq\exp(-C/|x|^l)$$

pour une certaine constante $C>0$. Comme f est C^k et $C^k(f)<+\infty$, on a $\left|\widehat{f}(x)\right|\leq\left\|f^{(k)}\right\|_1/|x|^k$. Il vient

$$\begin{aligned}\left|\int_{|x|>x_0}\widehat{f}(x)\widehat{\mu}(x)^n dx\right| &\leq \int_{|x|>x_0}\frac{C^k(f)}{|x|^k}\exp(-Cn/|x|^l)dx \\ &\leq C^k(f)\frac{C'}{n^{(k-1)/l}}\int_0^{+\infty}\frac{e^{-u}}{u^{(l+1-k)/l}}du\end{aligned}\tag{4.3}$$

où C' est une constante indépendante de f. En choisissant $k>lr+1$, la quantité ci-dessus est au plus de l'ordre de $C^k(f)\cdot o(1/n^r)$ quand n croît.

Passons maintenant à la partie de l'intégrale où $|x|<\varepsilon$. Puisque μ a un moment d'ordre 2 fini, on peut trouver $\varepsilon>0$ et $c>0$ tels que si $|x|<\varepsilon$, on a $|\widehat{\mu}(x)|\leq\exp(-cx^2)$. Fixons $D>r/c$. Alors si $\sqrt{D\log n}<|x|<\varepsilon\sqrt{n}$, on a

$$\left|\widehat{\mu}(\frac{x}{\sqrt{n}})^n\right|\leq\exp(-cD\log n)\leq\frac{1}{n^{cD}}=o(\frac{1}{n^r}).$$

Il vient

$$\left|\int_{\sqrt{D\log n}<|x|<\varepsilon\sqrt{n}}\widehat{f}(\frac{x}{\sqrt{n}})\widehat{\mu}(\frac{x}{\sqrt{n}})^n\frac{dx}{\sqrt{n}}\right|\leq C^k(f)\cdot o(\frac{1}{n^r})$$

□

On peut donc maintenant se concentrer sur la partie de l'intégrale où $|x|<\sqrt{D\log n}$. On peut écrire au voisinage de zéro

$$\widehat{f}(x)=\sum_{j=0}^r\frac{\widehat{f}^{(j)}(0)}{j!}x^j+\frac{x^{r+1}}{(2r+1)!}\phi(x)$$

où ϕ est une fonction bornée par $\left\|\widehat{f}^{(r+1)}\right\|_\infty\leq C_r^k(f)$. Il vient

$$\int_{|x|<\sqrt{D\log n}}\widehat{f}(\frac{x}{\sqrt{n}})\widehat{\mu}(\frac{x}{\sqrt{n}})^n dx=\sum_{j=0}^r\frac{\widehat{f}^{(j)}(0)}{j!n^{j/2}}\Phi_j+\frac{1}{n^{(r+1)/2}}\Phi_{r+1}\tag{4.4}$$

où l'on a noté

$$\Phi_j = \int_{|x|<\sqrt{D\log n}} x^j \widehat{\mu}(\frac{x}{\sqrt{n}})^n dx$$

et

$$\begin{aligned}\Phi_{r+1} &= \frac{1}{(r+1)!}\int_{|x|<\sqrt{D\log n}} x^{r+1}\phi(\frac{x}{\sqrt{n}})\widehat{\mu}(\frac{x}{\sqrt{n}})^n dx \\ |\Phi_{r+1}| &\leq C_r^k(f)\int |x|^{r+1} e^{-cx^2} dx\end{aligned}$$

Comme μ est centrée et possède un moment d'ordre $r+2$ fini, il existe une fonction ψ_1 bornée continue et valant 0 en 0, telle qu'au voisinage de 0

$$\log \widehat{\mu}(x) = x^2(-\frac{\sigma_2}{2!} + P_r(x)) + x^{r+2}\psi_1(x)$$

où P_r est un polynôme de degré $\leq r$ tel que $P_r(0) = 0$ dont les coefficients dépendent seulement des moments de μ jusqu'à l'ordre $r+2$. Il vient,

$$\widehat{\mu}(\frac{x}{\sqrt{n}})^n = e^{-\frac{\sigma_2 x^2}{2}} \exp\left(x^2 P_r(\frac{x}{\sqrt{n}}) + \frac{x^{r+2}}{n^{r/2}}\psi_1(\frac{x}{\sqrt{n}})\right) \tag{4.5}$$

Si $r = 0$, on obtient directement le résultat du théorème en appliquant le théorème de convergence dominée de Lebesgue. Si $r \geq 1$, pour tout x tel que $|x| < \sqrt{D\log n}$, l'expression ci-dessus à l'intérieur de exp est un $O(\log^{3/2}(n)/\sqrt{n})$, d'où

$$e^{\frac{\sigma_2 x^2}{2}}\widehat{\mu}(\frac{x}{\sqrt{n}})^n = \sum_{m=0}^{r}\frac{1}{m!}\left[x^2 P_r(\frac{x}{\sqrt{n}})\right]^m + \frac{x^{r+2}}{n^{r/2}}\psi_1(\frac{x}{\sqrt{n}}) + O_{\mu,r,D}(\frac{\log^{3(r+1)/2}(n)}{n^{(r+1)/2}})$$

En regroupant les termes de la somme ci-dessus qui sont en facteur de $1/n^{i/2}$, on obtient des polynômes A_i (de degré au plus $3i$) tels que

$$e^{\frac{\sigma_2 x^2}{2}}\widehat{\mu}(\frac{x}{\sqrt{n}})^n = \sum_{i=0}^{r}\frac{1}{n^{i/2}}A_i(x) + \frac{x^{r+2}}{n^{r/2}}\psi_1(\frac{x}{\sqrt{n}}) + O_{\mu,r,D}(\frac{\log^{3(r+1)/2}(n)}{n^{(r+1)/2}}) \tag{4.6}$$

En tenant compte de (4.6), l'expression (4.4) devient

$$\sum_{i=0}^{2r}\sum_{j=0}^{2r}\frac{1}{n^{(i+j)/2}}\frac{\widehat{f}^{(j)}(0)}{j!}\int_{|x|<\sqrt{D\log n}} x^j A_i(x) e^{-\frac{\sigma_2 x^2}{2}} dx + C_r(f)\cdot o_{\mu,r,D}(\frac{1}{n^{r/2}})$$

En prenant D assez grand, on peut remplacer l'intégrale ci-dessus par une intégrale sur $\mathbb{R}$ tout entier. Notons maintenant que le polynôme A_i a même parité que i et dépend des moments de μ jusqu'à l'ordre $i+2$. Donc lorsque $i+j$ est impair,

$$\int_{\mathbb{R}} x^j A_i(x) e^{-\frac{\sigma_2 x^2}{2}} dx = 0$$

Il ne reste plus que des puissances entières dans le développement. Grâce au lemme ci-dessus, la preuve du théorème est maintenant complète. $\square$

Le théorème qui suit donne l'analogue pour les mesures diophantiennes du développement asymptotique classique pour les densités donné dans [57]. La suite de polynômes (P_p) ci-dessous est celle de [57] XVI Theorem 2 (pour le calcul des premiers termes, on renvoie à [45]) :

Théorème 2.9 *Soit r un entier ≥ 0 et l un réel ≥ 0. Supposons que la mesure de probabilité μ est centrée et possède un moment d'ordre $r+2$ fini. Si de plus μ est l-diophantienne, alors pour toute fonction f de classe C^k avec $k > l(r+1)/2+1$ et telle que $C^k(f) < +\infty$, on a le développement asymptotique (de Edgeworth) suivant*

$$\mathbb{E}(f(S_n)) = \sum_{p=0}^{r} \frac{1}{n^{p/2}} \int_{\mathbb{R}} f(x\sqrt{n}) P_p(x) g(x) dx + C^k(f) \cdot o(\frac{1}{n^{(r+1)/2}})$$

où le $o()$ ne dépend que de r et de la loi μ, où g est la densité de la gaussienne associée à μ dans le théorème limite central et où les P_p sont des polynômes de degré $\leq 3p$ dont les coefficients ne dépendent que des moments de μ jusqu'à l'ordre $p+2$.

La preuve reprend les mêmes étapes que celle du théorème précédent. En particulier, elle résulte aisément de la combinaison du lemme 2.8 avec le lemme suivant :

Lemme 2.10 *Soit μ une mesure centrée sur $\mathbb{R}$ admettant un moment d'ordre $r+2$. Soit D un réel assez grand ($D > r/\sigma_2$ par exemple suffit). Alors on a, uniformément pour toute fonction intégrable f,*

$$\int_{|t| \leq \sqrt{\frac{D \log n}{n}}} \widehat{f}(t) \widehat{\mu}^n(t) dt = \sum_{p=0}^{r} \frac{1}{n^{p/2}} \int f(x\sqrt{n}) P_p(x) g(x) dx + \|f\|_1 \cdot o_{r,D,\mu}(\frac{1}{n^{(r+1)/2}})$$

où l'on a gardé les notations du théorème 2.9.

Preuve: On peut reprendre l'estimée (4.6) qui s'écrit, sous les hypothèses du lemme :

$$\widehat{\mu}(\frac{t}{\sqrt{n}})^n = e^{-\frac{\sigma_2 t^2}{2}} \sum_{p=0}^{r} \frac{1}{n^{p/2}} A_p(t) + \frac{t^{r+2}}{n^{r/2}} e^{-\frac{\sigma_2 t^2}{2}} \psi_1(\frac{t}{\sqrt{n}}) + O(\frac{\log^{3(r+1)/2}(n)}{n^{(r+1)/2}}) e^{-\frac{\sigma_2 t^2}{2}}$$

Il vient alors, comme $\widehat{g}(t) = e^{-\frac{\sigma_2 t^2}{2}}$,

$$\int_{|t| \leq \sqrt{D \log n}} \widehat{f}(\frac{t}{\sqrt{n}}) \widehat{\mu}^n(\frac{t}{\sqrt{n}}) dt = \sum_{p=0}^{r} \frac{1}{n^{p/2}} \int_{|t| \leq \sqrt{D \log n}} \widehat{f}(\frac{t}{\sqrt{n}}) A_p(t) \widehat{g}(t) dt + o(\frac{1}{n^{r/2}}) \|f\|_1$$

On sait que le degré de A_p ne dépasse pas $3p$, donc d'après le choix de D, pour tout $p \leq r$

$$\left| \int_{|t| \geq \sqrt{D \log n}} \widehat{f}(\frac{t}{\sqrt{n}}) A_p(t) \widehat{g}(t) dt \right| \leq \|f\|_1 \, o_r(\frac{1}{n^{r/2}})$$

On peut donc substituer l'intégrale sur $|t| \leq \sqrt{D\log n}$ ci-dessus par une intégrale sur tout $\mathbb{R}$. Mais on a $t^p\widehat{g}(t) = (-i)^p\widehat{g^{(p)}}(t)$ pour tout entier $p \geq 0$, d'où

$$\int_{\mathbb{R}} \widehat{f}(\frac{t}{\sqrt{n}})A_p(t)\widehat{g}(t)dt = \sqrt{n}\int_{\mathbb{R}} f(x\sqrt{n})P_p(x)g(x)dx$$

où $P_p(x)$ est le polynôme tel que $2\pi A_p(-i\frac{d}{dx})g(x) = P_p(x)g(x)$. On a donc bien le résultat souhaité. $\square$

4.2.4 Caractérisation des mesures diophantiennes par la vitesse de convergence

Dans ce paragraphe, on démontre que, pour une mesure de probabilité μ centrée, la convergence dans le théorème limite local est en général plus lente que ou égale à $O(1/n)$ pour une fonction f à support compact fixée. On montre aussi que cet ordre de grandeur est atteint pour des fonctions suffisamment régulières si et seulement si la mesure μ est diophantienne. On va démontrer ainsi une réciproque au théorème 2.7. Plus précisément, on a la caractérisation suivante des mesures diophantiennes :

Théorème 2.11 *Soit μ une mesure de probabilité sur $\mathbb{R}$ centrée, et ayant un moment d'ordre 3 fini. Soit ν la mesure gaussienne qui lui est associée par le théorème de la limite centrale (i.e. μ est dans le domaine d'attraction de ν). Alors μ est diophantienne si et seulement s'il existe un entier $k_0 \geq 0$ tel que pour toute fonction f à support compact et de classe C^k $(k \geq k_0)$ sur $\mathbb{R}$, on ait*

$$\sup_{t\in\mathbb{R}}\left|\int f(t+\cdot)d\mu^n - \int f(t+\cdot)d\nu^n\right| = O(\frac{1}{n})$$

où $f(t+\cdot)$ désigne la translatée de f par le réel t et où O dépend de μ et de f.

Preuve: Supposons d'abord μ diophantienne. Elle est donc apériodique. On peut reprendre à l'identique le début de la preuve du théorème (2.7). Notons pour simplifier $f_t = f(t+\cdot)$. On remarque d'abord que $\widehat{f_t}(x) = e^{ixt}\widehat{f}(x)$. On obtient successivement pour $\varepsilon > 0$ et $x_0 > 0$ comme dans la preuve ci-dessus

$$\left|\int_{\varepsilon<|x|<x_0} \widehat{f_t}(x)\widehat{\mu}(x)^n dx\right| \leq c^n \|f\|_1$$

pour $c \in]0,1[$

$$\left|\int_{|x|>x_0} \widehat{f_t}(x)\widehat{\mu}(x)^n dx\right| \leq \sup_{0\leq p\leq k}\left\|f^{(p)}\right\|_1 \frac{cste}{n^{(k-1)/l}}$$

et

$$\left|\int_{\sqrt{D\log n}<|x|<\varepsilon\sqrt{n}} \widehat{f_t}(\frac{x}{\sqrt{n}})\widehat{\mu}(\frac{x}{\sqrt{n}})^n dx\right| \leq \|f\|_1 \cdot o(\frac{1}{n}) \quad (4.7)$$

Ces estimations sont valides uniformément en $t \in \mathbb{R}$ et sont aussi satisfaites par ν à la place de μ. On peut donc se concentrer sur le terme

$$\int_{|x|<\sqrt{D\log n}} \widehat{f}_t(\frac{x}{\sqrt{n}}) \left(\widehat{\mu}(\frac{x}{\sqrt{n}})^n - e^{-\frac{\sigma_2 x^2}{2}}\right) dx$$

que l'on majore uniformément en t par

$$\|f\|_1 \int_{|x|<\sqrt{D\log n}} \left|\widehat{\mu}(\frac{x}{\sqrt{n}})^n - e^{-\frac{\sigma_2 x^2}{2}}\right| dx$$

Mais, puisque μ a un moment d'ordre 3 fini, on peut écrire

$$\widehat{\mu}(x) = \exp\left(-\frac{\sigma_2}{2!}x^2 - i\frac{\sigma_3}{3!}x^3 + x^3\psi_0(x)\right)$$

pour une certaine fonction ψ_0 qui tend vers 0 quand x tend vers 0. Alors lorsque $|x| < \sqrt{D\log n}$,

$$\widehat{\mu}(\frac{x}{\sqrt{n}})^n - e^{-\frac{\sigma_2 x^2}{2}} = \frac{1}{\sqrt{n}} e^{-\frac{\sigma_2 x^2}{2}} \left(-i\frac{\sigma_3}{3!}x^3 + x^3\psi_1(\frac{x^3}{\sqrt{n}}, \frac{x}{\sqrt{n}})\right) \tag{4.8}$$

pour une certaine fonction ψ_1, avec $\psi_1(u) \to 0$ si $u \to 0$. Il vient

$$\int_{|x|<\sqrt{D\log n}} \left|\widehat{\mu}(\frac{x}{\sqrt{n}})^n - e^{-\frac{\sigma_2 x^2}{2}}\right| dx \leq \frac{c_0(\mu)}{\sqrt{n}} \tag{4.9}$$

où $c_0(\mu)$ est une constante ne dépendant que de μ. Donc il vient

$$\sup_{t\in\mathbb{R}} \left|\int f(t+\cdot)d\mu^n - \int f(t+\cdot)d\nu^n\right| \leq \frac{c(\mu)}{n}(\|f\|_1 + \sup_{0\leq p\leq k} \left\|f^{(p)}\right\|_1)$$

pour une certaine constante $c(\mu)$ ne dépendant que de μ.

Maintenant, passons à la réciproque. Nous allons d'abord établir la proposition suivante qui concerne le cas particulier des mesures symétriques :

Proposition 2.12 *Soit ν une mesure de probabilité sur $\mathbb{R}$ qui possède un moment d'ordre 3 fini. Soit $\mu = \nu * \nu^{-1}$. On suppose de plus que μ n'est pas diophantienne. Soit S_n la marche aléatoire associée à μ. Alors pour tout entier $p \geq 0$ et tout $\varepsilon > 0$, il existe une fonction f de classe C^p et à support compact telle que*

$$\lim_{n\to+\infty} \sup n^\varepsilon \left|\sqrt{n}\sqrt{2\pi\sigma_2}\mathbb{E}(f(S_n)) - \int f\right| > 0$$

Preuve: On fixe l'entier p et on considère la fonction β du paragraphe 4.2.2 associée à la suite $(a_n)_{n\in\mathbb{Z}}$ définie par $a_0 = 0$ et $a_n = \frac{1}{|n|^p}$ pour $n \neq 0$. On prend $f = \widehat{\beta}$. C'est une fonction de classe C^p à support compact et d'intégrale nulle. De plus $\widehat{f}(x) = 2\pi\beta(x)$ est réel et $> \frac{1}{[x]^p}$ pour tout x tel que $[x] \neq 0$. Comme dans la preuve du théorème ci-dessus, on écrit

$$\sqrt{n}\sqrt{2\pi\sigma_2}\mathbb{E}(f(S_n)) = \sqrt{n}\sqrt{\frac{\sigma_2}{2\pi}}\int \widehat{f}(x)\widehat{\mu}(x)^n dx$$

Comme ci-dessus (μ a un moment d'ordre 3), on voit que

$$\int_{|x|\leq\varepsilon} \widehat{f}(x)\widehat{\mu}(x)^n dx = o(\frac{1}{n})$$

Maintenant, fixons un entier $l \geq \frac{p}{\varepsilon}$. Remarquons que $\widehat{\mu}(x) \geq 0$, car on a supposé $\mu = \nu * \nu^{-1}$. Puisque μ n'est pas diophantienne, il existe une suite non bornée $(x_m)_m$ telle que

$$\widehat{\mu}(x_m) \geq e^{-\frac{1}{|x_m|^l}}$$

Montrons que chaque x_m est inclus dans un intervalle I_m de longueur au moins $\frac{\sqrt{2/\sigma_2}}{|x_m|^{l/2}}$ tel que tout $x \in I_m$ vérifie $\widehat{\mu}(x) \geq e^{-\frac{4}{|x_m|^l}}$. En effet, supposons que ce ne soit pas le cas, alors $\widehat{\mu}(x)$ admet un maximum dans l'intervalle $]a_m, b_m[$ où $b_m - a_m \leq \frac{\sqrt{2/\sigma_2}}{|x_m|^{l/2}}$ et $x_m \in]a_m, b_m[$ et $\widehat{\mu}(a_m) \leq e^{-\frac{4}{|x_m|^l}}$, $\widehat{\mu}(b_m) \leq e^{-\frac{4}{|x_m|^l}}$. Soit y_m l'abscisse de ce maximum. Alors

$$\widehat{\mu}(x) \geq \widehat{\mu}(y_m) - \frac{\sigma_2}{2}|x - y_m|^2$$

car la dérivée seconde $\widehat{\mu}(x)''$ est bornée par $\sigma_2 = \mathbb{E}(X^2)$. D'où $\widehat{\mu}(a_m) \geq e^{-\frac{1}{|x_m|^l}} - \frac{1}{|x_m|^l} \geq e^{-\frac{3}{|x_m|^l}}$, dès que m est assez grand. Ce qui fournit une contradiction.

Finalement, il vient

$$\begin{aligned}\int_{|x|\geq\varepsilon} \widehat{f}(x)\widehat{\mu}(x)^n dx &\geq \int_{x\in I_m} \widehat{f}(x)\widehat{\mu}(x)^n dx \\ &\geq \frac{1}{[x_m]^p} e^{-\frac{4n}{|x_m|^l}} \frac{\sqrt{2/\sigma_2}}{|x_m|^{l/2}}\end{aligned}$$

En prenant $n = [x_m]^l$, on obtient

$$\sqrt{n}\int \widehat{f}(x)\widehat{\mu}(x)^n dx \geq \frac{C}{n^{p/l}} \geq \frac{C}{n^\varepsilon}$$

pour une certaine constante $C > 0$. $\square$

Passons maintenant au cas général. La mesure μ n'est pas diophantienne, mais est centrée et possède un moment d'ordre 3 fini. Alors la mesure symétrisée $\widetilde{\mu} = \mu * \mu^{-1}$ non plus n'est pas diophantienne, mais elle possède aussi un moment d'ordre 3 fini. D'après la proposition ci-dessus, quel que soit $\varepsilon > 0$ et quel que soit l'entier $k \geq 0$, on peut trouver une fonction f à support compact et de classe C^k sur $\mathbb{R}$, que l'on choisit aussi d'intégrale nulle, telle que

$$\lim_{n\to+\infty} \sup n^{\frac{1}{2}+\varepsilon} \left| \int f d\widetilde{\mu}^n \right| > 0 \tag{4.10}$$

Or si μ satisfait la conclusion du théorème (2.11), alors

$$\sup_{t\in\mathbb{R}} \left| \int f(t+\cdot) d\mu^n \right| = O(\frac{1}{n})$$

car, puisque f est d'intégrale nulle, on a automatiquement

$$\sup_{t\in\mathbb{R}} \left| \int f(t+\cdot) d\nu^n \right| \leq \|xf\|_1 \, O(\frac{1}{n})$$

Mais alors

$$\left| \int f d\widetilde{\mu}^n \right| = \left| \int \mathbb{E}(f(S_n - t)) d\mu^n(t) \right| \leq \sup_{t\in\mathbb{R}} |\mathbb{E}(f(S_n - t))| = O(\frac{1}{n})$$

ce qui contredit (4.10) dès que $\varepsilon < \frac{1}{2}$. $\square$

On peut construire des mesures non diophantiennes qui fournissent des vitesses de convergence aussi lentes que l'on veut. A cet égard, on renvoie à l'exemple décrit dans la proposition (4.10) plus bas.

4.3 Théorème local pour les écarts modérés

4.3.1 Loi des écarts modérés

Le développement asymptotique précédent permet d'obtenir simplement un résultat de grandes déviations valable sans faire l'hypothèse habituelle d'existence d'un moment exponentiel. Ce résultat, qui traite des écarts modérés (i.e. du comportement de S_n dans un domaine de taille $\sqrt{cn \log n}$), a été démontré par Rubin et Sethuraman dans [145] (voir aussi [Nag2,3], [118] et [[158]]). Cependant, comme nous en aurons besoin dans la suite de cet article et que l'argument que nous présentons est très différent de l'argument original et se généralise automatiquement en dimension supérieure, nous choisissons de l'écrire ici.

Théorème 3.1 *([145]) Soit r un entier ≥ 1. Supposons que μ est une mesure de probabilité centrée qui possède un moment d'ordre $r+2$ fini. Alors pour tout c, $0 < c < r$, on a uniformément en x pour $1 \leq x \leq \sqrt{c \log n}$*

$$\lim_{n \to +\infty} \frac{\mathbb{P}\left(|S_n| > x\sqrt{\sigma_2 n}\right)}{\mathbb{P}\left(|N| > x\right)} = 1 \tag{4.11}$$

où N est une variable gaussienne normalisée. En particulier

$$\mathbb{P}\left(|S_n| > \sqrt{c\sigma_2 n \log n}\right) \sim \frac{4}{n^{c/2}\sqrt{2\pi c \log n}}$$

Preuve: On peut supposer $\sigma_2 = 1$. La preuve qui suit illustre bien le principe classique de lissage et “délissage” (voir par exemple [164]). Supposons d'abord que la fonction caractéristique $\widehat{\mu}$ est intégrable. Dans ce cas la variable $S_n/\sqrt{n}$ admet une densité continue f_n et l'on dispose du développement 2.9 qui n'est autre que le développement de Edgeworth classique pour les densités (cf. [45] ou [57] XVI, 2) et s'écrit :

$$f_n(t) - g(t) = \sum_{p=1}^{r} \frac{1}{n^{p/2}} P_p(t) g(t) dt + o_{r,\mu}(\frac{1}{n^{r/2}})$$

où g est la densité gaussienne normalisée. En intégrant sur $|t| < x$, il vient

$$\mathbb{P}\left(|S_n^\mu| > x\sqrt{n}\right) - \mathbb{P}\left(|N| > x\right) = K_n^\mu(x) + o(\frac{\sqrt{\log n}}{n^{r/2}}) \tag{4.12}$$

où $K_n^\mu(x) = \sum_{p=1}^{r} \int_{|t|>x} P_p g/n^{p/2}$ ne dépend que des moments de μ jusqu'à l'ordre $r+2$. Or on peut trouver une autre probabilité μ' dont la fonction caractéristique est intégrable, qui possède un moment exponentiel fini et dont tous les moments jusqu'à l'ordre $r+2$ coïncident avec ceux de μ. Pour μ' on dispose alors de la théorie des grandes déviations classiques due à Cramér. En particulier la limite uniforme (4.11) a bien lieu pour S_n' à la place de S_n et on a (voir [44] ou [57] XVI, 6) :

$$\mathbb{P}\left(|N| > \sqrt{c \log n}\right) \sim \frac{4}{n^{c/2}\sqrt{2\pi c \log n}} = o(\frac{1}{n^{c/2}})$$

Mais puisque (4.12) est aussi valide pour μ' et que $K_n^\mu = K_n^{\mu'}$ d'après le choix de μ', on obtient uniformément en x pour $1 \leq x \leq \sqrt{c \log n}$

$$\lim_{n \to +\infty} \frac{K_n^\mu(x)}{\mathbb{P}\left(|N| > x\right)} = 0. \tag{4.13}$$

En revenant à (4.12) pour μ cette fois, on obtient bien la limite (4.11) souhaitée. Ceci termine la preuve dans le cas où $\widehat{\mu}$ est intégrable.

Dans le cas général, il faut reprendre l'argument précédent pour la famille de mesures $\mu_\varepsilon = \mu * g_\varepsilon$ où g_ε est la gaussienne d'écart type ε. On remarque d'abord que l'estimation (4.12) ci-dessus reste valable uniformément quand appliquée aux mesures μ_{ε_n} si l'on a $1 > \varepsilon_n > \sqrt{D_0 \log n}/\sqrt{n}$ pour une certaine constante $D_0 = D_0(\mu) > 0$. Pour voir cela, il suffit de reprendre la preuve du théorème 2.9, c'est-à-dire celle des lemmes 2.10 et 2.8. Au lemme 2.10 appliqué à μ_{ε_n}, on remarque que le reste est un $o(1/n^{D_0/2})$ dès que $\varepsilon_n > \sqrt{D_0 \log n}/\sqrt{n}$. De même, au lemme 2.8, le reste est uniforme quand ε est petit. On obtient donc (4.12) uniformément pour les μ_{ε_n}. Dans la suite, on pose $\mu_n = \mu_{\varepsilon_n}$ avec $\varepsilon_n = \sqrt{D_0 \log n}/\sqrt{n}$.

Ensuite on vérifie que $K_n^\mu(x)$ et $K_n^{\mu_n}(x)$ sont comparables. Plus précisément on remarque que

$$K_n^{\mu_n}(x) = \int \int_{|t|\leq x} h_\mu(t - \varepsilon_n y) g(y) dy + o_{r,\mu}(\frac{1}{n^{r/2}})$$

où $h_\mu(t) = \sum_{p=1}^r P_p(t) g(t)/n^{p/2}$, $K_n^\mu(x) = \int_{|t|\leq x} h_\mu(t) dt$. Grâce à (4.13), on conclut que pour tout $c < r$ on a uniformément en x pour $1 \leq x \leq \sqrt{c \log n}$

$$\lim_{n\to+\infty} \frac{K_n^{\mu_n}(x)}{\mathbb{P}(|N| > x)} = 0$$

et

$$\lim_{n\to+\infty} \frac{\mathbb{P}\left(|S_n^{\mu_n}| > x\sqrt{\sigma_2(\mu_n) n}\right)}{\mathbb{P}(|N| > x)} = 1$$

Comme $\sigma_2(\mu_n) = 1 + \varepsilon_n^2$, et d'après le choix de ε_n on peut remplacer $\sigma_2(\mu_n)$ par 1 dans la limite précédente.

Finalement, on écrit

$$\mathbb{P}\left(|S_n^{\mu_n}| > x\sqrt{n}\right) = \int_{|t|\leq 2x} \mathbb{P}\left(|S_n^\mu| > (x + \varepsilon_n t)\sqrt{n}\right) g(t) dt + o(\mathbb{P}(|N| > x))$$

D'où

$$\liminf \frac{\mathbb{P}(|S_n^\mu| > (1 - 2\varepsilon_n) x\sqrt{n})}{\mathbb{P}(|N| > x)} \geq 1$$

et

$$\limsup \frac{\mathbb{P}(|S_n^\mu| > (1 + 2\varepsilon_n) x\sqrt{n})}{\mathbb{P}(|N| > x)} \leq 1$$

Mais d'après le choix de ε_n, $\mathbb{P}(|N| > x(1 + 2\varepsilon_n))/\mathbb{P}(|N| > x)$ tend vers 1 uniformément quand $1 \leq x \leq \sqrt{c \log n}$. De plus (4.11) est valable uniformément pour x borné d'après le théorème limite central. Cela termine la preuve. $\square$

4.3.2 Une version uniforme du théorème local

Le théorème limite local classique (voir [30]) donne un équivalent quand n tend vers l'infini de la probabilité $\mathbb{P}(S_n \in I)$ où I est un intervalle borné de $\mathbb{R}$. Analogue local du théorème de Rubin et Sethuraman, le théorème suivant permet d'estimer $\mathbb{P}(S_n \in I + x)$ lorsque x varie de façon quelconque dans une boule de rayon au plus $\sqrt{cn \log n}$, où la constance c dépend du plus grand moment fini de μ. On obtient donc un théorème local pour les écarts modérés en supposant seulement un moment fini d'ordre assez grand. En particulier, dans le résultat suivant, on ne fait pas d'hypothèse sur la régularité de μ. Les cas où μ est supportée sur $\mathbb{Z}$ ou possède une densité bornée ont déjà été traités dans [11] et [[158]]. Notons ν la mesure gaussienne sur $\mathbb{R}$ associée à μ dans le théorème limite central

Théorème 3.2 *Si μ est une mesure centrée et apériodique sur $\mathbb{R}$ de variance σ_2 et admettant un moment d'ordre $r \geq 2$. Soit $s > 0$ et $I = I_s = [-s, s[$ un intervalle borné centré en 0. Alors pour tout $x \in \mathbb{R}$*

$$\lim_{n \to +\infty} \frac{\mu^n(I + x)}{\nu^n(I + x)} = 1. \tag{4.14}$$

De plus, si $r > 2$ (resp. $r = 2$) pour tout $c \in]0, r-2[$, la limite (4.14) est uniforme en x et s quand $|x| + s \leq \sqrt{c\sigma_2 n \log n}$ (resp. $|x| + s = O(\sqrt{n})$) et s est minoré par un réel > 0.

Si de plus $\widehat{\mu}$ vérifie la condition de Cramér, alors il existe $\delta = \delta(\mu) > 0$ tel que la limite (4.14) est uniforme quand $|x| + s \leq \sqrt{c\sigma_2 n \log n}$ (resp. $|x| + s = O(\sqrt{n})$ si $r = 2$) mais $s > e^{-\delta n}$.

Cet énoncé sera la clé des théorèmes qui vont suivre dans l'étude de l'asymptotique de $\mathbb{E}(f(S_n))$ pour des fonctions qui ne sont pas à support compact ou bien pas intégrables.

Dans le cas $r = 2$ ce résultat est dû à Ch. Stone (voir [Sto]). Plus précisément :

Théorème 3.3 ***(Stone)*** *Si μ est une mesure centrée et apériodique sur $\mathbb{R}$ admettant un moment d'ordre 2, alors il existe une suite $(\varepsilon_n(\mu))_{n\geq 0}$ ne dépendant que de μ et tendant vers 0 telle que, pour tout intervalle fermé I de $\mathbb{R}$, on a*

$$\sup_{x \in \mathbb{R}} |\mu^n(I + x) - \nu^n(I + x)| \leq \frac{\varepsilon_n(\mu)}{\sqrt{n}}(1 + |I|)$$

où ν est la mesure gaussienne sur $\mathbb{R}$ associée à μ dans le théorème limite central et $|I|$ la mesure de Lebesgue de I.

On remarque en passant que le théorème limite central pour une telle mesure μ est un corollaire immédiat. Pour des valeurs de x plus grandes que $\sqrt{n}$ Stone a aussi démontré un théorème local pour les grandes déviations sous l'hypothèse habituelle d'existence d'un moment exponentiel fini (voir [163]). Des théorèmes semblables au théorème 3.3

ont été démontrés dans le cadre de la théorie du renouvellement (voir par exemple [39], [124], [159]) et notamment par Höglund qui obtient dans [89] un théorème de renouvellement pour les chaînes mixtes possédant à la fois une partie apériodique et une partie arithmétique, cadre qui généralise celui du théorème local. Keener dans [97] obtient un analogue du théorème de Stone pour le renouvellement sous la condition de Cramér et avec des conditions de moments légèrement différentes. Passons à la preuve du théorème 9.3 :

Preuve : La preuve qui suit reprend certaines idées de la preuve de (4.11) ainsi que de la preuve de Stone dans le cas $r = 2$. Notons $P_n^I(x) = \mu^n(x+I)$ et $Q_n^I(x) = \nu^n(x+I)$. On peut supposer que $\sigma_2 = 1$. On fixe une fonction intégrable paire, positive et continue $h \geq 0$ sur $\mathbb{R}$ dont la transformée de Fourier $\widehat{h}$ est de classe C^∞ et à support compact inclus dans $[-1, 1]$ et vérifie de plus $\int h = 1$ (voir le paragraphe 4.2.2). On introduit aussi les $h_\varepsilon(x) = \frac{1}{\varepsilon}h(\frac{x}{\varepsilon})$ qui forment une famille de Dirac quand ε tend vers 0. La preuve s'effectue en trois étapes : lissage, preuve pour la quantité lissée, et "délissage". Pour plus de clarté, on écrit ces étapes sous la forme de lemmes.

Lemme 3.4 *Il existe une suite $\varepsilon_n(\mu) \to 0$ ne dépendant que de μ telle que pour tout $\varepsilon > 0$ fixé on ait, uniformément en x et s :*

$$\left|P_n^I * h_\varepsilon(x) - Q_n^I * h_\varepsilon(x)\right| \leq \varepsilon_n(\mu) Q_n^I * h_\varepsilon(x) + |I| \cdot o_\varepsilon(\frac{1}{n^{(r-1)/2}}).$$

De plus si μ satisfait la condition de Cramér, alors on peut choisir ε de la forme e^{-an} pour un certain $a > 0$ petit et on obtiendra toujours un $o(1/n^{(r-1)/2})$ dans le reste ci-dessus.

Etant donnés $\varepsilon > 0$ et $x \in \mathbb{R}$, on pose $f_x(u) = \chi_{I+x} * h_\varepsilon(u)$. Alors $P_n^I * h_\varepsilon = \int f_x d\mu^n$. On peut appliquer à f_x le lemme 2.10 et on obtient pour $D > r/\sigma_2$

$$\int_{|t| \leq \sqrt{\frac{D\log n}{n}}} \widehat{f_x}(t)\widehat{\mu}^n(t)dt = 2\pi Q_n^I * h_\varepsilon + \sum_{p=1}^{r-2} \frac{1}{n^{p/2}} \int f_x(u) P_p(\frac{u}{\sqrt{n}}) d\nu^n(u) + |I| \cdot o_{r,D,\mu}(\frac{1}{n^{(r-1)/2}}) \tag{4.15}$$

Comme les polynômes P_p ne dépendent que de μ et sont de degré au plus $3p$, il existe une constante $C_1 = C_1(\mu)$ telle que

$$\int_{|u| \leq \sqrt{Dn\log n}} f_x(u) P_p(\frac{u}{\sqrt{n}}) d\nu^n(u) \leq C_1(\sqrt{D\log n})^{3p} Q_n^I * h_\varepsilon(x)$$

d'autre part, comme $P_p(u)g(u) \leq e^{-u^2/4}$ dès que u est assez grand (g est la densité de la loi normale),

$$\int_{|u| \geq \sqrt{Dn\log n}} f_x(u) P_p(\frac{u}{\sqrt{n}}) d\nu^n(u) \leq \frac{1}{n^{D/4}} \|f_x\|_1 \leq \frac{1}{n^{D/4}} |I|$$

D'où l'on conclut que pour $D \geq 2r$ et uniformément en $x \in \mathbb{R}$, (4.15) devient

$$\frac{1}{2\pi} \int_{|t| \leq \sqrt{\frac{D\log n}{n}}} \widehat{f_x}(t)\widehat{\mu}^n(t)dt = (1 + \varepsilon_n(x)) Q_n^I * h_\varepsilon(x) + |I| \cdot o_{r,D,\mu}(\frac{1}{n^{(r-1)/2}})$$

où la suite $\varepsilon_n(x)$ ne dépend que de μ et tend vers 0 uniformément en x.

Passons maintenant aux grandes valeurs de t. Notons que $\widehat{f_x}(t) = \widehat{\chi_{I+x}}(t)\widehat{h}(\varepsilon t)$:

$$\begin{aligned}\left|\int_{|t|\geq\sqrt{\frac{D\log n}{n}}}\widehat{f_x}(t)\widehat{\mu}^n(t)dt\right| &= \left|\int_{\sqrt{\frac{D\log n}{n}}\leq|t|\leq 1/\varepsilon}\widehat{f_x}(t)\widehat{\mu}^n(t)dt\right| \\ &\leq |I|\int_{\sqrt{\frac{D\log n}{n}}\leq|t|\leq 1/\varepsilon}|\widehat{\mu}(t)|^n\,dt = o_\varepsilon(\frac{1}{n^r})\end{aligned}$$

dès que $D > r/c$ où c est la constante définie dans la preuve du lemme 2.8. On remarque que si μ satisfait la condition de Cramér, alors on peut choisir ε de la forme e^{-an} pour un certain $a > 0$ petit et on obtiendra toujours un $o(1/n^r)$ dans la ligne de calcul ci-dessus. En prenant D assez grand et en combinant les majorations ci-dessus on obtient l'estimation du lemme 3.4. □

Lemme 3.5 *Pour tout $\delta > 0$, il existe $\varepsilon_0 > 0$ tel que si x et s sont des $O(\sqrt{n\log n})$, alors pour tout n assez grand et tout $\varepsilon \in]0,\varepsilon_0[$ on a uniformément en x et $I = I_s = [-s,s[$*

$$\left|Q_n^I * h_\varepsilon(x) - Q_n^I(x)\right| \leq \delta\cdot Q_n^I(x)$$

Soit $c > 0$ une constante telle que $|x| + s \leq \sqrt{cn\log n}$. Notons que pour tous réels u et t

$$|g(u+\varepsilon t) - g(u)| = g(u)(e^{\varepsilon ut}e^{-\varepsilon^2t^2/2} - 1) \leq \frac{\delta}{2}g(u)$$

dès que $|ut| \leq 1$, $|t| \leq 1$ et pour tout ε assez petit (disons $\leq \varepsilon_0$). Il vient

$$\left|Q_n^I(x-\varepsilon y) - Q_n^I(x)\right| \leq \frac{\delta}{2}Q_n^I(x) \tag{4.16}$$

tant que $\varepsilon < \varepsilon_0$ et $|y| \leq \sqrt{n}/\sqrt{c\log n}$. D'où

$$\left|Q_n^I * h_\varepsilon(x) - Q_n^I(x)\right| \leq \frac{\delta}{2}\cdot Q_n^I(x) + 2\frac{|I|}{\sqrt{n}}\int_{|y|\geq\sqrt{n}/\sqrt{c\log n}}h(y)dy.$$

Mais $\widehat{h}$ est de classe C^∞ donc $h(y) = o(\frac{1}{|y|^k})$ au voisinage de l'infini pour tout $k \geq 0$. Donc

$$\int_{|y|\geq\sqrt{n}/\sqrt{c\log n}}h(y)dy = o(\frac{1}{n^k}). \tag{4.17}$$

Mais

$$\sqrt{2\pi}Q_n^I(x) \geq \frac{|I|}{\sqrt{n}}\frac{1}{n^{c/2}} \tag{4.18}$$

donc pour tout n assez grand on a le résultat souhaité. □

Nous allons maintenant terminer la preuve du théorème. On suppose désormais que $s = |I|/2 > \varepsilon^{1/3}$, et l'on note $I_\varepsilon = [-s-\sqrt{\varepsilon}, s+\sqrt{\varepsilon}[$ et $I^\varepsilon = [-s+\sqrt{\varepsilon}, s-\sqrt{\varepsilon}[$. Remarquons d'abord qu'il existe $\theta(\varepsilon)$ satisfaisant $\theta(\varepsilon) \to 0$ si $\varepsilon \to 0$ telle que, dans les conditions du lemme 3.5,

$$\left|Q_n^{I_\varepsilon}(x) - Q_n^{I^\varepsilon}(x)\right| \leq \theta(\varepsilon) \cdot Q_n^I(x). \tag{4.19}$$

Ensuite on a $P_n^{I_\varepsilon} * h_\varepsilon(x) \geq P_n^I(x) \int_{|y|\leq 1/\sqrt{\varepsilon}} h(y)dy$ puis d'après les lemmes 3.4 et 3.5 pour ε assez petit

$$\frac{P_n^I(x)}{1-\varepsilon} \leq P_n^{I_\varepsilon} * h_\varepsilon(x) \leq (1+\varepsilon_n(\mu))(1+\delta)Q_n^{I_\varepsilon}(x) + |I_\varepsilon| \cdot o_\varepsilon(\frac{1}{n^{(r-1)/2}}).$$

Puis en tenant compte de (4.18) (on a supposé $c \leq r-2$) et de (4.19) on obtient que pour tout n assez grand

$$P_n^I(x) \leq (1+2\delta)Q_n^I(x). \tag{4.20}$$

On remarque aussi que si la condition de Cramér est satisfaite, on peut prendre $\varepsilon = e^{-3an}$ pour un certain $a > 0$ donc, pour tout n grand, (4.20) est vraie uniformément quand $s + |x| \leq \sqrt{cn\log n}$ et $s > e^{-an}$.

De façon similaire, on a

$$P_n^{I^\varepsilon} * h_\varepsilon(x) \leq P_n^I(x) \int_{|y|<1/\sqrt{\varepsilon}} h(y)dy + \int_{|y|\geq 1/\sqrt{\varepsilon}} P_n^{I^\varepsilon}(x-\varepsilon y)h(y)dy$$

et d'après (4.20), (4.19) et (4.16), en posant $A_n = [1/\sqrt{\varepsilon}, \sqrt{n}/\sqrt{c\log n}]$ dès que ε est assez petit,

$$\begin{aligned}
\int_{|y|\in A_n} P_n^{I^\varepsilon}(x-\varepsilon y)h(y)dy &\leq (1+2\delta)\int_{|y|\in A_n} Q_n^{I^\varepsilon}(x-\varepsilon y)h(y)dy \\
&\leq (1+2\delta)(1+\theta(\varepsilon))\int_{|y|\in A_n} Q_n^I(x-\varepsilon y)h(y)dy \\
&\leq Q_n^I(x)(1+2\delta)(1+\theta(\varepsilon))(1+\delta)\int_{|y|\geq 1/\sqrt{\varepsilon}} h(y)dy \\
&\leq \delta Q_n^I(x)
\end{aligned}$$

et par (4.17) pour tout $k \geq 0$ fixé, pour n assez grand

$$\int_{|y|\geq\sqrt{n}/\sqrt{c\log n}} P_n^{I^\varepsilon}(x-\varepsilon y)h(y)dy \leq \frac{1}{n^k}.$$

Finalement par (4.18) pour ε assez petit et si $s + |x| \leq \sqrt{cn\log n}$, $s > \sqrt{\varepsilon}$, on a pour n grand

$$P_n^{I^\varepsilon} * h_\varepsilon(x) \leq P_n^I(x) + \delta Q_n^I(x). \tag{4.21}$$

Mais d'après le lemme 3.4 et (4.18)

$$P_n^{I^\varepsilon} * h_\varepsilon(x) = (1+o_\varepsilon(1))Q_n^{I^\varepsilon} * h_\varepsilon(x).$$

puis en appliquant le lemme 3.5 puis (4.19)

$$P_n^{I^\varepsilon} * h_\varepsilon(x) \geq (1 - |o_\varepsilon(1)|)(1-\delta)Q_n^{I^\varepsilon}(x) \geq (1-2\delta)Q_n^I(x)$$

pour n assez grand. En combinant ceci avec (4.21) on obtient

$$P_n^I(x) \geq (1-3\delta)Q_n^I(x). \tag{4.22}$$

On a donc bien la majoration (4.20) et la minoration (4.22) souhaitées. La preuve du théorème est complète. $\square$

Remarque 3.6 *Lorsque μ est l-diophantienne, il est possible de préciser le comportement asymptotique de la suite $(\varepsilon_n(\mu))_n$ intervenant dans (3.3). Le calcul donne $\varepsilon_n(\mu) = o(\frac{1}{n^{\delta/2}})$ pour tout $0 < \delta < \frac{1}{l}$ si μ possède un moment d'ordre 3. De même, si μ est l-diophantienne, l'équivalent est valable uniformément quand $s > \frac{1}{n^{\delta/2}}$ pour tout $0 < \delta < \frac{1}{l}$.*

Remarque 3.7 *On donne plus bas un exemple (c.f. proposition 4.10) qui montre que la convergence vers 0 de $\varepsilon_n(\mu)$ peut être aussi lente que l'on veut si l'on ne fait pas d'hypothèse sur μ. Plus exactement, pour toute suite de réels strictement positifs $\varepsilon_n > 0$ tendant vers 0, on peut trouver une mesure μ (non diophantienne en général) telle que $\limsup \varepsilon_n(\mu)/\varepsilon_n > 0$.*

4.4 Un théorème limite pour les fonctions bornées

Nous allons maintenant passer au second volet de cet article et nous intéresser au comportement asymptotique de $\mathbb{E}(f(S_n))$ lorsque la fonction f est seulement supposée bornée et assez régulière. Commençons par remarquer que pour toute fonction bornée f ayant une limite l en $\pm\infty$, les moyennes $\mathbb{E}(f(S_n))$ convergent vers l. Cela résulte en effet immédiatement du théorème limite central. Dans les paragraphes suivants, nous allons montrer que, en supposant μ ou f assez régulière mais sans aucune hypothèse sur le comportement à l'infini de f, ces moyennes sont asymptotiquement indépendantes de la marche aléatoire centrée de variance 1 que l'on considère (théorème 4.5).

Nous commençons dans le premier paragraphe par étendre le théorème local aux fonctions directement Riemann intégrables.

4.4.1 Fonctions directement Riemann intégrables

Dans ce paragraphe, on s'intéresse à la validité du théorème limite local pour des fonctions qui ne sont plus nécessairement à support compact mais toujours intégrables.

Théorème 4.1 *Soit μ une mesure de probabilité centrée et apériodique sur $\mathbb{R}$ avec moment d'ordre 2 fini. Soit f une fonction Riemann intégrable sur $\mathbb{R}$ telle que*

$$\sum_{n\in\mathbb{Z}} \max_{[x]=n} |f(x)| < +\infty \tag{4.23}$$

Alors on a le théorème limite local pour f, c'est-à-dire que

$$\lim_{n\to+\infty} \sqrt{2\pi\sigma_2 n} \int f d\mu^n = \int f(x)dx \tag{4.24}$$

Preuve: On reprend la preuve classique du théorème local (voir [30]). Notons $a_n = \max_{[x]=n} |f(x)|$ pour tout entier $n \in \mathbb{Z}$, et considérons la fonction

$$\beta(x) = \frac{1}{2^2} \sin^2(\pi x) \sum_{n\in\mathbb{Z}} \frac{a_n}{(x-n)^2}$$

On peut supposer que la suite (a_n) n'est pas identiquement nulle. D'après le paragraphe (4.2.2) β est continue, strictement positive sur $\mathbb{R}$, et appartient à $\mathbb{L}^1(\mathbb{R})$. De plus $\beta(x) \geq a_{[x]} \geq |f(x)|$. Enfin $\widehat{\beta}$ est à support compact. Considérons la suite de mesures positives $d\nu_n(x) = \sqrt{2\pi\sigma_2 n}\beta(x)d\mu^n(x)$. Puisque $\widehat{\beta}$ est à support compact, la suite des masses totales $\nu_n(\mathbb{R})$ est bornée. De plus, pour tout réel t_0,

$$\begin{aligned} \int e^{it_0 x} d\nu_n(x) &= \sqrt{2\pi\sigma_2} \int \widehat{\beta}(\frac{t}{\sqrt{n}} - t_0)\widehat{\mu}^n(\frac{t}{\sqrt{n}})dt \\ &\longrightarrow {}_{n\to+\infty}\widehat{\beta}(-t_0) = \int e^{it_0 x}\beta(x)dx \end{aligned}$$

On en déduit que la suite de mesures $(\nu_n)_n$ converge (convergence usuelle des mesures) vers la mesure βdx.

La fonction f/β est bien définie, Riemann intégrable et bornée par 1, donc

$$\lim_{n\to+\infty} \sqrt{2\pi\sigma_2 n} \int f d\mu^n = \lim_{n\to+\infty} \int \frac{f}{\beta} d\nu_n(x) = \int f(x)dx$$

ce qu'il fallait démontrer. □

Remarque 4.2 *On qualifie de directement Riemann intégrables les fonctions f Riemann intégrables satisfaisant la condition (4.23). La notion est introduite par Feller dans [57] et étudiée en détail (la définition de cette notion peut prendre plusieurs formes équivalentes) dans [88]. Le théorème de renouvellement de Höglund [89] est énoncé pour toute fonction directement Riemann intégrable. Le théorème 4.1 est donc un cas particulier du résultat de Höglund, bien que la méthode de preuve soit ici différente (comparer prop. 2.2.).*

Remarque 4.3 *En général, il y a des exemples de fonctions f intégrables sur $\mathbb{R}$ et de classe C^p, $p \geq 1$, dont toutes les dérivées jusqu'à l'ordre de p tendent vers 0 à l'infini, qui ne vérifient pas (4.24). Plus bas, on donne un exemple d'une telle fonction pour une distribution μ qui est même diophantienne. Cependant, on verra au paragraphe 4.4.3 que si μ satisfait la condition de Cramér, alors (4.24) a bien lieu dès que f est intégrable et hölderienne sur $\mathbb{R}$.*

4.4.2 Fonctions intégrables à variation bornée

Le résultat précédent montre en particulier que le théorème local est valide pour les fonctions continues intégrables et décroissantes au loin quand $|x| \to +\infty$. En fait, par une méthode différente, on a aussi le résultat suivant, valide pour les fonctions à variation bornée.

Théorème 4.4 *Soit à nouveau μ une mesure de probabilité centrée et apériodique sur $\mathbb{R}$ ayant moment d'ordre 2 fini. Soit $(\varepsilon_n(\mu))_n$ la suite qui tend vers 0 intervenant dans l'énoncé du théorème de Stone (théorème 4.1). Alors pour toute fonction f intégrable et à variation bornée sur $\mathbb{R}$,*

$$\left| \int f d\mu^n - \int f d\nu_n \right| \leq \frac{\varepsilon_n(\mu)}{\sqrt{n}} \left(\int |f| + Var(f) \right)$$

où $Var(f)$ est la variation totale de f, et $d\nu_n(x) = \frac{\exp(-x^2/2\sigma_2 n)}{\sqrt{2\pi\sigma_2 n}} dx$ est la suite de mesures gaussiennes associée à μ.

Comme on l'a remarqué (cf. remarque 3.6), le comportement de la suite $(\varepsilon_n(\mu))_n$ peut être précisé lorsque l'on suppose que μ est diophantienne. Passons à la preuve du théorème.

Preuve: Clairement, on peut supposer que f est à support compact (considérer $f(x)\chi_{|x|<M}$ et faire tendre M vers l'infini à n fixé et remarquant que $Var(f\chi_{|x|<M}) \to Var(f)$). De même, en approximant f uniformément par des fonctions en escalier de la forme $\sum a_i \chi_{A_i}$ où $A_i = [x_i, x_{i+1})$ pour une certaine subdivision $x_0 < ... < x_N$, on se ramène au cas où f a précisément cette forme. Finalement, puisque $Var(f) = Var(f^+) + Var(f^-)$ et $|f| = f^+ + f^-$, où f^+ et f^- sont respectivement les parties positives et négatives de f, on peut supposer que f ne prend que des valeurs positives, i.e. $a_i \geq 0$ pour tout i.

Soit $0 = f_0 < f_1 < ... < f_K$ la suite ordonnée des valeurs prises par la fonction f. Maintenant considérons la famille $(B_j)_j$ de toutes les composantes connexes des ensembles de niveau $\{x | f(x) \geq f_i\}$ pour i variant de 1 à K. Alors

$$f(x) = \sum (f_i - f_{i-1}) \chi_{f(x) \geq f_i} = \sum \beta_j \chi_{B_j}(x)$$

où l'on définit β_j par $f_i - f_{i-1} \geq 0$ si B_j est une composante de l'ensemble $\{x | f(x) \geq f_i\}$. Remarquons que chaque B_j est une union finie d'intervalles A_i adjacents.

D'après le théorème de Stone (3.3), on obtient

$$\left| \int f d\mu^n - \int f d\nu_n \right| \leq \frac{\varepsilon_n(\mu)}{\sqrt{n}} \sum \beta_j (|B_j| + 1) \leq \frac{\varepsilon_n(\mu)}{\sqrt{n}} \left(\int f + \sum \beta_j \right)$$

Mais la construction des intervalles B_j est de telle sorte que

$$\sum \beta_j = \frac{1}{2} Var(f)$$

En fait, parmi toutes les représentations de f en sommes de la forme $\sum \beta_j \chi_{B_j}$ avec des $\beta_j \geq 0$ et des intervalles B_j, celle definie ci-dessus minimise $\sum \beta_j$. $\square$

4.4.3 Fonctions bornées

Ici nous allons obtenir l'équivalent annoncé dans l'introduction pour des fonctions bornées, sous l'hypothèse que μ est diophantienne (resp. Cramér). Plus précisément, nous avons :

Théorème 4.5 *Soit μ une mesure de probabilité centrée sur $\mathbb{R}$ et l un réel ≥ 0. On suppose de plus que μ possède un moment fini d'ordre 4 et on note ν la mesure gaussienne associée. Enfin on suppose que μ est l-diophantienne. Soit k_0 un entier tel que $k_0 > 3l/2+1$. Alors toute fonction f définie sur $\mathbb{R}$ et telle que f et toutes ses dérivées jusqu'à l'ordre k_0 sont bornées vérifie la relation suivante*

$$\int f d\mu^n - \int f d\nu^n = o(\int |f| d\nu^n) \tag{4.25}$$

En particulier, si $f \geq 0$ non identiquement nulle

$$\frac{\int f d\mu^n}{\int f d\nu^n} \to 1 \tag{4.26}$$

Si l'on suppose de plus que μ vérifie la condition de Cramér, alors (4.25) et (4.26) sont vérifiées pour toute fonction f höldérienne bornée sur $\mathbb{R}$.

Preuve: Comme on le vérifie aisément, pour chaque n, on peut trouver deux fonctions f_n et ϕ_n de classe C^{k_0} telles que $f = f_n + \phi_n$ avec f_n à support dans $[-s_n - 1, s_n + 1]$ et $\phi_n(x) = 0$ si $|x| \leq s_n$ où $s_n = 2\sqrt{n \log n}$. On peut de plus choisir f_n telle que $C^{k_0}(f_n) \leq C\sqrt{n \log n}$ pour une certaine constante C. Appliquons alors le théorème 2.9 aux fonctions f_n. Il vient aussitôt

$$\left| \int f_n d\mu^n - \int f_n d\nu^n \right| \leq \varepsilon_n \int |f_n| d\nu^n + C\frac{\sqrt{\log n}}{n}$$

où ε_n est un $O(\frac{\log^3 n}{n})$ indépendant de f. De plus, d'après la loi des écarts modérés (théorème 4.11) $\mu^n(|x| \geq s_n)$ et $\nu^n(|x| \geq s_n)$ sont des $o(1/n^{3/2})$. Enfin, d'après le théorème local lui-même, il existe $c > 0$ tel que $\int |f| d\nu^n \geq c/\sqrt{n}$ si f n'est pas identiquement nulle. Le théorème s'ensuit aussitôt.

Si l'on suppose que μ satisfait la condition de Cramér, alors on écrit $f = \sum f_i + R_n$ où f_i est la restriction de f à l'intervalle I_i et R_n la restriction au complémentaire de $[-s_n, s_n]$. Ces intervalles I_i sont de longueur $e^{-\delta n}$ et subdivisent $[-s_n, s_n]$ (où $\delta > 0$ est la constante intervenant dans le théorème 3.2). Choisissons dans chaque I_i un point x_i. Alors

$$\int_{I_i} |f - f(x_i)| \, d\mu^n \leq Lip_\alpha(f) |I_i|^\alpha \mu^n(I_i)$$

et de même pour ν^n. D'après le théorème 3.2, il existe une suite $\varepsilon_n \to 0$ telle que pour tout i

$$|\mu^n(I_i) - \nu^n(I_i)| \leq \varepsilon_n \nu^n(I_i)$$

Enfin

$$\int |R_n| d\mu^n \leq \|f\|_\infty \mu^n(x, |x| \geq s_n)$$

Ainsi d'après la loi des écarts modérés 4.11, pour tout n assez grand,

$$\begin{aligned}\left|\int f\mu^n - \int f\nu^n\right| &\leq \varepsilon_n \sum_i f(x_i)\nu^n(I_i) + 2Lip_\alpha(f)e^{-\delta\alpha n} + \|f\|_\infty o(1/n^{3/2}) \\ &\leq 2\varepsilon_n \int |f| d\nu^n\end{aligned}$$

□

Remarque 4.6 *Remarquons que le théorème reste valable dans le cas où μ satisfait la condition de Cramér pour les fonctions f qui sont höldériennes par morceaux sur $\mathbb{R}$, si ces morceaux sont de longueur bornée inférieurement par rapport à* 0 *et si les constantes de Hölder sur les morceaux sont uniformément bornées.*

Remarque 4.7 *Si l'on suppose seulement μ diophantienne et si la fonction f possède un nombre p insuffisant de dérivées successives bornées, alors l'équivalent* (4.26) *n'est pas valable. On donne au paragraphe suivant un tel exemple avec p arbitrairement grand et μ diophantienne, dépendant de p et supportée par* 3 *points.*

Remarque 4.8 *En reprenant les preuves ci-dessus et celle du théorème 3.2, on constate que tout ce dont on a besoin pour obtenir l'équivalent* (4.26) *pour les fonctions höldériennes bornées est un moment d'ordre* 4 *et une condition suffisamment forte sur l'éloignement par rapport à* 0 *de $1 - |\widehat{\mu}(t)|$ quand t est grand. La condition μ diophantienne n'est pas suffisante, mais par exemple, $\exists q > 0$, $1 - |\widehat{\mu}(t)| > 1/(\log|t|)^q$ pour t grand, suffit.*

4.4.4 Un exemple

Ici on va construire un exemple d'une mesure centrée diophantienne et à support fini, pour laquelle on peut trouver des fonctions f continues et intégrables sur $\mathbb{R}$ telles que

$$\lim_{n\to+\infty} \sqrt{2\pi\sigma_2 n} \int f d\mu^n = +\infty$$

et qui par suite ne satisfont pas le théorème local. On verra qu'on peut même choisir f de classe C^p et telle que f et ses dérivées jusqu'à l'ordre p soient bornées sur $\mathbb{R}$. La construction de μ nous sera aussi utile à la fin de cet article (section 4.5) pour donner un autre exemple relatif au théorème de cette section-là.

Considérons

$$\mu_\alpha = \frac{1}{2}\delta_0 + \frac{1-\alpha}{2}\delta_\alpha + \frac{\alpha}{2}\delta_{\alpha-1}$$

où $\alpha \in (0,1)$ et $\alpha \notin \mathbb{Q}$. Alors $\mu = \mu_\alpha$ est apériodique. De plus $S_N = X_1 + ... + X_N = p\alpha - q(1-\alpha)$ où $p, q \geq 0$ et $n = p + q \leq N$. C'est-à-dire

$$\begin{aligned} S_N &= n\alpha - q \qquad (4.27) \\ n &= \#\{i \leq N,\ X_i \neq 0\} \\ q &= \#\{i \leq N,\ sX_i = \alpha - 1\} \end{aligned}$$

On fixe dès maintenant deux suites de nombres strictement positifs décroissantes vers 0, $(\varepsilon_n)_{n\geq 0}$ et $(h_n)_{n\geq 0}$. On va choisir α sous la forme $\sum\limits_{i\geq 1} 10^{-M_i}$ où les M_i sont définis plus bas.

$$\alpha = \sum_{i\geq 1} 10^{-M_i} \qquad (4.28)$$

On définit récursivement les suites $(N_i)_i$ et $(M_i)_i$ de la façon suivante :
• $N_0 = 1$, $M_0 = 0$ et $M_1 = 1$
• Pour $i \geq 1$, on choisit N_i assez grand pour que $\varepsilon_{N_i} \leq 10^{-M_i}$ et $N_i > 10^{2M_i}$ et $N_{i+1} > 5N_i$. Alors on choisit M_{i+1} de sorte que $10^{M_i - M_{i+1}} \leq \frac{h_{N_i}}{N_i}$ et $M_{i+1} - M_i > M_i - M_{i-1}$.

La dernière condition implique que α n'est pas rationnel, et donc que μ_α est apériodique. De plus,

$$\left\|10^{M_i}\alpha\right\| \leq 2 \cdot 10^{M_i - M_{i+1}} \leq 2\frac{h_{N_i}}{N_i} \qquad (4.29)$$

où on rappelle que $\{x\} = d(x, \mathbb{Z})$ pour tout $x \in \mathbb{R}$, et $[x]$ est l'entier le plus proche de x. Donc pour tout n entre 0 et N_i on a $\left\|n10^{M_i}\alpha\right\| \leq 2h_{N_i}$. Et il vient

$$d(S_{N_i}, 10^{-M_i}\mathbb{Z}) \leq 2 \cdot 10^{-M_i} h_{N_i} \qquad (4.30)$$

Avec ces notations, pour N assez grand (dès que $h_N < 1$), on a aussi

$$|S_N| \leq h_N \Leftrightarrow |n\alpha - q| \leq h_N \Leftrightarrow \|n\alpha\| \leq h_N \text{ et } q = [n\alpha] \qquad (4.31)$$

Remarque 4.9 *Remarquons que μ_α est diophantienne si et seulement si le réel α est diophantien. De plus, on constate d'après la définition de α en* (4.28) *que α est diophantien si et seulement si la suite $(M_i)_i$ croît au plus exponentiellement, i.e. $\exists \rho > 1$ tel que $M_i \leq \rho^i$. Enfin, on vérifie que si les suites (ε_n) et (h_n) sont de la forme $\varepsilon_n = 1/n^a$ et $\varepsilon_n = 1/n^b$ avec $a > 0$ et $b > 0$, alors on peut choisir la suite M_i comme ci-dessus et vérifiant $M_i \leq \rho^i$ pour un certain ρ assez grand, donc μ_α est alors diophantienne.*

La construction précédente permet en particulier de montrer le résultat suivant :

Proposition 4.10 *Pour tout choix de suites $(\varepsilon_n)_n$ et $(h_n)_n$ de nombres strictement positifs décroissantes vers 0, il existe un $\alpha \in \mathbb{R}$ et une mesure μ_α définie comme plus haut, telle que*

$$\limsup_{n\to+\infty} \frac{\sqrt{n}}{\varepsilon_n}\mu_\alpha^n([-h_n, h_n]) > 0$$

Ainsi la suite $(\varepsilon_n(\mu))_n$ intervenant dans le théorème local de Stone (3.3) peut décroître vers 0 aussi lentement que l'on veut.

Preuve: On note E_n l'événement $\{X_i = 0 \ N-n \text{ fois}\} \cap \{X_i = \alpha - 1 \ [n\alpha] \text{ fois}\} \cap \{\|n\alpha\| \leq h_N\}$. Alors pour tout N assez grand

$$\begin{aligned}
\mathbb{P}(|S_N| &\leq h_N) \geq \sum_{n=0}^{N} \mathbb{P}(E_n) \\
&\geq \sum_{n=0}^{N} \frac{1}{2^N} C_N^n \alpha^{[n\alpha]}(1-\alpha)^{n-[n\alpha]} C_n^{[n\alpha]} \chi_{\|n\alpha\|\leq h_N} \\
&\geq \sum_{N/2-\sqrt{N}\leq n\leq N/2+\sqrt{N}} \frac{1}{2^N} C_N^n \alpha^{[n\alpha]}(1-\alpha)^{n-[n\alpha]} C_n^{[n\alpha]} \chi_{\|n\alpha\|\leq h_N} \\
&\geq \frac{c^2}{\sqrt{N}} \sum_{N/2-\sqrt{N}\leq n\leq N/2+\sqrt{N}} \frac{1}{\sqrt{n}} \chi_{\|n\alpha\|\leq h_N}
\end{aligned}$$

Dans la dernière inégalité, $c > 0$ est une constante, telle que $\frac{1}{2^N} C_N^n > c$ pour tout N et $n \in [N/2 - \sqrt{N}, N/2 + \sqrt{N}]$ et $\alpha^{[n\alpha]}(1-\alpha)^{n-[n\alpha]} C_n^{[n\alpha]} > c$ pour tout n. Elle est déterminée par le théorème limite local pour les distributions non apériodiques (voir [65] chap. 9 §49) ou bien directement à partir de la formule de Stirling. Ainsi pour tout N assez grand

$$\mathbb{P}(|S_N| \leq h_N) \geq \frac{c^2}{N} A(N)$$

où

$$A(N) = \#\left\{n \in [N/2 - \sqrt{N}, N/2 + \sqrt{N}] \ s.t. \ \|n\alpha\| \leq h_N\right\}$$

Avec le réel α défini plus haut à partir des suites $(\varepsilon_n)_n$ et $(h_n)_n$, on a vu en (4.29) que $\left\|10^{M_i}\alpha\right\| \leq 2\frac{h_{N_i}}{N_i}$. Mais, puisque $N_i 10^{M_i}/2 \geq N_i$ et $10^{M_i} < \sqrt{N_i}$, on obtient $A(N_i) \geq \frac{2\sqrt{N_i}}{10^{M_i}}$ et

$$\frac{A(N_i)}{\sqrt{N_i}} \geq \varepsilon_{N_i}$$

Comme cette inégalité est vraie pour tout $i \geq 1$, on a bien la conclusion attendue. □

Passons maintenant à la construction de la fonction f annoncée au début de ce paragraphe. Soit p un entier ≥ 1 donné. Fixons $\varepsilon > 0$ tel que $\varepsilon < 1/(10 + 6p)$ et notons $a = \frac{1}{2} - 2\varepsilon$ et $b = \frac{a}{p+1}$. Ainsi $b > \varepsilon$. Fixons la suite $(\varepsilon_n)_n$ en posant $\varepsilon_n = n^{-\varepsilon}$ et la suite $(h_n)_n$ par $h_n = 1/n$. D'après la remarque 4.9, la mesure μ_α est diophantienne. Soit $a_i = N_i^{-a}$ et $b_i = N_i^{-b}$. Définissons maintenant la fonction f_i en lui donnant la valeur a_i aux points d'abscisse $k10^{-M_i}$ pour tous les entiers k tels que $k10^{-M_i} \in [\frac{1}{2}\sqrt{N_i}, \sqrt{N_i}]$ et posons autour de ces points (i.e. pour $|x| < 10^{-M_i}$) $f_i(k10^{-M_i} + x) = a_i\phi(x/b_i)$ où ϕ est une "fonction plateau" de classe C^∞ à support dans $[-1, 1]$, à valeurs dans $[0, 1]$ et égale à 1 dans un voisinage de 0. On suppose de plus que f_i est nulle partout ailleurs. La définition est cohérente car $b_i = N_i^{-b} \leq N_i^{-\varepsilon} = \varepsilon_{N_i} \leq 10^{-M_i}$. Enfin, on pose $f = \sum_i f_i$. La fonction f est C^∞, à valeurs positives ou nulles, et f ainsi que ses dérivées jusqu'à l'ordre p tendent vers 0 car $a_i/b_i^{p+1} = 1$. Remarquons aussi que f est intégrable car

$$\begin{aligned} \sum a_i b_i \frac{\sqrt{N_i}}{10^{-M_i}} &\leq \sum a_i b_i N_i^{\frac{1}{2}+\varepsilon} \\ &\leq \sum N_i^{-(a+b-\varepsilon-\frac{1}{2})} < \infty \end{aligned}$$

et d'après le choix de ε et a et b ci-dessus, $a + b > \varepsilon + \frac{1}{2}$ et de plus $N_i \geq 5^i$.

D'autre part, la relation (4.30) montre que $f(S_{N_i}) = a_i$ dès que $S_{N_i} \in [\frac{1}{2}\sqrt{N_i}, \sqrt{N_i}]$. Donc

$$\begin{aligned} \mathbb{E}(f(S_{N_i})) &\geq a_i \mathbb{P}(S_{N_i} \in [\frac{1}{2}\sqrt{N_i}, \sqrt{N_i}]) \\ &\geq c a_i \end{aligned}$$

où $c > 0$ est une constante dépendant de μ_α qui est obtenue par le théorème limite central. Mais

$$a_i\sqrt{N_i} = N_i^{2\varepsilon} \to +\infty$$

ce qui fournit le contre-exemple annoncé en début de paragraphe.

Remarque 4.11 *Un exemple du même ordre est présenté dans [38] (Example 1) dans l'étude de la vitesse de convergence pour le théorème de renouvellement classique en dimension un.*

4.5 Théorème limite pour les fonctions asymptotiquement constantes en moyenne

Pour terminer, nous présentons dans cette section un théorème limite valable sans hypothèse sur μ (autre que celles du théorème local) pour les fonctions dont les moyennes de Cesaro sur de grands intervalles sont asymptotiquement constantes. On en déduit ensuite une application à l'équidistribution des marches aléatoires. A nouveau, la clé du théorème qui suit est le théorème local de Stone :

Théorème 5.1 *Soit μ une mesure de probabilité sur $\mathbb{R}$ centrée, apériodique et ayant un moment d'ordre 2 fini. On suppose que f est une fonction uniformément continue et bornée sur $\mathbb{R}$ telle que la limite suivante existe*

$$\lim \frac{1}{T} \int_0^T f(t)dt = l$$

quand $|T| \to +\infty$. Alors

$$\lim_{n\to+\infty} \int f d\mu^n = l$$

La preuve résulte des deux observations suivantes, que l'on présente sous la forme de deux lemmes.

Lemme 5.2 *En reprenant les notations du théorème, pour toute fonction f uniformément continue sur $\mathbb{R}$, on a*

$$\lim_{n\to+\infty} \left| \int f d\mu^n - \int f d\nu^n \right| = 0$$

où ν est la mesure gaussienne $d\nu(x) = \frac{\exp(-x^2/2\sigma_2)}{\sqrt{2\pi\sigma_2}}dx$ associée à μ.

Preuve: Soit $\varepsilon > 0$ et $\omega > 0$ un module de continuité pour f relativement à ε, i.e. $|f(x+u) - f(x)| \leq \varepsilon$ si $|u| \leq \omega$. Fixons $C > \log \frac{1}{\varepsilon}$, et $A_n = \{x, |x| \leq C\sqrt{n}\}$. En appliquant le théorème limite central à la somme de variables aléatoires S_n, on obtient pour tout n assez grand : $\mu^{*n}(A_n^c) \leq \varepsilon$ et de façon similaire $\nu^{*n}(A_n^c) \leq \varepsilon$. On peut subdiviser A_n en $O(\sqrt{n}/\omega)$ intervalles $h_i + I$ de longueur $|I| = \omega$. On obtient

$$\begin{aligned} \left| \int f d\mu^{*n} - \int f d\nu^{*n} \right| &\leq \left| \int_{A_n} f d\mu^{*n} - \int_{A_n} f d\nu^{*n} \right| + 2\varepsilon \\ &\leq \sum_i f(h_i) \left| \mu^{*n}(h_i + I) - \nu^{*n}(h_i + I) \right| + 4\varepsilon \\ &\leq \|f\|_\infty O\left(\frac{\sqrt{n}}{\omega}\right) \sup_{h\in\mathbb{R}} \left| \mu^{*n}(h + I) - \nu^{*n}(h + I) \right| + 4\varepsilon \end{aligned}$$

On peut maintenant appliquer la version uniforme du théorème limite local de Stone (théorème 3.3), qui affirme que

$$\lim_{n\to\infty} \sqrt{n} \sup_{h\in\mathbb{R}} \left| \mu^{*n}(h + I) - \nu^{*n}(h + I) \right| = 0$$

On obtient ainsi le résultat escompté. □

Le théorème ci-dessus découle aussitôt de la combinaison du lemme précédent et de l'observation suivante :

Lemme 5.3 *Soit p une densité de probabilité continue sur* $\mathbb{R}$. *On suppose que f est une fonction uniformément continue et bornée sur* $\mathbb{R}$ *telle que la limite suivante existe*

$$\lim \frac{1}{T} \int_0^T f(t)dt = l$$

quand $|T| \to +\infty$. *Alors*

$$\lim_{T \to +\infty} \int_{\mathbb{R}} f(Tx)p(x)dx = l \tag{4.32}$$

Preuve: On choisit une fonction ϕ de classe C^1 sur $\mathbb{R}$ telle que $\|p - \phi\|_1 \leq \varepsilon$. Soit $F(x) = \int_0^x f(t)dt$ une primitive de f. On écrit pour tout $A > 0$

$$\int_{\mathbb{R}} f(Tx)\phi(x)dx = \int_{-A}^{A} f(Tx)\phi(x)dx + \varepsilon_1$$

où $|\varepsilon_1| \leq \|f\|_\infty \int_{|x| \geq A} p(x)dx$. Puis on intègre par parties

$$\int_{-A}^{A} f(Tx)\phi(x)dx = [\frac{1}{T}F(Tx)\phi(x)]_{-A}^{A} - \int_{-A}^{A} \frac{1}{T}F(Tx)\phi'(x)dx$$

et on fait tendre T vers $+\infty$. Il vient

$$\begin{aligned} \lim_{T \to +\infty} \int_{-A}^{A} f(Tx)\phi(x)dx &= l[x\phi(x)]_{-A}^{A} - l\int_{-A}^{A} x\phi'(x)dx \\ &= l\int_{-A}^{A} \phi(x)dx \end{aligned}$$

Comme A est arbitrairement grand, ε arbitrairement petit, et f bornée, on conclut aussitôt. □

4.5.1 Exemple

Nous allons maintenant montrer à travers un exemple que les hypothèses du théorème ne peuvent être affaiblies. Plus précisément, nous donnons un exemple de fonction C^∞ sur $\mathbb{R}$ vérifiant toutes les hypothèses du théorème sauf la continuité uniforme, et pour laquelle le résultat prédit par le théorème n'a pas lieu.

Nous renvoyons à la section précédente pour la définition de la mesure μ_α et du nombre α obtenu à partir de deux suites $(\varepsilon_n)_n$ et $(h_n)_n$ décroissantes vers 0. Nous reprenons ici la même construction pour la fonction $f = \sum f_i$, mais cette fois-ci avec $a_i = N_i^{-a} = 1$ et $b_i = \frac{1}{\sqrt{N_i}} = N_i^{-b}$, soit $a = 0$ et $b = \frac{1}{2}$. On fixe aussi $\varepsilon = \frac{1}{4}$. Ainsi $b > \varepsilon$. Et la suite $(\varepsilon_n)_n$ est à nouveau définie en posant $\varepsilon_n = n^{-\varepsilon}$ tout comme la suite $(h_n)_n$, définie par $h_n = e^{-n}$.

Soit $a_i = N_i^{-a}$ et $b_i = N_i^{-b}$. La fonction f est bornée, de classe C^∞, et les moyennes de Cesaro tendent vers 0 asymptotiquement car, quand j tend vers $+\infty$,

$$\frac{1}{\sqrt{N_j}} \sum_{i \leq j} a_i b_i \frac{\sqrt{N_i}}{10^{-M_i}} \leq \frac{1}{\sqrt{N_j}} \sum_{i \leq j} a_i b_i N_i^{\frac{1}{2}+\varepsilon} \leq \frac{1}{\sqrt{N_j}} \sum_{i \leq j} N_i^{\frac{1}{4}} \longrightarrow 0$$

Mais f n'est pas uniformément continue sur $\mathbb{R}$. Cependant,

$$\mathbb{E}(f(S_{N_i})) \geq \mathbb{P}(S_{N_i} \in [\frac{1}{2}\sqrt{N_i}, \sqrt{N_i}]) \geq c$$

pour une constante $c > 0$ déterminée par le théorème central limit. D'où

$$\lim_{n \to +\infty} \sup \int f d\mu^n > 0$$

bien que quand $|T| \to +\infty$, $\lim \frac{1}{T} \int_0^T f(t)dt = 0$.

4.5.2 Equidistribution de marches aléatoires

Le théorème 5.1 permet d'obtenir un résultat d'équidistribution probabiliste dès que l'on dispose du résultat déterministe correspondant. Plus précisément, soit X un espace localement compact et $(\phi_t)_t$ un flot agissant continûment sur X en y préservant une mesure borélienne finie m. Le corollaire qui suit montre que si la trajectoire d'un point $x \in X$ par le flot est équidistribuée par rapport à m, alors toute marche aléatoire centrée le long de cette trajectoire s'équidistribue de la même façon. On fixe une mesure de probabilité μ centrée et apériodique sur $\mathbb{R}$ avec un moment d'ordre 2 fini et S_n la marche aléatoire de loi μ^n associée.

Corollaire 5.4 *Supposons que la trajectoire $(\phi_t)_{t\in\mathbb{R}} \cdot x$ soit équidistribuée par rapport à m, c'est-à-dire*

$$\lim_{|T| \to +\infty} \frac{1}{T} \int_0^T f(\phi_t \cdot x)dt = \int_X f(y)dm(y)$$

pour toute fonction continue à support compact f sur X. Alors on a aussi

$$\lim_{n \to +\infty} \mathbb{E}(f(\phi_{S_n} \cdot x)) = \int_X f(y)dm(y)$$

La preuve découle aussitôt du théorème 5.1 appliquée à la fonction $t \mapsto f(\phi_t \cdot x)$ qui est bornée et uniformément continue sur $\mathbb{R}$, puisque f est à support compact. Dans cet énoncé, le fait que μ est centrée est essentiel.

4.6 Cas multi-dimensionnel

Tous les résultats précédents s'étendent sans difficulté au cas multi-dimensionnel. On énonce ici principaux théorèmes valides sur $\mathbb{R}^d$, $d \geq 1$. Le cas échéant, on indique les modifications à apporter aux démonstrations.

On se place sur $\mathbb{R}^d$ où l'on note $\|x\|$ la norme euclidienne et $x \cdot y$ le produit scalaire de deux vecteurs. On considère une mesure de probabilité μ sur $\mathbb{R}^d$ qui est centrée et possède un moment d'ordre 2, c'est-à-dire

$$\int \|x\|^2 \, d\mu(x) < \infty$$

La loi μ se trouve donc dans le bassin d'attraction d'une certaine loi gaussienne, que l'on note ν. De plus on fera l'hypothèse que μ est *apériodique* sur $\mathbb{R}^d$, c'est-à-dire que le support de μ n'est pas contenu dans un sous-groupe fermé propre de $\mathbb{R}^d$. Cette condition équivaut à la suivante

$$\forall t \in \mathbb{R}^d \backslash \{0\} \;\; |\widehat{\mu}(t)| < 1$$

où $\widehat{\mu}(t) = \int e^{it \cdot x} d\mu(x)$.

De manière analogue, on note $\{x\}$ la distance du point $x \in \mathbb{R}^d$ au réseau $\mathbb{Z}^d$ et $[x]$ un point de $\mathbb{Z}^d$ qui minimise la distance à x.

Enfin $S_n = X_1 + ... + X_n$ est à nouveau la marche aléatoire somme des variables indépendantes X_i qui sont toutes de même loi μ.

La définition des mesures diophantiennes s'étend aisément au cas multi-dimensionnel. Soit l un réel ≥ 0 :

Définition 6.1 *Une mesure de probabilité sur $\mathbb{R}^d$ est dite l-**diophantienne** s'il existe un réel $C > 0$ tels que pour tout $x \in \mathbb{R}^d$ assez grand en norme,*

$$\inf_{y \in \mathbb{R}} \int \{x \cdot a + y\}^2 d\mu(a) \geq \frac{C}{\|x\|^l}$$

*et μ est dite **diophantienne** si elle est l-diophantienne pour un certain l.*

De même, μ est l-diophantienne si et seulement s'il existe $C > 0$ telle que pour tout x assez loin de zéro

$$|\widehat{\mu}(x)| \leq 1 - \frac{C}{\|x\|^l}$$

Ainsi μ est diophantienne si et seulement si $\mu * \mu^{-1}$ est diophantienne. De même, μ est diophantienne si elle (ou une de ses puissances) n'est pas étrangère à la mesure de Lebesgue. Et lorsque μ est à support fini, la condition devient une condition sur le support S de μ uniquement.

L'analogue multi-dimensionnel de la construction faite au paragraphe (4.2.2) consiste à considérer les fonctions suivantes. Soit $x = (x_1, ..., x_d) \in \mathbb{R}^d$, on pose pour $p \geq d$

$$\beta(x) = \left(\sum_{i=1}^{d} \frac{1}{4} \sin^2(\pi x_i) \right)^p \sum_{n \in \mathbb{Z}^d} \frac{a_n}{\|x - n\|^{2p}}$$

Comme auparavant, β est dans $\mathbb{L}^1(\mathbb{R}^d)$ dès que la suite $(a_n) \in \ell^1(\mathbb{Z}^d)$ et dans ce cas $\widehat{\beta}$ est à support compact d'après le théorème de Paley-Wiener. De plus, si a_n est un $O(\frac{1}{\|n\|^{2p}})$ au voisinage de l'infini, alors $\beta(x)$ est un $O(\frac{1}{\|x\|^{2p-d}})$. Ainsi $\widehat{\beta}$ est de classe C^l dès que $2(p - d) > l$.

De plus, si les a_n sont tous ≥ 0 et non tous nuls, alors β est strictement positive sur $\mathbb{R}^d$ et de plus on a $\beta(x) \geq a_{[x]}$.

Pour toute fonction f on définit $C^k(f)$ et $C^k_r(f)$ par

$$C^k(f) = \max_{|\beta| \leq k} \left\| D^\beta f \right\|_1, \; C^k_r(f) = \max_{|\alpha| \leq r+1 \; |\beta| \leq k} \left\{ \int |x_\alpha f|, \left\| D^\beta f \right\|_1 \right\}$$

où $\alpha = (i_1, ..., i_l)$ avec chaque i_j entier compris entre 1 et d et $|\alpha| := l$.

L'analogue des théorèmes (2.7) et 2.9 s'énonce de la façon suivante. La preuve est la même.

Théorème 6.2 *Soit μ une mesure de probabilité centrée sur $\mathbb{R}^d$ ayant un moment d'ordre $r + 2$ fini et une matrice de covariance K. Soit $l \geq 0$. On suppose μ l-diophantienne. Alors il existe un réel $k_0 = k_0(l) \geq 0$ tel que pour toute fonction f de classe C^k avec $k > k_0$*

(i) si $C^k(f) < +\infty$, alors

$$\mathbb{E}(f(S_n)) = \sum_{p=0}^{r} \frac{1}{n^{p/2}} \int_{\mathbb{R}^d} f(x\sqrt{n}) P_p(x) g(x) dx + C^k(f) \cdot o(\frac{1}{n^{(r+d)/2}})$$

(ii) si $C^k_r(f) < +\infty$, alors

$$(2\pi n)^{d/2} \sqrt{\det K} \mathbb{E}(f(S_n)) = \sum_{p=0}^{[r/2]} \frac{1}{n^p} \langle f, Q_p \rangle_{L^2(\mathbb{R}^d)} + C^k_r(f) \cdot o(\frac{1}{n^{r/2}})$$

où les P_p (resp. Q_p) sont des polynômes de $\mathbb{R}[x_1, ..., x_d]$ de degré total $\leq 3p$ (resp. $2p$) qui ne dépendent que des moments de μ d'ordre $\leq p + 2$, et les constantes intervenant dans les $o()$ ne dépendent que de r et de μ (g est la densité de la gaussienne ν associée à μ). On a $P_0 = Q_0 = 1$.

De même le théorème (2.11) a la généralisation ci-dessous. Le contre-exemple intervenant dans la preuve se généralise à l'identique en faisant intervenir la fonction β introduite ci-dessus.

Théorème 6.3 *Soit μ une mesure de probabilité sur $\mathbb{R}^d$ centrée, et ayant un moment d'ordre 3 fini. Alors μ est diophantienne si et seulement s'il existe un entier $k_0 \geq 0$ tel que pour toute fonction f à support compact et de classe C^k $(k \geq k_0)$ sur $\mathbb{R}^d$, on a*

$$\sup_{t\in\mathbb{R}^d}\left|\int f(t+\cdot)d\mu^n - \int f(t+\cdot)d\nu^n\right| = O(\frac{1}{n^{(d+1)/2}})$$

Le résultat sur les écarts modérés et le théorème local s'étendent eux aussi avec la même preuve au cas multi-dimensionnel.

Théorème 6.4 *Soit r un entier ≥ 1. Supposons que μ est une mesure de probabilité sur $\mathbb{R}^d$, centrée, de matrice de covariance $K = Id$ et qui possède un moment d'ordre $r+2$ fini. Alors pour tout c, $0 < c < r$, on a uniformément en x pour $1 \leq x \leq \sqrt{c\log n}$*

$$\lim_{n\to+\infty}\frac{\mathbb{P}\left(\|S_n\| > x\sqrt{n}\right)}{\mathbb{P}\left(\|N\| > x\right)} = 1$$

où N, de loi ν, est une variable gaussienne centrée de matrice de covariance K.

Le théorème de Stone (3.3) est valide sur $\mathbb{R}^d$ (voir [Sto]) et énonce que pour tout rectangle $R = \prod[-s_i, s_i]$ $(s_1, ..., s_d > 0)$

$$\sup_{x\in\mathbb{R}^d}|\mu^n(R+x) - \nu^n(R+x)| \leq \frac{\varepsilon_n(\mu)}{n^{d/2}}(1+|R|)$$

où $\varepsilon_n(\mu)$ est une suite tendant vers 0 et ne dépendant que de μ. La preuve est parfaitement analogue au cas $d = 1$. Le théorème 3.2 reste aussi vrai sur $\mathbb{R}^d$. Ainsi

Théorème 6.5 *Si μ est une mesure centrée et apériodique sur $\mathbb{R}^d$, de matrice de covariance $K = Id$ et admettant un moment d'ordre $r \geq 2$ et ν la loi gaussienne associée. Soit $R = \prod[-s_i, s_i]$ un rectangle borné centré en 0 et $s = \max s_i$. Alors pour tout $x \in \mathbb{R}^d$*

$$\lim_{n\to+\infty}\frac{\mu^n(R+x)}{\nu^n(R+x)} = 1. \tag{4.33}$$

De plus, si $r > 2$ (resp. $r = 2$) pour tout $c \in]0, r-2[$, la limite (4.33) est uniforme en x et s quand $\|x\| + s \leq \sqrt{cn\log n}$ (resp. $\|x\| + s = O(\sqrt{n})$) et tous les s_i sont minorés par un réel > 0.

Si de plus $\widehat{\mu}$ vérifie la condition de Cramér, alors il existe $\delta = \delta(\mu) > 0$ tel que la limite (4.33) est uniforme quand $|x| + s \leq \sqrt{cn\log n}$ (resp. $|x| + s = O(\sqrt{n})$ si $r = 2$) et tous les s_i sont minorés par $e^{-\delta n}$.

Le théorème (4.1) admet une généralisation évidente à $\mathbb{R}^d$, et la preuve est identique en substituant la fonction β ci-dessus à celle sur $\mathbb{R}$. Ainsi

Théorème 6.6 *Soit f une fonction Riemann intégrable sur $\mathbb{R}^d$ telle que*

$$\sum_{n\in\mathbb{Z}^d} \max_{[x]=n} |f(x)| < +\infty$$

Alors on a le théorème limite local pour f, c'est-à-dire que

$$\lim_{n\to+\infty} (2\pi n)^{d/2}\sqrt{\det K} \int f d\mu^n = \int f(x)dx$$

L'analogue du théorème 4.5 est le

Théorème 6.7 *Soit μ une mesure de probabilité centrée sur $\mathbb{R}^d$ et l un réel ≥ 0. On suppose de plus que μ possède un moment fini d'ordre $3 + d$ et on note ν la mesure gaussienne associée. Enfin, on suppose que μ est l-diophantienne. Soit k_0 un entier tel que $k_0 > 3l/2 + 1$. Alors toute fonction f définie sur $\mathbb{R}^d$ et telle que f et toutes ses dérivées jusqu'à l'ordre k_0 sont bornées vérifie la relation suivante*

$$\int f d\mu^n - \int f d\nu^n = o(\int |f| d\nu^n)$$

En particulier, si $f \geq 0$ non identiquement nulle

$$\frac{\int f d\mu^n}{\int f d\nu^n} \to 1$$

Si l'on suppose de plus que μ vérifie la condition de Cramér, alors (4.25) et (4.26) sont vérifiées pour toute fonction f höldérienne bornée sur $\mathbb{R}^d$.

Enfin, le théorème de la dernière section s'étend facilement à $\mathbb{R}^d$ en procédant de façon similaire en deux étapes. Le premier lemme utilise la version multi-dimensionnelle du théorème local uniforme de Stone, rappelée au début de cette section. On a

Théorème 6.8 *Soit μ une mesure de probabilité sur $\mathbb{R}^d$ centrée, apériodique et ayant un moment d'ordre 2 fini. On suppose que f est une fonction uniformément continue et bornée sur $\mathbb{R}^d$ telle que la limite suivante existe*

$$\lim \frac{1}{T_1 \cdots T_d} \int_0^{T_1} \cdots \int_0^{T_d} f(t)dt = \ell$$

quand $|T_i| \to +\infty$ pour tout $i = 1, ..., d$. Alors

$$\lim_{n\to+\infty} \int f d\mu^n = \ell$$

L'application à l'équidistibution des marches aléatoires s'étend verbatim aux flots multi-dimensionnels.

Remerciements 6.9 *Je remercie vivement Martine Babillot pour ses remarques enrichissantes sur une version antérieure de l'article et pour l'attention qu'elle a portée, malgré la maladie, à ce travail.*

Deuxième partie

Sous-groupes libres des groupes algébriques

Chapitre 5

Introduction

The second part of this dissertation is devoted to the study of some aspects of dense subgroups of Lie groups. We present here two articles written jointly with Tsachik Gelander (Hebrew University, Jerusalem) about dense free subgroups of Lie groups. Needless to say, dense subgroups of Lie groups and more generally dense semi-groups have not been the object of as many studies as discrete subgroups of Lie groups. Nevertheless, they appear in different parts of mathematics, most notably in the Banach-Tarski paradox and related problems [107], [84], the theory of amenable actions [178], of Riemannian foliations [39], [80], and of profinite groups [50]. Other examples are Mosher's theory of laminable actions [122], or the existence of Kazhdan groups with no uniform Kazhdan constant [61].

The main result here is what we call "a topological Tits' alternative". Given a real Lie group G it says that any dense subgroup of G contains a non-commutative free subgroup which is still dense in G, unless there is an obvious reason why this may not happen, that is when the connected component of the identity G° is solvable. Actually, our full result generalizes this to an arbitrary closed subgroup G of $GL_n(k)$ where k is any local field, i.e. the field of real numbers, complex numbers, a finite extension of the field of p-adic numbers, or a field of formal Laurent series with coefficients in a finite field. This question was first asked by Carrière and Ghys in [40], and was communicated to us by Lubotzky. As I will try to show, this result has many applications to different questions arising in the subjects mentioned above. In particular, it provides a short proof, as well as a generalization, of a theorem of Zimmer [178] which was first conjectured by Connes and Sullivan. As an easy corollary, we also answer a question of Carrière [39] about the so-called "local growth" of dense subgroups. Finally we prove a conjecture of Dixon, Pyber, Seress and Shalev about the profinite completion of linear groups.

The goal of this introduction is two-fold. First, I would like to present the main results as well as some of the main ideas of the proof by giving a short and, as far as it is possible, self-contained sketch of the proof of a somewhat weaker statement, i.e. the existence of a non-discrete free group on two generators. Second, I want to show in more details how the result can be applied to the different issues mentioned above. In the end, I ask a few questions and problems related to dense subgroups.

This work consists of two articles. The first deals exclusively with the real connected case. In the second paper, we handle the general case, in which additional difficulties arise, and provide some applications. The two papers are appended at the end of this introduction.

5.1 Statement of the main results

In his 1972 paper [167], Tits showed his famous alternative about linear groups. Let k be any field and Γ be a finitely generated subgroup of $GL_n(k)$, then either Γ contains a solvable subgroup of finite index (i.e. is virtually solvable), or it contains a non-commutative free group on two generators. The assumption that Γ is finitely generated can be removed when the characteristic of k is 0. This important theorem was first conjectured by Bass and Serre. Among other things, it implies that finitely generated amenable linear groups are virtually solvable (see [47]) and it shows the dichotomy polynomial growth vs. exponential growth for linear groups (see [176]). We will come back to these two facts in later sections.

In fact, what Tits really showed is the following theorem :

Theorem 1.1 *Suppose Γ is finitely generated and let G be the Zariski closure of Γ in $GL_n(k)$. Then either the Zariski connected component G° of G is solvable, or Γ contains a non-commutative free subgroup on finitely many generators which is Zariski dense in G°.*

This can be deduced from the methods of [167], where G was assumed Zariski-connected semi-simple ([167] Theorem 3, see also 5.2.3 below). What we propose here is a topological analogue of this statement where the Zariski topology is replaced by the standard topology arising from the field k which is assumed to be local. More precisely, let k be *a local field*, i.e. $\mathbb{R}$, $\mathbb{C}$, a finite extension of $\mathbb{Q}_p$, or $\mathbb{F}_q((t))$. Let Γ be any subgroup of $GL_n(k)$. We consider Γ as a topological group with topology induced by the natural topology on $GL_n(k)$ coming from the field k. Then our main result reads :

Theorem 1.2 *Either Γ contains an open solvable subgroup, or Γ contains a dense free subgroup.*

It should be emphasized here that the terms "open" and "dense" refer to the topology on the group Γ that was just defined above. It may happen that no finitely generated subgroup of Γ is dense in Γ, in this case the free subgroup we obtain has to have infinitely many generators. But when Γ itself is assumed to be finitely generated, then the dense free subgroup can be taken to be finitely generated too. We can also give some bound on the minimal number of generators of a dense free subgroup (see [32] Theorem 4.7.). Finally we also have the following easier result :

Theorem 1.3 *The following two conditions are equivalent :*

(i) Γ *has no open solvable subgroup.*

(ii) Γ *contains a free group* F_2 *on two generators which is non-discrete in* $GL_n(k)$

Let us specialize Γ in the last theorem in order to fix ideas and give some consequences. Let G be the closure of Γ in $GL_n(k)$.

- Suppose Γ is closed (i.e. $\Gamma = G$) and $k = \mathbb{R}$. Then Theorem 1.2 says that if the connected component G° of the Lie group G is not solvable (this is equivalent to saying that G contains no open solvable subgroup), then G contains a dense free subgroup. Moreover if G/G° is finitely generated, then we can find in G a free group on finitely many generators. The existence of dense free subgroups in real Lie groups was first studied in some detail by Kuranishi in [102], where it is shown that a connected perfect real Lie group whose Lie algebra is generated by 2 elements contains two-generated non-commutative dense free subgroups. Using Tits' alternative and an easy Baire category argument already present in [102], it is easy to show that any non solvable connected real Lie group contains a dense free subgroup on finitely many generators (i.e. of finite rank). Nevertheless, if G is not connected this argument fails and Theorem 1.2 is needed.
- We get that $SL_n(\mathbb{Q}_p)$ contains a dense free subgroup of finite rank. This fact has an interesting application found in [4], namely that the free group F_2 contains a proper subgroup H which surjects on every proper quotient of F_2.
- Suppose now that Γ is finitely generated and that G, the closure of Γ, is a connected non-solvable real Lie group. Then by Theorem 1.2, we can find finitely many elements in Γ which generate a free group which is still dense in G. In [31] we show that we can even take less that $2 \dim G$ generators. This was asked in [40] and proved in the same note for $G = SL_2(\mathbb{R})$. The proof in the $SL_2(\mathbb{R})$ case follows easily from the ideas of the proof of the Tits alternative by taking the two free generators to be two conjugate elliptic elements (and there are many such in $SL_2(\mathbb{R})$ since the set of elliptic elements is open). In the general case this method fails and other arguments are needed, because the free group obtained in the proof of Tits' alternative (generated by large powers of semi-simple elements to produce so-called proximal elements) is in general discrete in G.
- When G is semi-simple, then two elements are enough to generate a dense free subgroup in a given dense subgroup Γ. In particular, there are two matrices A and B with determinant 1 and with rational coefficients, which generate a dense free group in $SL_n(\mathbb{R})$.

5.2 About the proofs

The proof results from the combination of the study of two relatively independent problems : find generators of a dense subgroup (this highly depends on the field k, with essentially three cases : archimedean, non-archimedean in $char(k) = 0$, and non-archimedean in $char(k) = p$), and find generators of a free group (the method works for an arbitrary field). The main strategy of the proof is to make use of the fact that in many cases, when we perturb generators of a dense subgroup of G (in the topology of $GL_n(k)$), the new elements continue to generate a dense subgroup. We then find a perturbation of the generators which gives rise to generators of an abstract free group.

5.2.1 Generating dense subgroups

Let G be a closed subgroup of $GL_n(k)$ where k is a given local field. We want to find a criterium for a given k-tuple $g_1, ..., g_k$ of elements in G to generate a dense subgroup of G. This depends on whether k is archimedean or not.

Suppose k is archimedean, that is, G is a real Lie group. The problem of determining when two given elements of G generate a non-discrete, or dense, subgroup is usually a hard question. For $SL_2(\mathbb{R})$ the monograph [64] gives a rather complete method to answer this question. The famous Jorgensen inequality [93] is a basic tool. Nevertheless, in general we have the following properties :

Proposition 2.1 *Let G be a connected semi-simple real Lie group. There exists a neighborhood of the identity U in G on which $\log = \exp^{-1}$ is a well-defined diffeomorphism, such that for any $x, y \in U$, the pair $\{x, y\}$ generates a dense subgroup of G if and only if the pair $\{\log(x), \log(y)\}$ generates $Lie(G)$ as a Lie algebra. In particular "generating a dense subgroup of G" is an open condition on $G \times G$.*

The main idea behind this lemma, already present in [102] is to use the Kazhdan-Margulis-Zassenhaus lemma about commutators in Lie groups. Note that semi-simple Lie algebras are always generated by two elements ([102]). The method actually yields a little more. We call G topologically perfect if the commutator group $[G, G]$ is dense in G.

Proposition 2.2 *Let G be a connected topologically perfect real Lie group with Lie algebra $\mathfrak{g}$. Then there is a neighborhood of the identity $\Omega \subset G$, on which $\log = \exp^{-1}$ is a well defined diffeomorphism, such that $g_1, \ldots, g_m \in \Omega$ generate a dense subgroup whenever $\log(g_1), \ldots, \log(g_m)$ generate $\mathfrak{g}$ as a Lie algebra. Hence "generating a dense subgroup of G" is an open condition on G^m.*

Observe (see [31] example 2.2) that the converse may not be true : there may exists dense proper immersed Lie subgroups in a topologically perfect Lie group. And note that if G is not topologically perfect, this proposition always fails since G has a non-trivial homomorphism to the circle. As a corollary, we also obtain :

Corollary 2.3 *Let G be a connected Lie group and $\Gamma \leq G$ a finitely generated dense subgroup. Then Γ contains a dense subgroup on $2 \dim G$ generators. If G is compact then Γ contains a dense subgroup on $\dim G$ generators.*

Suppose now that k is non archimedean and $\mathcal{O}_k$ is its ring of integers. Let G be a closed subgroup of $GL_n(k)$. Then $G(\mathcal{O}_k) := GL_n(\mathcal{O}_k) \cap G$ is a profinite group. If k is a finite extension of $\mathbb{Q}_p$, then $G(\mathcal{O}_k)$ is an analytic pro-p group (see [49]). In particular, it is topologically finitely generated (i.e. contains a finitely generated dense subgroup). In such groups, it is easy to generate dense subgroups. Let H be a topologically finitely generated pro-p group. Let F be its Frattini subgroup, that is the intersection of all maximal open subgroups of H. Then we have the following fact (see [49]) :

Proposition 2.4 *The Frattini subgroup F of H is open (hence of finite index) and normal in H. Moreover, if $\{x_1, ..., x_n\}$ is a set of representatives of the cosets of F in H, then $\{x_1, ..., x_n\}$ generates a dense subgroup in H. In particular "generating a dense subgroup of H" is an open condition in H^n, $n = [H : F]$.*

It follows that when k is non-archimedean of characteristic 0 and G is a closed subgroup of $GL_n(k)$, then "generating a dense subgroup of G" is an open condition.

When $char(k) = p > 0$, then $G(\mathcal{O}_k)$ is still a pro-p group, but it may not be topologically finitely generated (for example the ring of integers $\mathcal{O}_k$ itself is not). This makes things a little more complicated, but it is possible to reduce to Proposition 2.4 by first considering a "sufficiently big" closed subgroup of $G(\mathcal{O}_k)$ which we choose topologically finitely generated (see [32] 4.3).

5.2.2 Generating free subgroups

This is by far the harder part of the proof. I will not try to explain it here, but rather refer to the next section where a sketch is given. Let me just mention that free subgroups are generated via the well known "ping-pong lemma" (see [167] prop. 1.1) applied to Γ acting on some projective space X via a suitable irreducible representation. Unlike in Tits' proof, where large powers of semi-simple elements are used to produce ping-pong players, we show that players can be obtained easily once we have constructed contracting elements and a separating set (see below). The main tool is the Cartan decomposition.

Lemma 2.5 ***(Ping-pong lemma)*** *Let G be a group acting on a set X. Let a and b be given in G. Suppose V_a, $V_{a^{-1}}$, V_b and $V_{b^{-1}}$ are subsets of X and $p \in X \backslash (V_a \cup V_{a^{-1}} \cup V_b \cup V_{b^{-1}})$. Suppose further that the condition $a(V_{a^{-1}} \cup V_b \cup V_{b^{-1}} \cup \{p\}) \subset V_a$ and the other three corresponding conditions for a^{-1}, b and b^{-1} hold. Then the pair $\{a, b\}$ generates a free subgroup of G.*

5.2.3 The non-connected case

In Theorem 1.1, Tits obtained a Zariski dense subgroup of the connected component of the identity G° of the Zariski closure G of Γ. As may be surprising at first glance, showing that we can actually find a Zariski dense subgroup in the whole of G represents a significant difficulty. In their remarkable paper [116] Margulis and Soifer, prove this result by first reducing to the case when G° is simple, and then treating each case independently (D_4 being the most delicate). They derive from it (more accurately from its proof) the following interesting fact : if Γ is a non virtually solvable finitely generated linear group, then Γ contains maximal subgroups of infinite index, even uncountably many such (if Γ is virtually solvable, then all maximal subgroups have finite index).

In the proof of 1.2, we face the same problem when the Zariski closure of Γ is not connected. Observe, that in this problem, we cannot reduce to a finite index subgroup. Most of the difficulty here is to pass from a free subgroup dense in some finite index subgroup of Γ to a free subgroup dense in the whole of Γ. This occupies a large part of [32]. Our proof does not use the classification of simple groups but uses finite dimensional representation theory of semi-simple algebraic groups (heighest weights). In particular we have to consider the problem of determining when a given irreducible algebraic representation of the Zariski connected component of the identity of some semi-simple algebraic group extends to the full group. As it turns out, such representations are easily characterized in terms of their Dynkin diagram and the action of automorphisms of the Zariski connected component on the roots.

The non-connected case is crucial in the proof of the conjecture of Dixon-Pyber-Shalev-Seress about the profinite completion of linear groups (see below Theorem 6.1).

5.3 Proof of existence of a non-discrete free subgroup

In order to illustrate a part of the proof of our main Theorem 1.2 given in the papers [31] and [32], we are going to explain here in detail a rather self-contained proof of the following weaker statement :

Theorem 3.1 *Let G be a real Lie group and Γ a finitely generated non-discrete subgroup of G such that $\overline{\Gamma}^\circ$ is not solvable. Then Γ contains a non-discrete free subgroup on two generators.*

We start by observing that in this statement we can assume G connected and Γ dense in G.

5.3.1 How to find a non-discrete subgroup

The following lemma is a straightforward consequence of the classical Zassenhaus lemma about commutators in Lie groups. Nevertheless, we present here another proof using a packing argument suggested to us by Y. Guivarc'h.

Lemma 3.2 *Let G be a connected Lie group. Then there exists a neighborhood of the identity U in G, such that if two elements x and y in U generate a free group, then the group $\langle x, y\rangle$ is not discrete.*

Proof: Let B_r be the ball of radius $r > 0$ around the identity in G according to some fixed left invariant Riemaniann metric. If $d = \dim G$, then there are two positive constants C_1 and C_2 such that for all $r < 1$

$$C_1 r^d \leq vol(B_r) \leq C_2 r^d$$

Let $\varepsilon > 0$ and suppose that x and y lie in the ball B_ε and generate a free subgroup F in G. Then, we claim that if ε is small enough, F contains a non trivial element inside the ball B_δ for $\delta = \varepsilon/3$. This follows from a simple packing argument : by contradiction suppose this were not the case, then the distance between two distinct elements of F would be $\geq 2\delta$. Let $B(n)$ be the ball of radius $n \geq 1$ in the word metric in F with respect to the generating set $\{x^{\pm 1}, y^{\pm 1}\}$. Then $B(n)$ contains $4 \cdot 3^{n-1} + 1$ elements and $B(n) \subset B_{n\varepsilon}$. Balls of radius δ around points of F are disjoint, hence as long as $n < 1/\varepsilon$, we have

$$C_1 3^n \delta^d \leq 3^n vol(B_\delta) \leq vol(B_{n\varepsilon}) \leq C_2(n\varepsilon)^d$$

We obtain a contradiction as soon as $3^n > \frac{C_2}{C_1}(3n)^d$. This can happen if we choose ε smaller than some $\varepsilon(G) > 0$ which can be explicitely computed from the constants d, C_1 and C_2. The claim is proved.

Let $a \in B_{\varepsilon/3}$ be a non trivial element of F. If ε is small enough, then, clearly, for any $h \in B_1$ we have $hB_{\varepsilon/3}h^{-1} \subset B_{\varepsilon/2}$. Now conjugating a by either x or y, we obtain a non trivial element b of F in the ball $B_{\varepsilon/2}$ which does not commute with a. Hence F contains two generators of a free subgroup which lie in the ball $B_{\varepsilon/2}$. Iterating this process, we see that F is not discrete. $\square$

Remark 3.3 *Since free groups are hopfian, and subgroups of free groups are again free groups, any two non-commuting elements in a free group are the two free generators of a free group F_2.*

With this lemma in hands, we can now concentrate on the problem of finding a pair of generators of a free group lying in a small neighborhood of the identity in G. Clearly, Theorem 3.1 follows from the above lemma and the following proposition :

Proposition 3.4 *Let G be a connected non solvable Lie group and Γ a finitely generated dense subgroup of G. Given any neighborhood U of the identity in G, one can find two elements x and y in $U \cap \Gamma$ which generate a free group F_2.*

We can obviously reduce to the case when G is semi-simple by dividing by the solvable radical of G. Taking the adjoint representation, we can further assume that G is a closed semi-simple algebraic subgroup of $SL_n(\mathbb{C})$. Finally, taking an irreducible composition factor if necessary, we can assume that G (hence Γ) acts irreducibly on $\mathbb{C}^n$ and $n \geq 2$. We can also assume that U is symmetric $U = U^{-1}$.

5.3.2 How to find a free subgroup

We say that a given subset Ω of $SL_n(\mathbb{C})$ is **strongly irreducible** if the subgroup it generates acts strongly irreducibly on $\mathbb{C}^n$ (i.e. does not fix any finite union of proper subspaces). This is equivalent to saying that the connected component $\mathbb{H}^\circ$ of the Zariski closure of the group $\langle\Omega\rangle$ generated by Ω acts irreducibly on $\mathbb{C}^n$. If $n \geq 2$, it implies easily that $\mathbb{H}^\circ$ is semi-simple (the radical R, being solvable, fixes at least one line in $\mathbb{C}^n$ by Lie's theorem and $[R, R]$ fixes a non zero vector ; since $[R, R]$ is normal, the set of fixed vectors is stabilized by the full group, hence is the whole of $\mathbb{C}^n$, hence $[R, R]$ is trivial and the group reductive ; by irreducibility, its center is trivial, so it is semi-simple).

Now observe that $\Omega := U \cap \Gamma$ is a strongly irreducible subset of $SL_n(\mathbb{C})$ which is clearly Zariski dense in G. Since Γ is finitely generated, we have reduced Proposition 3.4 to the following one :

Proposition 3.5 ***(main proposition)*** *Let R be a finitely generated subring of $\mathbb{C}$ and Ω a strongly irreducible subset of $SL_n(\mathbb{C})$ such that $\Omega = \Omega^{-1}$. Suppose elements of Ω have all their matrix entries in R. If Ω is Zariski dense in the group $\langle\Omega\rangle$ it generates, then Ω^5 contains two elements which are generators of a free group F_2.*

This proposition is the bulk of the proof. Observe that Tits' alternative for finitely generated subgroups of $GL_n(\mathbb{C})$ is a consequence of this statement. Like in the standard Schottky-Tits argument, in order to produce a free group we have to "play ping-pong" on a suitable projective space. Players of this game are usually called **proximal elements**. They are projective transformations with a particularly nice action of the projective space : they contract nearly the entire projective space (everything but the ε-neighborhood of some projective hyperplane, "the repelling neighborhood") to a small ε-ball ("the attracting neighborhood") disjoint from the repelling neighborhood. In [167], Tits produces proximal elements by taking large powers of semi-simple elements. We cannot apply this method here. In [1], Abels, Margulis and Soifer give another way to produce proximal elements : via the Cartan decomposition. Our method is inspired from theirs. I will not give the full proof of Proposition 3.5 here (see [31]), but I want to sketch it and give the main ideas.

We need a little terminology. Let V be a finite dimensional vector space over a local field. Let $\|\cdot\|$ be the standard norm on k^n, i.e. the Euclidean norm if k is Archimedean and $\|(x_1, ..., x_n)\| = \max_i |x_i|$ when k is non-Archimedean. This norm extends in the usual way to $\Lambda^2 k^n$. We define the *standard metric* on $\mathbb{P}(k^n)$ by

$$d([v], [w]) = \frac{\|v \wedge w\|}{\|v\| \, \|w\|}.$$

A pair of projective transformations $a, b \in PGL(V)$ is called a **ping-pong pair** if a, a^{-1}, b and b^{-1} are proximal elements such that the attracting neighborhoods of a and a^{-1} (resp. of b and b^{-1}) are disjoint from the repelling neighborhoods of b and b^{-1} (resp.

of a and a^{-1}). By the well-known "ping-pong lemma", if two elements form a ping-pong pair, they generate a free group F_2.

By a **contracting element** (more precisely ε**-contracting**, $\varepsilon > 0$), we mean a projective transformation that maps the complementary of a repelling neighborhood (i.e. an ε-neighborhood around some projective hyperplane) to an attracting neighborhood (i.e. an ε-ball around some point), with no assumption of disjointness of the neighborhoods. Finally, we shall say that a finite set $S \subset PGL(V)$ is an r-**separating** set (for some $r > 0$) if $S = S^{-1}$ and for every choice of 4 points $v_1, ..., v_4$ in $\mathbb{P}(V)$ and 4 projective hyperplanes $H_1, ..., H_4$ there exists $\gamma \in F$ such that

$$\min_{1 \le i,j \le 4} \{d(\gamma v_i, H_j), d(\gamma^{-1} v_i, H_j)\} > r.$$

where d is the standard distance on the projective space.

The following easy geometrical lemma (see [31] prop. 3.8 and 3.11) says that a sequence of ε_n-contracting elements ($\varepsilon_n \to 0$) and a separating set are the only two ingredients needed to produce a ping-pong pair.

Lemma 3.6 *Let $r > 0$ and S be an r-separating set. Then there is $\varepsilon_0 = \varepsilon_0(r) > 0$ such that for any ε-contracting element γ with $0 < \varepsilon < \varepsilon_0$, one can find $h_1, h_2 \in S^2$ and $g_1, g_2 \in S^2$ such that $\{h_1\gamma g_1, h_2\gamma g_2\}$ form a ping-pong pair and hence generate a free group F_2.*

Hence if we can find a separating set S inside Ω and ε-contracting elements inside Ω too, with ε arbitrary small, then Proposition 3.5 is proved.

The main geometrical idea behind this lemma is the following principle : if g is a contracting element (with v_g a point in the attracting neighborhood and H_g a projective hyperplane in the repelling neighborhood) and f is some other projective transformation, then fg (resp. gf) will also be a contracting element : its attracting neighborhood will be a small ball around fv_g (resp. v_g) and its repelling neighborhood will be around H_g (resp. $f^{-1}H_g$). It is now clear why a separating set is needed : to isolate attracting and repelling neighborhoods from one another, hence giving rise to a ping-pong pair. Let us now focus on how to produce a separating set and contracting elements inside Ω.

5.3.3 Change of field

The existence of a finite separating set in Ω is easily granted by the assumptions that Ω is a strongly irreducible subset and is Zariski dense in the group $\langle \Omega \rangle$: basically if no separating set existed, then Ω would have to lie in a proper Zariski closed subset (see [31] Lemma 4.3).

To get ε-contracting elements, we first need to observe the following fact, which allows to encode this geometrical information into an algebraic quantity (see [31] prop. 3.3) :

Lemma 3.7 *A projective transformation $g \in PSL_n(k)$ is contracting if and only if $|a_1(g)/a_2(g)|$ is large, where $(a_1(g), ..., a_n(g))$ are the diagonal coefficients of the A-component a_g of g in its Cartan decomposition $g = k_1 a_g k_2$.*

The Cartan decomposition for SL_n reads $SL_n(k) = KAK$ where K is a maximal compact subgroup and A consists of diagonal matrices with ordered coefficients ; it is valid over any local field (see [31] section 3).

In order to produce a sequence of ε_n-contracting elements inside Ω, we may have to change field. Indeed, ε-contracting elements $g \in PGL_n(k)$ are large (because $|a_1(g)|$ is large) in the topology of $PGL_n(k)$, but Ω may be bounded in $SL_n(\mathbb{C})$. Thus if we want contracting elements in Ω, we should first be concerned with making Ω unbounded. This is done via the following lemma (see [32] prop. 2.1), which provides a handy generalization of Lemma 4.1. in [167].

Lemma 3.8 *Let R be a finitely generated subring of $\mathbb{C}$, F the subfield of $\mathbb{C}$ it generates, and I an infinite subset of R. Then there is a field embedding $\sigma : F \hookrightarrow k$ where k is a local field, such that $\sigma(I)$ is unbounded in k.*

In the proof of this lemma[1], there are two extreme cases whose combination leads to a proof of the general case : that is when R is integral over $\mathbb{Z}$, and when R is isomorphic to the ring of polynomials $\mathbb{Z}[X]$. The first case is easily dealt with if we consider the diagonal embedding of R into the ring of adèles of F. Our proof of the second case reduces to the following striking lemma, first due to Polya (see [133], his proof used potential theory) :

Lemma 3.9 *Let $P \in \mathbb{C}[X]$ be a monic polynomial, then*

$$area\{x \in \mathbb{C}, \ |P(x)| \leq 1\} \leq \pi$$

where the area is the standard Lebesgue measure on $\mathbb{C}$.

In [32], Lemma 2.2, we give an elementary proof that this area is uniformly bounded. The constant obtained, πe, is not the best possible but the method works for an arbitrary local field and a uniform bound independent of the degree of the polynomial is all we need.

Let us see how to apply Lemma 3.8 to our situation. Since Ω is infinite and all its matrix coefficients lie in a finitely generated subring of $\mathbb{C}$, Lemma 3.8 shows that there is some local field k (of characteristic 0) such that if we apply this field isomorphism σ to all matrix coefficients of Ω, we get a set, that we still call Ω which is unbounded in $SL_n(k)$ (also note that Ω remains strongly irreducible on k^n and Zariski dense in

[1]Observe that, as an immediate consequence of this lemma, we obtain the following fact due to Zimmer (see [181] and [84]) : if a countable subgroup of $SU_2(\mathbb{C})$ (resp. $SO_3(\mathbb{R})$, $SO_4(\mathbb{R})$, or $SL_2(\mathcal{O}_k)$ for k local non archimedean) has Kazhdan property (T) then it is finite. Indeed, every Kazhdan group acting by isometries on a tree or on the real (or complex) hyperbolic plane fixes a point. The lemma can also be used as a substitute for Lemma VII.6.1 in [114].

$\langle\Omega\rangle$). This implies that the set $\{|a_1(g)|_k,\ g \in \Omega\}$ is unbounded. If, additonally, the set $\{|a_1(g)/a_2(g)|_k,\ g \in \Omega\}$ were unbounded, then we would be done : by Lemma 3.7 we would have obtained ε-contracting elements in Ω with ε arbitrary small. If not, then we have to consider one of the n wedge power representations : i.e. make Ω act on $\Lambda^i(k^n)$ for $1 \leq i \leq n$ instead of just k^n. Then for at least one i, the set $\{|a_1(\Lambda^i g)/a_2(\Lambda^i g)|_k,\ g \in \Omega\}$ is unbounded. However, Ω may not be strongly irreducible when acting on $\Lambda^i(k^n)$. We then have to find an irreducible subrepresentation ρ for which $\{|a_1(\rho(g))/a_2(\rho(g))|_k,\ g \in \Omega\}$ is unbounded. There always exists one such, because the Zariski closure of Ω is a semi-simple algebraic group defined over k (which we can always assume Zariski connected), hence is completely reducible (see [31] Lemma 4.2).

The proof of Proposition 3.5 and Theorem 3.1 is now complete.

5.4 Applications to amenable actions

A famous conjecture of Von Neumann from 1929 was to determine whether a finitely generated non-amenable group contains a free subgroup F_2. Tits alternative gives a positive answer for linear groups. However, as Olshanskii first proved in [129], there are examples of non-amenable finitely generated groups with no non abelian free subgroups.

Similarly, an important corollary of the Tits alternative is the characterization of finitely generated amenable linear groups as virtually solvable linear groups. This fact is actually easier than Tits' alternative. In [154], Shalom gave a very slick argument, which, by means of Lemma 3.8 can even be shortened as follows :

Theorem 4.1 *Let $\Gamma \subset GL_n(K)$ be a finitely generated linear group. Suppose Γ is amenable, then Γ is virtually solvable.*

Proof: Suppose Γ is not virtually solvable but is amenable. Let $\mathbb{G}$ be its Zariski closure with connected component of the identity $\mathbb{G}^\circ$, and let $\Gamma^\circ = \Gamma \cap \mathbb{G}^\circ$. This is a subgroup of finite index of Γ, hence still amenable and non virtually solvable, hence we can assume $\Gamma = \Gamma^\circ$. The semisimple part of $\mathbb{G}$ is non trivial and has an irreducible representation on some $\overline{K}^n$, $n \geq 2$. Hence we can suppose that $\Gamma \subset SL_n(\overline{K})$ is strongly irreducible. Let R be the ring generated by the matrix entries of the (finitely many) generators of Γ. According to Lemma 3.8 there is an embedding of R into a local field k such that Γ becomes unbounded in $PGL_n(k)$. But Γ is amenable, hence fixes a probability measure μ on $\mathbb{P}(k^n)$. By Furstenberg lemma (cf. [179] Lemma 3.2.1), μ must be supported on two projective hyperplanes. This contradicts the strong irreducibility of Γ. $\square$

5.4.1 Definitions

We are now going to present the original motivation for our main Theorem 1.2. Considering that the last statement is a direct corollary of the Tits alternative, it is not surprising that Theorem 1.2 also has a corollary of a similar nature. Namely, we obtain a characterization of amenable pairs (Γ, G) (where Γ is a countable subgroup of a closed subgroup $G \subset GL_n(k)$) as pairs (Γ, G) where $\overline{\Gamma}$ contains an open solvable subgroup.

Let us recall a few definitions. Let (S, μ) be a regular Borel space and G a separable locally compact group together with a Borel right action of G on S by Borel automorphisms of S which preserve the measure class of μ. The action is said to be **amenable** (equivalently S is an amenable G-space) if one of the following equivalent conditions holds :

(i) every Borel bundle over S with compact affine fibers has a G-equivariant section.

(ii) there exists a G-invariant mean $\mathbb{L}^\infty(G \times S) \to \mathbb{L}^\infty(S)$.

(iii) the representation of G on $\mathbb{L}^2(S)$ is weakly contained in the regular representation of G and there exists a G-invariant mean from $\mathbb{L}^\infty(S \times S) \to \mathbb{L}^\infty(S)$.

Let us be a little more precise. Condition (i) was historically the first definition of amenability of an action as it was introduced by Zimmer (see [180] and [179]). By a Borel bundle with compact affine fibers, we mean the data given by a separable Banach space E, together with a cocycle $\alpha : S \times G \to Iso(E)$ into the isometries of E and a Borel field $\{A_s\}_{s \in S}$ of compact convex subsets of the unit ball of the dual E^* (varying in a Borel manner) endowed with the weak-$*$ topology such that for every $g \in G$ we have $\alpha(s, g)A_{sg} = A_s$ for almost all $s \in S$. A G-equivariant section is a Borel map $\phi : S \to \{A_s\}_{s \in S}$ such that $\phi(s) \in A_s$ for almost all $s \in S$ and $\alpha(s, g)\phi(sg) = \phi(s)$ for every $g \in G$ and almost all $s \in S$.

If $X \to Y$ is a Borel G-map preserving the measure class between two Borel G-spaces X and Y, we define a G-invariant mean $m : \mathbb{L}^\infty(X) \to \mathbb{L}^\infty(Y)$ to be a positive $\mathbb{L}^\infty(Y)$-linear map (i.e. linear with respect to those elements of $\mathbb{L}^\infty(X)$ which are pull-backs of elements in $\mathbb{L}^\infty(Y)$ via the above G-map) which sends the constant 1 on X to the constant 1 on Y and is G-equivariant in the sense that $m(g \cdot f) = g \cdot m(f)$ for all $g \in G$ and $f \in \mathbb{L}^\infty(X)$. The equivalence between (ii) and (i) was first proved for discrete G by Zimmer (see [182]) and subsequently in full generality by Adams, Eliott and Giordano (cf. [2]). Finally the equivalence between (i) and (iii) was proved first by Nevo (cf. [128]) in a special case, then by Anantharaman (cf. [12]) in full generality.

For basic properties of amenable actions see the books [179] and [121] II.5. Among these, we shall underline the following : let G be a separable locally compact group endowed with its standard Borel structure given by a Haar measure, and let Γ be a subgroup of G endowed with the induced topology,

- A group G is amenable if and only if it acts amenably on a point. If so, it acts amenably on any G-space.
- If H is a closed subgroup of G then every amenable G-space is also, by the restricted action, an amenable H-space.
- If $T \to S$ is a G-equivariant map between Borel G-spaces and S is amenable, then T is amenable.
- Any subgroup Γ of G acts amenably on G by left translations if and only it acts amenably on G/P, where P is a given closed amenable subgroup of G.
- Any subgroup Γ of G acts amenably on G if and only if it acts amenably on its

closure $\overline{\Gamma}$.

– Suppose S is a Borel G-space and $N \triangleleft G$ is the kernel of the action. Then G acts amenably on S if and only if G/N acts amenably on S and N is amenable.

Zimmer also defined the concept of an amenable pair of G-spaces (X, Y). In our situation, saying that the pair (Γ, G) is amenable is equivalent to saying that Γ acts amenably on G.

5.4.2 The Connes-Sullivan conjecture

Let us now come back to the main point of this section. As was already mentioned above, Theorem 1.2 answers a question first asked by Carrière and Ghys in [40] in the case when $G = \overline{\Gamma}$. In their note, they made the following observation :

Proposition 4.2 ***(Carrière-Ghys*** *[40]****)*** *Let G be a locally compact group and Γ a subgroup of G. Suppose Γ contains a free subgroup on 2 generators which is not discrete in G. Then Γ does not act amenably by right translations on G.*

Clearly, the analogous statement holds if we consider the action of Γ by left translations. For completeness, let us give a different argument from that in [40] using Zimmer's definition of amenability.

Proof: By the property of amenable action stated in second above we can assume that Γ itself is the non-discrete free group. Suppose $\Gamma = \langle x, y \rangle$ acts amenably on G. By Proposition 4.3.9 in [179], it follows that there exists a Γ-equivariant Borel map $g \mapsto m_g$ from G to the space of probability measures on the boundary $\partial\Gamma$, i.e. the set of infinite reduced words (read from right to left), with its standard topology and Borel structure. Let X (resp. Y) be the set of infinite words starting with a non trivial power of x (resp. y). Let (ξ_n) (resp. θ_n) be a sequence of elements of Γ tending to the identity element in G and consisting of reduced words starting with y (resp. y^{-1}). By the converse to Lebesgue's dominated convergence theorem, up to passing to a subsequence of $(\xi_n)_n$ if necessary, we have that for almost all $g \in G$, $m_{g\xi_n}(X)$ and $m_{g\theta_n}(X)$ converge to $m_g(X)$. However, for almost every $g \in G$, $m_{g\xi_n}(X) = m_g(\xi_n X)$ and $m_{g\theta_n}(X) = m_g(\theta_n X)$. Moreover $X\xi_n$ and $X\theta_n$ are disjoint subsets of Y. Hence, for almost every $g \in G$, $2m_g(X) \leq m_g(Y)$. Since we also have

$$m_g(X) + m_g(Y) = 1 \tag{5.1}$$

we obtain $m_g(X) \leq 1/3$ for almost every $g \in G$. Reversing the roles of X and Y we get $m_g(Y) \leq 1/3$. This is a contradiction to (5.1). $\square$

Hence, to show the non amenability of the action, it is enough to exhibit a non-discrete free subgroup. Carrière and Ghys' goal was to prove the following result which was first conjectured by Connes and Sullivan, then subsequently proved by Zimmer in [178] using a completely different method which made full use of ideas from Margulis' super-rigidity theory :

Theorem 4.3 ***(Zimmer)*** *Let G be a connected Lie group and Γ a countable subgroup of G. Then Γ acts amenably on G by left translations if and only if $\overline{\Gamma}^{\circ}$, the connected component of the identity of the closure of Γ, is solvable.*

In this theorem, the "if" part follows easily from the basic properties of amenable actions. In the same note, Carrière and Ghys show this theorem in the case $G = \overline{\Gamma} = SL_2(\mathbb{R})$ by showing this existence of non-discrete free group inside a given dense countable subgroup.

Now Theorem 1.3 shows that the Carrière-Ghys approach to the Connes-Sullivan conjecture can be carried out successfully. Moreover, we can generalize this to an arbitrary locally compact group (see [32] Theorem 7.3.) :

Theorem 4.4 *Let G be a locally compact group and Γ a countable subgroup of G. Then Γ acts amenably on G (or, equivalently on G/P, with P closed amenable) by left translations if and only if Γ contains a relatively open subgroup which is amenable as an abstract group.*

For the special case of actions by isometries, this reads :

Corollary 4.5 *A countable subgroup of isometries of a symmetric space, or a locally compact Bruhat-Tits building, acts amenably on it if and only if it contains an open solvable subgroup.*

As was already noted in [178], this theorem implies a classical result of Auslander (see [137] 8.24) : if Γ is a discrete subgroup of a Lie group G and $R \leq G$ a closed normal solvable subgroup, then $\overline{\pi(\Gamma)}^{\circ}$ is solvable, where $\pi : G \to G/R$. In particular if Γ is Zariski dense in G (G algebraic) and R is the radical of G, then $\pi(\Gamma)$ is still discrete. In [136] the case of algebraic groups over a non-archimedean local field is treated. More generally, we obtain :

Corollary 4.6 *Let G be a locally compact group and R a closed normal amenable subgroup. If a countable subgroup Γ in G contains an open subgroup which is amenable (as a discrete group), then the image of Γ in G/R also contains a open subgroup which is amenable (as a discrete group).*

5.5 Applications to Riemannian foliations and local growth

The notion of growth of a finitely generated group was first studied in some detailed by Milnor and Wolf (see [119], [176]), where it is shown among other things that the growth of a finitely generated solvable group is either polynomial or exponential. It is also asked whether this dichotomy holds for an arbitrary finitely generated group. As is well-known it does not hold, and examples of groups with intermediate growth were found by Grigorchuk (see [69]). Nevertheless, as it follows from the Tits alternative (combined with Milnor-Wolf's result for solvable groups), the dichotomy holds for linear groups. In this section, we will present a similar corollary of Theorem 1.2 which echoes this dichotomy in the context of dense subgroups.

5.5.1 Local growth

Given a (non-discrete) finitely generated subgroup Γ of a given connected real Lie group, one can define a notion of "local growth" of Γ in G. Following Carrière (see [39], [41]), we define the **local growth** of a finitely generated subgroup Γ in a given connected real Lie group G in the following way. Fix a left-invariant Riemannian metric on G and consider the ball B_R of radius $R > 0$ around the identity. Suppose that S is a finite symmetric set of generators of Γ. Let $B(n)$ be the ball of radius n in Γ for the word metric determined by S. And let $B_R(n)$ be the subset of $B(n)$ consisting of those elements $\gamma \in B(n)$ which can be written as a product $\gamma = \gamma_1 \cdot \ldots \cdot \gamma_k$, $k \leq n$, of generators $\gamma_i \in S$ in such a way that whenever $1 \leq i \leq k$ the element $\gamma_1 \cdot \ldots \cdot \gamma_i$ belongs to B_R. In this situation, we will say that γ can be written as a word with letters in S which "stays all its life in B_R". Let $f_{R,S}(n) = card(B_R(n))$. As it is easy to check, if S_1 and S_2 are two symmetric sets of generators of Γ, then there exist integers $N_0, N_1 > 0$ such that $f_{R,S_1}(n) \leq f_{N_0 R,S_2}(N_1 n)$. Analogously, if R_1 and R_2 are two positive radii, and S a finite symmetric generating set, there is another bigger finite symmetric set S' and an integer N such that $f_{R_1,S}(n) \leq N \cdot f_{R_2,S'}(n)$.

Recall that two non-decreasing functions f and $g : \mathbb{N} \to \mathbb{N}$ are said to have the same growth type if there are positive integer constants C_i's such that $g(n) \leq C_0 f(C_1 n)$ and $f(n) \leq C_2 g(C_3 n)$ for all $n \in \mathbb{N}$. A function has polynomial growth of order d (resp. bounded or exponential growth type) if it has the same growth type as n^d (resp. 1 or e^n). We also say that f has simply "polynomial growth" if $f(n) \leq C_0 n^k$ for some positive constants C_0 and k.

Definition 5.1 *The local growth of Γ in G with respect to a set S of generators and a ball B_R of radius R is the growth type of $f_{R,S}(n)$.*

Note that Γ is discrete in G if and only if the local growth is bounded for any S and R. For connected Lie groups, the notion of growth is also well-defined, Lie groups of polynomial growth have been characterized by Guivarc'h and independently by Jenkins (see [73], [92]) and nilpotent Lie groups have polynomial growth of order d, where d can be explicited in terms of the ranks of the quotients in the central descending series of the Lie algebra. However, the following result shows that in a sense the local growth of Γ depends more on Γ itself than on the ambiant Lie group.

The main statement of this paragraph is the following :

Theorem 5.2 *Let Γ be a finitely generated dense subgroup of a connected real Lie group G. If G is nilpotent then Γ has polynomial local growth (for any choice of S and R). If*

G is not nilpotent, then Γ has exponential local growth (for any choice of S and any R big enough).

In particular a finitely generated dense subgroup Γ has either polynomial or exponential local growth, independently of the choice of a set of generators S and a large enough radius R. So one can talk about *the local growth* of Γ, without any reference to a set S or a ball B_R. This dichotomy mirrors the well-known dichotomy about growth of linear groups.

If G is nilpotent, then Γ is nilpotent too and is of polynomial growth, hence only the second half of the theorem requires proof. This will follow from the combination of the following two lemmas (cf. [32]) :

Recall that a subset of a given metric space X is called ε-discrete if it does not contain two distinct elements that are less than ε apart from each other. A maximal ε-discrete set is an ε -discrete set which ceases to be ε-discrete as soon as any point of X is added to it, i.e. any point of X is at distance less than ε for some element of the ε-discrete set.

Lemma 5.3 *Let G be a connected real Lie group endowed with a left-invariant metric. Let B_R be the open ball of radius R centered at the identity. Let $S = \{s_1, ..., s_k\}$ be a finite subset of B_R which forms a maximal ε-discrete subset. Moreover, assume that the elements of S are free generators of a free semi-group. If $\varepsilon > 0$ is small enough, then any finitely generated subgroup of G containing S has exponential local growth.*

Lemma 5.4 *Let G be a non-nilpotent connected real Lie group and Γ a finitely generated dense subgroup. For any finite set $S = \{s_1, ..., s_k\}$ of generators of Γ, and for every $\varepsilon > 0$, one can find perturbations $t_1, ..., t_k$ of the s_i's such that $t_i \in s_i B_\varepsilon$ and the t_i's are free generators of a free semi-group on k generators.*

This second lemma easily follows from the proof of Theorem 1.2 (the strategy in Theorem 1.2 was precisely to perturb generators into generators of a free group) in the case when G is not solvable. When G is solvable and not nilpotent, one can adapt the argument to exhibit a free semi-group.

5.5.2 Riemannian foliations

The notion of local growth was introduced by Carrière in [39] in the context of Riemannian foliations for the purpose of characterizing the growth of leaves of Riemannian foliations in terms of the underlying structural Lie algebra.

Let M be a smooth compact connected manifold. Recall that a *foliation* $\mathcal{F}$ of codimension q on M is defined in the following way. We start with an open cover $(U_i)_{i \in I}$ of M and another manifold T (the transverse manifold) of dimension q together with local submersions $f_i : U_i \to T$ with connected fibers. We also assume that transition maps $h_{i,j} : f_i(U_i \cap U_j) \to f_j(U_i \cap U_j)$ are diffeomorphisms such that $f_j = h_{j,i} \circ f_i$. A leaf of $\mathcal{F}$ is a connected subset of M which coincides with a fiber of f_i when intersected with any U_i.

A *Riemannian foliation* on M is a foliation for which one can find a Riemannian metric on T which turns the local diffeomorphisms $h_{i,j}$'s into local isometries. This means that the leaves of $\mathcal{F}$ are equidistant.

A *Lie foliation* is a foliation on M such that the transverse manifold T is a simply connected real Lie group G and the transition maps $h_{i,j}$'s are restrictions of left translations by elements of G. Since one can always endow G with a left invariant metric, every Lie foliation is Riemannian.

Naturally associated to a Lie foliation is the developing map $D : \widetilde{M} \to G$ from the universal cover of M to the group G obtained in the usual way by patching together the transition maps $h_{i,j}$ along a given path. This also yields a homomorphism $\delta : \pi_1(M) \to G$ which makes D an equivariant fibration over G for the action of $\pi_1(M)$ on $\widetilde{M}$. The image group $\Gamma := \delta(\pi_1(M))$ is called the *holonomy group* of the Lie foliation. Subgroups Γ appearing as holonomy groups of foliations on compact manifolds are called *realizable*. As shown by Haefliger (see [117]) they must satisfy the following important condition called "compact generation" : a finitely generated subgroup Γ of a connected Lie group G is said to be *compactly generated* if there is a finite generating set S of Γ and a bounded open subset U of G and a compact subset K of G such that $U\Gamma = G$ and every element in $\Gamma \cap U$ can be written as a word with letters in S which stays all its life in the compact K (in the sense defined at the beginning of this section). Clearly if Γ is a dense free subgroup in G (not compact), then Γ is not compactly generated, hence not realizable.

It is still not known whether all compactly generated subgroups are realizable. Haefliger showed that dense subgroups in compact Lie groups and nilpotent Lie groups are realizable, but already in the solvable case things get more complicated (see [62], [81], [63]). Sufficient conditions have been obtained by Meigniez and a classification of compactly generated subgroups of the affine group of the real line is available (see [117]).

By the structure theorem of Molino, every Riemannian foliation gives rise to a Lie algebra $\mathfrak{g}$, called the structural Lie algebra of the foliation, which is an invariant of the foliation (for a definition of $\mathfrak{g}$ and background about Molino's theorem, see [120], and [80]). Moreover, to every Riemannian foliation $\mathcal{F}$ on M one can associate in a natural way another Riemannian foliation $\widehat{\mathcal{F}}$ on the bundle $\widehat{M}$ over M whose fibers above a given point consist of all orthogonal frames which are orthogonal to the leaf at this point. By Molino's theorem, the restriction of $\widehat{\mathcal{F}}$ to the closure of a leaf of $\widehat{\mathcal{F}}$ is a G-Lie foliation (where G is the simply connected Lie group with Lie algebra $\mathfrak{g}$).

A Riemannian foliation is said to have *polynomial growth* (resp. exponential growth) if the volume of a ball of radius r inside a leaf grows polynomially for any leaf (resp. exponentially for a generic leaf) with r (see [40] and [39]).

In his 1988 paper [39], Carrière shows the following theorem :

Theorem 5.5 *Let $\mathcal{F}$ be a Riemannian foliation on a compact manifold M with structural Lie algebra $\mathfrak{g}$. Then $\mathcal{F}$ has polynomial growth if and only if $\mathfrak{g}$ is nilpotent.*

Using Molino's structure theorem, Carrière first reduces to the case of G-Lie foliations with dense leaves. In this case the growth of the leaves can be read off on the holonomy group Γ of the foliation. In particular $\mathcal{F}$ has polynomial (resp. exponential) growth if and only if Γ has polynomial (resp. exponential) local growth in G. Carrière proves Theorem 5.5 in two steps : using the Zimmer's Theorem 4.3 (i.e. the Connes-Sullivan conjecture) he first reduces to the case when G is solvable then completes the proof by showing that if G is not nilpotent then the local growth of Γ is super-polynomial. He then asks whether it is in fact exponential. In Theorem 5.2, we answered this question positively. This implies :

Theorem 5.6 *If $\mathfrak{g}$ is not nilpotent, then the growth of the foliation is exponential.*

5.6 Applications to profinite groups

In recent years, there has been quite a lot of interest in questions related to dense and free subgroups of profinite groups. Recall that a profinite group is the inverse limit of a system of finite groups, endowed with the inverse limit topology. Equivalently, it is a compact totally disconnected topological group (see [49] or [144] for background on profinite groups). Many of these questions are tackled from a "probabilistic" point of view. The group G, being compact, is endowed with a unique invariant probability measure, making it a probability space in a natural way. In what follows, by profinite group we understand "topologically finitely generated" profinite group (i.e. there exists a dense finitely generated subgroup).

Given a profinite group G, one can ask : does k independent randomly chosen elements in G generate a dense subgroup, or a free subgroup, with positive probability ? does G contains a dense free subgroup of finite rank ? Observe that if the answer to the first question is positive in both cases (i.e. free subgroup, and dense subgroup) then the answer to the second question is also positive. The answers to these questions are known in some cases, widely open in others, but a general theory is not yet available. For instance, Tits' alternative for profinite groups is still an open problem (i.e. does a non-virtually solvable profinite group contain a non abelian free group ?). Let us give below a sample of known examples, before stating our contribution.

Profinite groups for which k elements generate a dense subgroup with positive probability for some integer $k \geq 1$ are called positively finitely generated (*PFG,* see [108]). Pro-p groups and pro-solvable groups are PFG. However free profinite groups (i.e. the profinite completion of a free group) are not PFG (see [94]). Mann and Shalev characterized PFG groups as profinite groups with polynomial maximal subgroup growth (cf. [109]). Similarly, one can ask when do two random elements generate topologically an open subgroup with probability 1. This is the case for p-adic analytic pro-p groups (see [108]) and for compact open subgroups of $\mathbb{G}(k)$ where $\mathbb{G}$ is a simply connected semi-simple algebraic group over a local field k of positive characteristic (see [18]).

Analogously, many examples have recently been found of profinite groups G for which k randomly chosen elements in G generate a free group F_k with probability 1. This is the case for the profinite completion of $SL_d(\mathbb{Z})$ and $Aut(F_d)$, (see [50] corollary 6), for the Nottingham group (see [166]), for $Aut(T)$ where T is a k-regular rooted tree and for profinite weakly branched groups (see [3] and [4]). Furthermore, the Nottingham group and profinite weakly branched groups (for example the profinite completion of a Grigorchuk group of intermediate growth) contain dense free subgroups of finite rank ([166], [3]).

A famous conjecture of Dixon (1969), which was proved partially by Kantor-Lubotzky in [94] and then completed by Liebeck and Shalev in [106], asserts that the probability that two random elements in a given finite simple group S generate S tends to 1 as $card(S)$ tends to infinity. In a recent paper (see [50]), Dixon, Pyber, Seress and Shalev show that given a non-trivial word $w \in F_2$, the probability that two random elements x, y in a given finite simple group S do not satisfy $w(x, y) = 1$ tends to 1 as $card(S)$ tends to infinity. The combination of these two results yields an elegant new proof of another well-known conjecture of Magnus, first proved by Weigel, which asserts that F_2 is residually $\mathcal{S}$ for any infinite collection $\mathcal{S}$ of pairwise non isomorphic finite simple groups (i.e. the intersection of all normal subgroups of F_2 whose quotient is in $\mathcal{S}$ is trivial). Let us mention that the proof of these last two statements is based on a case by case study using the classification of finite simple groups.

In the same paper the authors infer the following corollary : if a profinite group G has infinitely many non abelian finite simple quotients then any k random elements in G generate a free group F_k with probability 1. Using then a deep result of Larsen and Pink, they show that a non-virtually solvable linear group has a finite index subgroup with infinitely many non abelian finite simple quotients. This result combined with Dixon's conjecture allows them to deduce that the profinite completion $\widehat{\Gamma}$ of a non-virtually solvable linear group contains a free subgroup of finite rank which is dense in a subgroup of finite index $G_0 \leq \widehat{\Gamma}$. The authors then conjecture that in fact one can take $G_0 = \widehat{\Gamma}$. Quite unexpectedly, it turns out that this conjecture in fact follows in a simple way from Theorem 1.2.

Theorem 6.1 *Let Γ be a finitely generated non-virtually solvable linear group over some field. Suppose Γ is generated by r elements. Then, for any $k \geq r$, its profinite completion $\widehat{\Gamma}$ contains a dense free subgroup F_k.*

The proof (see [32] Theorem 6.1.) goes like this : one can always imbed Γ into $GL_n(\mathcal{O}_k)$ for some local field k, giving rise to a natural map $\widehat{\Gamma} \to \overline{\Gamma} \subset GL_n(\mathcal{O}_k)$. Then by the classical Gaschütz's lemma (see [144] Prop 2.5.4), one can lift the generators of the dense free group in $\overline{\Gamma}$ obtained from Theorem 1.2 to generators of a dense free group on $\widehat{\Gamma}$.

Let us observe here that this result uses the "non-connected case" in Theorem 1.2 : as already mentioned, the passage from a virtually dense subgroup to a dense subgroup requires work.

Additionally, we can give a partial answer to a conjecture of Shalev about coset identities (see [32] Theorem 7.7.4). For an account on these questions and related problems see [152] and [153].

5.7 Concluding Remarks

Let us consider again Proposition 3.5. We would like to relax the assumption that Ω is Zariski dense in $\langle\Omega\rangle$. In fact the methods of the proof presented above allow to show the following :

Proposition 7.1 *Let R be a finitely generated subring of an algebraically closed field k, I an infinite subset of $SL_n(k)$ and Ω a strongly irreducible subset of $SL_n(k)$ such that $\Omega = \Omega^{-1}$. Suppose the elements of Ω and I have all their matrix entries in R. Then, there exists an integer $N(n)$ depending only on n, and an element $x \in I$ such that $\Omega^N x \Omega^N$ contains two elements which are generators of a free group F_2.*

Indeed, it is possible to find a separating set inside Ω^N, $N \leq N(n)$, then we can apply Lemma 4.3. This result gives a lot of freedom in the choice of generators of a free group. However, we conjecture that even more is true, namely :

Conjecture 7.2 ***(Effective Tits' alternative)**. Let Γ be a non virtually solvable finitely generated linear group. Then there exists an integer $N = N(\Gamma)$ depending on Γ only such that, for any symmetric set of generators Ω ($\Gamma = \langle\Omega\rangle$, $\Omega = \Omega^{-1}$), Ω^N contains two generators of a free group F_2.*

Proposition 7.1 shows that this conjecture holds for infinite generating sets Ω. In their remarkable paper on uniform exponential growth [56], Eskin, Mozes and Oh show that a weaker property holds in characteristic 0, namely that in Ω^N (for bounded $N \leq N(\Gamma)$) one can find two generators of a free semi-group. This, in fact, should hold as long as Γ is not virtually nilpotent. At any case, it is enough to yield uniform exponential growth for non virtually solvable linear groups. For solvable groups of exponential growth, this was showed independently by different authors (Alperin in [8] (polycyclic case), Osin in [131], and Wilson). Another quick proof of this fact (for solvable linear groups) can be derived easily by making use of Lemma 9.6 below.

This effective Tits' alternative is known to hold in some cases, most notably for Gromov hyperbolic groups (see [70], [48], [101]), finitely generated subgroups of $GL_2(K)$ where $char(K) > 0$ (see [9]) and geometrically finite groups of isometries of pinched Hadamard manifolds (see [10]). In [33] we sketch another approach for Zariski-dense subgroups of rank-1 lattices.

Chapitre 6

On dense free subgroups of Lie groups[1]

6.1 Introduction

The main purpose of this paper is to give a proof of the following statement :

Theorem 1.1 *Let G be a connected semi-simple real Lie group and Γ a dense subgroup of G. Then Γ contains two elements which generate a dense free subgroup of G.*

We shall also give the following generalizations of this result :

Theorem 1.2 *Let G be a connected real Lie group which is topologically perfect (i.e. the commutator $[G, G]$ is dense in G). If its Lie algebra $\mathfrak{g}$ is generated by l elements, then any dense subgroup contains a dense free subgroup of rank l.*

Theorem 1.3 *Let G be a connected non solvable real Lie group of dimension d. Then any finitely generated dense subgroup of G contains a dense free subgroup of rank $2d$.*

The finite generation assumption in 1.3 is crucial. Moreover, as it will be clear from the proofs, one can find a free dense subgroup of rank k, for any integer $k \geq l$ in theorem 1.2 (resp. $k \geq 2d$ in theorem 1.3). For some particular groups G one can give a smaller bound for the rank of the dense free subgroup than the bounds l and $2d$ given in 1.2 and 1.3. We will give examples to illustrate these facts.

Theorem 1.1 was originally motivated by a conjecture in the theory of amenable actions of groups on measure spaces. It answers a question first raised by Carrière and Ghys (see [40]) in a note from 1985 on amenable actions. The authors learned about this question from A. Lubotzky. Combined with [40] theorem 2, it gives a direct proof of the following theorem :

[1]Joint work with T. Gelander [31], appeared in the Journal of Algebra, March 2003

Theorem 1.4 *Let G be a real Lie group and Γ a countable subgroup. Then the action of Γ on G (or on G/P for P closed amenable) by left translations is amenable if and only if the connected component of the closure of Γ in G is solvable.*

This result was first conjectured by Connes and Sullivan (see [40]) and subsequently proved by Zimmer in [178] using the techniques of super-rigidity theory. In their note, Carrière and Ghys show that a non-discrete free subgroup of a real Lie group cannot act amenably by left translations on the Lie group (see [40], theorem 2) and then give a quick proof of theorem 1.1 in the case $G = SL_2(\mathbb{R})$, hence proving theorem 1.4 for $SL_2(\mathbb{R})$.

A celebrated theorem of Tits (see [167]), asserts that any Zariski-dense subgroup of a semisimple algebraic group G over a field of characteristic zero contains a Zariski-dense free subgroup on two generators. Theorem 1.1 can then be viewed as a kind of topological version of Tits' theorem. The difficulty of the problem we are considering comes from the fact that the free group obtained in Tits' proof is in general discrete. In this paper, we give a practical method for constructing free generators of a free group in Zariski dense subsets of G (see theorem 4.5). As an application, we then obtain theorems 1.1 to 1.3. One can also easily derive from it the original statement of Tits' alternative. The proof relies on a careful study of the so-called "proximal" elements in Γ acting on a projective space over some local field. Unlike in Tits' paper, where powers of suitable semisimple elements are used to produce proximal elements, our method is inspired from that of Abels-Margulis-Soifer (see [1]), which is based on the Cartan decomposition of projective transformations.

In section 6.2, we give a simple method for producing finitely generated dense subgroups in connected real Lie groups. Section 6.3 is devoted to the study of proximal transformations in projective spaces over local fields. We give some quantitative estimates that enable us to exhibit proximal elements with nice properties. These two sections can be read independently. In section 6.4 we explain how to find a suitable linear representation of Γ and elements in Γ which have a nice proximal action on the corresponding projective space. Finally, the last section contains the proof of the three theorems above.

Let us make one remark about terminology.

Remark 1.5 *In this paper, we use both the usual real topology (or the Hausdorff topology induced by a local field k) and the Zariski topology on the groups considered. To avoid confusion, we shall add the prefix Zariski- to any topological notion regarding the Zariski topology (Zariski-connected, Zariski-dense, etc.). For the real topology, however, we shall plainly say "dense" or "open" without further notice. Note that the Zariski topology on rational points does not depend on the field of definition, i.e. if V is an algebraic variety defined over a field K and if L is any extension of K, then the K-Zariski topology on $V(K)$ coincide with the trace of the L-Zariski topology on it.*

6.2 Generating dense subgroups in Lie groups

We call a connected Lie group G **topologically perfect** if its commutator group $[G, G]$ is dense, or equivalently, if G has no continuous surjective homomorphism to the circle.

The following result shows that, in a topologically perfect group, elements which lie near the identity generate a dense subgroup, unless they have some algebraic reason not to. Similar statements were established by Kuranishi (see [102]).

Theorem 2.1 (Generating dense subgroups in topologically perfect groups) *Let G be a connected topologically perfect real Lie group with Lie algebra $\mathfrak{g}$. Then there is an identity neighborhood $\Omega \subset G$, on which $\log = \exp^{-1}$ is a well defined diffeomorphism, such that $g_1, \dots, g_m \in \Omega$ generate a dense subgroup whenever $\log(g_1), \dots, \log(g_m)$ generate $\mathfrak{g}$.*

Proof: Recall that a Zassenhaus neighborhood of a real Lie group is an identity neighborhood Ω for which the intersection $\Omega \cap \Sigma$ with any discrete subgroup Σ is contained in a connected nilpotent Lie subgroup. A classical theorem of Kazhdan and Margulis (see [137] 8.16) says that a Zassenhaus neighborhood always exists. Checking the details of the proof of the Kazhdan-Margulis theorem, one can easily verify that the identity neighborhood Ω established there, satisfies the stronger property that its image under any surjective homomorphism f is a Zassenhaus neighborhood in the image group (this is because the properties of the neighborhoods established in 8.18 and 8.20 in [137] are inherited to their images). Moreover, as in [137], one writes $\Omega = \exp V$ for a suitable neighborhood V of 0 in $\mathfrak{g}$, and then, a subset $\{X_1, \dots, X_p\} \subset (df)(V)$ generates a nilpotent Lie subalgebra iff $\exp(X_1), \dots, \exp(X_p)$ generate a nilpotent subgroup in $f(G)$. We shall call such an Ω a *strongly Zassenhaus neighborhood.*

Let Ω be a strongly Zassenhaus neighborhood in G, and assume that it is small enough so that $\log|_\Omega$ is a well defined diffeomorphism. Pick $g_1, \dots, g_m \in \Omega$ for which $\log(g_1), \dots, \log(g_m)$ generate $\mathfrak{g}$. Denote by A the closure of $\langle g_1, \dots, g_m \rangle$, and by A^0 its identity component. We need to show that $A^0 = G$.

First observe that A^0 is normal in G, i.e. its Lie subalgebra $\mathfrak{a}$ is an ideal in $\mathfrak{g}$. Indeed, since $\mathfrak{g} = \langle \log(g_i) \rangle$, it is enough to show that $\mathrm{ad}\big(\log(g_i)\big)(\mathfrak{a}) = \mathfrak{a}$ for any $1 \le i \le m$, but

$$\mathrm{ad}(\log g_i)(\mathfrak{a}) = \big(\log \mathrm{Ad}(g_i)\big)(\mathfrak{a}),$$

and since $g_i \in A$, it normalizes A^0, i.e. $\mathrm{Ad}(g_i)(\mathfrak{a}) = \mathfrak{a}$, which implies also that $\big(\log \mathrm{Ad}(g_i)\big)(\mathfrak{a}) = \mathfrak{a}$.

Next consider the quotient map $f : G \to G/A^0$. Clearly, $f\big(\langle g_1, \dots, g_m \rangle\big) = f(A)$ is discrete, and generated by the set $\{f(g_1), \dots, f(g_m)\}$ which is contained in the Zassenhaus neighborhood $f(\Omega)$. Therefore $f\big(\langle g_1, \dots, g_m \rangle\big)$ is a nilpotent group, but this means that

$$\log f(g_i) = (df)(\log g_i), \; i = 1, \dots, m$$

generate a nilpotent Lie algebra. As $\{(df)(\log g_i)\}_{i=1}^m$ generate $\mathfrak{g}/\mathfrak{a}$, this implies that G/A^0 is nilpotent. Since G is assumed to be topologically perfect, any nilpotent quotient is trivial. Thus $G = A^0$ and $\langle g_1, \dots, g_m\rangle$ is dense.

In fact, it is enough to require that $\log(g_1), \dots, \log(g_m)$ generate a Lie subalgebra which corresponds to a dense Lie subgroup. This phenomenon is illustrated by the following :

Example 2.2 *Let $\widetilde{SL}_2(\mathbb{R})$ be the universal covering of $SL_2(\mathbb{R})$, and let $a \in \widetilde{SL}_2(\mathbb{R})$ be a central element of infinite order. Let α be an irrational rotation in the circle group $\mathbb{T}$. Consider the group*

$$G = \big(\widetilde{SL}_2(\mathbb{R}) \times \mathbb{T}\big)/\langle (a, \alpha)\rangle.$$

G is an example of a topologically perfect non-perfect group. Its Lie algebra $\mathfrak{g} \cong sl_2(\mathbb{R}) \oplus \mathbb{R}$ is not generated by 2 elements, but the 2-generated Lie subalgebra $sl_2(\mathbb{R})$ corresponds to a dense Lie subgroup in G. Theorem 1.2 guarantees only that any dense subgroup of G contains a dense F_3, however, it is possible to show (by the same argument which proves theorem 1.1) that there is always a dense F_2.

However, in general, 2 generators are not enough. This phenomenon is illustrated by the following :

Example 2.3 *Consider the complex Lie group $G = SL_2(\mathbb{C}) \cdot (\mathbb{C}^2)^5$, with the ordinary action on $SL_2(\mathbb{C})$ on each of the 5 copies of $\mathbb{C}^2$. It is easy to verify that*

1. *G is perfect, i.e. $[G, G] = G$.*
2. *The exponential map is onto.*
3. *Any 2 elements in G's Lie algebra $\mathfrak{g} = sl_2(\mathbb{C}) \cdot (\mathbb{C}^2)^5$ generate a Lie subalgebra for which the closure of the corresponding Lie subgroup is proper.*

These properties imply that there is no 2-generated dense subgroup in G.

It is well known that a real semisimple Lie algebra is generated by 2 elements (cf. [102] or [29] VIII, 3, ex. 10). In fact $V = \{(X, Y) \in \mathfrak{g} \times \mathfrak{g} : \langle X, Y\rangle \neq \mathfrak{g}\}$ is a proper algebraic subset. We therefore obtain :

Theorem 2.4 (Generating dense subgroups in semisimple groups) *Let G be a connected semisimple Lie group. Then there exists an identity neighborhood $\Omega \subset G$, and a proper exponential algebraic subvariety $R \subset \Omega \times \Omega$ such that $\langle x, y\rangle$ is dense in G whenever $(x, y) \in \Omega \times \Omega \setminus R$.*

Theorem 2.1 yields a lot of freedom in the procedure of generating dense subgroups in topologically perfect groups. In fact, if G is a topologically perfect Lie group and $g_1, \dots, g_m \in G$ are elements close enough to 1 for which $\log(g_1), \dots, \log(g_m)$ generate the Lie algebra $\mathfrak{g}$, and if $U \subset G$ is a sufficiently small identity neighborhood, then for any selections $h_i \in g_i \cdot U$, $\log(h_1), \dots, \log(h_m)$ generate $\mathfrak{g}$, and $h_1, \dots, h_m$ generate a dense subgroup in G. In particular, we obtain :

Corollary 2.5 *If G is a topologically perfect Lie group for which the Lie algebra is generated by m elements, then any dense subgroup in G contains a m-generated dense subgroup.*

Proof: Let $g_1, \ldots, g_m$ and U be as in the last paragraph above. Clearly, if $\Gamma \leq G$ is a dense subgroup then $\Gamma \cap g_i \cdot U \neq \emptyset$. Pick $h_i \in \Gamma \cap g_i \cdot U,\ i = 1, \ldots, m$ then $\langle h_1, \ldots, h_m \rangle$ is dense.

The requirement that G is topologically perfect is crucial for corollary 2.5 as explained by the following remark :

Remark 2.6 *If a connected Lie group G is not topologically perfect then it has a surjective homomorphism to the circle. Let $\Gamma \leq G$ be the pre-image of the group of rational rotations. Clearly Γ is dense in G but, as any finitely generated group of rational rotations is finite, Γ has no finitely generated dense subgroup.*

This remark explains why, when G is not topologically perfect, we shall consider only *finitely generated* dense subgroups. Then we can control the number of generators, by the following reasoning.

Proposition 2.7 *Let G be a connected Lie group and $\Gamma \leq G$ a finitely generated dense subgroup. Then Γ contains a dense subgroup on $2 \dim G$ generators. If G is compact then Γ contains a dense subgroup on $\dim G$ generators.*

Proof: Assume first that G is abelian. Then $G = \mathbb{R}^{d_1} \times \mathbb{T}^{d_2}$ and we can find d_1 elements in Γ which generate a discrete cocompact subgroup. Dividing by this subgroup, we may assume that G is a torus. Then we argue by induction on $\dim G$. As Γ is finitely generated and dense it contains an element γ of infinite order. By replacing γ by some power γ^j if necessary, we obtain an element which generates a subgroup with connected closure C of positive dimension. By induction, the proposition holds for G/C. Lift arbitrarily to G a set of $\dim G/C$ generators for a dense subgroup of the image of Γ in G/C. Together with γ they generate a dense subgroup in G.

For the general case, we argue as follows. Consider the sequence $G_{i+1} = \overline{[G_i, G_i]}$, where $[G_i, G_i]$ is the commutator group of G_i, starting at $G_0 = G$. As $\dim G$ is finite this sequence stabilizes at some finite step k. Then G_k is a connected closed normal topologically perfect subgroup. Moreover the commutator $[\Gamma, \Gamma]$ is a finitely generated dense subgroup in G_1, and similarly, the i'th commutator of Γ is a finitely generated dense subgroup in G_i. Thus, by the above paragraph, we can find, for each i, $0 \leq i < k$, a set Σ_i of $2 \dim G_i/G_{i+1}$ (or $\dim G_i/G_{i+1}$ in the compact case) elements in $G_i \cap \Gamma$ which projects to generators of a dense subgroup in the abelian group G_i/G_{i+1}. Additionally, by corollary 2.5 there is a set Σ_k of $\dim G_k$ elements in $\Gamma \cap G_k$ which generate a dense subgroup of G_k (in the non-abelian compact case $G_k = G_1$ is semisimple and we can take Σ_k of size 2).

Clearly $\bigcup_{i=0}^{k} \Sigma_i$ generates a dense subgroup and has the required cardinality.

6.3 Projective transformations, proximality, ping-pong

Let k be a local field equipped with an absolute value $|\cdot|$. We wish to investigate the action of projective transformations in $\mathrm{PSL}_n(k)$ on the projective space $\mathbb{P}^{n-1}(k)$. We shall start by recalling the Cartan decomposition, i.e. the KAK decomposition in $\mathrm{SL}_n(k)$, and introducing a nice metric on the projective space $\mathbb{P}^{n-1}(k)$.

1. Consider first the case when k is archimedean, i.e. $k = \mathbb{R}$ or $\mathbb{C}$. We shall denote by $\|\cdot\|$ the canonical euclidean (resp. hermitian) norm on k^n, i.e.$\|x\|^2 = \sum |x_i|^2$ if $x = \sum x_i e_i$, where $(e_1, ..., e_n)$ is the canonical basis of k^n. The Cartan decomposition in $\mathrm{SL}_n(k)$ reads
$$\mathrm{SL}_n(k) = KAK$$
where,
$$\begin{aligned} K &= \mathrm{SO}_n(\mathbb{R}) \text{ or } \mathrm{SU}_n(\mathbb{C}) \text{ according to whether } k = \mathbb{R} \text{ or } \mathbb{C}, \\ \text{and } A &= \{\mathrm{diag}(a_1, \ldots, a_n) : a_1 \geq \ldots \geq a_n > 0, \prod a_i = 1\}. \end{aligned}$$
Any element $g \in \mathrm{SL}_n(k)$ can be decomposed as a product $g = k_g a_g k'_g$, where $k_g, k'_g \in K$ and $a_g \in A$. We remark that a_g is uniquely determined by g, but k_g, k'_g are not. Even so, we shall use the subscript g to indicate the relation to g.

2. Next, consider the case where k is non-archimedean with valuation ring $\mathcal{O}_k$ and uniformizer π. We endow k^n with the canonical norm defined by $\|x\| = \max |x_i|$, where $x = \sum x_i e_i$ is the expansion of x with respect to the canonical basis $(e_1, ..., e_n)$ of k^n. Then we get
$$\mathrm{SL}_n(k) = KAK$$
where
$$\begin{aligned} K &= \mathrm{SL}_n(\mathcal{O}_k), \\ \text{and } A &= \{\mathrm{diag}(\pi^{j_1}, \ldots, \pi^{j_n}) : j_i \in \mathbb{Z}, j_i \leq j_{i+1}, \sum j_i = 0\}. \end{aligned}$$

Any $g \in \mathrm{SL}_n(k)$ can be decomposed as a product $g = k_g a_g k'_g$, where $k_g, k'_g \in K$ and $a_g \in A$ (see [132] page 150, or [35] 4.4.3.). Again a_g is uniquely determined, but k_g, k'_g are not.

In both cases, the canonical norm on k^n gives rise to the associated canonical norm on $\bigwedge^2 k^n$. We shall then define the *standard metric* on $\mathbb{P}^{n-1}(k)$ by the formula
$$d\big([v], [w]\big) = \frac{\|v \wedge w\|}{\|v\| \cdot \|w\|}$$
This is well defined and satisfies the following properties :

– d is a distance on $\mathbb{P}^{n-1}(k)$ which induces the canonical topology inherited from the local field k.
– d is a ultra-metric distance if k is non-archimedean, i.e.

$$d\big([v],[w]\big) \leq \max\{d\big([v],[u]\big), d\big([u],[w]\big)\}$$

for any non-zero vectors u, v and w in k^n.
– If f is a linear form $k^n \to k$, then for any non-zero vector $v \in k^n$,

$$d\big([v],[\ker f]\big) = \frac{|f(v)|}{\|f\| \cdot \|v\|} \tag{6.1}$$

– the compact group K acts by isometries on $\big(\mathbb{P}^{n-1}(k), d\big)$.

In the sequel, we shall denote by $a_1(g), ..., a_n(g)$ the coefficients of the diagonal matrix a_g corresponding to g in the Cartan decomposition. For further use, we note that in the above notations, for any matrix $g \in \mathrm{SL}_n(k)$,

$$\begin{aligned} \|g\| &= |a_1(g)| \\ \left\|\bigwedge^2 g\right\| &= |a_1(g)a_2(g)|. \end{aligned}$$

Let us observe a few properties of projective transformations. For $g \in \mathrm{SL}_n(k)$ we denote by $[g]$ the corresponding projective transformation $[g] \in \mathrm{PSL}_n(k)$.

Lemma 3.1 *Every projective transformation is bi-Lipschitz on the entire projective space for some constant depending on the transformation.*

Proof: As the compact group K acts by isometries on $\mathbb{P}^{n-1}(k)$, the KAK decomposition allows us to assume that $g = a_g = diag(a_1, ..., a_n) \in A$. Thus, one only needs to check the easy fact that a_g is $\left|\frac{a_1}{a_n}\right|^2$-Lipschitz. □

Definition 3.2 *Let $\epsilon \in (0,1)$. A projective transformation $[g] \in PSL_n(k)$ is called* ***ϵ-contracting*** *if there exist a point $v_g \in \mathbb{P}^{n-1}(k)$, called an* ***attracting point*** *of $[g]$, and a projective hyperplane H_g, called a* ***repulsive hyperplane*** *of $[g]$, such that $[g]$ maps the complement of the ϵ-neighborhood of $H_g \subset \mathbb{P}^{n-1}(k)$ into the ϵ-ball around v_g. We say that $[g]$ is* ***ϵ-very contracting*** *if both $[g]$ and $[g^{-1}]$ are ϵ-contracting.*

Note that in general v_g and H_g are not necessarily unique. Nevertheless, all the statements we shall make about them will be valid for any choice.

The following proposition shows that "ϵ-contraction" is equivalent to "a big ratio between the first and second diagonal terms in the KAK decomposition."

Proposition 3.3 *Let $\epsilon < \frac{1}{4}$. If $|\frac{a_2(g)}{a_1(g)}| \leq \epsilon^2$, then $[g]$ is ϵ-contracting. More precisely, writing $g = k_g a_g k'_g$, one can take H_g to be the projective hyperplane spanned by $\{k'^{-1}_g(e_i)\}_{i=2}^n$, and $v_g = [k_g(e_1)]$.*

Conversely, suppose $[g]$ is ϵ-contracting and k is non-archimedean with uniformizer π (resp. archimedean), then $|\frac{a_2(g)}{a_1(g)}| \leq \frac{\epsilon^2}{|\pi|}$ (resp. $|\frac{a_2(g)}{a_1(g)}| \leq 4\epsilon^2$).

Proof: Suppose that $|\frac{a_2(g)}{a_1(g)}| \leq \epsilon^2$. Using the Cartan decomposition and the fact that K acts by isometries, we can again assume that

$$g = a_g = diag(a_1, ..., a_n)$$

Let $[v] \in \mathbb{P}^{n-1}(k)$ be outside the ϵ-neighborhood of H_g, which, in our case, simply means that (cf. (6.1))

$$d([v], H_g) = \frac{|v_1|}{\|v\|} \geq \epsilon$$

Then

$$d([gv], [e_1]) = \frac{\|gv \wedge e_1\|}{\|gv\|} \leq \frac{|a_2| \cdot \|v\|}{|a_1| \cdot |v_1|} \leq \epsilon$$

since $\|gv \wedge e_1\| \leq |a_2| \, \|v\|$ and $\|gv\| \geq |a_1 v_1|$.

The converse is more delicate. We shall describe the non-archimedean case, and remark that the archimedean case can be dealt with in an analogous way. Again we can take $g = a_g$. Let $f = (f_1, ..., f_n)$ be a linear form of norm 1 such that $\ker(f) = H_g$ and let $w \in k^n$ be a normalized representative of the attracting point v_g. The fact that $[g]$ is ϵ-contracting means that for every non-zero vector $v \in k^n$

$$|f(v)| \geq \epsilon \, \|v\| \Rightarrow \|gv \wedge w\| \leq \epsilon \, \|gv\| . \tag{6.2}$$

Suppose that for some index i_0, $|f_{i_0}| \geq \epsilon$. Take $v = e_{i_0}$ and get that

$$|a_{i_0}| \, \|e_{i_0} \wedge w\| \leq \epsilon \, |a_{i_0}|$$

So $\|e_{i_0} \wedge w\| = \max_{i \neq i_0} |w_i| \leq \epsilon$ and hence $|w_{i_0}| = 1$, and i_0 is unique. Since $\|f\| = \max |f_i| = 1$, we must then have $|f_{i_0}| = 1$ and $\max_{i \neq i_0} |f_i| < \epsilon$.

We now claim that $i_0 = 1$. Suppose the contrary and take $v = e_1 + e_{i_0}$. Then $\|v\| = 1$ and $f(v) = f_1 + f_{i_0}$. Since $|f_{i_0}| = 1 \geq \epsilon > |f_1|$ we indeed have

$$|f(v)| \geq \epsilon \, \|v\|$$

Hence, by (6.2), we obtain

$$\|a_1 e_1 \wedge w + a_{i_0} e_{i_0} \wedge w\| \leq \epsilon \, \|a_1 e_1 + a_{i_0} e_{i_0}\|$$

which also reads

$$|a_1| \leq \epsilon \, |a_1|$$

and gives the desired contradiction.

Finally take $v = xe_1 + e_2$, where $x \in k$ is chosen such that $|x|$ is least possible and $\geq \epsilon$. Then $\|v\| = 1$ and $f(v) = xf_1 + f_2$. Since $|xf_1| \geq \epsilon > |f_2|$ we have

$$|f(v)| \geq \epsilon \|v\|$$

and, again, by (6.2), we obtain

$$\|xa_1e_1 \wedge w + a_2e_2 \wedge w\| \leq \epsilon \|xa_1e_1 + a_2e_2\|$$

Suppose $|a_2| > \epsilon |x| |a_1|$ then the last inequality translates to

$$|a_2| \leq \epsilon \max\{|xa_1|, |a_2|\}$$

which leads to either $|a_2| \leq \epsilon |a_2|$, which is absurd, or $|a_2| \leq \epsilon |x| \cdot |a_1|$, which contradicts the assumption.

Therefore we have obtained that $|a_2| \leq \epsilon |x| |a_1|$, which, considering the choice of x, implies the desired conclusion :

$$\left|\frac{a_2}{a_1}\right| \leq \frac{\epsilon^2}{|\pi|}$$

□

Note that the factor $|\pi|$ is not necessary when ϵ belongs to the value group of k.

Lemma 3.4 *Let $r, \epsilon \in (0,1]$. If $|\frac{a_2(g)}{a_1(g)}| \leq \epsilon$, then $[g]$ is $\frac{\epsilon}{r^2}$-Lipschitz outside the r-neighborhood of the repulsive hyperplane $H_g = [span\{k'^{-1}_g(e_i)\}_{i=2}^n]$.*

Proof: Again, we can assume $g = a_g$, $v_g = [e_1]$, $H_g = [span(e_i)_{i\geq 2}]$. Let v, w be arbitrary non-zero vectors in k^n. Then $\|gv \wedge gw\| \leq |a_1a_2| \|v \wedge w\|$ and $\|gv\| \geq |a_1v_1|$ and $\|gw\| \geq |a_1w_1|$. Therefore

$$d\big([gv],[gw]\big) \leq \frac{|a_2| \cdot \|v\| \cdot \|w\|}{|a_1| \cdot |v_1| \cdot |w_1|} d\big([v],[w]\big). \tag{6.3}$$

We conclude by observing that the r-neighborhood of $H_g = [span(e_i)_{i\geq 2}]$ corresponds to non-zero vectors v for which $|v_1| \leq r \|v\|$. □

Note in particular that if $\delta > 0$, any projective transformation $[g] \in \mathrm{PSL}_n(k)$ is $(1+\delta)$-Lipschitz in some open set of $\mathbb{P}^{n-1}(k)$.

The following lemma gives a converse statement to the last result as well as a handy criterion for ϵ-contraction.

Lemma 3.5 *Let $[g]$ be a projective transformation corresponding to $g \in SL_n(k)$, and assume that the restriction of $[g]$ to some open set $O \subset \mathbb{P}^{n-1}(k)$ is ϵ-Lipschitz for some $\epsilon < 1$, then $|\frac{a_2(g)}{a_1(g)}| \leq \epsilon$ when k is non-archimedean, and $|\frac{a_2(g)}{a_1(g)}| \leq \frac{\epsilon}{\sqrt{1-\epsilon^2}}$ when k is archimedean.*

Proof: Once again, we may assume $g = a_g$. Let v be a non-zero vector in k^n such that $[v] \in O$. For some $\delta \in k\backslash\{0\}$ small enough, both $[w_1] = [v+\delta e_1]$ and $[w_2] = [v+\delta e_2]$ belong to O. Hence, for w_1

$$|\delta a_1| \frac{\|av \wedge e_1\|}{\|av\| \cdot \|aw_1\|} = d\big([av],[aw_1]\big) \leq \epsilon \cdot d\big([v],[w_1]\big) = \epsilon\,|\delta| \frac{\|v \wedge e_1\|}{\|v\| \cdot \|w_1\|} \qquad (6.4)$$

thus

$$\frac{\|av \wedge e_1\|}{\|av\|} \leq \epsilon$$

In the non-archimedean case, this implies $|a_1 v_1| = \|av\|$, and in the archimedean case, $\|av \wedge e_2\| \geq |a_1 v_1| \geq \sqrt{1-\epsilon^2}\,\|av\|$. Now, expressing the Lipschitz condition for v and w_2 as in (6.4) yields

$$|a_2| \frac{\|av \wedge e_2\|}{\|av\|} \leq \epsilon\,|a_1|$$

which, in the non-archimedean case gives $|a_2/a_1| \leq \epsilon$ and in the archimedean case, $|a_2/a_1| \leq \epsilon/\sqrt{1-\epsilon^2}$. □

Definition 3.6 *A projective transformation $[g] \in PSL_n(k)$ is called (r,ϵ)-**proximal** ($r > 2\epsilon > 0$) if it is ϵ-contracting with respect to some attracting point $v_g \in \mathbb{P}^{n-1}(k)$ and some repulsive hyperplane H_g, such that $d(v_g, H_g) \geq r$. The transformation $[g]$ is called (r,ϵ)-**very proximal** if both $[g]$ and $[g]^{-1}$ are (r,ϵ)-proximal.*

Similar notions of (r,ϵ)-proximality were defined in [1] and [21].

Definition 3.7 *A finite subset $F \subset PSL_n(k)$ is called (m,r)-**separating** ($r > 0$, $m \in \mathbb{N}$) if for every choice of $2m$ points $v_1, ..., v_{2m}$ in $\mathbb{P}^{n-1}(k)$ and $2m$ projective hyperplanes $H_1, ..., H_{2m}$ there exists $\gamma \in F$ such that*

$$\min_{1 \leq i,j \leq 2m} \{d(\gamma v_i, H_j), d(\gamma^{-1} v_i, H_j)\} > r.$$

We use the properties described above in order to construct ϵ-very contracting and (r,ϵ)-very proximal elements from two basic ingredients : an ϵ-contracting element and an r-separating set.

Proposition 3.8 *Let F be a $(1,r)$-separating set ($r < 1$) in $PSL_n(k)$ with uniform bi-Lipschitz constant $C = \max\{biLip\big([f]\big) : [f] \in F\}$ (see lemma 3.1).*

(i) If $\epsilon < \frac{r}{2C}$ and $[g] \in PSL_n(k)$ is an ϵ-contracting transformation, then one can find $[f] \in F$ such that $[fg]$ is $(r, C\epsilon)$-proximal.

(ii) Let $d = 4$ when k is archimedean and $d = \frac{1}{|\pi|}$ when k is non-archimedean. If $\epsilon < \frac{r}{\sqrt{2Cd}}$ and $[g] \in PSL_n(k)$ is ϵ-contracting, then there is an element $[f] \in F$, such that

$$[gfg^{-1}]$$

is $\frac{\sqrt{2Cd}}{r}\epsilon$*-very contracting.*

(iii) If $\epsilon < \frac{r}{2C^2}$ *and* $[g] \in PSL_n(k)$ *is an* ϵ*-very contracting transformation, then there is an element* $[f] \in F$*, such that* $[gf]$ *is* $(\frac{r}{C}, C\epsilon)$*-very proximal.*

Proof: Let v_g (resp. H_g) be an attracting point (resp. repulsive hyperplane) corresponding to the transformation $[g]$. Let $[f] \in F$ be such that $d([f]v_g, H_g) > r$. Then $[fg]$ is clearly $C\epsilon$-contracting with attracting point $[f]v_g$ and repulsive hyperplane H_g. The choice of f guarantees that $[fg]$ is $(r, C\epsilon)$-proximal. This proves (i).

By lemma 3.5 and the remark following it, there is an open set O on which $[g^{-1}]$ is, say, $\sqrt{2}$-Lipschitz. Let $[u]$ be some point inside O. We can find $[f] \in F$ such that $d([fg^{-1}u], H_g) > r$ and $d([f^{-1}g^{-1}u], H_g) > r$. Since $[g]$ is ϵ-contracting, by proposition 3.4 we have $\left|\frac{a_2(g)}{a_1(g)}\right| \leq d\epsilon^2$. Hence by lemma 3.5, $[g]$ is $\frac{d\epsilon^2}{r^2}$-Lipschitz outside the r-neighborhood of H_g. Therefore both $[gfg^{-1}]$ and $[gf^{-1}g^{-1}]$ are $\frac{dC\sqrt{2}\epsilon^2}{r^2}$–Lipschitz in some open neighborhood of $[u]$, which implies by lemma 3.5, that they both satisfy $\frac{|a_2|}{|a_1|} \leq 2\frac{dC\epsilon^2}{r^2}$. By proposition 3.4 it now follows that they are both $\frac{\sqrt{2Cd}}{r}\epsilon$-contracting transformations. This proves (ii).

Find $[f] \in F$ which takes v_g at least r away from H_g and for which $[f^{-1}]$ takes $v_{g^{-1}}$ at least r away from $H_{g^{-1}}$. Then $[gf]$ is $(\frac{r}{C}, C\epsilon)$-proximal and $[f^{-1}g^{-1}]$ is $(r, C\epsilon)$-proximal. Hence $[gf]$ is $(\frac{r}{C}, C\epsilon)$-very proximal. This proves (iii).

□

We end this section by showing how to obtain generators of a free group via the so-called ping-pong lemma, once we are given a very contracting element and a separating set. The following definition and lemma are classical (cf. [35]).

Definition 3.9 (Ping-pong pair) *Let k be a local field and V a finite dimensional k-vector space. A pair of projective transformations $a, b \in PSL(V)$ is called a* ***ping-pong pair*** *if both a and b are (r, ϵ)-very proximal, with respect to some $r > 2\epsilon > 0$, and if the attracting points of a and a^{-1} (resp. of b and b^{-1}) are at least r-apart from the repulsive hyperplanes of b and b^{-1} (resp. of a and a^{-1}).*

More generally, a m-tuple of projective transformations $a_1, \ldots, a_m$ is called a ***ping-pong m-tuple*** *if all a_i's are (r, ϵ)-very proximal (for some $r > 2\epsilon > 0$) and the attracting points of a_i and a_i^{-1} are at least r-apart from the repulsive hyperplanes of a_j and a_j^{-1}, for any $i \neq j$.*

We have the following variant of the ping-pong lemma (see [35] 1.1).

Lemma 3.10 *If $a_1, \ldots, a_m \in PSL(V)$ form a ping-pong m-tuple, then they are free generators of a free group F_m.*

The following proposition explains how the above ingredients can be combined to yield a free group.

Proposition 3.11 *Given $a_1, \ldots, a_m \in PSL(V)$, an (m,r)-separating set F, and an ϵ-very contracting element γ (with $\epsilon < \frac{r}{2c^4}$, where c is the maximal bi-Lipschitz constant of $\{a_1, ..., a_m\} \cup F$), there are $h_1, \ldots, h_m \in F$ and $g_2, \ldots, g_m \in F$ such that*

$$\gamma a_1 h_1, \; g_2 \gamma a_2 h_2, \ldots, \; g_m \gamma a_m h_m$$

form a ping-pong m-tuple of $(\frac{r}{c}, c^3\epsilon)$-very proximal transformations and hence generate a free group F_m.

Proof: In this proof, whenever we speak about an ϵ-contracting or (r,ϵ)-proximal transformation, we shall choose and fix a point and a projective hyperplane with respect to which it is ϵ-contracting or (r,ϵ)-proximal and shall refer to them as the attracting point and the repulsive hyperplane of that transformation.

As γ is ϵ-very contracting, γa_1 is $c\epsilon$-very contracting. Now by proposition 3.3 (*iii*), we can find $h_1 \in F$ such that $\gamma a_1 h_1$ is $(\frac{r}{c}, c^2\epsilon)$-very proximal.

We proceed by induction and suppose that $x_1 = \gamma_1 a_1 h_1$ and $x_j = g_j \gamma a_j h_j$ have been constructed for all $j < i$ $(i = 2, ..., m)$. Let's construct h_i and g_i. First note that γa_i is $c\epsilon$-very contracting. We take $h_i \in F$ such that h_i^{-1} maps the attracting point of $(\gamma a_i)^{-1}$ at a distance at least r from the repulsive hyperplanes of all transformations x_j and x_j^{-1} for $j < i$, and h_i maps the attracting point of any x_j and x_j^{-1} $(j < i)$ at least r away from the repulsive hyperplane of γa_i. This is possible since F is an (m,r)-separating set. The second requirement for h_i implies that the repulsive hyperplane $H_{(\gamma a_i h_i)} = h_i^{-1}(H_{(\gamma a_i)})$ of $\gamma a_i h_i$ is at least r/c away from the attracting points of the x_j's and x_j^{-1}'s.

Then pick $g_i \in F$ which takes the attracting point of $\gamma a_i h_i$ at a distance at least r from the repulsive hyperplanes of all x_j's and x_j^{-1}'s for all $j < i$ and also from the repulsive hyperplane of $\gamma a_i h_i$, and whose inverse g_i^{-1} takes the attracting point of any x_j and x_j^{-1} $(j < i)$ at least r away from the repulsive hyperplane of $(\gamma a_i h_i)^{-1}$. This means that the repulsive hyperplane $g_i(H_{(\gamma a_i h_i)^{-1}}) = H_{(g_i \gamma a_i h_i)^{-1}}$ of $g_i \gamma a_i h_i$ is at least r/c away from the attracting points all x_j and x_j^{-1}. Additionally we can require that g_i^{-1} takes the attracting point of $(\gamma a_i h_i)^{-1}$ at a distance at least r from the repulsive hyperplane of $(\gamma a_i h_i)^{-1}$.

Then clearly, $x_i = g_i \gamma a_i h_i$ is $(\frac{r}{c}, c^3\epsilon)$-very proximal and $(x_1, ..., x_i)$ form a ping-pong i-tuple of $(\frac{r}{c}, c^3\epsilon)$-very proximal transformations. $\square$

6.4 Constructing very proximal elements in Γ

Our aim in this section is to establish an action of Γ, on some projective space over a local field, which has many very proximal elements. Proposition 3.3 tells us how to get very proximal elements out of contracting elements when a separating set is available. We shall show how to construct an action of Γ, on a projective space over a local field, with both ingredients : contracting elements and separating set.

For an algebraic number field there are naturally associated local fields - its completions. The following lemma reduces the general case to this.

Lemma 4.1 *Let $\mathbb{G}$ be a semisimple algebraic group defined over $\mathbb{Q}$, and $G = \mathbb{G}(\mathbb{R})^0$ be the corresponding connected semisimple Lie group. Let Γ be a finitely generated dense subgroup in G. Then, there exists a number field $K \subset \mathbb{R}$ and a group homomorphism $\pi : \Gamma \to \mathbb{G}$ such that $\pi(\Gamma) \subset \mathbb{G}(K)$ and $\pi(\Gamma)$ is still dense in G.*

Proof: Choose a finite set of generators $\gamma_1, \gamma_2, \ldots, \gamma_k$ such that γ_1, γ_2 are close to 1 and (γ_1, γ_2) belongs to the complement of the proper exponential algebraic set introduced in theorem 2.4. Let $\Gamma = \{\gamma_1, ..., \gamma_k \mid (r_\alpha)_\alpha\}$ be a presentation of Γ. The variety of representations

$$\mathrm{Hom}\big(\Gamma, \mathbb{G}\big) = \{(g_1, ..., g_k) \in \mathbb{G}^k : \forall \alpha,\ r_\alpha(g_1, ..., g_k) = 1\}$$

is an algebraic subvariety of $\mathbb{G}^k$ which is defined over $\mathbb{Q}$. Its set of algebraic points is therefore dense (for the real topology) in its set of real points. Choosing an algebraic point in $\mathrm{Hom}\big(\Gamma, \mathbb{G}(\mathbb{R})\big)$ very close to the original $(\gamma_1, ..., \gamma_k)$, we obtain a representation π of Γ into G such that $\pi(\Gamma)$ is contained in $\mathbb{G}(\overline{\mathbb{Q}})$ and is still dense in G, as small deformations of γ_1, γ_2 generate dense subgroups. Finally, since it is finitely generated, it is contained in $\mathbb{G}(K)$ for some number field K. $\square$

As explained in the last section, "ε-contraction" is equivalent to "big ratio between the first and the second diagonal entries in the KAK decomposition."

Lemma 4.2 (Constructing contracting elements) *Let K be a number field and $\mathbb{G}$ be a non-trivial semisimple Zariski-connected algebraic group defined over K. Let R be a finitely generated subring of K and I an infinite subset of $\mathbb{G}(R)$. Then for some completion k of K there is an irreducible rational representation $\rho : \mathbb{G} \to \mathbb{SL}(V)$ defined over k, with $\dim_k V \geq 2$, such that the set*

$$\left\{ \frac{|a_1\big(\rho(g)\big)|}{|a_2\big(\rho(g)\big)|} : g \in I \right\}$$

is unbounded, where $a_1\big(\rho(g)\big)$ and $a_2\big(\rho(g)\big)$ are the first and second diagonal terms in a Cartan decomposition of $\rho(g)$ in $\mathbb{SL}(V)$.

Proof: As R is finitely generated, the discrete diagonal embedding of K in its adèle group gives rise to a discrete embedding of R into a product of finitely many places. Explicitly, there is a finite set S of places of K, including all archimedean ones, such that R is contained in the ring of S-integers $\mathcal{O}_K(S)$. Projecting to the finite product of places corresponding to S, we get a discrete embedding

$$R \hookrightarrow \prod_{\nu \in S} K_\nu,$$

which gives rise to a discrete embedding

$$\mathbb{G}(R) \hookrightarrow \prod_{\nu \in S} \mathbb{G}(K_\nu).$$

The infinite set I is mapped to some unbounded set in the latter product. Hence for some place $\nu \in S$, the embedding $\mathbb{G}(R) \hookrightarrow \mathbb{G}(K_\nu)$ sends I to an unbounded set. Take $k = K_v$. Since $\mathbb{G}$ is also defined over k, there exists a k-rational faithful representation of $\mathbb{G}$ on some k-vector space V_0. Under this representation, the set I is sent to some unbounded set in $\mathbb{SL}(V_0)$ which we shall continue to denote by I.

Fix a k-basis of V_0 and consider the Cartan decomposition in $\mathbb{SL}(V_0)$ corresponding to this basis. Recall the notations of section 3.30. For $x \in \mathbb{SL}(V_0)$, $|a_1(x)| \geq |a_2(x)| \geq \ldots \geq |a_d(x)| > 0$ denote the corresponding diagonal coefficients in the Cartan decomposition of x in $\mathbb{SL}(V_0)$. Since their product is 1, there is some i $(1 \leq i < d)$ for which the set $\{\frac{|a_i(g)|}{|a_{i+1}(g)|}\}_{g\in I}$ is unbounded.

Considering the i'th wedge product representation of the above representation of $\mathbb{G}$, we obtain a rational representation

$$\rho_0 : \mathbb{G} \to \mathbb{SL}(\bigwedge^i V_0)$$

defined over k. Looking at the Cartan decomposition in $\mathbb{SL}(\bigwedge^i V_0)$ with respect to the basis of $\bigwedge^i V_0$ induced by the given basis of V_0, we see that $a_1\big(\rho_0(x)\big) = a_1(x)...a_i(x)$ and $a_2\big(\rho_0(g)\big) = a_1(x)...a_{i-1}(x)a_{i+1}(x)$, and hence, under this representation, the set

$$\left\{\frac{|a_1\big(\rho_0(g)\big)|}{|a_2\big(\rho_0(g)\big)|}\right\}_{g\in I}$$

is unbounded.

The representation ρ_0 might be reducible. However, as $\mathbb{G}(k)$ is a connected semisimple algebraic group, the k-rational representation ρ_0 is completely reducible. Let $W = \bigwedge^i V_0$ be the representation space. Decompose it into $\mathbb{G}(k)$-irreducible spaces $W = W_1 \oplus ... \oplus W_q$. Any $x = \rho_0(g) \in \rho_0\big(\mathbb{G}(k)\big) \leq \mathbb{SL}(W)$ stabilizes each W_i, and we shall write $x_i \in \mathbb{SL}(W_i)$ for its restriction to the subspace W_i (note that the determinant of x_i is 1 as $\mathbb{G}(k)$ is semisimple).

The space W is endowed with the norm $\|\cdot\|$ corresponding to the above choice of a basis of V_0 (as in section 3.30) for which $|a_1(g)| = \|g\|$ and $|a_1(g)a_2(g)| = \|\bigwedge^2 g\|$ for any $g \in \mathbb{SL}(W)$. Similarly we can choose a basis and corresponding norms $\|\cdot\|_i$ for each W_i such that

$$|a_1(x_i)| = \|x_i\|_i \text{ and } |a_1(x_i)a_2(x_i)| = \left\|\bigwedge^2 x_i\right\|_i.$$

As any two norms on the finite dimensional vector space $M_n(k)$, $\big(n = \dim(W)\big)$ are equivalent, there is $c_1 > 0$ with respect to which

$$\frac{1}{c_1} \max_{1\leq j\leq q}\{\|x_j\|_j\} \leq \|x\| \leq c_1 \max_{1\leq j\leq q}\{\|x_j\|_j\}.$$

Similarly, for some constant $c_2 > 0$

$$\frac{1}{c_2}\max_{i,j}\{\left\|(\bigwedge^2 x)_{ij}\right\|_{ij}\} \leq \left\|\bigwedge^2 x\right\| \leq c_2 \max_{i,j}\{\left\|(\bigwedge^2 x)_{ij}\right\|_{ij}\}.$$

Replacing I by an infinite subset if necessary, we can find an index $i_0 \leq q$ with $\|x\| \leq c_1 \|x_{i_0}\|_{i_0}$ for any x of the form $x = \rho_0(g)$, $g \in I$. As $\{\|\rho_0(g)\|\}_{g\in I}$ is unbounded, and each x_i belongs to $\mathbb{SL}(W_i)$, we must have $\dim W_{i_0} \geq 2$. Moreover, we have

$$\left\|\bigwedge^2 x\right\| \geq \frac{1}{c_2}\left\|(\bigwedge^2 x)_{ii}\right\|_{ii} = \frac{1}{c_2}\left\|\bigwedge^2 x_i\right\|_i,$$

which implies

$$\frac{|a_1(x)|}{|a_2(x)|} \leq c_1^2 c_2 \frac{|a_1(x_{i_0})|}{|a_2(x_{i_0})|}.$$

Hence $\rho = (\rho_0)_{|W i_0}$ is the desired rational irreducible representation.

□

Lemma 4.3 (Constructing separating sets) *Let m be a positive integer, k a local field, $\mathbb{G}$ a Zariski-connected k-algebraic group, V a k-vector space, and $\rho : \mathbb{G} \to \mathbb{SL}(V)$ a k-irreducible rational representation. Then for any Zariski-dense subset $\Omega \subset \mathbb{G}(k)$, there exists a finite subset $F \subset \Omega$ and a positive real number $r > 0$, which gives rise under ρ to an (m, r)-separating set (cf. 3.7) on $\mathbb{P}(V)$.*

Proof: For each $\gamma \in \Gamma$, let M_γ be the set of all tuples $(v_i, H_i)_{i=1}^{2m}$ of $2m$ points $v_i \in \mathbb{P}(V)$ and $2m$ projective hyperplanes $H_i \subset \mathbb{P}(V)$, not necessarily different, such that for some $1 \leq i, j \leq 2m$, $\gamma \cdot v_i \in H_j$ or $\gamma^{-1} \cdot v_i \in H_i$.

We first claim that

$$\bigcap_{\gamma\in\Omega} M_\gamma = \emptyset. \tag{6.5}$$

Suppose this were not the case. Then there would exist $(v_i, H_i)_{i=1}^{2m}$ such that

$$\Omega \subset \bigcup_{1\leq i,j\leq 2m} \left\{\gamma \in \mathbb{G}(k), \gamma \cdot v_i \in H_j \text{ or } \gamma^{-1} \cdot v_i \in H_j\right\}.$$

The sets $\{\gamma \in \mathbb{G}(k) : \gamma \cdot v_i \in H_j\}$ and $\{\gamma^{-1} \in \mathbb{G}(k) : \gamma \cdot v_i \in H_j\}$ are clearly Zariski-closed and proper, since $\mathbb{G}(k)$ acts irreducibly on V. As $\mathbb{G}(k)$ is Zariski-connected, this yields a contradiction. Thus we have (6.5).

Second, as the sets M_γ are compact in the appropriate product of grassmannians, there is a finite subset $F \subset \Omega$, such that

$$\bigcap_{\gamma\in F} M_\gamma = \emptyset. \tag{6.6}$$

The following MaxMin of continuous functions

$$\max_{\gamma\in F}\Big(\min_{1\le i,j\le 2m}\{d(\gamma\cdot v_i,H_j),d(\gamma^{-1}\cdot v_i,H_j)\}\Big) \tag{6.7}$$

depends continuously on $(v_i,H_i)_{i=1}^{2m}$, and by (6.6) never vanishes. Finally, the compactness of the set of all tuples $(v_i,H_i)_{i=1}^{2m}$ in

$$\big(\mathbb{P}(V)\times\mathbb{G}r_{\dim(V)-1}(V)\big)^{2m}$$

implies that (6.7) attains a positive minimum r. Thus F gives rise to the desired (m,r)-separating set. $\square$

Putting together proposition 3.3 with lemmas 4.2 and 4.3, we obtain a great liberty in the choice of very contracting (or very proximal) elements in Γ.

Lemma 4.4 *Let m be a positive integer, K a number field, $\mathbb{G}$ a Zariski-connected semisimple K-algebraic group, $R\subset K$ a finitely generated subring, $I\subset\mathbb{G}(R)$ an infinite subset, and $\Omega\subset\mathbb{G}(K)$ a Zariski-dense subset of $\mathbb{G}$. Then there exist :*

- *An embedding $\sigma: K\hookrightarrow k$ of K into some local field k.*
- *An irreducible rational representation $\rho:\mathbb{G}\to\mathbb{SL}(V)$ defined over k,*
- *A finite (m,r)-separating set $F\subset\Omega$ (for some $r>0$) for the action induced by ρ on the projective space $\mathbb{P}(V)$,*

such that for any $\varepsilon>0$, there is $g\in I$ and $f\in F$ for which gfg^{-1} acts as an ε-very contracting transformation on $\mathbb{P}(V)$.

Proof:

Lemma (4.2) yields the local field k, the embedding $K\hookrightarrow k$, and the representation ρ, while lemma (4.3) yields the r-separating set $F\subset\Omega$.

For any $\varepsilon>0$, we can find $g\in I$ with

$$\frac{a_1(\rho(g))}{a_2(\rho(g))}\ge\frac{r}{\epsilon\sqrt{2Cd}}$$

where $C=\max\{\mathrm{biLip}(f): f\in F\}$ and d is as in proposition (3.3) ($=4$ in archimedean case, $\frac{1}{|\pi|}$ in the non-archimedean case). By lemma (3.4), g acts as a $\frac{\epsilon\sqrt{2Cd}}{r}$-contracting transformation on $\mathbb{P}(V)$. Finally, for ϵ small enough, proposition (3.3) provides $f\in F$ such that gfg^{-1} is ϵ-very contracting.

$\square$

Finally, we obtain the following result :

Theorem 4.5 *Let $\mathbb{G}$ be a Zariski connected semisimple algebraic K-group, where K is a number field. Let R be a finitely generated subring of K, and $\Omega\subset\mathbb{G}(R)$ a Zariski dense subset of $\mathbb{G}$ with $\Omega=\Omega^{-1}$. Then for any $m\in\mathbb{N}$ and any $a_1,\dots,a_m\in\mathbb{G}(K)$ there are $x_1,\dots,x_m$ with $x_i\in\Omega^4a_i\Omega$ which are generators of a free group F_m.*

Proof: This follows from the combination of lemma 4.4 and proposition 3.11. □

Remark 4.6 *Note that the original statement of Tits' alternative ([35], theorem 1) can be easily derived from the last theorem, first by taking a homomorphism of* Γ *into* $\mathbb{GL}_n(K)$*, for some number field* K*, whose image is not almost solvable (this is possible as the index of the solvable subgroup of an almost solvable group and the length of the derived series of solvable groups are both uniformly bounded for groups contained in* $\mathbb{GL}_n$*) and then projecting to a semisimple factor of the Zariski closure. Then take* $\Omega = \Gamma$ *in the last theorem.*

6.5 Proofs of the main theorems

To complete the proofs of the main results of this paper, we shall need some preliminary lemmas.

Lemma 5.1 *Let* G *be a connected real Lie group and* $\Gamma \subset G$ *a dense subgroup. Then* Γ *is generated by* $\Gamma \cap U$*, for any identity neighborhood* U *in* G*.*

Proof: Let H be the subgroup of G generated by $\Gamma \cap U$. Since Γ is dense and G connected, H is again dense in G. If $\gamma \in \Gamma$ we can thus find $h \in H$ such that $h \in \gamma U$. But then $\gamma^{-1}h \in \Gamma \cap U \subset H$. Hence $\Gamma = H$. □

Lemma 5.2 *Let* $\mathbb{G}$ *be a Zariski-connected algebraic group defined over* $\mathbb{R}$ *and let* $G = \mathbb{G}(\mathbb{R})^0$ *be the corresponding connected real Lie group. Let* Γ *be a dense subgroup in* G *and* $\pi : \Gamma \to G$ *a group homomorphism such that* $\pi(\Gamma)$ *is Zariski-dense. Then for any identity neighborhood* $U \subset G$*,* $\pi(U \cap \Gamma)$ *is Zariski-dense in* $\mathbb{G}$*.*

Proof: Let $(U_n)_n$, $U_n \supset U_{n+1}$, be a nested sequence of identity neighborhoods in G, such that for any identity neighborhood U, $U_n \subset U$ for large n's. Let $U'_n = U_n \cap \Gamma$ and let X_n be the Zariski closure of $\pi(U'_n)$ in G. Since $(X_n)_n$ is a decreasing sequence of Zariski-closed subsets of G, it stabilizes at some finite step n_0. For any $g \in U'_{n_0}$, $gU'_n \subset U'_{n_0}$, for n large enough, hence $\pi(g)X_{n_0} = \pi(g)X_n \subset X_{n_0}$. By lemma 5.1, U'_{n_0} generates Γ, hence $\pi(\Gamma)X_{n_0} \subset X_{n_0}$. Since $\pi(\Gamma)$ is Zariski dense, this implies $X_{n_0} = G$. Thus $\pi(U \cap \Gamma)$ is Zariski-dense for any identity neighborhood $U \subset G$. □

Proof: [**Proof of theorem 1.1**] The adjoint group $H = \mathrm{Ad}(G)$ is a center free connected semisimple Lie group, hence is of the form $\mathbb{H}(\mathbb{R})^0$ where $\mathbb{H}$ is some Zariski-connected semisimple $\mathbb{Q}$-algebraic group.

By Corollary 2.5, we can assume that Γ is finitely generated. By lemma 4.1, there is a homomorphism π of Γ into $\mathrm{Ad}(G)$ with dense, hence Zariski-dense, image which lies in $\mathbb{H}(K)$ for some number field K. Since Γ is finitely generated, there is a finitely generated subring R of K, such that $\pi(\Gamma) \subset \mathbb{H}(R)$.

Consider two elements a and b in Γ and a sufficiently small identity neighborhood U' in G such that for any $x \in V(a) := U'aU'$ and $y \in V(b) := U'bU'$, $\langle x, y\rangle$ is a dense subgroup of G. By theorem 2.4 this is always possible. Let U be a smaller symmetric identity neighborhood in G such that $U^4 \subset U'$ and denote

$$\Omega = \pi(\Gamma \cap U).$$

By lemma 5.2, the conditions of theorem 4.5 are fulfilled. We thus obtain the desired dense free subgroup.

□

The proof of theorem 1.2 is basically the same as the proof of theorem 1.1 and we shall not go over it again, but only describe the needed modifications in the above argument.

Proof: [**Proof of theorem 1.2 (outline)**]

Pick $a_1, \ldots, a_m \in \Gamma$ close to the identity, and a small symmetric identity neighborhood U such that for any selection $x_i \in U^4 a_i U$, $i = 1, \ldots, m$, the group $\langle x_1, \ldots, x_m \rangle$ is dense in G. The existence of such a_i's and U is an easy consequence of theorem 2.1.

Now project G on the semisimple quotient $S = G/R$ where R is the radical of G, and find as above a projective space over a local field and an action of Γ (via S) for which there is

– a (m, r)-separating set $F \subset U \cap \Gamma$ (for some $r > 0$), and

– for $\epsilon < \frac{r}{2c^4}$, (where c is the maximal bi-Lipschitz constant of $\{a_1, ..., a_m\} \cup F$ when acting on the projective space), an element $\gamma_\epsilon \subset U^3 \cap \Gamma$ acting as an ϵ-very contracting element.

Then there are

$$h_i \in F \ (1 \le i \le m) \text{ and } g_i \in F \ (2 \le i \le m),$$

such that

$$\langle \gamma_\epsilon a_1 h_1, \ g_2 \gamma_\epsilon a_2 h_2, \ldots, \ g_m \gamma_\epsilon a_m h_m \rangle$$

is dense and free. □

Proof: [**Proof of theorem 1.3**] Recall that G is a connected non-solvable Lie group and $\Gamma \le G$ is a finitely generated dense subgroup. Define $G_0 = G$ and $G_n = \overline{[G_{n-1}, G_{n-1}]}$. Then as $\dim(G) < \infty$ the sequence G_n stabilizes at some finite step k, $H = G_k$ is a connected topologically perfect closed normal subgroup of G, and the quotient G/H is solvable. Moreover it follows by induction on n that $\Gamma \cap G_n$ is dense in G_n.

Let l be the codimension of H and let m be the minimal number of generators of the Lie algebra of H, then $l + m \le \dim(G)$. Since the projection of Γ to G/H is finitely generated and dense, proposition 2.7 implies that it contains $2l$ elements which generate a dense subgroup in G/H. Let $a_1, \ldots, a_{2l} \in \Gamma$ be arbitrary lifts of these elements. Additionally, as in the proof of theorem 1.2, let $a_{2l+1}, \ldots, a_{2l+m}$ be m elements in $H \cap \Gamma$ near the identity, and let $U \subset H$ be a small identity neighborhood of H, such that for any selection $x_i \in U^4 a_i U$ $(2l + 1 \le i \le 2l + m)$, the group $\langle x_{2l+1}, \ldots, x_{2l+m} \rangle$ is dense in H. Then for any selection $x_i \in U^4 a_i U$ $(1 \leq i \leq 2l+m)$ the group $\langle x_1, \ldots, x_{2l+m} \rangle$ is dense in G.

From this point, we can continue exactly as in the proof of theorem 1.2 in order to find x_i's such that $\langle x_1, \ldots, x_{2l+m} \rangle$ is free.

Remark 5.3 *The bound* $2\dim(G)$ *is not sharp as the latter proof shows. But* $2l+m$ *can also be decreased in some cases. For example, it is easy to see that if* G *is a direct product of a semisimple group with* $\mathbb{R}^d$*, then* $\max\{2, 2d\}$ *is the sharp bound for theorem 1.3.*

Acknowledgments 5.4 *We would like to thank A. Lubotzky, from whom we learned about the question, as well as G.A. Margulis for their help and support while working on this paper. We are also thankful to Y. Benoist and S. Mozes for many valuable discussions and comments and to Y. Guivarc'h for pointing out to us the work of Carrière and Ghys and the link with the theory of amenable actions.*

Chapitre 7

A topological version of the Tits alternative[1]

7.1 Introduction

In his celebrated 1972 paper [167] J. Tits proved the following fundamental dichotomy for linear groups :

Theorem 1.1 *(Tits alternative) Let K be a field and Γ a subgroup of $GL_n(K)$. In case $char(K) > 0$ assume further that Γ is finitely generated. Then either Γ contains a solvable subgroup of finite index, or Γ contains a non-commutative free subgroup on two generators.*

This theorem answered a conjecture of Bass and Serre and is an important step in the understanding of linear groups. The purpose of the present paper is to give a topological analogue of this dichotomy and to provide various applications of it.

Let k be a local field, that is $\mathbb{R}$, $\mathbb{C}$, a finite extension of $\mathbb{Q}_p$, or a field of formal power series in one variable over a finite field. $\mathrm{GL}_n(k)$ is endowed with the standard topology, i.e. the topology induced from the local field k. Now, let Γ be a group and $\sigma : \Gamma \to \mathrm{GL}_n(k)$ some injective homomorphism. We view Γ as a topological group endowed with the topology obtained by pulling back by σ the standard topology on $\mathrm{GL}_n(k)$ (i.e. the open sets are $\sigma^{-1}(O)$ where O is some open set in $\mathrm{GL}_n(k)$). We then prove the following topological version of the Tits alternative :

Theorem 1.2 *Suppose Γ contains no open solvable subgroup. Then Γ contains a dense free subgroup.*

Note that Γ may contain both a dense free subgroup and an open solvable subgroup : in this case Γ has to be discrete and free. For non discrete groups however, the two cases are mutually exclusive.

[1]joint work with T. Gelander [32]

In general, the dense free subgroup from Theorem 1.2 may have an infinite (but countable) number of free generators. However, in many cases (see below Theorems 5.1 and 5.7), for example when Γ itself is finitely generated, then we can find a dense free subgroup on finitely many free generators. For another example consider the group $\mathrm{SL}_n(\mathbb{Q})$, $n \geq 2$. It is not finitely generated, yet, we show that it contains a free subgroup of rank 2 which is dense with respect to the topology induced from $\mathrm{SL}_n(\mathbb{R})$. Similarly, for any prime $p \in \mathbb{N}$, we show that $\mathrm{SL}_n(\mathbb{Q})$ contains a free subgroup of finite rank $r = r(p) \geq 2$ which is dense with respect to the topology induced from $\mathrm{SL}_n(\mathbb{Q}_p)$.

When $char(k) = 0$, the linearity assumption can be replaced by the weaker assumption that Γ is contained in some second-countable k-analytic Lie group G. In particular, Theorem 1.2 applies to subgroups of any real Lie group with countably many connected components, and to subgroups of any group containing a p-adic analytic pro-p group as an open subgroup of countable index.

In [31] we proved Theorem 1.2 in the special case where $k = \mathbb{R}$ and Γ is dense in a connected real Lie group : if Γ is a finitely generated dense subgroup of a non solvable connected real Lie group, then Γ contains a free subgroup of finite rank which is also dense. The main difficulty in the general case lies in the study of representations of algebraic groups which are not Zariski connected. A similar situation arose in [116].

Theorem 1.2 has various applications. In sections 7.7, 7.8 and 7.9 we shall concentrate on developing some of them.

When k is non-Archimedean, Theorem 1.2 provides some new results about pro-finite groups (see Section 7.7). In particular, we answer a conjecture of Dixon, Pyber, Seress and Shalev (cf. [50] and [135]), by proving :

Theorem 1.3 *Let Γ be a finitely generated linear group over an arbitrary field. Suppose that Γ is not virtually solvable, then its pro-finite completion $\hat{\Gamma}$ contains a dense free subgroup of finite rank.*

In [50], using the classification of finite simple groups, the weaker statement, that $\hat{\Gamma}$ contains a free subgroup whose closure is of finite index, was established. Let us remark again that the passage from a subgroup whose closure is of finite index, to a dense subgroup is a central part in the proof of Theorem 1.2. We also note that Γ itself may not contain a pro-finitely dense free subgroup of finite rank. It was shown in [148] that surface groups have the property that any proper finitely generated subgroup is contained in a proper subgroup of finite index (see also [155]).

In Section 7.7 we also answer a conjecture of Shalev about coset identities for p-adic analytic pro-p groups.

The question of existence of a free subgroup is indeed closely related to questions concerning amenability. It follows from the Tits alternative that a finitely generated linear group is amenable if and only if it contains no free subgroups and if and only if it contains a solvable subgroup of finite index. The topology enters the game when considering actions of subgroups on the full group. It follows from Theorem 1.2 that the action by left multiplications of a countable subgroup Γ of $G = \mathrm{GL}_n(k)$ on G, where k is a local field, is amenable if and only if Γ contains no relatively dense free subgroup and if and only if Γ contains a relatively open solvable subgroup. We refer the reader to [179], Chapter 4, for an introduction and background on amenable actions. For $k = \mathbb{R}$ this answers a question of Carrière and Ghys [40] and provides a short proof for a conjecture of Connes and Sullivan which was first proved by Zimmer [178] by means of super rigidity methods. We obtain here the following generalization of Zimmer's theorem for arbitrary locally compact groups (see Section 7.8) :

Theorem 1.4 *Let Γ be a countable subgroup of a locally compact topological group G. Then the action of Γ on G (as well as on G/P for $P \leq G$ closed amenable) by left multiplication is amenable, if and only if Γ contains a relatively open subgroup which is amenable as an abstract group.*

Theorem 1.4 implies the following generalization of Auslander's theorem (see [137] Theorem 8.24) :

Theorem 1.5 *Let G be a locally compact topological group, let $P \leq G$ be a closed normal amenable subgroup, and let $\pi : G \to G/P$ be the canonical projection. Suppose that $H \leq G$ is a subgroup which contains a relatively open amenable subgroup. Then $\pi(H)$ also contains a relatively open amenable subgroup.*

We also apply Theorem 1.2 to prove the following dichotomy between polynomial and exponential growth for leaves of Riemannian foliations (see [80], [39], [120]) :

Theorem 1.6 *Let $\mathcal{F}$ be a Riemannian foliation on a compact manifold M. The leaves of $\mathcal{F}$ have polynomial growth if and only if the structural Lie algebra of $\mathcal{F}$ is nilpotent. Otherwise, generic leaves have exponential growth.*

The first half of Theorem 1.6 was actually proved by Carrière in [39]. Using Zimmer's proof of the Connes-Sullivan conjecture, he first reduced to the solvable case, then he proved the nilpotency of the structural Lie algebra of $\mathcal{F}$ by a delicate direct argument (see also [80]). He then asked whether the second half of this theorem holds. Both parts of Theorem 1.6 follow from Theorem 1.2 and the methods developed in its proof.

The strategy used in this article to prove Theorem 1.2 consists in perturbing the generators γ_i of Γ within Γ and in the topology of $\mathrm{GL}_n(k)$, in order to obtain (under the assumption that Γ has no solvable open subgroup) free generators of a free subgroup, which is still dense in Γ. As it turns out, there exists an identity neighborhood U of some non virtually solvable subgroup $\Delta \leq \Gamma$, such that any selection of points x_i in $U\gamma_i U$ generate a dense subgroup in Γ. The argument used here to prove this claim depends on whether k is Archimedean, p-adic or of positive characteristic.

In order to find a free group, we use a variation of the so-called ping-pong method used by Tits. The original method of Tits (via the use of high powers of semisimple elements to produce ping-pong players) is not applicable to our situation and a more careful study of the contraction properties of projective transformations is necessary. In a first, more algebraic step, we construct a representation ρ of Γ into some $\mathrm{GL}_n(K)$ for some other local field K such that the Zariski closure of $\rho(\Delta)$ acts strongly irreducibly (i.e. fixes no finite union of proper vector subspaces). We then make use of the ideas developed in [31] and inspired from [1], where it is shown how the dynamical properties of a projective transformation can be read off on its Cartan decomposition. This allows, provided that some additional assumption on ρ holds, to produce a set of elements in U which "play ping-pong" on the projective space $\mathbb{P}(K^n)$, and hence generate a free group (see Theorem 4.3). Theorem 4.3 provides a very handy way to generate free subgroups, as soon as some infinite subset of matrices with entries in a given finitely generated ring (e.g. an infinite subset of a finitely generated linear group) is given. The required additional assumption on ρ is the existence of so-called proximal elements lying in $\rho(U)$ (a proximal transformation is a transformation of $\mathbb{P}(K^n)$ which contracts almost all $\mathbb{P}(K^n)$ into a small ball). In [31] we developed a method to produce such a representation when the Zariski closure of Γ is connected. But in case it is not connected, a much more careful study of the possible candidates for ρ has to be made. This is performed in Section 7.4.

One of the main ingredients in the proof of the existence of the representation ρ with suitable proximal elements is a result (generalizing a lemma of Tits) asserting that in an arbitrary finitely generated integral domain, any infinite set can be sent to an unbounded set under an appropriate embedding of the ring into some local field (see Section 7.2). This result is crucial in particular when dealing with non finitely generated subgroups. Our proof uses a striking simple fact, originally due to Pólya in the case $k = \mathbb{C}$, about the inverse image of the unit disc under polynomial transformations (see Lemma 2.3).

The paper is organized as follows. In Section 7.2, we give a proof of the above mentioned lemma. In Section 7.3 we recall and complete some of the results from [31] concerning the action of projective transformations and the Cartan decomposition, then move on proving Theorem 4.3 in Section 7.4. In Section 7.5, we prove our main theorem in the case when Γ is finitely generated. In Section 7.6, using the result of the preceding section, we prove the existence of dense free subgroups on infinitely many free generators. We then devote Sections 7.7, 7.8 and 7.9 to prove the main corollaries and applications of Theorem 1.2. In the last section, we discuss some related problems.

Finally, let us give here a few notations that will be used throughout the paper. The notation $H \leq G$ means that H is a subgroup of the group G. By $[G, G]$ we denote the derived group of G, i.e. the group generated by commutators. Given a group Γ, we denote by $d(\Gamma) \in \mathbb{N}$ the minimal size of a generating set of Γ. If $\Omega \subset G$ is a subset if G, then $\langle\Omega\rangle$ denotes the subgroup of G generated by Ω. If Γ is a subgroup of an algebraic group, we denote by $\overline{\Gamma}^z$ its Zariski closure. Note that the Zariski topology on rational points does not depend on the field of definition, that is if V is an algebraic variety defined over a field K and if L is any extension of K, then the K-Zariski topology on $V(K)$ coincides with the trace of the L-Zariski topology on it. To avoid confusion, we shall always add the prefix "Zariski" to any topological notion regarding the Zariski topology (e.g. "Zariski dense", "Zariski open"). For the topology inherited from the local field k, however, we shall plainly say "dense" or "open" without further notice (e.g. $\mathrm{SL}_n(\mathbb{Z})$ is open and Zariski dense in $\mathrm{SL}_n(\mathbb{Z}[1/p])$, where $k = \mathbb{Q}_p$).

7.2 A generalization of a lemma of Tits

In the original proof of the Tits alternative, Tits used an easy but crucial lemma saying that given a finitely generated field K and an element $\alpha \in K$ which is not a root of unity, there always is a local field k and an embedding $f : K \to k$ such that $|f(\alpha)| > 1$. A natural and useful generalization of this statement is the following lemma, which may also be considered for its own sake.

Lemma 2.1 *Let R be a ring. Suppose that R is finitely generated (as a ring) and has no zero divisors. Let I be an infinite subset of R. Then there exists a local field k and an embedding $i : R \hookrightarrow k$ such that $i(I)$ is unbounded.*

As explained below, this lemma provides a straightforward way to build the proximal elements needed in the construction of dense free subgroups. It is more general and more natural than the argument used in [31].

Before getting into the proof of Lemma 2.1 let us demonstrate its usefulness by proving a straightforward consequence :

Corollary 2.2 (Zimmer [180], Theorems 6 and 7, and [84] 6.26) *There is no faithful conformal action of an infinite Kazhdan group on the Euclidean 2-sphere S^2.*

Proof: Suppose there is an infinite Kazhdan subgroup Γ in $\mathrm{SL}_2(\mathbb{C})$, the group of conformal transformations of S^2. Since Γ has property (T), it is finitely generated, and hence, Lemma 2.1 could be applied to yield a faithful representation of Γ into $\mathrm{SL}_2(k)$ for some local field k, with unbounded image. However $\mathrm{PSL}_2(k)$ acts faithfully with compact isotropy groups by isometries on the hyperbolic space $\mathbb{H}^3$ if k is Archimedean, and on a tree if it is not. As Γ has property-(T), it must fix a point (c.f. [84] 6.4 and 6.23 or [180] Prop. 18) and hence lie in some compact group. A contradiction. $\square$

When R is integral over $\mathbb{Z}$, the lemma follows easily by considering the diagonal embedding of R into a product of finitely many completions of its field of fractions. The main difficulty comes from the possible presence of transcendental elements. Our proof of Lemma 2.1 relies on the following interesting fact. Let k be a local field, and let $\mu = \mu_k$ denote the standard Haar measure on k, i.e. the Lebesgue measure if k is Archimedean, and the Haar measure giving measure 1 to the ring of integers $\mathcal{O}_k$ of k when k is non-Archimedean. Given a polynomial P in $k[X]$, let

$$A_P = \{x \in k, |P(x)| \leq 1\}.$$

Lemma 2.3 *For any local field k, there is a constant $c = c(k)$ such that $\mu(A_P) \leq c$ for any monic polynomial $P \in k[X]$.*

Proof: Let $\overline{k}$ be an algebraic closure of k, and P a monic polynomial in $k[X]$. We can write $P(X) = \prod(X - x_i)$ for some $x_i \in \overline{k}$. The absolute value of k extends uniquely to an absolute value in $\overline{k}$ (see [103] XII, 4, Theorem 4.1 p. 482). Now if $x \in A_P$ then $|P(x)| \leq 1$, and hence

$$\sum \log|x - x_i| = \log|P(x)| \leq 0.$$

But A_P is measurable and bounded, therefore integrating with respect to μ, we obtain

$$\sum \int_{A_P} \log|x - x_i|\, d\mu(x) = \int_{A_P} \sum \log|x - x_i|\, d\mu(x) \leq 0.$$

The lemma will now follow from the following **claim** : *for any measurable set $B \subset k$ and any point $z \in \overline{k}$,*

$$\int_B \log|x - z|\, d\mu(x) \geq \mu(B) - c, \tag{7.1}$$

where $c = c(k) > 0$ is some constant independent of z and B.

Indeed, let $\widetilde{z} \in k$ be such that $|\widetilde{z} - z| = \min_{x \in k}|x - z|$, then $|x - z| \geq |x - \widetilde{z}|$ for all $x \in k$, so

$$\int_B \log|x - z|\, d\mu(x) \geq \int_B \log|x - \widetilde{z}|\, d\mu(x) = \int_{B - \widetilde{z}} \log|x|\, d\mu(x).$$

Therefore, it suffices to show (1) when $z = 0$. But a direct computation for each possible field k shows that $-\int_{|x|\leq 1} \log|x| d\mu(x) < \infty$. Therefore taking $c = \mu\{x \in k, |x| \leq e\} + |\int_{|x|\leq 1} \log|x| d\mu(x)|$ we obtain (1). This concludes the proof of the lemma. □

Lemma 2.3 was proved by Pólya in [133] for the case $k = \mathbb{C}$ by means of potential theory. Pólya's proof gives the best constant $c(\mathbb{C}) = \pi$. For $k = \mathbb{R}$ one can show that the best constant is $c(\mathbb{R}) = 4$ and that it can be realized as the limit of the sequence of lengths of the pre-image of $[-1, 1]$ by the Chebyshev polynomials (under an appropriate normalization of these polynomials). In the real case, this result admits generalizations to arbitrary smooth functions such as the Van der Corput lemma (see [36] for a multi-dimensional analog). For k non-Archimedean, the constant is always < 2 and it tends to 1 as the residue field $\mathcal{O}_k/\pi\mathcal{O}_k$ gets larger.

Let us just explain how, with a little more consideration, one can improve the constant c in the above proof[2]. We wish to find the minimal $c > 0$ such that for every compact subset B of k whose measure is $|B| \geq c$ we have

$$\int_B \log|x| d\mu(x) \geq 0.$$

[2]Let us also remark that there is a natural generalization of Lemma 2.3 to higher dimension which follows by an analogous argument : For any local field k and $n \in \mathbb{N}$, there is a constant $c(k, n)$, such that for any finite set $x_1, \ldots, x_m \in k^n$, we have $\mu\big(\{y \in k^n : \prod_1^m \|y - x_i\| \leq 1\}\big) \leq c(k, n)$.

Suppose $k = \mathbb{C}$. Since $\log |x|$ is increasing with $|x|$, for any B

$$\int_B \log |x| d\mu(x) \geq \int_C \log |x| d\mu(x)$$

where C is a ball around 0 ($C = \{x \in k : |x| \leq t\}$) with the same volume as B. Therefore $c = \pi t^2$ where t is such that $2\pi \int_0^t r \log(r) dr = 0$. The unique positive root of this equation is $t = \sqrt{e}$. Thus we can take

$$c = \pi e.$$

For $k = \mathbb{R}$ the same argument gives a possible constant $c = 2e$, while for k non-Archimedean it gives $c = 1 + \frac{1}{f(q-1)}$ where $q = p^f$ is the size of the residue class field and f is its dimension over its prime field $\mathbb{F}_p$. In particular, the constant $c(k)$ can be chosen to be independent of the local field k.

As in the proof of Lemma 2.3, there is a positive constant c_1 such that the integral of $\log |x|$ over a ball of measure c_1 centered at 0 is at least 1. This implies :

Corollary 2.4 *For any monic polynomial $P \in k[X]$, the integral of $\log |P(x)|$ over any set of measure grater then c_1 is at least the degree $d^\circ P$.*

We shall also need the following two propositions :

Proposition 2.5 *Let k be a local field and k_0 its prime field. If $(P_n)_n$ is a sequence of monic polynomials in $k[X]$ such that the degrees $d^\circ P_n \to +\infty$ as $n \to \infty$, and $\xi_1, \dots, \xi_m$ are given numbers in k, then there exists a number $\xi \in k$, transcendental over $k_0(\xi_1, \dots, \xi_m)$, such that $(|P_n(\xi)|)_n$ is unbounded in k.*

Proof: Let T be the set of numbers in k which are transcendental over $k_0(\xi_1, \dots, \xi_m)$. Then T has full measure. For every $r > 0$ we consider the compact set

$$K_r = \{x \in k, \forall n \ \ |P_n(x)| \leq r\}.$$

We now proceed by contradiction. Suppose $T \subset \bigcup_{r>0} K_r$. Then for some large r, we have $\mu(K_r) \geq c_1$, where $c_1 > 0$ is the constant from Corollary 2.4. This implies

$$d^\circ P_n \leq \int_{K_r} \log \big(|P_n(x)|\big) d\mu(x) \leq \mu(K_r) \log r,$$

contradicting the assumption of the proposition. $\square$

Proposition 2.6 *If $(P_n)_n$ is a sequence of distinct polynomials in $\mathbb{Z}[X_1, \dots, X_m]$ such that $\sup_n d^\circ P_n < \infty$, then there exist algebraically independent numbers $\xi_1, \dots, \xi_m$ in $\mathbb{C}$ such that $(|P_n(\xi_1, \dots, \xi_m)|)_n$ is unbounded in $\mathbb{C}$.*

Proof: Let $d = \sup_n d^\circ P_n$ and T be the set of all m-tuples of complex numbers algebraically independent over $\mathbb{Z}$. The P_n's lie in

$$\{P \in \mathbb{C}[X_1, \ldots, X_m] : d^\circ P \leq d\}$$

which can be identified, since T is dense and polynomials are continuous, as a finite dimensional vector subspace V of the $\mathbb{C}$-vector space of all functions from T to $\mathbb{C}$. Let $l = \dim_{\mathbb{C}} V$. Then, as it is easy to see, there exist $(\overline{x}_1, \ldots, \overline{x}_l) \in T^l$, such that the evaluation map $P \mapsto \big(P(\overline{x}_1), \ldots, P(\overline{x}_l)\big)$ from V to $\mathbb{C}^l$ is a linear isomorphism from V to $\mathbb{C}^{\dim V}$. Since the P_n's belong to a $\mathbb{Z}$-lattice in V, so does their image under the evaluation map. Since the P_n's are all distinct, $\{P_n(\overline{x}_i)\}$ is unbounded for an appropriate $i \leq l$. $\square$

Proof: [Proof of Lemma 2.1]Let us first assume that the characteristic of the field of fractions of R is 0. By Noether's normalization theorem, $R \otimes_{\mathbb{Z}} \mathbb{Q}$ is integral over $\mathbb{Q}[\xi_1, \ldots, \xi_m]$ for some algebraically independent elements $\xi_1, \ldots, \xi_m$ in R. Since R is finitely generated, there exists an integer $l \in \mathbb{N}$ such that the generators of R, hence all elements of R, are roots of monic polynomials with coefficients in $S = \mathbb{Z}[\frac{1}{l}, \xi_1, \ldots, \xi_m]$. Hence $R_0 = R[\frac{1}{l}, \xi_1, \ldots, \xi_k] = R[\frac{1}{l}]$ is integral over S. Let F be the field of fractions of R_0 and K that of S. Then F is a finite extension of K and there are finitely many embeddings $\sigma_1, \ldots, \sigma_r$ of F into some algebraic closure $\overline{K}$ of K. Note that S is integrally closed. Therefore if $x \in R$, the characteristic polynomial of x over F belongs to $S[X]$ and equals

$$\prod_{1 \leq i \leq r} (X - \sigma_i(x)) = X^r + \alpha_r(x) X^{r-1} + \ldots + \alpha_1(x)$$

where each $\alpha_i(x) \in S$. Since I is infinite, we can find i_0 such that $\{\alpha_{i_0}(x)\}_{x \in I}$ is infinite. This reduces the problem to the case $R = S$, for if S can be embedded in a local field k such that $\{|\alpha_{i_0}(x)|\}_{x \in I}$ is unbounded, then for at least one i, the $|\sigma_i(x)|$'s will be unbounded in some finite extension of k in which F embeds.

So assume $I \subset S = \mathbb{Z}[\frac{1}{l}, \xi_1, \ldots, \xi_m]$ and proceed by induction of the transcendence degree m.

The case $m = 0$ is easy since $S = \mathbb{Z}[\frac{1}{l}]$ embeds discretely (by the diagonal embedding) in the product of finitely many copies of $\mathbb{R}$ and $\mathbb{Q}_p$'s.

Now assume $m \geq 1$. Suppose first that the total degrees of the x's in I are unbounded. Then, for say ξ_m, $\sup_{x \in I} d^\circ_{\xi_m} x = +\infty$. Let $a(x)$ be the dominant coefficient of x in its expansion as a polynomial in ξ_m. Then $a(x) \in \mathbb{Z}[\frac{1}{l}, \xi_1, \ldots, \xi_{m-1}]$ and is non zero. If $\{a(x)\}_{x \in I}$ is infinite, then we can apply the induction hypothesis and find an embedding of $\mathbb{Z}[\frac{1}{l}, \xi_1, \ldots, \xi_{m-1}]$ into some local field k for which $\{|a(x)|\}_{x \in I}$ is unbounded. Hence $I' := \{x \in I, |a(x)| \geq 1\}$ is infinite. Now $x/a(x)$ is a monic polynomial in $k[\xi_m]$, so we can then apply Proposition 2.5 and find an embedding of $\mathbb{Z}[\frac{1}{l}, \xi_1, \ldots, \xi_{m-1}][\xi_m] = S$ in k, such that $\{x/a(x)\}_{x \in I'}$ is unbounded in k. Hence, under this embedding, I is unbounded in k.

Suppose now that $\{a(x)\}_{x \in I}$ is finite. Then either $a(x) \in \mathbb{Z}[\frac{1}{l}]$ for all but finitely many x's or not. In the first case we can embed $\mathbb{Z}[\frac{1}{l}, \xi_1, \ldots, \xi_{m-1}]$ into either $\mathbb{R}$ or $\mathbb{Q}_p$ (for some prime p dividing l) so that $|a(x)| \geq 1$ for infinitely many x's, while in the second case we can find $\xi_1, \dots, \xi_{m-1}$ algebraically independent in $\mathbb{C}$, such that $|a(x)| \geq 1$ for infinitely many x in I, Then, the same argument as above, using Proposition 2.5 applies.

Now suppose that the total degrees of the x's in I are bounded. If for some infinite subset of I, the powers of l in the coefficients of x (lying in $\mathbb{Z}[\frac{1}{l}]$) are bounded from below, then we can apply Proposition 2.6 to conclude. If not, then for some prime factor p of l, we can write $x = \frac{1}{p^{n(x)}}\tilde{x}$ where $\tilde{x} \in \mathbb{Z}_p[\xi_1, \dots, \xi_m]$ with at least one coefficient of p-adic absolute value 1, and the $n(x) \in \mathbb{Z}$ are not bounded from above. By compactness, we can pick a subsequence $(\tilde{x})_{x \in I'}$ which converges in $\mathbb{Z}_p[\xi_1, \dots, \xi_m]$, and we may assume that $n(x) \to \infty$ on this subsequence. The limit will be a non-zero polynomial $\tilde{x}_0$. Pick arbitrary algebraically independent numbers $z_1, \dots, z_m \in \mathbb{Q}_p$. The limit polynomial $\tilde{x}_0$ evaluated at the point $(z_1, \dots, z_m) \in \mathbb{Q}_p^m$ is not 0, and the sequence of polynomial $(\tilde{x})_{x \in I'}$ evaluated at $(z_1, \dots, z_m)$ tends to $\tilde{x}_0(z_1, \dots, z_m) \neq 0$. Hence $\big(x(z_1, \dots, z_m)\big)_{x \in I'}$ tends to ∞ in $\mathbb{Q}_p$. Sending the ξ_i's to the z_i's we obtain the desired embedding. (Note that in this case, after p is selected, the specific values of the z_i's are not important.)

Finally, let us turn to the case when $char(k) = p > 0$. The first part of the argument remains valid : R is integral over $S = \mathbb{F}_q[\xi_1, \dots, \xi_m]$ where $\xi_1, \dots, \xi_m$ are algebraically independent over $\mathbb{F}_q$ and this enables to reduce to the case $R = S$. Then we proceed by induction on the transcendence degree m. If $m = 1$, then the assignment $\xi_1 \mapsto \frac{1}{t}$ gives the desired embedding of S into $\mathbb{F}_q((t))$. Let $m \geq 2$ and note that the total degrees of elements of I are necessarily unbounded. From this point the proof works verbatim as in the corresponding paragraph above. □

7.3 Contracting projective transformations

In this section and the next, unless otherwise stated, k is assumed to be a local field, with no assumption on the characteristic.

7.3.1 Proximality and ping-pong

Let us first recall some basic facts about projective transformations on $\mathbb{P}(k^n)$, where k is a local field. For proofs and a detailed (and self-contained) exposition, see [31], Section 3. We let $\|\cdot\|$ be the standard norm on k^n, i.e. the standard Euclidean norm if k is Archimedean and $\|x\| = \max_{1 \leq i \leq n} |x_i|$ where $x = \sum x_i e_i$ when k is non-Archimedean and $(e_1, \dots, e_n)$ is the canonical basis of k^n. This norm extends in the usual way to $\Lambda^2 k^n$. Then we define the *standard metric* on $\mathbb{P}(k^n)$ by

$$d([v], [w]) = \frac{\|v \wedge w\|}{\|v\| \, \|w\|}.$$

With respect to this metric, every projective transformation is bi-Lipschitz on $\mathbb{P}(k^n)$. For $\epsilon \in (0, 1)$, we call a projective transformation $[g] \in \mathrm{PGL}_n(k)$ **ϵ-contracting** if there exist a point $v_g \in \mathbb{P}^{n-1}(k)$, called an attracting point of $[g]$, and a projective hyperplane H_g, called a repelling hyperplane of $[g]$, such that $[g]$ maps the complement of the ϵ-neighborhood of $H_g \subset \mathbb{P}(k^n)$ (the repelling neighborhood of $[g]$) into the ϵ-ball around v_g (the attracting neighborhood of $[g]$). We say that $[g]$ is **ϵ-very contracting** if both $[g]$ and $[g^{-1}]$ are ϵ-contracting. A projective transformation $[g] \in \mathrm{PGL}_n(k)$ is called **(r,ϵ)-proximal** ($r > 2\epsilon > 0$) if it is ϵ-contracting with respect to some attracting point $v_g \in \mathbb{P}(k^n)$ and some repelling hyperplane H_g, such that $d(v_g, H_g) \geq r$. The transformation $[g]$ is called **(r,ϵ)-very proximal** if both $[g]$ and $[g]^{-1}$ are (r,ϵ)-proximal. Finally $[g]$ is simply called **proximal** (resp. **very proximal**) if it is (r,ϵ)-proximal (resp. (r,ϵ)-very proximal) for some $r > 2\epsilon > 0$**.**

The attracting point v_g and repelling hyperplane H_g of an ϵ-contracting transformation are not uniquely defined. Yet, if $[g]$ is proximal we have the following nice choice of v_g and H_g.

Lemma 3.1 *Let $\epsilon \in (0, \frac{1}{4})$. There exist two constants $c_1, c_2 \geq 1$ (depending only on the local field k) such that if $[g]$ is an (r,ϵ)-proximal transformation with $r \geq c_1\epsilon$ then it must fix a unique point $\overline{v}_g$ inside its attracting neighborhood and a unique projective hyperplane $\overline{H}_g$ lying inside its repelling neighborhood. Moreover, if $r \geq c_1\epsilon^{2/3}$, then all positive powers $[g^n]$, $n \geq 1$, are $(r - 2\epsilon, (c_2\epsilon)^{\frac{n}{3}})$-proximal transformations with respect to these same $\overline{v}_g$ and $\overline{H}_g$.*

Let us postpone the proof of this lemma till the next paragraph.

An m-tuple of projective transformations $a_1, \ldots, a_m$ is called a **ping-pong m-tuple** if all the a_i's are (r,ϵ)-very proximal (for some $r > 2\epsilon > 0$) and the attracting points of a_i and a_i^{-1} are at least r-apart from the repelling hyperplanes of a_j and a_j^{-1}, for any $i \neq j$. Ping-pong m-tuples give rise to free groups by the following variant of the *ping-pong lemma* (see [167] 1.1) :

Lemma 3.2 *If $a_1, \ldots, a_m \in PGL_n(k)$ form a ping-pong m-tuple, then $\langle a_1, \ldots, a_m \rangle$ is a free group of rank m.*

A *finite* subset $F \subset \mathrm{PGL}_n(k)$ is called **(m,r)-separating** ($r > 0$, $m \in \mathbb{N}$) if for every choice of $2m$ points $v_1, \ldots, v_{2m}$ in $\mathbb{P}(k^n)$ and $2m$ projective hyperplanes $H_1, \ldots, H_{2m}$ there exists $\gamma \in F$ such that

$$\min_{1 \leq i,j \leq 2m} \{d(\gamma v_i, H_j), d(\gamma^{-1} v_i, H_j)\} > r.$$

A separating set and an ϵ-contracting element for small ϵ are precisely the two ingredients needed to generate a ping-pong m-tuple. This is summarized by the following proposition (see [31] Propositions 3.8 and 3.11).

Proposition 3.3 *Let F be an (m,r)-separating set ($r < 1$, $m \in \mathbb{N}$) in $PGL_n(k)$. Then there is $C \geq 1$ such that for every ϵ, $0 < \epsilon < 1/C$, we have* (*i*) *If* $[g] \in PGL_n(k)$ *is an* ϵ*-contracting transformation, one can find an element* $[f] \in F$*, such that* $[gfg^{-1}]$ *is* $C\epsilon$*-very contracting.*

(*ii*) *If* $a_1, \ldots, a_m \in PGL_n(k)$*, and* γ *is an* ϵ*-very contracting transformation, then there are* $h_1, \ldots, h_m \in F$ *and* $g_1, \ldots, g_m \in F$ *such that*

$$(g_1\gamma a_1 h_1, g_2\gamma a_2 h_2, \ldots, g_m\gamma a_m h_m)$$

forms a ping-pong m*-tuple and hence are free generators of a free group.*

7.3.2 The Cartan decomposition

Now let $\mathbb{H}$ be a Zariski connected reductive k-split algebraic k-group and $H = \mathbb{H}(k)$. Let $\mathbb{T}$ be a maximal k-split torus and $T = \mathbb{T}(k)$. Fix a system Φ of k-roots of $\mathbb{H}$ relative to $\mathbb{T}$ and a basis Δ of simple roots. Let $\mathbb{X}(\mathbb{T})$ be the group of k-rational multiplicative characters of $\mathbb{T}$ and $V' = \mathbb{X}(\mathbb{T}) \otimes_{\mathbb{Z}} \mathbb{R}$ and V the dual vector space of V'. We denote by C^+ the positive Weyl chamber :

$$C^+ = \{v \in V : \forall \alpha \in \Delta,\ \alpha(v) > 0\}.$$

The Weyl group will be denoted by W and is identified with the quotient $N_H(T)/Z_H(T)$ of the normalizer by the centralizer of T in H. Let K be a maximal compact subgroup of H such that $N_K(T)$ contains representatives of every element of W. If k is Archimedean, let A be the subset of T consisting of elements t such that $|\alpha(t)| \geq 1$ for every simple root $\alpha \in \Delta$. And if k is non-Archimedean, let A be the subset of T consisting of elements such that $\alpha(t) = \pi^{-n_\alpha}$ for some $n_\alpha \in \mathbb{N} \cup \{0\}$ for any simple root $\alpha \in \Delta$, where π is a given uniformizer for k (i.e. the valuation of π is 1). Then we have the following *Cartan decomposition* (see Bruhat-Tits [35])

$$H = KAK. \tag{7.2}$$

In this decomposition, the A component is uniquely defined. We can therefore associate to every element $g \in H$ a uniquely defined $a_g \in A$.

Then, in what follows, *we define* $\chi(g)$ to be equal to $\chi(a_g)$ for any character $\chi \in \mathbb{X}(\mathbb{T})$ and element $g \in H$. Although this conflicts with the original meaning of $\chi(g)$ when g belongs to the torus $\mathbb{T}(k)$, we will keep this notation throughout the paper. Thus we always have $|\alpha(g)| \geq 1$ for any simple root α and $g \in H$.

Let us note that the above decomposition (7.2) is no longer true when $\mathbb{H}$ is not assumed to be k-split (see Bruhat-Tits [35] or [136] for the Cartan decomposition in the general case).

If $\mathbb{H} = \mathbb{GL}_n$ and α is the simple root corresponding to the difference of the first two eigenvalues $\lambda_1 - \lambda_2$, then a_g is a diagonal matrix $diag(a_1(g), \ldots, a_n(g))$ where $|\alpha(g)| = |\frac{a_1(g)}{a_2(g)}|$. Then we have the following nice criterion for ϵ-contraction, which justifies the introduction of this notion (see [31] Proposition 3.3).

Lemma 3.4 *Let $\epsilon < \frac{1}{4}$. If $|\frac{a_1(g)}{a_2(g)}| \geq 1/\epsilon^2$, then $[g] \in PGL_n(k)$ is ϵ-contracting on $\mathbb{P}(k^n)$. Conversely, suppose $[g]$ is ϵ-contracting on $\mathbb{P}(k^n)$ and k is non-Archimedean with uniformizer π (resp. Archimedean), then $|\frac{a_1(g)}{a_2(g)}| \geq \frac{|\pi|}{\epsilon^2}$ (resp. $|\frac{a_1(g)}{a_2(g)}| \geq \frac{1}{4\epsilon^2}$).*

The proof of Lemma 3.1, as well as of Proposition 3.3, is based on the latter characterization of ϵ-contraction and on the following crucial lemma (see [31] Lemmas 3.4 and 3.5) :

Lemma 3.5 *Let $r, \epsilon \in (0, 1]$. If $|\frac{a_2(g)}{a_1(g)}| \leq \epsilon^2$, then $[g]$ is ϵ-contracting with respect to the repelling hyperplane*

$$H_g = [span\{k'^{-1}(e_i)\}_{i=2}^n]$$

and the attracting point $v_g = [ke_1]$, where $g = ka_gk'$ is a Cartan decomposition of g. Moreover, $[g]$ is $\frac{\epsilon^2}{r^2}$-Lipschitz outside the r-neighborhood of H_g. Conversely assume that the restriction of $[g]$ to some open set $O \subset \mathbb{P}(k^n)$ is ϵ-Lipschitz, then $|\frac{a_2(g)}{a_1(g)}| \leq 2\epsilon$.

7.3.3 The proof of Lemma 3.1

Given a projective transformation $[h]$ and $\delta > 0$, we say that (H, v) is a δ*-related pair* of a repelling hyperplane and attracting point for $[h]$, if $[h]$ maps the complementary of the δ-neighborhood of H inside the δ-ball around v.

The attracting point and repelling hyperplane of an δ-contracting transformation $[h]$ are not uniquely defined. However, note that if $\delta < \frac{1}{4}$ then for any two δ-related pairs of $[h]$ $(H_h^i, v_h^i),\ i = 1, 2$, we have $d(v_h^1, v_h^2) < 2\delta$. Indeed, since $\delta < \frac{1}{4}$, the union of the δ -neighborhoods of the H_h^i's does not cover $\mathbb{P}(k^n)$. Let $p \in \mathbb{P}(k^n)$ be a point lying outside this union, then $d([h]p, v_h^i) < \delta$ for $i = 1, 2$.

Now consider two δ-related pairs $(H_h^i, v_h^i),\ i = 1, 2$ of some projective transformation $[h]$, satisfying $d(v_h^1, H_h^1) \geq r$ and no further assumption on the pair (H_h^2, v_h^2). Suppose that $1 \geq r > 4\delta$. Then we claim that $\mathrm{Hd}(H_h^1, H_h^2) \leq 2\delta$, where Hd denotes the standard distance between hyperplanes, i.e. the Hausdorff distance. (Note that $\mathrm{Hd}(H^1, H^2) = max_{x \in H^1}\{\frac{|f_2(x)|}{\|x\|}\}$ where f_2 is the unique (up to sign) norm one functional whose kernel is the hyperplane H_2 (for details see [31] section 3).) To see this, notice that if $\mathrm{Hd}(H_h^1, H_h^2)$ were greater than 2δ then any projective hyperplane H would contain a point outside the δ-neighborhood of either H_h^1 or H_h^2. Such a point is mapped under $[h]$ to the δ-ball around either v_h^1 or v_h^2, hence to the 3δ-ball around v_h^1. This in particular applies to the hyperplane $[h^{-1}]H_h^1$. A contradiction to the assumption $d(H_h^1, v_h^1) > 4\delta$. We also conclude that when $r > 8\delta$, then for any two δ-related pairs (H^i, v^i) $i = 1, 2$ of $[h]$, we have $d(v^i, H^j) > \frac{r}{2}$ for all $i, j \in \{1, 2\}$.

Let us now fix an arbitrary ϵ-related pair (H, v) of the (r, ϵ)-proximal transformation $[g]$ from the statement of Lemma 3.1. Let also (H_g, v_g) be the hyperplane and point introduced in Lemma 3.5. From Lemmas 3.4 and 3.5, we see that the pair (H_g, v_g) is a $C\epsilon$-related pair for $[g]$ for some constant $C \geq 1$ depending only on k. Assume

$d(v,H) \geq r > 8C\epsilon$. Then it follows from the above that the ϵ-ball around v is mapped into itself under $[g]$, and that $d(v,H_g) > \frac{r}{2}$. From Lemma 3.5, we obtain that $[g]$ is $(\frac{4C\epsilon}{r})^2$-Lipschitz in this ball, and hence $[g^n]$ is $(\frac{4C\epsilon}{r})^{2n}$-Lipschitz there. Hence $[g]$ has a unique fixed point $\overline{v}_g$ in this ball which is the desired attracting point for all the powers of $[g]$. Note that $d(v,\overline{v}_g) \leq \epsilon$.

Since $[g^n]$ is $(\frac{4C\epsilon}{r})^{2n}$-Lipschitz on some open set, it follows from Lemma 3.5 that $|\frac{a_2(g^n)}{a_1(g^n)}| \leq 2(\frac{4C\epsilon}{r})^{2n}$, and from Lemma 3.4 that $[g^n]$ is $2(\frac{4C\epsilon}{r})^n$-contracting. Moreover, it is now easy to see that if $r > (4C)^2\epsilon$, then for every $2(\frac{4C\epsilon}{r})^n$-related pair (H_n,v_n) for $[g^n]$ $n \geq 2$, we have $d(\overline{v}_g,v_n) \leq 4(\frac{4C\epsilon}{r})^n$. (To see this apply $[g^n]$ to some point of the ϵ-ball around v which lies outside the $2(\frac{4C\epsilon}{r})^n$-neighborhood of H_n). Therefore $(H_n,\overline{v}_g)$ is a $6(\frac{4C\epsilon}{r})^n$-related pair for $[g^n]$, $n \geq 2$.

We shall now show that the ϵ-neighborhood of H contains a unique $[g]$-invariant hyperplane which can be used as a common repelling hyperplane for all the powers of $[g]$. The set $\mathcal{F}$ of all projective points at distance at most ϵ from H is mapped into itself under $[g^{-1}]$. Similarly the set $\mathfrak{H}$ of all projective hyperplanes which are contained in $\mathcal{F}$ is mapped into itself under $[g^{-1}]$. Both sets $\mathcal{F}$, and $\mathfrak{H}$ are compact with respect to the corresponding Grassmann topologies. The intersection $\mathcal{F}_\infty = \cap[g^{-n}]\mathcal{F}$ is therefore non empty and contains some hyperplane $\overline{H}_g$ which corresponds to any point of the intersection $\cap[g^{-n}]\mathfrak{H}$. We claim that $\mathcal{F}_\infty = \overline{H}_g$. Indeed, the set $\mathcal{F}_\infty$ is invariant under $[g^{-1}]$ and hence under $[g]$ and $[g^n]$. Since $(H_n,\overline{v}_g)$ is a $6(\frac{4C\epsilon}{r})^n$-related pair for $[g^n]$, $n \geq 2$, and since $\overline{v}_g$ is "far" (at least $r - 2\epsilon$ away) from the invariant set $\mathcal{F}_\infty$, it follows that for large n, $\mathcal{F}_\infty$ must lie inside the $6(\frac{4C\epsilon}{r})^n$-neighborhood of H_n. Since $\mathcal{F}_\infty$ contains a hyperplane, and since it is arbitrarily close to a hyperplane, it must coincide with a hyperplane. Hence $\mathcal{F}_\infty = \overline{H}_g$. It follows that $(\overline{H}_g,\overline{v}_g)$ is a $12(\frac{4C\epsilon}{r})^n$-related pair for $[g^n]$ for any large enough n. Note that then $d(\overline{v}_g,\overline{H}_g) > r - 2\epsilon$, since $d(\overline{v}_g,v) \leq \epsilon$ and $\mathrm{Hd}(\overline{H}_g,H) \leq \epsilon$. This proves existence and uniqueness of $(\overline{H}_g,\overline{v}_g)$ as soon as $r > c_1\epsilon$ where $c_1 \geq (4C)^2 + 8C$.

If we assume further that $r^3 \geq 12(4C\epsilon)^2$, then $\mathcal{F}_\infty$ lies inside the $6(\frac{4C\epsilon}{r})^n$-neighborhood of H_n as soon as $n \geq 2$. Then $(\overline{H}_g,\overline{v}_g)$ is a $12(\frac{4C\epsilon}{r})^n$-related pair for $[g^n]$, hence a $(c_2\epsilon)^{n/3}$-related pair for $[g^n]$ whenever $n \geq 1$, where $c_2 \geq 1$ is a constant easily computable in terms of C. This finishes the proof of the lemma. □

In what follows, whenever we add the article *the* to an attracting point and repelling hyperplane of a proximal transformation $[g]$, we shall mean these fixed point $\overline{v}_g$ and fixed hyperplane $\overline{H}_g$ obtained in Lemma 3.1.

7.3.4 The case of general semisimple group

Now let us assume that $\mathbb{H}$ is a Zariski connected semisimple k-algebraic group, and let (ρ, V_ρ) be a finite dimensional k-rational representation of $\mathbb{H}$ with highest weight χ_ρ. Let Θ_ρ be the set of simple roots α such that χ_ρ/α is again a non-trivial weight of ρ

$$\Theta_\rho = \{\alpha \in \Delta : \chi_\rho/\alpha \text{ is a weight of } \rho\}.$$

It turns out that Θ_ρ is precisely the set of simple roots α such that the associated fundamental weight π_α appears in the decomposition of χ_ρ as a sum of fundamental weights. Suppose that the weight space V_{χ_ρ} corresponding to χ_ρ has dimension 1, then we have the following lemma.

Lemma 3.6 *There are positive constants $C_1 \leq 1 \leq C_2$, such that for any $\epsilon \in (0,1)$ and any $g \in \mathbb{H}(k)$, if $|\alpha(g)| > \frac{C_2}{\epsilon^2}$ for all $\alpha \in \Theta_\rho$ then the projective transformation $[\rho(g)] \in PGL(V_\rho)$ is ϵ-contracting, and conversely, if $[\rho(g)]$ is ϵ-contracting, then $|\alpha(g)| > \frac{C_1}{\epsilon^2}$ for all $\alpha \in \Theta_\rho$.*

Proof: Let $V_\rho = \bigoplus V_\chi$ be the decomposition of V_ρ into a direct sum of weight spaces. Let us fix a basis $(e_1, \ldots, e_n)$ of V_ρ compatible with this decomposition and such that $V_{\chi_\rho} = ke_1$. We then identify V_ρ with k^n via this choice of basis. Let $g = k_1 a_g k_2$ be a Cartan decomposition of g in H. We have $\rho(g) = \rho(k_1)\rho(a_g)\rho(k_2) \in \rho(K)D\rho(K)$ where $D \subset SL_n(k)$ is the set of diagonal matrices. Since $\rho(K)$ is compact, there exists a positive constant C such that if $[\rho(g)]$ is ϵ-contracting then $[\rho(a_g)]$ is $C\epsilon$-contracting, and conversely if $[\rho(a_g)]$ is ϵ-contracting then $[\rho(g)]$ is $C\epsilon$-contracting. Therefore, it is equivalent to prove the lemma for $\rho(a_g)$ instead of $\rho(g)$. Now the coefficient $|a_1(\rho(a_g))|$ in the Cartan decomposition on $SL_n(k)$ equals $\max_\chi |\chi(a_g)| = |\chi_\rho(g)|$, and the coefficient $|a_2(\rho(a_g))|$ is the second highest diagonal coefficient and hence of the form $|\chi_\rho(a_g)/\alpha(a_g)|$ where α is some simple root. Now the conclusion follows from Lemma 3.4. $\square$

7.4 Irreducible representations of non-Zariski connected algebraic groups

In the process of constructing dense free groups, we need to find some suitable linear representation of the group Γ we started with. In general, the Zariski closure of Γ may not be Zariski connected, and yet we cannot pass to a subgroup of finite index in Γ in Theorem 1.2. Therefore we will need to consider representations of non Zariski connected groups.

Let $\mathbb{H}^\circ$ be a connected semisimple k-split algebraic k-group. The group $Aut_k(\mathbb{H}^\circ)$ of k-automorphisms of $\mathbb{H}^\circ$ acts naturally on the characters $\mathbb{X}(\mathbb{T})$ of a maximal split torus $\mathbb{T}$. Indeed, for every $\sigma \in Aut_k(\mathbb{H}^\circ)$, the torus $\sigma(T)$ is conjugate to $T = \mathbb{T}(k)$ by some element $g \in H = \mathbb{H}(k)$ and we can define the character $\sigma(\chi)$ by $\sigma(\chi)(t) = \chi(g^{-1}\sigma(t)g)$. This is not well defined, since the choice of g is not unique (it is up to multiplication by an element of the normalizer $N_H(T)$). But if we require $\sigma(\chi)$ to lie in the same Weyl chamber as χ, then this determines g up to multiplication by an element from the centralizer $Z_H(T)$, hence it determines $\sigma(\chi)$ uniquely. Note also that every σ sends roots to roots and simple roots to simple roots.

In fact, what we need are representations of algebraic groups whose restriction to the connected component is irreducible. As explained below, it turns out that an irreducible representation ρ of a connected semisimple algebraic group $\mathbb{H}^\circ$ extends to the full group $\mathbb{H}$ if and only if its highest weight is invariant under the action of $\mathbb{H}$ by conjugation.

We thus have to face the problem of finding elements in $\mathbb{H}^\circ(k)\Gamma$ which are ε-contracting under such a representation ρ. By Lemma 3.6 this amounts to finding elements h such that $\alpha(h)$ is large for all simple roots α in the set Θ_ρ defined in Paragraph 7.3.4. As will be explained below, we can find such a representation ρ such that all simple roots belonging to Θ_ρ are images by some outer automorphisms σ's of $\mathbb{H}^\circ$ (coming from conjugation by an element of $\mathbb{H}$) of a single simple root α. But $\sigma(\alpha)(h)$ and $\alpha(\sigma(h))$ are comparable. The idea of the proof below is then to find elements h in $\mathbb{H}^\circ(k)$ such that all relevant $\alpha(\sigma(h))$'s are large. But, according to the converse statement in Lemma 3.6, this amounts to finding elements h such that all relevant $\sigma(h)$'s are ε-contracting under a representation ρ_α such that $\Theta_{\rho_\alpha} = \{\alpha\}$. This is the content of the forthcoming proposition.

Before stating the proposition, let us note that, $\mathbb{H}^\circ$ being k-split, to every simple root $\alpha \in \Delta$ corresponds an irreducible k-rational representation of $\mathbb{H}^\circ(k)$ whose highest weight χ_{ρ_α} is the fundamental weight π_α associated to α and has multiplicity one. In this case the set Θ_{ρ_α} defined in Paragraph 7.3.4 is reduced to the singleton $\{\alpha\}$.

Proposition 4.1 *Let α be a simple root. Let I be a subset of $\mathbb{H}^\circ(k)$ such that $\{|\alpha(g)|\}_{g\in I}$ is unbounded in $\mathbb{R}$. Let $\Omega \subset \mathbb{H}^\circ(k)$ be a Zariski dense subset. Let $\sigma_1, \ldots, \sigma_m$ be algebraic k-automorphisms of $\mathbb{H}^\circ$. Then for any arbitrary large $M > 0$, there exists an element $h \in \mathbb{H}^\circ(k)$ of the form $h = f_1\sigma_1^{-1}(g)\ldots f_m\sigma_m^{-1}(g)$ where $g \in I$ and the f_i's belong to Ω, such that $|\sigma_i(\alpha)(h)| > M$ for all $1 \le i \le m$.*

Proof: Let $\epsilon \in (0,1)$ and $g \in I$ such that $|\alpha(g)| \ge \frac{1}{\epsilon^2}$. Let (ρ_α, V) be the irreducible representation of $\mathbb{H}^\circ(k)$ corresponding to α as described above. Consider the weight space decomposition $V_{\rho_\alpha} = \bigoplus V_\chi$ and fix a basis $(e_1, \ldots, e_n)$ of $V = V_{\rho_\alpha}$ compatible with this decomposition and such that $V_{\chi_{\rho_\alpha}} = ke_1$. We then identify V with k^n via this choice of basis, and in particular, endow $\mathbb{P}(V)$ with the standard metric defined in the previous section. It follows from Lemma 3.6 above that $[\rho_\alpha(g)]$ is ϵC-contracting on $\mathbb{P}(V)$ for some constant $C \ge 1$ depending only on ρ_α. Now from Lemma 3.5, there exists for any $x \in \mathbb{H}^\circ(k)$ a point $u_x \in \mathbb{P}(V)$ such that $[\rho_\alpha(x)]$ is 2-Lipschitz over some open neighborhood of u_x. Similarly there exists a projective hyperplane H_x such that $[\rho_\alpha(x)]$ is $\frac{1}{r^2}$-Lipschitz outside the r-neighborhood of H_x. Moreover, combining Lemmas 3.4 and 3.5 (and up to changing C if necessary to a larger constant depending this time only on k), we see that $[\rho_\alpha(g)]$ is $\frac{\epsilon^2 C^2}{r^2}$-Lipschitz outside the r-neighborhood of the repelling hyperplane H_g defined in Lemma 3.5. We pick u_g outside this r-neighborhood.

By modifying slightly the definition of a finite (m,r)-separating set (see above Paragraph 7.3.1), we can say that a finite subset F of $\mathbb{H}^\circ(k)$ is an (m,r)-separating set *with respect to* ρ_α and $\sigma_1, \ldots, \sigma_m$ if for every choice of m points $v_1, \ldots, v_m$ in $\mathbb{P}(V)$ and m projective hyperplanes $H_1, \ldots, H_m$ there exists $\gamma \in F$ such that

$$\min_{1\le i,j,k\le m,} d(\rho_\alpha(\sigma_k(\gamma))v_i, H_j) > r > 0.$$

Claim : The Zariski dense subset Ω contains a finite (m,r)-separating set with respect to ρ_α and $\sigma_1,\dots,\sigma_m$, for some positive number r.

Proof of claim : For $\gamma\in\Omega$, we let M_γ be the set of all tuples $(v_i,H_i)_{1\le i\le m}$ such that there exists some i,j and l for which $\rho_\alpha(\sigma_l(\gamma))v_i\in H_j$. Now $\bigcap_{\gamma\in\Omega}M_\gamma$ is empty, for otherwise there would be points $v_1,\dots,v_m$ in $P(V)$ and projective hyperplanes $H_1,\dots,H_m$ such that Ω is included in the union of the closed algebraic k-subvarieties $\{x\in\mathbb{H}^\circ(k),\ \rho_\alpha(\sigma_l(x))v_i\in H_j\}$ where i,j and l range between 1 and m. But, by irreducibility of ρ_α each of these subvarieties is proper, and this would contradict the Zariski density of Ω or the Zariski connectedness of $\mathbb{H}^\circ$. Now, since each M_γ is compact in the appropriate product of Grassmannians, it follows that for some finite subset $F\subset\Omega$, $\bigcap_{\gamma\in F}M_\gamma=\emptyset$. Finally, since $\max_{\gamma\in F}\min_{1\le i,j,l\le m}d(\rho_\alpha(\sigma_l(\gamma)v_i,H_j)$ depends continuously on $(v_i,H_i)_{i=1}^m$ and never vanishes, it must attain a positive minimum r, by compactness of the set of all tuples $(v_i,H_i)_{i=1}^m$ in

$$\big(\mathbb{P}(V)\times\mathbb{G}r_{\dim(V)-1}(V)\big)^{2m}.$$

Therefore F is the desired (m,r)-separating set.

Up to taking a bigger constant C, we can assume that C is larger than the bi-Lipschitz constant of every $\rho_\alpha(x)$ on $\mathbb{P}(k^n)$ when x ranges over the finite set $\{\sigma_k(f),f\in F,1\le k\le m\}$.

Now let us explain how to find the element $h=f_m\sigma_m^{-1}(g)\dots f_1\sigma_1^{-1}(g)$ we are looking for. We shall choose the f_j's recursively, starting from $j=1$, in such a way that all the elements $\sigma_i(h)$, $1\le i\le m$, will be contracting. Write

$$\begin{aligned}\sigma_i(h)\ &=\ \sigma_i(f_m\sigma_m^{-1}(g)\dots f_1\sigma_1^{-1}(g))=\\ \big(\sigma_i(f_m)\sigma_i\sigma_m^{-1}(g)\cdot\ldots\cdot\sigma_i(f_i)\big)\quad g&\quad\big(\sigma_i(f_{i-1})\sigma_i\sigma_{i-1}(g)\cdot\ldots\cdot\sigma_i(f_1)\sigma_i\sigma_1^{-1}(g)\big).\end{aligned}$$

In order to make $\sigma_i(h)$ contracting, we shall require that :

– For $m\ge i\ge 2$, $\sigma_i(f_{i-1})$ takes the image under $\sigma_i\sigma_{i-1}(g)\cdot\ldots\cdot\sigma_i(f_1)\sigma_i\sigma_1^{-1}(g)$ of some open set on which $\sigma_i\sigma_{i-1}(g)\cdot\ldots\cdot\sigma_i(f_1)\sigma_i\sigma_1^{-1}(g)$ is 2-Lipschitz, e.g. a small neighborhood of the point

$$u_i:=\big(\sigma_i\sigma_{i-1}(g)\cdot\ldots\cdot\sigma_i(f_1)\sigma_i\sigma_1^{-1}(g)\big)(u_{\sigma_i\sigma_{i-1}(g)\cdot\ldots\cdot\sigma_i(f_1)\sigma_i\sigma_1^{-1}(g)})$$

at least r apart from the hyperplane H_g, and

– For $m>j\ge i$, $\sigma_i(f_j)$ takes the image of u_i under $\big(\sigma_i\sigma_j^{-1}(g)\cdot\ldots\cdot\sigma_i(f_i)\big)g$ at least r apart from the hyperplane $H_{\sigma_i\sigma_{j+1}^{-1}(g)}$ of $\sigma_i\sigma_{j+1}^{-1}(g)$ (i.e. of the next element on the left in the expression of $\sigma_i(h)$).

Assembling the conditions on each f_i we see that there are $\le m$ points that the $\sigma_j(f_i)$'s, $1\le j\le m$ should send r apart from $\le m$ projective hyperplanes.

This appropriate choice of $f_1,\dots,f_m$ in F forces each of $\sigma_1(h),\dots,\sigma_m(h)$ to be $\frac{2C^{m+2}\epsilon^2}{r^{2m}}$-Lipschitz in some open subset of $\mathbb{P}(V)$. Lemma 3.5 now implies that $\sigma_1(h),\dots,\sigma_m(h)$ are $C_0\epsilon$-contracting on $\mathbb{P}(V)$ for some constant C_0 depending only on (ρ,V).

Moreover $h \in K a_h K$ and each of the $\sigma_i(K)$ is compact, we conclude that $\sigma_1(a_h), \ldots, \sigma_m(a_h)$ are also $C_1\epsilon$-contracting on $\mathbb{P}(V)$ for some constant C_1. But for every σ_i there exists an element $b_i \in \mathbb{H}^\circ(k)$ such that $\sigma_i(T) = b_i T b_i^{-1}$ and $\sigma_i(\alpha)(t) = \alpha(b_i^{-1}\sigma_i(t)b_i)$ for every element t in the positive Weyl chamber of the maximal k-split torus $T = \mathbb{T}(k)$. Up to taking a larger constant C_1 (depending on the b_i's) we therefore obtain that $b_1^{-1}\sigma_1(a_h)b_1, \ldots, b_m^{-1}\sigma_m(a_h)b_m$ are also $C_1\epsilon$-contracting on $\mathbb{P}(V)$ via the representation ρ_α. Finally Lemma 3.6 yields the conclusion that $|\sigma_i(\alpha)(h)| = |\alpha(b_i^{-1}\sigma_i(a_h)b_i)| \geq \frac{1}{C_2\epsilon^2}$ for some other positive constant C_2. Since ϵ can be chosen arbitrarily small, we are done. $\square$

Now let $\mathbb{H}$ be an arbitrary algebraic k-group, whose identity connected component $\mathbb{H}^\circ$ is semisimple. Let us fix a system Σ of k-roots for $\mathbb{H}^\circ$ and a simple root α. For every element g in $\mathbb{H}(k)$ let σ_g be the automorphism of $\mathbb{H}^\circ(k)$ which is induced by g under conjugation, and let $\mathcal{S}$ be the group of all such automorphisms. As was described above, $\mathcal{S}$ acts naturally on the set Δ of simple roots. Let $\mathcal{S}\cdot\alpha = \{\alpha_1, \ldots, \alpha_p\}$ be the orbit of α under this action. Suppose $I \subset \mathbb{H}^\circ(k)$ satisfies the conclusion of the last proposition for $\mathcal{S}\cdot\alpha$, that is for any $\epsilon > 0$, there exists $g \in I$ such that $|\alpha_i(g)| > 1/\epsilon^2$ for all $i = 1, \ldots, p$. Then the following proposition shows that under some suitable irreducible projective representation of the full group $\mathbb{H}(k)$, for arbitrary small ϵ, some elements of I act as ϵ-contracting transformations.

Proposition 4.2 *Let $I \subset \mathbb{H}^\circ(k)$ be as above. Then there exists a finite extension K of k, $[K : k] < \infty$, and a non-trivial finite dimensional irreducible K-rational representation of $\mathbb{H}^\circ$ into a K-vector space V which extends to an irreducible projective representation $\rho : \mathbb{H}(K) \to PGL(V)$, satisfying the following property : for every positive $\epsilon > 0$ there exists $\gamma_\epsilon \in I$ such that $\rho(\gamma_\epsilon)$ is an ϵ-contracting projective transformation of $\mathbb{P}(V)$.*

Proof: Up to taking a finite extension of k, we can assume that $\mathbb{H}^\circ$ is k-split. Let (ρ, V) be an irreducible k-rational representation of $\mathbb{H}^\circ$ whose highest weight χ_ρ is a multiple of $\alpha_1 + \ldots + \alpha_p$ and such that the highest weight space V_{χ_ρ} has dimension 1 over k. Burnside's theorem implies that, up to passing to a finite extension of k, we can also assume that the group algebra $k[\mathbb{H}^\circ(k)]$ is mapped under ρ to the full algebra of endomorphisms of V, i.e. $End_k(V)$. For a k-automorphism σ of $\mathbb{H}^\circ$ let $\sigma(\rho)$ be the representation of $\mathbb{H}^\circ$ given on V by $\sigma(\rho)(g) = \rho(\sigma(g))$. It is a k-rational irreducible representation of $\mathbb{H}^\circ$ whose highest weight is precisely $\sigma(\chi_\rho)$. But $\chi_\rho = d(\alpha_1 + \ldots + \alpha_p)$ for some $d \in \mathbb{N}$, and is invariant under the action of $\mathcal{S}$. Hence for any $\sigma \in \mathcal{S}$, $\sigma(\rho)$ is equivalent to ρ. So there must exists a linear automorphism $J_\sigma \in$GL(V) such that $\sigma(\rho)(h) = J_\sigma \rho(h) J_\sigma^{-1}$ for all $h \in \mathbb{H}^\circ(k)$. Now set $\widetilde{\rho}(g) = [\rho(g)] \in PGL(V)$ if $g \in \mathbb{H}^\circ(k)$ and $\widetilde{\rho}(g) = [J_{\sigma_g}] \in$PGL$(V)$ otherwise. Since the $\rho(g)$'s when g ranges over $\mathbb{H}^\circ(k)$ generate the whole of $End_k(V)$, it follows from Schur's lemma that $\widetilde{\rho}$ is a well defined projective representation of the whole of $\mathbb{H}(k)$. Now the set Θ_ρ of simple roots α such that χ_ρ/α is a non-trivial weight of ρ is precisely $\{\alpha_1, \ldots, \alpha_p\}$. Hence if $\gamma_\epsilon \in I$ satisfies $|\alpha_i(\gamma_\epsilon)| > \frac{1}{\epsilon^2}$ for all $i = 1, \ldots, p$, then we have by Lemma 3.6 that $\widetilde{\rho}(\gamma_\epsilon)$ is $C_2\epsilon$-contracting on $\mathbb{P}(V)$ for some constant C_2 independent of ϵ. $\square$

We can now state and prove the main result of this paragraph, and the only one which will be used in the sequel. Let here K be an arbitrary field and $\mathbb{H}$ an algebraic K-group such that its connected component $\mathbb{H}^\circ$ is semisimple and non-trivial. Let $\mathbb{H} \hookrightarrow \mathbb{GL}_d$ be some faithful K-rational representation of $\mathbb{H}$ and let R be a finitely generated subring of K. We shall denote by $\mathbb{H}(R)$ (resp. $\mathbb{H}^\circ(R)$) the subset of points of $\mathbb{H}(K)$ (resp. $\mathbb{H}^\circ(K)$) which are mapped into $\mathbb{GL}_d(R)$ under the latter embedding.

Theorem 4.3 *Suppose K is finitely generated and $\Omega_0 \subset \mathbb{H}^\circ(R)$ is a Zariski dense subset of $\mathbb{H}^\circ$ with $\Omega_0 = \Omega_0^{-1}$. Suppose $\{g_1, \ldots, g_m\}$ is a finite subset of $\mathbb{H}(K)$ exhausting all cosets of $\mathbb{H}^\circ$ in $\mathbb{H}$ and let $\Omega = g_1\Omega_0 g_1^{-1} \cup \ldots \cup g_m\Omega_0 g_m^{-1}$. Then we can find a number $r > 0$, a local field k, an embedding $K \hookrightarrow k$, and an irreducible projective representation $\rho : \mathbb{H}(k) \to PGL_d(k)$ defined over k with the following property. If $\epsilon \in (0, \frac{r}{2})$ and $a_1, \ldots, a_n \in \mathbb{H}(K)$ are n arbitrary points ($n \in \mathbb{N}$), then there exist n elements $x_1, \ldots, x_n$ with $x_i \in \Omega^{4m+2} a_i \Omega$ such that the $\rho(x_i)$'s form a ping-pong n-tuple of (r, ϵ)-very proximal transformations on $\mathbb{P}(k^d)$, and in particular are generators of a free group F_n.*

Proof: Up to enlarging the subring R if necessary, we can assume that K is the field of fractions of R. We shall make use of Lemma 2.1. Since Ω_0 is infinite, we can apply this lemma and obtain an embedding of K into a local field k such that Ω_0 becomes an unbounded set in $\mathbb{H}(k)$. Up to enlarging k if necessary we can assume that $\mathbb{H}^\circ(k)$ is k-split. We fix a maximal k-split torus and a system of k-roots with a base Δ of simple roots. Then, in the corresponding Cartan decomposition of $\mathbb{H}(k)$ the elements of Ω_0 have unbounded A component (see Paragraph 7.3.2). Therefore, there exists a simple root α such that the set $\{|\alpha(g)|\}_{g \in \Omega_0}$ is unbounded in $\mathbb{R}$. Let σ_{g_i} be the automorphism of $\mathbb{H}^\circ(k)$ given by the conjugation by g_i. The orbit of α under the group generated by the σ_{g_i}'s is denoted by $\{\alpha_1, \ldots, \alpha_p\}$. Now it follows from Proposition 4.1 that for every $\epsilon > 0$ there exists an element $h \in \Omega^{2p}$ such that $|\alpha_i(h)| > 1/\epsilon^2$ for every $i = 1, \ldots, p$. We are now in a position to apply the last Proposition 4.2 and obtain (up to taking a finite extension of k if necessary) an irreducible projective representation $\rho : \mathbb{H}(k) \to \mathrm{PGL}(V)$, such that the restriction of ρ to $\mathbb{H}^\circ(k)$ is also irreducible and with the following property : for every positive $\epsilon > 0$ there exists $h_\epsilon \in \Omega^{2p}$ such that $\rho(h_\epsilon)$ is an ϵ-contracting projective transformation of $\mathbb{P}(V)$. Moreover, since $\rho_{|\mathbb{H}^\circ}$ is also irreducible and Ω_0 is Zariski dense in $\mathbb{H}^\circ$ we can find an (n, r)-separating set with respect to $\rho_{|\mathbb{H}^\circ}$ for some $r > 0$ (for this terminology, see definitions in Paragraph 7.3.1). This follows from the proof of the claim in Proposition 4.1 above (see also Lemma 4.3. in [31]). By Proposition 3.3 (i) above, we obtain for every small $\epsilon > 0$ an ϵ-very contracting element γ_ϵ in $h_\epsilon \Omega_0 h_\epsilon^{-1} \subset \Omega^{4p+1}$. Similarly, statement ($ii$) of the same Proposition gives elements $f_1, \ldots, f_n \in \Omega_0$ and $f_1', \ldots, f_n' \in \Omega_0$ such that, for ϵ small enough, $(x_1, \ldots, x_n) = (f_1'\gamma_\epsilon a_1 f_1, \ldots, f_n'\gamma_\epsilon a_n f_n)$ form under ρ a ping-pong n-tuple of proximal transformations on $\mathbb{P}(V)$. Then each x_i lies in $\Omega^{4p+2} a_i \Omega$ and together the x_i's form generators of a free group F_n of rank n. $\square$

7.4.1 Further remarks

For further use in later sections we shall state two more facts. Let $\Gamma \subset \mathbb{G}(K)$ be a Zariski dense subgroup of some algebraic group $\mathbb{G}$. Suppose Γ is not virtually solvable and let $\Delta \leq \Gamma$ be a subgroup of finite index. Taking the quotient by the solvable radical of $\mathbb{G}^\circ$, we obtain a homomorphism π of Γ into an algebraic group $\mathbb{H}$ whose connected component is semisimple. Let $g_1, \ldots, g_m \in \Gamma$ be representatives of all different cosets of Δ in Γ. Then $\Omega_0 = \pi(\cap_{i=1}^m g_i \Delta g_i^{-1}) \cap \mathbb{H}^0$ is clearly Zariski dense in $\mathbb{H}^0$ and satisfies the conditions of Theorem 4.3. Hence taking $a_i = \pi(g_i)$ in the theorem, we obtain :

Corollary 4.4 *Let Γ be a linear group which is not virtually solvable, and let $\Delta \subset \Gamma$ be a subgroup of finite index. Then there is some choice of coset representatives for Γ/Δ which generate a free group.*

The following lemma will be useful when dealing with the non-Archimedean case.

Lemma 4.5 *Let k be a non-Archimedean local field. Let $\Gamma \leq GL_n(k)$ be a linear group over k which contains no open solvable subgroup. Then there exists a homomorphism ρ from Γ into a k-algebraic group $\mathbb{H}$ such that the Zariski closure of the image of any open subgroup of Γ contains the connected component of identity $\mathbb{H}^\circ$. Moreover, we can take $\rho : \Gamma \to \mathbb{H}(k)$ to be continuous in the topology induced by k, and we can find $\mathbb{H}$ such that $\mathbb{H}^\circ$ is semisimple and* $\dim(\mathbb{H}^\circ) \leq \dim \overline{\Gamma}^z$.

Proof: Let U_i be a decreasing sequence of open subgroups in $\mathrm{GL}_n(k)$ forming a base of identity neighborhoods. Consider the decreasing sequence of algebraic groups $\overline{\Gamma \cap U_i}^z$. This sequence must stabilize after a finite step s. The limiting group $\mathbb{G} = \overline{\Gamma \cap U_s}^z$ must be Zariski connected. Indeed, the intersection of $\Gamma \cap U_s$ with the Zariski connected component of identity of $\mathbb{G}$ is a relatively open subgroup and contains $\Gamma \cap U_t$ for some large t. If $\mathbb{G}$ were not Zariski connected, then $\overline{\Gamma \cap U_t}^z$ would be a smaller algebraic group. Moreover, from the assumption on Γ, we get that $\mathbb{G}$ is not solvable.

Note that the conjugation by an element of Γ fixes $\mathbb{G}$, since $\gamma U_i \gamma^{-1} \cap U_i$ is again open if $\gamma \in \Gamma$ and hence contains some U_j. Since the solvable radical $Rad(\mathbb{G})$ of $\mathbb{G}$ is a characteristic subgroup of $\mathbb{G}$, it is also fixed under conjugation by elements of Γ. We thus obtain a homomorphism ρ from Γ to the k-points of the group of k-automorphisms $\mathbb{H} = Aut(\mathbb{S})$ of the Zariski connected semisimple k-group $\mathbb{S} = \mathbb{G}/Rad(\mathbb{G})$. This homomorphism is clearly continuous. Since the image of $\Gamma \cap U$ is Zariski dense in $\mathbb{G}$ for all open $U \subset \mathrm{GL}_n(k)$, $\Gamma \cap U$ is mapped under this homomorphism to a Zariski dense subgroup of the group of inner automorphisms $Int(\mathbb{S})$ of $\mathbb{S}$. But $Int(\mathbb{S})$ is a semisimple algebraic k-group which is precisely the Zariski connected component of identity of $\mathbb{H} = Aut(\mathbb{S})$ (see for example [26], 14.9). Finally, it is clear from the construction that $\dim \mathbb{H} \leq \dim \overline{\Gamma}^z$. $\square$

7.5 The proof of Theorem 1.2 in the finitely generated case

In this section we prove our main result, Theorem 1.2, in the case when Γ is finitely generated. We obtain in fact a more precise result which yields some control on the number of generators required for the free group.

Theorem 5.1 *Let $\Gamma \leq GL_n(k)$ be a finitely generated linear group over a local field k. Suppose Γ contains no solvable open subgroup. Then, there is a constant $h(\Gamma) \in \mathbb{N}$ such that for any integer $r \geq h(\Gamma)$, Γ contains a dense free subgroup of rank r. Moreover, if $char(k) = 0$ we can take $h(\Gamma) = d(\Gamma)$ (i.e. the minimal size of a generating set for Γ), while if $char(k) > 0$ we can take $h(\Gamma) = d(\Gamma) + n^2$.*

In the following paragraphs we split the proof to three cases (Archimedean, non-Archimedean of characteristic zero, and positive characteristic) which have to be dealt with independently.

7.5.1 The Archimedean case

Consider first the case $k = \mathbb{R}$ or $\mathbb{C}$. Let G be the linear Lie group $G = \overline{\Gamma}$, and let G° be the connected component of the identity in G. The condition "Γ contains no open solvable subgroup" means simply "G° is not solvable". Note also that $d(G/G^\circ) \leq d(\Gamma) < \infty$.

Define inductively $G^\circ_0 = G^\circ$ and $G^\circ_{n+1} = \overline{[G^\circ_n, G^\circ_n]}$. This sequence stabilizes after some finite step t to a normal topologically perfect subgroup $H := G^\circ_t$ (i.e. the commutator group $[H, H]$ is dense in H). As was shown in [31] Theorem 2.1, any topologically perfect group H contains a finite set of elements $\{h_1, \ldots, h_l\}$, $l \leq \dim(H)$ and a relatively open identity neighborhood $V \subset H$ such that, for any selection of points $x_i \in V h_i V$, the group $\langle x_1, \ldots, x_l\rangle$ is dense in H. Moreover, H is clearly a characteristic subgroup of G°, hence it is normal in G. It is also clear from the definition of H that if Γ is a dense subgroup of G then $\Gamma \cap H$ is dense in H.

Let $r \geq d(\Gamma)$, and let $\{\gamma_1, \ldots, \gamma_r\}$ be a generating set for Γ. Then one can find a smaller identity neighborhood $U \subset V \subset H$ such that for any selection of points $y_j \in U\gamma_j U,\ j = 1, \ldots, r$, the group they generate $\langle y_1, \ldots, y_r\rangle$ is dense in G. Indeed, as $\Gamma \cap H$ is dense in H, there are l words w_i in r letters such that $w_i(\gamma_1, \ldots, \gamma_r) \in V h_i V$ for $i = 1, \ldots, l$. Hence, for some smaller neighborhood $U \subset V \subset H$ and for any selection of points $y_j \in U\gamma_j U,\ j = 1, \ldots, r$, we will have $w_i(y_1, \ldots, y_r) \in V h_i V$ for $i = 1, \ldots, l$. But then $\langle y_1, \ldots, y_r\rangle$ is dense in G, since its intersection with the normal subgroup H is dense in H, and its projection to G/H coincides with the projection of Γ to G/H.

Let $R \leq G^\circ$ be the solvable radical of G°. The group G/R is a semisimple Lie group with connected component G°/R and H clearly projects onto G°/R. Composing the projection $G \to G/R$ with the adjoint representation of G/R on its Lie algebra $\mathfrak{v} = Lie(G^\circ/R)$, we get a homomorphism $\pi : G \to \mathrm{GL}(\mathfrak{v})$. The image $\pi(G)$ is open in the group of real points of some real algebraic group $\mathbb{H}$ whose connected identity component $\mathbb{H}^\circ$ is semisimple. Moreover $\pi(\Gamma)$ is dense in $\pi(G)$. Let $m = |\mathbb{H}/\mathbb{H}^\circ|$ and let $g_1, \ldots, g_m \in \Gamma$ be elements which are sent under π to representatives of all cosets of $\mathbb{H}^\circ$ in $\mathbb{H}$. Let U_0 be an even smaller symmetric identity neighborhood $U_0 \subset U \subset H$ such that $U_1^{4m+2} \subset U$ where $U_1 = \cup_{j=1}^m g_j U_0 g_j^{-1}$, and set $\Omega_0 = \pi(U_0 \cap \Gamma)$. Then the conditions of Theorem 4.3 are satisfied, since Ω_0 is Zariski dense in $\mathbb{H}^\circ(\mathbb{R})$ (see [31] Lemma 5.2 applied to H). Thus we can choose $\alpha_i \in \Gamma \cap U_1^{4m+2}\gamma_i U_1$ which generate a free group $\langle \alpha_1, \ldots, \alpha_r \rangle$. It will also be dense by the discussion above.

7.5.2 The p-adic case

Suppose now that k is a non-Archimedean local field of characteristic 0, i.e. it is a finite extension of the field of p-adic numbers $\mathbb{Q}_p$ for some prime $p \in \mathbb{N}$. Let $\mathcal{O} = \mathcal{O}_k$ be the valuation ring of k and $\mathfrak{p}$ its maximal ideal. Let $\Gamma \leq \mathrm{GL}_n(k)$ be a finitely generated linear group over k, let G be the closure of Γ in $\mathrm{GL}_n(k)$. Let $G(\mathcal{O}) = G \cap \mathrm{GL}_n(\mathcal{O})$ (and $\Gamma(\mathcal{O}) = \Gamma \cap G(\mathcal{O})$) and denote by $\mathrm{GL}_n^1(\mathcal{O})$ the first congruence subgroup, i.e. the kernel of the homomorphism $\mathrm{GL}_n(\mathcal{O}) \to \mathrm{GL}_n(\mathcal{O}/\mathfrak{p})$. The subgroup $G^1(\mathcal{O}) = G \cap \mathrm{GL}_n^1(\mathcal{O})$ is an open compact subgroup of G and is a p-adic analytic pro-p group. The group $\mathrm{GL}_n(\mathcal{O})$ has finite rank (i.e. there is an upper bound on the minimal number of topological generators for all closed subgroups of $\mathrm{GL}_n(\mathcal{O})$) as it follows for instance from Theorem 5.2 in [49]. Consequently, $G(\mathcal{O})$ itself is finitely generated as a pro-finite group and it contains the finitely generated pro-p group $G^1(\mathcal{O})$ as a subgroup of finite index. This implies that the Frattini subgroup $F \leq G(\mathcal{O})$ (the intersection of all maximal open subgroups of $G(\mathcal{O})$) is open (normal), hence of finite index in $G(\mathcal{O})$ (Proposition 1.14 in [49]). In this situation, generating a dense group in G is an open condition. More precisely :

Lemma 5.2 *Suppose $x_1, \ldots, x_r \in G$ generate a dense subgroup of G, then there is a neighborhood of identity $U \subset G$, such that for any selection of points $y_i \in Ux_iU$, $1 \leq i \leq r$, the y_i's generate a dense subgroup of G.*

Proof: Note that a subgroup of the pro-finite group $G(\mathcal{O})$ is dense if and only if it intersects every coset of the Frattini subgroup F. Now since $\langle x_1, \ldots, x_r \rangle$ is dense, there are $l = [G(\mathcal{O}) : F]$ words $\{w_i\}_{i=1}^l$ on r letters, such that the $w_i(x_1, \ldots, x_r)$'s are representatives of all cosets of F in $G(\mathcal{O})$. But then, if U is small enough, and $y_i \in Ux_iU$, the elements $w_i(y_1, \ldots, y_r)$ form again a full set of representatives for the cosets of F in $G(\mathcal{O})$. This implies that $\langle y_1, \ldots, y_r \rangle \cap G(\mathcal{O})$ is dense in $G(\mathcal{O})$. Now if we assume further that U lies inside the open subgroup $G(\mathcal{O})$, then we have $x_i \in G(\mathcal{O})y_iG(\mathcal{O})$, hence $x_i \in \overline{\langle y_1, \ldots, y_r \rangle}$ for $i = 1, \ldots, r$. This implies that $G = \overline{\langle y_1, \ldots, y_r \rangle}$. $\square$

The proof of the theorem now follows easily. Let $\{x_1, \ldots, x_r\}$ be a generating set for Γ. Choose U as in the lemma, and take it to be an open subgroup. Hence it satisfies $U^l = U$ for all $l \geq 1$. By Lemma 4.5 we have a representation $\rho : \Gamma \to \mathbb{H}$ into some semisimple k-algebraic group $\mathbb{H}$ such that the image of $U \cap \Gamma$ is Zariski dense in $\mathbb{H}^0$.

Thus we can use Theorem 4.3 in order to find elements $\alpha_i \in \Gamma \cap Ux_iU$ that generate a free group. It will be dense by Lemma 5.2.

7.5.3 The positive characteristic case

Finally, consider the case where k is a local field of characteristic $p > 0$, i.e. a field of formal power series $\mathbb{F}_q[[t]]$ over some finite field extension $\mathbb{F}_q$ of $\mathbb{F}_p$. First, we do not suppose that Γ is finitely generated (in particular in the lemma below). We use the same notations as those introduced at the beginning of the last Paragraph 7.5.2 in the p-adic case. In particular G is the closure of Γ, $G(\mathcal{O})$ is the intersection of G with $\mathrm{GL}_n(\mathcal{O})$ where $\mathcal{O}$ is the valuation ring of k. In positive characteristic, we have to deal with the additional difficulty that, even when Γ is finitely generated, $G(\mathcal{O})$ may not be topologically finitely generated.

However, when $G(\mathcal{O})$ is topologically finitely generated, then the argument used in the p-adic case (via Lemmas 4.5 and 5.2) applies here as well without changes. In particular, if $\overline{\Gamma}$ is compact, then we do not have to take more than $d(\Gamma)$ generators for the dense free subgroup. This fact will be used in Section 7.7. We thus have :

Proposition 5.3 *Let k be a non-Archimedean local field and let $\mathcal{O}$ be its valuation ring. Let $\Gamma \leq GL_n(\mathcal{O})$ be a finitely generated group which is not virtually solvable, then Γ contains a dense free group F_r for any $r \geq d(\Gamma)$.*

Moreover, it is shown in [18] that if k is a local field of positive characteristic, and $G = \mathbb{G}(k)$ for some semisimple simply connected k-algebraic group $\mathbb{G}$, then G and $G(\mathcal{O})$ are finitely generated. Thus, the above proof applies also to this case and we obtain :

Proposition 5.4 *Let k be a local field of positive characteristic, and let G be the group of k points of some semisimple simply connected k-algebraic group. Let Γ be a finitely generated dense subgroup of G, then Γ contains a dense F_r for any $r \geq d(\Gamma)$.*

Let us now turn to the general case, when $G(\mathcal{O})$ is not assumed topologically finitely generated. As above, we denote by $\mathrm{GL}_n^1(\mathcal{O})$ the first congruence subgroup $\mathrm{Ker}\big(\mathrm{GL}_n(\mathcal{O}) \to \mathrm{GL}_n(\mathcal{O}/\mathfrak{p})\big)$. This group is pro-$p$ and, as it is easy to see, the elements of torsion in $\mathrm{GL}_n^1(\mathcal{O})$ are precisely the unipotent matrices. In particular the order of every torsion element is $\leq p^n$. Moreover, every open subgroup of $\mathrm{GL}_n^1(\mathcal{O})$ contains elements of infinite order. Hence the torsion elements are not Zariski dense in $\mathrm{GL}_n^1(\mathcal{O})$. More generally we have :

Lemma 5.5 *Let k be a non-Archimedean local field of arbitrary characteristic and n a positive integer. There is an integer m such that the order of every torsion element in $\mathbb{GL}_n(k)$ divides m. In particular, if $\mathbb{H}$ is a semisimple algebraic k-group, then the set of torsion elements in $\mathbb{H}(k)$ is contained in a proper subvariety.*

Proof: Let $\mathrm{char}(k) = p \geq 0$. Suppose $x \in \mathbb{GL}_n(k)$ is an element of torsion, then x^{p^n} (resp. x if $char(k) = 0$) is semisimple. Since the minimal polynomial of x is of degree at most n, its eigenvalues, which are roots of unity, lie in an extension of degree at most n of k. But, as k is a local field, there are only finitely many such extensions. Moreover, in a given non-Archimedean local field, there are only finitely many roots of unity. Hence there is an integer m such that $x^m = 1$.

The last claim follows from the obvious fact that if $\mathbb{H}$ is semisimple then there are elements if infinite order in $\mathbb{H}(k)$. $\square$

Let $\rho : \Gamma \to \mathbb{H}$ be the representation given by Lemma 4.5. Then for any sufficiently small open subgroup U of $\mathrm{GL}_n(\mathcal{O})$ (for instance some small congruence subgroup), $\rho\big(\Gamma \cap U\big)$ is Zariski dense in $\mathbb{H}^\circ$. We then have (Γ is not assumed finitely generated) :

Lemma 5.6 *There are* $t := \dim(\mathbb{H})$ *elements* $x_1, \ldots, x_t \in \Gamma(\mathcal{O})$ *such that* $\rho\big(\langle x_1, \ldots, x_t\rangle\big)$ *is Zariski dense in* $\mathbb{H}^0$.

Proof: Let U be an open subgroup of $\mathrm{GL}_n(\mathcal{O})$ so that $\rho\big(\Gamma \cap U\big)$ lies in $\mathbb{H}^\circ$ and is Zariski dense in it. It follows from the above lemma that there is $x_1 \in \Gamma \cap U$ such that $\rho(x_1)$ is of infinite order. Then the algebraic group $A = \overline{\langle x_1\rangle}^z$ is at least one dimensional.

Let the integer i, $1 \leq i \leq t$, be maximal for the property that there exist i elements $x_1, \ldots, x_i \in \Gamma \cap U$ whose images in $\mathbb{H}$ generate a group whose Zariski closure is of dimension $\geq i$. We have to show that $i = t$. Suppose this is not the case. Fix such $x_1, \ldots, x_i$ and let A be the Zariski connected component of identity of $\overline{\langle \rho(x_1), \ldots, \rho(x_i)\rangle}^z$. Then for any $x \in \Gamma \cap U$, $\overline{\langle \rho(x_1), \ldots, \rho(x_i), \rho(x)\rangle}^z$ is i-dimensional. This implies that $\rho(x)$ normalizes A. Since $\rho\big(\Gamma \cap U\big)$ is Zariski dense in $\mathbb{H}^0$, we see that A is a normal subgroup of $\mathbb{H}^\circ$. Dividing $\mathbb{H}^\circ$ by A we obtain a Zariski connected semisimple k-group of positive dimension, and a map from $\Gamma \cap U$ with Zariski dense image into the k-points of this semisimple group. But then, again by Lemma 5.5 above, there is an element $x_{i+1} \in \Gamma \cap U$ whose image in $\mathbb{H}^\circ/A$ has infinite order — a contradiction to the maximality of i. $\square$

Suppose now that Γ is finitely generated and let Δ be the closure in $\mathrm{GL}_n(\mathcal{O})$ of the subgroup generated by $x_1, \ldots, x_t$ given by Lemma 5.6 above. It is a topologically finitely generated pro-finite group containing the pro-p subgroup of finite index $\Delta \cap G^1(\mathcal{O})$. Hence its Frattini subgroup F is open and of finite index ([49] Proposition 1.14). In particular, $\rho(F \cap \langle x_1, \ldots, x_t\rangle)$ is Zariski dense in $\mathbb{H}^\circ$, and we can use Theorem 4.3 with $\Omega_0 = \rho(F \cap \langle x_1, \ldots, x_t\rangle)$. Note that F, being a group, satisfies $F^m = F$ for $m \in \mathbb{N}$. Also F is normal in Δ. Let $\gamma_1, \ldots, \gamma_r$ be generators for Γ. By Theorem 4.3, we can choose $\alpha_i \in F\gamma_i F$, $i = 1, \ldots, r$, and $\alpha_{i+r} \in x_i F$, $i = 1, \ldots, t$, so that $D = \langle \alpha_1, \ldots, \alpha_{r+t}\rangle$ is isomorphic to the free group F_{r+t} on $r + t$ generators. Clearly $D \cap \Delta$ is dense in Δ and $D \cap F$ is dense in F. This implies that each γ_i lies in $F\alpha_i F \subset \overline{D}$. As the γ_i's generate Γ, we see that D is dense in $\overline{\Gamma}$ and this finishes the proof.

7.5.4 A stronger statement

The argument above combined with the argument of [31] Section 2 provides the following generalization :

Theorem 5.7 *Let k be a local field and let $G \leq GL_n(k)$ be a closed linear group containing no open solvable subgroup. Assume also that $G(\mathcal{O})$ is topologically finitely generated in case k is non-Archimedean of positive characteristic. Then there is an integer $h(G)$ which satisfies*

- *$h(G) \leq 2\dim(G) - 1 + d(G/G^\circ)$ if k is Archimedean, and*
- *$h(G)$ is the minimal cardinality of a set generating a dense subgroup of G if k is non-Archimedean,*

such that any finitely generated dense subgroup $\Gamma \leq G$ contains a dense F_r, for any $r \geq \min\{d(\Gamma), h(G)\}$. Furthermore, if k is non-Archimedean or if G° is topologically perfect (i.e. $\overline{[G^\circ, G^\circ]} = G^\circ$) and $d(G/G^\circ) < \infty$, then we can drop the assumption that Γ is finitely generated. In these cases, any dense subgroup Γ in G contains a dense F_r for any $r \geq h(G)$.

Remark 5.8 *The interested reader is referred to [31] for a sharper estimation of $h(G)$ in the Archimedean case. For instance, if G is a connected and semisimple real Lie group, then $h(G) = 2$.*

Let us also remark that in the characteristic zero case, we can drop the linearity assumption, and assume only that Γ is a subgroup of some second countable k-analytic Lie group. To see this, simply note that the procedure of generating a dense subgroup does not rely upon the linearity of $G = \overline{\Gamma}$, and for generating a free subgroup, we can look at the image of G under the adjoint representation which is a linear group. The main difference in the positive characteristic case is that we do not know in that case whether or not the image $\mathrm{Ad}(G)$ is solvable. For this reason we make the additional linearity assumption in positive characteristic.

7.6 Dense free subgroups with infinitely many generators

In this section, we let k be any local field and $\Gamma \leq \mathrm{GL}_n(k)$ be any linear group over k. We shall prove the following :

Theorem 6.1 *Assume that Γ contains no open solvable subgroup. Then there is a dense countable subset $X \subset \Gamma$ which forms a free set. In particular the subgroup $\langle X \rangle \leq \Gamma$ is a dense free subgroup of infinite rank.*

As above, we denote by G the closure of Γ. We can assume that Γ is countable. If k is non-Archimedean we let $\mathcal{O}$ denote the valuation ring of k, and $G(\mathcal{O}) := G \cap$

$\mathrm{GL}_n(\mathcal{O})$ be the corresponding open pro-finite group and let $\Gamma(\mathcal{O}) = \Gamma \cap \mathrm{GL}_n(\mathcal{O})$. Set $G_j := G \cap \mathrm{GL}_n^j(\mathcal{O})$ where $\mathrm{GL}_n^j(\mathcal{O}) = \mathrm{Ker}\big(\mathrm{GL}_n(\mathcal{O}) \to \mathrm{GL}_n(\mathcal{O}/\mathfrak{p}^j)\big)$ is the j'th congruence subgroup, and write $\Gamma_j := \Gamma \cap G_j$. In order to treat both the Archimedean and the non-Archimedean cases at the same time, we will say *by convention* that in the Archimedean case, Γ_j denotes always the same group $H \cap \Gamma$ where H is the limit of the sequence of closed commutators introduced in Paragraph 7.5.1.

We now fix once and for all a sequence (x_j) of elements of Γ which is dense in G. In the non-Archimedean case, we can also require that the $G_{k_j} x_j G_{k_j}$'s form a base for the topology of G for some choice of a sequence of integers (k_j). We are going to perturb the x_i's by choosing elements y_i's inside $\Gamma_{k_j} x_j \Gamma_{k_j}$ which will all play ping-pong together on some projective space, hence generate a dense free group.

From Lemma 4.5 in the non-Archimedean case, and from the discussion in Paragraph 7.5.1 in the Archimedean case, we have a homomorphism $\pi : \Gamma \to \mathbb{H}(k)$, where $\mathbb{H}$ is an algebraic k-group with $\mathbb{H}^\circ$ semisimple, such that the Zariski closure of $\pi(\Gamma_j)$ contains $\mathbb{H}^\circ$ for all $j \geq 1$. It now follows from Lemma 5.6 when $char(k) > 0$ and from the discussion in Paragraphs 7.5.1 and 7.5.2 in the other cases (i.e. from the fact that H and G_j are topologically finitely generated) that Γ_1 contains a finitely generated subgroup Δ_1 such that $\pi(\Delta_1)$ is also Zariski dense in $\mathbb{H}^\circ$ (we also take Δ_1 to be dense in H when k is Archimedean). From Theorem 4.3 we can find a local field k' and an irreducible projective representation on $\mathbb{P}(V_{k'})$ of $\mathbb{H}$ defined over k' such that, under this representation, some elements of Δ_1 play ping-pong in the projective space $\mathbb{P}(V_{k'})$. In particular, for some $r > 0$ and for every positive $\epsilon < \frac{r}{2}$, there is an element in Δ_1 acting on $\mathbb{P}(V_{k'})$ by an (r, ϵ)-very proximal transformation (c.f. Paragraph 7.3.1). Furthermore, there is a field extension K of k' such that under this representation the full group Γ is map into $\mathrm{PGL}(V_K)$ where $V_K = V_{k'} \otimes K$. This field extension may not be finite, nor finitely generated. Nevertheless, the absolute value on k' extends to an absolute value on K (see [103] XII, 4, Theorem 4.1 p. 482) and the projective space $\mathbb{P}(V_K)$ is still a metric space (although not compact in general) for the metric introduced in Paragraph 7.3.1. Moreover, if $[g] \in \mathrm{PGL}(V_{k'})$ is ϵ-contracting on $\mathbb{P}(V_{k'})$, it is $c\epsilon$-contracting on $\mathbb{P}(V_K)$ for some constant $c = c(k', K) \geq 1$. Similarly, if $[g] \in \mathrm{PGL}_n(V_{k'})$ is (r, ϵ)-proximal transformation on $\mathbb{P}(V_{k'})$ then it is $(\frac{r}{c}, c\epsilon)$-proximal on $\mathbb{P}(V_K)$. Let $\rho : \Gamma \to \mathrm{PGL}(V_K)$ be this representation. (The reason why we may not assume that K is local is that there may not be a finitely generated dense subgroup in Γ.)

In the Archimedean case, the discussion in Paragraph 7.5.1 shows that we can find inside Δ_1 elements $z_1, \ldots, z_l$, generating a dense subgroup of $\Gamma_1 = H \cap \Gamma$, and such that, under the above representation, they act as a ping-pong l-tuple of projective transformations. We can find another element $g \in \Delta_1$ such that $(z_1, \ldots, z_l, g)$ acts as a ping-pong $(l+1)$-tuple and g acts as an (r, ϵ)-very proximal transformation on $\mathbb{P}(V_{k'})$ where the pair (r, ϵ) satisfies the conditions of Lemma 3.1 with respect to k'. In the non-Archimedean case, let simply g be some element of Δ_1 acting as an (r, ϵ)-very proximal transformation on the projective space $\mathbb{P}(V_{k'})$ with (r, ϵ) as in Lemma 3.1. As follows from Lemma 3.1, g (resp. g^{-1}) fixes an attracting point $\overline{v}_g$ (resp. $\overline{v}_{g^{-1}}$) and a repelling hyperplane $\overline{H}_g$ (resp.

$\overline{H}_{g^{-1}}$) and the positive (resp. negative) powers g^n behave as $(\frac{r}{C}, (C\epsilon)^{\frac{n}{3}})$-very proximal transformations with respect to these same attracting points and repelling hyperplanes. Note that in the non-Archimedean case, if n_j is the index of the j'th congruence subgroup G_j in $G(\mathcal{O})$, then $g^{n_j} \in G_j$ and in particular $g^{n_j} \to 1$ as j tends to infinity.

We are now going to construct an infinite sequence (g_j) of elements in Δ_1 acting on $\mathbb{P}(V_{k'})$ by very proximal transformations and such that they all play ping-pong together on $\mathbb{P}(V_{k'})$ (and also together with $z_1, \ldots, z_l$ in the Archimedean case). Since $\pi(\Delta_1)$ is Zariski dense in $\mathbb{H}^\circ$ and the representation of $\mathbb{H}^\circ$ is irreducible, we may pick an element $\gamma \in \Delta_1$ such that

$$\{\rho(\gamma)\overline{v}_g,\ \rho(\gamma)\overline{v}_{g^{-1}},\ \rho(\gamma^{-1})\overline{v}_g,\ \rho(\gamma^{-1})\overline{v}_{g^{-1}}\} \cap \left(\overline{H}_g \cup \overline{H}_{g^{-1}} \cup \{\overline{v}_g, \overline{v}_{g^{-1}}\}\right) = \emptyset.$$

Now consider the element $\delta_{m_1} = g^{m_1}\gamma g^{m_1}$. When m_1 is large enough, δ_{m_1} acts on $\mathbb{P}(V_{k'}^n)$ under ρ as a very proximal transformation, whose repelling neighborhoods lie inside the ϵ-repelling neighborhood of g and whose attracting points lies inside the ϵ-attracting neighborhood of g. We can certainly assume that $\rho(\delta_{m_1})$ satisfies the conditions of Lemma 3.1. Hence δ_{m_1} fixes some attracting points $\overline{v}_{\delta_{m_1}}, \overline{v}_{\delta_{m_1}^{-1}}$ which are close to, but distinct from $\overline{v}_g, \overline{v}_{g^{-1}}$ respectively. Similarly the repelling neighborhoods of $\delta_{m_1}, \delta_{m_1}^{-1}$ lie inside the ϵ-repelling neighborhood of g, g^{-1}, and the repelling hyperplanes $\overline{H}_{\delta_{m_1}}, \overline{H}_{\delta_{m_1}^{-1}}$ are close to that of g. We claim that for all large enough m_1

$$\{\overline{v}_g, \overline{v}_{g^{-1}}\} \cap \left(\overline{H}_{\delta_{m_1}} \cup \overline{H}_{\delta_{m_1}^{-1}}\right) = \emptyset, \text{ and } \{\overline{v}_{\delta_{m_1}}, \overline{v}_{\delta_{m_1}^{-1}}\} \cap \left(\overline{H}_g \cup \overline{H}_{g^{-1}}\right) = \emptyset. \tag{7.3}$$

Let us explain, for example, why $\overline{v}_{g^{-1}} \notin \overline{H}_{\delta_{m_1}}$ and why $\overline{v}_{\delta_{m_1}} \notin \overline{H}_{g^{-1}}$ (the other six conditions are similarly verified). Apply δ_{m_1} to the point $\overline{v}_{g^{-1}}$. As g stabilizes $\overline{v}_{g^{-1}}$ we see that

$$\delta_{m_1}(\overline{v}_{g^{-1}}) = g^{m_1}\gamma g^{m_1}(\overline{v}_{g^{-1}}) = g^{m_1}\gamma(\overline{v}_{g^{-1}}).$$

Now, by our assumption, $\gamma(\overline{v}_{g^{-1}}) \notin \overline{H}_g$. Moreover when m_1 is large, g^{m_1} is a ϵ_{m_1}-contracting with $\overline{H}_{g^{m_1}} = \overline{H}_g$, $\overline{v}_{g^{m_1}} = \overline{v}_g$ and ϵ_{m_1} arbitrarily small. Hence, we may assume that $\gamma(\overline{v}_{g^{-1}})$ is outside the ϵ_{m_1} repelling neighborhood of g^{m_1}. Hence $\delta_{m_1}(\overline{v}_{g^{-1}}) = g^{m_1}\big(\gamma(\overline{v}_{g^{-1}})\big)$ lie near $\overline{v}_g$ which is far from $\overline{H}_{\delta_{m_1}}$. Since $\overline{H}_{\delta_{m_1}}$ is invariant under δ_{m_1}, we get that $\delta_{m_1}\overline{v}_{g^{-1}} \notin \overline{H}_{\delta_{m_1}}$.

To show that $\overline{v}_{\delta_{m_1}} \notin \overline{H}_{g^{-1}}$ we shall apply g^{-2m_1} to $\overline{v}_{\delta_{m_1}}$. If m_1 is very large then $\overline{v}_{\delta_{m_1}}$ is very close to $\overline{v}_g$, and hence also $g^{m_1}(\overline{v}_{\delta_{m_1}})$ is very close to $\overline{v}_g$. As we assume that γ takes $\overline{v}_g$ outside $\overline{H}_{g^{-1}}$, we get (by taking m_1 sufficiently large) that γ also takes $g^{m_1}\overline{v}_{\delta_{m_1}}$ outside $\overline{H}_{g^{-1}}$. Taking m_1 even larger if necessary we get that g^{-m_1} takes $\gamma g^{m_1}\overline{v}_{\delta_{m_1}}$ to a small neighborhood of $\overline{v}_{g^{-1}}$. Hence

$$g^{-2m_1}\overline{v}_{\delta_{m_1}} = g^{-2m_1}\delta_{m_1}\overline{v}_{\delta_{m_1}} = g^{-m_1}\gamma g^{m_1}\overline{v}_{\delta_{m_1}}$$

lies near $\overline{v}_{g^{-1}}$. Since $\overline{H}_{g^{-1}}$ is g^{-2m_1} invariant and is far from $\overline{v}_{g^{-1}}$, we obtain that $\overline{v}_{\delta_{m_1}} \notin \overline{H}_{g^{-1}}$.

Now it follows from (7.3) and Lemma 3.1 that for every $\epsilon_1 > 0$ we can take j_1 sufficiently large so that g^{j_1} and $\delta_{m_1}^{j_1}$ are ϵ_1-very proximal transformations, and the ϵ_1-repelling neighborhoods of each of them are disjoint from the ϵ_1-attracting points of the other, and hence they form a ping-pong pair. Set $g_1 = \delta_{m_1}^{j_1}$.

In a second step, we construct g_2 in an analogous way to the first step, working with g^{j_1} instead of g. In this way we would get g_2 which is ϵ_2-very proximal, and play ping-pong with $g^{j_1 j_2}$. Moreover, by construction, the ϵ_2-repelling neighborhoods of g_2 lie inside the ϵ_1-repelling neighborhoods of g^{j_1}, and the ϵ_2-attracting neighborhoods of g_2 lie inside the ϵ_1-attracting neighborhoods of g^{j_1}. Hence the three elements g_1, g_2 and $g^{j_1 j_2}$ form a ping-pong 3-tuple.

We continue recursively and construct the desired sequence (g_n). Note that in the Archimedean case, we have to make sure that the g_n's form a ping-pong $\aleph_0$-tuple also when we add to them the finitely many z_i's. This can be done by declaring $g_i = z_i$ for $i = 1, \dots l$, and starting the recursive argument by constructing g_{l+1}.

Now since all $\Gamma_{k_j} = G_{k_j} \cap \Gamma$'s are mapped under the homomorphism π to Zariski dense subsets of $\mathbb{H}^\circ$, we can multiply x_j on the left and on the right by some elements of Γ_{k_j} so that, if we call this new element x_j again, $\rho(x_j)\overline{v}_{g_j} \notin \overline{H}_{g_j}$ and $\rho(x_j^{-1})\overline{v}_{g_j^{-1}} \notin \overline{H}_{g_j^{-1}}$. Considering the element $y_j = g_j^{l_j} x_j g_j^{l_j}$ for some positive power l_j, we see that it lies in $\Gamma_{k_j} x_j \Gamma_{k_j}$. Moreover, if we take l_j large enough, it will behave on $\mathbb{P}(V_K)$ like a very proximal transformation whose attracting and repelling neighborhoods are contained in those of g_j. Therefore, the y_j's also form an infinite ping-pong tuple and in the Archimedean case they do so together with $z_1, \dots, z_l$. Hence the family $(y_j)_j$ (resp. $(z_1, \dots, z_l, (y_j)_j)$) generate a free group.

In the non-Archimedean case, the y_j's are already dense in G since we selected them from sets which form a base for the topology. In the Archimedean case, the elements $z_1, \dots, z_k$ already generate a dense subgroup of H, and since the g_j's belong to H and the x_j's are dense, we see that the group generated by the z_i's and y_j's is dense in G. This completes the proof of Theorem 6.1.

7.7 Applications to pro-finite groups

We derive two conclusions in the theory of pro-finite groups. The following was conjectured by Dixon, Pyber, Seress and Shalev (see [50]) :

Theorem 7.1 *Let Γ be a finitely generated linear group over some field. Assume that Γ is not virtually solvable. Then, for any integer $r \geq d(\Gamma)$, its pro-finite completion $\hat{\Gamma}$ contains a dense free subgroup of rank r.*

Proof: Let R be the ring generated by the matrix entries of the elements of Γ. It follows from the Noether normalization theorem that R can be embedded in the valuation ring $\mathcal{O}$ of some local field k. Such an embedding induces an embedding of Γ in the pro-finite group $\mathrm{GL}_n(\mathcal{O})$. By the universal property of $\hat{\Gamma}$ this embedding induces a surjective map $\hat{\Gamma} \to \overline{\Gamma} \leq \mathrm{GL}_n(\mathcal{O})$ onto the closure of the image of Γ in $\mathrm{GL}_n(k)$. Since Γ is not virtually solvable, $\overline{\Gamma}$ contains no open solvable subgroup, and hence by Theorem 1.2 (see also Proposition 5.3), $\overline{\Gamma}$ contains a dense F_r (in fact we can find such an F_r inside Γ). By Gaschütz's lemma (see [144], Proposition 2.5.4) it is possible to lift the r generators of this F_r to r elements in $\hat{\Gamma}$ generating a dense subgroup in $\hat{\Gamma}$. These lifts, thus, generate a dense F_r in $\hat{\Gamma}$. $\square$

Let now H be a subgroup of a group G. Following [152] we define the notion of coset identity as follows :

Definition 7.2 *A group G satisfies a coset identity with respect to H if there exist*
- *a non-trivial reduced word W on l letters,*
- *l fixed elements $g_1, \ldots, g_l$,*

such that the identity

$$W(g_1h_1, \ldots, g_lh_l) = 1$$

holds for any $h_1, \ldots, h_l \in H$.

It was conjectured by Shalev [152] that if there is a coset identity with respect to some open subgroup in a pro-p group G, then there is also an identity in G. The following immediate consequence of Corollary 4.4 settles this conjecture in the case where G is an analytic pro-p group, and in fact, shows that a stronger statement is true in this case :

Theorem 7.3 *Let G be an analytic pro-p group. If G satisfies a coset identity with respect to some open subgroup, then G is virtually solvable.*

Proof: If G is not virtually solvable, then by Corollary 4.4, we can choose coset representatives for H in G which are free generators of a free group. $\square$

In fact the analogous statement holds also for finitely generated linear groups :

Theorem 7.4 *Let Γ be a finitely generated linear group over any field. If Γ satisfies a coset identity with respect to some finite index subgroup Δ, then Γ is virtually solvable.*

7.8 Applications to amenable actions

For convenience, we introduce the following definition :

Definition 8.1 *We shall say that a topological group G has property (OS) if it contains an open solvable subgroup.*

Our main result, Theorem 1.2, states that if Γ is a (finitely generated) linear topological group over a local field, then either Γ has property (OS) or Γ contains a dense (finitely generated) free subgroup. In the previous section we proved the analogous statement for pro-finite completions of linear groups over an arbitrary field. For real Lie groups, property (OS) is equivalent to "the identity component is solvable".

It was conjectured by Connes and Sullivan and proved subsequently by Zimmer [178] that if Γ is a countable subgroup of a real Lie group G, then the action of Γ on G by left multiplications is amenable if and only if Γ has property (OS). We refer the reader to [179] Chapter 4 for an introduction and background on amenable actions. The harder part of the equivalence is to show that if Γ acts amenably then it has (OS). As noted by Carrière and Ghys [40], the Connes-Sullivan conjecture is a straightforward consequence of Theorem 1.2. Let us reexplain this claim : by Theorem 1.2, it is enough to show that if Γ contains a non-discrete free subgroup, then it cannot act amenably.

Proof: (*non-discrete free subgroup$\Rightarrow$action is non-amenable*). By contradiction, if Γ were acting amenably, then any subgroup would do so too, hence we can assume that Γ itself is a non-discrete free group $\langle x, y\rangle$. By Proposition 4.3.9 in [179], it follows that there exists a Γ-equivariant Borel map $g \mapsto m_g$ from G to the space of probability measures on the boundary $\partial\Gamma$. Let X (resp. Y) be the set of infinite words starting with a non trivial power of x (resp. y). Let (ξ_n) (resp. θ_n) be a sequence of elements of Γ tending to the identity element in G and consisting of reduced words starting with y (resp. y^{-1}). By the converse to Lebesgue's dominated convergence theorem, up to passing to a subsequence of $(\xi_n)_n$ if necessary, we have that for almost all $g \in G$, $m_{g\xi_n}(X)$ and $m_{g\theta_n}(X)$ converge to $m_g(X)$. However, for almost every $g \in G$, $m_{g\xi_n}(X) = m_g(\xi_n X)$ and $m_{g\theta_n}(X) = m_g(\theta_n X)$. Moreover $\xi_n X$ and $\theta_n X$ are disjoint subsets of Y. Hence, for almost every $g \in G$, $2m_g(X) \leq m_g(Y)$. Reversing the roles of X and Y we get a contradiction. $\square$

A theorem of Auslander (see [137] 8.24) states that if G is a real Lie group, R a closed normal solvable subgroup, and Γ a subgroup with property (OS), then the image of Γ in G/R also has property (OS). Taking G to be the group of Euclidean motions, R the subgroup of translations, and $\Gamma \leq G$ a torsion free lattice, one obtains the classical theorem of Bieberbach that any compact Euclidean manifold is finitely covered by a torus.

Following Zimmer ([178]), we remark that Auslander's theorem follows from Zimmer's theorem. To see this, note that G (hence also Γ) being second countable, we can always replace Γ by a countable dense subgroup of it. Then, if Γ has property (OS) it must act amenably on G. As R is closed and amenable, this implies that $\Gamma R/R$ acts amenably on G/R (see [179]), which in turn implies, by Zimmer's theorem, that $\Gamma R/R$ has property (OS).

A discrete linear group is amenable if and only if it is virtually solvable. It follows that for a countable linear group over some topological field, being (OS) is the same as containing an open amenable subgroup.

Definition 8.2 *We shall say that a countable topological group Γ has property (OA) if it contains an open subgroup, which is amenable in the abstract sense (i.e. amenable with respect to the discrete topology).*

The following is a generalization of Zimmer's theorem :

Theorem 8.3 *Let G be a locally compact group, and let $\Gamma \leq G$ be a countable subgroup. Then the action of Γ on G by left multiplications is amenable if and only if Γ has property (OA).*

As an immediate corollary we obtain the following generalization of Auslander's theorem :

Corollary 8.4 *Let G be locally compact group, $R \leq G$ a closed normal amenable subgroup, and $\Gamma \leq G$ a subgroup with property (OA). Then the image of Γ in G/R has also (OA).*

Remark 8.5 *The original statement of Auslander follows easily from 8.4.*

Proof of 8.3. The proof makes use of the structure theory for locally compact groups (see [123]). We shall reduce the general case to the already known case of real Lie groups.

The "if" side is clear.

Assume that Γ acts amenably. Let G^0 be the identity connected component of G. Then G^0 is normal in G and $F = G/G^0$ is a totally disconnected locally compact group, and as such, has an open profinite subgroup. Since Γ acts amenably, its intersection with an open subgroup acts amenably on the open subgroup. Therefore we can assume that G/G^0 itself is profinite. By [123] Theorem 4.6, there is a compact normal subgroup K in G such that the quotient G/K is a Lie group. Up to passing to an open subgroup of G again, we can assume that G/K is connected. Since $\Gamma \cap K$ acts amenably on K and K is amenable, $\Gamma \cap K$ is amenable (see [179], Chapter 4). Moreover, as K is amenable, Γ, and hence also $\Gamma K/K$, acts amenably on the connected Lie group G/K. We conclude that $\Gamma K/K$ has property (OS) and hence Γ has property (OA). $\square$

7.9 The growth of leaves in Riemannian foliations

The main result of this section is the following theorem which answers a question of Carrière [39] (see also [80]) :

Theorem 9.1 *Let $\mathcal{F}$ be a Riemannian foliation on a compact manifold M. The leaves of $\mathcal{F}$ have polynomial growth if and only if the structural Lie algebra of $\mathcal{F}$ is nilpotent. Otherwise, generic leaves have exponential growth.*

For background and definitions about Riemannian foliations, the structural Lie algebra, and growth of leaves see [39], [80] and [120].

Following Carrière [39],[41] we define the **local growth** of a finitely generated subgroup Γ in a given connected real Lie group G in the following way. Fix a left-invariant Riemannian metric on G and consider the open ball B_R of radius $R > 0$ around the identity. Suppose that S is a finite symmetric set of generators of Γ. Let $B(n)$ be the ball of radius n in Γ for the word metric determined by S, and let $B_R(n)$ be the subset of $B(n)$

consisting of those elements $\gamma \in B(n)$ which can be written as a product $\gamma = \gamma_1 \cdot \ldots \cdot \gamma_k$, $k \leq n$, of generators $\gamma_i \in S$ in such a way that whenever $1 \leq i \leq k$ the element $\gamma_1 \cdot \ldots \cdot \gamma_i$ belongs to B_R. In this situation, we say that γ can be written as a word with letters in S which *stays all its life in* B_R. Let $f_{R,S}(n) = card(B_R(n))$. As it is easy to check, if S_1 and S_2 are two symmetric sets of generators of Γ, then there exist integers $N_0, N_1 > 0$ such that $f_{R,S_1}(n) \leq f_{R+N_0,S_2}(N_1 n)$.

Definition 9.2 *The local growth of* Γ *in* G *with respect to a set* S *of generators and a ball* B_R *of radius* R *is the growth type of* $f_{R,S}(n)$.

The growth type of $f_{R,S}(n)$ is *polynomial* if there are positive constants A and B such that $f_{R,S}(n) \leq An^B$ and is *exponential* if there are constants $C > 0$ and $\rho > 1$ such that $f_{R,S}(n) \geq C\rho^n$. It can be seen that Γ is discrete in G if and only if the local growth is bounded for any S and R.

Carrière [39], [80] showed that Theorem 9.1 is a consequence of the following :

Theorem 9.3 *Let* Γ *be a finitely generated dense subgroup of a connected real Lie group* G*. If* G *is nilpotent then* Γ *has polynomial local growth (for any choice of* S *and* R*). If* G *is not nilpotent, then* Γ *has exponential local growth (for any choice of* S *and any* R *big enough).*

The rest of this section is therefore devoted to the proof of Theorem 9.3. As it turns out, Theorem 9.3 is an immediate corollary of Theorem 1.2 in the case when G is not solvable. When G is solvable we can adapt the argument as shown below. The main proposition is the following :

Proposition 9.4 *Let* G *be a non-nilpotent connected real Lie group and* Γ *a finitely generated dense subgroup. For any finite set* $S = \{s_1, \ldots, s_k\}$ *of generators of* Γ*, and any* $\varepsilon > 0$*, one can find perturbations* $t_i \in \Gamma$ *of the* s_i*,* $i = 1, \ldots, k$ *such that* $t_i \in s_i B_\varepsilon$ *and the* t_i*'s are free generators of a free semi-group on* k *generators.*

Before going through the proof of this proposition, let us explain how we deduce from it a proof of Theorem 9.3.

Proof: [Proof of Theorem 9.3.] Suppose that $\Sigma := \{g_1, \ldots, g_N, h_1, \ldots, h_N\}$ is a subset of B_R consisting of pairwise distinct elements such that both $\{g_1, \ldots, g_N\}$ and $\{h_1, \ldots, h_N\}$ are maximal $R/2$-discrete subsets of $\overline{B_R}$ (that is $d(g_i, g_j), d(h_i, h_j) \geq R/2$ if $i \neq j$). Then

$$\overline{B_R} \subset \bigcup_{1 \leq i,j \leq N} (g_i B_{R/2} \cap h_j B_{R/2}).$$

Lemma 9.5 *Let* G *be a connected real Lie group endowed with a left-invariant Riemannian metric. Let* B_R *be the open ball of radius* R *centered at the identity. Let* $\Sigma = \{s_1, \ldots, s_k\}$ *be a finite subset of pairwise distinct elements of* B_R *such that*

$$\overline{B_R} \subset \bigcup_{i<j} (s_i^{-1} B_{R/2} \cap s_j^{-1} B_{R/2}). \tag{7.4}$$

Assume also that the elements of Σ *are free generators of a free semi-group. Then any finitely generated subgroup of* G *containing* Σ *has exponential local growth.*

Proof: Let $S(n)$ be the sphere of radius n in the free semi-group for the word metric determined by the generating set Σ. Let $w \in S(n) \cap B_R$. By (7.4) there are indices $i \neq j$ such that $w \in s_i^{-1} B_{R/2}$ and $w \in s_j^{-1} B_{R/2}$. This implies that $s_i w$ and $s_j w$ belong to $S(n+1) \cap B_R(n+1)$. All elements obtained in this way are pairwise distinct, hence $card(S(n+1) \cap B_R(n+1)) \geq 2 \cdot card(S(n) \cap B_R(n))$. This yields $card(S(n) \cap B_R(n)) \geq 2^n$ for all $n \geq 0$. $\square$

Now observe that any small enough perturbation of a finite set Σ in B_R satisfying (7.4) still satisfies (7.4). Hence exponential local growth for dense subgroups in non-nilpotent connected real Lie groups follows from the combination of Lemma 9.5 and Proposition 9.4. $\square$

Proof of Proposition 9.4. When G is not solvable, we already know this fact from the proof of Theorem 1.2 for connected Lie groups (see Paragraph 7.5.1). In that case, we showed that we could even take the t_i's to generate a free subgroup. Thus we may assume that G is solvable. By Ado's theorem it is locally isomorphic to a subgroup of $\mathrm{GL}_n(\mathbb{C})$, and it is easy to check that the property to be shown in Proposition 9.4 does not change by local isomorphisms. Thus, we may also assume that $G \leq \mathrm{GL}_n(\mathbb{C})$. Let $\mathbb{G}$ be the Zariski closure of G in $GL_n(\mathbb{C})$. It is a Zariski connected solvable algebraic group over $\mathbb{C}$ which is not nilpotent. We need the following elementary lemma for $k = \mathbb{C}$.

Lemma 9.6 *Let* $\mathbb{G}$ *be a solvable connected algebraic* k*-group which is not nilpotent. Suppose it is* k*-split, then there is an algebraic* k*-morphism from* $\mathbb{G}(k)$ *to* $\mathbb{GL}_2(k)$ *whose image is the full affine group*

$$\mathbb{A}(k) = \left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} a, b \in k \right\}. \tag{7.5}$$

Proof: We proceed by induction on $\dim \mathbb{G}$. We can write $G := \mathbb{G}(k) = T \cdot N$ where $T = \mathbb{T}(k)$ is a split torus and $N = \mathbb{N}(k)$ is the unipotent radical of G (see [26], Chapter III). Let Z be the center of N. It is a non trivial normal algebraic subgroup of G. If T acts trivially on Z by conjugation then G/Z is again non-nilpotent k-split solvable k-group and we can use induction. thus we may assume that T acts non-trivially on Z by conjugation. As T is split over k, its action on Z also splits, and there is a non-trivial algebraic multiplicative character $\chi : T \to \mathbb{G}_m(k)$ defined over k and a 1-dimensional subgroup Z_χ of Z such that, identifying Z_χ with the additive group $\mathbb{G}_a(k)$, we have $tzt^{-1} = \chi(t)z$ for all $t \in T$ and $z \in Z_\chi$. It follows that Z_χ is a normal subgroup in G, and we can assume that T acts trivially on N/Z_χ, for otherwise we could apply the induction assumption on G/Z_χ. For all $\gamma \in T$, this yields a homomorphism $\pi_\gamma : N \to Z_\chi$ given by the formula $\pi_\gamma(n) = \gamma n \gamma^{-1} n^{-1}$. Since T and N do not commute, π_γ is non trivial for at least one $\gamma \in T$. Fix such a γ and let N act on Z_χ by left multiplication by $\pi_\gamma(n)$. Let T act on Z_χ by conjugation. One can verify that this yields an algebraic action of the whole of G on Z_χ. Identifying Z_χ with the additive group $\mathbb{G}_a(k)$, we have that N acts unipotently and non-trivially and T acts via the non-trivial character χ. We have found a k-algebraic affine action of G on the line, and hence a k-map $\mathbb{G} \to \mathbb{A}$. Clearly this map is onto. $\square$

By Lemma 9.6, $\mathbb{G}$ surjects onto the affine group of the complex line, which we denote by $A = \mathbb{A}(\mathbb{C})$. The image of G is a real connected subgroup of A which is Zariski dense. Hence it is enough to prove Proposition 9.4 for Zariski dense connected subgroups of A. We need the following technical lemma :

Lemma 9.7 *Let Γ be a non-discrete finitely generated Zariski dense subgroup of $\mathbb{A}(\mathbb{C})$ with connected closure. Let $R \subset \mathbb{C}$ be the subring generated by the matrix entries of elements in Γ. Then there exists a sequence $(\gamma_n)_n$ of points of Γ, together with a ring embedding $\sigma : R \hookrightarrow k$ into another local field k, such that $\gamma_n = (a_n, b_n) \to (1, 0)$ in $\mathbb{A}(\mathbb{C})$ and $\sigma(\gamma_n) = \big(\sigma(a_n), \sigma(b_n)\big) \to (0, \sigma(\beta))$ in the topology of k for some number β in the field of fractions of R.*

Proof:

Let $g_n = (a_n, b_n)$ be a sequence of distinct elements of Γ converging to identity in $\mathbb{A}(\mathbb{C})$ and such that $|a_n|_{\mathbb{C}} \leq 1$ and $a_n \neq 1$ for all integers n. From Lemma 2.1 one can find a ring embedding $\sigma : R \hookrightarrow k$ for some local field k such that, up to passing to a subsequence of g_n's, we have $\sigma(a_n) \to 0$ in k. We can assume $|\sigma(a_n)|_k < 1$ for all n. Now let $\xi = (a, b) := g_0$ and consider the element

$$\xi^m g_n \xi^{-m} = (a_n, \frac{1 - a^m}{1 - a} b(1 - a_n) + a^m b_n).$$

Since $|a|_{\mathbb{C}} \leq 1$, the second component remains $\leq \frac{2}{|1-a|_{\mathbb{C}}} |b|_{\mathbb{C}} |1 - a_n|_{\mathbb{C}} + |b_n|_{\mathbb{C}}$ for all m, and tends to 0 in $\mathbb{C}$ when $n \to \infty$ uniformly in m. Applying the isomorphism σ, we have :

$$\sigma(\xi^m g_n \xi^{-m}) = \big(\sigma(a_n), \frac{1 - \sigma(a)^m}{1 - \sigma(a)} \sigma(b)(1 - \sigma(a_n)) + \sigma(a)^m \sigma(b_n)\big). \tag{7.6}$$

Since $|\sigma(a)|_k < 1$, for any given n, choosing m large, we can make $|\sigma(a)^m \sigma(b_n)|_k$ arbitrarily small. Hence for some sequence $m_n \to +\infty$ the second component in (7.6) tends to $\sigma(\beta)$ where $\beta := \frac{b}{1-a}$ as n tends to $+\infty$. $\square$

We shall now complete the proof of Proposition 9.4. Note that if k is some local field and $\gamma = (a_0, b_0) \in \mathbb{A}(k)$ with $|a_0|_k < 1$, then γ acts on the affine line k with a fixed point $x_0 = b_0/(1 - a_0)$ and it contracts the disc of radius R around x_0 to the disc of radius $|a_0|_k \cdot R$. Therefore, if we are given t distinct points $b_1, \ldots, b_t$ in k, there exists $\varepsilon > 0$ such that for all $a_1, \ldots, a_t \in k$ with $|a_i|_k \leq \varepsilon$, $i = 1, \ldots, t$, the elements (a_i, b_i)'s play ping-pong on the affine line, hence are free generators of a free semi-group. The group $G = \overline{\Gamma}$ is a connected and Zariski dense subgroup of $\mathbb{A}(\mathbb{C})$: it follows that we can find arbitrary small perturbations $\widetilde{s_i}$ of the s_i's within Γ such that the $a(\tilde{s}_i)\beta + b(\tilde{s}_i)$'s are pairwise distinct complex numbers. If $(\gamma_n)_n$ is the sequence obtained in the last Lemma, then for some n large enough the points $t_i := \widetilde{s}_i\gamma_n$ will be small perturbations of the s_i's (i.e. belong to s_iB_ε) and the $\sigma(t_i) = \big(\sigma(a(\tilde{s}_i)a_n), \sigma(a(\tilde{s}_i)b_n) + \sigma(b(\tilde{s}_i))\big)$ will play ping-pong on k for the reason we just explained (the $\sigma(a(\tilde{s}_i))\sigma(\beta) + \sigma(b(\tilde{s}_i))$'s are all distinct). $\square$

7.10 Some concluding remarks

It is natural to ask : *To what extent the analog of Theorem 1.2 holds ?*

Using our methods, one can prove for example the analogous result for products of finitely many simple linear Lie groups over various local fields, e.g. any dense subgroup of $\mathrm{SL}_n(\mathbb{R}) \times \mathrm{SL}_n(\mathbb{Q}_p)$ has a dense free subgroup of finite rank. In some cases, using the same methods we can show that any subgroup which does not have an open solvable subgroup has a non-discrete (not necessarily dense) free subgroup. This for example holds for the group of adèles of a semisimple algebraic group over an algebraic number field, e.g. for $\mathrm{SL}_n(\mathbb{A}_\mathbb{Q})$. However, in general, not every subgroup of a locally compact group which has no open solvable subgroup contains a non-discrete free subgroup. For instance, take a finitely generated non virtually solvable group which has no non-abelian free subgroups, and embed it into its pro-finite completion. It would be interesting to know if every locally compact group with no open solvable subgroup contains a dense (or at least non-discrete) free subgroup. In particular, does any pro-finite group which is not virtually solvable has a non-abelian free subgroup ?

Acknowledgments 10.1 *We thank the following persons for their help and the interesting conversations we had about various points in this article : Y. Barnea, D. Getz, E. Ghys, M. Larsen, A. Lubotzky, G.A. Margulis, N. Nikolov, J.F. Quint, A. Salehi-Golsefidi, G.A. Soifer. We also thank Olivia Barbarroux for her hospitality during our stay at Luminy in July 2002 and H. Abels for inviting us to Bielefeld in July 2003, where part of this work was conducted.*

Annexe A

An effective Tits alternative[1]

A.1 statements

We prove

Theorem 1.1 *Let G be a connected rank-1 simple real Lie group, and let Γ be an arithmetic lattice in G. Then there is an integer m such that if $\Delta \leq \Gamma$ is a Zariski dense subgroup and if Σ is a generating set for Δ, then the m-ball Σ^m contains two elements which generate a free group. Moreover, m depends only on vol(G/Γ).*

As noted in the Introduction, Chapter 5, this result actually holds for non-elementary discrete groups of isometries of geometrically finite manifolds of strictly negative curvature (see [10]). However, our methods allow us to prove the following :

Proposition 1.2 *If G be a connected semisimple Lie group, and $\Delta \leq G_{\mathbb{Z}}$ a Zariski dense subgroup. Then there is an integer m, such that if Σ is a generating set for Δ which satisfies $\Sigma = \Sigma^{-1} = {}^t\Sigma$ then Σ^m contains generators of a free group F_2.*

Moreover, we reduce the proof of the effective Tits alternative (i.e. the validity of last proposition without assuming that $\Sigma = \Sigma^t$) to the following conjectural lemma which we still cannot prove :

Let S be a symmetric space of non-compact type and c a unit speed oriented regular geodesic in S. The geodesic c determines uniquely a Cartan subgroup of Isom(S) and an order on its tangent Lie algebra $\mathfrak{h}$. There is some $H \in \mathfrak{h}$ and $x \in S$ for which $c(t) = \exp(tH) \cdot x$. We say that c is ϵ-regular, if $\alpha(H) \geq \epsilon$ for any positive root α.

Conjecture 1.3 *Let G be a center-free connected semisimple Lie group and Γ an arithmetic lattice in G. Then there are constants $m \in \mathbb{N}$, $\epsilon > 0$ (depending on Γ), such that if $\Sigma \subset \Gamma$ is a finite set which generates a Zariski dense subgroup, them Σ^m contains a regular element whose axis is ϵ-regular.*

[1]joint work with T. Gelander [33]

The idea of the proof of Theorem 1.1 is to play ping-pong both in the symmetric space and in the projective space; when it turns out that it is not possible to play ping pong in the symmetric space, we show that we can in fact do so in some projective space.

This is performed as follows. First we can easily reduce to the case where $G \leq \mathrm{GL}_n(\mathbb{R})$, $\Gamma \leq \mathrm{GL}_n(\mathbb{Z})$ and G is invariant under transpose. Let $K = G \cap \mathrm{SO}_n(\mathbb{R})$. We identify the symmetric space G/K to the space of normalized Euclidean norms on $\mathbb{R}^n$ by sending any point $x \in G/K$ to the natural Euclidean norm $\| \, \|_x$ on $\mathbb{R}^n$ given by the conjugate of the standard Euclidean norm on $\mathbb{R}^n$ by a representative of x in G. It induces a standard metric d_x on the projective space $P(\mathbb{R}^n)$. Using a generalized version of Bezout theorem we can find in Σ^m two hyperbolic elements a, b whose axis are of positive distance from each other. The arithmeticity gives a lower bound on the displacements of a and b. Hence if their axis are farther than some fixed constant then some bounded power of them plays ping-pong on the ideal boundary of G/K. Otherwise the axis are close, but then we can find a point in a compact set whose translation by some element of Γ is close to both axes. Using this point and translation we can define a Euclidean norm on $\mathbb{R}^n$ and a standard metric on the projective space $P(\mathbb{R}^n)$ with respect to which $\|a\|$, $\|b\|$ are comparable to their diagonal forms, and some bounded powers of a and bab^{-1} form a pair of proximal elements which plays ping-pong on $P(\mathbb{R}^n)$.

A.2 sketch of proofs

There exists a semisimple group G' defined over $\mathbb{Q}$ and an epimorphism $(G'_{\mathbb{R}})^0 \to G$ with compact kernel such that some finite index subgroup of $G'_{\mathbb{Z}}$ projects onto Γ. We shall identify Γ with this subgroup of $G'_{\mathbb{Z}}$. Clearly, we may assume that Γ is Zariski dense in G'. By Selberg's lemma, we can also assume that Γ is torsion free.

Via the adjoint representation we embed G' in $\mathrm{SL}_n(\mathbb{R})$. Then we choose a $\mathbb{Z}$-structure on $\mathbb{R}^n$ so that $\Gamma \leq \mathrm{SL}_n(\mathbb{Z})$. As $G \leq G'$ are reductive, it follows from a theorem of Mostow that we can assume that they are both invariant under transpose, i.e. ${}^tG = G, {}^tG' = G$. The symmetric space S of G embeds naturally as a totally geodesic submanifold of the symmetric space $P_n(\mathbb{R})$ of $\mathrm{SL}_n(\mathbb{R})$. $K = \mathrm{SO}_n(\mathbb{R}) \cap G$ is a maximal compact subgroup, and the embedding is given by

$$S = G/K \leq \mathrm{SL}_n(\mathbb{R})/\mathrm{SO}_n(\mathbb{R}) = P_n(\mathbb{R}).$$

We look at $P_n(\mathbb{R})$ as the set of normalized Euclidean norms on $\mathbb{R}^n$. As $\mathrm{SL}_n(\mathbb{R})$ is contained in the algebra of linear transformations of $\mathbb{R}^n$, any norm $\| \cdot \|$ on $\mathbb{R}^n$ induces an operator norm on $\mathrm{SL}_n(\mathbb{R})$ given by

$$\|g\| = \max_{v \in \mathbb{R}^n \setminus \{0\}} \frac{\|g(v)\|}{\|v\|}.$$

For $x \in P_n(\mathbb{R})$ we denote by $\| \cdot \|_x$ the corresponding norm on $\mathrm{SL}_n(\mathbb{R})$ and call it the x-norm. Observe that if $g \cdot x = y$ then $\|h\|_y = \|g^{-1}hg\|_x$ for any $h \in \mathrm{SL}_n(\mathbb{R})$. We denote by $I \in P_n(\mathbb{R})$ the point which corresponds to the standard Euclidean norm in $\mathbb{R}^n$ (i.e. the coset of 1 in $\mathrm{SL}_n(\mathbb{R})/\mathrm{SO}_n(\mathbb{R})$). Then S coincides with the orbit $G \cdot I$.

Since $\mathbb{R}\text{-rank}(G) = \mathbb{R}\text{-rank}(G') = 1$, any hyperbolic element $a \in G'$ has a unique eigenvalue outside the unit disk. We shall denote this maximal eigenvalue by $\alpha(a)$. Then $\alpha(a)$ is real and has the same multiplicity as the minimal eigenvalue $1/\alpha(a)$. For $a \in \mathrm{SL}_n(\mathbb{R})$ we denote by $\tau(a)$ the minimal displacement :

$$\tau(a) = \inf_{x \in P_n(\mathbb{R})} d(x, a \cdot x).$$

Lemma 2.1 *If $a \in G'$ is semisimple then* $\log|\alpha(a)| \leq \tau(a) \leq n^{\frac{1}{2}} \log|\alpha(a)|$.

The following is a consequence of the arithmeticity of Γ

Lemma 2.2 *There exist $\delta = \delta(\Gamma) > 0$ such that $\log \alpha(\gamma) \geq \delta$ for any semisimple element $\gamma \in \Gamma$.*

We shall use the following lemma (see [56])

Lemma 2.3 (Generalized Bezout theorem) *For any proper Zariski closed subset $X \subset G'$ there is a constant $k = k(\Delta, X)$ such that if Σ is any generating set for Δ, then Σ^k is not contained in X.*

Let now Σ be a given generating set for Δ. It follows from Lemma 2.3 that for some constant m_1, independent of Σ, the set Σ^{m_1} contains two semisimple elements a_*, b_* with no common eigenvectors. By replacing $\mathbb{R}^n$ with an appropriate wedge power we can assume that the multiplicity of the maximal eigenvalue $\alpha(a_*)$ of a_* is 1.

Let $\|\cdot\|$ (resp. $d(\cdot,\cdot)$) denote the norm on $\mathrm{SL}_n(\mathbb{R})$ (resp. the canonical metric on the projective space $\mathbb{P}(\mathbb{R}^n)$) which is induced from the standard Euclidean norm on $\mathbb{R}^n$. For $v, w \in \mathbb{R}^n \setminus \{0\}$, the distance $d\big([v],[w]\big)$ is defined by

$$d\big([v],[w]\big) = \frac{\|v \wedge w\|}{\|v\| \cdot \|w\|},$$

where $\|\cdot \wedge \cdot\|$ is the Euclidean norm on $\wedge^2(\mathbb{R}^n)$ induced from the standard Euclidean norm on $\mathbb{R}^n$.

For a polynomial $P(x) = \sum_{i=0}^n a_i x^i$ we denote $|P| = \max|a_i|$. If $P(x)$ is monic with integral coefficients and $a_0 \neq 0$, then any root α of P satisfies $\frac{1}{n|P|} \leq |\alpha| \leq n|P|$. Moreover for any rational number $\frac{p}{q}$ we have $|\alpha - \frac{p}{q}| \geq \frac{C}{q^n}$ for some constant $C > 0$ depending on P, as long as α is irrational (Liouville). This implies

Lemma 2.4 *If v be an eigenvector of a_*, then*

$$d\big([v],[b_*(v)]\big) \geq \max\{\|a_*\|, \|b_*\|\}^{-m_2}$$

for some constant $m_2 \in \mathbb{N}$ depending only on n.

Corollary 2.5 *For any constants l_3, c_3, there is another constant $m_3 = m_3(l_3, c_3)$, such that if $\max\{\|a_*\|, \|b_*\|\} \leq \big(\alpha(a_*)\big)^{l_3}$, then ${a_*}^{m_3}$ and $b_* {a_*}^{m_3} {b_*}^{-1}$ form a ping-pong pair when acting on $\mathbb{P}(\mathbb{R}^n)$.*

Proof: We use Proposition 3.3 from [31]. $\square$

So the game now is to find a point $x \in S$ so that, with respect to the norm $\|\cdot\|_x$, our chosen elements a_*, b_* satisfy $\max\{\|a_*\|_x, \|b\|_x\} \leq \alpha^l(a_*)$ for some fixed constant $l = l(n)$. This is not necessarily possible. However, we shall show that when there is no such $x \in S$ then a_*, b_* (or a bounded power of them) form a ping-pong pair with respect to the action on the boundary of the symmetric space $S(\infty)$.

We shall use the following comparison between the norm induced by $x \in S$ and the evaluation of the displacement at x.

Lemma 2.6 *For $x \in P_n(\mathbb{R})$ and $g \in SL_n(\mathbb{R})$ we have*

$$e^{\frac{1}{n}d(x,g\cdot x)} \leq \|g\|_x \leq n^{\frac{1}{2}} e^{d(x,g\cdot x)}.$$

Our elements a_*, b_* are semisimple, and since Γ is torsion free they are hyperbolic. We denote by c_{a_*} and c_{b_*} their corresponding axes. Observe that if $d(x, c_{a_*}) \leq 1$ then $d(x, a_* \cdot x) \leq \tau(a_*) + 2$, and hence

$$\|a_*\|_x \leq n^{\frac{1}{2}} e^{\tau(a_*)+2} \leq C_4 \cdot \big(\alpha(a_*)\big)^{m_4}$$

for some constants C_4, m_4. Therefore, if $d(c_{a_*}, c_{b_*}) \leq 1$ then there is $x \in S$ such that $d(x, c_{a_*}), d(x, c_{b_*}) \leq 1$, and hence the x-norm of each of them is bounded by C_4 times the maximal eigenvalue to the m_4. On the other hand, if the geodesics c_{a_*}, c_{b_*} are far from each other, then since $\tau(a_*), \tau(b_*) \geq \delta$, we have

Lemma 2.7 *If $d(c_{a_*}, c_{b_*}) \geq 1$ then for some constant m_5, the couple ${a_*}^{m_5}, {b_*}^{m_5}$ form a ping-pong pair with respect to the action on the ideal boundary $S(\infty)$ of S.*

There is the following alternative for the couple a_*, b_* :

1. $d(c_{a_*}, c_{b_*}) \geq 1$, in which case ${a_*}^{m_5}, {b_*}^{m_5}$ generate a free group by Lemma 2.7.
2. $d(c_{a_*}, c_{b_*}) \leq 1$, in which case there is a point $x \in S$ such that $\|a_*\|_x \leq C_4\big(\alpha(a_*)\big)^{m_4}$ and $\|b_*\|_x \leq C_4\big(\alpha(b_*)\big)^{m_4}$.

Below we prove the remaining case, that is we assume that we are in case (2). Let us first explain the proof of the theorem in the case G/Γ is compact.

Proof: [**The proof in the compact case**] As S/Γ is assumed to be compact, for some $r < \infty$ the ball B_r around $I \in S$ contains a fundamental domain for S/Γ so the translations $\Gamma \cdot B_r$ cover S. In particular, there is $y \in B_r$ and $\gamma \in \Gamma$ with $\gamma \cdot y = x$. We will choose the norm which corresponds to $\gamma \cdot I$. This amounts to consider the standard norm $\|\cdot\| = \|\cdot\|_I$ composed with the conjugation by γ^{-1}. Recall that

$$\|g\|_{\gamma \cdot I} = \|\gamma^{-1} g \gamma\|$$

for all $g \in \mathrm{SL}_n(\mathbb{R})$.

As $d(x, c_{a_*}), d(x, c_{b_*}) \leq 1$ we have $d(I, \gamma^{-1} \cdot c_{a_*}), d(I, \gamma^{-1} \cdot c_{b_*}) \leq r + 1$. Since $\gamma^{-1} \cdot c_{a_*}$ and $\gamma^{-1} \cdot c_{b_*}$ are the axes of $a_*^{\gamma^{-1}}$ and $b_*^{\gamma^{-1}}$ respectively, it follows from Lemmas 2.6 and 2.1 that

$$\|a_*^{\gamma^{-1}}\| \leq n^{\frac{1}{2}} e^{d(I, a_*^{\gamma^{-1}} \cdot I)} \leq n^{\frac{1}{2}} e^{2(r+1) + \tau(a_*)} \leq C_6 \big(\alpha(a_*)\big)^{m_6}$$

and similarly

$$\|b_*^{\gamma^{-1}}\| \leq C_6 \big(\alpha(b_*)\big)^{m_6}$$

Without loss of generality we may assume that $\alpha(a_*) \geq \alpha(b_*)$. Then by corollary 2.5 $a_*^{m_7}$ and $(a_*^{m_7})^{b_*}$ form a ping-pong pair with respect to the action on the projective space $\mathbb{P}(\mathbb{R}^n)$. $\square$

Proof: [**The proof in the non-compact case**] Let $M = \Gamma \setminus S$ denote the corresponding manifold, let ϵ be the constant from the Margulis lemma, and let $M_{\geq \epsilon}$ be the ϵ-thick part.

We need the following lemma

Lemma 2.8 *Any closed geodesic in M intersects $M_{\geq \epsilon}$.*

Let now $B_r \subset S$ be a ball centered at I whose image in $M = \Gamma \setminus S$ covers $M_{\geq \epsilon}$. Let $\pi(x)$ denote the projection of the point x to M. If $\pi(x)$ lies inside $M_{\geq \epsilon}$, then we can apply the same argument as in the compact case, and in fact we can do this whenever $\pi(x)$ lies in the image of (say) B_{r+2}. Assume now that $\pi(x)$ is not in $\Gamma \setminus \Gamma \cdot B_{r+2}$. Let l be the distance of $\pi(x)$ from $M_{\geq \epsilon}$, then $l \geq 2$. Since c_{a_*} and c_{b_*} pass at distance ≤ 1 from x and since their projection $\pi(c_{a_*}), \pi(c_{a_*})$ are closed geodesics in M, and hence intersect $M_{\geq \epsilon}$, it follows that the lengths of $\pi(c_{a_*})$ and $\pi(c_{a_*})$, which are precisely $\tau(a_*)$ and $\tau(b_*)$, are at least $2\big(d(\pi(x), M_{\geq \epsilon}) - 1\big) = 2(l-1)$. Now from 2.1 we have that

$$\alpha(a_*) \geq \exp\left(n^{-\frac{1}{2}} \tau(b_*)\right) \geq \exp\left(n^{-\frac{1}{2}} 2(l-1)\right)$$

and similarly

$$\alpha(b_*) \geq \exp\left(n^{-\frac{1}{2}} 2(l-1)\right).$$

On the other hand, let $y \in B_r$ be a point for which the distance to the orbit $\Gamma \cdot x$ is minimal, i.e. y projects to a closest point to $\pi(x)$ in $\pi(B_r)$. Then $d(\gamma \cdot y, x) = l$ for some

$\gamma \in \Gamma$. We shall choose the norm $\|\cdot\|_{\gamma\cdot I}$ which is given by

$$\|g\|_{\gamma\cdot I} = \|g^{\gamma^{-1}}\|.$$

Since $d(\gamma \cdot y, x) = l$ we have $d(\gamma \cdot y, c_{a_*}), d(\gamma \cdot y, c_{b_*}) \leq l+1$ and thus

$$d\big(a_* \cdot (\gamma \cdot y), \gamma \cdot y\big) \leq \tau(a_*) + 2(l+1)$$

which together with Lemma 2.6 gives that

$$\|a^*\|_{\gamma\cdot I} \leq n^{\frac{1}{2}} e^{\tau(a_*)+2(l+1)} \leq C_7\big(\alpha(a_*)\big)^{m_7},$$

and similarly

$$\|b^*\|_{\gamma\cdot I} \leq C_7\big(\alpha(b_*)\big)^{m_7}.$$

This along with corollary 2.5 implies that the couple $a_*^{M_8}, (a_*^{M_8})^{b_*}$ form a ping-pong pair with respect to the action on $\mathbb{P}(\mathbb{R}^n)$ equipped with the norm which is induced from the norm $\|\cdot\|_{\gamma\cdot I}$ on $\mathbb{R}^n$ (note that we can assume $\alpha(a_*) \geq \alpha(b_*)$). $\square$

Bibliographie

[1] H. Abels, G.A. Margulis, G. A. Soifer, *Semigroups containing proximal linear maps*, Israel J. Math. **91** (1995), no. 1-3, p. 1-30.

[2] S. Adams, G.A. Eliott, T. Giordano, *Amenable actions of groups*, Trans. Amer. Math. Soc. **344** (1994), no. 2, p. 803-822.

[3] M. Abert, *Group laws and free subgroups in topological groups* (Univ. Chicago preprint 2003).

[4] M. Abert, Y. Glasner, (Univ. Chicago preprint 2003)

[5] G. Alexopoulos, *Centered sub-laplacians with drift on Lie groups of polynomial volume growth*, Mem. Amer. Math. Soc. **155**, n. 739, (2002)

[6] G. Alexopoulos, *Random walks on discrete groups of polynomial volume growth*, Annals of Proba., vol **30**, n. 2, (2002) p. 723-801.

[7] G. Alexopoulos, *Centered densities on Lie groups of polynomial volume growth*, Probab. Theory Relat. Fields **124**, (2002) p. 112-150.

[8] R. Alperin, *Uniform exponential growth of polycyclic groups*, Geom. Dedicata **92** (2002), p. 105–113.

[9] R. Alperin, G. Noskov, *Uniform growth, action on trees and GL_2*, Computational and statistical group theory (Las Vegas, 2001), p. 1-5, Contemp. Math., **298**, Amer. Math. Soc., 2002

[10] R. Alperin, G. Noskov, *Non vanishing of algebraic entropy for geometrically finite groups of isometries of Hadamard manifolds*, (UC Berkeley preprint).

[11] N. Amosova, *Local limit theorems for probabilities of moderate deviations*, Lit. Mat. Sb. **14**, no3. (1974), p. 401-409.

[12] C. Anantharaman, *On spectral characterisations of amenability*, (preprint 2003, univ. Orléans).

[13] V.I. Arnol'd et 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.

[14] A. Avez, *Limite de quotients pour des marches aléatoires sur des groupes,* C. R. Acad. Sci. Paris Sér. A-B **276** (1973), A317–A320.

[15] M. Babillot, *Points entiers et groupes discrets, de l'analyse aux systèmes dynamiques*, Panoramas et synthèses, (2002).

[16] P. Baldi, *Caractérisation des groupes de Lie connexes récurrents,* Ann. Inst. H. Poincaré Sect. B (N.S.) 17 (1981), no. **3**, 281–308.

[17] P. Baldi, L. Caramelino, *Large and moderate deviations for random walks on nilpotent groups*, J. Theoret. Probab. **12** (1999), no. 3, p. 779–809.

[18] Y. Barnea, M. Larsen, *Random generation in semi-simple algebraic groups over local fields* (Univ. Indiana preprint).

[19] C. Béguin, A. Valette, A. Zuk, *On the spectrum of a random walk on the discrete Heisenberg group and the norm of Harper's operator*, J. Geom. and Phys. (1997), vol **21**, n. 4, p. 337-356.

[20] J. Bellissard, B. Simon, *Cantor spectrum for the almost Mathieu Equation*, J. Funct. Anal. (1982), vol **48**, n. 3, p. 408-419.

[21] Y. Benoist, *Propriétés asymptotiques des groupes linéaires*, GAFA, Geom. Funct. Anal. **7** (1997), p. 1-47.

[22] V. Bentkus, G. Pap, *The accuracy of Gaussian approximations in nilpotent Lie groups*, J. Theoret. Probab. **9** (1996), no. 4, p. 995–1017.

[23] P. Billingsley, *Measure and Probability*, 2ème édition. Wiley Series in Probability and Statistics : Probability and Statistics. John Wiley & Sons, 1999

[24] F. Boca, A. Zaharescu, *Norm estimates of almost Mathieu operators*, preprint ArXiv 2003.

[25] A. Borel, *Introduction aux groupes arithmétiques*, Hermann, Paris (1969).

[26] A. Borel, *Linear Algebraic Groups*, 2nd edition, GTM **126**, Springer-Verlag, (1991) 288 pp.

[27] P. Bougerol, *Théorème central limite local sur certains groupes de Lie*, Ann. Sci. École Norm. Sup. (4) 14 (1981), no. **4**, 403–432 (1982).

[28] P. Bougerol, J. Lacroix, *Products of random matrices with applications to Schrödinger operators*, Birkhäuser, Prog. in Probab. & Stat. (1985).

[29] N. Bourbaki, *Groupes et algèbres de Lie*, chapitres 7 et 8, Hermann, (1975).

[30] L. Breiman, *Probability*, Addison-Wesley (1968).

[31] E. Breuillard, T. Gelander, *On dense free subgroups of Lie groups*, J. Algebra **261** (2003), no. 2, p. 448-467.

[32] E. Breuillard, T. Gelander, *A topological version of Tits' alternative*, (preprint 2003, chapter 7 in this dissertation).

[33] E. Breuillard, T. Gelander, *Effective Tits' alternative for some linear groups*, (see Appendix in this dissertation).

[34] I. D. Brown, *Dual topology of a nilpotent Lie group*, Ann. Sci. École Norm. Sup. (4) **6** (1973), p. 407–411.

[35] F. Bruhat, J. Tits, *Groupes réductifs sur un corps local, I. Données radicielles valuées*, Publ. math. IHES, **41** (1972).

[36] A. Carbery, M. Christ, J. Wright, *Multidimensional van der Corput and sublevel set estimates*, J. Amer. Math. Soc. **12** (1999), no. 4, p. 981-1015

[37] H. Carlsson, *Error estimates in d-dimensional renewal theory*, Compositio Math. **46** (1982), no. 2, 227–253.

[38] H. Carlsson, *Remainder term estimates of the renewal function* Ann. Probab. **11** (1983), no. 1, p. 143-157.

[39] Y. Carrière *Feuilletages riemanniens à croissance polynômiale*, Comment. Math. Helv. **63** (1988), no. 1, p. 1-20.

[40] Y. Carrière, E. Ghys, *Relations d'équivalence moyennables sur les groupes de Lie*, C. R. Acad. Sci. Paris Sér. I Math. **300** (1985), no. 19, p. 677-680.

[41] Y. Carrière, F. Dal'bo, *Généralisations du premier théorème de Bieberbach sur les groupes cristallographiques*, Enseign. Math. (2) **35** (1989), no. 3-4, p. 245-262.

[42] G. Choquet, J. Deny, *Sur l'équation de convolution* $\mu = \mu * \sigma$, C. R. Acad. Sc. Paris, t. **250** (1960) p. 799-801.

[43] L. Corwin et F.P. Greenleaf, *Representations of nilpotent Lie groups and their applications*, Cambridge studies in advanced mathematics 18, CUP (1990).

[44] H. Cramér, *Les sommes et les fonctions de variables aléatoires*, Paris Hermann, (1938).

[45] H. Cramér, *Elements of probability theory and some of its applications*, Cambridge (1962).

[46] P. Crépel et A. Raugi, *Théorème central limite sur les groupes nilpotents*, Ann. Inst. H. Poincaré sect. B, Prob. & Stat., vol XIV, 2, (1978) 145-164.

[47] M. Day, *Amenable semigroups*, Illinois J. Math. **1** (1957), p.509-544

[48] T. Delzant, *Sous-groupes distingués et quotients des groupes hyperboliques*, Duke Math. J. **83** (1996), no. 3, p.661-682.

[49] J.D. Dixon, A. Mann, M.P du Sautoy, D. Segal, *Analytic pro-p groups*, London Mathematical Society Lecture Note Series, 157. CUP, Cambridge, 1991, 251 pp.

[50] J.D. Dixon, L. Pyber, A. Seress, A. Shalev, *Residual properties of free groups and probabilistic methods*, J. Reine Angew. Math. **556** (2003), p.159-172.

[51] D. Dolgopyat, *On mixing properties of compact group extensions of hyperbolic systems*, Israel J. Math. **130** (2002), p. 157–205.

[52] E. Dynkin, M. Malioutov, *Random walk on groups with a finite number of generators*, Dokl. Akad. Nauk SSSR **137** (1961) p.1042–1045.

[53] D. Epstein, *Almost all subgroups of a Lie group are free*, J. of Algebra. **19** (1971) p. 261-262.

[54] A. Eskin, G.A. Margulis, Recurrence properties of random walks on homogeneous spaces, (preprint E. Schrodinger Institute, Vienna 2001).

[55] A. Eskin, G.A. Margulis, S. Mozes, *Upper bounds and asymptotics in a quantitative version of the Oppenheim conjecture*, Ann. of Math. (2) **147** (1998), no. 1, p. 93–141.

[56] A. Eskin, S. Mozes, H. Oh, *Uniform exponential growth for linear groups*, (preprint 2003, Princeton).

[57] W. Feller, *An introduction to probability theory and its applications*, Vol. II. 2nd ed. John Wiley Sons, (1971).

[58] H. Furstenberg, *The unique ergodicity of the horocycle flow*, Recent Advances in Topological Dynamics (A. Besk ed.) Springer Verlag, 1972, p. 95-115.

[59] H. Furstenberg, *Stiffness of group actions*, in Lie groups and ergodic theory (Mumbai, 1996), 105–117, Tata Inst. Fund. Res. Stud. Math., 14, 1998.

[60] A. Gamburd, D. Jakobson, P. Sarnak, *Spectra of elements in the group ring of* $SU(2)$, J. Eur. Math. Soc. (JEMS) **1** (1999), no. 1, 51–85.

[61] T. Gelander, A. Zuk *Dependence of Kazhdan constants on generating subsets*, Israel J. Math. **129** (2002) p. 93-98,

[62] E. Ghys, *Groupes d'holonomie des feuilletages de Lie*, Nederl. Akad. Wetensch. Indag. Math. **47** (1985), no. 2, p.173-182.

[63] E. Ghys, *Riemmanian foliations : examples and problems*, Appendix to the book by P. Molino *Riemannian foliations* Birkäuser, (1988).

[64] J. Gilman, *Two-generator discrete subgroups of* $\mathrm{PSL}(2, R)$, Mem. Amer. Math. Soc. **117** (1995), no. 561, 204 pp.

[65] B. Gnedenko, A. Kolmogorov, *Limit distributions for sums of independent random variables*, Revised edition Addison-Wesley, (1968) 293 pp.

[66] R. W. Goodman, *Nilpotent Lie groups : structure and applications to analysis*, LNM 562, Springer-Verlag (1976).

[67] L. Gorostiza, *The central limit theorem for random motions of* d*-dimensional Euclidean space*, Ann. of Proba. (1973).

[68] U. Grenander, *Probabilities on algebraic structures,* John Wiley & Sons ; Almqvist & Wiksell, Stockholm-Göteborg-Uppsala (1963), 218 pp.

[69] 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.

[70] M. Gromov, *Hyperbolic groups*, in Essays in Group Theory, p. 263, Math. Sci. Res. Inst. Publ., Springer, (1987).

[71] M. Gromov, *Groups of polynomial growth and expanding maps*, Publ. Math. IHES **53** (1981).

[72] Y. Guivarc'h, *Générateurs des groupes résolubles*, Publications des Séminaires de Mathématiques de l'Université de Rennes, 1967–1968, Exp. No. 1, 17 pp., Rennes, (1968).

[73] Y. Guivarc'h, *Croissance polynomiale des groupes de Lie et périodes des fonctions harmoniques*, Bull. Soc. Math. France **101**, (1973), p. 333-379.

[74] Y. Guivarc'h, *Equirépartition dans les espaces homogènes,* Théorie ergodique (Actes Journées Ergodiques, Rennes, 1973/1974), pp. 131–142. Lecture Notes in Math., Vol. 532, Springer, Berlin, (1976).

[75] Y. Guivarc'h, M. Keane, P. Roynette, *Marches aléatoires sur les groupes de Lie,* LNM Vol. 624, Springer-Verlag, (1977).

[76] Y. Guivarc'h, *Sur la loi des grands nombres et le rayon spectral des marches aléatoires*, Conf. on Random Walks (Kleebach, 1979), 3, Astérisque, **74**, (1980) p. 47–98.

[77] Y. Guivarc'h, *Marches aléatoires sur les groupes et problèmes connexes*, Fascicule de probabilités, 39 pp., Publ. Inst. Rech. Math. Rennes, Univ. Rennes I, Rennes, (1993).

[78] Y. Guivarc'h, *Marches aléatoires sur les groupes*, Development of mathematics 1950–2000, 577–608, Birkhäuser, Basel, (2000).

[79] Y. Guivarc'h, *Limit theorems for random walks and products of random matrices*, Lectures held at the Tata Institute, Summer School in Probab. Theory, Sept. 2002.

[80] A. Haefliger, *Feuilletages Riemanniens*, Séminaire Bourbaki, Vol. 1988/89. Astérisque No. **177-178** (1989), Exp. No. 707, p. 183–197.

[81] A. Haefliger, *Groupoïdes d'holonomie et classifiants*, Transversal structure of foliations (Toulouse, 1982). Astérisque No. **116** (1984), p. 70-97.

[82] P. de la Harpe,*Topics in Geometric Group Theory*, Chicago Univerity Press, (2001).

[83] De la Harpe, P., *Free Groups in Linear Groups*, L'Enseignement Mathématique, **29** (1983), 129-144

[84] P. de la Harpe, A. Valette, *La propriété* (T) *de Kazhdan pour les groupes localement compacts, (avec un appendice de Marc Burger)*, Astérisque No. **175** (1989), 158 pp.

[85] W. Hebisch, L. Saloff-Coste, *Gaussian estimates for Markov chains and random walks on groups*, Annals of Proba. **21** (1993) p. 673-709.

[86] W. Hebisch, A. Sikora, *A smooth subadditive homogeneous norm on a homogeneous group*, Studia Math. **96** (1990), no. 3, p. 231–236.

[87] H. Heyer, *Probability measures on locally compact groups*, Ergeb. **94** Spinger Verlag (1977).

[88] K. Hinderer, *Remarks on Directly Riemann Integrable Functions*, Math. Nachr. **130** (1987) p. 225-230.

[89] T. Höglund, *A multi-dimensional renewal theorem*, Bull. Sc. Math, 2ème série, **112** (1988) p. 111-138.

[90] G.A. Hunt, *Semi-groups of measures on Lie groups*, Trans. Amer. Math. Soc. **81** (1956), p. 264–293.

[91] I. Ibragimov, *A central limit theorem for a class of dependent random variables*, Teor. Verojatnost. i Primenen, **8**, (1963), p. 89–94.

[92] J. W. Jenkins, *Growth of locally compact groups*, J. Functional Analysis **12** (1973), p. 113-127.

[93] T. Jorgensen, *On discrete groups of Möbius transformations.* Amer. J. Math. **98** (1976), no. 3, p. 739-749.

[94] W. Kantor, A. Lubotzky, *The probability of generating a finite classical group*, Geom. Dedicata **36** (1990), no. 1, p. 67-87.

[95] Y. Kawada, K. Itô, *On the probability distribution on a compact group,* I. Proc. Phys.-Math. Soc. Japan (3) **22**, (1940) p. 977-998.

[96] D.A. Kazhdan, *Uniform distribution on a plane*, Trudy Moskov. Mat. Ob. **14**, (1965) , p. 299-305.

[97] R. Keener, *Asymptotic expansions in multivariate renewal theory*, Stochastic Process. Appl. 34 (1990), no. 1, p. 137-153.

[98] H. Kesten, *Symmetric random walks on groups*, Trans. Amer. Math. Soc. **92**, (1959), p. 336-354.

[99] A. Kirillov, *Unitary representations of nilpotent Lie groups*, Uspehi Mat. Nauk **17** (1962) no. 4 (106), 57–110.

[100] A. Kirillov, *Éléments de la théorie des représentations*, Éditions Mir, Moscow, (1974) 347 pp.

[101] M. Koubi, *Croissance uniforme dans les groupes hyperboliques*, Ann. Inst. Fourier (Grenoble) **48** (1998), no. 5, p. 1441-1453.

[102] M. Kuranishi, *On everywhere dense imbedding of free groups in Lie groups.* Nagoya Math. J. **2**, (1951). p. 63-71.

[103] S. Lang, *Algebra*, Third Edition, Addison-Wesley (1994).

[104] E. Le Page, *Théorèmes quotients pour certaines marches aléatoires*, Comptes Rendus Acad. Sc., **279**, série A, n. 2, (1974).

[105] E. Le Page, *Un théorème local sur le groupe de Heisenberg*, preprint IRMAR, Univ. Rennes 1.

[106] M. Liebeck, A. Shalev, *The probability of generating a finite simple group*, Geom. Dedicata **56** (1995), no. 1, p. 103-113.

[107] A. Lubotzky, *Discrete groups, expanding graphs and invariant measures*, with an appendix by J. D. Rogawski. Progress in Math. **125**, Birkhäuser, 1994, 195 pp.

[108] A. Mann, *Positively finitely generated groups,*. Forum Math. **8** (1996), no. 4, p. 429-459.

[109] A. Mann, A. Shalev, *Simple groups, maximal subgroups, and probabilistic aspects of profinite groups*, Israel J. Math. **96** (1996), part B, p. 449-468.

[110] G.A. Margulis, *Positive harmonic functions on nilpotent groups,* Dokl. Akad. Nauk SSSR **166** (1966) p. 1054–1057

[111] G.A. Margulis, *On the action of unipotent subgroups in the space of lattices*, Mat. Sb. 86 (1971), 552-556.

[112] G.A. Margulis, *Oppenheim conjecture*, Fields Medallists' lectures, 272–327, World Sci. Ser. 20th Century Math., 5, World Sci. Publishing, River Edge, NJ, 1997

[113] G.A. Margulis, Random walks and Borel Harish-Chandra theorem, preprint.

[114] G. A. Margulis, *Discrete subgroups of semi-simple Lie groups*, Ergeb. Math. Grenz. (3) **17** Springer Verlag, (1991), 388 pp.

[115] G.A. Margulis, G.M. Tomanov, *Invariant measures for actions of unipotent groups over local fields on homogeneous spaces*, Invent. Math. **116** (1994), no. 1-3, p. 347-392.

[116] G. A. Margulis, G. A. Soifer, *Maximal subgroups of infinite index in finitely generated linear groups*, J. Algebra **69**, (1981), no. 1, p. 1-23.

[117] G. Meigniez, *Holonomy groups of solvable Lie foliations*, in *Integrable systems and foliations* (Montpellier, 1995), p. 107-146, Progr. Math., **145**, Birkhäuser, (1997).

[118] R. Michel, *Results on probabilities of moderate deviations*, Ann. Prob., no. 2 (1974), p. 349-353.

[119] J. Milnor, *Growth of finitely generated solvable groups*, J. Diff. Geometry **2** (1968) p. 447-449.

[120] P. Molino, *Riemannian foliations*, Prog. in Math. **73**, Birkhäuser, (1988), 339 pp.

[121] N. Monod, *Continuous bounded cohomology of locally compact groups*, LNM **1758**, Springer Verlag, (2001), 214 pp.

[122] L. Mosher, *Indiscrete representations, laminations, and tilings*, Geometric group theory down under (Canberra, 1996), p. 225-259, de Gruyter, (1999).

[123] D. Montgomery, L. Zippin, *Topological transformation groups*, Reprint of the 1955 original. R. E. Krieger Publ. Co., (1974) 289 pp.

[124] A. V. Nagaev, *Renewal theorems in* $\mathbb{R}^d$, Teor. Veroyatnost. i Primenen. **24** (1979), no. 3, p. 565-573.

[125] A. V. Nagaev, *Integral limit theorems with regards to large deviations when Cramér's condition is not satisfied*, Prob. Th. Appl. 14 (1969), p.51-63

[126] A. V. Nagaev, *Large deviations of sums of independant random variables*, Ann. of Prob. **7**, no 5., (1979), p. 745-789.

[127] D. Neuenschwander, *Probabilities on the Heisenberg group*, LNM 1630, Springer-Verlag (1996).

[128] A. Nevo, preprint quoted in [12].

[129] A. Yu. Olshanskii, *On the question of the existence of an invariant mean on a group*, Uspekhi Mat. Nauk **35** (1980), no. 4 (214), p. 199-200.

[130] V.I. Oseledec, *Markov chains, skew-products, and ergodic thoerems for "general" dynamical systems*, Th. Prob. & App. (1965), **10**, 3, 551-557.

[131] D. Osin, *Exponential growth, solvable groups*, Erg. Theory. Dyn. Sys. **23**, no. 3, (2003).

[132] V. Platonov, A. Rapinchuk, *Algebraic Groups and Number Theory*, Academic Press, (1994).

[133] G. Pólya, *Beitrag zur Verallgemeinerung des Verzerrungssatzes auf mehrfach Zusammenhangende Gebiete*, S.B. Preuss. Akad. Wiss., Berlin, K.L. Math. Phys. Tech. (1928), p. 228-232 and p. 280-282.

[134] H. Poincaré, Calcul des Probabilités, 2ème éd., Gauthier Villars, (1912)

[135] L. Pyber, *Group enumeration and where it leads us*, Prog. in Math. Vol 169, Birkhaüser.

[136] J. F. Quint, *Sous-groupes discrets des groupes semi-simples*, thèse Paris VII, 2001.

[137] M. S. Raghunathan, *Discrete Subgroups of Discrete Groups*, Springer Verlag (1972).

[138] M. Ratner, *Interaction between Ergodic Theory, Lie Groups, and Number Theory*, in Proc. ICM (Zurich, 1994) 157-182, Birkhaüser, 1995.

[139] M. Ratner, *On Raghunathan's measure conjecture*, Ann. Math. **134** (1991), 545-607.

[140] M. Ratner, *Raghunathan's topological conjecture and distributions of unipotent flows*, Duke Math. J. **63** (1991), 235-290.

[141] M. Ratner, *Interaction between Ergodic Theory, Lie Groups, and Number Theory*, in Proc. ICM (Zurich, 1994) 157-182, Birkhaüser, 1995.

[142] A. Raugi, *Théorème de la limite centrale sur les groupes Nilpotents,* Z. Wahrsch. Verw. Gebiete **43** (1978), no. 2, 149–172.

[143] A. Raugi, *A general Choquet-Deny theorem for a large class of locally compact second countable groups*, preprint IRMAR, Univ. Rennes 1, (2002).

[144] L. Ribes, P. Zalevsskii, *Profinite groups*, Ergeb. Math. vol. **40**, Springer Verlag (2000).

[145] H. Rubin, J. Sethuraman, *Probabilities of moderate deviations*, Sankhyā Ser. A, **27** (1965), p. 325-346.

[146] W. Rudin, *Real and Complex Analysis*, Addison-Wesley.

[147] P. Sarnak, *Some applications of modular forms*, Cambridge Tracts in Mathematics, **99** CUP, Cambridge, (1990)

[148] P. Scott, *Subgroups of surface groups are almost geometric*, J. London Math. Soc. (2) 17 (1978), no. 3, p. 555–565 and erratum J. London Math. Soc. (2) 32 (1985), no. 2, p. 217–220.

[149] J.P. Serre, *Lie algebras and Lie groups*, Springer LNM 1500.

[150] N. Shah, *Limit distributions of polynomial trajectories on homogeneous spaces*, Duke Math. J. **75**, 3, (1994), 711-732.

[151] N. Shah, *Invariant measures and orbit closures on homogeneous spaces for actions of subgroups generated by unipotent elements.* Lie groups and ergodic theory (Mumbai, 1996), 229–271, Tata Inst. Fund. Res. Stud. Math., **14**, Bombay, 1998

[152] A. Shalev, *Simple groups, permutation groups, and probability*, Proc. of the Inter. Cong. of Math., Vol. II (Berlin, 1998), Doc. Math. 1998, Extra Vol. II, p. 129-137.

[153] A. Shalev, *Lie methods in the theory of pro-p groups*, in *New horizons in pro-p groups*, p. 1-54, Progr. Math., **184**, Birkhäuser (2000).

[154] Y. Shalom, *The growth of linear groups*, J. Algebra **199** (1998), no. 1, p. 169-174.

[155] G. Soifer, T. Venkataramana, *Finitely generated pro-finitely dense free groups in higher rank semi-simple groups*, Transform. Groups **5** (2000), no. 1, p. 93-100.

[156] E. Siebert, *Absolute continuity, singularity, and supports of Gauss semigroups on a Lie group*, Monat. für Math. **93**, (1982), p. 239-253.

[157] E. Siebert, *Densities and differentiability properties on Gauss semigroups on a Lie group*, Proc. Amer. Math. Soc. vol **91**, n. 2, (1984), p.298-305.

[158] A. D. Slastnikov, *Limit theorems for moderate deviations probabilities*, Probab. Theo. and Appl. **23** (1978), no2., p. 325-340.

[159] A. J. Stam, *Renewal theory in r dimensions I.* Compositio Math. **21** (1969) p. 383-399 and **23** (1971) p. 1-13.

[160] A. N. Starkov, *Dynamical systems on homogeneous spaces*, Transl. of Math. Monographs, **190**, AMS Prov. RI, (2000).

[161] Ch. Stone, *A local limit theorem for nonlattice multi-dimensional distribution functions*, Ann. Math. Stat, **36** (1965) 546-551.

[162] Ch. Stone, *Ratio limit theorems for random walks on groups*, Trans. Amer. Math. Soc. **125** (1966) p. 86–100.

[163] Ch. Stone, *Ratio local and ratio limit theorems*, 5th Berkeley Symp. on Math. Stat. and Prob., Berkeley and Los Angeles, UCP (1966), Vol II, 2, p. 217-224

[164] Ch. Stone, *Application of unsmoothing and Fourier Analysis to random walks*, Markov Processes and Potential Theory, Madison Wis, (1967), p. 165-192.

[165] D. Stroock, S.R.S Varadhan, *Limit theorems for random walks on Lie groups*, Indian J. of Stat. Sankhyā, Ser. A, **35**, (1973) p. 277-294.

[166] B. Szegedy, Communication at the conference “Groups and Probability”, Budapest July 2003.

[167] J. Tits, *Free subgroups in linear groups*, J. Algebra **20** (1972) p. 250-270.

[168] J. Tits, *Représentations linéaires irréductibles d'un groupe algébrique sur un corps quelconque*, J. Reine Angew. Math. **247** (1971) p. 196-220.

[169] V.N. Tutubalin, *Compositions of measures on the simplest nilpotent group*, Teor. Verojatnost. i Primenen **9**, 1964, p. 531-539.

[170] N. Varopoulos, *Wiener-Hopf theory and nonunimodular groups*, J. Funct. Anal., **120,** (1994) p.467-483.

[171] N. Varopoulos, L. Saloff-Coste, et Th. Coulhon, *Analysis and geometry on groups*, Cambridge Tracts in Mathematics, **100**, CUP, Cambridge, 1992.

[172] D. Wehn, *Probabilities on Lie groups*, Proc. Nat. Acad. Sci. USA, vol **48**, (1962) p.791-795

[173] D. Wehn, *Limit distributions on Lie groups*, Yale thesis, 1960.

[174] D. Wehn, *Some remarks on Gaussian Distributions on a Lie Group*, Z. Wahrsch. Geb. **30**, p. 255-263.

[175] W. Woess, *Random walks on infinite graphs and groups*, Cambridge Tracts in Mathematics, **138**, CUP, Cambridge, (2000)

[176] J. Wolf, *Growth of finitely generated solvable groups and curvature of Riemanniann manifolds*, J. Diff. Geom. **2** (1968), p. 421-446.

[177] J. Wilson, *On uniform exponential growth for solvable groups* (preprint 2000).

[178] R. Zimmer, *Amenable actions and dense subgroups of Lie groups*, J. Funct. Anal. **72** (1987), no. 1, p. 58-64.

[179] R. Zimmer, *Ergodic theory and semi-simple groups*, Monographs in Math., **81**, Birkhäuser Verlag, (1984), 209 pp.

[180] R. Zimmer, *Amenable ergodic group actions and an application to Poisson boundaries of random walks*, J. Funct. Anal. **27** (1978), no. 3, p. 350-372.

[181] R. Zimmer, *Kazhdan groups acting on compact manifolds*, Invent. Math. **75** (1984), no. 3, p. 425-436.

[182] R. Zimmer, *Amenable pairs of groups and ergodic actions and the associated von Neumann algebras*, Trans. Amer. Math. Soc. 243 (1978), 271–286.

Résumé : La thèse comprend deux parties relativement indépendantes. La première, plus probabiliste, traite des marches aléatoires sur les groupes de Lie et en particulier des problèmes d'équirépartition des marches aléatoires après un temps très long. Le chapitre 2 est consacré à l'étude de l'équirépartition des marches symétriques à support fini dans les groupes de Lie nilpotents. Au chapitre 3, on démontre un théorème limite local pour les produits de matrices aléatoires sur le groupe de Heisenberg et un théorème de comparaison avec le noyau de la chaleur associé. On en déduit un équivalent probabiliste du théorème d'équirépartition de Ratner pour les marches aléatoires unipotentes dans les espaces homogènes. Le chapitre 4 est indépendant et entièrement consacré au théorème limite local sur $\mathbb{R}^d$ et sa vitesse de convergence.

La deuxième partie, plus algébrique, traite des sous-groupes denses et libres des groupes de Lie réels et p-adiques. On démontre une version topologique de l'alternative de Tits qui affirme que tout sous-groupe de $GL(n,k)$, pour un corps local k, admet ou bien un sous-groupe résoluble relativement ouvert, ou bien un sous-groupe libre relativement dense. On illustre ensuite ce théorème par diverses applications aux théories des groupes profinis, des actions moyennables et des feuilletages Riemanniens.

Abstract : This dissertation consists of two relatively independent parts. The first part, more probabilistic in nature, deals with random walks on Lie groups and especially with equidistribution properties of random walks after a very large time. Chapter 2 is devoted to the study of equidistribution of finitely supported symmetric walks on nilpotent Lie groups. In Chapter 3, we prove a local limit theorem for product of random matrices on the Heisenberg group and a comparison theorem with the associated heat kernel. We then derive a probabilistic equivalent of Ratner's equidistribution theorem for unipotent random walks on homogeneous spaces. Chapter 4 is independent and entirely devoted to the local limit theorem on $\mathbb{R}^d$ with a study of the speed of convergence.

The second part, of a more algebraic flavor, deals with dense free subgroups of real and p-adic Lie groups. We show a topological version of Tits' alternative asserting that any subgroup of $GL(n.k)$, where k is a local field, contains either a relatively open solvable subgroup, or a relatively dense free subgroup. We then provide several applications of this theorem to the theory of profinite groups, of amenable actions and of Riemannian foliations.