université PARIS-SACLAY | École doctorale de mathématiques Hadamard (EDMH)

UNIVERSITÉ PARIS SUD — FACULTÉ DES SCIENCES D'ORSAY

THÈSE DE DOCTORAT

de

L'UNIVERSITÉ PARIS-SACLAY

École doctorale de mathématiques Hadamard (EDMH, ED 574)

Établissement d'inscription :
Université Paris-Sud

Laboratoire d'accueil :
Laboratoire de mathématiques d'Orsay, UMR 8628 CNRS

Spécialité de doctorat :
Mathématiques fondamentales

Çağrı SERT

Joint Spectrum and Large Deviation Principles for Random Matrix Products

Après avis des rapporteurs :
YVES GUIVARC'H (Université de Rennes 1)
JEAN-FRANÇOIS QUINT (Université Bordeaux 1)

Jury de soutenance :

PHILIPPE BOUGEROL	(Université Paris 6)	Examinateur
EMMANUEL BREUILLARD	(Université Paris 11)	Directeur de thèse
YVES GUIVARC'H	(Université Rennes 1)	Rapporteur
JEAN-FRANÇOIS LE GALL	(Université Paris 11)	Examinateur
FRÉDÉRIC PAULIN	(Université Paris 11)	Examinateur
MARC PEIGNÉ	(Université de Tours)	Examinateur
JEAN-FRANÇOIS QUINT	(Université Bordeaux 1)	Rapporteur

RDMath IdF — Domaine d'Intérêt Majeur (DIM) en Mathématiques — îledeFrance

MATHEMATIQUES ORSAY

Fondation mathématique FMJH Jacques Hadamard

Résumé

Après une introduction et la présentation d'un exemple explicite pour illustrer notre étude dans le chapitre 1, nous exposons certains outils et techniques généraux dans le chapitre 2. Ensuite,

- Dans le chapitre 3, nous démontrons l'existence d'un principe de grandes déviations (PGD) avec une fonction de taux convexe, pour les composantes de Cartan le long des marches aléatoires sur les groupes linéaires semisimples G. L'hypothèse principale porte sur le support S de la mesure de la probabilité en question et demande que S engendre un semi-groupe Zariski dense.
- Dans le chapitre 4, nous introduisons un objet limite (une partie de la chambre de Weyl) que l'on associe à une partie bornée S de G et que nous appelons le spectre joint $J(S)$ de S. Nous étudions ses propriétés et démontrons que $J(S)$ est une partie convexe compacte d'intérieur non-vide dès que S engendre un semi-groupe Zariski dense. Nous relions le spectre joint avec la notion classique du rayon spectral joint et la fonction de taux du PGD pour les marches aléatoires sur G.
- Dans le chapitre 5, nous introduisons une fonction de comptage exponentiel pour un S fini dans G, nous étudions ses propriétés que nous relions avec $J(S)$ et démontrons un théorème de croissance exponentielle dense.
- Dans le chapitre 6, nous démontrons le PGD pour les composantes d'Iwasawa le long des marches aléatoires sur G. L'hypothèse principale demande l'absolue continuité de la mesure de probabilité par rapport à la mesure de Haar.
- Dans le chapitre 7, nous développons des outils pour aborder une question de Breuillard sur la rigidité du rayon spectral d'une marche aléatoire sur le groupe libre. Nous y démontrons un résultat de rigidité géométrique.

Abstract

After giving a detailed introduction and the presentation of an explicit example to illustrate our study in Chapter 1, we exhibit some general tools and techniques in Chapter 2. Subsequently,

- In Chapter 3, we prove the existence of a large deviation principle (LDP) with a convex rate function, for the Cartan components of the random walks on linear semi-simple groups G. The main hypothesis is on the support S of the probability measure in question, and asks S to generate a Zariski dense semigroup.
- In Chapter 4, we introduce a limit object (a subset of the Weyl chamber) that we associate to a bounded subset S of G. We call this the joint spectrum $J(S)$ of S. We study its properties and show that for a subset S generating a Zariski dense semi-group, $J(S)$ is convex body, i.e. a convex compact subset of non-empty interior. We relate the joint spectrum with the classical notion of joint spectral radius and the rate function of LDP for random walks on G.
- In Chapter 5, we introduce an exponential counting function for a finite S in G. We study its properties, relate it to joint spectrum of S and prove a dense exponential growth theorem.
- In Chapter 6, we prove the existence of an LDP for Iwasawa components of random walks on G. The hypothesis asks for a condition of absolute continuity of the probability measure with respect to the Haar measure.
- In Chapter 7, we develop some tools to tackle a question of Breuillard on the rigidity of spectral radius of a random walk on a free group. We prove a weaker geometric rigidity result.

Remerciements

Ceci est tout d'abord une occasion d'exprimer ma reconnaissance envers mon professeur de thèse, Emmanuel Breuillard. Sa patience et sa bienveillance m'ont permis d'avoir de très favorables conditions d'études et particulièrement, une grande liberté mathématique. Chacune de nos discussions a été, pour moi, une source avant tout d'inspiration et de motivation, non seulement par le contenu, mais aussi par ce qu'elle m'a permis de percevoir son attitude vis-à-vis les mathématiques. Je salue son effort, lors de ces discussions, de me transmettre son art et de me faire sentir l'odeur de la recherche mathématique. Ainsi, ces années de travail avec lui, ont été les plus enrichissantes de mes études supérieures.

En deuxième lieu, je remercie sincèrement à mes rapporteurs de thèse, Yves Guivarc'h et Jean-François Quint, d'avoir accepté cette besogne de lire en détail mon premier écrit mathématique. Je remercie de surcroît à J.-F. Quint de son soin et de son invitation à Bordeaux, qui m'ont encouragé à améliorer divers points de ce texte. Cela est une heureuse "coïncidence" (les événements n'ont pourtant pas l'air indépendants !) que mes rapporteurs sont des mathématiciens dont la leçon de mathématiques au M1 (J.-F. Quint) sur les systèmes dynamiques dans les espaces homogènes a joué un rôle dans mon orientation mathématique et dont les articles (Y. Guivarc'h et J.-F. Quint) m'ont appris une partie importante du sujet de ce texte.

Je serai honoré par la présence de chaque membre de jury : P. Bougerol, E. Breuillard, Y. Guivarc'h, J.-F. Le Gall, F. Paulin, M. Peigné, J.F.-Quint. Je leur sais gré d'avoir accepté d'y être. Cela est encore une belle coïncidence d'avoir parmi eux les deux mathématiciens dont les polycopiés (Intégration, Probabilités et Processus aléatoires de J.-F. Le Gall et Topologie, analyse et calcul différentiel de F. Paulin) ont été mes livres de référence au tout début de mon séjour en France, et P. Bougerol, dont le joli livre "Products of Random Matrices" m'a appris une bonne partie de cette théorie.

Pendant la réalisation de cette thèse, suite à l'invitation de E. Breuillard, j'ai passé sept mois sous des conditions très favorables dans Universität Münster. Je remercie pour l'hospitalité de cette dernière.

Je salue aussi tous mes collègues de l'Université Paris-Sud, d'Universität Münster et d'ailleurs, dont la présence a été préciseuse durant ces années. Un merci particulier à mes «cobureaux» de 112 et 114 du bâtiment 430.

Un merci à Pierre Pansu, jadis directeur des études de DMA/ENS, qui a joué un rôle en ce que mon agréable séjour en France (initialement prévu pour six mois) dépasse maintenant six ans.

En les personnes de la Fondation Sciences Mathématiques de Paris et de la Fondation Mathématique Jacques Hadamard, j'exprime ma reconnaissance à tous, au sens large, dont l'organisation a rendu possible la réalisation de cette thèse, et ceci sous des conditions propices.

Finalement, un merci à ma famille (Nevin, Yasemin, Yüksel, oncles Hakan, Otobüs et les autres), et à Céline (Baykuş) et à Lisa pour leur soutien et de bons moments au long de ces années.

Table des matières

Première partie

Large Deviation Principles For Random Matrix Products And The Joint Spectrum

Chapitre 1

INTRODUCTION

1. Soient G un groupe (algébrique) réel linéaire semisimple connexe (comme $\mathrm{SL}(d, \mathbb{R})$) et S une partie de G. Dans ce texte, on étudie l'asymptotique de la suite $S, S^2, S^3, \ldots$ des parties de G, où pour $n \in \mathbb{N}$, S^n est la partie $\{g_1 . \ldots . g_n \mid g_i \in S\}$. Les propriétés de cette suite seront étudiées via des décompositions classiques du groupe G, les composants desquelles donnent de l'information sur l'asymptotique dans G.

2. Introduisons tout d'abord l'une de ces décompositions classiques, la décomposition de Cartan, et l'application de «projection», qui en résulte et qui sera d'un intérêt central pour nous : dans le cas de $\mathrm{SL}(d, \mathbb{R})$, elle correspond à la décomposition polaire selon laquelle un élément g dans $\mathrm{SL}(d, \mathbb{R})$ s'écrit $g = k_1 a k_2$, où $k_1, k_2 \in \mathrm{SO}(d, \mathbb{R})$ et a est une matrice diagonale avec des coefficients strictement positifs. Contrairement à k_1 et k_2, les coefficients de a sont déterminés de manière unique (à l'ordre près) et s'appellent les valeurs singulières de g. Pour un groupe G plus général comme ci-dessus, on peut écrire $G = K \exp(\mathfrak{a}^+) K$, où K est un sous-groupe compact maximal et $\mathfrak{a}^+$ est une chambre de Weyl choisie de G. L'application $\kappa : G \longrightarrow \mathfrak{a}^+$ qui associe à un élément g de G l'unique a_g dans $\mathfrak{a}^+$ tel que $g \in K \exp(a_g) K$ s'appelle la projection de Cartan. Pour $g \in \mathrm{SL}(d, \mathbb{R})$, on peut écrire $\kappa(g) = (\log ||g||, \log \frac{||\wedge^2 g||}{||g||}, \ldots, \log \frac{||\wedge^d g||}{||\wedge^{d-1} g||})$, où $||.||$ note la norme d'opérateur.

3.1 Dans le Chapitre 3, pour notre étude, on adopte une approche probabiliste et on consière les produits aléatoires indépendants des éléments de S suivant une loi μ de support S. Ceci s'appelle une marche aléatoire dans le groupe G, et plus précisément, nous étudierons des *événements rares*, i.e. des grandes déviations, sur ces marches aléatoires.

3.2 Dans cette perspective probabiliste, commençons par décrire le résultat fondamental de Furstenberg-Kesten [59] (plus tard, un corollaire du théorème ergodique sous-additif de Kingman), que l'on peut considérer comme une première version non-commutative de la loi des grands nombres. Pour l'énoncer, notons G le groupe $\mathrm{SL}(d, \mathbb{R})$ ou $\mathrm{GL}(d, \mathbb{R})$ et soit μ une mesure de probabilité borélienne sur G. On dit que μ a un moment d'ordre un fini, si, en posant $N(g) = \max\{||g||, ||g^{-1}||\}$, on a $\int \log N(g)\, \mu(dg) < \infty$. En notant par S_n le $n^{\text{ième}}$ pas de la μ-marche aléatoire dans G (autrement dit, $S_n = X_n . \ldots . X_1$, où les X_i sont des variables aléatoires indépendantes et identiquement distribuées de loi μ), le théorème de Furstenberg-Kesten affirme que, presque sûrement, la norme moyenne $\frac{1}{n}\log||S_n||$ du produit aléatoire S_n, converge vers une constante $\lambda_1(\mu)$ que l'on appelle le premier exposant de Lyapunov de μ. On peut considérer que cette propriété et la constante $\lambda_1(\mu)$ donnent une première information asymptotique sur les projections de Cartans des produits (aléatoires) des éléments de S, dépendant, bien entendu, de μ.

3.3 Dans notre premier résultat, on obtient une version non-commutative du résultat classique de Cramér [41] sur les probabilités de grandes déviations. Compte tenu du paragraphe précédent, on verra que notre résultat a la même correspondance avec celui de Cramér, que possède le théorème de Furstenberg-Kesten avec la loi des grands nombres : soient X un espace topologique et $\mathcal{F}$ une tribu sur X ; on rappelle qu'une suite μ_n de mesures de probabilités (ou de manière équivalente, une suite de variables aléatoires Z_n de lois μ_n) sur $(X,\mathcal{F})$ est dite satisfaire un principe de grandes déviations (PGD) avec la fonction de taux/profil $I : X \longrightarrow [0,\infty]$, si pour toute partie R mesurable de X, on a

$$-\inf_{x\in\overset{\circ}{R}} I(x) \le \liminf_{n\to\infty}\frac{1}{n}\log\mu_n(R) \le \limsup_{n\to\infty}\frac{1}{n}\log\mu_n(R) \le -\inf_{x\in\overline{R}} I(x)$$

Le théorème de Cramér affirme que la suite des moyennes $S_n = \frac{1}{n}\sum_{i=1}^n X_i$ des variables aléatoires X_i indépendantes et identiquement distribuées à valeurs dans $\mathbb{R}$ et de moment exponentiel fini, satisfait un PGD avec une fonction de taux I propre et convexe, qui est donnée par la transformée de Fenchel-Legendre du logarithme de la fonction génératrice des moments des X_i. On précise que dans le cas précédent de $\mathrm{SL}(d,\mathbb{R})$, un moment exponentiel fini pour μ signifie qu'il existe une constante $C > 0$ telle que $\int N(g)^c\mu(dg) < \infty$ (pour le cas général, voir la section 3.3). Notre premier résultat est le suivant

Théorème 1.1. (Version simplifiée) *Soit G un groupe (algébrique) réel linéaire semisimple connexe, μ une mesure de probabilité de moment exponentiel fini sur G, dont le support engendre un semi-groupe Zariski dense dans G. Alors, la suite des variables aléatoires $\frac{1}{n}\kappa(S_n)$ satisfait un PGD avec une fonction de taux propre et convexe.*

Remarque 1.2. *1. On obtient aussi un analogue de la généralisation du théorème de Cramér par Bahadur [8] sans condition de moment sur la mesure de probabilité μ (voir le théorème 3.1).*

2. Dans le cas où l'on possède d'une condition de moment plus fort (voir la section 3.3), en exploitant la convexité, on arrive à identifier la fonction de taux I avec la transformée de Fenchel-Legendre d'une fonction apparaissant comme limite des logarithmes des fonctions génératrices des moments (voir le théorème 3.2).

3. Avec une étude détaillée du support effectif $D_I := \{x \in \mathfrak{a} \mid I(x) < \infty\}$ de la fonction de taux I et en utilisant la convexité, on démontre aussi l'existence des limites pour les probabilités de grandes déviations dans des parties R suffisamment régulières de $\mathfrak{a}$ (voir le corollaire 4.24).

3.4 Disons aussi quelques mots sur l'hypothèse principale de notre théorème, portant sur le semi-groupe engendré par le support de la mesure de probabilité et mentionnons quelques antécédents de notre résultat : le théorème de Furstenberg-Kesten a été suivi par des résultats remarquables de Furstenberg ([58], [57], [56] ...) sur les produits aléatoires des matrices et, plus généralement, sur les marches aléatoires dans les groupes. En particulier, il a, par exemple, considéré les variables $\log ||S_n v||$ avec $v \in \mathbb{R}^d \setminus \{0\}$: en réalisant ces dernières comme des fonctionnelles d'une chaîne de Markov sur un espace d'états comprenant l'espace projetif $\mathbb{P}(\mathbb{R}^d)$ et en accomplissant une étude fine de l'action projective (aléatoire), il a obtenu des résultats cruciaux sur le comportement asymptotique des produits aléatoires de matrices (par exemple, sur l'exposant de Lyapunov $\lambda_1(\mu)$ introduit en-haut). Concernant les extensions des théorèmes limites classiques aux produits aléatoires des matrices, un autre pas important a été fait par Tutubalin ([110], ...) qui, en exploitant encore la structure markovienne mise en avant par Furstenberg, est parvenu à appliquer les méthodes générales de Nagaev ([90]) reposant sur l'étude spectrale des opérateurs, et a établi, par exemple, un théorème central limite (TCL) pour les variables $\log ||S_n v||$. Néanmoins, pour pouvoir appliquer ces méthodes, Tutubalin a introduit des hypothèses restrictives sur la mesure de probabilité μ, notamment l'absolue continuité de cette dernière par rapport à la mesure de Haar.

Le deuxième pas fondamental après Furstenberg a été fait dans les années 80 par l'école française : Le Page [84], Guivarc'h [64] [65], Raugi [67], Bougerol [30], ..., qui ont été capables d'étendre des résultats initiaux de Furstenberg sur la contraction exponentielle et les mesures stationnaires sur les espaces projectifs à un cadre plus général et, plus remarquablement, à des mesures de probabilité d'une généralité considérablement plus grande. Dans ce cas général, par exemple, Le Page a obtenu un résultat de trou spectral, et conséquemment par la méthode de Nagaev-Guivarc'h, a établi les versions non-commutatives de plusieurs théorèmes limites classiques (TCL, loi du logarithme itéré, et par exemple, le théorème 1.3 en-bas). Ladite généralité des mesures de probabilité contient seulement des conditions algébriques, notamment celle de proximalité (contraction) et irréductibilité forte (pour les définitions, voir la section 4.2), sur le semi-groupe Γ engendré par le support de la mesure en question. Parmi ces extensions des théorèmes limites classiques, le théorème suivant de Le Page est un premier résultat sur les estimées de grandes déviations et très remarquablement, il établit la décroissance exponentielle des probabilités de grandes déviations :

Théorème 1.3. (Le Page [84]) *Soit μ une mesure de probabilité de moment exponentiel fini sur $\mathrm{GL}(d,\mathbb{R})$ dont le support engendre un semi-groupe contractant et fortement irréductible. Alors, il existe une constante $B > 0$ telle que pour tout vecteur $x \in \mathbb{R}^d$ non nul et tout $0 < \epsilon < B$, on a*

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(\log ||S_n x|| - n\lambda_1(\mu) > n\epsilon) = \phi(\epsilon) \tag{1.1}$$

où $\phi : [0, B[\to \mathbb{R}$ est une fonction concave telle que pour tout $\epsilon \in]0, B[$, $\phi(\epsilon) < 0$ et $\phi(0) = 0$ (on a aussi le résultat correspondant pour $\mathbb{P}(\log ||S_n x|| - n\lambda_1(\mu) < -n\epsilon)$).

À ce stade, on tient à mentionner le fait que le résultat de décroissance exponentielle de ce théorème, et ses généralisations par Benoist-Quint [14], [15], ont permis à Bourgain-Furman-Lindenstrauss-Mozes [32] et Benoist-Quint [21], [22], [23] d'obtenir des théorèmes surprenants sur les actions des sous-groupes des groupes de Lie et leurs espaces homogènes. On utilisera aussi ce phénomène de décroissance pour décrire l'ensemble de zéros de la fonction de taux I apparaissant dans notre théorème en haut.

En continuant notre description, on indique que certains de ces résultats pour les produits aléatoires de matrices avaient été obtenus dans une plus grande généralité, notamment celle des groupes linéaires semi-simples, suivant encore les travaux initiaux de Furstenberg, par Raugi, Guivarc'h, Goldsheid et Margulis [61]. Ce dernier travail de Goldsheid-Margulis a remarquablement clarifié l'essence algébrique des dites hypothèses sur le support de la mesure de probabilité, mettant en évidence le rôle de l'adhérence de Zariski du semi-groupe engendré par le support. Notre hypothèse principale, la densité au sens de Zariski de Γ, apparaît parallèlement, notamment en sa relation avec la notion de proximalité. Mais nous tirons parti aussi de, et construisons sur, les ultérieurs résultats et constructions algébriques d'Abels-Margulis-Soifer [16] et de Benoist [10], [11] sur les éléments loxodromiques et les semi-groupes de Schottky. On voudrait finalement mentionner que récemment, la théorie des produits aléatoires des matrices et des marches aléatoires sur les groupes semi-simples et réductifs a été réétudiée en profondeur par Benoist-Quint [14], qui ont aussi utilisé, en partie, des méthodes différentes pour établir des théorèmes limites. Une conséquence concrète de l'utilisation de ces différentes méthodes est la démonstration par Benoist-Quint [15], du TCL sous condition de moment L^2, à la place de l'hypothèse précédente de moment exponentiel fini, de, par exemple, Le Page.

3.5 Pour terminer la discussion de ce chapitre, on souhaite énoncer notre résultat suivant comme un théorème à part pour le lecteur qui serait plus intéressé par les produits aléatoires des matrices. On souligne que quand G est un groupe semi-simple comme avant et que V est un G-module irréductible, ce résultat découle immédiatement du théorème précédent en utilisant une technique générale dans la théorie de grandes déviations, le principe de contraction (voir le lemme 3.29). La généralité supplémentaire dans l'affirmation ci-dessous ne cause pas de difficulté particulière et est traitée dans la section 3.4.

Théorème 1.4. (Version simplifiée) *Soit μ une mesure de probabilité de moment exponentiel fini sur $\mathrm{GL}(V)$ telle que l'adhérence de Zariski du semi-groupe engendré par le support de μ soit une extension centrale et triviale d'un groupe algébrique réel linéaire semi-simple dans $\mathrm{GL}(V)$ (e.g. $GL(V)$). Alors, la suite des variables aléatoires $\frac{1}{n}\log||S_n||$ satisfait un PGD avec une fonction de taux propre et convexe.*

4.1 Dans cette introduction au chapitre 4, soit S une partie bornée d'un groupe G (algébrique) réel linéaire semi-simple connexe. Cette fois-ci, on entreprend une étude déterministe de la suite $S, S^2, \ldots$, encore une fois via les décompositions classiques de G, comme celles de Cartan et de Jordan. En résumé, à toute telle partie S de G, on associe une partie compacte de la sous-algèbre $\mathfrak{a}$ de l'algèbre de Lie $\mathfrak{g}$ de G, que l'on appelle le spectre joint de S. Celle-ci encode les comportements asymptotiques des projections de Cartan et de Jordan des éléments de S^n. En utilisant, en partie, cette notion de spectre joint, on effectue aussi une étude détaillée de la fonction de taux apparue dans le théorème 1.1. Avant d'expliciter ceux-ci et énoncer nos résultats, commençons par mentionner quelques notions proches, afin de mettre nos résultats en perspective.

4.2 Rappelons d'abord la décomposition de Jordan et l'application de «projection »correspondante pour un tel groupe G. La décomposition de Jordan affirme que tout élément $g \in G$ peut être factorisé de manière unique en un produit $g = g_e g_h g_u$ d'un élément elliptique g_e, hyperbolique g_h et unipotent g_u, commutant entre eux. L'élément hyperbolique g_h est conjugué à un unique élément de type $\exp(x_g)$ où x_g appartient à une chambre de Weyl choisie $\mathfrak{a}^+$. L'application $\lambda : G \longrightarrow \mathfrak{a}^+$, qui associe ce $x_g \in \mathfrak{a}^+$ à g s'appelle la projection de Jordan. Dans le cas particulier où l'on a $G = \mathrm{SL}(d, \mathbb{R})$, la projection de Jordan $\lambda(g)$ correspond aux valeurs propres de g ; plus précisément, pour un $g \in G$, en notant par $\lambda_i(g)$ pour $i = 1, \ldots, d$ les valeurs propres de g telles que $|\lambda_1(g)| \geq \ldots \geq |\lambda_d(g)| > 0$, on peut écrire $\lambda(g) = (\log|\lambda_1(g)|, \ldots, \log|\lambda_1(g)|)$.

Par ailleurs, on rappelle la notion de rayon spectral joint, que l'on note $r(S)$, d'une partie bornée S d'une algèbre normée A introduite par Rota-Strang dans [114] comme $r(S) := \lim_{n\to\infty} \sup\{||x||^{\frac{1}{n}} \mid x \in S^n\}$. Dans le cas où S est un singleton, $r(S)$ est bien-entendu le rayon spectral de l'élément correspondant par la formule classique de rayon spectral. Contrairement au cas d'un singleton, pour une partie S arbitraire bornée, il existe d'autres valeurs numériques, comme par exemple le rayon sous-spectral joint $r_{sub}(S) := \lim_{n\to\infty} \inf\{||x||^{\frac{1}{n}} \mid x \in S^n\}$, qui expriment l'asymptotique en norme de la suite $S, S^2, \ldots$. Ces caractéristiques numériques ont été étudiées par plusieurs auteurs, on envoie le lecteur à [114], [43], [24], [31], [25], ...

4.3 Dans ce travail, pour commodité, on se restreint à un groupe G comme précédemment, et à une partie bornée S de G que l'on suppose engendrer un semi-groupe Zariski dense dans cette introduction. On définit le spectre joint $J(S)$ de S comme la limite pour la distance de Hausdorff de la suite $K_n(S) = \{\frac{\kappa(g)}{n} \mid g \in S^n\} \subset \mathfrak{a}^+$. Ce faisant, on unifie et étend les dites caractéristiques numériques à une partie compacte $J(S)$ de $\mathfrak{a}^+$, qui, comme on le verra, contient clairement plus d'information sur l'asymptotique de la suite $S, S^2, \ldots$. L'auteur tient à signaler que cette notion lui a été suggérée par Emmanuel Breuillard et qu'elle apparaît aussi en fort lien avec nos considérations de grandes déviations de la partie précédente (voir le théorème 1.7). On voudrait aussi souligner que la notion du spectre joint peut être définie de même manière dans une plus grande généralité, notamment pour une partie bornée d'une algèbre de matrices sur un corps local.

Avant d'énoncer notre résultat sur le spectre joint, rappelons un dernier invariant numérique pour une partie S bornée de l'algèbre $M_n(\mathbb{C})$: celui-ci est le rayon spectral généralisé de S et est liée aux valeurs propres des éléments de S^n. Elle a été introduite par Daubechies-Lagarias dans [43] comme

$$\rho(S) = \limsup_{n\to\infty} \sup\{\lambda_1(x)^{\frac{1}{n}} \mid x \in S^n\}$$

Conjecturé par Daubechies-Lagarias, un résultat important de Berger-Wang [24] (plus tard démontré avec un énoncé plus précis par des moyens différents par Bochi [25], Breuilllard [33]) affirme que pour les parties bornées S de $M_n(\mathbb{C})$, on a l'égalité $r(S) = \rho(S)$. Revenant à notre cadre, définissons aussi $\Lambda_n(S) := \{\frac{\lambda(g)}{n} \mid g \in S^n\}$ de manière similaire aux $K_n(S)$ avec la projection de Jordan.

Énonçons maintenant notre deuxième résultat principal. Il résume quelques propriétés principales du spectre joint $J(S)$. On notera que le premier point du résultat suivant généralise l'égalité de Berger-Wang dans notre cadre.

Théorème 1.5. *Soient G un groupe (algébrique) réel linéaire semi-simple connexe et S une partie bornée de G engendrant un semi-groupe Zariski dense dans G.*

1. Les limites, pour la distance de Hausdorff dans $\mathfrak{a}^+$, suivantes existent et l'on a les égalités :

$$\lim_{n\to\infty} K_n(S) = J(S) = \lim_{n\to\infty} \Lambda_n(S)$$

2. $J(S)$ est compact, convexe et d'intérieur non-vide dans $\mathfrak{a}^+$.

Remarque 1.6. *Dans la section 4.1, on précise une région (dans $\mathfrak{a}^+$), optimale en certains aspects, contenant le spectre joint $J(S)$, en utilisant des hyperplans que l'on définit en considérant le rayon spectral joint classique dans de diverses représentations de G (voir la Fig. 4.1).*

4.4 À ce stade, on souhaite indiquer une notion proche, celle du cône limite, introduit par Benoist dans [11]. Il associe à chaque semi-groupe Γ Zariski-dense dans G comme auparavant, un cône fermé de $\mathfrak{a}^+$, que l'on appelle le cône limite de Benoist B_Γ de Γ. Ceci décrit les directions asymptotiques des éléments de Γ dans leurs projections de Cartan et de Jordan. Benoist démontre que pour un tel Γ, B_Γ est un cône fermé d'intérieur non-vide. Ce résultat est en fait un précurseur de notre théorème 1.5 dans la mesure où l'on voit facilement que pour une partie S bornée engendrant Γ, B_Γ est exactement le cône engendré par le spectre joint $J(S)$ de sorte que ces propriétés de B_Γ découlent du théorème 1.5. On souhaite faire remarquer au lecteur que bien que le fait que $J(S)$ est d'intérieur non-vide puisse être déduit du résultat correspondant de Benoist pour B_Γ, ici, on en donne une autre démonstration rapide (dépendant partiellement du resultat de Benoist, voir ci-dessous) en utilisant la théorie des produits aléatoires des matrices, notamment en combinant le théorème central limite de Goldsheid-Guivarc'h [64] (pour $G = \mathrm{SL}(d,\mathbb{R})$) et Guivarc'h [65] (plus généralement, pour un groupe G linéaire semisimple comme avant) avec le résultat de finitude d'Abels-Margulis-Soifer [16]. Ceci donne aussi une nouvelle preuve de la propriété correspondante du cône de Benoist pour $G = \mathrm{SL}(d,\mathbb{R})$, mais on souligne que pour un G plus général comme ci-dessus, le TCL de Guivarc'h dépend de cette propriété du cône de Benoist. On indique que cet utilisation éventuelle du TCL pour établir ce fait avait été mentionnée par Guivarc'h [38]. Par ailleurs, on signale aussi que d'abord Quint [104] et ensuite Guivarc'h [65] ont donné d'autres démonstrations de cette propriété du cône de Benoist, comme une conséquence des études plus précises qu'ils mènent, le premier sur la densité du groupe additif engendré par les projections de Jordan d'un semi-groupe Zariski dense, et le deuxième sur les projections des éléments loxodromiques sur le centralisateur d'un tore maximal déployé.

4.5 Comme il a été déjà mentionné, le spectre joint d'une partie bornée S de G est en lien étroit avec des considérations de grandes déviations pour les produits aléatoires des éléments de S. Dans la deuxième partie du chapitre 4, on effectue une étude détaillée de la fonction de taux apparaissant dans le théorème 1.1, et en particulier, on met en évidence son lien avec le spectre joint. Pour donner une idée de ce lien, soit μ une mesure de probabilité de support $S \subset \mathrm{GL}(d,\mathbb{R})$ et supposons que les moyennes $\frac{1}{n}\log||S_n||$ le long de la μ-marche aléatoire satisfont un PGD avec une fonction de taux I (voir le théorème 1.4). Au vu de la définition du rayon spectral joint $r(S)$ de S, on s'aperçoit que ce dernier est un majorant dans $\mathbb{R}$ du support effectif de I, c'est-à-dire l'ensemble $\{x \in \mathbb{R} \mid I(x) < \infty\}$. En ce sens, le point 1 du théorème suivant est une traduction extensive de cette observation au spectre joint, comprenant aussi une affirmation «réciproque». Les propriétés de continuité du point 2. découlent principalement de la convexité de I et du point 1., notamment du fait que D_I est d'intérieur non-vide. Finalement, la propriété d'unicité de zéro est une conséquence (en fait, équivalente à) du résultat de décroissance exponentielle de Le Page dans le théorème 1.3.

Théorème 1.7. *Soient G et μ comme dans le théorème 1.1, notons S le support de μ et soit $D_I = \{x \in \mathfrak{a} \mid I(x) < \infty\}$ le support effectif de la fonction de taux fournie par Théorème 1.1. Alors, on a*

1. L'ensemble D_I est convexe, d'intérieur non-vide dans $\mathfrak{a}^+$, et il satisfait les égalités $\overline{D}_I = J(S)$, $\mathring{D}_I = \mathring{J}(S)$ si S est bornée, et finalement, l'égalité $D_I = J(S)$ si S est finie.

2. La fonction I est convexe, donc localement lipschitzienne sur $\mathring{D}_I$, et possède un unique zéro correspondant au vecteur de Lyapunov $\vec{\lambda}_\mu \in \mathfrak{a}^+$ de μ.

5.1 Le chapitre 5 est dans un même esprit que le chapitre 4 : cette fois-ci, on suppose que la partie bornée S de G est finie, et on entreprend une étude de la croissance exponentielle du nombre des éléments dans les S^n en fonction de leurs comportements asymptotiques dans les projections de Cartan et de Jordan. On introduit des fonctions indicatrices de croissance exponentielle pour les parties finies de S ; elle généralisent la notion classique du taux de croissance exponentielle (voir ci-dessous).

5.2 Pour expliquer notre approche plus précisément, commençons par rappeler une notion classique : soit T un ensemble fini dans un semi-groupe Γ et notons $|T|$ le nombre des éléments dans T. La limite $v_T := \lim_{n\to\infty} |T^n|^{\frac{1}{n}}$ existe par sousmultiplicativité et est appelée le taux de croissance exponentielle de T. Le semi-groupe Γ est dite de croissance exponentielle s'il contient une partie génératrice finie T avec $v_T > 1$. En réalité, cette dernière condition ne dépend pas du choix de T, c'est une propriété de Γ. En s'approchant de notre cadre, dans le cas d'un groupe linéaire Γ (i.e. $\Gamma \leq GL(d, \mathrm{k})$ pour un $d \in \mathbb{N}$, et un corps k), il découle des résultats classiques de Milnor-Wolf et de l'alternative de Tits que Γ est à croissance exponentielle dès qu'il n'est pas nilpotent-par-fini. De plus, par des variantes et versions uniformes de l'alternative de Tits obtenues par Eskin-Mozes-Oh [51], Breuillard-Gelander [34] et Breuillard [35], la croissance exponentielle de Γ jouit d'une propriété plus forte, dite de croissance exponentielle uniforme (cf. la Section 5.1).

5.3 Supposons, dans cette introduction, que la partie finie S engendre un semi-groupe Zariski dense dans un groupe G comme ci-dessus. Inspiré de l'expression à la Ruelle-Lanford [106], [83] d'une fonction de taux d'un PGD (voir le théorème 2.4) et le travail précédent de Quint dans un cadre très proche [100], pour S, on introduit les fonctions de comptage suivantes, qui étendent la donnée numérique du taux de croissance v_S à, essentiellement, une fonction réelle sur le spectre joint $J(S)$. On précise aussi que l'on a emprunté la terminologie suivante de Quint [100].

Définition 1.8 *On appelle indicateur de croissance (de Cartan) de S la fonction* $\phi_S : \mathfrak{a} \longrightarrow \mathbb{R}_+ \cup \{-\infty\}$, *définie par*

$$\phi_S(x) := \inf_{\substack{O \text{ ouvert de } \mathfrak{a} \\ x \in O}} \limsup_{n \to \infty} \frac{1}{n} \log \#\{g \in S^n \mid \frac{1}{n}\kappa(g) \in O\}$$

L'indicateur de croissance de Jordan ψ_S *de* S *est défini de la même manière en utilisant la projection de Jordan* λ.

Remarque 1.9 *1. L'extension de la donnée numérique du taux de croissance exponentielle* v_S *de* S *par les indicateurs de croissance de* S *devrait être comparée à l'extension de la donnée numérique des rayons spectraux joints à une partie convexe, i.e. le spectre joint.*

2. Il s'avère que l'indicateur de croissance ϕ_S *est, en fait, intimement lié à des considérations de grandes déviations pour une suite de probabilités «de nature déterministe», plus précisément, la suite des images des mesures de probabilités uniformes sur les* S^n*'s par les projections de Cartan normalisées (cf. la Section 5.1).*

5.4 On résume nos résultats sur les indicateurs de croissance dans le théorème suivant. On voudrait souligner qu'avec ces fonctions s'apparaissent plusieurs questions naturelles et ouvertes (voir ci-dessous ainsi que le chapitre concerné). Dans le résultat suivant, on note que la deuxième affirmation dit, en particulier, que l'on peut lire la valeur numérique v_S sur les indicateurs de croissances de S, et la dernière affirmation signifie que l'on possède une croissance exponentielle sur un ensemble dense de comportements asymptotiques pour S, ces derniers sont naturellement paramétrés par des éléments du spectre joint $J(S)$. Brièvement, on appelle ce phénomène la croissance exponentielle dense de S.

Théorème 1.10 *Soit S une partie finie d'un groupe G (algébrique) réel linéaire semi-simple connexe, engendrant un semi-groupe Zariski dense dans G. Alors,*

1. Les indicateurs de croissance ϕ_S et ψ_S sont semi-continues supérieurement ayant pour maximum $\log v_S$.

2. On a l'inégalité $\phi_S \leq \psi_S$.

3. On a les égalités suivantes entre parties de $\mathfrak{a}$: $\{\phi_S \geq 0\} = \{\psi_S \geq 0\} = J(S) = \overline{\{\phi_S > 0\}}$

5.5 Dans la seconde partie du Chapitre 5, on exhibe des résultats antérieurs de Benoist [11] et de Quint [100] : pour un semi-groupe Γ Zariski dense dans un G comme avant, Benoist a introduit et étudié la notion d'un cône limite B_Γ de G dans une chambre de Weyl fixée $\mathfrak{a}^+$. Quint, à son tour, pour un tel Γ, cette fois-ci de plus un groupe discret, a introduit une fonction de comptage exponentiel sur $\mathfrak{a}$, l'indicateur de croissance ψ_Γ de Γ, qui est intimement liée au cône B_Γ, et a étudié ses propriétés (comme la concavité). Ces deux notions de Benoist et de Quint sont en analogie étroite avec, respectivement, notre spectre joint et indicateur de croissance (de Cartan). On y précise cette analogie. Pour le cône de Benoist, comme il a été déjà mentionné, nos résultats du chapitre 4 nous permettent de démontrer à nouveau (voir le corollaire 5.15) quelques propriétés de B_Γ, et d'autres se traduisent en des questions ouvertes. Pour l'indicateur de croissance ψ_Γ de Quint, on voudrait souligner qu'une différence de notre indicateur de croissance de cette dernière, provient d'une manière de comptage différente, de celle de Quint, du même objet. En utilisant cette observation, on établit quelques relations entre les deux indicateurs de croissance, de Quint et le nôtre, en utilisant une fonction de rayon spectral joint directionnelle que l'on introduit. Finalement, on souhaite préciser que la concavité remarquable de ψ_Γ, qui est en accord avec la convexité des fonctions de taux des PGD ci-dessus, se traduit dans notre cadre en une question ouverte que l'on examinera dans un travail futur.

5.6 La dernière partie du chapitre 5 est une collection de divers résultats. Dans un premier temps, on étudie les propriétés du PGD des projections de Jordan des marches aléatoires dans le même cadre que le théorème 1.1 : on montre qu'en utilisant le PGD fourni par ce théorème pour les projections de Cartan et combinant le résultat de finitude d'Abels-Margulis-Soifer avec les estimées, sur les projections de Cartan et de Jordan des éléments loxodromiques, de Benoist, on peut déduire la minoration dans la définition d'un PGD, avec la même fonction de taux que dans le théorème 1.1, pour les projections normalisées de Jordan le long des marches aléatoires (ceci est très similaire au point 2. du théorème 1.10). En outre, on établit l'analogue exacte du résultat du théorème 1.1 pour les projections de Jordan, mais dans un cadre très particulier, notamment celui des marches aléatoires (r, ϵ)-Schottky (cf. la partie concernée). Dans un second temps, en étudiant un exemple particulier, on arrive à une stratégie pour améliorer le théorème de croissance exponentielle dense de la Section 5.1. Et dernièrement, en utilisant la notion de rayon sous-spectral joint, on présente un critère de discrétude pour un semi-groupe Γ de type fini dans un G comme avant, et on montre que ce critère s'applique à des semi-groupes (r, ϵ)-Schottky de type finis.

6.1 Dans la dernière partie de ce texte, le chapitre 6, on adopte de nouveau une perspective probabiliste et, en peu de mots, on cherche un PGD pour la suite des variables aléatoires $\frac{1}{n}\log||S_nv||$ où S_n est le $n^{\text{ième}}$ pas d'une marche aléatoire sur $GL(d,\mathbb{R})$ et $v \in \mathbb{R}^d \setminus \{0\}$. N'ayant pas été capable d'utiliser nos techniques des parties précédentes, on suit l'idée initiale de Fustenberg qui voit ces variables en tant que des fonctionnelles sur une chaîne de Markov, et on tâche d'appliquer la théorie générale des chaînes de Markov dans notre situation. À la fin, sous des conditions considérablement restrictives (d'absolue continuité d'une puissance de convolution, comme Tutubalin, et une condition de bornitude que l'on explicite), on établit un PGD pour ces variables et leurs généralisations multi-dimensionnelles (cocycle d'Iwasawa) comme un corollaire d'une conclusion générale que l'on atteint.

6.2 Expliquons notre stratégie plus précisément : elle part d'un théorème (le théorème 6.2) de Stroock [109] et Ellis [50], qui donne une condition suffisante, appelée l'uniformité, sur le noyau de transition d'une chaîne de Markov M, pour que la suite de mesures empiriques de M satisfasse un PGD (sur l'espace de mesures de probabilités sur l'espace d'états de M). Comme dans l'article [14] de Benoist-Quint, par la généralité des techniques, on suit d'abord un cadre général, y établit une affirmation et déduit nos résultats comme des applications dans les cas particuliers : pour un groupe G localement compact dénombrable à l'infini, agissant continûment et transitivement sur un espace X métrisable compact, on considère une chaîne de Markov sur $G \times X$ associée à une marche aléatoire sur G gouvernée par une mesure de probabilité μ sur G. On traduit alors la condition d'uniformité de Stroock-Ellis à une condition **(D)**, plutôt technique, sur μ. Pour clarifier **(D)**, on montre aussi que, par exemple, elle est satisfaite dès que μ est de support compact et possède une puissance de convolution μ^{*n} absolument continue par rapport à la mesure de Haar et est minorée par un $\alpha > 0$ sur un voisinage de l'identité engendrant G (pour une description plus générale, voir la fin de la Section 6.1). Finalement, en utilisant une technique générale de la théorie de grandes déviations, le principe de contraction, dans un premier lieu (1.), on prend $G = GL(V)$, $X = \mathbb{P}(V)$ et obtient un PGD en faisant un transfert par une fonction que l'on construit en utilisant le cocycle de norme, et en deuxième lieu (2.), on prend G un groupe algébrique réel linéaire semisimple connexe, $X = \mathcal{F}_G$ sa variété de drapeaux et cette fois-ci, on fait un transfert en utilisant le cocycle d'Iwasawa σ, que l'on explicite dans la Section 6.2. En conséquence, dans ces cas particuliers, on obtient le résultat suivant.

Théorème 1.11. *Pour G et X comme dans* (1.) *et* (2.) *ci-dessus, et μ une mesure de probabilité sur G satisfaisant à la condition* ***(D)****, les suites des variables aléatoires $\frac{1}{n}\log||S_nv||$ et $\frac{1}{n}\sigma(S_n,\eta)$ satisfont des PGD, uniformément en v avec $||v|| = 1$ et $\eta \in \mathcal{F}_G$, avec des fonctions de taux propres et convexes, respectivement, sur $\mathbb{R}$ et $\mathfrak{a}$.*

6.3 À la fin, on accomplit une première étude des fonctions de taux apparaissant dans le théorème précédent, notamment de leur supports effectifs $\{x \mid I(x) < \infty\}$. On se contente, dans ce texte, de fournir des régions précises comprenant ces ensembles convexes, dans, respectivement, $\mathbb{R}$ et $\mathfrak{a}$, et on note quelques pistes de recherche.

Directions de recherches futures

7.1 Comme conséquence de la nouveauté des objets que l'on introduit, plusieurs questions ouvertes et directions de recherche émergent dans ce texte. La plupart de celles-ci est indiquée dans le texte, dans les parties concernées. En outre, comme il y est indiqué, ces questions s'illustrent dans "l'exemple du groupe libre", que l'on traite ci-dessous, à travers les résultats correspondants dans cet exemple. Dans les paragraphes suivants, on résume ces questions ouvertes et directions de recherches.

7.2 Dans le chapitre 3, la première question est naturellement de se poser si on peut se passer de l'hypothèse de Zariski-densité dans le théorème 1.1. Il paraît plausible qu'un PGD faible existe (pour la suite $\frac{1}{n}\kappa(S_n)$) sans aucune hypothèse sur la mesure de probabilité (voir le théorème 3.1). Une deuxième question directe concerne la régularité de la fonction de taux I apparaissant dans ces théorèmes : est-ce qu'elle est strictement convexe ou dérivable/analytique à l'intérieur de D_I.

Concernant le PGD pour les projections de Jordan normalisées le long des marches aléatoires $(\frac{1}{n}\lambda(S_n))$, nous présumons qu'au moins sous l'hypothèse de Zariski-densité, son existence peut être démontrée (voir la remarque 4.16). Nous examinons cette question, et donnons des réponses partielles dans la section 5.3. Une question ultérieure sur ce sujet sera alors sur la relation entre les fonction de taux de Cartan et de Jordan (voir la proposition 5.22, le corollaire 5.23, le corollaire 2.29 ainsi que "l'exemple du groupe libre"). Néanmoins, notons qu'on ne peut pas s'attendre à ce que le PGD existe pour la suite $\frac{\lambda(S_n)}{n}$ sans aucune hypothèse sur le support de la mesure de probabilité : d'une manière intéressante, cela est une conséquence d'un exemple de Breuillard sur la non-convergence à la Hausdorff au spectre joint de Jordan (voir le point 2. de la remarque 4.15 et le point 1. de la remarque 4.16).

Une autre question que l'on peut se poser sur le Théorème 1.1 concerne l'affaiblissement des hypothèses dans une autre direction : celle des marches aléatoires non-indépendantes. Dans cette direction, une hypothèse de mélange exponentiel, ou au moins celle de mélange super-exponentiel, serait vraisemblablement suffisante pour établir l'existence d'un PGD.

Encore une autre direction concerne les PGD pour les processus à temps continu (sur les groupes de Lie semi-simples ou sur leurs espaces symétriques), c'est-à-dire, le mouvement brownien ou en tant que généralisations directes des marches aléatoires à temps discret, les processus de Lévy (voir [85]). Nos techniques peuvent être utiles pour ces études et cela sera considéré dans un travail ultérieur.

Finalement, une question que l'on examine dans un travail en progrès concerne l'équivalent du théorème 1.1 d'une part pour les groupes algébriques linéaires semi-simples définis sur les autres corps locaux que $\mathbb{R}$ (voir la remarque 2.26), et d'autre part, pour les groupes algébriques linéaires réductifs.

7.3 Concernant le chapitre 4 sur les spectres joints, les questions immédiates concernent l'existence des limites pour les distances de Hausdorff sous des hypothèses moins fortes (que la Zariski-densité) : pour le cas de Cartan, il paraît de nouveau plausible que la suite $\frac{\kappa(S^n)}{n}$ converge sans aucune hypothèse sur S. Pour les projections de Jordan, bien qu'une condition simple comme $e \in S$, ou la Zariski-densité de $\cup_{n\geq 1} S^n$ soient suffisantes (voir le point 2. du théorème 4.4) pour la convergence de $\frac{\lambda(S^n)}{n}$, la convergence n'a pas lieu sans aucune hypothèse sur S (voir les points 1. et 2. de la remarque 4.15). On note aussi que sans hypothèse de Zariski-densité, on ne peut pas s'attendre à avoir les mêmes propriétés du spectre joint (e.g. le fait que $J(S)$ est d'intérieur non-vide) : par exemple, on peut facilement trouver S telle que $J(S)$ soit contenu dans un mur de la chambre de Weyl.

Un deuxième type de question sur le spectre joint concerne les formes possibles que l'on peut atteindre : plus précisément, quelles parties compactes convexes d'intérieur non vides de $\mathfrak{a}^+$ d'un groupe G peuvent être le spectre joint d'une partie S de G (voir la deuxième sous-section de la Section 4.1). Cette question apparaît en relation avec les résultats correspondants sur le cône de Benoist ([11]) (voir aussi la Proposition 5.13).

Une autre direction que l'auteur souhaite approfondir concerne la considération du spectre joint dans un cadre plus général. Comme il est mentionné dans le texte, on peut facilement observer que la définition du spectre joint se transpose mutatis mutandis aux $M_n(k)$ pour un corps local k.

Finalement, dans un travail en cours, on étudie un troisième type de spectre en relation avec la décomposition d'Iwasawa. Nous le définissons d'une manière similaire, cette fois-ci, en utilisant la projection d'Iwasawa (voir "l'exemple du groupe libre" et la remarque 6.31) et examinerons ses propriétés ainsi que sa relation avec le spectre joint (comparer la Fig. 4.1 et la Fig. 6.1).

7.4 Dans le chapitre 5, l'étude des indicateurs de croissance est en fait d'un caractère plutôt inachevé, elle consiste en des résultats de l'auteur (le théorème 1.10) d'un projet en cours. Quelques-unes des questions que l'on aborde dans ce projets sont :

1. Est-ce que les indicateurs de croissance sont des "fonctions de taux" ? Par cela, on entend : est-ce qu'on obtient les mêmes fonctions si on remplace $\limsup$ par $\liminf$ dans leurs définitions ? On rappelle que ceci est équivalent à se demander si les poussés en avant des mesures de probabilités uniformes sur les S^n par les projections de Cartan/Jordan normalisées satisfont un PGD (voir la remarque 5.2).
2. Est-ce que les indicateurs de croissance sont concaves (voir la remarque 5.9, la remarque 5.18 et le paragraphe qui la suit) ? Cette question est en relation avec le résultat correspondant pour l'indicateur de croissance ψ_Γ de Quint (voir le théorème 5.17 et [100]).
3. Une troisième question concerne l'étude de l'ensemble des maxima des indicateurs de croissance (voir la remarque 5.7 et "l'exemple du groupe libre"). On note que dans le cadre probabiliste, cela correspond à l'étude de l'ensemble des zéros des fonctions de taux des PGD's : pour la projection de Cartan, ceci est essentiellement le théorème 1.3 de Le Page.
4. Une quatrième question est sur la compréhension de la relation entre les deux indicateurs de croissance ϕ_S et ψ_S ainsi que la relation entre ϕ_S et la fonction de taux I du théorème 1.1 (voir la remarque 5.2). Cette question semble être liée au contenu de la remarque 7.5.

7.5 Les questions qui surgissent sur le contenu du chapitre 6 sont naturellement similaires à celles du chapitre 3. On note que dans le chapitre 6, nous avons dû poser une hypothèse **(D)** beaucoup plus forte sur la mesure de probabilité, pour établir le PGD pour le cocycle d'Iwasawa le long des marche aléatoires. Pour un premier affaiblissement de cette hypothèse, les techniques récemment développées par Guivarc'h-Le Page [66] peuvent être utiles. Quelques autres questions que l'on a, portent sur la relation entre les fonctions de taux apparaissant dans ce chapitre et le rayon sous-spectral joint, et plus généralement, le spectre joint d'Iwasawa (voir la remarque 6.27).

L'exemple du groupe libre

Dans cet exemple, suivant la suggestion d'Emmanuel Breuillard, nous traitons un cas très spécifique pour illustrer, à travers une analogie, bon nombre des résultats et nouveaux objets dont on vient de donner un aperçu. Il va aussi nous permettre d'illustrer de nombreuses questions ouvertes, que l'on se pose dans ce texte, en comparant ce que l'on (sait faire) fait dans ce cas spécifique et ce que l'on ne (sait pas faire) fait pas dans notre étude d'asymptotique, déterministe ou probabiliste, dans le cas des groupes G, qui dans cet exemple signifie un groupe algébrique réel linéaire semi-simple connexe comme auparavant. L'exemple auquel on pense est, sans surprise, celui du groupe libre F_q.

Soit donc F_q le groupe libre de rang q, $S = \{a_1^{\pm 1}, \ldots, a_q^{\pm 1}\}$ une partie génératrice libre dans F_q et μ la mesure de probabilité uniforme sur S. On considère F_q muni de sa structure métrique provenant de notre choix de S : pour $g, h \in F_q$, $d(g,h) = l(g^{-1}h)$ où $l(.)$ note la fonction de longueur des mots (réduits) en S. Une première comparaison que l'on va considérer est entre la projection de Cartan $\kappa(.)$ d'un groupe G semi-simple (ou réductif...) et la fonction de longueur $l(.)$ sur F_q. Ainsi, pour notre analogie, on est dans le cas d'un groupe G de rang réel un (e.g. $\mathrm{SL}(2,\mathbb{R})$) et $l(.)$ 'projette' sur la 'demi-droite' $\mathbb{N} \subset \mathbb{Z}$ (chambre de Weyl de dimension 1).

Notons par T_{2q} le graphe de Cayley (à droite) de F_q associé à S. Bien entendu, il s'agit de l'arbre homogène de degré $2q$ avec sa structure étiquettée. Le groupe F_q agit naturellement (à gauche) sur T_{2q} par 'multiplication à gauche' et chaque élément $g \in F_q \setminus \{e\}$ a un unique axe (géodésique bi-infinie dans T_{2q}) sur lequel g induit une translation (pour plus d'informations voir la partie 2 de ce texte qui est entièrement sur les groupes libres et leurs actions). Cette distance de translation sera notée $\tau(g)$ et s'appelle la distance de translation de g. Elle correspond à la longueur de la partie cycliquement réduite de g. La deuxième comparaison que l'on va considérer est entre la projection de Jordan $\lambda(.)$ d'un groupe G comme avant et cette application $\tau(.)$ à valeurs dans la «demi-droite» $\mathbb{N} \subset \mathbb{Z}$.

Soit enfin ∂T_{2q} (ou également ∂F_q) le bord de l'arbre T_{2q} (ou groupe F_q) que l'on va identifier à l'ensemble des mots (réduits) infinis en S. L'action de F_q sur T_q s'étend à une action sur la compactification $T_{2q} \cup \partial T_{2q}$. Pour préciser, l'axe de translation d'un élément $g \in F_q \setminus \{e\}$ du paragraphe précédent, correspond à deux points de ∂T_{2q}, un attactif et un répulsif (voir la partie 2). Pour un élément $\xi \in \partial T_{2q}$, considérons la fonction de Busemann de point base $e \in F_q \simeq T_{2q}$, c'est-à-dire, la fonction définie, pour un $g \in F_q$, par $b_\xi(g) = \lim_{n\to\infty} d(g, z_n) - d(e, z_n)$ où z_n est une suite dans F_q dont les sommets correspondants dans T_q tendent vers ξ dans $T_{2q} \cup \partial T_{2q}$. La limite ne dépend pas du choix de $(z_n)_{n\in\mathbb{N}}$. La troisième analogie que l'on va considérer est entre cette fonction de Busemann $b_\xi(.)$ et le cocycle d'Iwasawa $\sigma(., \xi)$ à valeurs dans $\mathfrak{a}$ d'un groupe G comme avant, où ξ est un élément de la variété des drapeaux de G, i.e. une classe de conjugaison d'un sous-groupe parabolique (pour la définition de $\sigma(., .)$ voir le chapitre 6, et pour l'analogie voir aussi les lemme 5.29 et lemme 7.2 dans [14]). On précise en finissant cette description que nous restons bien au niveau d'une analogie formelle et que ces notions ne se correspondent pas d'une manière évidente dans le cas où, par exemple, on considère une injection de F_q dans un G (voir aussi le paragraphe suivant le théorème 7.11).

On va donc mener une étude de l'asymptotique de $S, S^2, \ldots$ dans F_q via les applications de longueur ('projection de Cartan'), de distance de translation ('projection de Jordan') et fonctions de Busemann ('cocycle d'Iwasawa'), comme nous l'avons décrite pour un groupe G dans l'introduction. En ce qui concerne les spectres joints, dans notre approche déterministe, la situation est simple : les limites pour la distance de Hausdorff des ensembles $\{\frac{l(g)}{n} \mid g \in S^n\}$ et $\{\frac{\tau(g)}{n} \mid g \in S^n\}$ existent et sont toutes les deux égales à $[0, 1]$. Cela correspond à notre spectre joint (voir le théorème 1.5). Par ailleurs, les ensembles $\{\frac{b_\xi(g)}{n} \mid g \in S^n\}$ convergent aussi pour tout $\xi \in \partial F_q$ vers l'intervalle $[-1, 1]$; cela correspondrait au spectre joint d'Iwasawa (voir la remarque 6.31) qui fera l'objet d'un travail futur dans le cadre d'un groupe linéaire G comme ci-dessus.

Soit $S_n = X_1.\ldots.X_n$ le $n^{\text{ième}}$ pas de la μ-marche aléatoire sur F_q. Pour l'aspect probabiliste de notre étude, on va donc se demander si (en fait montrer que) chacune des suites des variables aléatoires $\frac{l(S_n)}{n}$, $\frac{\tau(S_n)}{n}$ et $\frac{b_\xi(S_n)}{n}$ pour un $\xi \in \partial F_q$ satisfait un principe de grandes déviations (PGD) ; et on trouvera l'expression exacte de leur fonctions de taux respectives I, J et K (cette dernière ne dépendra pas de $\xi \in \partial F_q$). À ce stade, rappelons que dans notre étude d'un groupe linéaire G, on possède un PGD pour les projections de Cartan le long des marches aléatoires (théorème 1.1), un PGD sous des hypothèses très restrictives pour le cocycle d'Iwasawa le long des marches aléatoires (dont la fonction de taux ne dépend pas de ξ dans la variété des drapeaux, voir le théorème 1.11) mais on ne possède pas d'un PGD pour les projections de Jordan (pour un résultat partiel voir la section 5.3). De plus, l'expression exacte de I, J et K dans cet exemple correspondrait au théorème 1.7 où l'on donne une description de nos fonctions de taux. En ce qui concerne l'étude de nos indicateurs de croissance (voir la définition 1.8), notons d'abord que comme la mesure de probabilité μ est uniforme sur une partie libre S de F_q, les convolutions μ^{*n} pour $n \in \mathbb{N}$, c'est-à-dire les lois des S_n induisent des lois uniformes sur les sphères de F_q (de centre $e \in F_q$) qu'elles chargent ; autrement dit la loi conditionnelle de $l(S_n) = k$ (n et k de même parité) est la mesure de probabilité uniforme sur la sphère $B_{=k} := \{x \in F_q \mid l(x) = k\}$ (on utilise la même notation pour la sphère d'origine le sommet e et de rayon k dans T_{2q}). Cette observation implique que, dans ce cas particulier, l'étude des indicateurs de croissance ϕ_S et ψ_S (voir le Théorème 1.10) est équivalente à celle de I et de J, car on a (voir la Section 5.3)

$$\phi_S(x) = \log(2q-1) - I(x) \quad \text{et} \quad \psi_S(x) = \log(2q-1) - J(x) \tag{1.2}$$

où $2q-1$ est le taux de croissance exponentielle de F_q. Dans la suite de cet exemple, pour deux suites a_n et b_n ($n \in \mathbb{N}$) de nombre réels strictement positifs, la notation $a_n \sim b_n$ indiquera l'équivalence exponentielle en infini, c'est-à-dire, $\frac{a_n^{\frac{1}{n}}}{b_n^{\frac{1}{n}}} \underset{n\to\infty}{\longrightarrow} 1$.

L'étude du PGD de la suite $\frac{l(S_n)}{n}$ et des fonctions I et ϕ_S

Signalons tout d'abord que pour éviter des problèmes de parité, nous traiterons le cas $\mathbb{P}(l(S_{2n}) = 2k)$ pour $0 \leq k \leq n$ (bien entendu, pour les autres k, cette quantité est nulle). Ceci n'est pas une restriction, car on a clairement $\mathbb{P}(l(S_{2n+1}) = 2k+1) = \mathbb{P}(l(S_{2n}) = 2k)\frac{2q-1}{2q} + \mathbb{P}(l(S_{2n}) = 2k+2)\frac{1}{2q}$. On commence alors par remarquer que par la propriété (de μ^{*n}, mentionnée dans le dernier paragraphe) d'équidistribution sur les sphères, on est ramené à des problèmes de pur comptage. Par cette propriété, de façon évidente, on a

$$\mathbb{P}(l(S_{2n}) = 2k) = \frac{\#(\text{chemins de longueur } 2n \text{ d'origine } e \in T_{2q} \text{ et d'extrémité sur } B_{=2k})}{\#(\text{chemins de longueur } 2n \text{ d'origine } e \in T_{2q})} \tag{1.3}$$

où par un chemin de longueur $2n$, on entend $2n$ choix successifs de sommets voisins de T_{2q}. On va utiliser le fait que le numérateur de (1.3) est $\sim \binom{2n}{n+k}(2q-1)^{n+k}$ (pour l'expression exacte de ce numérateur, voir l'appendice A.2). Par ailleurs, il est clair que le dénominateur est égal à $(2q)^{2n}$. Il découle de la définition du PGD (où du théorème 2.4) que, pour démontrer qu'il existe, il suffit de voir que pour tout $\alpha \in \mathbb{R}$, la limite

$$\lim_{\epsilon\to 0}\lim_{n\to\infty} \frac{1}{2n} \log \mathbb{P}(\alpha - \epsilon < \frac{l(S_{2n})}{2n} < \alpha + \epsilon) \tag{1.4}$$

existe dans $\mathbb{R} \cup \{-\infty\}$, et alors, cette limite sera (définira) $-I(\alpha)$. Or, on a

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{2n} \log \mathbb{P}(\alpha - \epsilon < \frac{l(S_{2n})}{2n} < \alpha + \epsilon) &= \lim_{n\to\infty} \frac{1}{2n} \log \sum_{k \geq (\alpha-\epsilon)n}^{(\alpha+\epsilon)n} \mathbb{P}(l(S_{2n}) = k) \\ &= \lim_{n\to\infty} \frac{1}{2n} \log \mathbb{P}(l(S_{2n}) = \beta 2n)\end{aligned} \tag{1.5}$$

où $\alpha - \epsilon \leq \beta \leq \alpha + \epsilon$ maximisant cette dernière expression. Cette dernière égalité suit des considérations élémentaires sur les limites. Maintenant, en utilisant l'équivalent du numérateur de (1.3) et ensuite la formule de Stirling, on a

$$\mathbb{P}(l(S_{2n}) = 2k) \sim \binom{2n}{n+k}\frac{(2q-1)^{n+k}}{(2q)^{2n}} \sim \frac{(2q-1)^{n+k}}{(2q)^{2n}} \frac{(2n)^{2n}}{(n+k)^{n+k}(n-k)^{n-k}} \tag{1.6}$$

et finalement, remplaçant k par βn, en prenant le logarithme et divisant par $\frac{1}{2n}$, maximisant par calcul différentiel en $\beta \in]\alpha - \epsilon, \alpha + \epsilon[$ et faisant ϵ tendre vers zéro, en posant $0\log(0) := 0$, on trouvera (d'après (1.5) et la ligne suivant (1.4))

$$I(\alpha) = \begin{cases} \frac{1+\alpha}{2}\log(1+\alpha) + \frac{1-\alpha}{2}\log(1-\alpha) + \log(q) - \frac{1+\alpha}{2}\log(2q-1) & \alpha \in [0,1] \\ \infty & \text{sinon} \end{cases} \tag{1.7}$$

On remarque que cette fonction est strictement convexe sur $[0,1]$, admet son unique zéro en $1-\frac{1}{q}$, qui correspond à la vitesse de fuite de la marche aléatoire, et $I(0)$ est égal au logarithme du rayon spectral de la marche aléatoire. On a l'égalité entre le support effectif $D_I = \{x \in \mathbb{R} \mid I(x) < \infty\}$ et «le spectre joint» $[0,1]$ (voir le théorème 1.7 (1.)). On notera aussi que sur la frontière de D_I, on a la dérivée $I'(0) < \infty$ et $\lim_{\alpha \to 1^-} I'(\alpha) = +\infty$. Finalement, pour l'indicateur de croissance ϕ_S, par (1.2), on a

$$\phi_S(\alpha) = \begin{cases} \log(\frac{(2q-1)^{\frac{3}{2}}}{q}) + \frac{\alpha}{2}\log(2q-1) - \frac{1+\alpha}{2}\log(1+\alpha) - \frac{1-\alpha}{2}\log(1-\alpha) & \alpha \in [0,1] \\ -\infty & \text{sinon} \end{cases}$$

une fonction strictement concave sur $[0,1]$ avec unique maximum en $1-\frac{1}{q}$ (comparer avec le théorème 1.10, et voir les remarques 5.7 et 5.9).

On indique aussi que si on considère une marche aléatoire S'_n associée à une autre mesure de probabilité ν de support S', disons fini et engendrant F_q, il est n'est pas difficile de voir que pour la fonction de longueur $l(.)$ associée à une partie génératrice S, la suite $\frac{l(S'_n)}{n}$ satisfait toujours un PGD avec une fonction de taux I' dont le support effectif $D_{I'}$ est égal au 'spectre joint de S'', c'est-à-dire, la limite Hausdorff de $\{\frac{l(g)}{n} \mid g \in (S')^n\}$. On omet la démonstration de ce fait, qui se fait de manière parallèle (mais plus facile) à la partie d'existence dans notre preuve du théorème 1.1, et la proposition 4.22.

L'étude de PGD de la suite $\frac{\tau(S_n)}{n}$ et des fonctions J et ψ_S

Pour cet étude, il s'agira de suivre l'argument précédent pour la quantité $\mathbb{P}(\tau(S_{2n}) = 2k)$ (on note que pour tous $k, n \in \mathbb{N}$, $\mathbb{P}(\tau(S_{2n}) = 2k+1) = 0$). Pour pouvoir calculer cette dernière, on la décomposera suivant la distance à l'identité : on a

$$\mathbb{P}(\tau(S_{2n}) = 2k) = \sum_{j \geq k}^{n} \mathbb{P}(\tau(S_{2n}) = 2k \mid l(S_{2n}) = 2j).\mathbb{P}(l(S_{2n}) = 2j) \tag{1.8}$$

Pour comprendre le terme $\mathbb{P}(\tau(S_{2n}) = 2k \mid l(S_{2n}) = 2j)$, en vue du fait que $\mathbb{P}(. \mid l(S_{2n})2j)$ est une loi uniforme sur la sphère $B_{=2j}$, il suffit d'estimer suffisamment précisément le nombre des éléments de F_q de longueur $2j$ et de distance de translation $2k$. Or, il n'est pas difficile de voir que ce nombre est compris entre $2q(2q-1)^{j+k-2}(2q-3)$ et $2q(2q-1)^{j+k-1}$ (pour le calcul exact de ce nombre, voir l'Appendice A.2). Ainsi, pour $j \geq k$, on a

$$\mathbb{P}(\tau(S_{2n} = 2k) \mid l(S_{2n}) = 2j) \sim \frac{1}{(2q-1)^{j-k}} \tag{1.9}$$

En remplaçant ceci et (1.6) dans (1.8), on a

$$\mathbb{P}(\tau(S_{2n}) = 2k) \sim \sum_{j \geq k}^{n} \frac{(2q-1)^{n+k}}{q^{2n}(1+\frac{j}{n})^{n+j}(1-\frac{j}{n})^{n-j}} \tag{1.10}$$

Comme dans (1.5) et le paragraphe qui la suit, pour calculer $J(\alpha)$ pour $\alpha \in [0,1]$ (noter que $J(\alpha) = \infty$ pour tout $\alpha \in \mathbb{R} \setminus [0,1]$), ce qu'on doit faire est de remplacer

$k = \alpha n$ et de chercher j de la forme $\beta_\alpha n$ avec $1 \geq \beta_\alpha \geq \alpha$ maximisant le terme à l'intérieur de la somme (1.10) pour notre α fixé. Pour alléger la notation, posons $\beta = \beta_\alpha$. En remplaçant ces valeurs de k et j, le terme à l'intérieur s'écrit

$$\frac{(2q-1)^{n+n\alpha}}{q^{2n}(1+\beta)^{n+n\beta}(1-\beta)^{n-n\beta}}$$

En y prenant le logarithme et divisant par $2n$, ceci donne $\frac{1}{2}(1+\alpha)\log(2q-1) - \log(q) - \frac{1+\beta}{2}\log(1+\beta) - \frac{1-\beta}{2}\log(1-\beta)$. Donc par calcul différentiel, on trouve que le β $(= \beta_\alpha)$ qui maximise cette expression est $\beta = \alpha$ et donc, en multipliant cette expression par (-1) et remplaçant β par α, on trouve

$$J(\alpha) = \begin{cases} \frac{1+\alpha}{2}\log(1+\alpha) + \frac{1-\alpha}{2}\log(1-\alpha) + \log(q) - \frac{1+\alpha}{2}\log(2q-1) & \alpha \in [0,1] \\ \infty & \text{sinon.} \end{cases} \tag{1.11}$$

Ainsi, on aboutit aux égalités remarquables de $I = J$, et aussi, par les égalités (1.2), de $\phi_S = \psi_S$.

L'étude de PGD de la suite $\frac{b_\xi(S_n)}{n}$ et des fonctions K

De même que dans les parties précédentes, on se demande s'il existe une fonction $K : \mathbb{R} \longrightarrow [0,\infty]$ satisfaisant $\mathbb{P}(b_\xi(S_n) \simeq \alpha n) \sim \exp(-nK(\alpha))$. Si elle existe (et elle existe), la fonction K peut, a priori, dépendre du point ξ du bord, mais cela ne sera pas le cas. Pour estimer donc $\mathbb{P}(b_\xi(S_{2n}) = 2k)$ pour $-n \leq k \leq n$ (notons que pour les autres valeurs de k, cette probabilité est nulle, et qu'il y a le même problème de parité que pour $l(.)$ et $\tau(.)$), on la décomposera comme dans (1.8), en y incluant cette fois aussi les valeurs négatives de k. Ainsi, pour tout $\xi \in \partial F_q$, on a

$$\mathbb{P}(b_\xi(S_{2n}) = 2k) = \sum_{j \geq |k|}^{n} \mathbb{P}(b_\xi(S_{2n}) = 2k \mid l(S_{2n}) = 2j).\mathbb{P}(l(S_{2n}) = 2j) \tag{1.12}$$

Par conséquent, comme pour $\tau(.)$, on est ramené à estimer le nombre des éléments g sur la sphère $B_{=2j}$ de F_q satisfaisant $b_\xi(g) = 2k$ avec $k \in [-j,j]$. Un calcul exact de ce nombre est facile à faire, et l'on a

$$|b_\xi^{-1}(2k) \cap S_{=2j}| = \begin{cases} (2q-2)(2q-1)^{j+k-1} & k \notin \{-j,j\} \\ 1 & k = -j \\ (2q-1)^{j+k} & k = j \end{cases} \tag{1.13}$$

Finalement, on obtient

$$\mathbb{P}(b_\xi(S_{2n}) = 2k \mid l(S_{2n}) = 2j) \sim \frac{1}{(2q-1)^{j-k}} \tag{1.14}$$

En vue de (1.9), on remarque donc que l'on a $\mathbb{P}(b_\xi(S_{2n}) = 2k \mid l(S_{2n}) = 2j) \sim \mathbb{P}(\tau(S_{2n} = 2k) \mid l(S_{2n}) = 2j)$. Maintenant, de la même manière que l'on a obtenu (1.10), on a

$$\mathbb{P}(b_\xi(S_{2n}) = 2k) \sim \sum_{j \geq |k|}^{n} \frac{(2q-1)^{n+k}}{q^{2n}(1+\frac{j}{n})^{n+j}(1-\frac{j}{n})^{n-j}} \tag{1.15}$$

Dès lors, pour trouver l'expression de $K(\alpha)$, on peut raisonner comme dans le paragraphe qui suit (1.10). Ou encore, on peut conclure directement à partir de (1.10), en remarquant la symétrie en j du terme dans la somme dans (1.15), que

$$K(\alpha) = \begin{cases} \frac{1+\alpha}{2}\log(1+\alpha) + \frac{1-\alpha}{2}\log(1-\alpha) + \log(q) - \frac{1+\alpha}{2}\log(2q-1) & \alpha \in [-1,1] \\ \infty & \text{sinon} \end{cases} \tag{1.16}$$

On conclut donc qu'on a les égalités $K_{|[0,1]} = I_{|[0,1]} = J_{|[0,1]}$ et que K est strictement convexe et décroissante sur $[-1, 1-\frac{1}{q}]$ avec $K(-1) = \log(2q)$, $K(1-\frac{1}{q}) = 0$ et $K(1) = \log(1+\frac{1}{2q-1})$, et sur la frontière $\{-1,1\}$ du support effectif de K, pour la dérivée de K, on a $\lim_{\alpha\to -1^+} K'(\alpha) = \lim_{\alpha \to 1^-} K'(\alpha) = \infty$.

Introduction

1. Let G be a connected semisimple linear real algebraic group (e.g. $SL(d,\mathbb{R})$) and S be a subset of G. This text is concerned with the study of the asymptotics of the sequence $S, S^2, S^3, \ldots$ of subsets of G, where for $n \in \mathbb{N}$, S^n denotes the set $\{g_1.\ldots.g_n \mid g_i \in S\}$. The properties of this sequence will be investigated through some classical decompositions of G, components of which reflect information about asymptotics in G.

2. Let us first introduce a such classical decomposition, the Cartan decomposition, and the resulting projection map which will be of central interest to us : in the case of $SL(d,\mathbb{R})$, it corresponds to the classical polar decomposition which allows one to write a $g \in SL(d,\mathbb{R})$ as $g = k_1 a k_2$, where $k_1, k_2 \in SO(d,\mathbb{R})$ and a is a diagonal matrix with strictly positive entries. In contrast to the k_i's, the coefficients of a are uniquely determined (up to order), they are called singular values of g. For a general G as before, we can write $G = K\exp(\mathfrak{a}^+)K$, where K is a maximal compact subgroup and $\mathfrak{a}^+$ is a chosen Weyl chamber of G. The mapping $\kappa : G \longrightarrow \mathfrak{a}^+$ which associates to a $g \in G$, the unique element $a_g \in \mathfrak{a}^+$ such that $g \in K\exp(a_g)K$ is called the Cartan projection. For $g \in SL(d,\mathbb{R})$ we can write $\kappa(g) = (\log||g||, \log\frac{||\wedge^2 g||}{||g||}, \ldots, \log\frac{||\wedge^d g||}{||\wedge^{d-1}g||})$, where $||.||$ stands for the operator norm.

3.1 In Chapter 3, for our study, we shall adopt a probabilistic approach and consider products of independent random elements of S with respect to some probability measure μ supported on the subset S. This is called a random walk on the ambient group G and, more precisely, we will be concerned with the *rare events*, i.e. large deviation considerations, on these random walks.

3.2 In this probabilistic perspective, let us start with describing the fundamental result of Furstenberg-Kesten [59] (later, a corollary of Kingman's subadditive ergodic theorem), which can be thought of as a non-commutative analogue of law of large numbers. To state it, let G denote $SL(d,\mathbb{R})$ or $GL(d,\mathbb{R})$, and μ be a Borel probability measure on G. We say that μ has a finite first order moment, if, setting $N(g) = \max\{||g||, ||g^{-1}||\}$, we have $\int \log N(g)\mu(dg) < \infty$. Denoting by S_n the n^{th} step of the μ-random walk on G (i.e. $S_n = X_n.\ldots.X_1$, where X_i are independent random variables with values in $S \subset G$ of law μ), the Furstenberg-Kesten theorem says that, almost surely, the average norm of the random product S_n, $\frac{1}{n}\log||S_n||$ will converge to a constant $\lambda_1(\mu)$ called the first Lyapunov exponent of μ. This property and the constant $\lambda_1(\mu)$ can be thought of as giving a first asymptotic information about the Cartan projections of (random) products of S, which of course also depends on μ.

3.3 In our first result, we obtain a non-commutative analogue of the classical result of Cramér [41] on the large deviation probabilities. In the light of the previous paragraph, our result will be seen to stand in the same relation to Cramér's result as Furstenberg-Kesten theorem stands to the classical law of large numbers : let X be a topological space and $\mathcal{F}$ be a σ-algebra on X ; recall that a sequence μ_n of probability measures (equivalently, random variables Z_n of laws μ_n) on $(X, \mathcal{F})$ is said to satisfy a large deviation principle (LDP) with rate function $I : X \longrightarrow [0, \infty]$, if for all measurable subset R of X, we have

$$-\inf_{x \in \overset{\circ}{R}} I(x) \leq \liminf_{n \to \infty} \frac{1}{n} \log \mu_n(R) \leq \limsup_{n \to \infty} \frac{1}{n} \log \mu_n(R) \leq -\inf_{x \in \overline{R}} I(x)$$

Cramér's theorem says that the sequence of averages $S_n = \frac{1}{n} \sum_{i=1}^{n} X_i$, of $\mathbb{R}$-valued independent identically distributed random variables X_i of finite exponential moment, satisfies an LDP with a proper convex rate function I, given by the convex conjugate of the logarithmic moment generating function of X_i's. We note that in the above setting of $SL(d, \mathbb{R})$, the finite exponential moment for μ means that there exists $c > 0$ with $\int N(g)^c \mu(dg) < \infty$ (for general case, see Section 3.3). Our first main result writes

Theorem 1.1 (Simplified)**.** *Let G be a connected semisimple linear real algebraic group, μ a probability measure of finite exponential moment on G, whose support generates a Zariski dense sub-semigroup of G. Then, the sequence random variables $\frac{1}{n}\kappa(S_n)$ satisfies an LDP with a proper convex rate function I.*

Remark 1.2.
1. *We also obtain an analogue of the generalisation of Cramér's result by Bahadur [8] (see also Bahadur-Zabell [9]) with no moment condition on the probability measure μ (see Theorem 3.1).*
2. *In the case of a stronger moment condition (see Section 3.3), by exploiting convexity, we are able to identify the rate function I with the convex conjugate (Fenchel-Legendre transform) of a limiting logarithmic moment generating function (see Theorem 3.2).*
3. *Through a detailed analysis of the effective support $D_I := \{x \in \mathfrak{a} \mid I(x) < \infty\}$ of the rate function I and the convexity of I, we also obtain the existence of limits in the LDP for sufficiently regular sets $R \subset \mathfrak{a}$ (see Corollary 4.28)*

3.4 Let us also say some words about the main assumption in our theorem on the support of the probability measure and mention some of the predecessors of our result : Furstenberg-Kesten result was followed by further remarkable results of Furstenberg ([58], [57], [56] ...) on random matrix products and, more generally, on random walks on groups. In particular, he considered different but in many ways closely related quantities $\log ||S_n v||$ for $v \in \mathbb{R}^d \setminus \{0\}$: realising these variables as functionals of a Markov chain on a state space including the projective space $\mathbb{P}(\mathbb{R}^d)$ and finely analysing the (random) projective action, he obtained crucial results on the behaviour of random matrix products (for example, on the Lyapunov exponent $\lambda_1(\mu)$ introduced above). Concerning the extension of classical limit theorems to random matrix products, another step was taken by Tutubalin ([110], ...), who, again exploiting the Markovian structure highlighted by Furstenberg, was able to apply general methods of Nagaev ([90]) based on spectral theory, to establish, for example, a central limit theorem (CLT) for $\log||S_n v||$. But to use these methods, Tutubalin assumed rather restrictive hypotheses on the probability measure μ, namely the absolute continuity of it with respect to the Haar measure.

The following fundamental step was taken in 80's by the French school Le Page [84], Guivarc'h [64] [65], Raugi [67], Bougerol [30], ..., who were able to extend Furstenberg's initial results on exponential contraction and stationary measures on the projective spaces to a more general setting (for example flag varieties) and, more remarkably, to probability measures of considerably greater generality. With these extensions, Le Page was able to settle spectral gap results, by which, using Nagaev's method, he obtained non-commutative versions of several classical limit theorems (CLT, law of iterated logarithm and for example Theorem 1.3 below). Aforementioned generality of the probability measures solely include algebraic-geometric conditions, namely proximality (contraction) and strong irreducibility (for definitions, see Section 4.2), on the semigroup Γ generated by the support of the probability measure. Among these extensions of classical limit theorems, the following theorem of Le Page is the first result on the large deviation estimates and most importantly, it establishes the exponential decay of large deviation probabilities :

Theorem 1.3 (Le Page [84]). *Let μ be probability measure of finite exponential moment on $GL(d,\mathbb{R})$ whose support generates a contracting and strongly irreducible semigroup. Then, there exists a constant $B > 0$ such that for every vector $x \neq 0$ and for every $0 < \epsilon < B$, we have*

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(\log||S_n x|| - n\lambda_1(\mu) > n\epsilon) = \phi(\epsilon) \tag{1.17}$$

where $\phi : [0, B[\to \mathbb{R}$ is a concave function such that for all $\epsilon \in]0, B[$, $\phi(\epsilon) < 0$ and $\phi(0) = 0$ (the corresponding result holds for $\mathbb{P}(\log||S_n x|| - n\lambda_1(\mu) < -n\epsilon)$).

We would like to mention at this point that the exponential decay result of this theorem, and its several generalisations by Benoist-Quint [14], [15], was recently put to good use, by Bourgain-Furman-Lindenstrauss-Mozes [32] and Benoist-Quint [21], [22], [23], in the study of discrete subgroups of Lie groups and their actions. We shall also make use of this result to describe, for example, the set of zeros of the rate function I in the above theorem.

Continuing the line of development, we note that some of these results for random matrix products were obtained in our general setup of semisimple linear groups again following the initial works of Furstenberg, by Raugi, Guivarc'h, Goldsheid and Margulis [61]. The latter important work of Goldsheid-Margulis remakably clarified the algebraic essence of the aforementioned assumptions on the support of the probability measure, highlighting the role of the Zariski closure of the semigroup Γ generated by the support. Our main assumption, the Zariski density of Γ, appears along the same lines, namely with relation to the notion of proximality, but we also build on the later algebraic results and constructions of Abels-Margulis-Soifer [16] and Benoist [10], [11], on loxodromic elements and Schottky semigroups. We finally would like to mention again that recently, the theory of random matrix products and random walks on semisimple and reductive groups were extensively restudied in a more general setup by Benoist-Quint [14], who also used, in parts, different methods to establish limit theorems. One concrete consequence of these different methods was the proof by Benoist-Quint [15], of the CLT under L^2 moment assumptions, instead of the previous finite exponential moment assumptions of Le Page.

3.5 To end the discussion of this part, we would like to announce our following result as a separate theorem for the reader who is only interested in the random matrix products. We emphasise that, when G is a semisimple group as above and V is irreducible as G-module, this result immediately follows as a corollary (Corollary 3.29) of our previous theorem by a general principle of large deviations theory called contraction principle (see Lemma 3.28). The additional generality of G in the below statement does not cause a further essential difficulty and is dealt with in Section 3.4.

Theorem 1.4 (Simplified)**.** *Let μ be a probability measure of finite exponential moment on $GL(V)$ such that the Zariski closure of the semigroup generated by the support of μ is a trivial central extension of a semisimple linear real algebraic group in $GL(V)$ (e.g. $GL(V)$). Then, the sequence of random variables $\frac{1}{n}\log ||S_n||$ satisfies an LDP with a proper convex rate function.*

4.1 To introduce Chapter 4, let S be a bounded subset of a connected semisimple linear real algebraic group G. This time, we undertake a deterministic study of the sequence $S, S^2, \ldots$, again through classical decompositions of G, such as Cartan and Jordan decompositions. In summary, for each such subset S of G, we associate a compact set in the subalgebra $\mathfrak{a}$ of the Lie algebra $\mathfrak{g}$, which we call the joint spectrum of S. It encodes asymptotic behaviour of Cartan and Jordan projections of elements the sets S^n. En passant, using this notion of joint spectrum and its properties, we give a detailed study of the rate function obtained in Theorem 1.1. Before making these explicit and stating our main results, let us begin by mentioning some related notions to put our results into perspective.

4.2 Let us start by recalling the Jordan decomposition and the corresponding projection for such a group G. The Jordan decomposition says that each element $g \in G$ can be written uniquely as a commuting product of an elliptic, hyperbolic and unipotent elements $g = g_e g_h g_u$. The hyperbolic element g_h is conjugate to a unique element $\exp(x_g)$ with x_g belonging to a chosen Weyl chamber $\mathfrak{a}^+$. The mapping $\lambda : G \longrightarrow \mathfrak{a}^+$, associating to $g \in G$, this element $x_g \in \mathfrak{a}^+$ is called the Jordan projection. For the particular case $G = SL(d, \mathbb{R})$, the Jordan projection $\lambda(g)$ is related to the eigenvalues of g ; more precisely, for $g \in G$, denoting by $|\lambda_1(g)| \geq \ldots \geq |\lambda_d(g)| > 0$ the eigenvalues of g, we can write $\lambda(g) = (\log|\lambda_1(g)|, \ldots, \log|\lambda_1(g)|)$.

Furthermore, recall the notion of joint spectral radius $r(S)$ of a bounded subset S in a normed algebra $\mathcal{A}$ with a submultiplicative norm $||.||$, introduced by Rota-Strang in [114] as $r(S) := \lim_{n\to\infty} \sup\{||x||^{\frac{1}{n}} \mid x \in S^n\}$. When S is a singleton, $r(S)$ is indeed the spectral radius of the corresponding element by the classical spectral radius formula. Unlike for a singleton, for an arbitrary bounded set S, there exist other numerical values such as joint spectral subradius, $r_{sub}(S) := \lim_{n\to\infty} \inf\{||x||^{\frac{1}{n}} \mid x \in S^n\}$ which expresses norm asymptotics of the sequence $S, S^2, \ldots$; these numerical characteristics have been considered by several authors [114], [43], [24], [31], [25], [33], ...

4.3 In this work, for convenience, we restrict to a group G as above, and a bounded subset S, which we suppose to generate a Zariski dense sub-semigroup in this introduction. We define the joint spectrum $J(S)$ of S as the Hausdorff limit of the sequence $K_n(S) = \{\frac{\kappa(g)}{n} \mid g \in S^n\} \subset \mathfrak{a}^+$. In this way, we unify and extend the above different numerical characteristics to a compact subset $J(S)$ of $\mathfrak{a}^+$, which, as we will see, clearly involves more comprehensive asymptotic information on the sequence $S, S^2, \ldots$. The author would like to point out that this notion was suggested to him by Emmanuel Breuillard and it also appears in close connection to the large deviations considerations of the first part (see Theorem 1.7). We also want to underline that a notion of joint spectrum can be defined in a similar way in a greater generality, namely for a bounded subset of a matrix algebra over a local field.

Before stating our main result on the joint spectrum, let us also recall a corresponding numerical invariant for a bounded subset S of a matrix algebra $M_n(\mathbb{C})$. This is called the generalised spectral radius of S, and is related to the eigenvalues of elements of S^n. It was introduced by Daubechies-Lagarias in [43], as $\rho(S) = \limsup_{n\to\infty} \sup\{\lambda_1(x)^{\frac{1}{n}} \mid x \in S^n\}$. Conjectured by Daubechies-Lagarias, an important result of Berger-Wang [24] (later proven with a more precise statement by different methods by Bochi [25], Breuilllard [33]) says that for bounded subsets S of $M_n(\mathbb{C})$, we have the equality $r(S) = \rho(S)$. In our setting, similarly to $K_n(S)$'s define, $\Lambda_n(S) := \{\frac{\lambda(g)}{n} \mid g \in S^n\}$.

We now state our second main result. It summarises the main properties of the joint spectrum $J(S)$. We note also that its first point generalises the Berger-Wang equality in our setting.

Theorem 1.5. *Let G be a connected semisimple linear real algebraic group and S a bounded subset of G generating a Zariski dense sub-semigroup.*

1. *The following Hausdorff limits exist, and we have the equality :*

$$\lim_{n\to\infty} K_n(S) = J(S) = \lim_{n\to\infty} \Lambda_n(S)$$

2. *$J(S)$ is a compact convex set of non-empty interior (i.e. a convex body) in $\mathfrak{a}^+$.*

Remark 1.6. *In Section 4.1 we also precise a region, optimal in some aspects, bounding the convex body $J(S)$, using hyperplanes defined by looking at the classical joint spectral radii in different representations of G (see Fig. 4.1).*

4.4 At this juncture, we would like to mention a closely related notion, that of limit cone, introduced by Benoist in [11]. Benoist associates to each Zariski dense sub-semigroup Γ in G as above, a closed cone in $\mathfrak{a}^+$, that we call the Benoist limit cone B_Γ of Γ (see Section 5.2). It describes asymptotic directions of elements of Γ in the Cartan and Jordan projections. He proves that for such a Γ, B_Γ is a closed convex cone of non-empty interior. This result of Benoist is in fact a precursor of Theorem 1.5 inasmuch as it is easily seen that for a bounded S generating Γ, B_Γ equals to the cone generated by the joint spectrum $J(S)$ of S, so that these properties of B_Γ follows from Theorem 1.5. At this point we wish to draw attention to the fact that, although that $J(S)$ is of non-empty interior can be deduced from the corresponding result for B_Γ of Benoist, here, we give another quick proof (partly dependent on Benoist's result, see below), making use of the theory of random matrix products, namely combining the central limit theorem of Goldsheid-Guivarc'h [64] (for $G = \mathrm{SL}(d, \mathbb{R})$) and Guivarc'h [65] (more generally for linear semisimple G as before) with the Abels-Margulis-Soifer [16] finiteness result. This also gives a new proof for the corresponding result on Benoist cone in the case of $G = SL(d, \mathbb{R})$, but for more general G, we stress that the CLT of Guivarc'h uses this property of Benoist cone. We are also informed that this possible use of CLT to establish this fact for the Benoist limit cone was mentioned by Guivarc'h [38]. Moreover, we would like to also signal that first Quint [104] and then Guivarc'h [65] have given other proofs of this non-empty-interior property of the Benoist cone, as a result of more precise studies they conduct, the first one, on the density of the abelian group generated by the Jordan projections of a Zariski dense semigroup, and the second, on the projections of loxodromic elements on the centraliser of a maximal split torus.

4.5 As mentioned earlier, the joint spectrum of a bounded set $S \subset G$ is closely related to large deviations considerations on products of random elements of S. In the second part of Chapter 4, we give a detailed study of the rate function appearing in Theorem 1.1, and in particular, put into evidence its relation to the joint spectrum. To give an idea of this relation, let μ be a probability measure of support $S \subset GL(d, \mathbb{R})$ and suppose that the normalised norms $\frac{1}{n} \log ||S_n||$ of the μ-random walk satisfy an LDP with a rate function I (see Theorem 1.4). We then easily observe upon its definition that the joint spectral radius $r(S)$ of S is an upper bound to the effective support of I, i.e. the set $\{x \in \mathbb{R} \mid I(x) < \infty\}$. In this sense, 1. of the following theorem is an extensive translation of this observation to joint spectrum, including also a 'converse statement'. The continuity properties of 2. follows basically from convexity of I and 1., namely from the fact that D_I is of non-empty interior. Lastly, the unique zero property is a consequence (in fact, equivalent to) of Le Page's exponential decay result in Theorem 1.3.

Theorem 1.7. *Let G and μ be as in Theorem 1.1, S denote the support of μ, and let $D_I = \{x \in \mathfrak{a} \mid I(x) < \infty\}$ be the effective support of the rate function given by Theorem 1.1. We then have*

1. *D_I is a convex set of non-empty interior in $\mathfrak{a}^+$, satisfying, if S is moreover bounded, $\overline{D}_I = J(S)$ and $\overset{\circ}{D}_I = \overset{\circ}{J(S)}$, and finally, if S is moreover finite, $D_I = J(S)$.*

2. *I is convex, hence locally Lipschitz on $\mathring{D}_I$, with a unique zero, corresponding to the Lyapunov vector $\vec{\lambda}_\mu \in \mathfrak{a}^+$ of μ.*

5.1 Chapter 5 is closely related to Chapter 4 : this time we suppose that the subset $S \subset G$, as therein, is finite, and we take on a study of exponential growth of the number of elements in S^n's with respect to the asymptotic behaviour of these elements in the Cartan and Jordan projections. We introduce exponential growth indicator functions for finite sets in a G as before ; they generalise the classical notion of exponential growth rate (see below).

5.2 To explain our approach more precisely, let us recall a classical notion : let T be a finite set in a semigroup Γ and $|T|$ denote the number of elements in T. The limit $v_T := \lim_{n\to\infty} |T^n|^{\frac{1}{n}}$ exists by submultiplicativity and is called the exponential growth rate of T. The semigroup Γ is said to be of exponential growth if there exists a finite generating set T with $v_T > 1$. In fact, this does not depend on a particular T ; it is a property of Γ. Closer to our setting, in the case of a linear group Γ (i.e. $\Gamma \leq GL(d, \mathrm{k})$ for some $d \in \mathbb{N}$, and field k), it follows by classical results of Milnor-Wolf and Tits alternative that Γ is of exponential growth whenever it is not nilpotent-by-finite. Moreover, by uniform versions of Tits alternative obtained by Eskin-Mozes-Oh [51], Breuillard-Gelander [34] and Breuillard [35], the exponential growth of Γ enjoys a stronger property called uniform exponential growth (see Section 5.1).

5.3 Let now, in this introduction, the finite set S generate a Zariski dense semigroup in a G as above. Inspired by the expression à la Ruelle-Lanford [106], [83] of the rate function of an LDP (see Theorem 2.4) and Quint's earlier work in a very similar setting [100], we introduce the following counting functions for the subset S, which extend the data of the numerical growth rate v_S to, basically, the data of a function on the joint spectrum $J(S)$. We borrowed the terminology from Quint [100].

Definition 1.8. *The function $\phi_S : \mathfrak{a} \longrightarrow \mathbb{R}_+ \cup \{-\infty\}$, defined by, $\phi_S(x) := \inf_{\substack{O \text{ open in } \mathfrak{a} \\ x\in O}} \limsup_{n\to\infty} \frac{1}{n} \log \#\{g \in S^n \mid \frac{1}{n}\kappa(g) \in O\}$ will be called the (Cartan) growth indicator of S. The Jordan growth indicator ψ_S of S is defined in the same manner with the Jordan projection λ.*

Remark 1.9. *1. The extension of the data of numerical exponential growth rate v_S of S by the growth indicators of S, should be compared to the extension of the data of numerical joint spectral radii to a convex body, i.e. the joint spectrum.*

2. It turns out (see Section 5.1) that the growth indicator ϕ_S is in fact, concretely related to large deviation considerations for a sequence of probability measures of deterministic nature, more precisely, the sequence of the images by normalised Cartan projection of the uniform probability measures on S^n's.

5.4 We summarise our findings on the growth indicators in the following theorem. We want to point out that these functions come also with several nonfindings, i.e.

natural open questions on them. In the following result, we note that the second assertion says in particular that we can read off the numerical value v_S on the growth indicators of S, and the last assertion signifies that we have an exponential growth for S on a dense set of asymptotic behaviours, which are naturally parametrised by the points of joint spectrum $J(S)$, to put it briefly, S has a dense exponential growth :

Theorem 1.10 (Simplified)**.** *Let S be a finite subset of a connected semisimple linear real algebraic group G, generating a Zariski dense semigroup in it. Then,*

1. *The growth indicators ϕ_S and ψ_S are upper semicontinuous with maximum* $\log v_S$.
2. *We have $\phi_S \leq \psi_S$.*
3. *We have the following equality of sets : $\overline{\{\phi_S \geq 0\}} = \overline{\{\psi_S \geq 0\}} = J(S) = \overline{\{\phi_S > 0\}}$*

5.5 In the second part of Chapter 5, we exhibit previous works of Benoist [11] and Quint [100] : for a Zariski dense semigroup Γ in a G as before, Benoist introduces and studies the notion of a limit cone B_Γ of G in a fixed Weyl chamber $\mathfrak{a}^+$; Quint in his turn, for a discrete such group Γ, introduces an exponential counting function on $\mathfrak{a}$, the growth indicator ψ_Γ of Γ, which is closely related to B_Γ, and studies its properties (such as concavity). These two notions are in close analogy with, respectively, our joint spectrum and growth indicators. We make these analogies more precise. For Benoist cone, as mentioned earlier, our results in Chapter 4, permits us to recover (Corollary 5.15) some of the properties of B_Γ, and some others transfers as open questions. For Quint's growth indicator ψ_Γ, we indicate that the difference of our growth indicators comes, partly, from a different way of counting the same object. Using this observation, we establish some relations between Quint's and our growth indicators, using a 'directional joint spectral radius' function that we introduce. Finally, we note also that ψ_Γ's remarkable concavity, which is in line with the convexity of rate functions of above LDP's, translates to the setting of our growth indicators as an open question, and will be investigated in a subsequent work.

5.6 The last section of Chapter 5 is a collection of miscellaneous results. First, we study LDP properties for Jordan projections of random walks in the setting of Theorem 1.1 : we indicate that, in fact, from the LDP for Cartan projections in this theorem, using Abels-Margulis-Soifer finiteness result together with Benoist's estimates for Cartan and Jordan projections of loxodromic elements (see the text for these results and notions), one can deduce the lower inequality in the definition of the LDP, with the same rate function as in Theorem 1.1, for the Jordan projections (this is very similar to 2. of Theorem 1.10). Furthermore, we establish the analogue of Theorem 1.1 for Jordan projections, but for the very particular setting of (r, ϵ)-Schottky random walks (see therein). Second, by studying a particular example, we reach to, and indicate a strategy to improve the dense exponential growth theorem of Section 5.1. And finally, using the notion of joint spectral subradius, we set forth a discreteness criterion for a finitely generated semigroup Γ in a G as usual, and show that this applies to finitely generated (r, ϵ)-Schottky semigroups.

6.1 In the last part, Chapter 6, of this text − mostly independent of the previous parts − we again adopt a probabilistic viewpoint and, to put it simply, seek an LDP for the sequence of random variables $\frac{1}{n}\log||S_n v||$ where S_n denotes the n^{th} step of a random walk on $GL(d,\mathbb{R})$ and $v \in \mathbb{R}^d \setminus \{0\}$. Not having been able to apply our techniques of the previous parts, we pursue Furstenberg's initial idea to see these variables as functionals over a Markov chain, and we set on to translate the general theory of Markov chains to our setting. At the end, under considerably restrictive conditions (of absolute continuity of a convolution power as Tutubalin above, plus a boundedness condition that we make explicit), we establish an LDP for these variables and their multidimensional generalisations (Iwasawa cocycle) as a corollary of a stronger conclusion that we reach.

6.2 Since it is brief, let us explain our strategy more precisely : it is based on a theorem (Theorem 6.2) of Stroock [109] and Ellis [50], which gives a sufficient condition, called uniformity, on the Markov transition kernel of a Markov chain M, for the empirical measures of M to satisfy an LDP (on the space of probability measures on the state space of M). As in Benoist-Quint's [14], we first follow a general setup and deduce our results as applications in particular cases ; for a locally compact second countable group G acting transitively and continuously on a metrisable compact space X, we consider a Markov chain on $G \times X$ associated to a random walk on G governed by a probability measure μ on G. We then transfer the uniformity condition of Stroock-Ellis to a rather technical condition **(D)** on μ. To clarify **(D)**, we also show that, for example, it is satisfied whenever μ is of compact support and has a convolution power μ^{*n} absolutely continuous with respect to the Haar measure on G with a density bounded below by an $\alpha > 0$ on a neighbourhood of identity generating the group G (for greater generality, see the end of Section 6.1). Finally, by using a general technique of large deviations theory, the contraction principle : 1. we take $G = GL(V)$, $X = \mathbb{P}(V)$ and transfer the LDP under **(D)** with a transfer function that we construct using the norm cocycle, 2. we take G a connected semisimple linear real algebraic group, $X = \mathcal{F}_G$ its flag variety, and this time we transfer the LDP using the Iwasawa cocycle σ, which we explain in Section 6.2. As a result, on these particular cases, we obtain

Theorem 1.11. *For G and X respectively as in* 1. *and* 2. *above and μ a probability measure on G satisfying **(D)**, the sequences of random variables $\frac{1}{n}\log||S_n v||$ and $\frac{1}{n}\sigma(S_n, \eta)$ satisfy LDP, uniformly in v with $||v|| = 1$ and $\eta \in \mathcal{F}_G$, with proper convex rate functions respectively on $\mathbb{R}$ and $\mathfrak{a}$.*

6.3 At the very end, we set on to a study of the rate functions appearing in the previous theorem, namely of their effective support $\{x \mid I(x) < \infty\}$. We content with providing some bounding regions for these convex sets respectively in $\mathbb{R}$ and $\mathfrak{a}$, and reach to further open questions.

Further directions

7.1 Due in part to the novelty of several objects that we introduce, this text comes with numerous open questions and directions for further research. Most of these are mentioned in the text, at the concerned places. Moreover, as indicated therein, many of these questions are illustrated in "L'exemple du groupe libre" through the corresponding results in this example. In the following, let us briefly summarise these open questions and further research directions.

7.2 In Chapter 3, the first natural question is indeed whether one can omit the main hypothesis of Zariski density in Theorem 1.1. It seems plausible that a weak LDP exists (for $\frac{1}{n}\kappa(S_n)$'s) without any hypothesis on the governing probability measure μ (see Theorem 3.1). A further direct question is about the regularity of the rate function I appearing in this theorems, i.e. whether it is strictly convex or differentiable/analytic in the interior of D_I.

Concerning the LDP for normalised Jordan projections along the random walks ($\frac{1}{n}\lambda(S_n)$), we suspect that at least under the same Zariski density assumption, one can prove its existence (see Remark 4.19). We address this question, and give partial answers in Section 5.3. A later question on this issue will be then about the relation of Cartan and Jordan rate functions (see Proposition 5.22, Corollary 5.23 and Corollary 5.29 as well as "L'exemple du groupe libre"). At the same time, let us note that one can not expect the existence of an LDP for $\frac{\lambda(S_n)}{n}$'s without any hypothesis on the support of the governing probability measure : this interestingly follows from an example of Breuillard on the non-Hausdorff convergence to joint jordan spectrum (see 2. of Example 4.17 and 1. of Remark 4.19).

Another question that one can ask related to Theorem 1.1 is about weakening the hypotheses in a different direction : that of non-independent random products. In this direction, it seems plausible that an exponential mixing condition, or at least a super-exponential one, would be sufficient to establish the existence of an LDP.

Yet another direction concerns LDP's for continuous time processes (on semisimple Lie groups and their symmetric spaces), i.e. Brownian motion or as a direct generalisation of discrete time independent random walks, Lévy processes (see [85]). Our techniques may apply in these studies and these will be investigated in a future work.

Finally, a question that we address in a work in progress concerns the equivalent of Theorem 1.1 on one hand for semisimple linear algebraic groups defined over other local fields than $\mathbb{R}$ (see Remark 2.26), on the other, for reductive linear algebraic groups.

7.3 Concerning Chapter 4, i.e. the joint spectra, immediate questions concern the existence of Hausdorff limits with weaker assumptions : for the Cartan case, it again seems plausible that the sequence $\frac{\kappa(S^n)}{n}$ converge without any hypothesis on S. For Jordan projections, even though a simple condition like $e \in S$, or Zariski density of $\cup_{n\geq 1} S^n$ are sufficient (see 2. of Theorem 4.4) for the convergence of $\frac{\lambda(S^n)}{n}$, we do not have convergence without any hypothesis on S (see the notable 1. and 2. of Example 4.17). We note also that without the Zariski density assumption, one can not expect same properties of the joint spectrum (e.g. $J(S)$ being of non-empty interior). One can easily arrange S such that, for example, $J(S)$ lives on a wall of the Weyl chamber.

A second kind of question about the joint spectrum concerns the possible shapes : more precisely, which convex bodies in $\mathfrak{a}^+$ of a G can be the joint spectrum of a subset S of G (see the second subsection of Section 4.1). This question appears in relation with the corresponding results on Benoist cone ([11]) (see also Proposition 5.13).

Another direction that the author plans to study concerns the consideration of joint spectrum in a more general setting. As mentioned in the text, it is easily seen that the actual definition transfers almost verbatim to $M_n(k)$'s for a local field k.

Finally, in a work in progress, we study a third kind of spectrum related to the Iwasawa decomposition. We define it in a similar way, this time using the Iwasawa projection (see "L'exemple du groupe libre" and Remark 6.30) and we shall study its properties as well as its relation to joint spectrum (compare Fig. 4.1 and Fig. 6.1).

7.4 In Chapter 5, the study of growth indicators is in fact of a rather unaccomplished character, it mostly consists of results of the author (dense exponential growth Theorem 1.10) from an ongoing project. Some of the main questions about the growth indicators that we are addressing in this project are as follows : 1. Are the growth indicators 'profile functions'? By this, we mean : do we get the same functions if we change $\limsup$ by $\liminf$ in their definition? We remind that this is equivalent to asking whether the normalised push-forwards by Cartan/Jordan projections of uniform probability measures on S^n's satisfy an LDP (see Remark 5.2). 2. Are the growth indicators concave? (see Remark 5.9, Remark 5.18 and the paragraph following it) This question is related to the corresponding concavity result for Quint's growth indicator ψ_Γ (see Theorem 5.17 and [100]). 3. A third question concerns the study of set of maxima of growth indicators (see Remark 5.7 and "L'exemple du groupe libre"). We note that in the probabilistic setting, this corresponds to the study of set of zeros of the rate functions of LDP's : for Cartan projections, this is basically Le Page's Theorem 1.3. 4. A fourth question is about understanding the relation between two growth indicators ϕ_S and ψ_S as well as ϕ_S's relation to the rate function I of Theorem 1.1 (see Remark 5.2). This question seems to be related to Remark 7.5.

7.5 The questions arising about the content of Chapter 6 are naturally similar to those of Chapter 3. We note that in Chapter 6, to establish LDP for Iwasawa cocycle along random walks we had to assume a much more stronger hypothesis **(D)** on the governing probability measure μ. For a first weakening of this hypothesis, the recent techniques developped by Guivarc'h-Le Page in [66] might be of use. Some further questions are about understanding the relation between the rate functions emerging in this chapter and the joint spectral subradius, more generally, the Iwasawa joint spectrum (see Remark 6.26).

Chapitre 2

PRELIMINARY TOOLS

In Section 2.1, we give some basic definitions of large deviations theory, which we need to state the more precise versions of large deviations theorems stated in the previous introduction chapter. We also include some fundamental results of this theory that will be used on several occasions.

In Section 2.2 we expose a quantitative theory of loxodromic/$\mathbb{R}$-regular elements and Schottky semigroups as it was developed by Abels-Margulis-Soifer and Benoist. We also single out some definitions and remarks that will be put to good use in the sequel. We note that we pay particular attention to offer the reader a parallel reading of the rest, in the particular case where the group G in question is taken to be $SL(n, \mathbb{R})$.

2.1 Definitions and tools from large deviations theory

We give the precise definitions of an LDP and its weak version, indicate main existence theorem of a weak LDP and note the notion of exponential tightness of sequences of probability measures, which permits one to strengthen a weak LDP to a full LDP. More particular results of large deviations theory will be mentioned in the related sections.

Let X be a topological space. In the sequel, for convenience and in its relation to an LDP, we refer to a lower semicontinuous mapping $I : X \longrightarrow [0, +\infty]$ as a rate function. Now, let $\mathcal{F}$ be a σ-algebra on X (not necessarily Borel), and $(\mu_n)_{n\in\mathbb{N}}$ a sequence of probability measures on $(X, \mathcal{F})$.

Definition 2.1. *The sequence of probability measures μ_n (or equivalently, a sequence of X-valued random variables Z_n of laws μ_n) is said to satisfy the large deviation principle (LDP) with rate function I, if for every $\mathcal{F}$-measurable set E, we have*

$$-\inf_{x\in\overset{\circ}{E}} I(x) \leq \liminf_{n\to\infty} \frac{1}{n} \log \mu_n(E) \leq \limsup_{n\to\infty} \frac{1}{n} \log \mu_n(E) \leq -\inf_{x\in\overline{E}} I(x)$$

For precision, we note that in the rest of this text, we only work with Borel σ-algebras and we will work with a probability space fixed once for all. Now, reinterpreting the above definition, an LDP with rate function I is equivalent to the following :

1. (Upper bound) For any closed set $F \subset X$, $\limsup_{n\to\infty} \frac{1}{n} \log \mu_n(F) \leq -\inf_{x\in F} I(x)$
2. (Lower bound) For any open set $O \subset X$, $\liminf_{n\to\infty} \frac{1}{n} \log \mu_n(O) \geq -\inf_{x\in O} I(x)$

The following is the definition of a weak LDP ; it only slightly differs from the LDP, that we also refer to as full LDP to distinguish, and this on the upper bound :

Definition 2.2. *A sequence of probability measures (μ_n) on X is said to satisfy the weak LDP with a rate function I if the upper bound 1. (above) holds for all compact sets and the lower bound 2. holds the same, for all open sets in X.*

The following remark settles the uniqueness of the rate function issue in our setting. We refer to [44] (Lemma 4.1.4 and the subsequent remark therein).

Remark 2.3. *If X is locally compact or a polish space and a sequence of probability measures μ_n on X satisfies a weak LDP with a rate function I, then I is unique.*

The following is the main existence theorem for a weak LDP. We cite it from Dembo-Zeitouni's [44] (Theorem 4.1.11). As the authors mention, variants of this theorem can be traced back to the works of Ruelle [106] and Lanford [83] in statistical mechanics. In Chapter 5, in a deterministic setting, we define some growth functions for groups, inspired by the expressions in this theorem (by also Quint's work in [100]).

Theorem 2.4 (Existence of a weak LDP theorem)**.** *[44] Let X be a topological space endowed with its Borel σ-algebra β_X, and μ_n be a sequence of probability measures on (X, β_X). Let $\mathcal{A}$ be a base of topology for X. For each $x \in X$, define :*

$$I_{li}(x) := \sup_{\substack{A\in\mathcal{A}\\ x\in A}} -\liminf_{n\to\infty} \frac{1}{n} \log \mu_n(A)$$

$$I_{ls}(x) := \sup_{\substack{A\in\mathcal{A}\\ x\in A}} -\limsup_{n\to\infty} \frac{1}{n} \log \mu_n(A)$$

Suppose that for all $x \in X$, we have $I_{li}(x) = I_{ls}(x)$. Then, the sequence μ_n satisfies a weak LDP with rate function I, where $I(x) := I_{li}(x) = I_{ls}(x)$

Remark 2.5. *In a polish space X, the hypothesis of the preceding theorem is actually equivalent to the existence of a weak LDP (see Theorem 4.1.18 and the following remark in [44]).*

As noted earlier, an interest of the following notion is that it enables one to formulate a sufficient condition (see Lemma 2.7) to strengthen a weak LDP to an LDP. We note that, in our setting, if the rate function I of the LDP is proper, this condition turns out to be also necessary.

Definition 2.6. *A sequence of probability measures μ_n on X is said to be exponentially tight, if for all $\alpha \in \mathbb{R}$, there exists a compact set $K_\alpha \subset X$ such that $\limsup_{n\to\infty} \frac{1}{n} \log \mu_n(K_\alpha^c) < -\alpha$.*

With this definition, we have the following result, for which, we refer the reader again to [44] (Lemma 1.2.18) :

Lemma 2.7. *If an exponentially tight sequence of probability measures on X satisfies a weak LDP with a rate function I, then it satisfies a (full) LDP with a proper rate function I.*

2.2 Tools from (r,ϵ)-Schottky semigroup theory

In this section, we mainly borrow from Benoist [10], [11], [12], [13] and Abels-Margulis-Soifer [16], providing also some new definitions and remarks. We highly recommend [13] for the reader who wishes to familiarise more in detail with the useful notions of this section. While most of the result of this section will be quoted from the above sources without proofs, we will include the proofs of some of them. The main reason for our doing so is to give a feeling of the way in which one deals with the general semisimple case, to the reader who might possibly be unfamiliar with this generality. Nevertheless, if the reader wishes, s/he can always suppose that the group G in question is $SL_d(\mathbb{R})$ and we will indicate the corresponding notions for this particular case, in Example 2.14.

Given a metric space (X,d) and two subsets Y,Z of X ; we denote $d(Y,Z) := \inf_{y\in Y, z\in Z} d(y,z)$, and $d_H(Y,Z)$ the Hausdorff distance between Y and Z ; $d_H(Y,Z) = \sup_{y\in Y} d(y,Z) \vee \sup_{z\in Z} d(z,Y)$.

Let V be a finite dimensional real vector space, $X = \mathbb{P}(V)$ its projective space. Endowing V with a euclidean norm $||.||$, we will work with the Fubini-Study metric on X : for $x,y \in X$, denoting by v_x and v_y any two vectors in V projecting respectively on x and y, we have $d(x,y) := \frac{||v_x \wedge v_y||}{||v_x||.||v_y||}$, where the Euclidean norm on $\bigwedge^2 V$ is defined in the usual manner from the scalar product on V. In the sequel, we will also denote by the same $||.||$ notation, the operator norm on the linear endomorphisms of V associated to the norm $||.||$ on V.

We first start with a brief discussion of the proximality of a linear transformation. This notion relates to an important property of the dynamics of projective actions of linear transformations and in that, it is, for example, of essential use in the Tits' original proof of the Tits alternative in [111] through the so called ping-pong lemma. It is also in close relation to Furstenberg's earlier (quasi-) projective transformations [56]. See also Breuillard-Gelander's [34] for a more detailed account and Quint's [101] and [100] for a generalisation.

For $g \in End(V)$, denote by $\lambda_1(g)$ the spectral radius of g. An element $g \in End(V)$ is said to be proximal in $\mathbb{P}(V)$ if it has a unique eigenvalue $\alpha \in \mathbb{C}$ (therefore, in $\mathbb{R}$) such that $|\alpha| = \lambda_1(g)$, and this eigenvalue is simple. Denote by x_g^+, the element of X corresponding to the one dimensional eigenspace corresponding to α. Let v_g^+ be a vector of norm 1 on this line, and $V_g^<$ the supplementary g-invariant hyperplane, and put $X_g^< := \mathbb{P}(V_g^<) \subset X$.

The following definition singles out special proximal elements : let $0 < \epsilon \leq r$ and set $b_g^\epsilon := \{x \in X \mid d(x, x_g^+) \leq \epsilon\}$ and $B_g^\epsilon := \{x \in X \mid d(x, X_g^<) \geq \epsilon\}$.

Definition 2.8 ([16],[10]). *Let $0 < \epsilon \leq r$. An element $g \in End(V)$ is said to be (r,ϵ)-proximal in $\mathbb{P}(V)$, if $d(x_g^+, X_g^<) \geq 2r$, $g(B_g^\epsilon) \subset b_g^\epsilon$, and $g_{|B_g^\epsilon}$ is an ϵ-Lipschitz mapping.*

Remark 2.9.
1. *The notion of (r,ϵ)-proximality, as well as the numbers $0 < \epsilon \leq r$ depend on the choice of the norm on V.*
2. *Nevertheless, it is not hard to see that for every proximal transformation g and for any choice of norm on V, there exists $r > 0$ such that for all $k \in \mathbb{N}$ large enough, g^k is (r,ϵ_k)-proximal with $\epsilon_k \underset{k\to\infty}{\longrightarrow} 0$.*

Two properties of (r,ϵ)-proximal linear transformations

The following lemma says that for $\epsilon > 0$ small enough, the spectral radius of an (r,ϵ)-proximal transformation can be controlled by the operator norm of this transformation :

Lemma 2.10. *Let V be a finite dimensional real vector space and $0 < \epsilon \leq r$. Then, there exist constants $c_{r,\epsilon} \in]0,1[$ such that, for each $r > 0$, we have $\lim_{\epsilon\to 0} c_{r,\epsilon} = \frac{1}{2r}$, and for every endomorphism g of V, (r,ϵ)-proximal in $\mathbb{P}(V)$, we have*

$$c_{r,\epsilon}||g|| \leq \lambda_1(g) \leq ||g||$$

Démonstration. One notes that if $(g_k)_{k\in\mathbb{N}}$ is a convergent sequence of (r,ϵ_k)-proximal transformations with for all $k \in \mathbb{N}$ $||g_k|| = 1$ and $\epsilon_k \underset{k\to\infty}{\longrightarrow} 0$, then $\lim_{k\to\infty} g_k = \alpha p$, where α is a positive constant and p is a projection satisfying - denoting by v_p a non-zero vector in its image, $x_p \in \mathbb{P}(V)$ its projective image, and by $X_p \subset \mathbb{P}(V)$ the projective image of $\ker p$ - we have $d(x_p, X_p) \leq \frac{1}{2r}$. Since $||\alpha p|| = 1$, it follows by elementary computations that we have $\alpha \geq \frac{1}{2r}$, and the conclusion of lemma follows from the continuity of the application $\lambda_1(.)$. □

The following important proposition, proved by Benoist [10] (see also [12] or [13]), says that one can have a fairly good control over the spectral radii of the products of (r,ϵ)-proximal elements in terms of the spectral radii of the factors, given that the successive factors satisfy a natural geometric condition (see the hypothesis of the lemma).

We first define a useful notion of angle that appears when considering the products of proximal elements : for two proximal elements g, h of $End(V)$, denote by $\nu_1(g,h)$ the absolute value of the unique real number β satisfying $v_h^+ - \beta v_g^+ \in V_g^<$ (well defined by proximality of g in $\mathbb{P}(V)$). Furthermore, for $g_1,\ldots,g_l$, proximal elements of $End(V)$, putting $g_l = g_0$, set

$$\nu_1(g_l,\ldots,g_1) := \prod_{1\le j\le l} \nu_1(g_j, g_{j-1}) \tag{2.1}$$

Then, we have the following

Proposition 2.11. *For all real numbers $0<\epsilon\le r$, there exist constants $C_{r,\epsilon}>0$ and $D_{r,\epsilon}>0$ with the property that, for each $r>0$, we have $\lim_{\epsilon\to 0} D_{r,\epsilon}=1$, and $\lim_{\epsilon\to 0} C_{r,\epsilon}=C_r$, where C_r is a positive constant depending only on r, and such that if $g_1,\ldots g_l$ are (r,ϵ)-proximal linear transformations of V satisfying (putting again $g_l=g_0$) $d(x^+_{g_{j-1}}, X^<_{g_j})\ge 6r$, for all $j=1,\ldots l$, we have that for all $n_1,\ldots,n_l\ge 1$, the linear transformation $g=g_l^{n_l}\ldots g_1^{n_1}$ is $(2r,2\epsilon)$-proximal, and*

$$D^{-l}_{r,\epsilon}\nu_1(g_l,\ldots,g_1) \le \frac{\lambda_1(g_l^{n_l}\ldots g_1^{n_1})}{\lambda_1(g_l)^{n_l}\ldots\lambda_1(g_1)^{n_1}} \le D^{l}_{r,\epsilon}\nu_1(g_l,\ldots,g_1)$$
$$C^{-l}_{r,\epsilon}\nu_1(g_l,\ldots,g_1) \le \frac{||g_l^{n_l}\ldots g_1^{n_1}||}{\lambda_1(g_l)^{n_l}\ldots\lambda_1(g_1)^{n_1}} \le C^{l}_{r,\epsilon}\nu_1(g_l,\ldots,g_1)$$

This proposition partly motivates the following definitions which we will be of important use to us in the sequel (see also Definition 1.7 in [12]) :

Definition 2.12.
1. *A subset E of $GL(V)$ is called an (r,ϵ)-Schottky family in $\mathbb{P}(V)$ if*
   (a) For all $\gamma\in E$, γ is (r,ϵ)-proximal in $\mathbb{P}(V)$, and
   (b) $d(x^+_\gamma, X^<_{\gamma'})\ge 6r$, for all $\gamma,\gamma'\in E$.
2. *Let $E\subset GL(V)$ be a subset consisting of proximal elements and $a,b\ge 0$ two real numbers. We say that the set E is (a,b)-narrow in $\mathbb{P}(V)$, if there exist subsets Y,Z of $\mathbb{P}(V)$ of diameters respectively less than a and b, and such that we have $x^+_\gamma\in Y$ and $\mathbb{P}((V^<_\gamma)^\perp)\in Z$ for each $\gamma\in E$.*
3. *A sub-semigroup Γ of $GL(V)$ is said to be ((a,b)-narrow) (r,ϵ)-Schottky in $\mathbb{P}(V)$, if there exists an ((a,b)-narrow) (r,ϵ)-Schottky family generating Γ.*

Remark 2.13.
1. *Observe that, since for the Fubini-Study metric, we have $diam(\mathbb{P}(V))=1$, any set of proximal elements is $(1,1)$-narrow ; said differently, in the definition of (a,b)-narrowness, if a or b is greater or equal to 1, this means that there is no restriction on the location of the corresponding directions in $\mathbb{P}(V)$. For the same reason, in the definition of an (r,ϵ)-Schottky family, we necessarily have $r\le\frac{1}{6}$.*
2. *Note that, by definition, a Schottky family (i.e. (r,ϵ)-Schottky family, for some $r\ge\epsilon>0$) can not contain an element $g\in GL(V)$ and its inverse g^{-1} at the same time.*

The notion of proximality relates, so to say, to only one special direction of the action of a linear transformation. We would like to have an equivalent property for the other/all eigenvalues and eigendirections. This property is reflected in the notion of a loxodromic transformation, which we shall shortly define.

To avoid repetitions, we move at this point to the case of a connected semisimple linear real algebraic group and define loxodromy in this setting. For this, we will need some standard notions and facts about the representation theory of such groups. We summarise these in the following paragraphs and in the next two lemmata. As mentioned earlier, in Example 2.14, we indicate all the corresponding objects in case where the group in question is $SL_d(\mathbb{R})$, so that one can continue reading this section, and the rest, by sticking to this example.

Let G be a connected semisimple linear real algebraic group, $A_G \leq G$ a Cartan subgroup, $A_G^+ \leq A_G$ a closed Weyl chamber, and K a maximal compact subgroup such that we have the Cartan decomposition $G = KA_G^+K$. One has $A_G = \exp(\mathfrak{a})$, $A_G^+ = \exp(\mathfrak{a}^+)$, where $\mathfrak{a}$ is a Cartan subalgebra of the Lie algebra $\mathfrak{g}$ of G, $\mathfrak{a}^+$ a chosen positive closed Weyl chamber. We also set $\mathfrak{a}^{++}$ as the interior of the closed Weyl chamber in $\mathfrak{a}$ and we note that exp is the exponential mapping of the abelian Lie algebra $\mathfrak{a}$, which is an isomorphism of the additive Lie group $\mathfrak{a}$ and the Cartan subgroup A_G. We denote its inverse by $\log : A_G \to \mathfrak{a}$ and move by this application from A_G to $\mathfrak{a}$, changing the multiplicative notation to the additive notation, whenever deemed convenient.

For each $g \in G$, there exists a unique element $a_g \in A_G^+$ such that $g \in Ka_gK$. The application $\kappa : G \longrightarrow \mathfrak{a}^+$, defined by $\kappa(g) = \log(a_g)$ is called the Cartan projection. It is a continuous proper application, conveying, in particular, information about asymptotics in G ; it is also closely related to the action of G on it symmetric space. Let also $\lambda : G \longrightarrow \mathfrak{a}^+$ be the Jordan projection defined as the application mapping an element g of G to the logarithm of the unique element in A_G^+ that is conjugated to the hyperbolic component of g in its Jordan decomposition.

For a character χ of A_G, $\chi : A_G \longrightarrow]0, \infty[$, we denote $\overline{\chi}$ the corresponding element in the dual space of the Cartan subalgebra $\mathfrak{a}$ of $\mathfrak{g}$, through the formula $\overline{\chi} = d_e\chi = \log \circ \chi \circ \exp$. Let R be the set of restricted roots of G, R^+ the positive roots, compatible with the choice of A_G^+, and $\pi = \{\alpha_1, \ldots, \alpha_r\}$ the set of simple roots, where $r \in \mathbb{N}$ stands for the real rank of G. Furthermore, denote by $\overline{\omega}_1, \ldots, \overline{\omega}_r$ the fundamental weights of $\mathfrak{g}$; they can be defined by the relation $\frac{2<\overline{\alpha}_i, \overline{\omega}_j>}{<\overline{\alpha}_i, \overline{\alpha}_i>} = \delta_{i,j}$ for all $i, j = 1, \ldots, r$, where the inner product is defined by the pairing by the restriction of the Killing form of $\mathfrak{g}$ on $\mathfrak{a}$. Any restricted dominant weight of $\mathfrak{g}$ (i.e. $\overline{w}$'s such that $\overline{w}(x) \geq 0$ for all $x \in \mathfrak{a}^+$) can be expressed as a non-negative integral linear combination of $\overline{w}_i$'s.

Let (V, ρ) be an irreducible rational representation of G in a finite dimensional real vector space. The choice of A^+ canonically induces a partial order on the set of characters of the abelian group A_G. By the representation theory of semisimple groups, the set $W(\rho)$ of restricted weights of A_G has a unique maximal element χ_ρ for this order, called the highest restricted weight of ρ. The representation (V, ρ) is said to be proximal if, denoting by V_{χ_ρ} the weight space of χ_ρ, we have $\dim V_{\chi_\rho} = 1$.

Example 2.14. *If one takes $G = SL_d(\mathbb{R})$, then we can write, $A_G = \{diag(\alpha_1, \ldots, \alpha_d) \in G \,|\, \alpha_i > 0$ for all $i = 1, \ldots, d\}$, $A^+ = \{g = diag(\alpha_1, \ldots, \alpha_d) \,|\, \alpha_1 \geq \alpha_2 \ldots \geq \alpha_d > 0\}$, $\exp(\mathfrak{a}^{++}) = A^{++} = \{g = diag(\alpha_1, \ldots, \alpha_d) \,|\, \alpha_1 > \alpha_2 \ldots >$*

$\alpha_d > 0\}$, and $K = SO_d(R)$. The exponential and the logarithm maps correspond to the application of the usual exponential and logarithm maps to the diagonal coefficients, so that, for instance, $\mathfrak{a} := \{(x_1, \ldots, x_d) \in \mathbb{R}^d \mid \sum_{i=1}^d x_i = 0)\}$. The Cartan projection $\kappa(.)$ associates to an element g of $SL_d(\mathbb{R})$, the element of $\mathfrak{a}$ consisting of the logarithms of the diagonal entries of the matrix A in KAK decomposition of g, i.e. it is the vector of logarithms of the singular values of g placed in decreasing order ; and the Jordan projection $\lambda(.)$, the same with logarithms of the modules of eigenvalues of g.

As examples of characters on A_G, we can exhibit L_i's for $i = 1, \ldots, d$, defined by $L_i(diag(a_1, \ldots, a_d)) = a_i$. The set of restricted roots are the restricted weights of the Ad representation of $SL(d, \mathbb{R})$, i.e. $R = \{\frac{L_i}{L_j} \mid i \neq j\}$. For our choice of A_G^+, the positive roots are $R^+ = \{\frac{L_i}{L_j} \mid i < j\}$ and the set of simple roots $\pi = \{\frac{L_i}{L_{i+1}} \mid i = 1, \ldots d-1\}$. On $\mathfrak{a}$, we have, for example, $\overline{(\frac{L_i}{L_j})}(x_1, \ldots, x_d) = x_i - x_j$. The fundamental weights are $\omega_i = \prod_{j=1}^i L_j$.

Some examples of proximal irreducible representations are $\sigma_1 = id$ or, more generally, $\sigma_i : SL(\mathbb{R}^d) \longrightarrow SL(\bigwedge^i \mathbb{R}^d)$ where $\sigma_i(g) := \bigwedge^i g$ for $i = 1, \ldots, d-1$. These are also the 'fundamental representations', meaning that their highest restricted weights are the fundamental weights ω_i's. The partial ordering corresponding to the choice of A_G^+ on the set of characters of A_G is simply described as : for $\chi_1, \chi_2 : A_G \to]0, \infty[$, we have $\chi_1 \geq \chi_2 \iff \chi_1(a) \geq \chi_2(a)$ for all $a \in A_G^+$. Finally, for a representation (V, ρ) and $\chi \in W(\rho)$, the corresponding weight space is $V_\chi = \{v \in V \mid \rho(a)v = \chi(a)v$ for all $a \in A_G\}$.

For the remaining part of this article, we will **fix** the family of representations given by the next lemma and refer to them as the distinguished representations of G. In the proof, for the existence of fundamental representations in a more general situation, we refer the reader to Tits' influential work [113].

Lemma 2.15. *Let G be a connected semisimple linear real algebraic group of real rank d. Then, there exist d proximal irreducible representations ρ_i of G in real vector spaces V_i of highest restricted weights $(\chi_i)_{i=1,\ldots,d}$, which are powers of the fundamental weights ω_i of G, and such that the mapping from $\mathfrak{a}$ to $\mathbb{R}^d$ defined by $a \mapsto (\overline{\chi}_1(a), \ldots, \overline{\chi_d}(a))$ is an isomorphism of real vector spaces.*

Démonstration. For $i = 1, \ldots, d$, let $\sigma_i : G \longrightarrow SL(V_i)$ be irreducible representations of G whose highest restricted weights are the dominant fundamental weights ϖ_i of G. Let $d_i \in \mathbb{N}$ be the multiplicity of weights ϖ_i and put $\chi_i = \sigma_i^{d_i}$. Then, take ρ_i to be the irreducible sub-representation of G in $\bigwedge^{d_i} V_i$ (obtained by composing σ_i with the exterior power representation of $SL(V_i)$), having the highest restricted weight χ_i. $\square$

The following very useful lemma is taken from [11] ; it essentially follows from the main theorem of Mostow's [88] and [89]. It enables us to treat the semi-simple case in the same manner as in the beginning of this section, in each of the d distinguished representations of the previous lemma. We will also **fix** the scalar products $< .,. >_i$ and the associated norms $||.||_i$ on V_i's, given by the next lemma, for the rest of the article.

Lemma 2.16. *[11] Let G be a connected semisimple linear real algebraic group. For every irreducible representation (V, ρ) of G of highest restricted weight χ, there exists a Euclidean norm on V, such that for every $g \in G$, we have*

1. $\lambda_1(\rho(g)) = \chi(\exp \lambda(g)) = \exp \overline{\chi}(\lambda(g))$
2. $||\rho(g)|| = \chi(\exp \kappa(g)) = \exp \overline{\chi}(\kappa(g))$.

Finally, we are in a position to define the notion of loxodromic element in G :

Definition 2.17.
1. *An element $g \in G$ is said to be loxodromic or $\mathbb{R}$-regular if $\rho_i(g)$ is proximal in $\mathbb{P}(V_i)$ for all $i = 1, \ldots, d$.*
2. *Let $0 < \epsilon \leq r$. A loxodromic element $g \in G$ is said to be (r, ϵ)-loxodromic, if for each $i = 1, \ldots, d$, $\rho_i(g)$ is (r, ϵ)-proximal in $\mathbb{P}(V_i)$.*

Note that a similar remark as Remark 2.9 applies to this definition as well.

We now start with the following sequence of known results about the Cartan and Jordan projections and their relations to (r, ϵ)-loxodromy ; we will repeatedly make good use of these in the sequel. The first one is the following transposition of the usual spectral radius formula to our multidimensional setting of Cartan and Jordan projections.

Lemma 2.18 (Spectral radius formula)**.** *For every $g \in G$, we have $\frac{1}{n}\kappa(g^n) \underset{n\to\infty}{\longrightarrow} \lambda(g)$.*

Démonstration. This follows from the usual spectral radius formula using the last two lemmata. Indeed, for each of the d distinguished representations ρ_i, by Lemma 2.16, we have $\overline{\chi}_{\rho_i}(\frac{1}{n}\kappa(g^n)) = \frac{1}{n}\log ||\rho_i(g)^n||_i \underset{n\to\infty}{\longrightarrow} \log \lambda_1(g)$, where the last convergence is the usual spectral radius formula. From this, one readily concludes using Lemma 2.15. □

Another immediate facilitating use of Lemma 2.16 can be observed in the proof of the next key lemma which expresses a good behaviour of Cartan projections under products ; it is closely related to the submultiplicativity of the operator norm on an algebra of linear transformations.

Lemma 2.19 (Uniform continuity of Cartan projection)**.** *Let G be a connected semisimple linear real algebraic group, $\kappa : G \to \mathfrak{a}$ be the Cartan projection and L a compact subset of G. Then, there exists a compact subset M of $\mathfrak{a}$ such that for each $g \in G$, we have $\kappa(LgL) \subset \kappa(g) + M$.*

Démonstration. Set $C = \max_{g\in L} \max_{i=1,\ldots,d}(||\rho_i(g)||_i \vee ||\rho_i(g^{-1})||_i)$. Since, by submultiplicativity of the associated operator norms, for all $x, y, u \in GL(V)$ for a normed vector space V, we have $||x^{-1}||^{-1}.||y^{-1}||^{-1}.||u|| \leq ||xuy|| \leq ||x||.||u||.||y||$, Lemma 2.16 implies

$$\log ||\rho_i(g)||_i - 2C \leq \overline{\chi}_i(\kappa(LgL)) \leq \log ||\rho_i(g)||_i + 2C$$

for all $i = 1, \ldots, d$. The result then follows from Lemma 2.15. □

We also have the following multidimensional counterparts of respectively Lemma 2.10 and Proposition 2.11 that we have stated in the beginning of this section. They both follow, respectively, from those two results by a straightforward use of Lemma 2.15 and Lemma 2.16 as in the proof of the previous lemma.

Proposition 2.20. *Let G be a connected semisimple linear real algebraic group and $r>0$. Then, there exist a compact set $M_r \subset \mathfrak{a}$ and for each $r \geq \epsilon > 0$, compact subsets $M_{(r,\epsilon)}$ of $\mathfrak{a}$ such that we have $\lim_{\epsilon\to 0} M_{(r,\epsilon)} \subseteq M_r$ (Hausdorff convergence), and for each (r,ϵ)-loxodromic element g of G, we have $\lambda(g)-\kappa(g) \in M_{(r,\epsilon)}$.*

For a set of loxodromic elements $g_1,\dots,g_l$ of G such that (noting $g_0=g_l$) $x^+_{\rho_i(g_j)} \notin X^<_{\rho_i(g_{j+1})}$ for all $j=0,\dots,l-1$ and for all $i=1,\dots,d$, denote by $\nu(g_l,\dots,g_1)$ the element of $\mathfrak{a}$ defined by

$$\chi_i(\exp \nu(g_l,\dots,g_1)) := \nu_1(\rho_i(g_l),\dots,\rho_i(g_1)) \quad \text{for all} \quad i=1,\dots,d \tag{2.2}$$

(cf. (2.1) for the definition of ν_1)

We then have the analogous result to Proposition 2.11 :

Theorem 2.21 ([11], [12])**.** *Let G be a connected semisimple linear real algebraic group. For every $r>0$, there exist a compact set N_r and for every $0<\epsilon\leq r$, compact sets $N_{(r,\epsilon)}$ of $\mathfrak{a}$ such that for each $r>0$, we have $\lim_{\epsilon\to 0} N_{(r,\epsilon)} \subseteq N_r$, and if $g_1,\dots,g_l$ are (r,ϵ)-loxodromic elements having the property that (noting as usual $g_0=g_l$) $d(x^+_{\rho_i(g_j)}, X^<_{\rho_i(g_{j+1})}) \geq 6r$ for all $j=0,\dots,l-1$ and for all $i=1,\dots,d$, then we have that for all $n_1,\dots,n_l\geq 1$, the linear transformation $g=g_l^{n_l}\dots g_1^{n_1}$ is $(2r,2\epsilon)$-loxodromic, and satisfies*

$$\lambda(g_l^{n_l}\dots g_1^{n_1}) - \sum_{i=1}^{l} n_i\lambda(g_i) \in l.N_{(r,\epsilon)} + \nu(g_l,\dots,g_0)$$
$$\kappa(g_l^{n_l}\dots g_1^{n_1}) - \sum_{i=1}^{l} n_i\lambda(g_i) \in l.N_{(r,\epsilon)} + \nu(g_l,\dots,g_0)$$

Analogously to Definition 2.12, we give

Definition 2.22.

1. *Let G be a connected semisimple linear real algebraic group and $0<\epsilon\leq r$. A subset E of G is said to be an (r,ϵ)-Schottky family, if for each $i=1,\dots,d$, $\rho_i(E)$ is an (r,ϵ)-Schottky family in $\mathbb{P}(V_i)$.*
2. *A subset E of G consisting of loxodromic elements is said to be (a,b)-narrow, if for each $i=1,\dots,d$, $\rho_i(E)$ is a (a,b)-narrow in $\mathbb{P}(V_i)$.*
3. *A semigroup $\Gamma \subset G$ is said to be ((a,b)-narrow) (r,ϵ)-Schottky, if there exists a ((a,b)-narrow) (r,ϵ)-Schottky family in Γ generating it.*

Remark 2.23.

1. *Similar remarks as 1. and 2. of Remark 2.13 apply to this definitions as well.*
2. *We also would like to note here that, later on in the text, for our purposes, we will construct free semigroups using Schottky families (Proposition 5.10), and, for example, show that a finite Schottky family generates a discrete semigroup (Proposition 5.38).*

The following important finiteness result of Abels-Margulis-Soifer [16] can be thought of as an quantitative refinement of the existence results for loxodromic elements studied by Goldsheid-Margulis [61], Prasad [97] and Benoist-Labourie [20]. It will be of crucial use in our considerations. We note that our Lemma 3.18 is also inspired by this theorem's proof, for which we refer the reader to the original [16] or for another treatment, to Benoist's [13] and [11].

Theorem 2.24 (Abels-Margulis-Soifer [16])**.** *Let G be a connected semisimple linear real algebraic group, Γ a Zariski dense sub-semigroup of G. Then, there exists $0 < r = r(\Gamma)$ such that for all $0 < \epsilon \leq r$, we can find a finite subset F of Γ with the property that for every $\gamma \in G$, there exists $f \in F$ such that γf is (r, ϵ)-loxodromic.*

Remark 2.25.

1. *While dealing with the probability measures of uncountable support, we will use the following extension of this result : there exists $0 < r = r(\Gamma)$ such that for all $0 < \epsilon \leq r$, we can find a finite subset F of Γ and neighbourhoods V_f in G of each $f \in F$, with the property that for each $\gamma \in G$, there exist a neighbourhood U_γ of γ in G, and $f \in F$ such that for all $f' \in V_f$ and $\gamma' \in U_\gamma$, $\gamma' f'$ is (r, ϵ)-loxodromic. Indeed, this extension readily follows by the following two facts : 1. The set of loxodromic elements in G is open in G. 2. The attracting direction $x_g^+ \in \mathbb{P}(V)$ and the repulsive hyperplane $X_g^< \subset \mathbb{P}(V)$ depend continuously on $g \in GL(V)$, where V is a finite dimensional vector space.*
2. *For the fixed distinguished representations $(\rho_i)_{i=1,\ldots,d}$ of Lemma 2.15, and Euclidean norms of Lemma 2.16 on V_i's, the number $r(\Gamma)$ of the previous theorem depends only on Γ, and we will denote by the same $r(\Gamma) > 0$, the constant defined by the extended result in 1. $r(\Gamma)$ will be used in the sequel as it is defined here.*

Remark 2.26. *To precise the paragraph preceding the statement of Abels-Margulis-Soifer result, in fact, the mere existence of loxodromic elements in Γ goes as an input to the proof of this Theorem 2.24, which then constructs many of them, and this, efficiently. The absence of this existence result is an obstruction, in this work, for most of our techniques to fail to yield equivalents of the following results for other local fields, such as the p-adics $\mathbb{Q}_p$, and its finite extensions. However, concerning LDP's and existence of joint spectrum, this obstruction may not be a profound one and in a future work, we shall deal with this question.*

Chapitre 3

PROOFS OF LARGE DEVIATION PRINCIPLES

Before starting, let us precise that in the rest of this text, S_n denotes the n^{th} step of a random walk associated to a probability measure μ on a group G, i.e. $S_n = X_n.\dots.X_1$, where X_i's are the random walk increments which are G-valued independent random variables with distribution (law) μ, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, henceforth fixed.

The aim of Section 3.1 and Section 3.2 is to prove the following theorem, which is a more general version of Theorem 1.1, stated in the introduction, with a slightly weaker conclusion. Namely, it does not suppose a finite exponential moment condition on μ and in its conclusion, it yields a weak LDP for the sequence $\frac{1}{n}\kappa(S_n)$.

Theorem 3.1. *Let G be a connected semisimple linear real algebraic group, μ a probability measure on G, whose support generates a Zariski dense sub-semigroup Γ of G. Then, the sequence of random variables $\frac{1}{n}\kappa(S_n)$ in $\mathfrak{a}^+$ satisfies a weak LDP with a convex rate function I.*

The aim of Section 3.3 is to strengthen the previous theorem, with an exponential moment hypothesis, to prove the following more precise version of Theorem 1.1. For the notion of strong exponential moment and the limiting exponential moment generating function $\overline{\Lambda}$, see Section 3.3.

Theorem 3.2. *Let G be a connected semisimple linear real algebraic group, μ a probability measure on G, whose support generates a Zariski dense sub-semigroup Γ of G. Suppose that μ has a finite exponential moment. Then, full LDP exists with a proper convex rate function I for the sequence $\frac{1}{n}\kappa(S_n)$ of random variables in $\mathfrak{a}^+$. Moreover, if μ has a strong exponential moment, then we can identify $I = \overline{\Lambda}^*$, where $\overline{\Lambda}^*$ is the convex conjugate of the limiting exponential moment generating function of the sequence of random variables $\frac{1}{n}\kappa(S_n)$.*

Finally, in Section 3.4, we shall prove the following theorem which is a more precise version of Theorem 1.4. For the setting of Theorem 3.3, Γ denotes the semigroup generated by the support of the probability measure μ on $GL(V)$, for some finite dimensional real vector space V, and the linear algebraic subgroup G of $GL(V)$ is the

Zariski closure of Γ. We suppose that G is a linear real algebraic group isomorphic with an isomorphism of algebraic groups to some $H\times T$, which we identify with G, and then, where H is a semisimple linear real algebraic group and T is a central subgroup of GL. Examples of such G indeed include all semisimple real linear algebraic groups (connected and non-connected) as well as some reductive groups such as $GL(V)$.

Theorem 3.3. *Let μ be a probability measure on $GL(V)$ and the semigroup Γ generated by the support of μ and its Zariski closure G be as above. Then, the sequence μ_n of laws of $\frac{1}{n}\log||S_n||$ satisfies a weak LDP with a convex rate function I. If, additionally, μ possesses a finite exponential moment, then the sequence μ_n satisfies a full LDP with a proper convex rate function I. Finally, if μ possesses a strong exponential moment, then we have $I=\overline{\Lambda}^*$, where $\overline{\Lambda}^*$ is the convex conjugate of the limiting exponential moment generating function of μ_n's.*

Remark 3.4.
1. *We observe in each of the three previous theorems that in fact, if the support of the measure μ generates a bounded sub-semigroup in G, then it is still true that the sequences of random variables in these theorems satisfy an LDP with a rate function I, which is obviously seen to take the value 1 on $0\in\mathfrak{a}$ and ∞ elsewhere.*
2. *For a discussion of several properties of the rate functions appearing in these three theorems, see Section 4.2.*
3. *See Corollary 4.28 and Proposition 4.32, respectively, for a slightly more precise large deviation results involving the existence of certain limits in Theorem 3.2 and Theorem 3.3.*

3.1 Existence of weak LDP

The following first lemma essentially relies on Theorem 2.24 and the uniform continuity of Cartan projections (Lemma 2.19), and it says that if at some step, the Cartan projection of the walk hits a certain region of the Weyl chamber with a certain probability, then after some bounded number of steps, it will hit some loxodromic elements whose Cartan projection is close to that region, and this with not arbitrarily small probability :

Lemma 3.5. *Let $0<\epsilon<r=r(\Gamma)$. There exist a compact set $C=C(\Gamma,\epsilon)\subset\mathfrak{a}$, a natural number $i_0=i_0(\epsilon,\Gamma,\mu)$, and a constant $d_1=d_1(\epsilon,\Gamma,\mu)>0$ such that for all $n_0\in\mathbb{N}$ and $R\subset\mathfrak{a}^+$, there exists a natural number $n_1\geq n_0$ with $n_1-n_0\leq i_0$ such that we have*

$$\mathbb{P}(\kappa(S_{n_1})\in R+C \textit{ and } S_{n_1} \textit{ is } (r,\epsilon)\textit{-loxodromic})\geq d_1.\mathbb{P}(\kappa(S_{n_0})\in R)$$

Démonstration. Let $F=F(r,\epsilon)$ denote the finite subset of Γ given by Theorem 2.24 and V_f denote the neighbourhoods in G of elements f of F given by Remark 2.25. Fix $i_0\in\mathbb{N}$ such that $F\subset\bigcup_{i=1}^{i_0}\operatorname{supp}(\mu^{*i})$, this is indeed possible since $\operatorname{supp}(\mu)$ generates $\Gamma\supset F$. Denote $F=\{f_1,\ldots,f_{|F|}\}$ and using Remark 2.25, define a covering of Γ by the subsets $\Gamma_i:=\{g\in\Gamma\,|\,gf'_i \textit{ is } (r,\epsilon)\text{-loxodromic for every } f'_i\in V_{f_i}\}$ for $i=1,\ldots,|F|$. Fix numbers $k_1,\ldots,k_{|F|}\leq i_0$ such that $\mu^{*k_i}(V_{f_i})=:\alpha_i>0$, where this latter inequality is strict by definition of support of a probability measure, here μ^{k_i}'s. Then, since, Γ_i's cover Γ, we have

$$\mathbb{P}(\kappa(S_{n_0}) \in R) \leq \sum_{j=1}^{|F|} \mathbb{P}(S_{n_0} \in \Gamma_j \cap \kappa^{-1}(R))$$

so that there exists $j_0 \in \{1, \ldots, |F|\}$ such that

$$\mathbb{P}(S_{n_0} \in \Gamma_{j_0} \cap \kappa^{-1}(R)) \geq \frac{\mathbb{P}(\kappa(S_{n_0}) \in R)}{|F|}$$

Now, as $|F|$ is finite and G is a σ-finite topological space, the set $\cup_{i=1}^{|F|} \overline{V}_{f_i}$ is a compact set in G, and denote by C the compact subset M of $\mathfrak{a}$ given by Lemma 2.19, in which we take $L = \cup_{i=1}^{|F|} \overline{V}_{f_i}$. Therefore, by this lemma, for every $g \in \Gamma$ such that $\kappa(g) \in R$ and for all $f' \in \cup_{i=1}^{|F|} \overline{V}_{f_i}$, we have $\kappa(gf') \in R + C$. Then, it follows by the independence of the random walk increments that

$$\begin{aligned}
&\mathbb{P}(\kappa(S_{n_0+k_{j_0}}) \in R + C \text{ and } S_{n_0+k_{j_0}} \text{ is } (r,\epsilon)\text{-loxodromic}) \\
&\geq \mathbb{P}(X_{n_0+k_{j_0}} \cdot \ldots X_{k_{j_0}+1} \in \Gamma_{j_0} \cap \kappa^{-1}(R) \text{ and } X_{k_{j_0}} \cdot \ldots X_1 \in V_{f_{j_0}}) \\
&= \mathbb{P}(S_{n_0} \in \Gamma_{j_0} \cap \kappa^{-1}(R)).\mathbb{P}(S_{k_{j_0}} \in V_{f_{j_0}}) \geq \frac{\mathbb{P}(\kappa(S_{n_0}) \in R)}{|F|}.\alpha_{j_0}
\end{aligned}$$

Now, putting $n_1 := n_0 + k_{j_0} \leq n_0 + i_0$ and $\alpha_0 := \min_{k=1,\ldots,|F|} \alpha_k > 0$, we have

$$\mathbb{P}(\kappa(S_{n_1}) \in R + C \text{ and } S_{n_1} \text{ is } (r,\epsilon) - \text{loxodromic}) \geq d_1 \mathbb{P}(\kappa(S_{n_0}) \in R)$$

where we have put $d_1 = \frac{\alpha_0}{|F|} = d_1(\epsilon, \mu, \Gamma)$. □

Next lemma is an obvious observation on the relation between narrowness and (r,ϵ)-Schottky properties of a set of loxodromic elements. It will prove to be useful in our considerations together with the lemma following it. In its proof and in what follows, recall that $(\rho_i, V_i)_{i=1,\ldots,d}$ are the distinguished representations of G.

Lemma 3.6. *Let ϵ and r be two real numbers such that $0 < 6\epsilon \leq r$. Then, a $(r,1)$-narrow set E of (r,ϵ)-loxodromic elements in G is a (r_1,ϵ)-Schottky family, where we can take $r_1 = \frac{r}{6}$.*

Démonstration. Observe first that, by definition, if γ is (r,ϵ)-loxodromic, then γ is also (r_1,ϵ_1)-loxodromic for all $r_1 \leq r$ and $\epsilon_1 \geq \epsilon$ such that $r_1 \geq \epsilon_1$. Therefore, to prove the lemma, one just notes that for all γ, $\gamma' \in E$, since $d(x^+_{\rho_i(\gamma)}, X^<_{\rho_i(\gamma)}) \geq 2r$ and $d(x^+_{\rho_i(\gamma)}, x^+_{\rho_i(\gamma')}) < r$, we have $d(x^+_{\rho_i(\gamma)}, X^<_{\rho_i(\gamma')}) \geq 2r - r = r$. Hence putting $r_1 = \frac{r}{6}$ we have by hypothesis, $r_1 \geq \epsilon$ and $d(x^+_{\rho_i(\gamma)}, X^<_{\rho_i(\gamma')}) \geq 6r_1$ as in the definition of a (r_1,ϵ)-Schotky family. □

We shall now proceed with the following lemma, which is basically a consequence of the compactness of projective spaces of V_i's. We will put it to good use on two occasions ; once, together with Lemma 3.6 to obtain a useful corollary, and once in the proof of convexity.

Lemma 3.7. *Let $r \geq \epsilon > 0$ and a, b two positive constants, be given. Then, there exists a strictly positive constant $d_2 = d_2(a,b)$ such that for every subset E of G consisting of (r,ϵ)-loxodromic elements, and for all $n \in \mathbb{N}$, there exists an (a,b)-narrow subset E_n of E such that, we have $\mathbb{P}(S_n \in E_n) \geq d_2 \mathbb{P}(S_n \in E)$.*

Démonstration. Indeed, for all $i = 1, \ldots, d$, by compactness of $\mathbb{P}(V_i)$, we can choose two partitions $Y_1^i, \ldots, Y_{s_i}^i$, $Z_1^i, \ldots, Z_{t_i}^i$ of $\mathbb{P}(V_i)$, such that $\text{diam}(Y_j^i) < a$ and $\text{diam}(Z_j^i) < b$, with $s_i = s_i(a)$ and $t_i = t_i(b)$ (Recall that we are working with the fixed set of representations $(\rho_i)_{i=1,\ldots,d}$ on V_i's with fixed Euclidean structures). Let $\underline{i}$, $\underline{j}$ denote multi-indices of the form $\underline{i} = (i_1, \ldots, i_d)$ and $\underline{j} = (j_1, \ldots, j_d)$ where, for each $k = 1, \ldots, d$, $i_k \in \{1, \ldots, s_k\}$ and $j_k \in \{1, \ldots, t_k\}$. Now, let $E \subset \Gamma$ be given as in the statement and for multi-indices $\underline{i}$, $\underline{j}$, denote by $E_{\underline{i}}^{\underline{j}}$ the following subset of E :

$$E_{\underline{i}}^{\underline{j}} := \{\gamma \in E \,|\, x_{\rho_k(\gamma)}^{+} \in Y_{i_k}^{k} \text{ and } (X_{\rho_k(\gamma)}^{<})^{\perp} \in Z_{j_k}^{k}\}$$

By the choice of Y_j^i's and Z_j^i's, the family $E_{\underline{i}}^{\underline{j}}$ partitions E and we thus have for every $n \in \mathbb{N}$

$$\mathbb{P}(S_n \in E) = \sum_{\underline{i},\underline{j}} \mathbb{P}(S_n \in E_{\underline{i}}^{\underline{j}})$$

It follows that for every $n \in \mathbb{N}$, there exist at least two multi-indices $\underline{i}_0$ and $\underline{j}_0$ such that $\mathbb{P}(S_n \in E_{\underline{i}_0}^{\underline{j}_0}) \geq \frac{\mathbb{P}(S_n \in E)}{s_1 \ldots s_d t_1 \ldots t_d}$. Hence, putting $d_2 = d_2(a,b) = \frac{1}{s_1 \ldots s_d t_1 \ldots t_d}$ and $E_n = E_{\underline{i}_0}^{\underline{j}_0}$, we have the result of the lemma. □

Corollary 3.8. *Let r and ϵ be two real numbers with $r \geq 6\epsilon > 0$. Then, there exists a constant $d_3 = d_3(r) > 0$ such that for every subset E of G consisting of (r,ϵ)-loxodromic elements and for all $n \in \mathbb{N}$, there exists an (r_1,ϵ)-Schottky family $E_n \subset E$ with $r_1 \geq \frac{r}{6} \geq \epsilon$ and such that $\mathbb{P}(S_n \in E_n) \geq d_3.\mathbb{P}(S_n \in E)$.*

Démonstration. In Lemma 3.7, choose $a = r$ and $b = 1$, and apply Lemma 3.6. □

We continue with the next proposition, which is essentially a consequence of Benoist's Theorem 2.21. It says that the images in $\mathfrak{a}^+$ of the Cartan projections of the n^{th}-power of an (r,ϵ)-Schottky family is contained, up to compact perturbation, in the n-dilation of the images in $\mathfrak{a}^+$ of the Cartan projections of that family. For simplicity, all the compact subsets C of $\mathfrak{a}$ appearing in its proof are supposed (up to enlarging) to be convex and containing $0 \in \mathfrak{a}$.

Proposition 3.9. *There exists a compact subset $K = K(r,\epsilon)$ of $\mathfrak{a}$, depending only on r and ϵ, with the property that for all (r,ϵ)-Schottky family E in G and $n \in \mathbb{N}$, we have $\kappa(E^n) \subset n.(co(\kappa(E)) + K)$, where $E^n := \{\gamma_1.\ldots.\gamma_n \,|\, \gamma_i \in E\}$, $\kappa(E) := \{\kappa(\gamma) \,|\, \gamma \in E\}$, $\kappa(E) + K := \{x + k \,|\, x \in \kappa(E), k \in K\}$ and $co(.)$ stands for the convex hull.*

Remark 3.10. *It will follow from the proof that for each $r > 0$, there exists a compact set $K(r)$ depending only on r, such that if an (r,ϵ)-Schottky family E is moreover $(a,1)$-narrow, then the compact set $K(r,\epsilon)$ of this proposition can be chosen in a more optimal way as $K(a,r,\epsilon)$ depending also on a, so as to satisfy $K(a,r,\epsilon) \subseteq K(r)$ for each a and ϵ small enough.*

We start first by establishing a lemma that gives a control over $\nu(g_l,\dots,g_1)$ for g_i's coming from an (r,ϵ)-Schottky family.

Lemma 3.11. *There exists a compact subset $\hat{K}$ of $\mathfrak{a}$ depending only on r with the property that for every (r,ϵ)-Schottky family E in G, setting for each $l\in\mathbb{N}$ the set $P_l(E):=\{\nu(g_l,\dots,g_1)\,|\,g_i\neq g_{i+1}\in E \text{ for } i=1,\dots,l-1\}$, $P_l(E)$ is contained in the l-dilation $l.\hat{K}$ of $\hat{K}$.*

Démonstration. Recall that $\nu(g_l,\dots,g_1)$ is the element of $\mathfrak{a}$ defined by $\chi_i(\exp\nu(g_l,\dots,g_1)) = \nu_1(\rho_i(g_l),\dots,\rho_i(g_1))$ for $i=1,\dots,d$. Therefore by Lemma 2.15 ; it suffices to show that for an (r,ϵ)-Schottky family $D\subset GL(V)$ in $\mathbb{P}(V)$, the set $P_l(D):=\{\nu_1(g_l,\dots,g_1)\,|\,g_i\neq g_{i+1}\in D \text{ for } i=1,\dots,l-1\}$ is contained in the l^{th}-power of a compact subset of the multiplicative group $]0,\infty[$, depending on r. But, recall that $\nu_1(g_l,\dots,g_1) = \prod_{1\leq i\leq l}\nu_1(g_i,g_{i-1})$ (putting $g_0=g_l$) where $\nu_1(g,h)=|\beta|$ with $\beta\in\mathbb{R}$ defined as $v_g^+-\beta v_h^+\in V_g^<$. Hence, it suffices to bound $\nu_1(g,h)$ uniformly in $g,h\in D$ away from 0 and ∞.

This follows easily from the definition of an (r,ϵ)-Schottky family in $\mathbb{P}(V)$: indeed, write $v_g^+ = av^\perp + bv_g^<$ and $v_h^+ = cv^\perp + dv_h^<$ where $v^\perp$ is a unit orthogonal to $V_g^<$ and $v_g^<$ and $v_h^<$ are unit vectors in $V_g^<$. We have $a^2+b^2=c^2+d^2=1$. Moreover, by definition of an (r,ϵ)-Schottky family, we have $1\geq|a|=d(v_g^+,V_g^<)\geq 2r$ and $1\geq |c| = d(v_h^+,V_g^<)\geq 6r$. From this, it immediately follows that we have $\frac{1}{2r}\geq\nu_1(g,h) = |\beta|\geq 6r$ (note that this inequality is indeed consistent, by 1. of Remark 2.13). □

Remark 3.12.

1. *If, moreover, as will be our case, the set D in the above proof is $(\eta,1)$-narrow in $\mathbb{P}(V)$ with $\eta\leq r$, then one can easily show that for all $g,h\in D$, we can bound $\nu_1(g,h)$ as : $1-\frac{\eta}{r}\leq\nu_1(g,h)\leq 1+\frac{\eta}{6r-\eta}$.*
2. *As a special case of 1., observe that if for all $g,h\in E$ and $i=1,\dots,d$ $x^+_{\rho_i(g)} = x^+_{\rho_i(h)}$, then E is $(0,1)$-narrow in $\mathbb{P}(V_i)$, so that $\nu_1(\rho_i(g),\rho_i(h))=1$ and one can take $\hat{K}=\{0\}\subset\mathfrak{a}$ in the lemma.*

Denoting by $\hat{K}_r$ the compact subset of $\mathfrak{a}$ given by the last lemma, put $\tilde{K}_{(r,\epsilon)} := \hat{K}_r + N_{(r,\epsilon)}$ a compact subset of $\mathfrak{a}$, where $N_{(r,\epsilon)}$ is the compact set as given by Theorem 2.21. This latter theorem now implies in particular that

Lemma 3.13. *For an (r,ϵ)-Schottky family $E\subset G$, for all $l\in\mathbb{N}$, $g_i\in E$ for $i=1,\dots,l$ and $n_1,\dots,n_l\geq 1$, we have*

$$\kappa(g_l^{n_l}.\dots.g_1^{n_1}) - \sum_{i=1}^{l} n_i\lambda(g_i)\in l.\tilde{K}_{(r,\epsilon)}$$

□

Putting $K_{(r,\epsilon)} := \tilde{K}_{(r,\epsilon)} + M_{(r,\epsilon)}$, a compact subset of $\mathfrak{a}$, where $M_{(r,\epsilon)}$ is as given by Proposition 2.20, this proposition together with the above corollary implies that

Corollary 3.14. *For an (r,ϵ)-Schottky family $E \subset G$, for all $l \in \mathbb{N}$, $g_i \in E$ for $i = 1,\dots,l$, $n_1,\dots,n_l \geq 1$ and $n = n_1 + \ldots + n_l$, we have*

$$\kappa(g_l^{n_l}.\dots.g_1^{n_1}) - \sum_{i=1}^{l} n_i\kappa(g_i) \in l.\tilde{K}_{(r,\epsilon)} + n.M_{(r,\epsilon)} \subset n.K_{(r,\epsilon)}$$

Démonstration. Writing

$$\kappa(g_l^{n_l}.\dots.g_1^{n_1}) - \sum_{i=1}^{l} n_i\kappa(g_i) = (\kappa(g_l^{n_l}.\dots.g_1^{n_1}) - \sum_{i=1}^{l} n_i\lambda(g_i)) + \sum_{i=1}^{l} n_i(\lambda(g_i) - \kappa(g_i))$$

we see that the first term on the right hand side is included in $l.\tilde{K}_{(r,\epsilon)}$ by the previous lemma and the second sum is in $n.M_{(r,\epsilon)}$ by Proposition 2.20, since $M_{(r,\epsilon)}$ is convex. The last inclusion follows by definition of $K_{(r,\epsilon)}$ recalling again that $\tilde{K}_{(r,\epsilon)}$, $M_{(r,\epsilon)}$ and $K_{(r,\epsilon)}$ are supposed to be convex and containing 0. □

Proposition 3.9 follows obviously from the last inclusion in previous corollary :

Proof of Proposition 3.9. Let $g_1,\dots,g_l \in E$, $n_1,\dots,n_l \geq 1$ and $n = n_1 + \ldots + n_l$. By the previous corollary, we have

$$\kappa(g_l^{n_l}\dots g_1^{n_1}) \in \sum_{i=1}^{l} n_i\kappa(g_i) + n.K_{(r,\epsilon)} \subset n(co(\kappa(E)) + K_{r.\epsilon)})$$

□

For later convenient use, we single out the following topological notion and note two obvious facts about it in the following lemma.

Definition 3.15. *Let X be a topological space and $O_1 \subset O_2$ two open subsets of X. We say that O_1 is super-strictly contained in O_2 if $\overline{O}_1 \subsetneq O_2$.*

Lemma 3.16.
1. *Let V be a finite dimensional normed vector space and O_1 and O_2 two open bounded subsets of V, O_1 super-strictly contained in O_2. Then, for all bounded set $K \subset V$, there exists a constant $R(O_1,O_2,K) \in \mathbb{R}^+$ such that for all $Q \geq Q(O_1,O_2,K)$, we have $Q.O_1 + K \subset Q.O_2$*
2. *Let O_1 and O_2 be as above. Then, there exists a real number $q(O_1,O_2) < 1$ such that for all $n_1,n_2 \in \mathbb{N}$ with $1 \geq \frac{n_1}{n_2} > q(O_1,O_2)$, we have $n_1O_1 \subset n_2O_2$.*

Démonstration. Both statements are obvious. Remark that the hypothesis implies that $d(O_1,O_2^c) > 0$ and one can take $Q(O_1,O_2,K)$ and $1 > q(O_1,O_2)$ any real numbers larger than respectively $\frac{diam(K)}{d(O_1,O_2^c)}$ and $1 - \frac{d(O_1,O_2^c)}{\sup_{x\in O_1} ||x||}$. □

We shall need one last lemma before proceeding to prove the theorem. It relies on the uniform continuity of the Cartan projections (Lemma 2.19) and says that if the averages of the Cartan projections of the random product hits a certain region of the Cartan subalgebra at periodic times, then it will hit any open neighbourhood of this region at any time with at least the same asymptotic exponential rate of probability :

Lemma 3.17. *Let O_1 and O_2 be two open bounded convex subsets of $\mathfrak{a}^+$, O_1 super-strictly contained in O_2. Suppose that there exist $n_0 \in \mathbb{N}$ and $\alpha \geq 0$ such that for all $k \geq 1$, we have $\mathbb{P}(\kappa(S_{n_0 k}) \in k n_0 O_1) \geq e^{-n_0 k \alpha}$. Then we have $\liminf_n \frac{1}{n} \log \mathbb{P}(\frac{1}{n}\kappa(S_n) \in O_2) \geq -\alpha$.*

Démonstration. For all $n \in \mathbb{N}$, let $k_n \in \mathbb{N}$ be defined by $n_0(k_n+1) > n \geq n_0 k_n$. By σ-compactness, we can choose a compact subset L_{n_0} of G containing $e \in G$ and such that $\mu^{*i}(L_{n_0}) \geq \frac{1}{2}$ for each $i = 1, \ldots, n_0$. Let M_{n_0} be the compact subset M of $\mathfrak{a}$ given by Lemma 2.19, by taking in it $L = L_{n_0}$.

By definition of super-strict inclusion and the fact that the ambient space is a normed real vector space, we can pick O_{12} such that each of the inclusions $O_1 \subset O_{12} \subset O_2$ is super-strict. Now, let $Q_{n_0} := Q(O_{12}, O_2, M_{n_0}) \in \mathbb{R}$ and $q := q(O_1, O_{12}) < 1$ where these last quantities are as defined in Lemma 3.16. Then, for all $n \in \mathbb{N}$ such that $n \geq Q_{n_0}$ and $1 - \frac{n_0}{n} > q$, we have the following sequence of inclusions of events :

$$\{\kappa(S_n) \in k_n n_0 O_1 + M_{n_0}\} \subset \{\kappa(S_n) \in n O_{12} + M_{n_0}\} \subset \{\kappa(S_n) \in n O_2\}$$

where the first inclusion is by 2. and the second by 1. of Lemma 3.16.

As a result, by independence of random walk increments, for all $n \in \mathbb{N}$, we have

$$\begin{aligned} &\mathbb{P}(\frac{1}{n}\kappa(S_n) \in O_2) \geq \mathbb{P}(\kappa(S_{k_n n_0 + (n - k_n n_0)}) \in k_n n_0 O_1 + M_{n_0}) \geq \\ &\mathbb{P}(\kappa(S_{k_n n_0}) \in k_n n_0 O_1) . \mathbb{P}(S_{n - k_n n_0} \in L_{n_0}) \geq e^{-n_0 k_n \alpha} \frac{1}{2} \end{aligned} \tag{3.1}$$

where the last inequality follows by hypothesis and the construction of L_{n_0}. Now, in (3.1), taking logarithm, dividing by n, and taking n to infinity, we obtain the result of the lemma. □

We are now ready to prove the existence statement in Theorem 3.1 :

Proof of Theorem 3.1 (Existence of LDP). For all $n \geq 1$, denote by μ_n, the law of the random variable $\frac{1}{n}\kappa(S_n)$. It is a probability measure supported on the closed subset $\mathfrak{a}^+$ of the vector space $\mathfrak{a}$. To establish the weak LDP for this sequence of probability measures, we use Theorem 2.4 and argue by contradiction.

Let I_{li} and I_{ls} denote the functions on $\mathfrak{a}$, associated to the sequence μ_n as in Theorem 2.4, where we take the norm-open balls in $\mathfrak{a}$ as a base of topology. Suppose now for a contradiction that there exists $x \in \mathfrak{a}$ such that $I_{li}(x) > I_{ls}(x) \geq 0$. We can in fact suppose that x is in the closed Weyl chamber $\mathfrak{a}^+$ by the previous remark that for all $n \in \mathbb{N}$, $\mathrm{supp}(\mu_n) \subset \mathfrak{a}^+$.

Recalling the definitions of the functions I_{li} and I_{ls}, this implies that there exists an open ball $O_5 \subset \mathfrak{a}$ with $x \in O_5$ and such that

$$-\liminf_n \frac{1}{n} \log \mu_n(O_5) > \sup_{\substack{O \subset \mathfrak{a} \\ x \in O}} -\limsup_n \frac{1}{n} \log \mu_n(O) + 4\eta \tag{3.2}$$

for some $\eta > 0$ small enough.

By definition of super-strict inclusion and the fact that the ambient space $\mathfrak{a}$ is a vector space over an Archimedean field, namely $\mathbb{R}$, we can choose $x \in O_1 \subset O_2 \subset O_3 \subset O_4 \subset O_5$ open balls around x, where each inclusion is super-strict, such that (3.2) yields

$$-\liminf_n \frac{1}{n} \log \mu_n(O_5) > -\limsup_n \frac{1}{n} \log \mu_n(O_1) + 3\eta$$

Now, let $r = r(\Gamma)$ be given by Theorem 2.24 and choose $\epsilon \leq \frac{r}{6}$. Let $d_1 = d_1(r, \epsilon, \Gamma)$ and $i_0 = i_0(\epsilon, \Gamma, \mu)$ be the constants given by Lemma 3.5, $C = C(\Gamma, \epsilon)$ be the compact subset of $\mathfrak{a}$ also given by Lemma 3.5, $d_3 = d_3(r)$ be the constant given by Corollary 3.8, $K = K(r, \epsilon)$ be the compact subset of $\mathfrak{a}$ given by Proposition 3.9. Let us also fix a real number $Q \geq \max_{i<j}(Q(O_i, O_j, C) \vee Q(O_i, O_j, K))$ where these latter quantities are as defined in Lemma 3.16 and let $q := q(O_1, O_5)$ where again this is defined as in Lemma 3.16. Choose $n_0 \in \mathbb{N}$ such that

1. $-\frac{1}{n_0} \log \mu_{n_0}(O_1) + 2\eta < -\liminf_n \frac{1}{n} \log \mu_n(O_5)$
2. $e^{-n_0\eta} \leq d_1 d_3$
3. $n_0 \geq Q$
4. $\frac{n_0}{n_0+i_0} > q$

Put $\alpha := -\frac{1}{n_0} \log \mu_{n_0}(O_1)$ and $\beta := -\liminf\limits_{n\to\infty} \frac{1}{n} \log \mu_n(O_5)$ so that by Item 1 in the choice of n_0,

$$\alpha + 2\eta < \beta \tag{3.3}$$

Setting $R = n_0 O_1$ in Lemma 3.5, we obtain that for some n_1 such that $n_1 - n_0 \leq i_0$

$$\mathbb{P}(\kappa(S_{n_1}) \in n_0 O_1 + C \text{ and } S_{n_1} \text{ is } (r, \epsilon)\text{-loxodromic}) \geq e^{-n_1\alpha}.d_1 \tag{3.4}$$

The choice of n_0 (respectively Item 3 and Item 4 above) implies by Lemma 3.16 that $n_0 O_1 + C \subset n_0 O_2$ and $n_0 O_2 \subset n_1 O_3$ so that (3.4) becomes

$$\mathbb{P}(\kappa(S_{n_1}) \in n_1 O_3 \text{ and } S_{n_1} \text{ is } (r, \epsilon)\text{-loxodromic}) \geq e^{-n_1\alpha}.d_1 \tag{3.5}$$

Applying Corollary 3.8 by taking $L = \kappa^{-1}(n_1 O_3) \cap \Gamma_{(r,\epsilon)}$, which is non-empty by (3.5), and where $\Gamma_{(r,\epsilon)}$ is the set of (r, ϵ)-loxodromic elements in Γ, using also (3.5), we obtain that there exists an (r_1, ϵ)-Schottky family $E \subset L \subset \Gamma$ such that we have

$$\mathbb{P}(\kappa(S_{n_1}) \in n_1 O_3 \text{ and } S_{n_1} \in E) \geq e^{-n_1\alpha} d_1 d_3 \geq e^{-n_1(\alpha+\eta)}$$

where the last inequality follows by the Item 2 of the choice of n_0 and since $n_1 \geq n_0$.

Next, observe that by the construction of L and since $E \subset L$, we have $\kappa(E) \subset n_1 O_3$ and therefore, as O_3 is convex, $co(\kappa(E)) \subset n_1 O_3$. Then, by Proposition 3.9, we obtain that for each $k \geq 1$, $\kappa(E^k) \subset k.(co(\kappa(E)) + K) \subset k.(n_1 O_3 + K) \subset k n_1 O_4$ where the last inclusion follows also from the Item 3 of the choice of n_0 and since $n_1 \geq n_0$.

Finally, for all $k \geq 1$, by the independence of the random walk increments, we have that $\mathbb{P}(S_{n_1 k} \in E^k) \geq \mathbb{P}(S_{n_1} \in E)^k$ and thus we obtain

$$\mathbb{P}(\kappa(S_{n_1 k}) \in k n_1 O_4) \geq \mathbb{P}(S_{n_1 k} \in E^k) \geq \mathbb{P}(S_{n_1} \in E)^k \geq e^{-n_1 k(\alpha+\eta)}$$

Therefore, Lemma 3.17 establishes that $\beta = -\liminf \frac{1}{n} \log \mathbb{P}(\kappa(S_n) \in O_5) \leq \alpha + \eta$ which together with (3.3) yields $\alpha + 2\eta < \beta \leq \alpha + \eta$, a contradiction. $\square$

3.2 Convexity of the rate function

Our first lemma in this section is a key dispersion result which is in fact a corollary of the proof of Theorem 2.24 in Abels-Margulis-Soifer's [16] (see also the exposition of Quint in [100], Proposition 2.3.4). Namely, it says that, by the Zariski density of Γ in G and connectedness of G, one can find finite sets in Γ such that for each point of the projective spaces of the distinguished representation spaces V_i's, some elements of these finite sets of Γ will, by their action, disperse that point in the projective spaces. It will be useful on several occasions, particularly by its relation to the 1. (b) of Definition 2.12.

Lemma 3.18 (Dispersion lemma). *For all $t \in \mathbb{N}$, there exist a strictly positive constant $\eta_t = \eta(t,\Gamma)$, depending only on t and Γ, and a finite set $M_t \subset \Gamma$ with the following properties : for all $\bar{x} = (x_1,\ldots,x_d) \in \prod_{i=1}^d \mathbb{P}(V_i)$, where V_i's are the distinguished representation spaces of G, there exist $\gamma_1,\ldots,\gamma_t \in M_t$ such that*

1. *For all $i = 1,\ldots,d$ and for all $j \neq k \in \{1,\ldots,t\}$,*
$$d_i(\rho_i(\gamma_j).B_i(x_i,\eta_t), \rho_i(\gamma_k).B_i(x_i,\eta_t)) > \eta_t$$
2. *For all $i = 1,\ldots,d$ and for every subset $\{\gamma_{i_1},\ldots,\gamma_{i_k}\}$ of $\{\gamma_1,\ldots,\gamma_t\}$ of cardinality less than $k \leq \dim V_i$, for all $y_i^1,\ldots,y_i^k, z_i \in B_i(x_i,\eta_t)$, denoting by $< \rho_i(\gamma_{i_1})y_i^1,\ldots,\rho_i(\gamma_{i_k})y_i^k >$ the projective image of the subspace generated by these lines, and for all $j \notin \{i_1,\ldots,i_k\}$, we have,*
$$d_i(< \rho_i(\gamma_{i_1})y_i^1,\ldots,\rho_i(\gamma_{i_k})y_i^k > , \rho_i(\gamma_j)z_i) > \eta_t$$

Démonstration. We start by inductively finding elements $\gamma_1^{\bar{x}},\ldots,\gamma_t^{\bar{x}} \in \Gamma$ for each element $\bar{x} = (x_1,\ldots,x_d)$ of $\prod_{i=1}^d \mathbb{P}(V_i)$: choose $\gamma_1^{\bar{x}} \in \Gamma$ arbitrarily. Having constructed $\gamma_1^{\bar{x}},\ldots\gamma_k^{\bar{x}}$ for some $k < t$, put

$$\begin{aligned} G_{i,k+1} := \{\gamma \in G \,|\, &\rho_i(\gamma).x_i \text{ does not belong to the proper subspaces of } V_i \\ &\text{generated by the lines } \rho_i(\gamma_j).x_i \text{ for } j \in \{1,\ldots,k\}\} \end{aligned}$$

Since there are finitely many such proper spaces of V_i, and the condition of not belonging to a proper subspace is a Zariski open condition in G, $G_{i,k+1}$ is a finite intersection of Zariski open sets which are also non-empty since the distinguished representations, ρ_i's are irreducible. Consequently, $G_{i,k+1}$ is a non-empty Zariski open set in G. Similarly, the set G_{k+1} defined by $G_{k+1} := \cap_{i=1}^d G_{i,k+1}$ is Zariski open. Γ being, by assumption, Zariski dense in G, the intersection $G_{k+1} \cap \Gamma$ is non-empty ; choose one element $\gamma_{k+1}^{\bar{x}} \in G_{k+1} \cap \Gamma$.

By induction, we then have constructed $\gamma_1^{\bar{x}},\ldots,\gamma_t^{\bar{x}} \in \Gamma$ for each $\bar{x} \in \prod \mathbb{P}(V_i)$ such that for each $i = 1,\ldots,d$, the elements of $\{\rho_i(\gamma_1^{\bar{x}}).x_i,\ldots,\rho_i(\gamma_t^{\bar{x}}).x_i\}$ are in general position. Now choose $\eta_t^{\bar{x}} > 0$, such that

$$d_i(< \rho_i(\gamma_{i_1}^{\bar{x}}).x_i,\ldots,\rho_i(\gamma_{i_k}^{\bar{x}}).x_i >, \rho_i(\gamma_j^{\bar{x}}).x_i) > 2\eta_t^{\bar{x}}$$

for all $i = 1,\ldots,d$, $k \leq \dim V_i - 1$, $i_1,\ldots,i_k \in \{1,\ldots,t\}$ and $j \notin \{i_1,\ldots,i_k\}$. Such an $\eta_t^{\bar{x}} > 0$ indeed exists by our construction of the $\gamma_i^{\bar{x}}$'s.

Now, by continuity of the action of G on $\mathbb{P}(V_i)$'s, for all $\bar{x} = (x_1, \ldots, x_d) \in \prod \mathbb{P}(V_i)$, there exists a neighbourhood $W^{\bar{x}} = W^{\bar{x}}_{x_1} \times \ldots \times W^{\bar{x}}_{x_d} \subset \prod \mathbb{P}(V_i)$ such that for all $i = 1, \ldots, d$, for all $k \leq \dim V_i - 1$, and for all $(y^i_1, \ldots, y^i_k) \in W^{\bar{x}}_i$, $z^i \in W^{\bar{x}}_i$ and γ_i's as above ; we have

$$d_i(< \rho_i(\gamma^{\bar{x}}_{i_1}).y^i_1, \ldots, \rho_i(\gamma^{\bar{x}}_{i_k}).y^i_k >, \rho_i(\gamma^{\bar{x}}_j).z^i) > \eta^{\bar{x}}_t \tag{3.6}$$

Up to reducing $\eta^{\bar{x}}_t$, we can suppose that for each $i = 1, \ldots, d$; $B_i(x_i, 2\eta^{\bar{x}}_t) \subset W_i$. Now, cover the compact set $\prod \mathbb{P}(V_i)$ by the open sets $\bigcup\limits_{\bar{x} \in \prod \mathbb{P}(V_i)} \prod_{i=1}^d B_i(x_i, \eta^{\bar{x}}_t)$ and extract a finite subcover. Let us call the elements $\bar{x}^1, \ldots, \bar{x}^n \in \prod \mathbb{P}(V_i)$ such that $(\prod_{i=1}^d B_i(x^j_i, \eta^{\bar{x}^j}_t))_{j=1,\ldots,n}$ is the extracted finite subcover, and put $\eta_t := \min_{j=1,\ldots,n} \eta^{\bar{x}^j}_t$ and $M_t := \bigcup_{j=1}^n \{\gamma^{\bar{x}^j}_1, \ldots, \gamma^{\bar{x}^j}_t\}$.

Then, the result of the lemma readily follows : as in the assertion of the lemma, let $\bar{x} = (x_1, \ldots, x_d) \in \prod \mathbb{P}(V_i)$. Let also, up to reindexing, $\bar{x}^1$ be such that for each $i = 1, \ldots, d$; $d_i(x_i, x^1_i) < \eta^{\bar{x}^1}_t$ and take $\gamma^{\bar{x}^1}_1, \ldots, \gamma^{\bar{x}^1}_t \in M_t$. Then,

1. To see the first statement, fix $i \in \{1, \ldots, d\}$ and $j \neq k \in \{1, \ldots, t\}$, and consider $y_i, z_i \in B_i(x_i, \eta_t)$. Since $d_i(x_i, x^1_i) < \eta^{\bar{x}^1}_t$, $\eta_t \leq \eta^{\bar{x}^1}_t$ and $B_i(x^1_i, 2\eta^{\bar{x}^1}_t) \subset W^{\bar{x}^1}_i$, we have $B_i(x_i, \eta_t) \subset B_i(x^1_i, 2\eta^{\bar{x}^1}_t) \subset W^{\bar{x}^1}_i$, so that by (3.6) $d_i(\rho_i(\gamma_j).y_i, \rho_i(\gamma_k).z_i) > \eta^{\bar{x}^1}_t \geq \eta_t$, establishing the claim.
2. The proof of the second statement is similar. Fix $i \in \{1, \ldots, d\}$ and $i_1, \ldots, i_k, j \in \{1, \ldots, t\}$ with $j \notin \{i_1, \ldots, i_k\}$ and set $k = \dim V_i - 1$. For all $y_{i_1}, \ldots, y_{i_k}, z_i \in B_i(x_i, \eta_t)$, exactly as above, we have $y_{i_1}, \ldots, y_{i_k}, z_i \in B_i(x_i, \eta_t) \in W^{\bar{x}^1}_i$ so that (3.6) again proves the claim.

□

Remark 3.19. *A similar observation as Remark 2.25 of the Abels-Margulis-Soifer finiteness result, clearly applies to this finiteness result as well. Namely, for all $t \in \mathbb{N}$, there exists a constant $\eta_t \in \Gamma$, a finite subset M_t of Γ and for each $\gamma \in M_t$, bounded neighbourhoods V_γ of γ in G such that we have the conclusions of the lemma for every $\gamma'_i \in V_{\gamma_i}$, instead of only γ_i's for $i = 1, \ldots, d$. We shall use the* ***same constants*** *η_t for this extended result and Lemma 3.18.*

Lemma 3.20. *Let V be a finite dimensional Euclidean space and $g \in GL(V)$. For the action of $GL(V)$ on $\mathbb{P}(V)$ (endowed with the Fubini-Study metric), g is a $||\Lambda^2 g||.||g^{-1}||^2$-Lipschitz transformation.*

Démonstration. Indeed, for $x, y \in \mathbb{P}(V)$, we have

$$d(gx, gy) = \frac{||gx \wedge gy||}{||gx||.||gy||} \leq \frac{||\Lambda^2 g||.||x \wedge y||}{||g^{-1}||^{-2}.||x||.||y||} = ||\Lambda^2 g||.||g^{-1}||^2 d(x, y)$$

□

Accordingly, for an element $\gamma \in G$, put

$$L(\gamma) := \max_{i=1,\ldots,d} ||\Lambda^2 \rho_i(\gamma)||.||\rho_i(\gamma)^{-1}||^2 \in [1, \infty[. \tag{3.7}$$

The next technical lemma is based on the observation that if a proximal element g, when multiplied on the left by an arbitrary element γ, gives a proximal element γg, then the projective hyperplane $X^<_{\gamma g}$ is close to that of g, while the attracting directions $x^+_{\gamma g}$ and x^+_g may differ arbitrarily. The rest of the proof is along the same lines as the so called Tits proximality criterion (See [111] 3.8, [16] 2.1, [11] Lemme 6.2).

Lemma 3.21. *Let g be an (r,ϵ)-loxodromic element of G and $\gamma \in G$ such that $L(\gamma).\epsilon < 1$. Put $1 > \epsilon_1 := L(\gamma)\epsilon \geq \epsilon$ and suppose there exists a δ with $\delta > 6\epsilon_1$ such that for each $i = 1,\dots,d$, we have $d_i(\rho_i(\gamma)x^+_{\rho_i(g)}, X^<_{\rho_i(g)}) > \delta$. Then, γg is $(\frac{\delta}{3}, 2\epsilon_1)$-loxodromic. Moreover, for each $i = 1,\dots,d$, we have $d(x^+_{\rho_i(\gamma g)}, \gamma x^+_{\rho_i(g)}) < \epsilon_1$ and $d_H(X^<_{\gamma g}, X^<_g) < \epsilon$.*

Démonstration. To ease the notation, we will dismiss the representations ρ_i. By our definition of $L(.)$ in (3.7), our reasonings apply simultaneously to each representation ρ_i for $i = 1,\dots,d$.

We first establish that γg is loxodromic. One first observes that we have

$$\gamma g B^\epsilon_g \subseteq \gamma b^\epsilon_g \subseteq B(\gamma x^+_g, \epsilon L(\gamma)) \subseteq B^{4\epsilon_1}_g \tag{3.8}$$

where the first inclusions is by (r,ϵ)-loxodromy of g and the last by out hypothesis that $d(\gamma x^+_g, X^<_g) > \delta \geq 6\epsilon_1$.

Moreover, the restriction of the action of γg on B^ϵ_g is $L(\gamma)\epsilon = \epsilon_1$ Lipschitz with, by hypothesis, $\epsilon_1 < 1$. Therefore, γg is a continuous contraction of the compact B^ϵ_g into $B^{4\epsilon_1}_g \subseteq \overset{\circ}{B^\epsilon_g}$ and thus, by Banach fixed point theorem, has a unique attracting fixed point, of basin of attraction containing B^ϵ_g. This indeed implies that γg is loxodromic. One also sees from (3.8) that we must have $x^+_{\gamma g} \in B(\gamma x^+_g, \epsilon_1)$ and $d_H(X^<_{\gamma g}, X^<_g) < \epsilon$.

To get the complete statement of the lemma, in view of the definition of a $(\frac{\delta}{3}, 2\epsilon_1)$-loxodromic element, one checks that

1. Since by above $x^+_{\gamma g} \in B(\gamma x^+_g, \epsilon_1)$ and $d_H(X^<_{\gamma g}, X^<_g) < \epsilon$, and by hypothesis $d(\gamma x^+_g, X^<_g) > \delta \geq 6\epsilon_1$, we have $d(x^+_{\gamma g}, X^<_{\gamma g}) \geq \delta - \epsilon - \epsilon_1 \geq \delta - 2\epsilon_1 > 2\frac{\delta}{3}$.
2. Similarly, we have $\gamma g B^{2\epsilon_1}_{\gamma g} \subseteq \gamma g B^\epsilon_g \subseteq B(\gamma x^+_g, \epsilon_1) \subseteq b^{2\epsilon_1}_{\gamma g}$.
3. Finally, the restriction of the action of γg on $B^{2\epsilon_1}_{\gamma g} \subseteq B(g)^\epsilon$ is $\epsilon_1 = \epsilon L(\gamma)$ Lipschitz, as observed above.

These establish our claim. □

Before proceeding with the next proposition, let us recall the following elementary fact about the Fubini-Study metric d on $\mathbb{P}(V)$: if f is a linear functional on V with hyperplane H and $w \in V \setminus \{0\}$, then we have $d(\overline{w}, \mathbb{P}(H)) = \frac{|f(w)|}{||f||.||w||}$. From this, it also follows that the Hausdorff distance between the projective images $\mathbb{P}(H_1), \mathbb{P}(H_2)$ of two hyperplanes H_1, H_2 in V, is equal to the distance between the projective images of orthogonal vectors to H_1 and H_2 : for $v_1 \perp H_1$ and $v_2 \perp H_2$; $d(\overline{v_1}, \overline{v_2}) = d_H(\mathbb{P}(H_1), \mathbb{P}(H_2))$.

In the next proposition, we exploit more deeply the observation mentioned before the last lemma, in its relation with the result of Lemma 3.18 and the notion of narrowness of a set of loxodromic elements. It says that the union of left translates by suitable elements of two sufficiently narrow and contracting Schottky families is a Schottky family. By its probabilistic Corollary 3.24, it will be of crucial use in proving the convexity of the rate function.

Let us fix some notation before stating it : let t be a fixed natural number with $t > 2\sum_{i=1}^{d}(\dim V_i - 1)$. Let $\eta_t > 0$ and the finite subset M_t of Γ be as given by Lemma 3.18. For a subset M of G, denote by $L(M) = \max_{\gamma \in M}(L(\gamma) \vee L(\gamma^{-1})) \in [1, \infty]$ where $L(\gamma)$ is defined as in (3.7). Observe that by Lemma 3.20, for any $M \subset G$ contained in a compact of G, we have $L(M) < \infty$. With these notations, we have :

Proposition 3.22. *Let E_1 and E_2 be two (r, ϵ)-Schottky families with $\epsilon < \frac{\eta_t}{96L(M_t)^2}$. Suppose also that E_1 and E_2 are $(\frac{\eta_t}{4L(M_t)^2}, \frac{\eta_t}{4L(M_t)^2})$-narrow. Then, there exist γ_1 and γ_2 in M_t such that $\gamma_1 E_1 \cup \gamma_2 E_2$ is (r_1, ϵ_1)-Schottky family and we can take $r_1 = \frac{\eta_t}{48L(M_t)}$ and $\epsilon_1 = 2\epsilon L(M_t)$.*

Démonstration. To simplify the notation, we will only work in one fixed representation (ρ, V) among $(\rho_i, V_i)_{i=1,\dots,d}$ and dismiss that from the notation as in the proof of the previous lemma. Our reasonings are such that they simultaneously apply to all representations $(\rho_i, V_i)_{i=1,\dots,d}$; except at one point at the very end of the proof, where of course we will take into account all representations (we explicitly indicate that point).

By hypothesis, there exist Y^1 and Y^2, subsets of $\mathbb{P}(V)$ of diameter less than $\frac{\eta_t}{4L(M_t)^2}$ and such that for $i = 1, 2$, for all $g \in E_i$, we have $x_g^+ \in Y^i$. Let y_1 and y_2 be respectively in Y^1 and Y^2 such that for $i = 1, 2$; $E_i^+ := \{x_g^+ \mid g \in E_i\} \subseteq B(y_i, \frac{\eta_t}{4L(M_t)^2})$. Take elements $\gamma_{1,1}, \dots, \gamma_{1,t}$ and $\gamma_{2,1}, \dots, \gamma_{2,t}$ from M_t satisfying the conclusions of Lemma 3.18 respectively for the points y_1 and y_2.

Reformulating the conclusion 2) of Lemma 3.18 ; we have that for each hyperplane $H \subset V$; there exist at most k distinct indices $i_1, \dots, i_k \subset \{1, \dots, t\}$ with $k \leq \dim V - 1 =: d-1$, such that for each $l = 1, \dots, k$, $\mathbb{P}(H) \cap \gamma_{1,i_l}.B(y_1, \eta_t) \neq \emptyset$. Indeed, otherwise there exist $u_1, \dots, u_d \in B(y_1, \eta_t)$ and $\gamma_{1,i_1}, \dots, \gamma_{1,i_d} \in M_t$ such that $\mathbb{P}(H)$ contains the projective image of the span of the lines $\{\gamma_{1,i_1}.u_1, \dots, \gamma_{1,i_d}.u_d\}$ contradicting the conclusion of Lemma 3.18 since $d = \dim V$. (Of course, the same conclusion holds true for γ_{1,i_j}'s replaced by γ_{2,i_j}'s and y_1 by y_2)

Meanwhile, note that for each $\gamma \in M_t$, $x \in \mathbb{P}(V)$ and $\delta \geq 0$, by definition of $L(M_t)$, we have

$$\gamma B(x, \delta) \subseteq B(\gamma.x, L(M_t\delta)) \subseteq \gamma B(x, L(M_t)^2\delta) \tag{3.9}$$

Now, we claim that at most $d-1$ distinct elements $\gamma_{1,i_1}, \dots, \gamma_{1,i_k}$ among $\{\gamma_{1,1}, \dots, \gamma_{1,t}\}$ such that

$$B(\gamma_{1,i_j} y_1, \frac{\eta_t}{2L(M_t)}) \cap E_1^< \neq \emptyset \tag{3.10}$$

where we have put $E_1^< = \bigcup_{g \in E_1} X_g^<$.

Indeed, if $i \in \{1,\dots,t\}$ is such that $B(\gamma_{1,i}y_1, \frac{\eta_t}{2L(M_t)}) \cap E_1^< \neq \emptyset$, then since by hypothesis for all $g,h \in E_1$, one has $d_H(X_g^<, X_h^<) < \frac{\eta_t}{4L(M_t)^2}$, we have that for each $g \in E_1$; $B(\gamma_{1,i}y_1, \frac{1+2L(M_t)}{4L(M_t)^2}\eta_t) \cap X_g^< \neq \emptyset$. But by (3.9), since $L(M_t) \geq 1$, this implies that $\gamma_{1,i}B(y_1, \frac{1+2L(M_t)}{4L(m_t)}\eta_t) \cap X_g^< \neq \emptyset$ for each $g \in E_1$. Therefore, as $E_1 \neq \emptyset$, we have found an hyperplane $\mathbb{P}(H)$ in $\mathbb{P}(V)$ (take $H = X_g^<$ for an element $g \in E_1$) such that for each $i \in \{1,\dots,t\}$ satisfying (3.10), we have $\gamma_{1,i}B(y_1, \frac{1+2L(M_t)}{4L(m_t)}\eta_t) \cap \mathbb{P}(H) \neq \emptyset$. Since $\frac{1+2L(M_t)}{4L(m_t)} < 1$, the above reformulation of the conclusion of Lemma 3.18 tells us that there are at most $\dim V - 1$ such indices $i \in \{1,\dots,t\}$. Put

$$D_1 := \{i \in \{1,\dots,t\} \mid B(\gamma_{1,i}y_1, \frac{\eta_t}{2L(M_t)}) \cap E_1^< \neq \emptyset\}$$

so that $|D_1| \leq \dim V - 1$.

Observe then that for each $i \in \{1,\dots,t\} \setminus D_1$, $g \in E_1$ and $x \in X_g^<$, we have

$$d(B(\gamma_{1,i}y_1, \frac{\eta_t}{4L(M_t)}), x) \geq \frac{\eta_t}{4L(M_t)} \tag{3.11}$$

Therefore, since $E_1^+ \subseteq B(y_1, \frac{\eta_t}{4L(m_t)^2})$, by (3.9) we have that for each $\gamma \in M_t$; $\gamma E_1^+ \subset B(\gamma.y_1, \frac{\eta_t}{4L(M_t)})$ so that (3.11) implies

$$d(\gamma_{1,i}x_g^+, X_h^<) \geq \frac{\eta_t}{4L(M_t)} \tag{3.12}$$

for all $g,h \in E_1$ and for each $i \in \{1,\dots,t\} \setminus D_1$.

As a consequence, since by hypothesis $\epsilon < \frac{1}{L(M_t)}$ and $6\epsilon L(M_t) < \frac{\eta_t}{4L(M_t)}$, Lemma 3.21 is in force and gives that for each $i \in \{1,\dots,t\} \setminus D_1$ and $g \in E_1$; $\gamma_{1,i}g$ is $(\frac{\eta_t}{12L(M_t)}, 2\epsilon L(M_t))$-loxodromic. Moreover, $d(x^+_{\gamma_{1,i}g}, \gamma_{1,i}x_g^+) < 2\epsilon L(M_t)$ and $d_H(X^<_{\gamma_{1,i}g}, X_g^<) < \epsilon$.

Combining these last two inequalities with (3.12), one sees that for all $g,h \in E_1$, and for each $i \in \{1,\dots,t\} \setminus D_1$, we have

$$d(x^+_{\gamma_{1,i}g}, X^<_{\gamma_{1,i}h}) \geq \frac{\eta_t}{2L(M_t)} - 2\epsilon L(M_t) - \epsilon \geq \frac{\eta_t}{8L(M_t)} \tag{3.13}$$

Hence, it follows that for each $i \in \{1,\dots,t\} \setminus D_1$, $\gamma_{1,i}E_1$ is a $(\frac{\eta_t}{48L(M_t)}, 2\epsilon L(M_t))$-Schottky family.

Repeating exactly the same argument for E_2, one finds a subset D_2 of $\{1,\dots,t\}$ such that $|D_2| \leq \dim V - 1$ and for each $i \in \{1,\dots,t\} \setminus D_2$, one has that $\gamma_{2,i}E_2$ is a $(\frac{\eta_t}{48L(M_t)}, 2\epsilon L(M_t))$-Schottky family.

Again, the same reasoning, replacing in (3.10) $E_1^<$ by $E_2^<$, allows us to see that there exist at most $\dim V - 1$ indices $i \in \{1, \dots, t\}$, denoting the set of these by D_{12}, such that for each $g \in E_1$, $h \in E_2$ and $i \in \{1, \dots, t\} \setminus D_{12}$; we have $d(\gamma_{1,i} x_g^+, X_h^<) \geq \frac{\eta_t}{4L(M_t)}$. By the same token, we get $D_{21} \subset \{1, \dots, t\}$ with the corresponding properties.

By consequent, it follows that for each $i_1 \in \{1, \dots, t\} \setminus D_1 \cup D_{12}$ and $i_2 \in \{1, \dots, t\} \setminus D_2 \cup D_{21}$, $\gamma_{1,i_1} E_1 \cup \gamma_{2,i_2} E_2$ is a $(\frac{\eta_t}{48L(M_t)}, 2\epsilon L(M_t))$-Schottky family in $\mathbb{P}(V)$.

At this point, as indicated at the beginning of the proof, regarding the construction of the index sets D_1, D_2, D_{12}, D_{21}, we must take into account each of the d representations $\rho_1, \dots, \rho_d$. Hence, repeating the same procedure for each ρ_i, we get index subsets $D_1^j, D_2^j, D_{12}^j, D_{21}^j$ of $\{1, \dots, t\}$ for each $j = 1, \dots, d$ with cardinality at most $\dim V_j - 1$.

Finally, denoting $\tilde{D}_1 := \bigcup_{j=1}^{d} (D_1^j \cup D_{12}^j)$ and $\tilde{D}_2 := \bigcup_{j=1}^{d} (D_2^j \cup D_{21}^j)$, since for $i = 1, 2$, $t > 2 \sum_{j=1}^{r} (\dim V_j - 1) \geq |\tilde{D}_i|$, we have $\{1, \dots, t\} \setminus \tilde{D}_i \neq \emptyset$. As a result, choosing $\gamma_i \in \{1, \dots, t\} \setminus \tilde{D}_i$ for $i = 1, 2$, we get that $\gamma_1 E_1 \cup \gamma_2 E_2$ is a $(\frac{\eta_t}{48L(M_t)}, 2\epsilon L(M_t))$-Schottky family, proving the proposition.

□

Remark 3.23. *One notes from the proof that this proposition is also true with γ_i replaced by any γ_i' in the neighbourhood V_{γ_i} of γ_i given by Remark 3.19 for $i = 1, 2$, and $L(M_t)$ by $L(\cup_{\gamma \in M_t} V_\gamma)$.*

Combining the previous proposition with Lemma 3.7 and Corollary 3.8, we obtain the following technical probabilistic corollary which will be an essential step in our proof of convexity of the rate function. In the corollary, we denote by L, the Lipschitz constant $L(\cup_{\gamma \in M_t} V_\gamma)$ of the union of neighbourhoods of elements of M_t given by Remark 3.19. Since M_t is a finite set and V_γ's are bounded, we have $L \in [1, \infty[$.

Corollary 3.24. *Let ϵ and r be given with $0 < \epsilon < \frac{r}{6} \wedge \frac{\eta_t}{96L^2}$. Then, there exist a natural number $i_1 = i_1(\mu, M_t)$, a constant $d_4 > 0$ depending on the probability measure and a compact subset $\tilde{K}$ of $\mathfrak{a}$ with the property that for all subsets E_1 and E_2 of Γ consisting of (r, ϵ)-loxodromic elements, for all $n_1, n_2 \in \mathbb{N}$ there exist two natural numbers $n_1 + i_1 \geq n_{1,1} \geq n_1$ and $n_2 + i_1 \geq n_{2,2} \geq n_2$, two (r_1, ϵ_1)-Schottky families $\tilde{E}_1$ and $\tilde{E}_2$ such that $\tilde{E}_1 \cup \tilde{E}_2$ is an (r_1, ϵ_1)-Schottky family and for $i = 1, 2$, $\mathbb{P}(S_{n_{i,i}} \in \tilde{E}_i) \geq \mathbb{P}(S_{n_i} \in E_i).d_4$. Moreover, we have $\kappa(\tilde{E}_i) \subset \kappa(E_i) + \tilde{K}$, and one can choose $r_1 = \frac{\eta_t}{48L}$ and $\epsilon_1 = 2\epsilon L$.*

Démonstration. Write $M_t = \{\gamma_1, \dots, \gamma_m\}$ and put $i_1 = i_1(\mu, M_t)$ a natural number such that $M_t \subset \bigcup_{i=1}^{i_1} (\mathrm{supp}(\mu^{*i}))$. For each $i = 1, \dots, m$, take neighbourhoods V_{γ_i} of γ_i's as in Remark 3.19, set $k_i \leq i_1$ such that $\mu^{*k_i}(V_{\gamma_i}) =: \beta_i > 0$ and finally put $\beta := \min_{1 \leq i \leq m} \beta_i > 0$. Furthermore, taking the compact subset $\cup_{i=1}^m \overline{V}_{\gamma_i}$ of G as L in Lemma 2.19, get a compact subset $\tilde{K}$ of $\mathfrak{a}$ satisfying the conclusion of Lemma 2.19. Let also $d_2 = d_2(t, \Gamma) > 0$ be the constant given by Lemma 3.7, in which we take $a = b = \frac{\eta_t}{4L^2}$, $d_3 = d_3(r) > 0$ be the constant given by Corollary 3.8 and finally set $d_4 = d_2 d_3 \beta > 0$.

Let now E_1 and E_2 be two given subsets of Γ consisting of (r,ϵ)-loxodromic elements and $n_1, n_2 \in \mathbb{N}$. Applying Corollary 3.8 for E_1 and E_2, there exist two $(\frac{r}{6},\epsilon)$-Schottky families, $E_1' \subset E_1$ and $E_2' \subset E_2$ such that for $i = 1,2$

$$\mathbb{P}(S_{n_i} \in E_i') \geq \mathbb{P}(S_{n_i} \in E_i).d_3 \tag{3.14}$$

Noting that subsets of (r,ϵ)-Schottky families are themselves (r,ϵ)-Schottky families, using (3.14) and applying Lemma 3.7 twice with $a = b = \frac{\eta_t}{4L^2}$ for respectively E_1', E_2' and n_1, n_2, we get two $(\frac{\eta_t}{4L},\frac{\eta_t}{4L})$-narrow $(\frac{r}{6},\epsilon)$-Schottky families $\hat{E}_1 \subset E_1'$ and $\hat{E}_2 \subset E_2'$ such that for $i = 1,2$

$$\mathbb{P}(S_{n_i} \in \hat{E}_i) \geq \mathbb{P}(S_{n_i} \in E_i)d_3 d_2 \tag{3.15}$$

Now applying Proposition 3.22 and Remark 3.23 to the $(\frac{r}{6},\epsilon)$-Schottky families $\hat{E}_1$ and $\hat{E}_2$, remarking that the hypotheses of that proposition is satisfied by the constructions of $\hat{E}_1$ and $\hat{E}_2$, we get that, up to reindexing, there exist γ_1, γ_2 in M_t such that, setting for $i = 1,2$, $\tilde{E}_i := V_{\gamma_i}\hat{E}_i$, $\tilde{E}_1 \cup \tilde{E}_2$ is an (r_1,ϵ_1)-Schottky family, where we can take $r_1 = \frac{\eta_t}{48L}$ and $\epsilon_1 = 2\epsilon L$.

Then, setting $n_{1,1} := n_1 + k_1 \leq n_1 + i_1$ and $n_{2,2} = n_2 + k_2 \leq n_2 + i_1$; by independence of random walk increments, for $i = 1,2$, we have

$$\begin{aligned}\mathbb{P}(S_{n_{i,i}} \in \tilde{E}_i) &\geq \mathbb{P}(X_{n_i+k_i}.\ldots.X_{n_i} \in V_{\gamma_i} \text{ and } S_{n_i} \in \hat{E}_i)\\ &= \mathbb{P}(S_{n_i} \in \hat{E}_i)\mathbb{P}(S_{k_i} \in V_{\gamma_i}) \geq \mathbb{P}(S_{n_i} \in E_i)\beta d_3 d_2 = \mathbb{P}(S_{n_i} \in E_i)d_4\end{aligned}$$

Finally, one remarks that for $i = 1,2$, we have $\tilde{E}_i \subset M_t\hat{E}_i \subset M_t E_i$ so that by choice of $\tilde{K}$, Lemma 2.19 implies that $\kappa(\tilde{E}_i) \subset \kappa(E_i) + \tilde{K}$, establishing the last claim. □

We are now in a position to prove the convexity result :

Proof of Theorem 3.1 (Convexity of the rate function). Denoting the rate function by I, start by observing that, by lower semi-continuity, it is sufficient to show that for all $x_1, x_2 \in \mathfrak{a}$, we have $I(\frac{x_1+x_2}{2}) \leq \frac{I(x_1)}{2} + \frac{I(x_2)}{2}$. For this, we can indeed suppose that x_1, x_2 belongs to the effective domain D_I of I, where $D_I := \{x \in \mathfrak{a} \,|\, I(x) < \infty\}$. We shall argue by contradiction.

Suppose there exists $x_1, x_2 \in D_I$ with $I(\frac{x_1+x_2}{2}) > \frac{I(x_1)}{2} + \frac{I(x_2)}{2} + 5\xi$ for some $\xi > 0$. By the weak LDP and Remark 2.5, I satisfies

$$I(x) = \sup_{\substack{O \text{ open} \\ x \in O}} -\limsup_n \frac{1}{n}\log\mu_n(O) = \sup_{\substack{O \text{ open} \\ x \in O}} -\liminf_n \frac{1}{n}\log\mu_n(O) \tag{3.16}$$

hence, we can find neighbourhoods $O_1^{12} \subset O_2^{12}$ of $\frac{x_1+x_2}{2}$; where the inclusions are super-strict and such that

$$-\limsup_n \frac{1}{n}\log\mu_n(O_2^{12}) \geq \frac{I(x_1)}{2} + \frac{I(x_2)}{2} + 4\xi. \tag{3.17}$$

By (3.16) and (3.17), for $i = 1, 2$, one can also find neighborhoods $x_i \subset O_1^i \subset O_2^i \subset O_3^i$ where the inclusions are super-strict and O_i^j's are such that $O_3^1 \cap O_3^2 = \emptyset$, $\frac{O_3^1 + O_3^2}{2} \subset O_1^{12}$ and

$$-\limsup_n \frac{1}{n} \log \mu_n(O_2^{12}) \geq \frac{1}{2} \sum_{i=1}^{2} -\liminf_n \frac{1}{n} \log \mu_n(O_1^i) + 3\xi \tag{3.18}$$

It follows from (3.18) that, there exists $N_0 \in \mathbb{N}$ such that for all $n \geq N_0$, we have

$$-\limsup_n \frac{1}{n} \log \mu_n(O_2^{12}) \geq \frac{1}{2} \sum_{i=1}^{2} -\frac{1}{n} \log \mu_n(O_1^i) + 2\xi \tag{3.19}$$

Now, let $r = r(\Gamma) > 0$ be as given by Theorem 2.24, $t = 1 + 2\sum_i^r (\dim V_i - 1)$, $\eta_t > 0$, the finite set $M_t \subset \Gamma$ as given by Lemma 3.18, for each $\gamma \in M_t$, its neighbourhood V_γ as in Remark 3.19 and set $L \geq 1$ to be the Lipschitz constant $L(\cup_{\gamma \in M_t} V_\gamma)$. Choose $\epsilon < \frac{r}{6} \wedge \frac{\eta_t}{96L^2}$. Put $r_1 = \frac{\eta_t}{48L}$ and $\epsilon_1 = 2\epsilon L$. Let also the constants $d_1 = d_1(\epsilon, \Gamma, \mu)$, $i_0 = i_0(\epsilon, \Gamma, \mu)$ and the compact subset $C = C(\epsilon, \Gamma)$ of $\mathfrak{a}$ be as given by Lemma 3.5. Denote by K the compact set $K(r_1, \epsilon_1) \subset \mathfrak{a}$ given by Proposition 3.9. Let also the compact set $\tilde{K}$ and the constants $d_4 > 0$, $i_1 = i_1(\mu, M_t)$ be as in Corollary 3.24. Finally, fix $Q \in \mathbb{N}$ with for $i = 1, 2$, $Q \geq Q(O_1^i, O_2^i, C + \tilde{K}) \vee Q(O_2^i, O_3^i, K)$ and $q = q(O_1^{12}, O_2^{12}) < 1$, where $Q(.,.,.)$ and $q(.,.)$ are as defined in Lemma 3.16.

Now, choose $n_0 \in \mathbb{N}$ with

1. $n_0 \geq N_0$
2. $e^{-n_0 \xi} \leq d_1 d_4$
3. $n_0 \geq Q$
4. $\frac{n_0}{n_0 + i_0 + i_1} > q$

and put for $i = 1, 2$, $\alpha_i = -\frac{1}{n_0} \log \mu_{n_0}(O_1^i)$ and $\beta = -\limsup_n \frac{1}{n} \log \mu_n(O_2^{12})$ so as to have by Item 1 of the choice of n_0 and (3.19) that

$$\beta \geq \frac{\alpha_1 + \alpha_2}{2} + 2\xi \tag{3.20}$$

Applying Lemma 3.5 twice, once with taking $A = n_0 O_1^1$ and the other $A = n_0 O_1^2$ in that lemma, one gets $n_1, n_2 \in \mathbb{N}$ with for $i = 1, 2$ $n_0 + i_0 \geq n_i \geq n_0$ and

$$\mathbb{P}(\kappa(S_{n_i}) \in n_0 O_1^i + C \text{ and } S_{n_i} \text{ is } (r, \epsilon)\text{-loxodromic}) \geq e^{-n_0 \alpha_i} d_1 \tag{3.21}$$

Setting for $i = 1, 2$; $E_i := \kappa^{-1}(n_0 O_1^i + C) \cap \Gamma_{(r,\epsilon)}$, where $\Gamma_{(r,\epsilon)}$ denotes as before (r, ϵ)-loxodromic elements of Γ, by (3.21) E_i's are non-empty and by our choices of r and ϵ, they satisfy the hypotheses of Corollary 3.24. This corollary therefore gives that for some $n_{11}, n_{22} \in \mathbb{N}$ with for $i = 1, 2$; $n_0 + i_0 + i_1 \geq n_{ii} \geq n_0$, there exist two (r_1, ϵ_1)-Schottky families $\tilde{E}_i$ such that $\tilde{E}_1 \cup \tilde{E}_2$ is also an (r_1, ϵ_1)-Schottky family with

$$\mathbb{P}(S_{n_{ii}} \in \tilde{E}_i \text{ and } \kappa(S_{n_{ii}}) \in n_0 O_1^i + C + \tilde{K}) \geq e^{-n_0 \alpha_i} d_1 d_4 \geq e^{-n_0(\alpha_i + \xi)} \tag{3.22}$$

by the definitions of E_i above and the last statement of Corollary 3.24 and where the last equality follows from the choice of n_0, namely Item 2. Furthermore, by Item 3 in the choice of n_0, (3.22), implies

$$\mathbb{P}(S_{n_{ii}} \in \tilde{E}_i \text{ and } \kappa(S_{n_{ii}}) \in n_0 O_2^i) \geq e^{-n_0\alpha_i} d_1 d_4 \geq e^{-n_0(\alpha_i+\xi)} \tag{3.23}$$

for $i = 1, 2$.

Observe now that by our initial choice of open sets, we have $O_3^1 \cap O_3^2 = \emptyset$, so that up to taking their intersections, respectively with $\kappa^{-1}(n_0 O_2^1)$ and $\kappa^{-1}(n_0 O_2^2)$, we can suppose that $\tilde{E}_1$ and $\tilde{E}_2$ are disjoint and are such that for $i = 1, 2$, $\kappa(\tilde{E}_i) \subseteq n_0 O_2^i$. Now, for all $k_1, k_2 \geq 0$ define the collection of subsets E^{k_1,k_2} of Γ by

$$E^{k_1,k_2} = \{\gamma_1 \dots \gamma_{k_1+k_2} \,|\, |\{i \,|\, \gamma_i \in \tilde{E}_j\}| = k_j \text{ for } j = 1, 2\}$$

Making key use of the fact that $\tilde{E}_1 \cup \tilde{E}_2$ is an (r_1, ϵ_1)-Schottky family, Proposition 3.9 implies that for all $k_1, k_2 \geq 0$,

$$\kappa(E^{k_1,k_2}) \subset k_1(n_0 O_2^1 + K) + k_2(n_0 O_2^2 + K) \subset k_1 n_0 O_3^1 + k_2 n_0 O_3^2 \tag{3.24}$$

where the last inclusion is due to Item 3 of the choice of n_0. Hence, for all $k \geq 0$, choosing $k = k_1 = k_2$, since $\frac{O_3^1+O_3^2}{2} \subseteq O_1^{12}$, it follows from (3.24) that $\kappa(E^{k,k}) \subseteq 2kn_0 O_1^{12}$. Moreover, Item 4 of the choice of n_0 implies by Lemma 3.16 that for all $k \geq 0$, we have $2kn_0 O_1^{12} \subseteq k(n_{11} + n_{22}) O_2^{12}$.

Consequently, we have the following inclusion of events for each $k \geq 0$:

$$\{S_{kn_{11}+kn_{22}} \in E^{k,k}\} \subset \{\frac{1}{kn_{11} + kn_{22}} \kappa(S_{kn_{11}+kn_{22}}) \in O_2^{12}\} \tag{3.25}$$

Now, using, respectively, (3.25), independence of random walk increments and (3.23), for all $k \geq 1$, we have

$$\begin{aligned}
\mathbb{P}(\frac{\kappa(S_{kn_{11}+kn_{22}})}{kn_{11} + kn_{22}} \in O_2^{12}) &\geq \mathbb{P}(S_{kn_{11}+kn_{22}} \in E^{k,k}) \\
&\geq \mathbb{P}(S_{n_{11}} \in \tilde{E}_1)^k \mathbb{P}(S_{n_{22}} \in \tilde{E}_2)^k \\
&\geq e^{-kn_0(\alpha_1+\xi)} e^{-kn_0(\alpha_2+\xi)}
\end{aligned}$$

As a result, in the above inequality, taking logarithm, dividing by k, it follows that

$$-\beta(n_{11} + n_{22}) \geq \limsup_{k\to\infty} \frac{1}{k} \log \mathbb{P}(\frac{\kappa(S_{k(n_{11}+n_{22})})}{k(n_{11} + n_{22})} \in O_2^{12}) \geq -2n_0(\frac{\alpha_1 + \alpha_2}{2} + \xi)$$

where the first inequality is immediate by definition of β above.

Finally, dividing this last inequality by $-(n_{11} + n_{22})$, using (3.20), we get $\frac{\alpha_1+\alpha_2}{2} + 2\xi \leq \beta \leq \frac{\alpha_1+\alpha_2}{2} + \xi$, a contradiction. $\square$

3.3 Proof of Theorem 3.2

Existence of full LDP under exponential moment condition

In our setting, the strengthening of a weak LDP to a full LDP with proper rate function, requires the exponential tightness (Definition 2.6) of the sequence of probability measures in question. The aim of this subsection is to show that the finite exponential moment condition (defined in the introduction, see also below) on the probability μ governing the random walk, suffices to get the exponential tightness of the sequence of laws μ_n of $\frac{1}{n}\kappa(S_n)$.

Before dealing with the general case, we first show in a quick corollary that, in case of a compactly supported measure μ, a full LDP with a proper convex rate function I can readily be established.

Corollary 3.25 (of Theorem 3.1)**.** *Assume the setting of Theorem 3.1, and suppose moreover that $supp(\mu)$ is contained in a compact set. Then, the sequence μ_n of probability measures satisfies a full LDP with the proper convex rate function.*

Démonstration. By Lemma 2.7, it suffices to have that the sequence μ_n of probability measures is exponentially tight. It is clear from Definition 2.6 that it suffices to show that there exists a compact set C_μ in $\mathfrak{a}$ such that $\operatorname{supp}(\mu_n) \subset C_\mu$ for every $n \geq 1$. Set $C_\mu := \{x \in \mathfrak{a} \mid -N_i \leq \overline{\chi}_{\rho_i}(x) \leq N_i\}$ where $\overline{\chi}_{\rho_i}$'s stand, as usual, for the derivatives at identity of the highest weights χ_{ρ_i} of the distinguished representations ρ_i's defined in Lemma 2.15, and for each $i = 1, \ldots, d$, $N_i := \sup_{\gamma \in supp(\mu)}(\log ||\rho_i(\gamma)||_i)$. Since by the Lemma 2.16, for all $\gamma \in G$, we have $\overline{\chi}_{\rho_i}(\kappa(\gamma)) = \log ||\rho_i(\gamma)||_i$, the fact that for every $n \geq 1$, $\operatorname{supp}(\mu_n) \subset C_\mu$ now follows from the submultiplicativity of the associated operators norm on V_i's. □

We now proceed to precise the aforementioned exponential moment condition : for all $g \in G$, put $M(g) := \max_{i=1,\ldots,d}(||\rho_i(g)||_i \vee ||\rho_i(g)^{-1}||_i) \geq 1$. Observe upon this definition that the submultiplicativity of the associated operator norms implies that for all $g, h \in G$, we have $M(gh) \leq M(g)M(h)$. A measure μ on G is then said to have a finite exponential moment if there exists a constant $c > 0$ with $\int M(g)^c \mu(dg) < \infty$. This notion is usual in the theory of random matrix products (see also 2. of Remark 4.10). We indeed fixed the Euclidean norms $||.||_i$ on V_i's, but it is also easy to show that this condition (existence of such a $c > 0$) does not depend on the choice of norms.

We shall now show that this condition is sufficient to get the exponential tightness of the sequence of laws μ_n of $\frac{1}{n}\kappa(S_n)$. Fixing such a constant $c > 0$ for the exponential moment of μ, we have :

Proposition 3.26. *If μ has a finite exponential moment, then the sequence of probability measures $(\mu_n)_{n\geq 1}$ is exponentially tight.*

Démonstration. The proof is based on the so-called Chernoff estimates and standard techniques of the large deviation theory, for a prior instance of such techniques, see for example the proof of Gärtner-Ellis theorem in [44].

For $\rho \geq 0$, set B_ρ to be the compact subset of $\mathfrak{a}$ defined by $B_t := \bigcap_{i=1}^d \{x \in \mathfrak{a} \mid -t \leq \overline{\chi}_i(x) \leq t\}$. By union of events bound, we have

$$\begin{aligned}&\mathbb{P}(\frac{1}{n}\kappa(S_n) \notin B_t) \leq \sum_{i=1}^{d} \mathbb{P}(|\overline{\chi}_i(\kappa(S_n))| \geq tn) = \\ &\sum_{i=1}^{d} \mathbb{P}(|\log ||\rho_i(S_n)||_i| \geq tn) \leq d.\mathbb{P}(\log M(S_n) \geq tn)\end{aligned} \tag{3.26}$$

where the last inequality follows from the union bound and the definition of $M(.)$.

Now, by Chebyshev inequality, for all $s \geq 0$, we have

$$\mathbb{P}(\log M(S_n) \geq tn) \leq \mathbb{E}[e^{s \log M(S_n)}]e^{-stn}$$

In this inequality, taking log, dividing by n and specializing to some $s_0 \in \mathbb{R}$ such that $c \geq s_0 > 0$, we obtain

$$\frac{1}{n}\log \mathbb{P}(\log M(S_n) \geq tn) \leq -(s_0 t - \frac{1}{n}\log \mathbb{E}[e^{s_0 \log M(S_n)}])$$

Consequently, we have

$$\limsup_{n\to\infty} \frac{1}{n}\log \mathbb{P}(\log M(S_n) \geq tn) \leq -(s_0 t - \limsup_{n\to\infty} \frac{1}{n}\log \mathbb{E}[e^{s_0 \log M(S_n)}])$$

By submultiplicativity of $M(.)$ and independence of random walk increments, it follows that for each $n \geq 1$, we have $\frac{1}{n}\log \mathbb{E}[e^{s_0 \log M(S_n)}] \leq \log \mathbb{E}[e^{s_0 \log M(X_1)}]$. But, this last quantity is a real number by the exponential moment condition and the choice of s_0 as $c \geq s_0 > 0$. As a result, we have

$$\lim_{t\to\infty} \limsup_{n\to\infty} \frac{1}{n}\log \mathbb{P}(\log M(S_n) \geq tn) \leq \lim_{t\to\infty} -(s_0 t - \log \mathbb{E}[e^{s_0 \log M(X_1)}]) = -\infty$$

establishing the exponential tightness in view of (3.26). $\square$

We can now write the

Proof of Theorem 3.2 (Existence of full LDP with proper rate function). Put together the last proposition and Theorem 3.1, then conclude by Lemma 2.7. $\square$

Identification of the rate function

The aim of this section is to prove the second assertion of Theorem 3.2, namely that in the case of a random walk associated to a probability measure μ with a strong moment condition, we can identify the obtained proper convex rate function I of the LDP with the Fenchel-Legendre transform of the limiting exponential moment generating function (see below) of the laws of $\frac{1}{n}\kappa(S_n)$.

Before expliciting the above statement, we want to point out that in the proof, we will follow a general scheme in large deviations theory (see Section 4.5.2 of [44] or 2.2 of [45] for a good account of this scheme). Namely, we exploit the convexity of the rate function I, with the Fenchel-Moreau duality and Varadhan's integral lemma, as main tools.

To proceed, define, in a similar fashion as in the definition of K in the proof of Proposition 3.26, $\overline{\Lambda} : \mathfrak{a}^* \longrightarrow \overline{\mathbb{R}}$ as

$$\overline{\Lambda}(\lambda) = \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}[e^{\lambda(\kappa(S_n))}]$$

In fact this function already provides large deviations upper bounds (see for example Section 4.5.1 of [44] and in case the limit in question exists, the limiting function is referred to as the limiting exponential moment generating function (of the sequence of laws μ_n) and nice properties (e.g. differentiability, steepness) of this function have implications for LDP (e.g. Gärtner-Ellis theorem).

We start by establishing a rather straightforward lemma, relating the exponential moment constant $c > 0$ with the locus of finiteness of $\overline{\Lambda}$.

Lemma 3.27. *Let μ be a probability measure on G of finite exponential moment with constant $c > 0$. Then in the above notation, we have*

$$D_{\overline{\Lambda}} := \{\lambda \in \mathfrak{a}^* \,|\, \overline{\Lambda}(\lambda) < \infty\} \supset \{\lambda = \sum_{i=1}^{d} \lambda_i \overline{\chi}_i \in \mathfrak{a}^* \mid |\lambda_i| \leq \frac{c}{r}\}$$

Proof of Lemma 3.27. For $\lambda = \sum_{i=1}^{d} \lambda_i \overline{\chi}_i \in \mathfrak{a}^*$; we have

$$\begin{aligned}
\overline{\Lambda}(\lambda) &= \limsup_{n} \frac{1}{n} \log \mathbb{E}[e^{\sum_{i=1}^{d} \lambda_i \overline{\chi}_i(\kappa(S_n))}] \\
&= \limsup_{n} \frac{1}{n} \log \mathbb{E}[e^{\sum_{i=1}^{d} \lambda_i \log ||\rho_i(S_n)||_i}] && \text{by Lemma 2.16} \\
&\leq \limsup \frac{1}{dn} \sum_{i=1}^{d} \log \mathbb{E}[e^{d\lambda_i \log ||\rho_i(S_n)||_i}] && \text{by Hölder inequality} \\
&= \max_{i=1,\dots,d} \limsup_{n} \frac{1}{dn} \log \mathbb{E}[e^{d\lambda_i \log ||\rho_i(S_n)||_i}]
\end{aligned} \tag{3.27}$$

Now fix $i_0 \in \{1, \dots, d\}$ and observe that for all $t \geq 0$, we have $M(g)^t \geq ||g||_{i_0}^t \geq M(g)^{-t}$. Hence, putting $a_n(t) := \log \mathbb{E}[e^{t \log M(S_n)}]$, $b_n(t) := \log \mathbb{E}[e^{t \log ||S_n||_{i_0}}]$ and $c_n(t) := \log \mathbb{E}[e^{-t \log M(S_n)}]$, it follows that for all $t \geq 0$, the sequences $a_n(t)$ and $b_n(t)$ are subadditive and $c_n(t)$ is superadditive. Therefore, in $\mathbb{R} \cup \{\pm\infty\}$, we have

$$a_1(t) \geq \lim_{n} \frac{a_n(t)}{n} \geq \lim_{n} \frac{b_n(t)}{n} \geq \lim_{n} \frac{c_n(t)}{n} \geq c_1(t)$$

But for $t \in [0, c]$, by the exponential moment condition $a_1(t) < \infty$ and for all $t \geq 0$, using Jensen inequality and again the moment condition, we have $c_1(t) = \log \mathbb{E}[M(X_1)^{-t}] = \log \mathbb{E}[(M(X_1)^c)^{\frac{-t}{c}}] \geq -\frac{t}{c} \log \mathbb{E}[M(X_1)^c] > -\infty$ from which the result of the lemma immediately follows. $\square$

Following this lemma, we now precise our strong moment condition : a probability measure μ on G is said to have strong exponential moment, if for all $c > 0$, $\int M(g)^c \mu(dg) < \infty$. This condition is clearly satisfied in case of a compactly supported μ. We now complete the

Proof of Theorem 3.2 (Identification of the rate function). It follows from Lemma 3.27 that if μ is of strong exponential moment, we have for all $\lambda \in \mathfrak{a}^*$, $\overline{\Lambda}(\lambda) < \infty$. Then, it follows from Varadhan's integral lemma (see [44] section 4.3) that we in fact have for all $\lambda \in \mathfrak{a}^*$

$$\overline{\Lambda}(\lambda) = \lim_n \frac{1}{n} \log \mathbb{E}[e^{<\lambda, \kappa(S_n)>}] = \sup_{x \in \mathfrak{a}}(< \lambda, x > - I(x))$$

where I is the proper rate function governing the LDP, whose existence under the finite exponential moment condition is proved in the previous subsection.

For a function f on $\mathfrak{a}$, denote its convex conjugate function (Fenchel-Legendre tranform) on $\mathfrak{a}^*$ by $f^*(.)$, where $f^*(\lambda) := \sup_{x \in \mathfrak{a}}(< \lambda, x > - f(x))$, so that the above conclusion of Varadhan's integral lemma writes $\overline{\Lambda}(\lambda) = I^*(\lambda)$. Now, since I is a convex rate function, Fenchel-Moreau duality (see Proposition 4.1 in [49] or Theorem 1.10 in [39]) tells us that $I(x) = I^{**}(x) = \overline{\Lambda}^*(x)$, identifying $I(x)$ with $\overline{\Lambda}^*(x)$ and completing the proof. □

3.4 LDP for norms of random matrix products

The aim of this section is to prove Theorem 3.3. This theorem is closely related to Theorem 3.1 and Theorem 3.2 : recall that (exponential of) the Cartan projection of a linear transformation $g \in GL(V)$, where V is a finite dimensional Euclidean space, is the diagonal matrix consisting of singular values of g (i.e. the eigenvalues of the positive definite symmetric transformation ${}^t gg$) positioned in the decreasing order ; i.e. it is the A matrix in the usual KAK decomposition of $GL(V)$. Note that the first (left-upmost) coefficient of the matrix A corresponds to the associated operator norm $||g||$ of g. The other coefficients can be given similar interpretations with the exterior power representations. Namely, for $i = 1, \ldots, d$, we have $|| \bigwedge^i g|| = a_1 \ldots a_i$ where, $g = K_1 A K_2$ and $A = $diag$(a_1, \ldots, a_d)$ and on $\bigwedge^i V$ we use the scalar product canonically associated to that of V. Note also that $|| \bigwedge^d g|| = det(g)$.

Consequently, as we shall see shortly, when the Zariski closure G of the semigroup Γ generated by the support of μ is Zariski connected, semisimple and V is irreducible as a G-module (e.g. $G = SL(V)$), Theorem 3.3 follows as a corollary of Theorem 3.1 and Theorem 3.2 using a general technique of large deviations theory, called contraction principle (see Lemma 3.28). The main novelty of Theorem 3.3 is that it is valid under a larger generality for the Zariski closure G of Γ, namely it includes non-connected semisimple groups and direct product of semisimple groups with central subgroups (see below for a precise description).

Let us then, first, state the contraction principle and indicate how we can use it to deal with the connected irreducible case. In fact, a reason for us to expose Corollary 3.29 as a separate result is to illustrate the use of the idea of contraction principle of LDP's, for we are inspired by and use this idea in Section 5.2, in a deterministic setting, while relating our and Quint's growth indicator functions. For a detailed account of the next lemma and the related techniques, see for example Chapter 4.2. in [44].

Lemma 3.28. *(Contraction principle) Let X and Y be Hausdorff topological spaces and $f : X \to Y$ be a continuous function. Consider a proper rate function $I : X \to [0, \infty]$.*

1. *For each $y \in Y$, define $I_f := \inf\{I(x) \mid x \in X, y = f(x)\}$. Then, I_f is a proper rate function on Y, where the infimum over the empty set is taken as ∞.*
2. *If I controls the LDP associated to a sequence μ_n of measures on X, then I_f controls the LDP associated to the sequence $f_*\mu_n$ of measures on Y.*

Using this, we can now prove the following :

Corollary 3.29. *(of Theorem 3.1 and Theorem 3.2) Let μ be a probability measure on $GL(V)$ such that the Zariski closure G of the semigroup Γ generated by the support of μ is Zariski connected, semisimple and V is irreducible under G action. Then, the sequence μ_n, of laws of $\frac{1}{n} \log ||S_n||$ satisfies an LDP with a convex rate function I. Moreover, if μ possesses a finite exponential moment, then the sequence μ_n satisfies a full LDP with a proper convex rate function I. Finally, if μ is of strong exponential moment, then we have $I = \overline{\Lambda}^*$, where $\overline{\Lambda}^*$ is the convex conjugate of the limiting exponential moment generating function of μ_n's.*

Démonstration. Recall first that the Zariski closure G of Γ being a subgroup is not an additional hypothesis, the Zariski closure of a semigroup is a group, this had been implicit also in the previous parts. In the second place, it is easily observed that the result of the current corollary does not depend on the scalar product chosen on V and the associated operator norm (one other way to see this is to invoke the notion of exponential equivalence Definition 5.25 and Theorem 5.26), so we may as well suppose that V is equipped with the Euclidean norm given by Lemma 2.16. Therefore, if we denote by $\overline{\chi}$ the highest weight in $\mathfrak{a}^*$ of our irreducible inclusion representation, then for all $g \in G$, we have $\log ||g|| = \overline{\chi}(\kappa(g))$ and hence for all $n \geq 1$, we have $\frac{1}{n} \log ||g|| = \overline{\chi}(\frac{1}{n}\kappa(g))$.

In view of, respectively, Theorem 3.1 and Theorem 3.2, we can now apply the contraction principle, Lemma 3.28 with $X = \mathfrak{a}$, $Y = \mathbb{R}$, $f = \overline{\chi}$ and obtain the weak LDP in the case of general μ and the full LDP with the proper rate function $I_{\overline{\chi}}$ in the case of a probability measure μ with finite exponential moment. Refer to Section 3.4 for some indications as how to transfer directly the arguments of Section 3.3, for instance to get exponential tightness in order to have a full LDP.

Finally, under the corresponding assumption, the convexity of the rate function $I_{\overline{\chi}}$ is clear from the convexity of I and its expression as $I_{\overline{\chi}}(x) = \inf\{I(y) \mid y \in \mathfrak{a},\ \overline{\chi}(y) = x\}$. □

Remark 3.30. 1. *We note that for a non-irreducible G, the injection $G \overset{i}{\hookrightarrow} GL(V)$ can be seen as a rational representation of the connected semisimple group G and therefore, by complete reducibility of such representations, V decomposes as a direct sum of irreducible representations of G ; $V = \bigoplus_{i=1}^{s} V_i$ and $gV_i = V_i$ for all $g \in G$ and $i = 1, \ldots, s$. Therefore, vaguely, by this argument, the random walk S_n reduces to several (non-independent among them !), random walks S_n^i on connected semisimple G_i's, and the norm of S_n basically corresponds to the maximum of $\overline{\chi}_i(\kappa(S_n^i))$'s, where $\overline{\chi}_i$ denotes the highest weights. As a result, the existence of an LDP is not clear by this reasoning, although our Theorem 3.3 deals with this case.*

2. *We wish also to point out that, the problem mentioned in 1. is also one main difficulty when trying to generalise our results to connected reductive affine real algebraic groups.*

Before proceeding to prove Theorem 3.3, we start by recalling the setting and the notation (see also introduction of the section) : μ denotes a probability measure on $GL(V)$, $S \subset GL(V)$ denotes the support of μ, Γ is the semigroup generated by S and the subgroup G of $GL(V)$ denotes the Zariski closure of Γ. We suppose that G is a linear real algebraic group isomorphic with an isomorphism of algebraic groups to some $H \times T$, which we shall identify with G, and where H is a semisimple linear real algebraic group and T is a central subgroup of GL. We precise that all topological notions should be thought of bearing the adjective 'Zariski' unless otherwise noted.

In addition to our previous results, we essentially need only one more lemma to deal with the possible non-connectedness of H, the central component T does not cause any further difficulty. The proof of the following lemma is, in spirit, very similar to that of Lemma 3.5 and it basically says that for every $n \in \mathbb{N}$, after a bounded number of steps, the random walk S_n will come back to $(H^\circ \times T) \cap \Gamma$, without much changing the average logarithmic norm $\frac{1}{n} \log ||S_n||$, and this will happen with a probability that is bounded below by a strictly positive constant.

Lemma 3.31. *With the above notations, there exist a natural number $j_0 = j_0(G, \mu)$ and constants $d_5 = d_5(\mu, G) > 0$ and $C = C(\mu, G) > 0$, such that if for some $n_0 \in \mathbb{N}$, $x \in \mathbb{R}$, $K \geq 0$ and $\alpha \geq 0$, we have $\mathbb{P}(\log ||S_{n_0}|| \in]x - K, x + K[) \geq \alpha$, then there exists $n_1 \geq n_0 \geq n_1 - j_0$ such that*

$$\mathbb{P}(\log ||S_{n_1}|| \in]x - K - C, x + K + C[\quad \textit{and} \quad S_{n_1} \in (H^\circ \times T) \cap \Gamma) \geq \alpha . d_5$$

As it should be expected, the proof relies on the submultiplicativity of the operator norm $||.||$ and finiteness considerations, as in the proof of Lemma 3.5, related to the classical fact that the Zariski topology is a Noetherian topological structure, implying that the number of connected components of H is finite.

Démonstration. Start by noting that since the projection $p_1 : H \times T \longrightarrow H$ is a continuous morphism, $p_1(\Gamma)$ is dense in H. Recall also the classical fact that H° is a closed normal subgroup of finite index in H and denote $t := [H : H^\circ]$. Let $h_1, \ldots, h_t \in H$ be such that $H = \dot{\bigcup}_{i=1}^{t} h_i H^\circ$. Since H° is closed and of finite index, it follows that it is also open, and consequently, h_iH° is open for each $i = 1, \dots, t$. As a result, since $p_1(\Gamma)$ is dense in H, for each $i = 1, \dots, t$, it is intersects h_iH° on a dense subset, in particular, $p_1(\Gamma) \cap h_iH^\circ \neq \emptyset$. It follows that the restriction of $\pi_1 : H \twoheadrightarrow H/H^\circ$ to $p_1(\Gamma)$ is surjective and we can therefore find elements $\tilde{\gamma}_1, \dots, \tilde{\gamma}_t \in p_1(\Gamma)$ such that, for each $i = 1, \dots, t$, we have $\tilde{\gamma}_i.h_iH^\circ = H^\circ$.

Now, let, for $i = 1, \dots, t$, $\tilde{V}_i$ be a neighbourhood of $\tilde{\gamma}_i$ for the usual topology of H, such that for all $\tilde{\gamma}_i' \in \tilde{V}_i$, we have $\tilde{\gamma}_i' h_iH^\circ = H^\circ$. This is indeed possible since, H° is open and the usual topology is finer than the Zariski topology. And, for $i = 1, \dots, t$, set $V_i := p_1^{-1}(\tilde{V}_i)$ an open neighbourhood (for both topologies) of fixed elements $\gamma_i \in p_1^{-1}(\tilde{V}_i \cap p_1(\Gamma))$ of Γ.

Now, fix $m_1, \dots, m_t \in \mathbb{N}$ such that $\gamma_i \in \text{supp}(\mu^{*m_i})$ and put $j_0 = j_0(\mu) := \max\{m_1, \dots, m_t\}$. Set $\mu^{*m_i}(V_i) = \beta_i > 0$ and $\beta_0 = \min\limits_{i=1,\dots,t} \beta_i$. Finally put $C = C(G, \mu) := \max\limits_{i=1,\dots,t} \sup_{\gamma \in V_i} \exp(||\gamma|| \vee ||\gamma^{-1}||)$ and $d_5 = \frac{\beta_0}{t}$.

We can then write

$$\alpha \leq \mathbb{P}(\log ||S_{n_0}|| \in]x-K, x+K[) = \sum_{i=1}^{t} \mathbb{P}(\log ||S_{n_0}|| \in]x-K, x+K[\text{ and } S_{n_0} \in (h_iH^\circ \times T))$$

As a result, there exists $t_0 \in \{1, \dots, t\}$ such that

$$\mathbb{P}(\log ||S_{n_0}|| \in]x-K, x+K[\text{ and } S_{n_0} \in (h_{t_0}H^\circ \times T)) \geq \frac{\alpha}{t} \tag{3.28}$$

Notice also that, since for linear transformations $h, g \in GL(V)$, $||h^{-1}||^{-1}.||g|| \leq ||hg|| \leq ||h||.||g||$, for all $i = 1, \dots, t$ and $\gamma_i' \in V_i$, by the choice of C, one has that if for some real number x, we have $x - K \leq \log ||g|| \leq x + K$, then $x - K - C \leq \log ||\gamma_i' g|| \leq x + K + C$. Then it follows by this, (3.28) and the previous choices that, setting $n_1 = n_0 + m_{t_0} \leq n_0 + j_0$, using the independence of random walk increments, we have

$$\begin{aligned}
&\mathbb{P}(\log ||S_{n_1}|| \in]x-K-C, x+K+C[\text{ and } S_{n_1} \in (H^\circ \times T) \cap \Gamma) \geq \\
&\mathbb{P}(\log ||S_{n_0}|| \in]x-K, x+K[\text{ and } S_{n_0} \in (h_{t_0}H^\circ \times T) \text{ and } X_{n_1} \dots X_{n_0+1} \in V_{t_0}) = \\
&\mathbb{P}(\log ||S_{n_0}|| \in]x-K, x+K[\text{ and } S_{n_0} \in (h_{t_0}H^\circ \times T)).\mathbb{P}(S_{m_{t_0}} \in V_{t_0}) \geq \alpha d_5
\end{aligned}$$

establishing the claim of the lemma. □

With this lemma, the proof of Theorem 3.3 now boils down to the proofs of Theorem 3.1 and Theorem 3.2 as indicated below :

Démonstration. (Proof of Theorem 3.3) The idea of the proof is now to start, both in the proofs of existence and convexity of Theorem 3.1, by applying Lemma 3.31 and follow through these proofs working in $H^\circ \times T$ and making the necessary changes (to be indicated).

While transferring the proof of (existence of weak LDP and convexity of the rate function) Theorem 3.1 to our case, perhaps the way in which the least number of changes will be necessary is the following : we combine Lemma 3.31 and Lemma 3.5 and obtain a new lemma, Lemma 3.32 that we state below, and apply this whenever Lemma 3.5 is used in those proofs.

Denote the continuous projection $H \times T \longrightarrow H$ again by p_1. The proof of the following lemma goes mutatis mutandis as the concatenation of, respectively, the proofs of Lemma 3.31 and Lemma 3.5 using the above observed fact that the semigroup $p_1(\Gamma)$ is Zariski dense in H° (namely, in order to use the Abels-Margulis-Soifer result Theorem 2.24 in Lemma 3.5). To avoid unnecessary lengthy repetitions, we state it without a proof :

Lemma 3.32. *Let $0 < \epsilon < r(p_1(\Gamma))$, where $r(\Gamma) > 0$ is as given by Theorem 2.24. Then, there exist a natural number $j_1 = j_1(\epsilon, \mu)$, constants $d_6 = d_6(\epsilon, \mu) > 0$ and $C = C(\epsilon, \mu)$ such that if for some $n_0 \in \mathbb{N}$, $x \in \mathbb{R}$, $K \geq 0$ and $\beta \geq 0$, we have $\mathbb{P}(\log ||S_{n_0}|| \in]x - K, x + K[) \geq \beta$, then there exist $n_1 \in \mathbb{N}$ with $n_0 + i_0 \geq n_1 \geq n_0$ such that*

$$\mathbb{P}(\log ||S_{n_1}|| \in]x-K-C, x+K+C[,\ S_{n_1} \in (H^\circ \times T) \cap \Gamma \textit{ and } p_1(S_{n_1}) \textit{ is } (r,\epsilon)\textit{-loxodromic}) \geq \beta d_6$$

□

Now, replacing the use of Lemma 3.5 by this lemma, the Cartan projection κ by the logarithm of the operator norm $\log ||.||$ and $\mathfrak{a}$ by $\mathbb{R}$, the proof of Theorem 3.1 applies verbatim, establishing the existence of a weak LDP and the convexity of the rate function.

Finally, the remaining assertions of Theorem 3.3 follows in the same manner as in the proof of Theorem 3.2, in Section 3.3 : the proofs are the same replacing as before the Cartan projection κ by $\log ||.||$, $\mathfrak{a}$ and $\mathfrak{a}^*$ by $\mathbb{R}$, whenever they occur. □

Chapitre 4

JOINT SPECTRUM AND RATE FUNCTION

In Section 4.1, we introduce the notion of joint spectrum of a bounded set S in a connected semisimple linear real algebraic group G, a generalisation of the notions of joint spectral radii introduced in the introduction. It is easily generalised to the matrix algebras over local fields but we restrict our attention to the aforementioned case ; define joint Cartan and Jordan spectra, precise subsets of $\mathfrak{a}^+$ that bounds them and finally we analyse the characteristics of joint spectra.

Joint spectrum also appears in close connection to the large deviation principle considered in the previous chapters. In Section 4.2 we undertake a study of the rate functions appearing in Chapter 3 and put in evidence this relation with the joint spectrum.

4.1 Joint spectrum of a bounded subset S in G

Definition and main properties

We start with the definitions of the classical notions of joint/generalised spectral radii. The joint Cartan and Jordan spectra that we introduce, are generalisations of multidimensional extensions of these joint/generalised spectral radii.

Let $\mathcal{A}$ be an algebra endowed with a submultiplicative norm $||.||$. For a subset S of $\mathcal{A}$, denote by $S^n := \{x \in \mathcal{A} \mid x = x_1. \dots .x_n \text{ where } x_i \in \mathcal{A} \text{ for all } i \leq n\}$. Let S be a bounded subset of $\mathcal{A}$, the quantity $r(S) := \lim_{n\to\infty} (\sup\{||x|| \mid x \in S^n\})^{\frac{1}{n}} \in \mathbb{R}_+^*$ generalises the classical notion of spectral radius of an element and is called the joint spectral radius of S. It was introduced by Rota-Strang in [114]. The limit in question exists by submultiplicativity and it only depends on the equivalence class of the norm, i.e. for finite dimensional algebras, it doesn't depend on the norm. An important point is that unlike the spectral radius of a single element, one can consider other natural constants describing the asymptotic behaviour of the products of a set of elements. One of them is $r_{sub}(S) := \lim_{n\to\infty} (\inf\{||x|| \mid x \in S^n\})^{\frac{1}{n}} \in \mathbb{R}^+$ the joint spectral subradius of S.

Recall also that the 'generalised spectral radius' $\rho(S)$ of a set S was defined by Daubechies-Lagarias in [43] as $\rho(S) := \limsup_n (\sup\{\lambda_1(x) \mid x \in S^n\})^{\frac{1}{n}}$, where $\lambda_1(.)$ stands for the classical spectral radius. Joint spectral radii have been subjects of several conjectures and studies in recent years, we refer the reader to Rota-Strang [114], Daubechies-Lagarias [43], Berger-Wang [24], Bochi [25], Bousch-Mairesse [31], Breuillard [33] and [36], Protasov-Voynov [99] etc. (the list is largely incomplete). Among those, we single out the following notable fact, conjectured by Daubechies-Lagarias in [43], first proved by Berger-Wang [24] and then whose several independent proofs appeared (for instance Bochi [25], Breuillard[33], etc.).

Theorem 4.1 ([24],[25],[33],...). *For a bounded set of $\mathbb{C}$-linear tranformations S, we have $r(S) = \limsup_{n\to\infty} (\sup\{\lambda_1(x) \mid x \in S^n\})^{\frac{1}{n}}$.*

We will obtain an analogous result to this theorem, for the joint Cartan and Jordan spectra, for a reasonably general class of subsets S (see Proposition 4.21 and 3. of Definition 4.2).

We now proceed to define the joint spectra. As mentioned in the introduction, for convenience, we restrict ourselves to the case of a connected semisimple linear real algebraic group G : let S be a bounded subset of G and fix a non-principal ultrafilter $\mathcal{U} \subset \mathcal{P}(\mathbb{N})$. We wish to emphasise at this point that the use of an ultrafilter is to give a valid definition for an arbitrary bounded subset of G. We will shortly see that, in fact, for a large class of S, the joint spectra appear as Hausdorff limits and the use of an ultrafilter is irrelevant (see Theorem 4.4, Proposition 4.15, Proposition 4.13). Let now, as usual, $\mathfrak{a}$ denote the Cartan subalgebra of the Lie algebra $\mathfrak{g}$ of G endowed with a norm $||.||$ and $\mathfrak{a}^+$ be a fixed Weyl chamber in $\mathfrak{a}$. For each $n \in \mathbb{N}$, define the sets $K_n(S) := \frac{\log \kappa(S^n)}{n} \subset \mathfrak{a}^+$ and $\Lambda_n(S) := \frac{\log \lambda(S^n)}{n} \subset \mathfrak{a}^+$, where $\kappa(S^n) := \{\kappa(g) \mid g \in S^n\}$ and similarly for $\lambda(S^n)$.

Definition 4.2.

1. *The joint Cartan spectrum of S is defined to be the closed subset $K(S)$ of $\mathfrak{a}^+$ such that $\lim_{n\to\mathcal{U}} K_n(S) = K(S)$, where the convergence is with respect to the Hausdorff distance.*
2. *The joint Jordan spectrum of S is similarly defined as the closed set $\Lambda(S)$ satisfying $\lim_{n\to\mathcal{U}} \Lambda_n(S) = \Lambda(S)$*
3. *Whenever $\lim_{n\to\infty} K_n(S) = \lim_{n\to\infty} \Lambda_n(S)$ for a bounded subset S of G, we call this common limit the joint spectrum of S, and denote it by $J(S)$.*

Remark 4.3.

1. *It is easily seen that the joint Cartan and Jordan spectra do not depend on the chosen norm $||.||$ on $\mathfrak{a}$.*
2. *As indicated above, the joint Cartan and Jordan spectra are indeed generalisations of multidimensional extensions of, respectively, joint and generalised spectral radii : recall that when $G = SL(d, \mathbb{R})$ and $g \in G$, we have $\kappa(g) = (\log ||g||, \log \frac{||\Lambda^2 g||}{||g||}, \ldots, \log \frac{||\Lambda^d g||}{||\Lambda^{d-1} g||})$ and $\lambda(g) = (\log |\lambda_1(g)|, \ldots, \log |\lambda_d(g)|)$, where $\lambda_i(g)$'s are eigenvalues of g with $|\lambda_i(g)| \geq |\lambda_{i+1}(g)|$ for each $i = 1, \ldots, d-1$. Therefore, for instance, we immediately see that the (logarithm of) joint spectral radius of the set S, is just the upper bound of the first coordinate of the joint Cartan and Jordan spectra.*

3. *Respectively, the functions which associate the Cartan and Jordan joint spectrum to a bounded subset S of G are invariant under the conjugation action of G on its subsets, i.e. for all $g \in G$ and bounded $S \subset G$, we have $K(S) = K(gSg^{-1})$ and $\Lambda(S) = \Lambda(gSg^{-1})$. For the Jordan joint spectrum, this is immediate from the relation $\lambda(gxg^{-1}) = \lambda(g)$ for all $g, x \in G$ and for the Cartan joint spectrum, it easily follows from the uniform continuity of Cartan projection (Lemma 2.19).*

In the following theorem, we summarise our major findings about the joint Cartan and Jordan spectra. For its proof, we will break its statement into several propositions and prove as such.

Theorem 4.4. 1. *For a bounded subset S of G, generating a Zariski dense sub-semigroup in G, $K(S)$ is the (closed) Hausdorff limit of $K_n(S)$, i.e. $\lim_{n\to\infty} K_n(S) = K(S)$. In particular, it does not depend on the ultrafilter $\mathcal{U}$.*

2. *Similarly, $\lim_{n\to\infty} \Lambda_n(S) = \Lambda(S)$ for the Hausdorff distance if, either*
   a) S is a bounded subset of G containing the identity element $e \in G$, or
   b) S is a bounded subset of G generating a Zariski dense sub-semigroup in G.
3. *For an arbitrary bounded subset S of G, we have $\Lambda(S) \subseteq K(S)$, and for a subset S as in 1., $\Lambda(S) = K(S) =: J(S)$, and the joint spectrum $J(S)$ of S is a compact, convex set of non-empty interior in $\mathfrak{a}$.*
4. *Let S be as in 1. and μ be an arbitrary probability measure whose support is S. Let I be the rate function corresponding to μ-random walk, given by Theorem 3.1 and $D_I = \{x \in \mathfrak{a} \mid I(x) < \infty\}$ its effective support. Then, we have $\overline{D}_I = J(S)$ and $\mathring{D}_I = \mathring{J}(S)$.*

Before starting with the sequence of propositions to prove the previous theorem, we first wish to continue with the following passage in which we single out a region containing the joint spectra and make some additional observations.

A bounding region for the joint spectra

The aim of this passage is to locate a bounding region inside $\mathfrak{a}^+$ for the joint spectra adapting in a straightforward manner the classical notions of joint spectral radii to our setting.

Let k be a local field (e.g. $\mathbb{R}$), $\mathbb{G}$ a linear algebraic group over k, $\mathbb{G} \xrightarrow{\rho} GL_d$ a finite dimensional representation of $\mathbb{G}$, and G the group of k points of $\mathbb{G}$. For a bounded subset S of G, define the joint spectral (sub)radius of S with respect to the representation ρ, $r_\rho(S)$ ($r_{sub,\rho}(S)$) to be the joint spectral (sub)radius of $\rho(S) \subset GL(d, \mathrm{k})$. Note that this doesn't depend on the chosen norm on k^d (i.e. the associated operator norm).

Let G be a connected semisimple linear real algebraic group and $\rho : G \longrightarrow GL(V)$, an irreducible rational representation of highest weight χ_ρ of G. For a bounded set S of G, define the ρ-joint spectral half space of S on $\mathfrak{a}$ as

$$H_\rho(S) := \{x \in \mathfrak{a} \mid \chi_\rho(\exp(x)) \leq r_\rho(S)\} = \{x \in \mathfrak{a} \mid \overline{\chi}_\rho(x) \leq \log r_\rho(S)\} \qquad (4.1)$$

Note at this point that for such a linear algebraic group G, the image of G by the rational representation (ρ, V) is contained in $SL(V)$ so that for all $g \in G$, we have $||\rho(g)|| \geq 1$. As a consequence, $r_\rho(S) \geq r_{sub,\rho}(S) \geq 1$ and $H_\rho(S) \cap \mathfrak{a}^+ \neq \emptyset$ (nevertheless, note that this intersection may well be the set $\{0\} \subset \mathfrak{a}^+$).

Let $R_{ir}(G)$ denote a set of representatives of isomorphism classes of all irreducible rational representations of G. With these notations, we can state the following proposition locating a region for the joint spectra (see also Fig. 4.1).

Proposition 4.5. *Let S be a bounded subset of a semisimple linear algebraic group G as above. Then,*

1. $K(S) \subseteq \bigcap_{\rho \in R_{ir}(G)} H_\rho(S) \cap \mathfrak{a}^+$ *and* $\Lambda(S) \subseteq \bigcap_{\rho \in R_{ir}(G)} H_\rho(S) \cap \mathfrak{a}^+$.
2. $K(S) \cap \partial H_\rho(S) \neq \emptyset$ *for all* $\rho \in R_{ir}(G)$, *and*
   $\Lambda(S) \cap \partial H_\rho(S) \neq \emptyset$ *for all* $\rho \in R_{ir}(G)$.
3. *If the identity element e is in the semigroup Γ generated by S, then* $0 \in \Lambda(S) \cap K(S)$.

Démonstration.
1. It follows from Proposition 4.21 that for all bounded $S \subset G$, we have $\Lambda(S) \subseteq K(S)$. Therefore, it suffices to show the first statement. Recall that the joint spectral radius of $\rho(S)$ in $GL(V)$ does not depend on the norm chosen on V. In particular, choosing on V the norm given by Lemma 2.16, for all $g \in G$, we have $||\rho(g)|| = \exp \overline{\chi}(\kappa(g))$, where χ is the highest weight of (ρ, V). From this relation and the definition of ρ-joint spectral half space $H_\rho(S)$ of S, the result is immediately seen to hold.
2. By the same Proposition 4.21, it suffices to show the second statement. This is also readily seen using, as above, Lemma 2.16 and more importantly Theorem 4.1 together with the fact that $\Lambda(S)$ is closed in $\mathfrak{a}$.
3. By hypothesis, there exists, $n_0 \in \mathbb{N}$ such that $e \in S^{n_0}$. Fix $g \in S$ and put $\tilde{C} = \max_{i=1,\ldots,n_0-1} ||\kappa(g^i)|| < \infty$. Since for all $k \geq 1$ and $0 \leq i \leq n_0 - 1$; $g^i = e^k g_i \in S^{kn_0+i}$, we have that for all such k and i, $\Lambda_{kn_0+i}(S)$ and $K_{kn_0+i}(S)$ intersect $B(0, \frac{\tilde{C}}{kn_0})$, the $||.||$-ball of radius $\frac{\tilde{C}}{kn_0}$ and center $0 \in \mathfrak{a}$. It follows that along any subsequence $n_k \underset{k\to\infty}{\longrightarrow} \infty$, $\Lambda_{n_k}(S)$ and $K_{n_k}(S)$ accumulates at 0 ; whence the result.

□

Remark 4.6. *Recall that by the so-called highest weight theorem, the irreducible representations of G stand in one-to-one correspondence (through their highest weights) with the set of dominant weights, i.e. in the level of Lie algebra, with the usual identification of $\mathfrak{a}$ and $\mathfrak{a}^*$, those weights that belong to the Weyl chamber $\mathfrak{a}^+$. Recall also that the distinguished irreducible representations $(\rho_i)_{i=1,\ldots,d}$ given by Lemma 2.15 are such that, in the Lie algebra, their highest weights $\overline{\chi}_i$ are integral multiples of the fundamental weights $\overline{\omega}_i$. Moreover, the dominant weights stand in one-to-one correspondence with the integral linear combinations of the fundamental weights. As a result, endowing $\mathfrak{a}$ with the restriction of the Killing form of $\mathfrak{g}$, we see that for each positive rational direction (in terms of the dual basis to $\overline{\chi}_i$'s), we have an affine hyperplane*

$\partial H_\rho(S)$ with $\rho \in R_{ir}(G)$, orthogonal to that direction, and such that these affine hyperplanes $\partial H_\rho(S)$ appear as bounds to the sets $K(S)$ and $\Lambda(S)$ and are 'reached' by those. Moreover, the distinguished spectral hyperplanes $\partial H_{\rho_i}(S)$ corresponding to the distinguished representations ρ_i appear as the extremal ones, in the sense that they intersect orthogonally the coordinates axes in $\mathfrak{a}$ defined by the same basis.

In the case of a bounded set S generating a Zariski dense sub-semigroup Γ containing identity in $SL(3,\mathbb{R})$, we summarise the observations of this remark, the previous proposition and Theorem 4.4 on a picture as in the following Fig. 4.1 (see also Lemma 4.12).

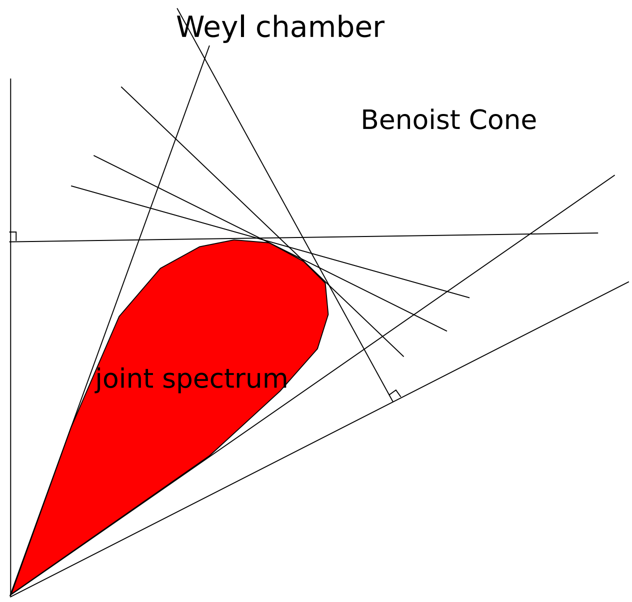

FIGURE 4.1 – The orthogonals to the walls of the Weyl chamber correspond to $\partial H_{\rho_i}(S)$'s, whereas the medium lines correspond to $\partial H_\rho(S)$ for ρ's in $\overset{\circ}{R}_{ir}(G)$ (see the following remark). For a discussion of Benoist cone, see Section 5.2.

Remark 4.7. *Let us also mention at this point a remark about the shape of the joint spectrum of an S as in Theorem 4.4 : it follows by 3. of that theorem that if the subset S in question is symmetric, i.e. $S = S^{-1} := \{g^{-1} \mid g \in S\}$, then clearly, $0 \in J(S)$ and $J(S)$ is invariant under opposition involution ι of $\mathfrak{a}^+$. We recall that for $x \in \mathfrak{a}^+$, $\iota(x)$ is the unique element of $\mathfrak{a}^+$ conjugated to $-x$ by the action of Weyl group of the root system of G.*

In passing, let us also note a simple observation (Proposition 4.9) on the joint spectral radius of a bounded subset S of G in different irreducible representations. Suppose for convenience that we have $\lim_{n\to\infty} K_n(S) = K(S)$. As usual, endow $\mathfrak{a}$ with the restriction of the Killing form of $\mathfrak{g}$, using this, identify $\mathfrak{a}$ and $\mathfrak{a}^*$, and denote by the same notation their identified elements. Let ρ be an irreducible rational representation of G and $\overline{\chi}_\rho$ denote its highest weight in $\mathfrak{a} \simeq \mathfrak{a}^*$. Then, by definition of ρ-joint spectral radius of S and using Lemma 2.16, we have

$$\begin{aligned}\log r_\rho(S) &= \lim_{n\to\infty}\sup_{g\in S^n}\frac{1}{n}\log\|\rho(g)\| = \lim_{n\to\infty}\sup_{g\in S^n}\frac{1}{n} < \overline{\chi}_\rho, \kappa(g) >= \lim_{n\to\infty}\sup_{g\in S^n} < \overline{\chi}_\rho, \frac{1}{n}\kappa(g) > \\ &= \lim_{n\to\infty}\sup_{x\in K_n(S)} < \overline{\chi}_\rho, x >= \sup_{x\in K(S)} < \overline{\chi}_\rho, x >\end{aligned} \tag{4.2}$$

where we use $K_n(S) \underset{n\to\infty}{\longrightarrow} K(S)$ in the last equality.

Remark 4.8. *Note at this point that in case the joint spectrum exists as a limit, i.e.* $\lim_{n\to\infty} K_n(S) = \lim_{n\to\infty}\Lambda_n(S)$ *(e.g. S generates a Zariski dense semigroup in G), the second statement of Proposition 4.5 is immediately deduced by the simple (4.2) (carrying out also the same calculation for the generalised spectral radius), and in turn proves the Berger-Wang Theorem 4.1 in our setting.*

Now, again for convenience, we suppose G to be simple and for $i = 1, \ldots, d$, let σ_i denote an irreducible rational representation of G with highest weight $\overline{\omega}_i \in \mathfrak{a}^*$, where $\overline{\omega}_i$'s are the fundamental weights of $\mathfrak{g}$. We recall that, by the classification of root systems of G, the angles between $\overline{\omega}_i$'s are acute. For $i = 1, \ldots, d$, denote the σ_i-joint spectral radius of S by $r_i(S)$. Then, we have the following estimations on the ρ-joint spectral radius of S :

Proposition 4.9. *In the setting of the previous paragraph, the ρ-joint spectral radius $r_\rho(S)$ of S satisfies*

$$\max_{i=1,\ldots,d} r_i(S)^{\left(\min_{j=1,\ldots,d}\frac{<\overline{\chi}_\rho,\overline{\omega}_j>}{<\overline{\omega}_i,\overline{\omega}_j>}\right)} \leq r_\rho(S) \leq \min_{i=1,\ldots,d} r_i(S)^{\left(\max_{j=1,\ldots,d}\frac{<\overline{\chi}_\rho,\overline{\omega}_j>}{<\overline{\omega}_i,\overline{\omega}_j>}\right)}$$

In view of Lemma 2.16, Remark 4.6 and (4.2), the proof relies on elementary Euclidean geometry considerations.

Démonstration. Let us start by proving the lower bound. For $i = 1, \ldots, d$, let x_i denote an element of $K(S) \cap \partial H_{\sigma_i}(S)$, which is non-empty by 2. of Proposition 4.5, i.e. we have $\overline{\omega}_i(x_i) = \log r_i(S)$. For $i, j = 1, \ldots, d$, let $x_{i,j}$ be the unique point of intersection of the line $\mathbb{R}\overline{\omega}_j$ with the hyperplane $\partial H_{\sigma_i}(S) = \{x \in \mathfrak{a} \mid \ < \overline{\omega}_i, x >= \log r_i(S)\}$. Correspondingly, let $\alpha_{i,j} \in \mathbb{R}$ be such that $x_{i,j} = \alpha_{i,j}\overline{\omega}_j$. $x_{i,j}$ and $\alpha_{i,j}$ are well-defined (since the angles between $\overline{\omega}_i$ and $\overline{\omega}_j$'s are acute) and since by construction $\overline{\omega}_i(\alpha_{i,j}\overline{\omega}_j) = \log r_i(S)$, we have

$$\alpha_{i,j} = \frac{\log r_i(S)}{< \overline{\omega}_i, \overline{\omega}_j >} \tag{4.3}$$

As a result, for each $i = 1, \ldots, d$, the simplex T_i which is the intersection of the hyperplane ∂H_{σ_i} with $\mathfrak{a}^+$, writes as $T_i = \{\sum_{j=1}^d \beta_j x_{i,j} \mid \sum_{j=1}^d \beta_j = 1\, ,\, \beta_j \geq 0\}$ and we have $x_i \in T_i$. Now, by (4.2), for each $i = 1, \ldots, d$, we have $\log r_\rho(S) = \sup_{x\in K(S)}\overline{\chi}_\rho(x) \geq \overline{\chi}_\rho(x_i) \geq \min_{x\in T_i}\overline{\chi}_\rho(x)$ and therefore

$$\log r_\rho(S) \geq \max_{i=1,\ldots,d}\min_{x\in T_i}\overline{\chi}_\rho(x) \tag{4.4}$$

But for each $i = 1, \ldots, d$, one has

$$\min_{x \in T_i} \overline{\chi}_\rho(x) = \min_{\substack{\sum \beta_j = 1 \\ \beta_j \geq 0}} \overline{\chi}_\rho(\sum_{j=1}^{d} \beta_j x_{i,j}) = \min_{\substack{\sum \beta_j = 1 \\ \beta_j \geq 0}} \sum_{j=1}^{d} \beta_j \overline{\chi}_\rho(x_{i,j}) = \min_{j=1,\ldots,d} \overline{\chi}_\rho(x_{i,j}) \qquad (4.5)$$

Now, we see that the desired lower bound follows combining (4.3), (4.4) and (4.5).

To see the upper bound, for each $j = 1, \ldots, d$, let $x_{\rho,j} \in \mathfrak{a}$ and $\alpha_{\rho,j} \in \mathbb{R}$ be defined by the property that $x_{\rho,j} := \alpha_{\rho,j}\overline{\omega}_j$ is the point of intersection of the line $\mathbb{R}\overline{\omega}_j$ with the hyperplane orthogonal to $\overline{\chi}_\rho$ and containing x_ρ, where x_ρ denotes an element of $K(S) \cap \partial H_\rho(S)$ (so that $\overline{\chi}_\rho(x_\rho) = \log r_\rho(S)$). By the previous remark on acute angles θ_{ij} between $\overline{\omega}_i$ and $\overline{\omega}_j$'s and the fact that $\overline{\chi}_\rho$ writes as $\overline{\chi}_\rho = \sum_{i=1}^{d} n_i \overline{\omega}_i$, with $n_i \in \mathbb{N}$ for each $j = 1, \ldots, d$, $\alpha_{\rho,j}$ is well-defined and by defining property of $\alpha_{\rho,j}$, we have $\log r_\rho(S) = \overline{\chi}_\rho(\alpha_{\rho,j}\overline{\omega}_j)$, so that $\alpha_{\rho,j} = \frac{\log r_\rho(S)}{<\overline{\chi}_\rho, \overline{\omega}_j>}$. Finally, reasoning as above with the simplex defined by $x_{\rho,j}$'s, one readily sees that for each $i = 1, \ldots, d$, we have

$$\overline{\omega}_i(x_\rho) \geq \min_{j=1,\ldots,d} < \overline{\omega}_i, x_{\rho,j} > \qquad (4.6)$$

Now, since by (4.2), for each $i = 1, \ldots, d$, $\log r_i(S) \geq \overline{\omega}_i(x_\rho)$, combining the expression of $\alpha_{\rho,j}$ and (4.6), we get the desired upper bound.

□

Remark 4.10. *1. Combining the estimates of the previous proposition by complete reducibility, one readily observes the joint spectral radius dichotomy for connected simple linear real algebraic groups G : for a bounded subset S of G, either there exists a rational representation ρ with $r_\rho(S) > 1$, in which case for all rational ρ', $r_{\rho'}(S) > 1$, or for all rational representation ρ, $r_\rho(S) = 1$.*

2. *Note that in view of the obvious extension of the previous proposition for semisimple groups, the exponential moment condition defined in Section 3.3 can be equivalently expressed as asking that there exists $(\rho, V) \in R_{ir}(G)$ whose highest weight $\overline{\chi}_\rho$ belongs to the interior of the Weyl chamber $\mathfrak{a}^+$, i.e. $\overline{\chi}_\rho = \sum_{i=1}^{d} n_i \overline{\omega}_i$, with $n_i > 0$ for each $i = 1, \ldots, d$, there exists a $c > 0$ with $\int (||\rho(g)|| \vee ||\rho(g)^{-1}||)^c \mu(dg) < \infty$.*
3. *We wish to mention that Breuillard's Proposition 3.4 in [36] is in the same lines with this geometric corollary.*
4. *See also the 'directional spectral radius function' $r(.)$ that we define in Section 5.2.*

Remark 4.11 (Subspectral half spaces of S)**.** *One can also define the joint subspectral half spaces $H_{sub,\rho}(S) := \{x \in \mathfrak{a} \mid \overline{\chi}_\rho(x) \geq \log r_{sub,\rho}(S)\}$, to bound the joint Cartan spectrum $K(S)$ of S 'from below', and hence, pinpoint a more precise region for the joint spectrum. Notice upon this definition that we have the corresponding statement as in 2. of Proposition 4.5, namely $\partial H_{sub,\rho}(S) \cap K(S) \neq \emptyset$, so that in case $K(S)$ is convex, one can argue as in the last corollary to obtain analogous upper and lower bounds for $r_{sub,\rho}(S)$ for an irreducible representation ρ of G. Notice, moreover, that the subspectral half spaces will give meaningful information about the location of $K(S)$ only if $r_{sub,\rho}(S) > 1$, which is, for example, the case for the proximal irreducible representations ρ of Schottky semigroups (see Proposition 5.38).*

In this paragraph and the following lemma, we will be concerned with the question of when for a $\rho \in R_{ir}(G)$, we have $r_\rho(S) > 1$. We first point out that for $(\rho, V) \in R_{ir}(G)$, the condition $r_\rho(S) > 1$ may not be satisfied (in which case $K(S) = \{0\}$) even though S generates an unbounded sub-semigroup Γ of G. An emblematic case for this situation is the unipotent subgroups : let $G = SL(2, \mathbb{R})$, $S = \{\begin{pmatrix} 1 & t \\ & 1 \end{pmatrix} \mid t \in B\}$, where B is any bounded subset of $\mathbb{R}$ and $\rho = id$. On the other hand, this condition is satisfied if Γ is unbounded and the restriction on Γ of the ρ-action of G on V is irreducible (this basically follows from Prasad [97], for an elementary treatment, see for example Protasov-Voynov's Proposition 2 in [99]). In the following lemma, using the techniques of Chapter 2, we give a short proof of the fact that this condition is satisfied for all $\rho \in R_{ir}(G)$, in the particular case of interest to us, namely when S generates a Zariski dense sub-semigroup Γ in G. It therefore gives information about the positions of affine hyperplanes $\partial H_\rho(S)$ for such S.

Lemma 4.12. *Let $S \subset \Gamma \subseteq G$ be as above and (ρ, V) an irreducible representation of G. Then, we have $r_\rho(S) > 1$. Equivalently, $0 \notin \partial H_\rho(S) \subset \mathfrak{a}$ and $\mathfrak{a}^{++} \cap H_\rho(S) \neq \emptyset$.*

Démonstration. By 2. of Remark 4.10, it suffices to show the claim for each of the distinguished representation ρ_i $i = 1, \ldots, d$. This basically follows from the existence of loxodromic elements in Γ (that we discussed previously to the statement of Theorem 2.24 and in Remark 2.26). Indeed, let $g \in \Gamma$ be a loxodromic element. Then, for each $i = 1, \ldots, d$ $\rho_i(g) \in SL(V)$ is a proximal element of $SL(V)$, and hence, denoting by $\lambda(.)_1$ the spectral radius of a linear transformation, we have $\lambda_1(\rho_i(g)) > 1$. Let now $n_0 \in \mathbb{N}$ be such that $g \in S^{n_0}$. Then, for all $k \geq 1$, we have $\sup\{||\rho_i(x)|| \mid x \in S^{kn_0}\} \geq ||\rho_i(g^k)|| \geq \lambda_1(\rho_i(g^k)) = \lambda_1(\rho_i(g))^k$, whence, $r_{\rho_i}(S) \geq \lambda_1(\rho_i(g))^{\frac{1}{n_0}} > 1$. The corresponding results for $\partial H_{\rho_i}(S)$ and $H_{\rho_i}(S)$ follows by definition of the joint spectral half spaces (4.1). □

Properties of the joint spectra

We now start to prove Theorem 4.4 in a series of propositions. For the following first proposition, we note that its proof is 'strictly included' in the existence part of the proof of Theorem 3.1. Here, although we use the same tools, we prefer to give a shorter proof which does not contain random walk considerations. We would like to stress out once again that, as a consequence, in defining the joint Cartan (Jordan) spectrum for such S as in Proposition 4.13 (Proposition 4.15), the use of the ultrafilter $\mathcal{U}$ is not needed and we can define the joint spectra as the closed Hausdorff limit sets.

Proposition 4.13. *For a bounded subset S of G, generating a Zariski dense semigroup in G, we have $\lim_{n\to\infty} K_n(S) = K(S)$ for the Hausdorff distance.*

To simplify the proofs of the assertions on Hausdorff convergence in this and the next proposition, we will make use of the following elementary lemma whose proof we provide for the sake of completeness.

Lemma 4.14. *Let (X, d) be a compact metric space and $(K_n)_{n\in\mathbb{N}}$ a sequence of subsets of X. Then, there exists a compact subset K of X such that $d_H(K_n, K) \underset{n\to\infty}{\longrightarrow} 0$ if and* *only if* $\lim_{n\to\infty} \limsup_{m\to\infty} \sup_{x\in K_n} d(x,K_m) = 0$. *Moreover in this last limit, one can switch the places of* $\limsup$ *and* $\sup$ *without altering the statement.*

Note that in case K_n's are singletons, the lemma is fairly obvious by sequential compactness of X. We will imitate the proof of this fact using thinner and thinner sequences of finite covers of X.

Démonstration. The 'only if' direction being clear, we only show the other direction. For this, let for each $i \in \mathbb{N}$, $(O_{i,n})_{n\in I_i}$ be a finite cover of X, where I_i is a finite index set in $\mathbb{N}$ and $O_{i,n}$'s are open balls of radius $\frac{1}{i}$ in X. Define CS, the set of compatible sequences with our sequence of covers, as $CS := \{(i,n_i)_{i\geq 1} \mid \text{for all } i,k \in \mathbb{N}, n_i \in I_i \text{ and } O_{i,n_i} \cap O_{i+k,n_{i+k}} \neq \emptyset\}$. Obviously $CS \neq \emptyset$ and $\bigcup_{(i,n_i)_{i\geq 1}\in CS} \bigcap_{i\geq 1} O_{i,n_i} = X$. Furthermore, for every $i \in \mathbb{N}$, define the index set $J_i := \{n \in I_i \mid \#\{m \in \mathbb{N} \mid O_{i,n} \cap K_m \neq \emptyset\} = \infty\}$ and note that since for each $i \geq 1$, $|I_i| < \infty$, it follows that for each $i \geq 1$, $J_i \neq \emptyset$. Now, set $\tilde{CS} := \{(i,n_i)_{i\geq 1} \in CS \mid \text{for all } i \geq 1, n_i \in J_i\}$. Observe that the hypothesis implies that for all m large enough, say $m \geq M_0 \in \mathbb{N}$, $K_m \neq \emptyset$. Therefore, choosing $x_m \in K_m$, by sequential compactness of X, there exists $x \in X$ such that $x_{m_l} \underset{l\to\infty}{\longrightarrow} x$ for some sequence m_l. As a result, considering the compatible sequence of covers for x, we see that $\tilde{CS} \neq \emptyset$. We now claim that the compact set $K = \overline{\bigcup_{(i,n_i)_{i\geq 1}\in \tilde{CS}} \bigcap_{i\geq 1} O_{i,n_i}}$ satisfies $d_H(K_n,K) \underset{n\to\infty}{\longrightarrow} 0$ (in fact, one does not need to take the completion in the definition of K, the union is automatically closed).

One first observes that for all $i \geq 1$, $n_i \in J_i$ and $\delta > 0$, we have, for all m large enough,

$$(O_{i,n_i})^\delta \cap K_m \neq \emptyset \tag{4.7}$$

where for a set O, O^δ denotes the open δ-blow up of O, i.e. $O^\delta := \{x \in X \mid d(x,O) < \delta\}$. By construction of K, this directly implies that for every $x \in K$, we have $d(x,K_n) \underset{n\to\infty}{\longrightarrow} 0$. To see (4.7), supposing the contrary, for some $i \geq 1$, $n_i \in J_i$ and $\delta > 0$, there exists a sequence $m_k \to \infty$ such that for each $k \geq 1$, $K_{m_k} \cap (O_{i,n_i})^\delta = \emptyset$. Moreover, by definition of J_i, there exists a sequence $p_l \to \infty$ such that for each $l \geq 1$, $K_{p_l} \cap O_{i,n_i} \neq \emptyset$. As a result, $\limsup_{n\to\infty} \limsup_{m\to\infty} \sup_{x\in K_n} d(x,K_m) \geq \limsup_{l\to\infty} \limsup_{k\to\infty} \sup_{x\in K_{p_l}} d(x,K_{m_k}) \geq \delta > 0$ contradicting the hypothesis.

On the other hand, suppose by way of contradiction that there exist a sequence $(n_k)_{k\geq 1}$, elements $x_k \in K_{n_k}$ such that for every $k \geq 1$, we have $d(x_k,K) \geq \epsilon$ for some $\epsilon > 0$. Then it follows by sequential compactness of X that there exists $x \in X$ such that, up to passing to a subsequence of x_k, $x_k \underset{k\to\infty}{\longrightarrow} x$ and $d(x,K) \geq \epsilon$. But then, for every $i \geq 1$, choose $n_i \in I_i$ so that $x \in O_{i,n_i}$. It follows that for each $i \geq 1$, $n_i \in J_i$ so that $(i,n_i)_{i\geq 1} \in \tilde{CS}$ and thus $x \in K$, yielding a contradiction to $d(x,K) \geq \epsilon$.

The last assertion of the lemma also follows easily by compactness of X, we omit the details. $\square$

Proof of Proposition 4.13. Since S is a bounded subset of G, by definition of $K_n(S)$'s, there exists a compact subset C of $\mathfrak{a}$ such that $K_n(S) \subset C$ for all $n \in \mathbb{N}$. As a result, to prove the claim of the proposition, by Lemma 4.14, it suffices to prove

$$\lim_{n\to\infty} \limsup_{m\to\infty} \sup_{x\in K_n(S)} d(x, K_m(S)) = 0$$

To do this, let $\delta > 0$ be given. Let then, $r = r(\Gamma)$ be as given by Theorem 2.24 and choose $0 < \epsilon < r$. Let $F = F_{(r,\epsilon)}$ be the finite subset of Γ given by Theorem 2.24. Let M_F be the compact subset of $\mathfrak{a}$ given by Lemma 2.19 for the compact $F \subset \Gamma$ and $K_{(r,\epsilon)}$ be the compact subset of $\mathfrak{a}$ given by Corollary 3.14. Finally, for each $f \in F$, fix $n_f \in \mathbb{N}$ such that $f \in S^{n_f}$, put $i_0 := \max_{f\in F} n_f$, and set $\tilde{C} := \max_{x\in C} ||x||$.

Now, let $x \in K_n(S)$ for some $n \in \mathbb{N}$ to be specified later. Then, by definition, $x = \frac{1}{n}\kappa(g)$ for some $g \in S^n$. Let, by Theorem 2.24, $f \in F$ be such that gf is (r,ϵ)-loxodromic, so that by Lemma 2.19, we have $\kappa(gf) \in \kappa(g) + M_F$. Observe moreover that, by Corollary 3.14, for each $k \geq 1$, we have

$$\kappa((gf)^k) \in k(\kappa(gf) + K_{(r,\epsilon)}) \tag{4.8}$$

Putting $n_1 = n + n_f \leq n + i_0$, (4.8) implies that for each $k \geq 1$,

$$\frac{1}{n_1 k}\kappa((gf)^k) \in \frac{n}{n_1}x + \frac{M_F}{n_1} + \frac{K_{(r,\epsilon)}}{n_1}$$

From this, a straightforward computation shows that if $n \geq N_\delta := \frac{2}{\delta}(diam(M_F) + diam(K_{(r,\epsilon)}) + \tilde{C} i_0)$, we get that for each $k \geq 1$, we have $d(x, \frac{1}{n_1 k}\kappa((gf)^k)) \leq \frac{\delta}{2}$. Finally, since $(gf)^k \in S^{kn_1}$, this implies that for all $n \geq N_\delta$ and $k \geq 1$; $d(x, K_{n_1 k}(S)) < \frac{\delta}{2}$. From this periodicity, one now concludes using uniform continuity of Cartan projection, Lemma 2.19, in a similar fashion as in the proof of Lemma 3.17. $\square$

The next proposition is analogous to the previous one, but for the joint Jordan spectrum. While its first condition is easily seen to be sufficient by the fact that for all $k \geq 1$ and $g \in G$, we have $\lambda(g^k) = \lambda(g)^k$, the verification of sufficiency of the second condition is slightly more involved and it heavily uses the techniques of Chapter 2 and Chapter 3. The idea is, in the absence of uniform continuity of the Jordan projection, to use the spectral radius formula to approximate the Jordan projection by Cartan projection and then using the uniform continuity of Cartan projections (Lemma 2.19) and AMS finiteness result (Theorem 2.24), to land on closeby loxodromic elements, and then finally, to obtain a certain stability after sufficiently many iterates (using in part Corollary 3.14) and sufficient dispersion of the attracting directions (see Lemma 4.20).

Proposition 4.15. $\lim_{n\to\infty} \Lambda_n(S) = \Lambda(S)$ *for the Hausdorff distance, if either*

1. *S is a bounded subset of G containing the identity element $e \in G$, or*
2. *S is a bounded subset of G, such that the semigroup Γ generated by S is Zariski dense in G.*

Remark 4.16. *We note that the main difference between the proofs of Proposition 4.13 and 2. of Proposition 4.15 - which is the use, in the latter, of spectral radius formula - which forces us to take an arbitrary power of elements, is responsible for our techniques to fail to yield an LDP as in Theorem 3.1, but for the Jordan projections. As mentioned previously to Proposition 4.15, we are in turn led to use this spectral radius formula of non-uniform nature due to the lack of uniform continuity of Jordan projections (cf. Lemma 2.19).*

Example 4.17. *Here we present an example of a set S such that the sequence $\Lambda_n(S)$ does not converge. This example was suggested to the author by Emmanuel Breuillard. Let $a>1$ and set $\alpha := \begin{pmatrix} a & \\ & a^{-1}\end{pmatrix}$, $u := \begin{pmatrix} & 1 \\ -1 & \end{pmatrix}$ and take the subset $S := \{\alpha u, u\}$ of $G = \mathrm{SL}(2,\mathbb{R})$. Let $\lambda : G \to [0,\infty[$ denote the Jordan projection, associating to an element the logarithm of its spectral radius and $\Lambda_n(S)$ be defined as before. Then, we claim that $\Lambda_{2n}(S) \underset{n\to\infty}{\longrightarrow} [0, \frac{\log a}{2}]$ and $\Lambda_{2n+1}(S) = \{0\}$ for each $n \geq 0$. Indeed, from the relations $(\alpha u)^2 = -id$ and $u^2 = -id$, it follows that we have a normal form for the elements of the semigroup generated by S, i.e. for each $n \geq 1$, each element of S^n can be written as $\alpha^k u^\epsilon$ or $-\alpha^k u^\epsilon$ for some $k \in \mathbb{Z}$ and with $\epsilon \in \{0,1\}$. Moreover, one sees that if n is even, then $\epsilon = 0$; and if n is odd, then $\epsilon = 1$. Since for each $k \geq 1$, $\lambda(\alpha^k) = \lambda(-\alpha^k) = k\log a$ and $\lambda(\alpha^k u) = \lambda(-\alpha^k u) = 0$, the claim follows by remarking that for each $n \geq 1$, and $n \geq k \geq 1$, we have $\{\alpha^k, -\alpha^k\} \cap S^{2n} \neq \emptyset$.*

Remark 4.18. *It is not hard to see that if there exists $n_0 \in \mathbb{N}^*$ such that S^{n_0} contains the identity element (or a central elliptic element as in Example 4.17), then the sequence $\Lambda_n(S)$ admits at most n_0 distinct Hausdorff limit points (i.e. subsets of $\mathfrak{a}^+$). Moreover, for each $r \in \mathbb{N}$, the sequence $\Lambda_{kn_0+r}(S)$ converges. (The first statement follows from the second and the proof of the second statement is similar to that of 1. of Proposition 4.15.)*

Remark 4.19.

1. *From the large deviations perspective, it is an easy observation (see also Remark 4.29) that when S is a finite set, the Hausdorff convergences of $K_n(S)$ and $\Lambda_n(S)$ are necessary conditions for an LDP to hold for, respectively, for Cartan and Jordan projections of random products of elements of S. From this point of view, the Proposition 4.15 is notable given that we do not know whether the Jordan projections of a random walk as in Theorem 3.1 satisfy an LDP.*
2. *In view of 1. of this remark and Example 4.17, for any probability measure μ supported on the set $S \subset \mathrm{SL}(2,\mathbb{R})$, denoting by S_n the n^{th} step of the μ-random walk in $\mathrm{SL}(2,\mathbb{R})$, the sequence of random variables $\frac{1}{n}\lambda(S_n)$ does not satisfy a principle of large deviations.*

To prove the previous proposition, we will need the following technical lemma which says that there exists a finite set in Γ such that given a loxodromic element $g \in G$ and an arbitrary $h \in G$, after taking a large power g^s of g, we can find an element from the finite set such that after multiplying g^s on the left by that element, the second left multiplication by h will preserve the loxodromy and this in a somewhat uniform manner. Two main points of its proof are : first the fact that for g loxodromic, g^s is (r,ϵ_s) loxodromic for some r and $\epsilon_s \underset{s\to\infty}{\longrightarrow} 0$, and the second is Lemma 3.18.

Lemma 4.20. *There exists a finite set M in Γ with the property that given a loxodromic element $g \in G$, there exist $\gamma_1, \ldots, \gamma_{\hat{d}}$ in M such that for all $h \in G$ and $L \in \mathbb{N}$, there exists r_1 such that for each $\hat{r}_1 \geq \hat{\epsilon} > 0$, for all $s \in \mathbb{N}$ large enough and $l \leq L$, there exists $i_l \in \{1, \ldots, \hat{d}\}$ such that $h^l \gamma_{i_l} g^s$ is $(\hat{r}_1, \hat{\epsilon})$-loxodromic.*

Démonstration. Let $(\rho_i, V_i)_{i=1,\ldots,d}$ be the distinguished irreducible representations of G given by Lemma 2.15. Fix $t > 2\sum_{i=1}^{d}(\dim V_i - 1)$ and let $M_t \subset \Gamma$ be as given by Lemma 3.18. To simplify the notation, we will work with only one representation and dismiss it from the notation. As usual, we will indicate explicitly the points where we have to take care of the multitude of representations.

First observe, by definition of loxodromy of g, that there exist $r > 0$ and $s_0 \in \mathbb{N}$ such that for all $s \geq s_0$, there exists $r \geq \epsilon_s > 0$ with $\epsilon_s \underset{s\to\infty}{\longrightarrow} 0$ and g^s is (r, ϵ_s)-loxodromic. Moreover, for each $s \geq 1$, we have $x_g^+ = x_{g^s}^+$ and $X_g^< = X_{g^s}^<$. Now, choose t elements $\gamma_1, \ldots, \gamma_t \in M_t$ for $x_g^+ \in \mathbb{P}(V)$ as in Lemma 3.18 with its conclusions.

Then it follows that there exist at most $\dim V - 1$ elements ($\sum_{i=1}^{d}(\dim V_i - 1)$, taking each of the d representations into account) among $\{\gamma_1, \ldots, \gamma_t\}$ such that, denoting the set of those by D, for all $\gamma \in \{\gamma_1, \ldots, \gamma_t\} \setminus D$, we have $\gamma.x_g^+ \in X_g^<$.

Now, using the last estimations of Lemma 3.21, one sees that there exists $r_1 > 0$ such that for all $s \in \mathbb{N}$ large enough, there exists $r_1 > \tilde{\epsilon}_s > 0$ and for each $\gamma \in \{\gamma_1, \ldots, \gamma_t\} \setminus D$, we have that γg^s is $(r_1, \tilde{\epsilon}_s)$-loxodromic.

Moreover, using again the fact that $x_{\gamma g^s}^+$ is close to γx_g^+ in terms of $\tilde{\epsilon}_s$ (Lemma 3.21), we see by the choice of γ's that for all $s \in \mathbb{N}$ large enough, the elements of $\{x_{\gamma g^s}^+ \mid \gamma \in \{\gamma_1, \ldots, \gamma_t\} \setminus D\}$ are in general position.

Since h acts by an invertible linear transformation on V, it follows by dimension considerations that for each $l \geq 1$, there exist at most $\dim V - 1$ elements (for the totality of the representations, at most $\sum_{i=1}^{d} \dim V_i - 1$ as above) in $\{\gamma_1, \ldots, \gamma_t\} \setminus D$, such that, calling the set of those by $\tilde{D}_l$ for each $\gamma \in \{\gamma_1, \ldots, \gamma_t \setminus D \cup \tilde{D}_l\}$, we have $h^l x_{\gamma g^s}^+ \notin X_g^<$.

Finally, using once more the last estimations in Lemma 3.21, we conclude that for each $L \in \mathbb{N}$, there exists $\hat{r}_1 = \hat{r}_1(L, g, h, M_t, G)$ such that for all $s \in \mathbb{N}$ large enough and each $l \leq L$, there exists $\gamma_{i_l} \in \{\gamma_1, \ldots, \gamma_t\}\ D \cup \tilde{D}_l$ with $h^l \gamma_{i_l} g^s$ is $(\hat{r}, \hat{\epsilon}_s)$-loxodromic with $\hat{\epsilon}_s \underset{s\to\infty}{\longrightarrow} 0$. Since, we chose $t > 2\sum_{i=1}^{d}(\dim V_i - 1) \geq |D \cup \tilde{D}_l|$, the existence of such γ_{i_l} is insured and the lemma follows. □

We are now in a position to prove the Proposition 4.15 following the strategy described before the statement of this proposition.

Proof of Proposition 4.15. 1. Since S is bounded, by definition of $\Lambda_n(S)$'s, $\bigcup_{n\geq 1} \Lambda_n(S)$ is contained in a compact subset C of $\mathfrak{a}$. As a result, to show that $\lim_{n\to\infty} d_H(\Lambda_n(S), \Lambda(S)) =$

0, by Lemma 4.14, it suffices to show that

$$\lim_{n\to\infty} \sup_{x\in\Lambda_n(S)} \limsup_{m\to\infty} d(x,\Lambda_m(S)) = 0 \tag{4.9}$$

This easily follows from the hypothesis that $e \in S$ and the fact that for each $g \in G$ and for all $k \geq 1$, $\lambda(g^k) = \lambda(g)^k$. Indeed, let $\epsilon > 0$ be given and fix $n_0 \in \mathbb{N}$ and $x_0 \in \Lambda_{n_0}(S)$ and put $\tilde{C} := \max_{x\in C}||x||$. Then, by definition, there exists $g_0 \in S^{n_0}$ with $x_0 = \frac{\lambda(g_0)}{n_0}$. But then for all $k \geq 1$ and $0 \leq l \leq n_0 - 1$, we have $g_0^k = g_0^k e^l \in S^{kn_0+l}$, so that $\frac{\lambda(g_0^k)}{kn_0+l} = \frac{k\lambda(g_0)}{kn_0+l} = \frac{kn_0x_0}{kn_0+l} \in \Lambda_{kn_0+l}(S)$. One then readily observes that for all $m \in \mathbb{N}$ such that $\lfloor \frac{m}{n_0} \rfloor \geq \frac{\tilde{C}}{\epsilon}$, we have $d(x_0, \Lambda_m(S)) \leq \epsilon$. This clearly proves (4.9) and establishes the claim.

2. Since the union of $\Lambda_n(S)$ is contained in a compact set, again by Lemma 4.14, it suffices to show that $\lim_{n\to\infty} \sup_{x\in\Lambda_{n_0}(S)} \limsup_{m\to\infty} d(x,\Lambda_m(S)) = 0$. Fix $n_0 \in \mathbb{N}$, as in the first part, we will in fact show that for each $x \in \Lambda_{n_0}(S)$

$$\lim_{m\to\infty} d(x,\Lambda_m(S)) = 0 \tag{4.10}$$

To do this, fix $\delta > 0$ and let $x_0 \in \Lambda_{n_0}(S)$ and $g_0 \in S^{n_0}$ such that $x_0 = \frac{\lambda(g_0)}{n_0}$.

By spectral radius formula, Lemma 2.18, there exists $K_\delta \in \mathbb{N}$ such that for each $k \geq K_\delta$, we have $||\lambda(g_0) - \frac{\kappa(g_0^k)}{k}|| \leq \frac{\delta}{5}$, fix such a k_0. We thus have

$$||x_0 - \frac{\kappa(g_0^{k_0})}{k_0 n_0}|| \leq \frac{\delta}{5n_0} \leq \frac{\delta}{5} \tag{4.11}$$

Now, let $r = r(\Gamma)$ and fix $\epsilon < r$. Let $F = F_{(r,\epsilon)}$ be the finite subset of Γ given by Theorem 2.24. Let M_F be the compact subset of $\mathfrak{a}$ given by Lemma 2.19 for the compact $F \subset \Gamma$. Denote $\tilde{M}_F := diam(M_F)$. Furthermore, for each $f \in F$, fix $n_f \in \mathbb{N}$ such that $f \in S^{n_f}$ and put $i_0 := \max_{f\in F} n_f$. Finally, let C be a compact in $\mathfrak{a}$ such that $K_n(S) \subseteq C$ for all $n \geq 1$ and put $\tilde{C} := \max_{x\in C}||x||$.

By Lemma 2.19, for all $f \in F$ we have $||\kappa(g_0^{k_0}f) - \kappa(g_0^{k_0})|| \leq \tilde{M}_F$ so that

$$\begin{aligned} &||\frac{\log\kappa(g_0^{k_0}f)}{k_0n_0+n_f} - \frac{\kappa(g_0^{k_0})}{k_0n_0}|| \leq \\ &||\frac{\kappa(g_0^{k_0}f)}{k_0n_0+n_f} - \frac{\kappa(g_0^{k_0}f)}{k_0n_0}|| + ||\frac{\kappa(g_0^{k_0}f)}{k_0n_0} - \frac{\kappa(g_0^{k_0})}{k_0n_0}|| \leq \frac{i_0\tilde{C}}{k_0n_0} + \frac{\tilde{M}_F}{k_0n_0} \end{aligned} \tag{4.12}$$

Combining this last inequality with (4.11), we get that for k_1 large enough (for example, $k_1 \geq K_\delta \vee \frac{5i_0\tilde{C}}{\delta} \vee \frac{5\tilde{M}_F}{\delta}$, see also (4.16)), we have

$$||x_0 - \frac{\kappa(g_0^{k_1}f)}{k_1n_0+n_f}|| \leq \frac{3\delta}{5} \quad \text{and} \quad g_0^{k_1}f \text{ is } (r,\epsilon)\text{-loxodromic} \tag{4.13}$$

where the last assertion follows by the choice of $f \in F$ for $g_0^{k_1}$ as in Theorem 2.24.

Now, take g as $g_0^{k_1} f$ in Lemma 4.20 and let $\gamma_1, \ldots, \gamma_{\hat{d}} \in \Gamma$ be given by this lemma. For $i = 1, \ldots, \hat{d}$, choose $m_i \in \mathbb{N}$ such that $\gamma_i \in S^{m_i}$ and put $j_0 := \max_{i=1,\ldots,\hat{d}} m_i$. Now, fix $h \in S$, take $L = j_0 + i_0 + k_1 n_0$ and apply Lemma 4.20 to get $\hat{r} > 0$ such that for each $\hat{r} > \hat{\epsilon} > 0$, for all $s \in \mathbb{N}$ large enough, and for each $l \leq L$, there exists $i_l \in \{1, \ldots, \hat{d}\}$ such that $h^l \gamma_{i_l} (g_0^{k_1} f)^s$ is $(\hat{r}, \hat{\epsilon})$-loxodromic.

First, observe that using Corollary 3.14, we get a compact $K_{(r,\epsilon)}$ in $\mathfrak{a}$ such that for all $s \geq 1$

$$||\kappa(g_0^{k_1} f) - \frac{\kappa((g_0^{k_1} f)^s)}{s}|| \leq diam(K_{(r,\epsilon)}) \tag{4.14}$$

Second, using Lemma 2.19 with the finite set $H = \{h^i \gamma_j \mid 1 \leq i \leq L,\ 1 \leq j \leq \hat{d}\}$, one gets a compact $M_H \subset \mathfrak{a}$ such that for all $l \leq L$ and $s \geq 1$

$$||\kappa((g_0^{k_1} f)^s) - \kappa(h^l \gamma_{i_l} (g_0^{k_1} f)^s)|| \leq diam(M_H) \tag{4.15}$$

Now, using the boundedness of $\bigcup_{n \geq 1} K_n(S)$ as in (4.12), putting together (4.13), (4.14) and(4.15), one gets that for all $l \leq L$ and $s \geq 1$

$$||x_0 - \frac{\kappa(h^l \gamma_{i_l} (g_0^{k_1} f)^s)}{l + m_l + s(k_1 n_0 + n_f)}|| \leq \frac{4\delta}{5} \tag{4.16}$$

given that k_1 is large enough in terms of $diam(K_{(r,\epsilon)})$, which, therefore, we can suppose without loss of generality before (4.13).

Finally using the fact that for all $l \leq L$ and s large enough, $h^l \gamma_{i_l} (g_0^{k_1} f)^s$ is $(\hat{r}, \hat{\epsilon})$-loxodromic, Proposition 2.20 yields a compact $M_{(\hat{r},\hat{\epsilon})}$ in A_G such that for all $l \leq L$ and s large enough $||\kappa(h^l \gamma_{i_l} (g_0^{k_1} f)^s) - \lambda(h^l \gamma_{i_l} (g_0^{k_1} f)^s)|| \leq diam(M_{(\hat{r},\hat{\epsilon})})$.

It then follows by (4.16) that for all $l \leq L$ and s large enough $||x_0 - \frac{\lambda(h^l \gamma_{i_l} (g_0^{k_1} f)^s)}{l + m_l + s(k_1 n_0 + n_f)}|| \leq \delta$. Since the term $\frac{\lambda(h^l \gamma_{i_l} (g_0^{k_1} f)^s)}{l + m_l + s(k_1 n_0 + n_f)}$ belongs to $\Lambda_{l + m_l + s(k_1 n_0 + n_f)}(S)$, this proves (4.10) and therefore establishes the claim of the proposition.

□

The next proposition is the generalisation of the equality of generalised and joint spectral radii (Theorem 4.1) to the notions of joint Cartan and Jordan spectra in the particular case where S generates a Zariski dense semigroup. It enables us to talk about the joint spectrum $J(S)$ for such S.

Proposition 4.21. *For any bounded subset S of G, we have $\Lambda(S) \subseteq K(S)$. If, moreover, S generates a Zariski dense semigroup in G, then $\Lambda(S) = K(S) =: J(S)$*

Démonstration. The first claim will be a consequence of the spectral radius formula, Lemma 2.18, and the uniform continuity of Cartan projections, Lemma 2.19. Let us show in fact that, for all $n \geq 1$ and $x \in \Lambda_n(S)$, we have $\lim_{m \to \infty} d(x, K_m(S)) = 0$. The first claim is then easily seen to follow from this.

Let $n_0 \in \mathbb{N}$, $\delta > 0$ be given. Fix an element $h \in S$ and put $L = \{h, h^2, \ldots, h^{n_0-1}\}$, a compact subset of G. Let also $x_0 \in \Lambda_{n_0}(S)$, say $x_0 = \frac{1}{n_0}\lambda(g_0)$ for some $g_0 \in S^{n_0}$. Let also C be a compact set in $\mathfrak{a}$ such that $K_n(S) \subseteq C$ for all $n \geq 1$, which exists since S is bounded, and put $\tilde{C} = \max_{x \in C} ||x||$.

By spectral radius formula, there exists $K_\delta \in \mathbb{N}$ such that for each $k \geq K_\delta$, we have $||\lambda(g_0) - \frac{\kappa(g_0^k)}{k}|| \leq \frac{\delta}{3}$, so that, $||x_0 - \frac{\kappa(g_0^k)}{kn_0}|| \leq \frac{\delta}{3n_0} \leq \frac{\delta}{3}$. On the other hand, observe that, by Lemma 2.19, there exists a compact set $M \subset \mathfrak{a}$ such that for all $k \geq 1$ and $l \in \{0, 1, \ldots, n_0 - 1\}$; we have $||\kappa(g_0^k h^l) - \kappa(g_0^k)|| \leq diam(M)$.

As a result, by elementary calculations, we see that for all $k \geq N_\delta := K_\delta \vee \frac{3diam(M)}{\delta} \vee \frac{3\tilde{C}}{\delta}$ and $l \in \{0, 1, \ldots, n_0 - 1\}$, we have $||x_0 \, \frac{\kappa(g_0^k h^l)}{kn_0 + l}|| \leq \delta$. The claim then follows since $g_0^k h^l \in S^{n_0 k + l}$.

For the second claim, we will make use of Cartan and Jordan spectra being limits (Proposition 4.13, Proposition 4.15), AMS finiteness result (Theorem 2.24), Cartan projection's uniform continuity (Lemma 2.19) and the closeness of Cartan and Jordan projections for a loxodromic element (Proposition 2.20). By the first claim, we only need to show $K(S) \subseteq \Lambda(S)$.

Let $x \in K(S)$. Since under the hypothesis, by Proposition 4.13, one has $\lim_{m \to \infty} K_m(S) = K(S)$ for the Hausdorff distance, we can choose $x_m \in K_m(S)$ with $||x_m - x|| = o(1)$. For each $m \geq 1$, fix $g_m \in S^m$ such that $x_m = \frac{1}{m}\kappa(g_m)$. Denote by Γ the semigroup generated by S, let $r = r(\Gamma)$ be as in Theorem 2.24, fix $0 < \epsilon \leq r$ so that by the same theorem, we have the corresponding finite set $F = F_{(r,\epsilon)} \subset \Gamma$. For each $m \geq 1$, fix $f_m \in F$ such that $g_m f_m$ is (r, ϵ)-loxodromic. Also fix $n_f \in \mathbb{N}$ such that $f \in S^{n_f}$ and put $i_0 = \max_{f \in F} n_f$.

Then, by Lemma 2.19, there exists a compact M_F in $\mathfrak{a}$ and by Proposition 2.20 a compact $M_{(r,\epsilon)}$ such that for each $m \geq 1$, we have $\lambda(g_m f_m) \in \kappa(g_m) + M_F + M_{(r,\epsilon)}$. Therefore, putting $q_m = m + n_{f_m} \leq m + i_0$, we have that

$$\frac{1}{q_m}\lambda(g_m f_m) \in (1 - \frac{m}{q_m})x_m + \frac{diam(M_F)}{q_m} + \frac{diam(M_{(r,\epsilon)})}{q_m} = O(\frac{1}{m})$$

where the equality follows by recalling that for all $m \in \mathbb{N}$, $x_m \in C$, where C is a compact set in $\mathfrak{a}$ as above. Finally, since for each $m \geq 1$, $g_m f_m \in S^{q_m}$, we have $\frac{1}{q_m}\lambda(g_m f_m) \in \Lambda_{q_m}(S)$ and thus $m \leq q_m \leq m + i_0$ implies that $d(x, \Lambda_{q_m}(S)) = O(\frac{1}{q_m})$. This completes the proof in view of the fact that under our hypothesis, by Proposition 4.15, $\lim_{n \to \infty} \Lambda_n(S) = \Lambda(S)$. $\square$

Coming back to the setting of random products of matrices and focusing on the joint spectral radius, we immediately see upon its definition that for the support of a probability measure μ governing the random walk S_n, the logarithm of its joint spectral radius appears as an upper bound to the first Lyapunov exponent $\lambda_1(\mu)$ of

μ and in case an LDP exists for the sequence of random variables $\frac{1}{n}\log||S_n||$, as an upper bound to the values contained in the effective support of the corresponding rate function. Therefore, the following proposition is an extensive translation of this observation for the joint spectrum, including also the 'converse statement'.

Proposition 4.22. *Let μ be a probability measure supported on a bounded subset S of a connected semisimple linear real algebraic group G. If S generates a Zariski dense sub-semigroup in G, then the closure of the effective support $\overline{D}_I$ of the rate function I given by Theorem 3.1 (or Theorem 3.2) is the joint spectrum $J(S)$ of S.*

Démonstration. We first show $\overline{D}_I \subseteq K(S)$ (Recall that by Proposition 4.21 $J(S) = K(S) = \Lambda(S)$). Since $K(S)$ is closed by definition, it obviously suffices to show $D_I \subseteq K(S)$. Let $x \in D_I$ and O_x be a neighbourhood of x in $\mathfrak{a}$. Then, by Theorem 3.1, the LDP inequality implies that

$$-\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}(\frac{1}{n}\kappa(S_n)\in O_x)\leq \inf_{y\in O_x} I(y)\leq I(x)<\infty$$

In particular, for all $n \in \mathbb{N}$ large enough, $\mathbb{P}(\frac{1}{n}\kappa(S_n) \in O_x) > 0$, implying that for all n large enough, $K_n(S) \cap O_x \neq \emptyset$. By definition of $K(S)$, since O_x is arbitrary, it follows that $x \in K(S)$.

To prove $\overline{D}_I \supseteq K(S)$, we shall show that for all $x \in K(S)$ and $\delta > 0$, we have $B(x,\delta) \cap D_I \neq \emptyset$. Let such x and δ be given. Since by Proposition 4.13, $K(S)$ is a Hausdorff limit of $K_n(S)$, there exists N_δ such that for each $n \geq N_\delta$, $K_n(S) \cap B(x, \frac{\delta}{4}) \neq \emptyset$. Let $n_0 \in \mathbb{N}$ be large enough (to be specified later), such that $x_{n_0} \in K_{n_0}(S)$ and $x_{n_0} \in B(x, \frac{\delta}{4})$. Denote by g_{n_0} an element of S^{n_0} such that $x_{n_0} = \frac{\kappa(g_{n_0})}{n_0}$, and let U_{n_0} be a neighbourhood of g_{n_0} in G such that $\frac{\kappa(U_{n_0})}{n_0} \subseteq B(x, \frac{\delta}{4})$. Take a compact C of $\mathfrak{a}$ such that $K_n(S) \subseteq C$ for each $n \geq 1$. This is indeed possible since S is bounded. Finally, put $\tilde{C} = \max_{x\in C}||x||$.

Denote by Γ the Zariski dense sub-semigroup of G generated by S and let $r = r(\Gamma)$ be as given by Theorem 2.24. Fix $0 < \epsilon \leq r$ such that $6\epsilon \leq r$ and let $F = F_{(r,\epsilon)}$ be the finite subset of Γ given by Theorem 2.24. For each $f \in F$, fix a neighbourhood V_f of f in G as in Remark 2.25. Let f_0 be an element of F such that $g_{n_0}f_0$ is (r,ϵ)-loxodromic. Up to reducing U_{n_0}, we can suppose by Remark 2.25 that for every $g \in U_{n_0}$ and $f' \in V_{f_0}$, gf' is (r,ϵ)-loxodromic.

Furthermore, let M be the compact subset of $\mathfrak{a}$ obtained by Lemma 2.19, applying it with $L = \overline{V}_f$. Put $K = K_{(\frac{r}{6},\epsilon)}$ the compact subset of $\mathfrak{a}$ given by Corollary 3.14. Fix $i_0 \in \mathbb{N}$ such that $f_0 \in S^{i_0}$, let $d_3 = d_3(r) > 0$ be as given by Corollary 3.8 and denote $d_7 = d_3\mathbb{P}(S_{i_0} \in V_{f_0}) > 0$. Finally, set $\beta_0 = \mathbb{P}(S_{n_0} \in U_{n_0}) > 0$.

In Corollary 3.8, taking $E = U_{n_0}V_{f_0}$ and using it with $n_1 = n_0 + i_0$, we get an $(\frac{r}{6},\epsilon)$-Schottky family $E_{n_1} \subseteq E$ such that $\mathbb{P}(S_{n_1} \in E_{n_1}) \geq d_3\mathbb{P}(S_{n_1} \in E)$. Now, arguing similarly as in the proof of Proposition 4.15, using Corollary 3.14, we see that if $n_0 \in \mathbb{N}$ satisfies $n_0 \geq 4\frac{i_0\tilde{C}+diam(M)+diam(K)}{\delta} \vee N_\delta$, then for all $k \geq 1$, and

$h_1, \ldots, h_k \in E_{n_1}$, we have $d(x_{n_0}, \frac{\kappa(h_1.\ldots.h_k)}{n_1 k}) < \frac{\delta}{4}$. Therefore, by this, the independence of random walk increments and the above observation, we have

$$\mathbb{P}(\frac{1}{n_1 k}\kappa(S_{n_1 k}) \in B(x_{n_0}, \frac{\delta}{2})) \geq \mathbb{P}(S_{kn_1} \in E_{n_1}^k) \geq \mathbb{P}(S_{n_1} \in E_{n_1})^k \geq$$
$$d_3^k \mathbb{P}(S_{n_1} \in E)^k \geq d_3^k \mathbb{P}(X_{n_1}.\ldots.X_{i_0+1} \in U_{n_0} \text{ and } S_{i_0} \in V_{f_0})^k =$$
$$d_3^k \mathbb{P}(S_{n_0} \in U_{n_0})^k \mathbb{P}(S_{i_0} \in V_{f_0})^k \geq (\beta_0 d_7)^k > 0$$

Then, this readily gives that

$$\limsup_{m\to\infty} \frac{1}{m} \log \mathbb{P}(\frac{1}{m}\kappa(S_m) \in B(x_n, \frac{\delta}{2})) \geq \frac{\log(\beta_0 d_7)}{n_1} > -\infty$$

Now, using LDP inequality, by Theorem 3.1, we get

$$\underline{\inf}_{y\in B(x,\frac{\delta}{2})} I(y) \leq -\frac{\log(\beta_0 d_7)}{n_1} < \infty$$

This implies in particular that $D_I \cap B(x,\delta) \neq \emptyset$, what we wanted to show. $\square$

The following corollary deduces, from the convexity of the rate function in Theorem 3.1 and the previous proposition, an important property of the joint spectrum of a bounded set $S \subset G$ generating a Zariski dense sub-semigroup in G; it says namely that $J(S)$ is a convex set.

Corollary 4.23. *Let S be a bounded subset of a linear algebraic group G as above, generating a Zariski dense sub-semigroup in G. Then, the joint spectrum $J(S)$ of S is a convex subset of $\mathfrak{a}^+$, and for I and μ as in the previous proposition, we have* $\overset{\circ}{D}_I = J(\overset{\circ}{S})$

Démonstration. For the first assertion, by the previous proposition, one just needs to observe that the set $D_I = \{x \in \mathfrak{a} \mid I(x) < \infty\}$ is convex. This follows immediately from the convexity of I. The second assertion now follows from the first one using again the previous proposition. $\square$

In the following proposition, we prove the remaining assertion of Theorem 4.4, i.e. that $J(\overset{\circ}{S}) \neq \emptyset$ for an S as in that theorem. One important observation is that yet another application of Theorem 2.24 boils down this problem to the problem of finding a point in $J(S)$ such that one can find, in each direction, elements of $K_n(S)$'s, arbitrarily far from that point. To do this, we use the theory of random matrix products and take this point as the Lyapunov vector $\vec{\lambda}_\mu$ and use the central limit theorem (CLT) of Goldsheid-Guivarc'h [64] (for $G = \mathrm{SL}(d,\mathbb{R})$) and Guivarc'h [65] (more generally, for linear semisimple G as before) to find arbitrarily far elements in $K_n(S)$'s, finally the non-degeneracy of the limiting Gaussian distribution in that theorem, to find them in each direction. We note that the fact that the Benoist cone (see Section 5.2) of the semigroup generated by S, is of non-empty interior was used in the proof of the more general CLT of Guivarc'h (the Goldsheid-Guivarc'h CLT is independent of this result). About the Lyapunov vector $\vec{\lambda}_\mu$, in the proof, apart from its role in the statement of the CLT, we only need to know that it belongs to D_I. For the precise definition and a discussion on the Lyapunov vector, see Section 4.2.

Proposition 4.24. *Let $S \subset G$ be as in the Theorem 4.4. Then $J(\overset{\circ}{S}) \neq \emptyset$.*

Démonstration. Remark at the outset that since $S' \supseteq S$ implies $J(S') \supseteq J(S)$, to simplify the proof, we can suppose that S is a countable subset of G. Let now, μ be a probability measure whose support is S, $\vec{\lambda}_\mu$ be the Lyapunov vector in $\mathfrak{a}^+$ of the corresponding μ-random walk S_n and D_I the effective support of the rate function given by Theorem 3.1. We obviously have $\vec{\lambda}_\mu \in D_I \subseteq J(S)$ where the last inclusion is by Proposition 4.22. Since $\mathfrak{a}$ is a finite dimensional vector space and by Corollary 4.23, K_s is a compact convex subset of $\mathfrak{a}$; it suffices to show that for all non-zero affine form ω on $\mathfrak{a}$ such that $\omega(\vec{\lambda}_\mu) = 0$, there exists $y \in K(S)$ with $\omega(y) > 0$.

Let ω be such an affine form and for each $\delta \geq 0$, define $H_\omega^{\geq \delta} := \{x \in \mathfrak{a} \mid \omega(x) \geq \delta\}$, the positive δ half-space of ω. Recall that by [64], [65] the sequence of $\mathfrak{a}$-valued random variables $\frac{\kappa(S_n) - n\vec{\lambda}_\mu}{\sqrt{n}}$ converges in distribution to a non-degenerate Gaussian law on $\mathfrak{a}$, that we shall denote by ν. By non-degeneracy, we have that for each $\delta \geq 0$, we have that $\nu(\partial H_\omega^{\geq \delta}) = 0$. This implies that

$$\begin{aligned}\lim_{n\to\infty} \mathbb{P}(\omega(\frac{\kappa(S_n) - n\vec{\lambda}_\mu}{\sqrt{n}}) \geq \delta) &= \lim_{n\to\infty} \mathbb{P}(\frac{\kappa(S_n) - n\vec{\lambda}_\mu}{\sqrt{n}} \in H_\omega^{\geq \delta}) \\ &= \int_{\mathfrak{a}} \mathbb{1}_{H_\omega^{\geq \delta}} d\nu =: \alpha_\delta > 0\end{aligned}$$

where the last strict inequality is again by non-degeneracy of the Gaussian law ν.

As a result, since by hypothesis $\omega(\vec{\lambda}_\mu) = 0$, we obtain that there exists $N_\delta \in \mathbb{N}$ such that for all $n \geq N_\delta$, we have $\mathbb{P}(\omega(\kappa(S_n)) \geq \delta\sqrt{n}) \geq \frac{\alpha_\delta}{2} > 0$. In particular, for all $n \geq N_\delta$

$$\kappa(S^n) \cap H_\omega^{\geq \delta\sqrt{n}} \neq \emptyset \tag{4.17}$$

where S^n denotes, as usual, the set of n-fold products of elements of S.

Now, denoting by Γ the Zariski dense sub-semigroup of G generated by S, let $r = r(\Gamma)$ and fix $\epsilon < r$. Let $F = F_{(r,\epsilon)}$ be the finite subset of Γ given by Theorem 2.24. Let $M(F)$ be the compact subset of $\mathfrak{a}$ given by Lemma 2.19 for the compact $F \subset \Gamma$ and put $M^\omega := \max_{x \in M_F} |\omega(x)|$. Furthermore, let $K_{(r,\epsilon)}$ be the compact subset of $\mathfrak{a}$ given by Corollary 3.14 and set $K^\omega := \max_{x \in K_{(r,\epsilon)}} |\omega(x)|$. Finally, set $C^\omega := \max_{x \in K(S)} |\omega(x)|$ and for each $f \in F$, choose $n_f \in \mathbb{N}$ such that $f \in S^{n_f}$ and set $i_0 := \max_{f \in F} n_f$. Then, choose $n_0 \in \mathbb{N}$ such that

1. $n_0 \geq N_\delta$
2. $n_0 \geq 4(\frac{i_0 C^\omega + M^\omega + K^\omega}{\delta})^2$

Let then, $g \in S^{n_0}$ such that $\kappa(g) \in H_\omega^{\geq \delta\sqrt{n}}$ and put $x := \frac{\kappa(g)}{n_0}$ so that $\omega(x) \geq \frac{\delta}{\sqrt{n_0}}$. The existence of such a g is indeed insured by (4.17) and Item 1 of the choice of n_0. Let, by Theorem 2.24, $f \in F$ be such that gf is (r, ϵ)-loxodromic and denote by $n_1 = n_0 + n_f \leq n_0 + i_0$.

Now, similar to the proof of Proposition 4.15, by Lemma 2.19 and Corollary 3.14, we have that for each $k \geq 1$

$$\frac{\kappa((gf)^k)}{kn_1} - x \in \frac{n_f x}{n_1} + \frac{M_F}{n_1} + \frac{K_{(r,\epsilon)}}{n_1}$$

Applying ω to this relation, since $\omega(x) \geq \frac{\delta}{\sqrt{n_0}}$, we readily get that for each $k \geq 1$

$$\omega\left(\frac{\kappa((gf)^k)}{kn_1}\right) \geq \frac{\delta}{\sqrt{n_0}} - \frac{i_0 C^\omega}{n_1} - \frac{M^\omega}{n_1} - \frac{K^\omega}{n_1} \geq \frac{\delta}{2\sqrt{n_0}}$$

where the last inequality follows by Item 2 of our choice of n_0.

Consequently, since $(gf)^k \in S^{n_1 k}$, we have that $H_\omega^{\geq \frac{\delta}{2\sqrt{n_0}}} \cap K_{n_1 k}(S) \neq \emptyset$ for each $k \geq 1$. Moreover, since by our hypothesis, by Proposition 4.13, $K_n(S)$ converge to the compact set $K(S)$ in Hausdorff distance, this proves that there exists $y \in K(S)$ with $\omega(y) \geq \frac{\delta}{2\sqrt{n_0}} > 0$, what we wanted to show. □

As an immediate corollary of the proof of the previous proposition, we get the following result about the locations of the Lyapunov vectors :

Corollary 4.25. *Let $S \subset G$ be as in Theorem 4.4 and $J(S) \subset \mathfrak{a}^+$ denote its joint spectrum. Then, for every probability measure μ on G whose support is S, we have that $\vec{\lambda}_\mu \in \mathring{J}(S)$.* □

4.2 Properties of the rate function

In this part, we investigate several properties of the rate functions of LDP's that are obtained in Chapter 3. In the main body, we focus on the rate functions given by Theorem 3.1 and Theorem 3.2, and in the last paragraph we indicate some analogous properties of the rate function of random matrix products as in Theorem 3.3.

Effective support of I and continuity properties

We unify our results about D_I and continuity of I in the following proposition (see also Remark 4.27 following it).

Proposition 4.26. *Let G be a connected semisimple linear real algebraic group, μ a probability measure whose support $S \subset G$ generates a Zariski dense sub-semigroup in G and let I be the corresponding rate function given by Theorem 3.1.*

1. *The effective support $D_I = \{x \in \mathfrak{a} \mid I(x) < \infty\}$ of I, is a convex set of non-empty interior in $\mathfrak{a}^+$.*
2. *If S is, moreover, a bounded subset of G, then we have $\overline{D_I} = J(S)$ and $\mathring{D}_I = \mathring{J}(S)$, where $J(S)$ stands for the joint spectrum of S.*
3. *If S is, moreover, a finite subset of G, then we have $D_I = J(S)$ and $I_{|D_I}$ is bounded above by $-\min_{g \in S} \log \mu(g) < \infty$. In particular, D_I is a convex body contained in $\mathfrak{a}^+$.*

In any case, I is a locally Lipschitz function (in particular continuous) on $\mathring{D}_I$.

Remark 4.27. 1. *In case $D_I \neq \overline{D}_I$, we have $Im(I) = [0, \infty]$. Indeed, that $0 \in Im(I)$ is a general property of rate functions controlling an LDP and is independent of the assumption $D_I \neq \overline{D}_I$: it follows immediately from the upper bound in Definition 2.1 of LDP, by, for example, taking in it E to be the whole set $\mathfrak{a}^+$ and using the lower semi-continuity of I. That $Im(I)$ fills the whole set $[0, \infty]$ now follows by convexity and lower semi-continuity using $D_I \neq \overline{D}_I$.*

2. *For $g \in G$, denote by τ_g the inner automorphism of G defined by $\tau_g(x) = gxg^{-1}$ for every $x \in G$ and by $\tau_{g*}\mu$ the push-forward of a probability measure μ on G by τ_g. Furthermore, for a probability measure μ whose support $S \subset G$ generates a Zariski dense sub-semigroup of G, denote by I_μ the corresponding rate function of LDP given by Theorem 3.1. Then, similar to 3. of Remark 4.3, for all $g \in G$, we have $I_\mu = I_{\tau_{g*}\mu}$. This also follows easily from Lemma 2.19 using the definition of I in Theorem 2.4*

Using the definition of LDP with a rate function I, we obtain the following result as an immediate corollary of the last (continuity) statement of the previous proposition :

Corollary 4.28. *Let R be a subset of $\mathfrak{a}$ intersecting the interior of D_I and such that $\overline{\mathring{R}} = \overline{R}$. Then, we have $\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(\frac{1}{n}\kappa(S_n) \in R) = -\inf_{x \in R} I(x)$.* □

Proof of Proposition 4.26. 1. As already mentioned in the proof of Corollary 4.23, the convexity of I immediately implies that of D_I. To see that $\mathring{D}_I \neq \emptyset$, observe that since G is σ-compact, we can find a bounded subset S_0 of S generating a Zariski dense sub-semigroup in G (see Remark 5.14) and such that $\mu(S_0) > 0$. Let then μ_0 denote the probability measure obtained by restricting μ to S_0 and let also I_0 denote the rate function given by Theorem 3.1 (or Theorem 3.2) corresponding to the μ_0-random walk on G. Now, by the defining property of I in Theorem 2.4 and definition of D_I, it is an easy matter to see that $D_{I_0} \subset D_I$. The claim then follows from Theorem 4.4, since we have $D_{I_0} = J(S_0)$ and $\mathring{J}(S_0) \neq \emptyset$.

2. This was already established in 4. of Theorem 4.4.

3. The first statement ($D_I = J(S)$) follows from the second statement by using 2. of this proposition and lower semi-continuity of I : indeed, if I is bounded above, say, by $M \in \mathbb{R}$, then for all $M' \geq M$, the set $\{x \in \mathfrak{a} \mid I(x) \leq M'\}$ is closed and equals to D_I, and hence 2. yields $D_I = \overline{D}_I = J(S)$.

Boundedness of I follows from the finiteness of S : we observe that any event defined by the random variable S_n for some $n \in \mathbb{N}$, will be either of probability zero or of probability greater than $(\min_{g \in S} \mu(g))^n$. More precisely, let $x \in D_I$ so that $I(x) < \infty$. It follows by the definition of $I(x)$ in Theorem 2.4, that there exists a neighbourhood O of x in $\mathfrak{a}$ such that $\mathbb{P}(\frac{\kappa(S_n)}{n} \in O) \neq 0$, for all n large enough. This means that for all such n, there exist $g_1, \ldots, g_n \in S$, such that $\frac{\kappa(g_n \cdot \ldots \cdot g_1)}{n} \in O$. But then, using the independence of random walk increments X_i's, we have $\mathbb{P}(\frac{\kappa(S_n)}{n} \in O) \geq \mathbb{P}(X_i = g_i$ for each $i = 1, \ldots, n) = \prod_{i=1}^{n} \mathbb{P}(X_i = g_i) \geq (\min_{g \in S} \mu(g))^n$. Now using again the definition of $I(x)$ in Theorem 2.4, we conclude that $I(x) \leq -\min_{g \in S} \log \mu(g)$.

Finally, the local Lipschitz property on the interior of its effective support of a proper convex function on a finite dimensional real vector space is a classical fact (see, for example, Corollary 2.4 in [49]). □

Remark 4.29. *An immediate observation in the proof of 3. of the previous proposition is the following : at least when S is a finite set, the Hausdorff convergence of $K_n(S)$ to the Cartan joint spectrum $K(S)$ is a necessary condition to generalise Theorem 3.1 to a random walk governed by a probability measure supported on S, not necessarily generating a Zariski dense sub-semigroup. Of course, when S does generate such a semigroup, the necessary condition is satisfied by Proposition 4.13.*

Set of zeros

The aim of this subsection is to point out that in the case of a probability measure μ on G with a finite exponential moment, more precisely, in the setting of Theorem 3.2, the classical large deviation estimates of Le Page [84] for random matrix products (see also Bourgerol's work in [30] and for more general results, Benoist-Quint's [14]) imply that the rate function I admits a unique zero in D_I. This unique zero is indeed the Lyapunov vector $\vec{\lambda}_\mu$ in $\mathfrak{a}^+$, and it in fact belongs to $\overset{\circ}{D}_I$ by Corollary 4.25 and Corollary 4.23.

Recall the classical result of Furstenberg and Kesten [59] (or later, a corollary of Kingman's subadditive ergodic theorem), a non-commutative analogue of classical law of large numbers, stating that for a probability measure μ on $GL(d, \mathbb{R})$ with a finite first order moment, i.e. $\int \log(||g|| \vee ||g^{-1}||)\mu(dg) < \infty$, there exist constants, $(\lambda_i(\mu))_{i=1,\ldots,d}$, called the Lyapunov exponents of μ, such that $\frac{1}{n} \log || \bigwedge^i S_n|| \xrightarrow[n \to \infty]{L^1 \text{ and a.s.}} \sum_{j=1}^{i} \lambda_j(\mu)$.

In our setting of a random walk on a connected semisimple linear real algebraic group G, recall that we define the Lyapunov vector $\vec{\lambda}_\mu \in \mathfrak{a}^+$ of a probability measure μ on G, as $\vec{\lambda}_\mu \overset{a.s.}{=} \lim_{n\to\infty} \frac{1}{n}\kappa(S_n)$. This is easily seen to be well-defined in case μ has finite first order moment, i.e. $\int \log M(g)\mu(dg) < \infty$ (see the paragraph following Corollary 3.25 for the definition of $M(.)$) by $\overline{\chi}_i(\vec{\lambda}_\mu) \overset{a.s.}{=} \lim_n \frac{1}{n} \log ||\rho_i(S_n)||_i$ where ρ_i's are the distinguished representations of G as in Lemma 2.15 and $\overline{\chi}_i$'s are, as before, the highest weights of ρ_i's on $\mathfrak{a}$. Moreover, under the hypotheses of Theorem 3.1 and the finite first order moment assumption, it is of course the case that $\vec{\lambda}_\mu$ is a zero of the rate function of the LDP given by Theorem 3.1.

To observe the uniqueness of this zero, we start by recalling some terminology in the setting of random matrix products. A subset S of $GL(d, \mathbb{R})$ is called contracting (or proximal) if there exists a sequence $(g_n)_{n\geq 1}$ in S such that $\frac{g_n}{||g_n||}$ converges to a rank one matrix. It is called strongly irreducible if S does not preserve a finite union of proper subspaces of $\mathbb{R}^d$. At this point, we also remind the important result of Goldsheid-Margulis [61] ; it says that a sub-semigroup Γ in $GL(d,\mathbb{R})$ is contracting if and only if the Zariski closure G of Γ contains a proximal element, provided that V is strongly irreducible as a G-module.

We now cite the following theorem of Le Page [84] (see also Bougerol's exponential decay Theorem 6.2 in [30] where the proximality assumption is removed, and we have a only slightly weaker conclusion, namely on the existence of limits ; and see Benoist-Quint's general results also on exponential decay for different quantities, i.e. general cocycles, and consequently, for Iwasawa, Cartan and Jordan projections Theorems 12.11, 12.17, 13.18, 13.20, 13.21, 13.22 in [14], related to these latter projections, see also Remark 5.28). We note that below in (4.18), the corresponding result holds for $\mathbb{P}(\log ||S_n x|| - n\lambda_1(\mu) < -n\epsilon)$ with a function $\tilde{\phi}$ with the same properties as ϕ.

Theorem 4.30 ([84])**.** *Let S_n denote the n^{th} step of a random walk associated to a probability measure μ on $GL(d,\mathbb{R})$. Suppose that the semigroup generated by supp(μ) is contracting, strongly irreducible, and that μ has a finite exponential moment. Then, there exists a constant $B > 0$ such that for every vector $x \neq 0$ and for every $0 < \epsilon < B$, we have*

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(\log ||S_n x|| - n\lambda_1(\mu) > n\epsilon) = \phi(\epsilon) \tag{4.18}$$

where $\phi : [0,B[\to \mathbb{R}$ is a concave function such that for all $\epsilon \in]0,B[$, $\phi(\epsilon) < 0$ and $\phi(0) = 0$.

First, as noted before the previous theorem, remark that using the theorem for $\mathbb{P}(\log ||S_n x|| - n\lambda_1(\mu) < -n\epsilon)$ with an $x \neq 0$ of norm 1, we immediately get that for all $\epsilon > 0$, $\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(\log ||S_n|| - n\lambda_1(\mu) < -n\epsilon) \leq \tilde{\phi}(\epsilon) < 0$. On the other hand, the following elementary reasoning shows that we also have $\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(\frac{1}{n} \log ||S_n|| > \lambda_1(\mu) + \epsilon) = \phi(\epsilon) < 0$. To see this latter, first, observe that if we fix a unit basis $e_1, \ldots, e_d$ of $\mathbb{R}^d$, for each $n \in \mathbb{N}$, we have the following inclusion of events :

$$\begin{aligned}\{\frac{1}{n} \log ||S_n e_1|| > \lambda_1(\mu) + \epsilon\} &\subset \{\frac{1}{n} \log ||S_n|| > \lambda_1(\mu) + \epsilon\} \\ &\subset \bigcup_i^d \{\frac{1}{n} \log ||S_n e_i|| > \lambda_1(\mu) + \epsilon - \frac{\log d}{2n}\}\end{aligned} \tag{4.19}$$

Using the union bound, for all $0 < \epsilon' < \epsilon < B$ Theorem 4.30 yields,

$$-\infty < \phi(\epsilon) \leq \limsup_n \frac{1}{n} \log \mathbb{P}(\frac{1}{n} \log ||S_n|| > \lambda_1(\mu) + \epsilon) \leq \phi(\epsilon') < 0 \tag{4.20}$$

Since ϕ is concave, it is continuous on $]0,B[$, and the claim follows.

Now, coming back to the setting of Theorem 3.2, with the distinguished representations $(\rho_i)_{i=1,\ldots,d}$ of Lemma 2.15 ; the probability measures $\rho_{i*}\mu$ on $GL(V_i)$ are such that the semigroup generated by supp$(\rho_{i*}\mu)$, i.e. $\rho_i(\Gamma)$, is strongly irreducible (since the representations ρ_i of G are irreducible and the Zariski closure G, of Γ is connected) and proximal (because Γ is Zariski dense in G of which ρ_i's are proximal representations ; so that, for example, Theorem 2.24 applies). Since, moreover, by our exponential moment definition, our finite exponential moment hypothesis implies those of $\rho_{i*}\mu$, Theorem 4.30 applies and as in (4.20), for each $i = 1, \dots, d$, there exists a constant B_i and a function ϕ_i (as in Theorem 4.30) such that for all $0 < \epsilon < B_i$ and all $0 < \epsilon < B_i$, we have

$$\phi_i(\epsilon) = \limsup_n \frac{1}{n} \log \mathbb{P}(\overline{\chi}_i(\frac{1}{n}\kappa(S_n) - \vec{\lambda}_\mu) > \epsilon) < 0$$

and similarly for $\mathbb{P}(\overline{\chi}_i(\frac{1}{n}\kappa(S_n) - \vec{\lambda}_\mu) < -\epsilon)$. Now, by definition of LDP and its rate function in Definition 2.1, with these conclusions, we have proved :

Proposition 4.31. *Under the hypotheses of Theorem 3.2, $\vec{\lambda}_\mu \in \mathfrak{a}^+$ is the unique zero of the rate function I.*

Note that these reasonings alone imply also that $\vec{\lambda}_\mu$ belongs to the interior of the Weyl chamber $\mathfrak{a}^+$, a fact which was observed by Goldsheid-Margulis (see [61], Theorem 6.1.). Compare this to our more precise result in Corollary 4.25, in this regard, see also Proposition 5.13.

Rate function for random matrix products

The aim of this part is to indicate some of the corresponding properties of the rate function appearing in the LDP for the random matrix products in Theorem 3.3. We note that due to our slightly different setting concerning $G \leq GL(d, \mathbb{R})$, some of the results (i.e. central limit theorem, Le Page's result) that we used to investigate the properties of the rate function I of Theorem 3.2, do not directly apply to this setting. As a consequence, we are not able to get the analogues of all the properties summarised in Proposition 4.26 and Proposition 4.31. We note that the confronted difficulties are of the same type as those described in Remark 3.30, i.e concerning large deviation considerations for random walks on reductive groups. These will be the subject of a subsequent study. For the moment, we content with the following :

Proposition 4.32. *In the same setting as in Theorem 3.3, the rate function I obtained in that theorem is a proper convex rate function, and therefore, locally Lipschitz (in particular continuous) on the interior of its effective support $D_I = \{x \in \mathbb{R} \mid I(x) < \infty\}$, where this latter satisfies the followings :*

1. *If S, denoting the support of μ, is a bounded countable subset of $GL(d, \mathbb{R})$, we have $]\log r_{sub}(S), \log r(S)[\subseteq D_I \subseteq [\log r_{sub}(S), \log r(S)]$*
2. *If S is finite, then $D_I = [\log r_{sub}(S), \log r(S)]$, and $I_{|D_I}$ is bounded above by $-\min_{g \in S} \log \mu(g) < \infty$*

Moreover, for all subset $R \subseteq \mathbb{R}$ intersecting the interior of D_I, and such that $\overline{\overset{\circ}{R}} = \overline{R}$ (e.g. intervals of non-empty interior), we have

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(\frac{1}{n} \log ||S_n|| \in R) = -\inf_{x \in R} I(x)$$

$\square$

Remark 4.33. *We would like to mention once more that, in the more particular setting of Corollary 3.29, as indicated in its proof, it follows by the contraction principle (Lemma 3.28) and our application of it, that the rate function of this corollary has the form* $I_{\overline{\chi}}(x) = \inf\{I(y) \mid y \in \mathfrak{a},\ \overline{\chi}(y) = x\}$*, where* I *is given by Theorem 3.2. Consequently, we can give a more explicit description of this rate function* $I_{\overline{\chi}}$ *using, and as in, Proposition 4.26 and Proposition 4.31 ; for example, its effective support is of non-empty interior, if the probability measure* μ *has a finite exponential moment, it has a unique zero, and if the support of* μ *is bounded, the joint spectral radii is contained in the closure of the effective support.*

Remark 4.34. *We note that, in the setup of the previous proposition, the joint spectral subradius* $r_{sub}(S)$ *of the support* S *of* μ *in Theorem 3.3 can be strictly smaller than* 1*, namely by the existence of central factor in* G *(Recall that, in the previous setting, for all* $g \in SL(d,\mathbb{R})$*, we have* $||g|| \geq 1$*, which is of course not true in* $GL(d,\mathbb{R})$*).*

We underline that the argument we give below does not use the particular structure of G, and applies whenever an LDP exists. It is based on continuity properties of joint spectral radius and subradius. We also note that the argument can be generalised, in one direction, to cover probability measures supported on uncountable sets, see Lemma 6.28.

Démonstration. By the proof of Proposition 4.26 and Corollary 4.28, we only have to prove 1., other assertions follow verbatim. The second inclusion in 1. follows by definitions. For the first inclusion, we make use of the continuity of joint spectral radius for the Hausdorff distance (this is known, for example, as mentioned in 1.2 of Bochi-Morris' [26], it follows from Berger-Wang's Theorem 4.1 : the joint spectral radius can be written, by submultiplicativity of the operator norm, as an infimum over continuous functions (of matrices), and hence is upper semicontinuous ; whereas the generalised spectral radius can be written as a supremum, and hence is lower semicontinuous ; thus the equality of these two yields the continuity) and of the joint spectral subradius (this is the main result of Bochi-Morris' recent [26]). Since S is bounded, by continuity, we can find a sequence of finite subsets T_n of S such that $\log r(T_m) \underset{m\to\infty}{\longrightarrow} \log r(S)$ (similarly for $\log r_{sub}(.)$). Note that since S is countable, for each $\gamma \in S$, we have $\mu(\gamma) > 0$. For all $m \geq 1$, consider the restriction of μ on T_m, i.e. denoting this restricted measure by ν_m, we have, for all $m \geq 1$ and $\gamma \in T_m$, $\nu_m(\gamma) = \frac{\mu(\gamma)}{\mu(T_m)}$.

Now, for each $m \geq 1$, denote by $(S_n^m)_{n\geq 1}$, the n^{th} step of the ν_m-random walk on G, and for $x \in \mathbb{R}$, define $I_{li}^{(m)}(x)$ as in Theorem 2.4 for ν_m :

$$I_{li}^{(m)}(x) := \sup_{\substack{O \text{ open in } \mathbb{R} \\ x\in O}} -\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(\frac{1}{n}\log ||S_n^m|| \in O)$$

Similarly to the proof of 3. in Proposition 4.26, since the support T_m of ν_m is finite, it is not hard to see that we have

$$D_m := \{x \in \mathbb{R} \mid I_{li}^{(m)}(x) < \infty\} = [\log r_{sub}(T_m), \log r(T_m)]$$

On the other hand, similarly to the proof of 1. in Proposition 4.26, we have $D_m \subseteq D_I$ (more precisely, in a similar way as in the proof of Lemma 5.33, by definitions of ν_m, one observes that we have $I_{li}^m(x) \geq I(x) - \log \mu(T_m)$). The result follows from these observations. $\square$

Chapitre 5

GROWTH INDICATOR

In Section 5.1, after a discussion of growth of semigroups, we introduce growth indicator functions for a finite set S in a group G as before. They generalise the exponential growth rate. After mentioning their relation to large deviations theory, we analyse their properties. Namely, we establish that, if such an S generates a Zariski dense semigroup in G, them S has en exponential growth in a dense set of points in the joint spectrum (see Theorem 5.8). Main tool is a precise construction of free (r, ϵ)-Schottky semigroups in the spirit of Benoist's earlier work.

In Section 5.2, in a group G as before, we introduce Benoist's limit cone B_Γ of a Zariski dense semigroup Γ, and for a discrete such Γ, Quint's growth indicator of Γ. They are in close analogy with, respectively, our joint spectrum and growth indicators of an S, generating Γ. We make these relations more explicit.

Section 5.3 is a section of miscellaneous results ; the chain of its three subsections do not possess a logical continuity : we first study LDP type properties of Jordan projections of random walks, and establish the analogue of Theorem 3.1 and Theorem 3.2 for a very particular class of random walks, namely of (r, ϵ)-Schottky type (see therein). The second subsection is an indication of a future study to improve the dense exponential growth theorem of Section 5.1. Finally, we set forth a criterion of discreteness for finitely generated semigroups of G as before, and apply this to finitely generated (r, ϵ)-Schottky semigroups.

5.1 Growth indicator of a finite subset of G

Let Γ be a finitely generated infinite semigroup and S a finite generating subset of Γ. One significant property of S is the growth type of S. S is said to have, respectively, polynomial, subexponential and exponential growth type if the sequence $(|S^n|)_{n\geq 1}$, is of polynomial growth (i.e. $O(n^d)$ for some $d \in \mathbb{N}$), dominates every polynomial but is in turn dominated by every sequence α^n where $\alpha > 1$, is of exponential growth. An important but simple observation is that these properties are in fact properties of Γ itself, i.e. they do not depend on the chosen finite generating set S. Consequently, the corresponding terminology is used for Γ.

For each finite generating set $S \subset \Gamma$, the limit $\lim_{n\to\infty} |S^n|^{\frac{1}{n}} =: v_S > 1$ exists by submultiplicativity and is called the exponential growth rate of S. We have indeed $v_S > 1$ if and only if Γ is of exponential growth. In the case of a linear group Γ (i.e. $\Gamma \leq GL(d, \mathrm{k})$ for some $d \in \mathbb{N}$, and field k), it follows by classical results of Milnor-Wolf and Tits that Γ is of exponential growth whenever it is not nilpotent-by-finite.

Recall also that a semigroup Γ is said to be of uniform exponential growth, if there exists a growth gap, i.e. $\inf\{v_S \mid S \text{ finite generating subset of } \Gamma\} > 1$. A further property of linear groups is that they are either solvable-by-finite or have uniform exponential growth. This follows by works of Eskin-Mozes-Oh [51], Breuillard-Gelander [34] and Breuillard [35], in which they establish (in the order, more and more) uniform versions of Tits alternative (with free semigroups in the first, and in full generality in the last two).

Coming back to our setting of a connected semisimple linear real algebraic group G and a finite subset S of G, we now proceed to define our counting functions : the growth indicators of S. By these, instead of assigning a single numerical value v_S to describe the growth asymptotics of the sequence $S, S^2, \ldots$, we assign functions, which describe the growth asymptotics in a more precise way, i.e. in terms of asymptotic behaviours of elements of S^n's with respect to Cartan and Jordan projections.

Definition 5.1. *We call the function $\phi_S : \mathfrak{a} \longrightarrow \mathbb{R}_+ \cup \{-\infty\}$, defined by, $\phi_S(x) := \inf_{\substack{O \text{ open in } \mathfrak{a} \\ x \in O}} \limsup_{n\to\infty} \frac{1}{n} \log \#\{g \in S^n \mid \frac{1}{n}\kappa(g) \in O\}$, the (Cartan) growth indicator of S. Let also ψ_S denote the Jordan growth indicator of S, defined in the same manner as ϕ_S, but with the Jordan projection $\lambda(.)$.*

Observe at the outset that if v_S denotes the exponential growth rate of S, then, for all $x \in \mathfrak{a}$, we have either $\phi_S(x) = -\infty$, or $0 \leq \phi(x) \leq \log v_S$, and similarly for ψ_S.

Remark 5.2. *Before delving into a study of the growth indicators, let us first point at their close relation to the large deviations theory. Denote by $\tilde{\nu}_S$ the uniform probability measure on S^n : for each $g \in S^n$, $\tilde{\nu}_S(g) = \frac{1}{|S^n|}$. Let ν_n be the image (push-forward) of $\tilde{\nu}_n$ by the normalised Cartan projection $\frac{1}{n}\kappa$, i.e. ν_n is a Borel probability measure on $\mathfrak{a}$ such that for all measurable set $O \subset \mathfrak{a}$, we have $\nu_n(O) = \frac{\#\{g \in S^n \mid \frac{1}{n}\kappa(g) \in O\}}{|S^n|}$. We observe that asking whether the sequence of probability measures ν_n satisfies an LDP is equivalent, by Theorem 2.4, to asking whether for all $x \in \mathfrak{a}$, we have*

$$\inf_{\substack{O \text{ open in } \mathfrak{a} \\ x \in O}} \liminf_{n\to\infty} \frac{1}{n} \log \nu_n(O) = \inf_{\substack{O \text{ open in } \mathfrak{a} \\ x \in O}} \limsup_{n\to\infty} \frac{1}{n} \log \nu_n(O) \tag{5.1}$$

But since, by definition of v_S, $\lim_{n\to\infty} \frac{1}{n} \log |S^n| = \log v_S$, the right-hand-side in (5.1) is equal to $\phi_S(x) - \log v_S$. As a consequence, if the sequence ν_S satisfies an LDP, again by Theorem 2.4, the corresponding rate function J will write, $J(x) = \log v_S - \phi_S(x) \in [0, \log v_S] \cup \{\infty\}$. Unlike, as in Chapter 3, the case of a sequence of probability measures on $\mathfrak{a}$ obtained from convolutions of a probability measure on S, the author ignores whether the sequence ν_n satisfies an LDP even in the case when S generates *a Zariski dense sub-semigroup in G (for a positive result in a very particular case, see Section 5.3). Note that, clearly, all this discussion is also valid for the Jordan growth indicator, replacing κ by the Jordan projection λ, except that this time we do not either possess the analogous results of Chapter 3 for the Jordan projection.*

We now proceed to study some properties of the growth indicators. In the next lemma, we note some straightforward relations between the joint spectra and growth indicators. In the sequel, we shorten $\{x \in \mathfrak{a} \mid \phi_S(x) \geq 0\}$ as $\{\phi_S \geq 0\}$, and similarly for $\{\psi_S \geq 0\}$.

Lemma 5.3. *1. The functions ϕ_S and ψ_S are upper semi-continuous.*

2. We have the inclusion $K(S) \subseteq \{\phi_S \geq 0\}$.

3. If have the Hausdorff convergence $\lim_{n\to\infty} K_n(S) = K(S)$, then $K(S) = \{\phi_S \geq 0\}$.

Same statements also hold when we replace $K(S)$ by $\Lambda(S)$ and ϕ_S by ψ_S. In particular, whenever the joint spectrum $J(S)$ exists (e.g. S generates a Zariski dense sub-semigroup in G), we have $J(S) = \{\phi_S \geq 0\} = \{\psi_S \geq 0\}$.

Démonstration. 1. The upper semi-continuity follows directly from the definition : let $\alpha \in \mathbb{R}$ and $x \in \mathfrak{a}$ such that $\phi_S(x) < \alpha$. Then, by definition of ϕ_S, there exists an open set $O \subset \mathfrak{a}$ containing x such that $\limsup_{n\to\infty} \frac{1}{n} \log \#\{g \in S^n \mid \frac{1}{n}\kappa(g) \in O\} < \alpha$. It follows from this, and again from the definition of ϕ_S that for every $y \in O$, we have $\phi_S(y) < \alpha$, which indeed proves that the set $\{x \in \mathfrak{a} \mid \phi_S < \alpha\}$ is open.

2. Let $x \in K(S)$ and O be an open set in $\mathfrak{a}$ containing x. It follows by definition of $K(S)$ (since a non-principal ultrafilter does not contain a finite set) that, in particular, there exists a sequence $n_k \underset{k\to\infty}{\longrightarrow} \infty$ such that for all $k \geq 1$, $K_{n_k}(S) \cap O \neq \emptyset$. Therefore, for all such O, we have $\limsup_{n\to\infty} \frac{1}{n} \log \#\{g \in S^n \mid \frac{1}{n}\kappa(g) \in O\} \geq 0$. Hence, the inclusion $K(S) \subseteq \{\phi_S \geq 0\}$ follows by definition of ϕ_S, by taking infimum over all such neighbourhoods O of x.

3. Observe that if $x \in \{\phi_S \geq 0\}$, then by definition of ϕ_S, for all open set O containing x, there exists a sequence $n_k \underset{k\to\infty}{\longrightarrow} \infty$ such that $K_{n_k}(S) \cap O \neq \emptyset$. If $K_n(S)$ convergence in the sense of Hausdorff to $K(S)$, this clearly implies that $x \in K(S)$. Thus we conclude by the second point. $\square$

Remark 5.4. *1. For a subset $R \subset \mathfrak{a}$ and $\delta > 0$, denote this time by $R^\delta = \{x \in \mathfrak{a} \mid d(x,R) \leq \delta\}$, the closed δ-blow up of R, where we chose an arbitrary norm on $\mathfrak{a}$. One can explicitly describe the set $\{\phi_S \geq 0\}$ in terms of the sets $K_n(S)$ as $\{\phi_S \geq 0\} = \bigcap_{\delta>0} \bigcap_{m\geq 0} (\bigcup_{n\geq m} K_n(S))^\delta$. And similarly for $\{\psi_S \geq 0\}$.*

2. Analogously to the Remark 4.29, the Hausdorff convergence of $K_n(S)$ to $K(S)$, and, a fortiori the equality $K(S) = \{\phi_S \geq 0\}$, stand as necessary conditions for an LDP to hold for the sequence of probability measures ν_n (Same remarks apply of course to the Jordan projections).

The next lemma is in line with the first assertion of Proposition 4.21. Unsurprisingly, its proof is similar, and simpler.

Lemma 5.5. *We have the inclusion* $\{\psi_S \geq 0\} \subseteq \{\phi_s \geq 0\}$.

Démonstration. This basically follows from the spectral radius formula : let $x \in \{\psi_S \geq 0\}$ and O be an open set in $\mathfrak{a}$ containing x. Then, by definition of ψ_S, there exists $n_0 \in \mathbb{N}$ and $g \in S^{n_0}$ such that $\frac{1}{n_0}\lambda(g) \in O$. Since O is open, it follows by spectral radius formula that for all $l \in \mathbb{N}$ large enough, $\frac{1}{n_0 l}\kappa(g^l) \in O$. But then, since $g^l \in S^{n_0 l}$, this shows that $\limsup_{n\to\infty} \#\{g \in S^n \mid \frac{1}{n}\kappa(g) \in O\} \geq 1$. The result follows by taking infimum over all such O. □

The next lemma says in particular that we can read off the growth exponent v_S of S from the growth indicators ϕ_S and ψ_S. Recall that $\log v_S > 0$ indicates by definition that S generates a semigroup Γ of exponential growth. This is the case whenever the Zariski closure G of Γ is for example semisimple linear real algebraic group. This follows for instance from Tits' work in [111], or later from Benoist's constructions of free Schottky semigroups in [11] ; for a slightly more precise version of this latter, see our Proposition 5.10.

Lemma 5.6. *The (attained) maximum of* ϕ_S *and* ψ_S *is* $\log v_S$.

Remark 5.7. *Recall by (5.1) above and paragraph following it that, the positive function* J *appearing there, has the form* $J(x) = \log v_S - \phi_S$. *As a result, the previous lemma shows that this positive function attains zero. Recall also that in Proposition 4.31, we showed that the zero of the rate function for the LDP as in Chapter 3 is attained on a unique value in the joint spectrum* $J(S)$. *Therefore, it would be interesting to study the loci of zeros of* J, *and ask whether this zero of* J *is unique.*

Démonstration. Let C be a compact set in $\mathfrak{a}$ such that $\bigcup_{m\geq 1} K_m(S) \subseteq C$. For each $n \geq 1$, fix a finite cover of C by open balls of radius $\frac{1}{n}$, $O_1^n, \ldots, O_{i_n}^n$. As $\mathfrak{a}$ is (finite dimensional) Euclidean space, we can take $i_n = O(n^d)$ for some $d \in \mathbb{N}$. Observe now that, since, by submultiplicativity of $|S^n|$, one has $\inf_{n\geq 1} |S^n|^{\frac{1}{n}} = v_S$, for each $n \geq 1$, we have

$$\sum_{k=1}^{i_n} \#\{g \in S^n \mid \frac{1}{n}\kappa(g) \in O_k^n\} \geq v_S^n$$

It follows that for each $n \geq 1$, there exists $j_n \in \{1, \ldots, i_n\}$ such that $\#\{g \in S^n \mid \frac{1}{n}\kappa(g) \in O_{j_n}^n\} \geq \frac{v_S^n}{i_n}$. Let now $\{x\} \subset C$ be the Hausdorff limit of a subsequence of $O_{j_n}^n$. As $i_n = O(n^d)$, it is easily checked by definition of ϕ_S that we have $\phi_S(x) = \log v_S$. The same reasoning applies mutatis mutandis to ψ_S and establishes that there exists $y \in \mathfrak{a}$ with $\psi_S(y) = \log v_S$. □

Using the tools of Chapter 2 and Chapter 3, we now prove a more precise result on the growth indicators ϕ_S and ψ_S, in case S generates a Zariski dense sub-semigroup. Parametrising naturally the asymptotic behaviours of elements of S^n for the Cartan projection, by points of the joint spectrum $J(S) = K(S)$, we obtain that there exists a dense set of asymptotic behaviours, such that the number of elements of S^n with a prescribed behaviour from that dense set, grows exponentially. For such an S, this gives, in particular, yet another characterisation of the joint spectrum $J(S)$.

Theorem 5.8. *Let G be a connected semisimple linear real algebraic group and S a finite subset of G generating a Zariski dense sub-semigroup in G. Then, we have*

1. $\phi_S \leq \psi_S$
2. $J(S) = \overline{\{\phi_S > 0\}}$

Remark 5.9.
1. *Notice that if ϕ_S was a concave function, then the set $\{\phi_S > 0\}$ would be a convex subset of $J(S)$ satisfying 2. of the previous theorem, and therefore, we would have $\overset{\circ}{\{\phi_S > 0\}} = \overset{\circ}{J(S)}$, and in particular, $\overset{\circ}{\{\phi_S > 0\}} \neq \emptyset$.*
2. *In Section 5.3, we indicate a way to possibly obtain this last result (in fact, it establishes that for some particular cases), hence to improve on 2. of the previous dense exponential growth theorem. This will be the subject of a future study.*

The first statement of the previous theorem will follow from an application of Abels-Margulis-Soifer finiteness result together with Lemma 2.19, and the fact for loxodromic elements, Cartan and Jordan projections are close (Proposition 2.20). For the second statement, we shall need a little more precision in Benoist's construction of free Schottky semigroups (see Quint's exposition in [100], Proposition II.2.7), which we establish in the following proposition.

Proposition 5.10. *Let G be as in the previous theorem and Γ a Zariski dense semigroup in G. Then, there exists a compact set $K \subset \mathfrak{a}$, depending on Γ, such that for every generating set S of Γ, there exists $m_0 = m_0(\Gamma, S) \in \mathbb{N}$ with the property that for every $g \in G$, there exist $i, j \in \mathbb{N}$ with $i + j \leq m_0$ such that there exist $a \neq b \in S^i g S^j$ freely generating a semigroup in G, and satisfying $\{\kappa(a), \kappa(b)\} \subseteq \kappa(g) + K$. Moreover, this free semigroup is (r, ϵ)-Schottky for some $0 < \epsilon < r$ and contained in Γ, if g belongs to Γ.*

To prove this proposition, we start by the following lemma, whose proof consists of an application of a variant of classical ping-pong lemma. We remind that in the sequel, we often use the notation of Chapter 2.

Lemma 5.11. *Let G be as in Theorem 5.8 and $\{a, b\} \subset G$ be an (r, ϵ)-Schottky family with $\epsilon < \frac{1}{8}$ satisfying $d_i(x^+_{\rho_i(a)}, x^+_{\rho_i(b)}) > 2\epsilon$ for each of the d distinguished representations ρ_i, $i = 1, \ldots, d$, of G. Then, $\{a, b\}$ freely generates a discrete (r, ϵ)-Schottky semigroup in G.*

Démonstration. Observe first that if the semigroup $< a, b > \subset G$ acts on a set X, and there exist two disjoint subsets $A_a, A_b \subseteq X$ and an element $x \in X \setminus (A_a \cup A_b)$ with the property that $a.(A_b \cup \{x\}) \subseteq A_a$ and $b.(A_a \cup \{x\}) \subseteq A_b$, then a, b freely generates the semigroup $< a, b >$.

Recall that for a Euclidean space V, endowing its projective space with the Fubini-Study metric, we have $diam(\mathbb{P}(V)) = 1$. As $\epsilon < \frac{1}{8}$, it follows that, denoting as usual by $(\rho_i, V_i)_{i=1,\ldots,d}$ the distinguished representations of G, we can find an element $\bar{x} = (x_1, \ldots, x_r) \in \prod_{i=1}^d \mathbb{P}(V_i)$ such that for each $i = 1, \ldots, d$, we have $x_i \in (B^\epsilon_{\rho_i(a)} \cap B^\epsilon_{\rho_i(b)}) \setminus (b^\epsilon_{\rho_i(a)} \cup b^\epsilon_{\rho_i(b)})$. Then, setting $X := \prod_{i=1}^d \mathbb{P}(V_i)$, $A_a := \prod_{i=1}^d b^\epsilon_{\rho_i(a)}$, $A_b :=$

$\prod_{i=1}^{d} b^{\epsilon}_{\rho_i(b)}$ and $\bar{x} := x$, it results from the definition of an (r, ϵ)-Schottky family that we are in the setting of the previous paragraph, thus the claim follows, except for the discreteness. This will be more generally proved later on in Proposition 5.38. $\square$

The proof of Proposition 5.10 now consists of, given an element $g \in G$, right multiplying g by an element f of a fixed finite set F given by the Abel-Margulis-Soifer's Theorem 2.24 to render it loxodromic, and then to use Proposition 3.22 appropriately with $E_1 = E_2 = \{gf\}$, i.e. left multiplying gf by elements of a fixed finite set given by Lemma 3.18, to obtain sufficiently many loxodromic elements with dispersed attracting directions so as to choose two of them and conclude by Lemma 5.11. We give the details below :

Proof of Proposition 5.10. To ease the notation, as in the proof of Proposition 3.22, we work in a single distinguished representation of G, and drop it out of the notation. By our choices, our reasonings apply simultaneously to all distinguished representations $(\rho_i, V_i)_{i=1,\dots,d}$. Let $t > 2\sum_{i=1}^{d}(\dim V_i - 1) + 1$ be fixed, $\eta_t > 0$, and the finite set $M_t \subset \Gamma$ be given by Lemma 3.18. Set $L = L(M_t) \geq 1$ as the Lipschitz constant of the set M_t (for its definition, see the paragraph preceding Proposition 3.22), fix $i_0 \in \mathbb{N}$ such that $\bigcup_{i=1}^{i_0} S^i \supseteq M_t$. Let $r = r(\Gamma)$ be given by Theorem 2.24 and fix $\epsilon < \frac{\eta_t}{96L^2} \wedge \frac{r}{6}$. Let $F = F_{(r,\epsilon)}$ be the finite subset of Γ given by Theorem 2.24 and fix $j_0 \in \mathbb{N}$ such that $\bigcup_{i=1}^{j_0} S^i \supseteq F$. Let K be the compact subset M of $\mathfrak{a}$ given by Lemma 2.19 in which we take the compact set L as $M_t \cup F$. Finally, put $m_0 = i_0 + j_0$.

Let now $g \in G$ be given. Then, by Theorem 2.24, there exists $f \in F$ such that gf is (r, ϵ)-loxodromic, let $j_f \leq j_0$ be such that $f \in S^{j_f}$. It is clear by definitions that the singleton $\{gf\}$ is an $(\frac{r}{6}, \epsilon)$-Schottky family which obviously satisfies the narrowness assumption of Proposition 3.22. As a result, applying Proposition 3.22 with $E_1 = E_2 = \{gf\}$, we get two elements $\gamma_1, \gamma_2 \in M_t$ such that $\{\gamma_1 gf, \gamma_2 gf\}$ is an $(\frac{\eta_t}{48L}, 2\epsilon L)$-Schottky family. By our choice of $t \in \mathbb{N}$ above (i.e. large enough), we can clearly take $\gamma_1 \neq \gamma_2 \in M_t$. This implies by Lemma 3.18 that we have

$$d(\gamma_1.x^+_{gf}, \gamma_2.x^+_{gf}) \geq \eta_t \tag{5.2}$$

Let $i_1, i_2 \leq i_0$ be such that for $k = 1, 2$, we have $\gamma_k \in S^{i_k}$.

Recalling now the use of Lemma 3.21 in the proof of Proposition 3.22 (namely, the first part of its second assertion : "$d(x^+_{\rho_j(\gamma g)}, \gamma.x^+_{\rho_j(g)}) < \epsilon_1$"), we see that for $i = 1, 2$, $\gamma_i gf$ satisfies $d(x^+_{\gamma_i gf}, \gamma_i.x^+_{gf}) \leq \epsilon L \leq \frac{\eta_t}{96L}$. Combining this with (5.2), since $L \geq 1$, we get

$$d(x^+_{\gamma_1 gf}, x^+_{\gamma_2 gf}) \geq \eta_t - 2\epsilon L \geq \eta_t - \frac{\eta_t}{48L} \geq \frac{\eta_t}{2} \tag{5.3}$$

As a result, setting $a := \gamma_1 gf$ and $b = \gamma_2 gf$, by constructions, we have $a \in S^{i_1} g S^{j_f}$ and $b \in S^{i_2} g S^{j_f}$ with $i_1, i_2 \leq i_0$, $j_f \leq j_0$, and $\{a, b\}$ is an $(\frac{\eta_t}{48L}, 2\epsilon L)$-Schottky family satisfying also, by (5.3), $d(x^+_a, x^+_b) \geq \frac{\eta_t}{2} \geq 2(2\epsilon L)$. Consequently, Lemma 5.11 is in force and yields that $\{a, b\}$ freely generates a free semigroup of rank 2. Finally, by our choice of the compact set $K \subset \mathfrak{a}$, γ_i's, f and the expressions of a, b, it follows by

Lemma 2.19 that for $i = 1, 2$, we have $\{\kappa(a), \kappa(b)\} \in \kappa(g) + K$, completing the proof of the proposition. □

We now give the proof of Theorem 5.8. For brevity, we shall be slightly less precise with the implied constants, namely in the use of Lemma 3.16, than in the proof of Theorem 3.1 in Chapter 3.

Proof of Theorem 5.8. 1. Let $x \in J(S)$ and $\alpha := \phi_S(x)$. Since, in our case by Lemma 5.3, we have $J(S) = \{\phi_S \geq 0\} = \{\psi_S \geq 0\}$, and therefore, for our purpose, we can suppose that $\alpha > 0$. By definition of ψ_S, the assertion $\psi_S(x) \geq \alpha$ will follow if we can show that for all $\delta > 0$ and neighbourhood O of x, we have

$$\limsup_{n\to\infty} \frac{1}{n} \log \#\{g \in S^n \mid \frac{1}{n}\lambda(g) \in O\} \geq \alpha - \delta \tag{5.4}$$

Let such a $\delta > 0$ (suppose without loss of generality $\delta < \alpha$) and $O \subset \mathfrak{a}$ be given and let $\delta_1, \delta_2 > 0$ be such that $\delta_1 + \delta_2 < \delta$. Then, by definition of ϕ_S, one can find open sets $x \in O_1 \subset O_2 \subset O_3 \subset O$, where the inclusions are super-strict in the sense of Definition 3.15, and such that we have

$$\limsup_{n\to\infty} \frac{1}{n} \log \#\{g \in S^n \mid \frac{1}{n}\kappa(g) \in O_1\} \geq \alpha - \delta_1 \tag{5.5}$$

For all $n \geq 1$, put $T_n := \{g \in S^n \mid \frac{1}{n}\kappa(g) \in O_1\}$. Observe that by Hausdorff convergence of $K_n(S)$ to $J(S)$ (Proposition 4.13) for all $n \geq 1$, T_n is non-empty and most importantly by (5.5), there exists a sequence $n_k \underset{k\to\infty}{\longrightarrow} \infty$ such that we have for all $k \geq 1$, we have

$$|T_{n_k}| \geq e^{n_k(\alpha - \delta_1 - \delta_2)} \tag{5.6}$$

Let now $r = r(\Gamma) > 0$ be as given by Theorem 2.24, choose $0 < \epsilon \leq r$ and let $F = F_{(r,\epsilon)}$ be the finite subset of Γ given by Theorem 2.24. Fix $i_0 \in \mathbb{N}$ satisfying $\bigcup_{i=1}^{i_0} S^i \supset F$. For each $k \geq 1$, construct a finite cover of T_{n_k} as follows : for each $f \in F$, set $T_{n_k,f} := \{g \in S^{n_k} \mid \frac{1}{n_k}\kappa(g) \in O_1 \text{ and } gf \text{ is } (r,\epsilon)\text{-loxodromic}\}$. Theorem 2.24 indeed implies that for each $k \geq 1$, we have $\bigcup_{f\in F} T_{n_k,f} = T_{n_k}$. In particular, for each $k \geq 1$, there exists $f_k \in F$, such that $|T_{n_k,f_k}| \geq \frac{|T_{n_k}|}{|F|}$. Since F is finite, up to passing to a subsequence, we can suppose that there exists $f \in F$ such that for all $k \geq 1$, we have

$$|T_{n_k,f}| \geq \frac{|T_{n_k}|}{|F|} \tag{5.7}$$

Now let $i_f \leq i_0 \in \mathbb{N}$ be such that $f \in S^{i_f}$. Firstly, it follows from Lemma 2.19 (by taking in it, the compact set L as $\{f\}$), and using Lemma 3.16 for the super-strict inclusion $O_1 \subset O_2$ that for all $k \in \mathbb{N}$ large enough, we have

$$T_{n_k,f} \subseteq \{g \in S^{n_k} \mid \frac{1}{n_k + i_f}\kappa(gf) \in O_2 \text{ and } gf \text{ is } (r,\epsilon)\text{-loxodromic}\} \tag{5.8}$$

Secondly, using the fact that Cartan and Jordan projections of loxodromic elements are close (Proposition 2.20), (5.8) and the super-strict inclusion $O_2 \subset O_3$ yields (again by Lemma 3.16) that for all $k \in \mathbb{N}$ large enough and $g \in T_{n_k,f}$, one has $\frac{1}{n_k+i_f}\lambda(gf) \in O_3$. But since for such a $g \in T_{n_k,f}$, we have $gf \in S^{n_k+i_f}$, the inclusion $\frac{1}{n_k+i_f}\lambda(gf) \in O_3$, (5.6) and (5.7) yield that for all $k \in \mathbb{N}$, one has

$$\#\{\gamma \in S^{n_k+i_f} \mid \frac{1}{n_k+i_f}\lambda(\gamma) \in O_3\} \geq \frac{e^{n_k(\alpha-\delta_1-\delta_2)}}{|F|}$$

But this clearly implies (5.4) and hence establishes the first assertion of the theorem.

2. Let $x \in J(S)$, O be an arbitrary bounded neighbourhood of x in $\mathfrak{a}$ and $x \in O_3$ be an open set with $\overline{O}_3 \subset O$. We aim at showing that

$$\limsup_{n\to\infty} \frac{1}{n} \log \#\{\gamma \in S^n \mid \frac{1}{n}\kappa(\gamma) \in O_3\} =: \alpha_0 > 0 \tag{5.9}$$

Indeed, by the same argument as in the proof Lemma 5.6, i.e. by using a sequence of finite (of cardinality $O(k^d)$ for some $d \in \mathbb{N}$) coverings of $\overline{O}_3$ with balls of radius $\frac{1}{k}$, (5.9) will imply that there exists $y \in \overline{O}_3 \subset O$ such that $\phi_S(y) = \alpha_0 > 0$ proving the claim.

We now prove (5.9) : let $x \in O_1 \subset O_2 \subset O_3$ be open sets in $\mathfrak{a}$, where the inclusions are super-strict in the sense of Definition 3.15. Let also $m_0 = m_0(S,\Gamma) \in \mathbb{N}$ and the compact set $K \subset \mathfrak{a}$ be as given by Proposition 5.10. As in 1. above, it results from the fact that $J(S) \subseteq \{\phi_S \geq 0\}$ (Lemma 5.3), that there exists a sequence $n_k \underset{k\to\infty}{\longrightarrow} \infty$ with $\frac{1}{n_k}\kappa(S^k) \cap O_1 \neq \emptyset$ for all $k \geq 1$. Observe at first place that Proposition 5.10 imply in particular that for all $k \geq 1$ and all $g \in S^{n_k}$ there exist $m_a, m_b \leq m_0$ and $a \in S^{n_k+m_a}$, $b \in S^{n_k+m_b}$ satisfying $\{\kappa(a), \kappa(b)\} \subseteq \kappa(g) + K$ and $\{a, b\}$ generates a free (r,ϵ)-Schottky semigroup of rank 2, for some $r \geq \epsilon > 0$.

Observe now that by the super-strict inclusion $O_1 \subset O_2$, Lemma 3.16 implies that if $k \in \mathbb{N}$ is large enough, for all $m \leq m_0$, one has $\frac{1}{n_k+m}(\kappa(g)+K) \subset O_2$. As a result, by above, for all $k \in \mathbb{N}$ large enough, we have

$$\{\frac{1}{n_k+m_a}\kappa(a), \frac{1}{n_k+m_b}\kappa(b)\} \subset O_2 \tag{5.10}$$

For $p \in \mathbb{N}$, let as usual $\{a,b\}^p$ denote the set of p-fold products of a and b. Note that since $\{a,b\}$ generates a free semigroup of rank 2, we have $|\{a,b\}^p| = 2^p$. Also, for all $p \geq 1$, and $\gamma \in \{a,b\}^p$, denote by $p_a(\gamma)$ and $p_b(\gamma)$, respectively, the number of a and b factors in γ. In particular, $p_a(\gamma) + p_b(\gamma) = p$.

Since $\{a,b\}$ is an (r,ϵ)-Schottky family, it follows from Proposition 3.9 that there exists a compact set $K_{(r,\epsilon)}$ in $\mathfrak{a}$, depending only on r and ϵ, such that for all $p \geq 1$ and $\gamma \in \{a,b\}^p$, we have

$$\kappa(\gamma) \in p(co\{\kappa(a), \kappa(b)\} + K_{(r,\epsilon)}) \tag{5.11}$$

As a result, combining (5.10) and (5.11), by the super-strict inclusion $O_2 \subset O_3$, using again Lemma 3.16, we get that for all $k \in \mathbb{N}$ large enough, $p \geq 1$ and $\gamma \in \{a, b\}^p$, one has

$$\gamma \in S^{p_a(\gamma)(n_k+m_a)+p_b(\gamma)(n_k+m_b)} \text{ and } \frac{\kappa(\gamma)}{p_a(\gamma)(n_k+m_a)+p_b(\gamma)(n_k+m_b)} \in O_3 \quad (5.12)$$

Finally, one observes that for all large enough fixed $k_0 \in \mathbb{N}$, setting for each $p \geq 1$ $A_p := \{q(n_{k_0} + m_a) + r(n_{k_0} + m_b) \mid q + r = p\}$, A_p is contained in an interval of size pm_0, in particular, depending linearly on p. Moreover, for all $m \in A_p$, we have $m \leq p(n_{k_0} + m_0)$. But then, since $|\{a, b\}^p| = 2^p$, it results from this observation and (5.12) that there exists a sequence $m_p \underset{p\to\infty}{\longrightarrow} \infty$ such that $\#\{\gamma \in S^{m_p} \mid \frac{1}{m_p}\kappa(\gamma) \in O_3\}$ grows exponentially in m_p. In particular, we have (5.9) and the result follows. $\square$

5.2 The Benoist limit cone and Quint's growth indicator

In this section, we recall two earlier results, of Benoist [11] and Quint [100], which are closely linked to, respectively, our joint spectrum and Cartan growth indicator. As indicated in the introduction, (part of) Benoist's result can be seen as a precursor of the notion of joint spectrum, in effect, the Benoist cone of a semigroup Γ (see below) turns out to be the cone generated by the joint spectrum of a generating set of Γ. Quint's growth indicator corresponds basically to another way of counting, and this in terms of only directions (see below) and for discrete groups, but with a notable concavity conclusion. We make these relations more precise.

Benoist cone

We start by briefly discussing the phenomenon discovered by Benoist [11]. We start by some definitions and notations mostly to state Benoist's result : let V be a topological vector space and E a subset of V. Define dl_E, directional limit cone of E as the set $\{x \in V \mid \exists t_n > 0 \text{ with } \lim_{n\to\infty} t_n = 0, \exists x_n \in E \text{ such that } t_n x_n \to x\}$. Note that this is automatically a cone in V for a non-empty E, and equals to $\{0\}$ if and only if E is bounded. For a Zariski dense semigroup Γ in G, where G is as usual, define the Cartan limit cone of Γ in $\mathfrak{a}^+$ as $Cl_\Gamma := dl(\kappa(\Gamma))$, i.e. Cl_Γ is the closed cone consisting of directions following which Cartan projections of elements of Γ goes to infinity in $\mathfrak{a}^+$. Finally define the limit cone of Γ, l_Γ, as the smallest closed cone in $\mathfrak{a}^+$ containing $\{\lambda(g) \mid g \in \Gamma\}$. We then have the following result of Benoist, of which we only cite a part :

Theorem 5.12 ([11]). *Let G be a semisimple connected linear real algebraic group and Γ a Zariski dense subsemigroup of G. We have*

1. $Cl_\Gamma = l_\Gamma$.
2. *l_Γ is convex and of non-empty interior.*

We note that Benoist also proves a converse to this statement, namely for any such cone as in 2. above, he finds a discrete Zariski dense semigroup in G admitting that cone as its limit cone (for a more precise formulation see Benoist [11], for a generalisation to linear algebraic groups over local fields, see Quint [101]). Henceforth, in the same setting, we shall call this cone the Benoist cone of Γ, and denote it by B_Γ.

In Benoist's original proof, the tools used to show 1. and the convexity statement, are (r,ϵ)-loxodromic elements and Schottky semigroup theory together with Abels-Margulis-Soifer finiteness result (Theorem 2.24), whereas the proof of the second part of 2. (non-empty interior), relies on the other algebraic tools as well (for a simpler proof of a stronger result, see Quint [104]). Through the following straightforward proposition and the subsequent remark, our corresponding result for joint spectrum is seen to provide yet another proof of this fact (new in the case of $G = \mathrm{SL}(d,\mathbb{R})$, for a general G, as indicated by J.F. Quint to the author, Guivarc'h uses Benoist's result to establish the TCL), combining the central limit theorem of Goldsheid-Guivarc'h and Guivarc'h with Abels-Margulis-Soifer result and Benoist's estimates (see the proof of Proposition 4.24).

Proposition 5.13. *Let G be a connected semisimple linear real algebraic group and Γ a boundedly generated Zariski dense semigroup in G. Let T be a bounded generating set of Γ. Then, the Benoist cone B_Γ of Γ equals to the cone generated by the joint spectrum $J(T)$ of T.*

Proof of Proposition 5.13. For a subset J of $\mathfrak{a}$, let $Cone(J)$ denote the cone generated by J, $\{tx \mid t \geq 0 \text{ and } x \in J\}$. To see $B_\Gamma \subseteq Cone(J(T))$, as by definition $B_\Gamma = \overline{Cone(\{\lambda(\gamma) \mid \gamma \in \Gamma\})}$ and $J(T)$ is closed, it suffices to show $\lambda(\gamma) \in Cone(J(T))$ for all $\gamma \in \Gamma$. But this is obvious : let $\gamma \in \Gamma$, and $n_0 \in \mathbb{N}$ be such that $\gamma \in T^{n_0}$. Then, by definition, $\frac{1}{n_0}\lambda(\gamma) \in \Lambda_{n_0}(T)$. Since for all $k \geq 0$, $\lambda(\gamma k) = k\lambda(\gamma)$, we have $\frac{1}{n_0}\lambda(\gamma) \in \Lambda_{kn_0}(S)$. Now, as $\Lambda_n(T) \underset{n\to\infty}{\longrightarrow} \Lambda(T) = J(T)$ by Proposition 4.21, it follows that $\frac{1}{n_0}\lambda(\gamma) \in J(T)$ and hence, $\lambda(\gamma) \in Cone(J(T))$.

To see the reverse inclusion, let $x \in Cone(J(T))$. Then, by Proposition 4.21, there exist $t \geq 0$ and $g_n \in T^n$ for $n \geq 1$, with $\frac{t}{n}\lambda(g_n) \longrightarrow x$ as $n \to \infty$. Since the rays $\mathbb{R}^+.\lambda(g_n) \subset B_\Gamma$, this immediately gives the other inclusion. $\square$

Remark 5.14.
1. *It is not true that in general, a Zariski dense semigroup Γ in G can be boundedly generated. An example can be constructed by using (r,ϵ)-Schottky semigroups. However, it follows, for example, from Tits' [111] that we can find a finite set in Γ generating a semigroup of the same Zariski closure as Γ, i.e. G.*
2. *In the previous proposition, by 1., the bounded generation hypothesis is indeed needed for the formulation of the statement, but as much as the application of this proposition to deduce some properties of Benoist cone of an arbitrary Zariski dense semigroup Γ in G is concerned, we observe, for example similar to the above proof, that if Γ is not boundedly generated, taking an appropriate sequence of larger and larger bounded subsets T_n of Γ, we will have $\overline{\bigcup_{n\geq 1} Cone(J(T_n))} = B_\Gamma$.*

As a result of this proposition and the previous remark, we can immediately recover some properties of the Benoist cone B_Γ of a Zariski dense semigroup Γ (already observed by Benoist [11]) from those of the joint spectrum of its generating set (see Theorem 4.4).

Corollary 5.15. *Let G be as usual and Γ be Zariski dense semigroup in G. Then, the Benoist cone B_Γ is convex and of non-empty interior in $\mathfrak{a}^+$.* □

Remark 5.16. *One sees in a similar way as in the proof of Proposition 5.13 that in fact, the equality $\Lambda(S) = K(S)$ also extends the corresponding result of Benoist, i.e. 1. of Theorem 5.12.*

Quint's growth indicator

In the sequel of this section, let, unless otherwise indicated, G denote a connected semisimple linear real algebraic group of finite center (e.g. $SL(d, \mathbb{R})$). In [100], for a Zariski dense discrete subgroup Γ of G, Quint introduced an exponential counting function ψ_Γ, the growth indicator of Γ, which is of a similar standing as our growth indicators. For this resemblance, we borrowed our terminology from [100]. We indicate the main difference and relations of Quint's and our growth indicators below. Let us first define Quint's function ψ_Γ : to begin, for a given open cone C in $\mathfrak{a}$ endowed with a norm $||.||$, consider the convergence exponent of $\sum_{\substack{\gamma\in\Gamma \\ \kappa(\gamma)\in C}} e^{-t||\kappa(\gamma)||}$, it is easily seen that we can write this exponent as $\limsup_{n\to\infty} \frac{1}{n} \log \#\{\gamma \in \Gamma \mid \kappa(\gamma) \in C \text{ and } ||\kappa(\gamma)|| \leq n\}$. Now for an $x \neq 0$ in $\mathfrak{a}$, consider the quantity, $\psi_0(x) := \inf_{\substack{C \text{ cone in } \mathfrak{a} \\ x\in C}} \limsup_{n\to\infty} \frac{1}{n} \log \#\{\gamma \in \Gamma \mid \kappa(\gamma) \in C \text{ and } ||\kappa(\gamma)|| \leq n\}$. The function ψ_0 indeed depends on the chosen norm $||.||$ on $\mathfrak{a}$. Moreover, since obviously $\psi_0(tx) = \psi_0(x)$ for all $t > 0$ and $x \in \mathfrak{a}$, it factorises through a function on a $||.||$-circle or more simply and essentially, since $\kappa(.)$ takes it values in $\mathfrak{a}^+$, through a function on $\mathbb{P}(\mathfrak{a}^+)$. We also point out at this point that by its expression à la Ruelle-Lanford, ψ_0 can also be seen as the unique candidate to the possible LDP of an explicit sequence of probability measures on Γ.

Quint proceeds to define a function $\psi_\Gamma : \mathfrak{a} \longrightarrow \mathbb{R} \cup \{-\infty\}$, the growth indicator of Γ as $\psi_\Gamma(x) := ||x||\psi_0(x)$ (for his exact formulation see [100]). This way, he obtains an homogeneous function which is easily seen to be independent of the chosen norm $||.||$ on $\mathfrak{a}$. We mention in passing that, one of Quint's main motivations is to generalise the Patterson type measures [91] to measures on the flag variety of higher rank G's (see Quint's [102] and also Albuquerque's [3]). The remarkable concavity of ψ_Γ is one of the central ingredients in this generalisation and it is proved in [100], whose main result is the following theorem. To state it, let also ρ denote the linear form on $\mathfrak{a}$ which writes as the sum of the roots of $\mathfrak{a}$ times their multiplicities (i.e. dimensions of their root spaces in $\mathfrak{g}$). We have :

Theorem 5.17 ([100]). *The growth indicator ψ_Γ satisfies $\psi_\Gamma \leq \rho$, it is concave and upper semicontinuous. Moreover, we have $\{x \in \mathfrak{a} \mid \psi_\Gamma(x) \geq 0\} = B_\Gamma$ and ψ_Γ is strictly positive in the interior of B_Γ.*

Remark 5.18. *1. One remarks the clear analogy between the second result of this theorem with our Theorem 1.10. The concavity statement of this theorem transfers to our setting as an open question. For an analysis of this in a very particular situation, see Section 5.3.*

2. It follows from Quint's proof (or of our Proposition 5.10 and proof of Theorem 5.8 ; both rely on Benoist's constructions) that in fact the last statement of this theorem is also valid for Zariski dense discrete semigroups Γ in G for which we can define ψ_Γ in the same manner.

Upon the definition of Quint's growth indicator, one sees that the main difference of Quint's and our growth indicators is that whereas our growth indicators come with a finite generating set S and counts the exponential rate (with n as parameter) of number of elements in S^n of a given position (norm and direction) in $\mathfrak{a}$ for their Cartan projection, Quint's function counts the exponential rate (with norm of the Cartan projections as parameter) of number of elements in Γ of a given direction for their Cartan projection. As a result, a moment of reflection suggests that for a finite set S generating a discrete Zariski dense semigroup Γ in G, one can relate the 'projectivized version' of our (Cartan) growth indicator ϕ_S with the Quint's ψ_Γ using the notions of joint spectral radii. This is the aim of the rest of this part.

In fact, we establish this relation with a more precise notion of joint spectral radii for S, one which depends on directions. Let us now make this vague statement more explicit : let G, $\mathfrak{g}$, $\mathfrak{a}$, $\mathfrak{a}^+$, S, Γ as before, endow $\mathfrak{a}$ with a Euclidean structure (e.g. with the Killing form) and identify $\mathfrak{a}$ and $\mathfrak{a}^*$ by this. To understand the growth of $\kappa(S^n)$ in terms of directions in $\mathfrak{a}$, set, for $x \in \mathfrak{a} \setminus \{0\}$, $r(x) := \limsup_{n\to\infty} \sup_{g\in S^n} < \frac{x}{||x||}, \frac{1}{n}\kappa(g) > \in \mathbb{R}$. $r(.)$ is easily seen to be continuous on $\mathfrak{a} \setminus \{0\}$ and for all $t > 0$, satisfies of course $r(tx) = r(x)$. Observe also that for an irreducible representation (V, ρ) of G, denoting by $\overline{\chi}_\rho$ the highest weight in $\mathfrak{a}^*$ as before, by Lemma 2.16, we have $r(\overline{\chi}_\rho) = r_\rho(S)^{\frac{1}{||\overline{\chi}_\rho||}}$ where $r_\rho(S)$ denotes the joint spectral radius of S for the representation ρ, as in Section 4.1.

Remark 5.19. *By the last sentence of the previous paragraph, one sees that in fact, one can, essentially, give an equivalent definition of the above function $r(.)$ as the extension, by continuity, of the function defined on the rational directions of $\mathfrak{a}^+$ (for the basis consisting of dominant fundamental weights, with the identification of $\mathfrak{a}$ and $\mathfrak{a}^*$), by $\frac{\overline{\chi}_\rho}{||\overline{\chi}_\rho||} \mapsto r_\rho(S)^{\frac{1}{||\overline{\chi}_\rho||}}$, for $\rho \in R_{ir}(G)$, the set of irreducible rational representations of G.*

In the same spirit as the expression of the (contracted) rate function in the contraction principle of LDP's (Lemma 3.28), set $\phi_{pr}(x) := \sup_{t\geq 0} \phi_S(tx)$. We then have the following lemma in line with the conclusion of the contraction principle.

Lemma 5.20. *For all $x \in \mathfrak{a} \setminus \{0\}$, we have*

$$\inf_{\substack{C \text{ open cone in } \mathfrak{a} \\ x \in C}} \limsup_{n\to\infty} \frac{1}{n} \log \#\{\gamma \in S^n \mid \kappa(\gamma) \in C\} = \phi_{pr}(x)$$

Démonstration. One first observes that the upper semicontinuity of ϕ_S (Lemma 5.3) together with the fact that the set $\{x \in \mathfrak{a} \mid \phi_S(x) \in \mathbb{R}\}$ is compact implies that we have

$$\sup_{t \geq 0} \phi_S(tx) =: \phi_{pr}(x) = \inf_{\substack{C \text{ cone in } \mathfrak{a} \\ x \in C}} \sup_{y \in \overline{C}} \phi_S(y)$$

As a result, from the expression in the assertion of the lemma, one sees that it suffices to show that for an open cone C in $\mathfrak{a}$

$$\limsup_{n \to \infty} \frac{1}{n} \log \#\{\gamma \in S^n \mid \kappa(\gamma) \in C\} = \sup_{y \in \overline{C}} \phi_S(y) \tag{5.13}$$

Using again the fact that $\{x \in \mathfrak{a} \mid \phi_S(x) \in \mathbb{R}\}$ is compact, (5.13) easily follows with a similar argument as in Lemma 5.6, i.e. taking a sequence of finer and finer finite (of polynomially growing cardinality) covers of $\overline{C} \cap \{x \in \mathfrak{a} \mid \phi_S(x) \in \mathbb{R}\}$. $\square$

We now state the first relation of Quint's and our growth indicators in the following proposition. To ease the notation in its proof, let us : for $n \geq 1$ and $x \in \mathfrak{a} \setminus \{0\}$, set $a_n(x) := \sup_{g \in S^n} < \frac{x}{||x||}, \kappa(g) >$, so that $\limsup_{n \to \infty} \frac{1}{n} a_n(x) = r(x)$ and for a cone D in $\mathfrak{a}$, set $D_{\leq n} := \{x \in D \mid ||x|| \leq n\}$.

Proposition 5.21. *For all $x \in \mathfrak{a} \setminus \{0\}$, we have $||x||.\phi_{pr}(x) \leq r(x).\psi_\Gamma(x)$.*

The proof is fairly intuitive as the statement itself, we give the details below :

Démonstration. The statement is equivalent to $\phi_{pr}(x) \leq r(x)\psi_0(x)$; let us show this : as a preliminary observation, start by noting that by definition of $a_n(.)$'s, for all $n \geq 1$ and all cone D in $\mathfrak{a}$, we have

$$\sup_{g \in S^n} \{||\kappa(g)|| \mid \kappa(g) \in D\} \leq \sup_{y \in D} a_n(y) \tag{5.14}$$

In a second step, note the following obvious consequence of (5.14) and definitions : for all $n \geq 1$ and cone D in $\mathfrak{a}^+$, we have

$$\#\{\gamma \in S^n \mid \kappa(\gamma) \in D\} \leq \#\{\gamma \in \Gamma \mid \kappa(\gamma) \in D_{\leq \sup_{y \in D} a_n(y)}\} \tag{5.15}$$

Now let an $\epsilon > 0$ and $x \in \mathfrak{a} \setminus \{0\}$ be given. It follows from Lemma 5.20 that for any open cone D containing x, there exists a sequence $m_k \underset{k \to \infty}{\longrightarrow} \infty$ such that for all $k \in \mathbb{N}$, (putting $e^{-\infty} = 0$) we have

$$e^{m_k(\phi_{pr}(x) - \epsilon)} \leq \#\{\gamma \in S^{m_k} \mid \kappa(\gamma) \in D\} \tag{5.16}$$

On the other hand, it is not hard to see, by definition of ψ_0, that we have

$$\begin{aligned} &\inf_{\substack{D \text{ open cone in } \mathfrak{a} \\ x \in D}} \limsup_{n \to \infty} \frac{1}{n} \log \#\{\gamma \in \Gamma \mid \kappa(\gamma) \in D_{\leq \sup_{y \in D} a_n(y)}\} \\ &\leq (\limsup_{n \to \infty} \frac{a_n(x)}{n}) \psi_0(x) = r(x)\psi_0(x) \end{aligned} \tag{5.17}$$

As a result, putting (5.15), (5.16) and (5.17) together, we get that for any open cone D small enough (in terms of ϵ), we have a sequence $n_k \underset{k\to\infty}{\longrightarrow} \infty$ such that for all $k \geq 1$, we have

$$e^{n_k(\phi_{pr}(x)-\epsilon)} \leq e^{n_k(r(x)\psi_0(x)+\epsilon)}$$

In particular, $\phi_{pr}(x) - \epsilon \leq r(x)\psi_0(x) + \epsilon$ and the result follows. □

5.3 LDP for Jordan projections and a discreteness criterion

As explained in the introduction of this chapter, this section consists of some miscellaneous results on LDP for Jordan projections of random walks, a strengthening of the dense exponential growth theorem of Section 5.1, and a discreteness criterion that we show to apply for the finitely generated (r,ϵ)-Schottky semigroups.

Two observations on the LDP for Jordan projections

One question which we are not able to answer in this text is the following : let G and μ be as in Theorem 3.1, S_n denote, as usual, the n^{th} step of the μ-random walk on G, and $\lambda : G \longrightarrow \mathfrak{a}^+$ be the Jordan projection ; does the sequence of random variables $\frac{1}{n}\lambda(S_n)$ satisfy an LDP ? (We remind the reader that we indicate in 2. of Remark 4.16 the obstruction for our techniques to settle this question.) In this section, our first observation (Corollary 5.23) is that we can control the lower bound of the LDP (see Definition 2.1) for the sequence $\frac{1}{n}\lambda(S_n)$ by the corresponding rate function of the LDP for Cartan projections, i.e. the one given by Theorem 3.1. Our second observation (Corollary 5.29) is that in the particular case when μ is supported on an (r,ϵ)-Schottky family for some $r \geq \epsilon > 0$, an LDP in fact holds for the sequence $\frac{1}{n}\lambda(S_n)$.

Large deviation lower bound for Jordan projections

Let G be a connected semisimple linear real algebraic group, μ be a probability measure on G whose support generates a Zariski dense semigroup in G, and $\mathfrak{g}$, $\mathfrak{a}$, $\mathfrak{a}^+$, $\lambda(.)$, $\kappa(.)$, S_n be as usual. Consider the following function on $\mathfrak{a}$ in relation to Theorem 2.4 : for $x \in \mathfrak{a}$, set

$$J_{li}(x) := \sup_{\substack{O \text{ open set in } \mathfrak{a} \\ x\in O}} -\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(\frac{1}{n}\lambda(S_n) \in O)$$

Let I be the rate function $\mathfrak{a}$ of the sequence $\frac{1}{n}\kappa(S_n)$, given by Theorem 3.1. Then the following proposition is proved in very much a similar way as 1. of Theorem 5.8 : for the current result, the estimation obtained by the use of Abels-Margulis-Soifer result (Theorem 2.24), on the number of elements in the proof of 1. of Theorem 5.8 is replaced by an estimation on probabilities of events, as in Lemma 3.5. To avoid unnecessary repetitions, we omit the details of the proof of :

Proposition 5.22. *For all $x \in \mathfrak{a}$, we have $I(x) \geq J_{li}(x)$.* □

The result on the lower bound in the definition of LDP (Definition 2.1) for the sequence $\frac{1}{n}\lambda(S_n)$ follows readily from the previous proposition :

Corollary 5.23. *For any subset R of $\mathfrak{a}$, we have*

$$\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}(\frac{1}{n}\lambda(S_n)\in R)\geq -\inf_{x\in \overset{\circ}{R}} I(x)$$

where the infimum over an empty set is set to be ∞.

Démonstration. It follows from the definition of $J_{li}(.)$ that for each $x\in\overset{\circ}{R}$, one has

$$\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}(\frac{1}{n}\lambda(S_n)\in R)\geq -J_{li}(x) \tag{5.18}$$

Therefore, taking the supremum over $x\in\overset{\circ}{R}$ in the right hand side of (5.18) and using Proposition 5.22, we get

$$\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}(\frac{1}{n}\lambda(S_n)\in R)\geq -\inf_{x\in \overset{\circ}{R}} J_{li}(x)\geq -\inf_{x\in \overset{\circ}{R}} I(x)$$

□

Remark 5.24. *In particular, denoting by S the support of the probability measure μ, by Proposition 4.26, for every subset R of $\mathfrak{a}$ such that $\overset{\circ}{R}\cap J(S)\neq\emptyset$ we have $\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}(\frac{1}{n}\lambda(S_n)\in R)>-\infty$.*

LDP for Jordan projections of Schottky random walks

We now indicate a rather straightforward deduction of the existence of LDP – for Jordan projections of random walks on an (r,ϵ)-Schottky semigroup (which will turn out to be the same LDP, i.e. with the same rate function, of Cartan projections) – by combining Theorem 3.2, the notion of exponential equivalence of sequences of random variables, and Benoist's Proposition 2.20, Theorem 2.21. We first wish to note that in case of a probability measure μ on G, supported on an (r,ϵ)-Schottky family S, Theorem 3.1 and Theorem 3.2 are still true even if S does not generate a Zariski dense semigroup in G. Indeed, in essence, the Zariski density hypothesis was there to be able to transfer the problem to such Schottky random walks. Moreover, in the same vein, we also wish to point out that a direct proof of the existence of LDP for Jordan projections can be given using Benoist's Theorem 2.21 and the simple Lemma 3.11. But let us illustrate the above mentioned method :

We first give a definition of exponential equivalence following Dembo-Zeitouni [44] :

Definition 5.25. *Let (X,d) be a separable metric space. The sequences of random variables (equivalently, their laws) (Z_n) and $(\tilde{Z}_n)$ defined on a probability space, of joint laws $\mathbb{P}_n$, and of marginals, respectively, ν_n and $\tilde{\nu}_n$, are called exponentially equivalent if, for each $\delta>0$, setting $D_\delta:=\{(x_1,x_2)\in X\times X\mid d(x_1,x_2)>\delta\}$, we have*

$$\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}_n(D_\delta)=-\infty$$

With this definition, we have the following expected large deviation result (see Theorem 4.2.13 or Theorem 4.2.16 in [44]) :

Theorem 5.26 ([44]). *If an LDP with a proper rate function holds for a sequence (Z_n) of random variables (as in Definition 5.25) which are exponentially equivalent to the sequence $(\tilde{Z}_n)$, then the same LDP holds for $(\tilde{Z}_n)$.*

Now, let G be a group as before and μ be a probility measure on G supported on an (r,ϵ)-Schottky family S for some $r \geq \epsilon > 0$. Then the following lemma is an immediate conclusion of Benoist's Theorem 2.21 and Proposition 2.20 :

Lemma 5.27. *The sequences of $\mathfrak{a}^+$-valued random variables $\frac{1}{n}\kappa(S_n)$ and $\frac{1}{n}\lambda(S_n)$ are exponentially equivalent.*

Démonstration. In Definition 5.25, take (X,d) to be the Cartan subalgebra endowed with a norm $||.||$, $(\mathfrak{a},||.||)$, and for each $n \geq 1$, take (Z_n) and $(\tilde{Z}_n)$, respectively, as $\frac{1}{n}\kappa(S_n)$ and $\frac{1}{n}\lambda(S_n)$. Then, note that for the corresponding set $D_\delta \subset \mathfrak{a} \times \mathfrak{a}$ as in Definition 5.25, for each $n \geq 1$, we have

$$\mathbb{P}_n(D_\delta) = \mathbb{P}(||\kappa(S_n) - \lambda(S_n)|| > \delta.n) \tag{5.19}$$

Since, by Theorem 2.21, for all $\delta \in \Gamma$, where Γ denotes the semigroup generated by the support S of μ, γ is $(2r,2\epsilon)$-loxodromic, it follows by Proposition 2.20 that for some compact set $M_{(2r,2\epsilon)} \subset \mathfrak{a}$ and for all $n \geq 1$, we have deterministically (in particular, with probability 1), $\lambda(S_n) - \kappa(S_n) \in M_{(2r,2\epsilon)}$. As a result, for all $\delta > 0$ and $n \in \mathbb{N}$ large enough, we have

$$\mathbb{P}(||\kappa(S_n) - \lambda(S_n)|| \geq n\delta) = 0 \quad \text{whence} \quad \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_n(D_\delta) = -\infty \tag{5.20}$$

establishing the assertion of the lemma. □

Remark 5.28.

1. *Compare (5.19) and (5.20) above with 3. of Remark 6.26.*
2. *We note that the type of statement as 3. of Remark 6.26 is key to transfer the analogues of some classical limit theorems (law of large numbers, central limit theorem etc.) for, for example, Iwasawa projection (see Section 6.2) to corresponding limit theorems for the Cartan projection (see for instance Chapter 12 and 13 of [14], for example, Theorems 13.10, 13.12, 13.15 therein). However, this kind of exponential decay is, in general, not sufficient to transfer the LDP, which is a more precise limit theorem for rare events, and one needs a stronger result such as (5.20), i.e. exponential equivalence.*

Now suppose that, moreover, the set S above generates a Zariski dense semigroup Γ in G (as mentioned before, one can dispense with this hypothesis without altering the statement below). Suppose also that μ is of finite exponential moment in the sense of Section 3.3 (we shall need this only for a direct application of Theorem 5.26, namely for the assumption of properness of the rate function. In fact, one can prove that the weak LDP for $\frac{1}{n}\lambda(S_n)$ exists without this assumption). We then have the following corollary of Theorem 3.2, Lemma 5.27 and Theorem 5.26 :

Corollary 5.29. *The sequence of random variables $\frac{1}{n}\lambda(S_n)$ satisfies an LDP with the same proper convex rate function as the LDP of the sequence $\frac{1}{n}\kappa(S_n)$, given by Theorem 3.2.* □

Remark 5.30. *In the same vein, one can show that for a finite (r,ϵ)-Schottky family S, for the growth indicators of S, one has $\phi_S(x) = \psi_S(x)$ for all $x \in \mathfrak{a}$. The proof is basically the same, we omit the details. For a direct deduction of this from Corollary 5.29 in a particular case, see the following section.*

A particular example and a general application of it

In this part, we study our rate functions and growth indicators for a very particular set T and probability measure μ supported on T. Then, we indicate an application of this study to an improvement of 2. of Theorem 5.8, which will be studied in more detail in a future work.

Let G be a group as before and $T = \{\gamma_1, \dots, \gamma_q\} \subset G$, $q \geq 2$, be a finite set generating a free (r,ϵ)-Schottky semigroup Γ in G, for some $r \geq \epsilon > 0$. Since, the semigroup Γ has no relation in its presentation, the probabilistic and deterministic studies of asymptotics of a set S in Γ, in a sense, differs only superficially, i.e. modulo combinatorics on finite alphabets. To set this difference aside, let μ be the uniform probability measure on T, i.e. $\mu = \frac{1}{q}\sum_{i=1}^{q}\delta_{\gamma_i}$. Then, the natural observation is that the n^{th} convolution μ^{*n} of μ (i.e. the law of S_n, denoting as usual by S_n the n^{th} step of the μ-random walk on $\Gamma \subset G$), is equal to the uniform measure on the set T^n. By consequent, this gives that, the laws of $\frac{1}{n}\kappa(S_n)$ are equal to the probability measures ν_n on $\mathfrak{a}$, of Remark 5.2. Similarly, the laws of $\frac{1}{n}\lambda(S_n)$ equal the push-forwards by $\frac{1}{n}\lambda(.)$ of the uniform measures on T^n. Therefore, for all $n \geq 1$, we have

$$\mathbb{P}(\frac{1}{n}\kappa(S_n) \in O) = \frac{\#\{\gamma \in T^n \mid \frac{1}{n}\kappa(\gamma) \in O\}}{q^n}$$

and similarly for $\kappa(.)$ replaced by $\lambda(.)$ on both sides. By consequent, for all $x \in \mathfrak{a}$, we have

$$-\inf_{\substack{O \text{ open in } \mathfrak{a} \\ x \in O}} \limsup_{n\to\infty} \frac{1}{n}\log\mathbb{P}(\frac{1}{n}\kappa(S_n) \in O) = \log q - \inf_{\substack{O \text{ open in } \mathfrak{a} \\ x \in O}} \limsup_{n\to\infty} \frac{1}{n}\#\{\gamma \in T^n \mid \frac{1}{n}\kappa(\gamma) \in O\} \tag{5.21}$$

(and similarly for $\lambda(.)$.) But, we observe that the left-hand-side of (5.21) is equal to the value at x of the proper convex rate function I given by Theorem 3.2 and the right-hand-side is equal to the value $\log q - \phi_T(x)$, where $\phi_T(.)$ stands for the (Cartan) growth indicator of S. (We remind the reader that for random products of elements of Schottky families, we can dispense with the Zariski density assumption in Theorem 3.2.) Moreover, in the corresponding version of (5.21), with $\kappa(.)$'s replaced by $\lambda(.)$'s, by Corollary 5.29, the left-hand-side also equals $I(x)$, and the right-hand-side equals $\log q - \psi_T(x)$, where $\psi_T(.)$ is the Jordan growth indicator of S.

A direct consequence of the above reasoning is the following

Proposition 5.31. *For the random walk S_n on G with respect to uniform probability measure on T, denoting by $I(.)$ the LDP rate function of the sequence $\frac{1}{n}\kappa(S_n)$ (equivalently, by Corollary 5.29, the LDP rate function of $\frac{1}{n}\lambda(S_n)$), we have*

1. *For all $x \in \mathfrak{a}$, $I(x) = \log q - \phi_T(x) = \log q - \psi_T(x)$, in particular, $\phi_T(x) = \psi_T(x)$.*
2. *$\phi_T = \psi_T$ is a concave function, which is locally Lipschitz on $\overset{\circ}{J(T)}$ and satisfies $J(T) = \{x \in \mathfrak{a} \mid \log q \geq \phi_T(x) \geq 0\}$ and $\phi_{|J(T)^c} = -\infty$.*
3. *If, furthermore, the semigroup generated by T is Zariski dense in G, then $\phi_T = \psi_T$ attains its maximum $\log q$ on a unique point in $\overset{\circ}{J(T)}$, which the Lyapunov vector of the μ-random walk S_n.*

Démonstration. 1. follows from the observations preceding the statement of the proposition. 2. follows from 1. and the convexity of I and 3. of Proposition 4.26. (We note that, similar to the LDP, under the Schottky assumption on T, $J(T)$ is well-defined without the Zariski density assumption of Theorem 4.4. But $\overset{\circ}{J(T)}$ may be empty if $< T >= \Gamma$ is not Zariski dense.) 3. also follows from the corresponding properties of $I(.)$ under Zariski density assumption, namely by Proposition 4.31. □

We note the following corollary of this result, which we will use in our indication of the improvement of 2. of Theorem 5.8.

Corollary 5.32. *For a finite set $T \subset G$ generating a free (r,ϵ)-Schottky semigroup Zariski dense in G, the set $\{x \in \mathfrak{a} \mid \phi_T(x) > 0\}$ is of non-empty interior in $\mathfrak{a}$.*

Démonstration. This follows directly from 3. and the continuity assertion in 2. of the above proposition □

We now proceed with an observation on the growth indicator function, which we shall put to good use in combination with the previous corollary. It can be seen as a generalisation of the argument in the proof of 1. of Proposition 4.26.

Lemma 5.33. *Let G be a group as before and S be an arbitrary finite subset of G. Then, for each $n_0 \in \mathbb{N}$, and every set $T \subseteq S^{n_0}$, and all $x \in \mathfrak{a}$, we have $\phi_S(x) \geq \frac{1}{n_0}\phi_T(n_0 x)$*

Démonstration. Let $n_0 \in \mathbb{N}$, $x \in \mathfrak{a}$ and $T \subseteq S^{n_0}$ be given. Then, for every neighbourhood O of x in $\mathfrak{a}$, for each $k \geq 1$, we clearly have

$$\#\{\gamma \in S^{n_0 k} \mid \frac{1}{n_0 k}\kappa(\gamma) \in O\} \geq \#\{\gamma \in T^k \mid \frac{1}{k}\kappa(\gamma) \in n_0 O\}$$

As a consequence, we have

$$\limsup_{k\to\infty} \frac{1}{n_0 k} \log \#\{\gamma \in S^{n_0 k} \mid \frac{1}{n_0 k}\kappa(\gamma) \in O\} \geq \frac{1}{n_0} \limsup_{k\to\infty} \frac{1}{k} \log \#\{\gamma \in T^k \mid \frac{1}{k}\kappa(\gamma) \in n_0 O\} \tag{5.22}$$

Now, in (5.22), since for neighbourhoods O of x shrinking to x, the neighbourhoods $n_0 O$ of $n_0 x$ shrinks to $n_0 x$, taking the infimum on both sides over the open subsets of $\mathfrak{a}$ containing x, we obtain the desired result by definitions of ϕ_S and ϕ_T. □

About the improvement of 2. of Theorem 5.8, for the time being, we content with the following result. See also the subsequent remark.

Proposition 5.34. *Let S be a finite set in a group G as before. Suppose that there exists $n_0 \in \mathbb{N}$ such that for some $r \geq \epsilon > 0$, there exists a subset T of S^{n_0} generating a free Zariski dense (r,ϵ)-Schottky semigroup in G. Then $\{\phi_S \overset{\circ}{>} 0\} \neq \emptyset$.*

Démonstration. This is an immediate consequence of Corollary 5.32 and Lemma 5.33. □

Remark 5.35. *It follows from Benoist's work [11] that, for example, if $e \in S$ then the hypothesis of the above proposition is satisfied.*

A sufficient condition for discreteness of a finitely generated semigroup in G

In this short section, using the notion of joint spectral subradius, we give an explicit sufficient condition for a finitely generated semigroup of a connected semisimple linear real algebraic group G to be discrete, and apply this criterion to deduce discreteness of finitely generated (r,ϵ)-Schottky semigroups.

Let as before $d \in \mathbb{N}$ denote the real rank of the group G and $(\rho_i, V_i)_{i=1,\dots,d}$ be the distinguished representations of G in $SL(V_i)$'s. Recall that for any representation $G \overset{\rho}{\longrightarrow} GL(V)$ of G in a Euclidean space V, and bounded subset S of G, $r_{sub,\rho}(S)$ denotes the joint spectral subradius of $\rho(S)$ in $GL(V)$. Then, we can write a criterion of discreteness using $r_{sub,\rho}(.)$ as follows :

Lemma 5.36. *(Discreteness criterion) Let S be a finite subset if G and denote by Γ the semigroup generated by S. If for each $i = 1,\dots,d$, $r_{sub,\rho_i}(S) > 1$, then Γ is a discrete sub-semigroup of G.*

Démonstration. Denoting, as usual, by $\overline{\chi}_i$ the highest weights of ρ_i's in $\mathfrak{g}$, we know by Lemma 2.15 that $(\overline{\chi}_i)_{i=1,\dots,d}$ is a real basis of $\mathfrak{a}^*$. It then follows by 2. of Lemma 2.16 and definition of joint spectral subradius that, if for each $i = 1,\dots,d$, $r_{sub,\rho_i}(S) > 1$, then the Cartan projections of elements of iterates of S, i.e. $\kappa(S^n)$, drifts to infinity at a linear speed in $\mathfrak{a}^+$; in other words, for any norm $||.||$ on $\mathfrak{a}$, we have $O(\min_{g \in S^n} ||\kappa(g)||) = n$. The result then follows easily by finiteness of S^n's and continuity of the Cartan projection $\kappa : G \longrightarrow \mathfrak{a}^+$. □

Remark 5.37. *We note that this discreteness criterion is, in a sense, of purely semigroup nature, inasmuch as it automatically fails whenever the identity element $e \in \Gamma$ or more generally $\Gamma \cap \Gamma^{-1} \neq \emptyset$, where $\Gamma^{-1} = \{\gamma^{-1} \mid \gamma \in \Gamma\}$.*

The following result will follow from an application of the above discreteness criterion.

Proposition 5.38. *Let S be a finite (r,ϵ)-Schottky family in G, for some $r \geq \epsilon > 0$, in the sense of Definition 2.22. Then, the semigroup Γ generated by S is discrete in G.*

The assertion of this proposition is implied by Lemma 5.36 as a particular consequence of the next lemma.

Lemma 5.39. *Let S be a (r, ϵ)-Schottky family in G, for some $r \geq \epsilon > 0$. Then for all irreducible rational representation (ρ, V) of G, we have $r_{sub,\rho}(S) > 1$.*

Démonstration. By Remark 4.6, since the dominant weights of irreducible rational representations of G are contained in the $\mathbb{Q}$-span (with positive coefficients) of distinguished highest weights $(\overline{\chi}_i)_{i=1,\dots,d}$, it suffices to prove, by Lemma 2.16, that for each $i = 1, \dots, d$, we have $r_{sub,\rho_i}(S) > 1$. Moreover, using Definition 2.22, to prove this latter, it suffices to show that for a Euclidean space V of dimension $\tilde{d} \geq 2$ and any subset $T \subset SL(V)$, which is an (r, ϵ)-Schottky family in $\mathbb{P}(V)$ in the sense of Definition 2.12, we have $r_{sub}(T) > 1$.

Now, let $n \in \mathbb{N}$, $g \in T^n$ and write $g = g_k^{n_k}. \dots .g_1^{n_1}$, where g_i's are in T, satisfy $g_{i+1} \neq g_i$ and with $\sum_{i=1}^{k} n_i = n$. Note that, g is automatically $(2r, 2\epsilon)$-proximal by Theorem 2.21. Using the notation of Chapter 2, by definition of an (r, ϵ)-Schottky family, it follows that we have

$$g.B_{g_1}^{\epsilon} = g_k^{n_k}. \dots .g_1^{n_1} B_{g_1}^{\epsilon} \subseteq g_k^{n_k}. \dots .g_2^{n_2} b_{g_1}^{\epsilon} \subseteq g_k^{n_k}. \dots .g_2^{n_2} B_{g_2}^{\epsilon} \subseteq \dots \subseteq g_k^{n_k} B_{g_k}^{n_k} \subseteq b_{g_k}^{\epsilon} \tag{5.23}$$

Since, by definition of an (r, ϵ)-Schottky family, for example, $b_{g_k}^{2\epsilon} \subset B_{g_1}^{\epsilon}$, this yields that $x_g^+ \in b_{g_k}^{\epsilon}$. More importantly (for our argument), since by definitions, for each $i = 1, \dots, k$, g_i is ϵ-Lipschitz on $B_{g_i}^{\epsilon}$, it follows from (5.23) that the restriction of g to $B_{g_1}^{\epsilon}$ is an ϵ^n-Lipschitz transformation, more precisely, $Lip(g_{|B_{g_1}^{\epsilon}}) \leq \epsilon^n$. In the following, we give a lower bound on $Lip(g_{|B_{g_1}^{\epsilon}})$ in terms of the operator norm $||g||$ of g, and our result will follow from this last inequality.

Let $q \geq 2$ denote the dimension of V, fix an ordered orthonormal basis $(e_1, \dots, e_q)$ of V and write $g = k_1 a k_2$, where $k_1, k_2 \in SO(\mathbb{R}^q)$ and a is the diagonal matrix with decreasing strictly positive coefficients $a_1 \geq \dots \geq a_q$ (singular values of g). Since $SO(\mathbb{R}^q)$ acts by isometries on $\mathbb{P}(\mathbb{R}^q)$ endowed with the Fubini-Study metric, we have $Lip(g_{|B_{g_1}^{\epsilon}}) = Lip(a_{|k_2 B_{g_1}^{\epsilon}})$. Now, noting that $e_1 = x_a^+ \in k_2 b_{g_k}^{2\epsilon} \subset k_2 B_{g_1}^{\epsilon}$, a quick computation with the Fubini-Study metric shows that we have $Lip(a_{|k_2 B_{g_1}^{\epsilon}}) \geq \frac{a_2}{a_1}$. By consequence, the previous paragraph yields that $a_1 \epsilon^n \geq a_2$ and since $a_1. \dots .a_q = 1$, it follows that $||g|| = ||a|| = a_1 \geq \frac{1}{(\epsilon^n)^{1-\frac{1}{q}}}$. In particular, this implies that

$$r_{sub}(T) = \lim_{n \to \infty} (\inf_{g \in T^n} ||g||)^{\frac{1}{n}} \geq \frac{1}{\epsilon^{1-\frac{1}{q}}} > 1$$

where the last strict inequality is implied by, for example 1. of Remark 2.13, proving our claim. □

Chapitre 6

LDP FOR IWASAWA DECOMPOSITION UNDER DENSITY

In Section 6.1, we start by introducing a uniformity condition of Stroock and Ellis ensuring the existence of an LDP for the empirical measures of a Markov chain. In the case of a random walk on a group G, looking at the Markov chain coming from a random walk and an action of the group G on a space X, we express this condition in terms of a condition **(D)** on the probability measure μ (Lemma 6.10) of the random walk and clarify a situation where **(D)** is satisfied (Lemma 6.12).

In Section 6.2, using the contraction principle, we transfer the LDP obtained in Section 6.1 for empirical measures on $G \times X$, by using functions that we construct from cocycles of G-action on X. We apply this general method to two particular cases, respectively with the norm and the Iwasawa cocycles, which we explain in detail. We note that a simple and useful example to keep in mind throughout the exposition is $G = GL(V)$ and $X = \mathbb{P}(V)$ for a Euclidean space V.

6.1 LDP for occupation times of random group actions

Uniformity assumption (U) and LDP for occupation times of Markov chains

In this part, we explain the uniformity assumption **(U)** for a Markov chain and state the result, of Stroock and Ellis [109], [50], on the existence of an LDP for sequence of laws of occupation times of Markov chains. We also set forth a corollary of the results of Donsker-Varadhan [46], [47] ,[48], whose work is the fundamental source of LDP considerations for empirical measures of Markov chains ; in our case, it gives a second expression for the occurring rate function. Finally we mention that in this part, we mainly follow the exposition of Dembo-Zeitouni in [44].

Let Σ be a Polish space, $M_1(\Sigma)$ denote the space of Borel probability measures on Σ equipped with the Lévy-Prokhorov metric, making it into a Polish space with convergence compatible with the weak convergence. Let $\pi(.,.)$ be a Markovian kernel on Σ, i.e. $\forall\sigma \in \Sigma$, $\pi(\sigma,.) \in M_1(\Sigma)$. Fixing π, let $(Z_n^\sigma)_{n\geq 1}$ denote the Markov chain in Σ, with initial state $Z_0 = \sigma$ (i.e. initial distribution δ_σ) and transition probability π. Define the occupation time of the Markov chain $(Z_n^\sigma)_{n\geq 1}$ at step $n \in \mathbb{N}$ as the (random) empirical probability measure $L_n^{Z^\sigma} := \frac{1}{n}\sum_{i=1}^n \delta_{Z_i^\sigma} \in M_1(\Sigma)$. This is a random variable taking values in the Polish space $M_1(\Sigma)$, denote by $\mu_{n,\sigma}$ its distribution (law).

The following assumption on a Markov chain was first introduced by Stroock [109] in a somewhat more restrictive form, and then slightly relaxed by Ellis [50] to obtain an LDP for sequences of laws of occupation times of Markov chains :

Uniformity assumption (U) : *There exist integers $0 < l \leq N$ and a constant $C > 0$, such that for all $\sigma, \tau \in \Sigma$*

$$\pi^l(\sigma,.) \leq C\sum_{m=1}^{N} \pi^m(\tau,.)$$

where $\pi^m(.,.)$ is the m^{th}-step transition probability, inductively defined by $\pi^{m+1}(\tau,.) = \int \pi^m(\xi,.)\pi(\tau,d\xi)$.

Remark 6.1. *1. This assumption clearly holds true for every finite state irreducible Markov chain.*

*2. Under **(U)**, the chain Z_n is uniformly ergodic, admits a unique stationary probability measure and the modified chain with transition probability $\frac{1}{N}\sum_{m=1}^N \pi^m(.,.)$ satisfies the Doeblin condition.*

We now state the main existence of LDP theorem that we shall use in the sequel of this chapter :

Theorem 6.2 (Stroock [109], Ellis [50])**.** *Assume **(U)** and let Z_i^σ be a Markov chain as above. Then, for every $f \in C_b(\Sigma)$, the following limits exist and is independent of σ :*

$$\Lambda(f) = \lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}[\exp(\sum_{i=1}^n f(Z_i^\sigma))],$$

Moreover, $\mu_{n,\sigma}$ satisfies an LDP in $M_1(\Sigma)$ with the proper convex rate function

$$I(\nu) = \Lambda^*(\nu) := \sup_{f\in C_b(\Sigma)}\{< f,\nu > -\Lambda(f)\},$$

Remark 6.3. *1. Actually, $\mu_{n,\sigma}$ satisfies the above LDP uniformly in $\sigma \in \Sigma$ in the following sense : For every closed set $F \subset M_1(\Sigma)$, we have*

$$\limsup_{n\to\infty}\frac{1}{n}\log\sup_{\sigma\in\Sigma}\mu_{n,\sigma}(F) \leq -\inf_{\nu\in F}\Lambda^*(\nu)$$

and for every open set $G \subset M_1(\Sigma)$,

$$\limsup_{n\to\infty}\frac{1}{n}\log\inf_{\sigma\in\Sigma}\mu_{n,\sigma}(G) \geq -\inf_{\nu\in G}\Lambda^*(\nu)$$

2. *The limit $\Lambda(f)$ in the theorem above, does not depend on the initial distribution of the chain.*

As mentioned above, the pioneering work of Donsker-Varadhan (for our setting see [45] Chapter 4 or Theorem 6.5.4 in [44]) gives an alternative expression for the rate function I controlling the above LDP, we cite it from [44] :

Proposition 6.4 ([44]). *Assume* ***(U)*** *holds. Then, for all $\nu \in M_1(\Sigma)$,*

$$I(\nu) = \sup_{f \in C_b(\Sigma), f \geq 1} - \int_\Sigma \log(\frac{\pi f}{f}) d\nu$$

where $\pi f(.) := \int_\Sigma f(\tau)\pi(., d\tau)$.

Group action Markov chains and transferring the assumption (U)

One fundamental way to study the groups is to analyse various properties of their actions. To study random walk on groups, a crucial idea of Furstenberg [57], which also turned out to be very useful, was to associate to random walks, a Markov chain whose state space includes the group and a set on which the group acts. The construction is equivalent to skew-products in dynamical systems. We shall pursue this idea and consider the corresponding Markov chains. In a second step, we investigate the question of which kind of probability measures will lead to uniform **(U)** Markov chains in the sense of last section. This is where we reach to our sufficient but considerably restrictive (compared to the generality in the theorems of Chapter 3) condition **(D)**, which includes absolute continuity (with respect to the Haar measure) and some further assumptions on the probability measure which governs the random walk.

Let G be a locally compact second countable topological group and X a compact metrisable space with a continuous transitive action of G, in this chapter we refer to this case by “as usual". Let also μ be a Borel probability measure on G. As before, we denote by S_n the n^{th} step of the μ-random walk $S_n = Y_n. \dots .Y_1$, i.e. Y_i's are independent random variables with law μ. To this data, we associate the following Markov chain : we take the state space Σ to be the Cartesian product $G \times X$, and for $x \in X$, we inductively define the Markov chain on Σ by setting $Z_1^x = (Y_1, x)$ and for $n \geq 2$, $Z_n^x = (Y_n, S_{n-1}.x)$. It is clear that this stochastic process is a Markov chain in the usual sense. To avoid repetitions, we shall refer to that as random walk action Markov chain.

Remark 6.5. *Depending on the support supp(μ) of the probability measure μ, we can of course restrict the state space Σ of this Markov chain to $\Sigma_\mu :=$ supp$(\mu) \times X$. This will not modify anything in the following considerations. We will use this remark later on in our application (namely of contraction principle for LDP's).*

We now investigate under what conditions on μ, the transition kernel π of this Markov chain will satisfy the assumption **(U)**. For a Borel probability measure ν on X, denote by $\mu * \nu$, the convolution probability measure resulting from the action : it is defined as the probability measure satisfying for every $f \in C(X)$, $\int f(x)\mu*\nu(dx) :=$

$\int\int f(g.y)\mu(dg)\nu(dy)$. Furthermore, as usual, we denote by μ^{*k} the k^{th} convolution of μ, which can be also seen as being equal to $\mu * \ldots * \mu$ (k-times), with the action of G on itself, say, by left multiplication. For $k = 0$, we set $\mu^{*0} = \delta_e$, it will be compatible with our calculations.

We start by the following elementary lemma which expresses in a handy form the transition kernel π of the above chain. Let us first note the basic fact that, by definition, π verifies **(U)** if there exist $0 < l \leq N$ and $C > 0$ such that for all $(A, x), (B, y) \in \Sigma$ and for all Borel measurable $M \subset G$ and $O \subset X$, $\pi^l((A, x), M \times O) \leq C \sum_{k=1}^{N} \pi^k((B, y), M \times O)$.

Lemma 6.6. *For all $l \geq 1$, we have $\pi^l((A, x), M \times O) = \mu(M)\mu^{*(l-1)} * \delta_{A.x}(O)$.*

Démonstration. By independence of random walk increments, we have

$$\begin{aligned}\pi^l((A, x), M \times O) &= \mathbb{P}(Y_l \in M \text{ and } S_{l-1}.A.x \in O) = \mu(M).P(S_{l-1}.A.x \in O)\\ &= \mu(M)\mu^{*(l-1)} * \delta_{A.x}(O)\end{aligned}$$

where the last equality follows by definition of the convolution measure $\mu^{*(l-1)} * \delta_{A.x}$ on X. □

In particular, using the expression of this lemma, we see that the assumption **(U)** is satisfied in case there exist $N \geq l \geq 1$ and $C > 0$ such that for all $x, y \in X$, for all $g \in G$ and for all Borel set $O \subseteq X$, we have :

$$\mu^{*l} * \delta_x(O) \leq C \sum_{k=0}^{N} \mu^{*k} * \delta_y(O). \tag{6.1}$$

(note that the condition is never satisfied for $l = 0$.)

Remark 6.7. *Considering, for example, a particular case of interest to us where $G = GL(V)$ for a Euclidean space V, and $X = \mathbb{P}(V)$, one sees that (6.1) can possibly not be satisfied for probability measure μ of countable support and is easily seen to be never satisfied for probability measures of finite support. Therefore, as mentioned before, we shall be focusing on diffuse (non-atomic) measures.*

We now set the following natural definition which will be useful in our application. Let G be a group acting on a set X,

Definition 6.8. *A subset K of G is called X-transitive, if for all $x \in X$, we have $\{k.x \mid k \in K\} = X$.*

Remark 6.9. *We note that, under our topological hypotheses on G, X and the action, it follows in particular that one can always find a compact X-transitive set K in G.*

Now, let G, X and μ be as usual. We state our condition **(D)** on the probability measure μ (and on the action of G on X) which will ensure **(U)** for the associated Markov chain (see Lemma 6.10) :

Condition (D) : *There exists $n_0 \geq 1$ such that μ^{*n_0} is absolutely continuous with respect to the (right invariant) Haar measure of G (denote by f, its density function, $f := \frac{d\mu}{dHaar_G} \in L^1(G)$). There exist integers $N \geq l \geq 1$, a constant $C > 0$ and an X-transitive set K in G with the property that for all $t \in K$ and for (Haar) almost all $g \in G$, we have $f^{*l}(g) \leq C \sum_{k=1}^{N} f^{*k}(gt)$.*

Above, f^{*k} denotes the k^{th} convolution of f with itself with respect to the Haar measure of G : setting $f^{*0} = \delta_e$, for $k \geq 1$, we have $f^{*k}(x) = \int f^{*(k-1)}(xy)f(y^{-1})dy$, where dy stands for the Haar measure of G.

The following lemma basically follows from the right invariance of the Haar measure :

Lemma 6.10. *The Markovian kernel π satisfies **(U)** if μ satisfies **(D)**.*

Démonstration. Let $n_0 \geq 1$, $f \in L^1(G)$, $N \geq l \geq 1$, $C > 0$, $K \subset G$ be as in the condition **(D)**. One observes that using the density function f of μ^{*n_0}, the probabilities $\mu^{*ln_0} * \delta_x(O)$ appearing in the sufficient condition (6.1) writes as

$$\mu^{*ln_0} * \delta_x(O) = \int 1_O(h.x)\mu^{*ln_0}(dh) = \int 1_O(h.x)f^{*l}(h)dh \tag{6.2}$$

Now, for each pair of elements $(x, y) \in X^2$, fix an element $\gamma_{x,y}$ in the X-transitive set K such that $\gamma_{x,y}.y = x$. Then, using **(D)** and (6.2), for all $x, y \in X$ and Borel set $O \subseteq X$, we have

$$\begin{aligned}
\mu^{*ln_0} * \delta_x(O) &= \int 1_O(h.x)f^{*l}(h)dh \leq \int 1_O(h.x)C\sum_{k=1}^{N} f^{*k}(h\gamma_{x,y})dh \\
&= C\sum_{k=1}^{N}\int 1_O(h.x)f^{*k}(h\gamma_{x,y})dh = C.\sum_{k=1}^{N}\int 1_O((h\gamma_{x,y}y)f^{*k}(h\gamma_{x,y})dh \\
&= C\sum_{k=1}^{N}\int 1_O(h.y)f^{*k}(h)dh = C\sum_{k=1}^{N}\mu^{*kn_0} * \delta_y(O)
\end{aligned} \tag{6.3}$$

where we used the right invariance of the Haar measure of G, namely on the before last equality. Now, one observes that (6.3) clearly establishes (6.1), and hence, the lemma. □

As an immediate corollary of this lemma and Theorem 6.2, we note the following evident fact. We are in the setting described in the beginning of this section :

Corollary 6.11. *(of Theorem 6.2) For a random walk on the group G governed by a probability measure μ satisfying **(D)**, the sequence of laws $\mu_{n,x}$ of occupation times $L_n^{Z^x} := \frac{1}{n}\sum_{i=1}^{n}\delta_{Z_i^x}$ of the associated random walk action Markov chain $(Z_n^x)_{n\geq 1}$ on $\Sigma = G \times X$ satisfies an LDP with a proper convex rate function as in Theorem 6.2 and Proposition 6.4.* □

A class of density functions satisfying (D)

In this paragraph, we give an explicit (see also Remark 6.13) condition on the density function (and on the group action) with which, our condition **(D)** is satisfied. This is expressed in the following lemma. We use the above notation and properties for G, X, X-transitive compact set K (see Remark 6.9). Let μ be a Borel probability measure on G.

Lemma 6.12. *Suppose that there exists a positive integer n_0 such that μ^{*n_0} is absolutely continuous with respect to Haar measure on G with a compactly supported L^∞ density function f on G. Suppose that there exists a neighbourhood V of identity and a constant $\alpha > 0$ with the property that the compact sets K and $supp(f)$ are contained in the semigroup generated by V, and f satisfies $f_{|V} \geq \alpha$. Then, the condition **(D)** is satisfied.*

Démonstration. Set $K_0 := supp(f).K$, a compact subset of G by continuity of the group operation. Since the compact sets K and $supp(f)$ are contained in the semigroup $\cup_{n\geq 1} V^n$, it follows that there exists an integer $N_0 \geq 1$ such that $K_0 \subseteq (\overset{\circ}{V^{N_0}})$. On the other hand, observe that by the regularizing property of Haar measure, and supposing $N_0 \geq 2$ without loss of generality, the function $(\alpha \mathbb{1}_V)^{*N_0}$ is continuous. Moreover, it is clearly positive on $(\overset{\circ}{V^{N_0}})$ and in particular on the compact set K_0. By consequent, there exists a positive constant β such that we have $(\alpha \mathbb{1}_V)^{*N_0}_{|K_0} \geq \beta$. But since, $f \geq \alpha \mathbb{1}_V$, we have $f^{*N_0} \geq (\alpha \mathbb{1}_V)^{*N_0}$, and restricting this inequality on K_0, we indeed have $f^{*N_0}_{|K_0} \geq \beta$. Now, let $A > 0$ be a constant such that $||f||_\infty \leq A$. Then, since by construction of K_0, for each $t \in K$ and $g \in supp(f)$, we have $gt \in K_0$, it follows that for Haar almost all $g \in G$, we have $f(g) \leq \frac{A}{\beta} f^{*N_0}(gt)$.

Now, observe by (6.1), (6.2) and (6.3) that **(D)** is satisfied with $n_0 = n_0$, $l = 1$, $N = N_0$, $C = \frac{A}{\beta}$. □

Remark 6.13. *A class of cases where the hypothesis of the above lemma is satisfied and is more explicit goes as follows : suppose that the connected component of identity G° is open in G (for example, if G is locally connected), and that it acts transitively on X ; then, the hypothesis of the above lemma just writes as : there exists a neighbourhood V of identity and a constant $\alpha > 0$ such that the compact support of the L^∞ function f is contained in G° and it satisfies $f_{|V} \geq \alpha > 0$. Indeed, in this case, the semigroup generated by V automatically contains G°, which in turn contains, by transitivity, a compact X-transitive set K. As noted in the next section, this readily applies, for instance, to $G = GL(d, \mathbb{R})$, $X = \mathbb{P}(\mathbb{R}^d)$.*

6.2 LDP for Iwasawa decomposition and the rate function

Transferring the LDP with continuous cocycles

In this part, we introduce the cocycle maps and state our main result for the applications. It will consist of transferring the LDP for the laws of occupation times via continuous cocycles, by the contraction principle.

Let G be a group acting on a set X, and let W be a real vector space. A W-valued cocycle for this action is a mapping $\sigma : G \times X \longrightarrow W$ satisfying the cocycle property $\sigma(gg', x) = \sigma(g, g'x) + \sigma(g', x)$, for all $g, g' \in G$ and $x \in X$. Note that if the action of G on X is trivial, this is just a choice, for each $x \in X$, of a group homomorphism from G to the abelian group W.

Example 6.14. *Keeping up with our example of the introduction of this chapter for $G = GL(V)$, $X = \mathbb{P}(V)$ and $W = \mathbb{R}$, the following cocycle will be of interest to us :*

$$\sigma : GL(V) \times \mathbb{P}(V) \longrightarrow \mathbb{R}$$
$$(M, \bar{x}) \mapsto \log \frac{||Mx||}{||x||}$$

where x denotes a non-zero vector on the line $\bar{x}$. It is immediate to see that this map satisfies the cocycle property for the canonical action of $GL(V)$ on $\mathbb{P}(V)$.

Let now G and X be as usual, μ a Borel probability measure on G, W a Euclidean space and $\sigma : G \times X \longrightarrow W$ be a continuous cocycle. Let Σ_μ denote the restricted state space as in Remark 6.5 of the random walk action Markov chain and $M_1(\Sigma_\mu)$ the convex set of Borel probability measures on Σ_μ. Denote by T_σ, the map $T_\sigma : M_1(\Sigma_\mu) \longrightarrow W$ defined by $T_\sigma(\nu) := \int \sigma d\nu$. Note that T_σ is continuous in case μ is of compact support in G. This follows by definition of the topology on the set $M_1(\Sigma_\mu)$ of Borel probability measures on the Polish space Σ_μ, induced by Lévy-Prokhorov metric, which is equivalent to the topology of weak convergence. With these preliminary observations, we state the following proposition, the main result of this subsection, which will be important to us for our applications in mind (namely, with norm or more generally Iwasawa cocycle which we shortly introduce) :

Proposition 6.15. *Let G, X, W and σ be as in the previous paragraph. Suppose that μ is a Borel probability measure on G, satisfying the condition **(D)** (see also Lemma 6.12). Then, the sequence of W-valued random variables $\frac{1}{n}\sigma(S_n, x)$ satisfies an LDP, uniformly for $x \in X$, with a proper convex rate function I_σ on W.*

Remark 6.16. *As in Remark 6.3, uniformity in $x \in X$ means precisely the following : for all subset B of W, we have*

$$\begin{aligned} -\inf_{v \in \overset{\circ}{B}} I_\sigma(v) &\leq \liminf_{n\to\infty} \inf_{x\in X} \frac{1}{n} \log P(\frac{1}{n}\sigma(S_n, x) \in B) \\ &\leq \limsup_{n\to\infty} \sup_{x\in X} \frac{1}{n} \log P(\frac{1}{n}\sigma(S_n, x) \in B) \leq -\inf_{v\in\overline{B}} I_\sigma(v) \end{aligned}$$

Proof of Proposition 6.15. We denote as before by Z_n^x the random walk action Markov chain defined by $Z_n^x = (Y_n, S_{n-1}.x)$. With our hypotheses, it follows by Corollary 6.11 that the sequence of $M_1(\Sigma_\mu)$-valued random variables $L_n^{Z^x} = \frac{1}{n}\sum_{k=1}^n \delta_{Z_k^x}$ satisfies an LDP with a proper convex rate function, that we denote by I, $I : M_1(\Sigma_\mu) \longrightarrow [0, \infty]$. Now, one notes that, by cocycle property, for each $n \geq 1$, we have

$$\frac{1}{n}\sigma(S_n, x) = \frac{1}{n}\sum_{k=1}^n \sigma(Y_k, S_{k-1}.x) = \frac{1}{n}\sum_{k=1}^n \sigma(Z_k^x) = \int \sigma dL_n^{Z^x} = T_\sigma(L_n^{Z^x}) \tag{6.4}$$

where T_σ is the continuous map defined on the last paragraph.

Now, we apply the contraction principle (Lemma 3.28) by taking in it, $X = M_1(\Sigma_\mu)$, $Y = W$, $f = T_\sigma$. Since by (6.4), the laws of the random variables $\frac{1}{n}\sigma(S_n, x)$ are the push-forwards by T_σ of the laws $\mu_{n,x}$ of $L_n^{Z^x}$, it follows that the sequence $\frac{1}{n}\sigma(S_n, x)$ satisfies an LDP with a proper rate function that we denote by I_σ. The uniformity of LDP follows by Remark 6.3.

The convexity of I_σ results from the convexity of I, linearity of T_σ and the expression of I_σ in terms of those two, given by Lemma 3.28 : $I_\sigma(v) = \inf\{I(x) \mid T_\sigma(x) = v\}$ (see also the following remark). □

Remark 6.17. *Note that Proposition 6.4 gives an alternative expression for I by the action of the Markov kernel on the functions in $C(\Sigma_\mu)$. Therefore by using it in contraction principle, we have an alternative expression for I_σ.*

LDP for norm and Iwasawa cocycles

This section consists of some applications of the previous results of this chapter. Namely, we shall apply these in two particular setups (for G and X), one with the norm cocycle and the other with the Iwasawa cocycle. The first setup is in fact the example we used throughout the preceding sections to illustrate our definitions and hypotheses. For the second, we first briefly describe the Iwasawa cocycle and then discuss our application.

The norm cocycle

With the notation of the previous section, for a Euclidean space V (identify with an $\mathbb{R}^d$), let $GL(V)$ be the locally compact second countable group G, $\mathbb{P}(V)$ be, with its usual topology, the metrisable compact set X and consider the natural action of $GL(V)$ on $\mathbb{P}(V)$ which is of course continuous and transitive. Let also μ be a Borel probability measure on $GL(V)$ satisfying the condition **(D)**. At this point, we would like to point out that the compact subgroup $SO(V)$ of $GL(V)$ contained in $GL(V)^\circ = GL^+(V)$ is a $\mathbb{P}(V)$-transitive set and therefore - as noted in Remark 6.13 - by Lemma 6.12, **(D)** is for example satisfied whenever μ is an absolutely continuous probability measure of compact support in $GL^+(V)$ and with density bounded below by a strictly positive constant on a neighbourhood of the identity in $GL(V)$. The following result follows from Proposition 6.15 by considering the norm cocycle σ defined in Example 6.14, which is clearly continuous.

Corollary 6.18 (of Proposition 6.15)**.** *In the setting of the previous paragraph, for all $x \in \mathbb{R}^d$, the sequence of random variables $\frac{1}{n}\log ||S_n x||$ satisfies an LDP with a proper convex rate function I on $\mathbb{R}$.* □

Remark 6.19.

1. *The rate function I does not depend on $x \in V \setminus \{0\}$. Furthermore, for any bounded set C in V, we can strengthen the LDP inequalities, as in Remark 6.16, to include suprema and infima over $x \in C \setminus \{0\}$.*
2. *Similar to Corollary 4.28 and Proposition 4.32, it follows by convexity of I and* 1. *that in fact, for any bounded subset C of V and any subset B of $\mathbb{R}$*

intersecting the interior of the effective support D_I of I (see the last part for a discussion), and satisfying $\overset{\circ}{\overline{B}} = \overline{B}$, one has

$$\lim_{n\to\infty} \sup_{\substack{x\in C\\ x\neq 0}} \frac{1}{n} \log \mathbb{P}(\frac{1}{n}\log ||S_n x|| \in B) = \lim_{n\to\infty} \inf_{\substack{x\in C\\ x\neq 0}} \frac{1}{n} \log \mathbb{P}(\frac{1}{n}\log ||S_n x|| \in B)$$
$$= -\inf_{x\in B} I(x)$$

Remark 6.20. *Before moving on to discuss the Iwasawa cocycle, we would like to point at the following relation between our previous result and Le Page's Theorem 4.30 : recall that under far less restrictive assumptions (of strong irreducibility and proximality, see Section 4.2) than **(D)**, Le Page obtained Theorem 4.30, in which he in fact gives an "LDP around the Lyapunov exponent ". Consequently, we see that our convex rate function I of Corollary 6.18 with values in $\mathbb{R}\cup\{\infty\}$ extends the function ϕ of Theorem 4.30, defined around the Lyapunov exponent $\lambda_1(\mu)$, to the entire $\mathbb{R}$ while preserving the LDP control properties : for $B > 0$ as in that result and for all $x \in \mathbb{R}$ satisfying $|\lambda_1(\mu) - x| < B$, we have $-\phi(x) = I(x - \lambda_1(\mu))$.*

The Iwasawa cocycle

In this discussion, we closely follow the exposition of Benoist-Quint in [14]. To minimise the technical preliminaries, we will work in the same setting as in Chapter 3 ; in the rest of this part, G will denote a connected semisimple linear real algebraic group.

We first introduce the Iwasawa decomposition for such a group G, another classical decomposition as the previously introduced Cartan and Jordan decompositions. For the following, we wish to recall Example 2.14, in which we explicitly described situation for $SL(d,\mathbb{R})$. One will remark that, in particular, for $SL(d,\mathbb{R})$ the Iwasawa decomposition corresponds basically to the Gram-Schmidt orthogonalisation of a basis. We now take over the discussion in Section 2.2 following Remark 2.13.

Recall that $\mathfrak{g}$ denotes the Lie algebra of G, $\mathfrak{a}$ a Cartan subalgebra in $\mathfrak{g}$ and $A_G = \exp(\mathfrak{a})$. Let as before R denote the set of restricted roots of G, $\pi \subset R$ be a choice of simple roots and R^+ denote the corresponding positive roots. For each $\alpha \in R$, denote by $\mathfrak{g}_\alpha$, the root space of α, the subspace $\{X \in \mathfrak{g} \mid Ad(a)X = \alpha(a)X \text{ for all } a \in A_g\}$ of $\mathfrak{g}$. For simplicity, let us also assume that G is *split* so that $\mathfrak{a}$ is equal to its centraliser in $\mathfrak{g}$.

Let $\mathfrak{p}(\pi) = \mathfrak{p}$ be the Borel subalgebra (minimal parabolic subalgebra) associated to our choice of simple roots, $p = \mathfrak{a} \oplus \bigoplus_{\alpha\in R^+} \mathfrak{g}_\alpha$. Let $\mathfrak{u}$ denote the derived subalgebra of $\mathfrak{p} : \mathfrak{u} = \bigoplus_{\alpha\in R^+} g_\alpha$. It is automatically a (maximal) nilpotent subalgebra in $\mathfrak{g}$. One other way to see the nilpotency is to recall the property that $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha+\beta}$ for $\alpha, \beta \in R$. Let U be the connected algebraic subgroup of G with Lie algebra $\mathfrak{u}$. Correspondingly, it is a maximal unipotent subgroup in G. Finally, choose a maximal compact subgroup K of G.

With these notations, the Iwasawa decomposition of G writes $G = KA_GU$. It also enjoys the nice property that the map $K \times A_G \times U \longrightarrow G$, associating to (k, a, u), the product kau is a homeomorphism.

We can now introduce the Iwasawa cocycle. For this, let P be the Borel subgroup corresponding to the choice of Borel subalgebra $\mathfrak{p}$. In the Iwasawa decomposition, one has $P = MA_GU$, where $M = K \cap C_G(A)$, the latter standing for the centraliser. The homogeneous space $\mathcal{F}_G = G/P$ is called the flag variety of G. With the quotient topology, it is compact and endowed with a continuous action of G, by 'multiplication on the left'. We define the Iwasawa cocycle σ using this action, and the Iwasawa decomposition as follows : for $(g, \eta) \in G \times \mathcal{F}_G$, letting $k \in K$ be an element such that $\eta = kP$, we set $\sigma(g, \eta)$ to be the element of $\mathfrak{a}$ satisfying $gk \in K \exp(\sigma(g, \eta))U$. The mapping σ is well-defined and continuous by the above (homeomorphism) property of the Iwasawa decomposition.

The following lemma justifies the terminology for σ. We take it from [14] (Lemma 5.29) :

Lemma 6.21 ([14]). *$\sigma : G \times \mathcal{F}_G \longrightarrow \mathfrak{a}$ satisfies, for all $g, h \in G$ and $\eta \in \mathcal{F}_G$, $\sigma(gh, \eta) = \sigma(g, h.\eta) + \sigma(h, \eta)$.*

Observe that we are now in our usual setting to apply Proposition 6.15 : we possess a group G, a compact space $\mathcal{F}_G$ as in that result, with a continuous action of G on $\mathcal{F}_G$ and a continuous cocycle σ for that action. But before that evident application, let us note an extension of Lemma 2.16 which relates the norm and Iwasawa cocycles. Together with Lemma 2.15, it allows us to interpret the Iwasawa cocycle as a multidimensional generalisation of the norm cocycle.

To this aim, let, as in Lemma 2.16, (V, ρ) be an irreducible rational representation of G with highest weight χ. As before, let V_χ denote the highest weight space $\{v \in V \mid \rho(a)v = \chi(a)v \text{ for all } a \in A_G\}$. V_χ is also equal to the set of fixed points in V of the G-action restricted to the maximal unipotent subgroup U of G, $V_\chi = \{v \in V \mid \rho(u)v = v\}$ (see 5.8 in [14]). It readily follows, from this, that the application associating to a $g \in G$, the subspace $\rho(g)V_\chi$ of V descends to a map from the flag variety $\mathcal{F}_G$ of G to the Grassmannian $\mathbb{G}_p(V)$, where $p = \dim V_\chi$ (recall that by definition, for a proximal ρ, we have $p = 1$, so that $\mathbb{G}_p(V) = \mathbb{P}(V)$ and for a split G, $\dim V_\chi$ is always one). Consequently, for $\eta \in \mathcal{F}_G$, we denote by V_η the subspace $\rho(g)V_\chi$, where g is any element of G satisfying $\eta = gP$. With these notations, we have the following extension of Lemma 2.16 that we take from [14] (Lemma 5.33) :

Lemma 6.22 ([14]). *The Euclidean norm of Lemma 2.16 can be chosen to additionally satisfy $\overline{\chi}(\sigma(g, \eta)) = \log \frac{||\rho(g)v||}{||v||}$ for all $(g, \eta) \in G \times \mathcal{F}_G$, and $v \in V_\eta$, and where as usual, $\overline{\chi} = \log \circ \chi \circ \exp \in \mathfrak{a}^*$*

Now, the application of Proposition 6.15 to the Iwasawa cocycle reads :

Corollary 6.23. *Let G be as before, μ be a Borel probability measure on G satisfying the condition **(D)**, and σ denote the Iwasawa cocycle for G. Then, for the μ-random* *walk S_n, the sequence of random variables $\frac{1}{n}\sigma(S_n, \eta)$ satisfies an LDP with a proper convex rate function on $\mathfrak{a}$, uniformly for $\eta \in \mathcal{F}_G$.* □

Remark 6.24. *1. Uniformity is to be understood as in Remark 6.16.*

2. The observation on the existence of limits for large deviation probabilities as in 2. of Remark 6.19 applies to this result as well.

Properties of the rate functions

We start by studying briefly the above rate function of the norm cocycle LDP, mostly the interplay between its effective support and the joint spectral radius of the support of the probability measure in question. We then transfer these observations, to the rate function of the Iwasawa cocycle LDP, using the relation between the two cocycles. Finally, we suggest a further notion of interest in relation with the joint spectra and the rate function of Iwasawa cocycle LDP.

• In the first place, let G, μ and I be as in Corollary 6.18. Denote by S_μ the support of μ, $S_\mu = \{x \in G \mid \mu(O) > 0$ for each neighbourhood O of x in $G\}$. Note that by its definition, S_μ is closed in G. For a subset T of G, denote by T^{-1} the set $\{g^{-1} \mid g \in T\}$. Recall also that $r(T)$ and $r_{sub}(T)$ denote, respectively, the joint spectral radius and joint spectral subradius of T.

Our main observation on the norm cocycle LDP rate function (i.e. that of Corollary 6.18) is the following proposition. One should compare it with Proposition 4.32 (and parallelly, compare Fig. 4.1 with Fig. 6.1). For 2. below, note that for any bounded set $T \subset GL(V)$, by definitions, we have $-\log r(T^{-1}) \leq -\log r_{sub}(T^{-1}) \leq \log r(T)$.

Proposition 6.25. *I admits a unique zero on the first Lyapunov exponent $\lambda_1(\mu)$ of μ, and we have*

1. $\lambda_1(\mu) < \log r(S_\mu)$.
2. *D_I is an interval of non-empty interior contained in* $[-\log r(S_\mu^{-1}), \log r(S_\mu)]$
3. $\log r(S_\mu) \in \partial D_I$*, of course, it is the upper boundary point of the interval.*

Remark 6.26. *1. Note first that under the assumptions of Corollary 6.18, we have, in particular, $\overset{\circ}{S}_\mu \neq \emptyset$; and therefore, it follows that $\log r_{sub}(S_\mu) < \log r(S_\mu)$ (think of eigenvalues and the usual Jordan decomposition). By the obvious relation between $||gv||$ and $||g||$, where $g \in GL(d, \mathbb{R})$ and $v \in \mathbb{R}^d$, it is suggested by Proposition 4.26 and its 'contracted version' Remark 4.33, that for the effective support of the rate function I, in fact, the following inclusion holds : $]\log r_{sub}(S_\mu), \log r(S_\mu)[\subseteq D_I$. It also seems probable that Abels-Margulis-Soifer type of finiteness result (see Theorem 2.24 and subsequent Remark 2.25), together with Benoist's estimates (Theorem 2.21) can be used to study such an inclusion. This, and relatedly, other properties of the Iwasawa rate function (i.e. that of Corollary 6.23), will be investigated in a future study (see also Remark 6.30).*

2. *More generally, a point of further interest concerns the quantity* $\inf\{x \in \mathbb{R} \mid x \in D_I\}$*. Recall that in Proposition 4.32, the corresponding quantity is equal to* $\log r_{sub}(S_\mu)$*. A question is whether this is also the case for* I_n *and if not, for example, whether this quantity only depends on the support of* μ*.*
3. *In this connection, we would like to recall the following result, valid in a more general setup, which expresses a close relation between the asymptotic behaviours of* $\frac{1}{n}\log||S_n||$ *and* $\frac{1}{n}\log||S_n v||$ *: it follows from Le Page's result, Theorem 4.30 (see also, Aoun's work [5], Proposition 2.4.14) that for every* $\epsilon > 0$*, there exists* $0 < \rho < 1$ *with* $\mathbb{P}((\log||S_n|| - \log||S_n v|| > \epsilon n)) = O(\rho^n)$*.*

Proof of Proposition 6.25. That the effective support is an interval follows directly from the convexity of the rate function I. The unique-zero property for I, and that the effective support is of non-empty interior are direct consequences of Le Page's Theorem 4.30, which of course applies to our situation. The inclusion $\overline{D_I} \subseteq [-\log r(S_\mu^{-1}), \log r(S_\mu)]$ follows readily by definition of joint spectral radius together with the obvious fact that for all $v \in V \setminus \{0\}$ and $g \in GL(V)$, we have $||g^{-1}||^{-1} \leq \frac{||gv||}{||v||} \leq ||g||$. Finally, 3. assertion will follow from the following two lemmata, which in turn, will imply 1. by 2. $\square$

We remind the reader that in the following two lemmata, even though we are in the setting of Corollary 6.18, the first one is valid whenever it makes sense, i.e. a rate function exists, and the second one is valid for any probability measure μ on $GL(V)$.

Lemma 6.27.

$$\sup\{\alpha \in \mathbb{R} \mid I(\alpha) < \infty\} = \inf\{\alpha \in \mathbb{R} \mid \limsup_{n\to\infty} \frac{1}{n}\log\mathbb{P}(\frac{1}{n}\log||S_n|| > \alpha) = -\infty\}$$

Démonstration. This follows easily by similar considerations as in the proof of Proposition 4.31 (see (4.19) there). To avoid repetitions, we omit the details. $\square$

As mentioned above, the following lemma applies to any probability measure μ on $GL(V)$: If μ is of finite support, it is easier and follows the same as 3. of Proposition 4.26. For the general case, we make use of a recent result of Bochi-Morris [26], which we combine with Berger-Wang's Theorem 4.1 (see also the paragraph following Remark 4.34).

Lemma 6.28. *We have*

$$\log r(S_\mu) = \inf\{x \in \mathbb{R} \mid \limsup_{n\to\infty} \frac{1}{n}\log\mathbb{P}(\frac{1}{n}\log||S_n|| > x) = -\infty\}$$

Démonstration. The inequality $\geq$ is clear by definition of joint spectral radius. For the other inequality, fix an $\alpha < \log r(S_\mu)$. Then by Theorem 4.1, there exist $n_0 \in \mathbb{N}$ and $g \in S_\mu^{n_0}$ with $\frac{1}{n_0}\log\lambda_1(g) > \alpha$. In particular, the singleton $\{g\}$ has joint spectral subradius larger than $e^{\alpha n_0}$. Now, by continuity of joint spectral subradius ([26]), there exists a neighbourhood U_g of g such that $\frac{1}{n_0}\log r_{sub}(U_g) > \alpha$. Write $g = g_{n_0}.\dots.g_1$ with g_i's in S_μ. By continuity of the group operation, there exist neighbourhoods V_i of g_i for $i = 1,\dots,n_0$ such that $V_{n_0}.\dots.V_1 \subseteq U_g$. In particular, for all $k \in \mathbb{N}$ large enough and $h_k \in (V_{n_0}.\dots.V_1)^k$, we have $\frac{1}{n_0}\log||h_k|| > \alpha$. Set then $\alpha_i = \mu(V_i) > 0$ for

$i = 1, \ldots, n_0$. Now using the independence of random walk increments, for all $k \in \mathbb{N}$ large enough, we have

$$\mathbb{P}(\frac{1}{n_0 k} \log ||S_{n_0 k}|| > \alpha) \geq \prod_{i=1}^{k-1} P(X_{(i+1)n_0} \in V_{n_0}, \ldots, X_{in_0+1} \in V_1) \geq (\alpha_{n_0} \ldots, \alpha_1)^k \tag{6.5}$$

Therefore, in (6.5), taking log and dividing by $n_0 k$ and taking lim sup, we get that $\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(\frac{1}{n} \log ||S_n|| > \alpha) \geq \frac{1}{n_0} \sum_{i=1}^{n_0} \log \alpha_i > -\infty$, so that $\alpha \leq \inf\{x \in \mathbb{R} \mid \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(\frac{1}{n} \log ||S_n|| > x) = -\infty\}$. Since $\alpha < \log r(S_\mu)$ is arbitrary, this proves the other inequality and hence equality of the lemma. □

- In the second place, let now G, μ and I be as in Corollary 6.23, S_μ denote the support of μ, d be the real rank of G, and ρ_i for $i = 1, \ldots, d$ denote the distinguished rational representations of the semisimple group G, of highest weights $\overline{\chi}_i$, given by Lemma 2.15. The study of Iwasawa cocycle LDP rate function I follows, for several aspects, from the study of norm cocycle LDP rate function, using Lemma 6.22. However, due to the multidimensional feature of the effective support of I (when G is of higher rank), further points of interest emerge (see also Remark 6.30 and Fig. 6.1 below). In this final part, we content with the following proposition, which follows as an immediate corollary of 2. and 3. of Proposition 6.25, Lemma 6.22 and Lemma 2.15 :

Proposition 6.29. *In the setting of the previous paragraph, the rate function $I : \mathfrak{a} \longrightarrow [0, \infty]$ has a unique zero, its effective support $D_I = \{x \in \mathfrak{a} \mid I(x) < \infty\}$ is of non-empty interior and satisfies*

1. $$D_I \subseteq \bigcap_{i=1}^{d} \{x \in \mathfrak{a} \mid -\log r_{\rho_i}(S_\mu^{-1}) \leq x \leq \log r_{\rho_i}(S_\mu)\}$$

2. $\overline{D}_I \cap \{x \in \mathfrak{a} \mid \overline{\chi}_i(x) = \log r_{\rho_i}(S_\mu)\} \neq \emptyset$ *for each ρ_i for $i = 1, \ldots, d$.*

Démonstration. According to the paragraph which precedes the statement of the proposition, the only point that needs to be clarified is the unique zero property. But this follows similarly as in Proposition 4.31 from Le Page's result, or more directly from Benoist-Quint's Theorem 12.11.(iii) in [14]. □

Based on the previous proposition and in the same setting, below, we include a suggestive picture for the effective support of I, for $G = SL(3, \mathbb{R})$ (Compare with Fig. 4.1).

Remark 6.30. *For each point $\eta \in \mathcal{F}_G$ of the flag variety of G, one can define, using the Iwasawa cocycle $\sigma(., \eta)$ in a similar way as the Cartan $\kappa(.)$ and Jordan $\lambda(.)$ projections, to define the joint Iwasawa spectrum in $\mathfrak{a}$ of a bounded subset S of G. In case S generates a semigroup Γ Zariski dense in G, one can expect to show some nice properties of the joint Iwasawa spectrum, such as its independence of $\eta \in \mathcal{F}_G$ and geometric properties as for the joint spectra of S. We would like finally to speculate that one may establish a correspondence between the joint Iwasawa spectrum of S and the notion of limit set of the semigroup Γ, although they do not, a priori, live on the same space as the do the Benoist cone of Γ and joint spectrum of S generating Γ.*

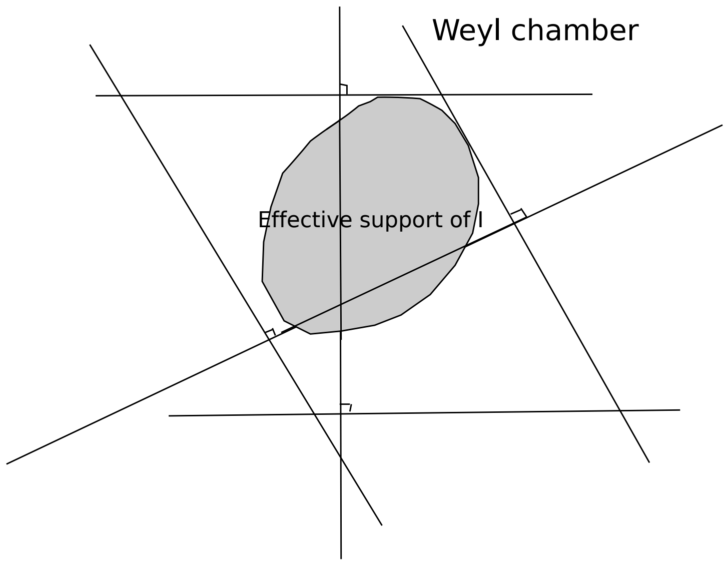

FIGURE 6.1 – The orthogonals to the walls of the Weyl chamber correspond to the hyperplanes described in 2. of Proposition 6.29.

Deuxième partie

Rigidity Results For Spectral Radius On Free Groups

Chapitre 7

7.1 Introduction : a '99 percent result' and a '1 percent question' of Breuillard

This text has two purposes : the first one is to discuss our partial results around a question of E. Breuillard about the rigidity of spectral radius of random walks on linear groups, in the case of a free group. This question and a related result of Breuillard will be introduced in this first Section 7.1. We explain our partial results in Section 7.3 (a quick description of our results can be found in the introduction of this section). The second purpose of this text, initially motivated by our tentative to solve the aforementioned question, is to develop a geometric understanding of a group acting by hyperbolic automorphisms on a tree. This will be mainly done in Section 7.2. This understanding will be used on one hand to prove results on geometric rigidity of spectral radius, and on the other hand, to obtain some algebraic results on the free group itself. This last one is an ongoing project, we only give a glimpse of it in Section 7.2 and Appendix A.1.

Let us start with a discussion on spectral radius : let Γ be a finitely generated group and μ be a symmetric probability measure on Γ. Here symmetric means that for each $g \in \Gamma$, one has $\mu(g) = \mu(g^{-1})$. μ is said to be adapted if its support $\{\gamma \in \Gamma \mid \mu(\gamma) > 0\}$ generates Γ. Let $S_n = X_n.\ldots, X_1$ denote the n^{th} step of the μ-random walk on Γ, i.e. X_i's are independent identically distributed Γ-valued random variables with law μ. The law of S_n will be denoted by μ^{*n}. In this setting, one notes that the laws of the left $(X_n.\ldots, X_1)$ and the right $(X_1.\ldots.X_n)$ μ-random walks are the same for each step. μ^{*n} is also called the n^{th} convolution power of μ and can be defined inductively by $\mu^{*(n+1)}(g) = \sum_{x\in\Gamma} \mu^{*n}(gx)\mu(x^{-1})$ for $n \in \mathbb{N}$ ($\mu^{*0} = \delta_e$, e standing for the identity element in Γ). To understand the dispersion of probability mass along a random walk, one notable quantity to look at is the decay rate of the return probability to identity $\mu^{*n}(e)$ (see also 3. of Remark 7.2).

Before moving on, let us also look at this from a different perspective : let $l^2\Gamma$ denote the Hilbert space of complex valued l^2 functions on Γ. It admits an orthonormal basis consisting of characteristic functions on elements of Γ : $\{\delta_\gamma \mid \gamma \in \Gamma\}$. Γ acts (on the left) naturally on this Hilbert space by unitary transformations : for $\gamma \in \Gamma$ and $f \in l^2\Gamma$, $(\gamma.f)(x) = f(\gamma^{-1}x)$. This action defines the left regular representation of Γ that we denote by $\lambda : \Gamma \longrightarrow \mathcal{U}(l^2\Gamma)$. Extending this last mapping linearly, we have a $*$-homomorphism from the complex group algebra $\mathbb{C}[\Gamma]$ of Γ into the Banach space $\mathcal{B}(l^2\Gamma)$ of bounded endomorphisms of $l^2\Gamma$, where the involution on $\mathbb{C}[\Gamma]$ is given by extending $\alpha.\delta_g \to \bar{\alpha}\delta_{g^{-1}}$ linearly. We shall denote this extended map by λ as well and let $||.||_{l^2\to l^2}$ denote the (operator) norm of $\mathcal{B}(l^2\Gamma)$. The completion $\overline{\lambda(\mathbb{C}[\Gamma])}$ of the image of $\mathbb{C}[\Gamma]$ in $\mathcal{B}(l^2\Gamma)$ has a C^*-algebra structure induced by $\mathcal{B}(l^2\Gamma)$, and it is denoted by $C^*_r(\Gamma)$ and called the reduced C^*-algebra of the group Γ. A probability measure $\mu = \sum_{g\in\Gamma}\mu(g)\delta_g \in \mathbb{C}[\Gamma]$ on Γ can be seen as an element of $C^*_r(\Gamma)$ by identifying it with its image $\lambda(\mu) = \sum_{g\in\Gamma}\mu(g)\lambda(g)$, the averaging operator with respect to μ. The spectral radius of the μ-random walk on Γ (or of the probability measure μ) is defined as the norm of $\lambda(\mu)$ in $C^*_r(\Gamma)$, i.e. the operator norm of $\lambda(\mu)$ on $l^2\Gamma$. We shall denote it by $r(\mu)$. One notes that μ is symmetric if and only if $\lambda(\mu)$ is a self-adjoint element of $C^*_r(\Gamma)$, so that $r(\mu)$ is actually the spectral radius of the operator $\lambda(\mu)$.

The following theorem of Kesten [76] relates the last two paragraphs and says that the exponential decay rate of the return probability is expressed by the spectral radius :

Theorem 7.1 (Kesten [76])**.** *Let μ be a symmetric probability measure on Γ. We have*

$$r(\mu) = \lim_{n\to\infty}\mu^{*2n}(e)^{\frac{1}{2n}}$$

Remark 7.2.
1. *The limit indeed exists in view of the obvious inequality $\mu^{*(n+m)}(e) \geq \mu^{*n}(e)\mu^{*m}(e)$, and it is in fact equal to $\sup_{n\geq 1}\mu^{*2n}(e)^{\frac{1}{2n}}$ and $\limsup_{n\to\infty}\mu^{*n}(e)^{\frac{1}{n}}$. In particular, one gets lower bounds to the operator norm $||\lambda(\mu)||_{l^2\to l^2} = r(\mu)$ by the return probabilities on finite steps, i.e. for each $n \geq 1$, we have $\mu^{*2n}(e)^{\frac{1}{2n}} \leq r(\mu)$ (this is sometimes referred to as the Kesten's bound).*
2. *Kesten has also calculated the spectral radius $r(\mu_q)$ of a uniform probability measure μ_q on a free generating set $\{a_1^{\pm},\dots,a_q^{\pm}\}$ of the free group F of rank q, it writes $r(\mu_q) = \frac{\sqrt{2q-1}}{q}$ (see [76]). It is also easy to see by probabilistic considerations that among the symmetric probability measures on groups, supported on $2q$ elements, the one with the smallest spectral radius is the one that is supported on a free set of elements $\{a_1^{\pm},\dots,a_q^{\pm}\}$ as before. Moreover, it was proved by Kesten in [76] that for a symmetric probability measure μ supported on $2q$ elements, one has $r(\mu) = \frac{\sqrt{2q-1}}{q}$ if and only if the support of μ is a free set of elements. By consequent, as it is also easily observed that one always has $r(\mu) \leq 1$, for any symmetric probability measure μ supported on $2q$ elements, one has $\frac{\sqrt{2q-1}}{q} \leq r(\mu) \leq 1$. One notes that the lower bound tends to zero, as q tends to infinity.*
3. *One remarks that the convolution powers of μ is simply its powers as an element of $\mathbb{C}[\Gamma]$ so that for $n \in \mathbb{N}$, one has $\lambda(\mu^{*n}) = \lambda(\mu)^n$. One also observes that for all $g,h,x \in \Gamma$, we have $<\lambda(g)\delta_h,\delta_x> = (\lambda(g)\delta_h)(x) = \delta_{gh}(x)$, so that one can write $\mu(g) = (\lambda(\mu)\delta_e)(g) = <\lambda(\mu)\delta_e,\delta_g>$. In view of $r(\mu) \leq 1$ and the fact that $\lambda(\mu)$ is a self-adjoint operator for a symmetric μ, using these former identities, one gets that for all $n \geq 1$, we have $\mu^{*2(n+1)}(e) \leq \mu^{*2n}(e)$ and for all $x \in \Gamma$, $\mu^{*2n}(x) \leq \mu^{*2n}(e)$. We omit the details of these simple facts (for*

a rapid treatment from which we are also inspired, see the lecture notes from a course of Breuillard [37], for an extensive treatment of several aspects of random walks on free groups, see [82], for a particular and in-depth treatment of spectral radius and related quantities, see [95], for a for a more general treatment of random walks on countable groups, see [117] and [94]).

At a first approach, it may be tempting to presume that the exponential decay rate of return probability will be determined by, or at least closely related to the exponential growth rate of number of elements in $\text{supp}(\mu^{*n}) = (\text{supp}(\mu))^n$, for a finitely supported symmetric μ. There is certainly a relation between these two : denoting by S the support of μ and using 3. of the previous remark, one sees that for each $n \in \mathbb{N}$, we have $1 = \mu^{*2n}(\Gamma) = \sum_{\gamma \in \Gamma} \mu^{*2n}(\gamma) \leq |S^{2n}|.\mu^{*2n}(e)$, so that one gets $\frac{1}{|S^{2n}|^{\frac{1}{2n}}} \leq \mu^{*2n}(e)^{\frac{1}{2n}} \leq r(\mu)$. In particular, taking limits as n tends to infinity, $r(\mu) \geq \frac{1}{v_S}$, where v_S is the exponential growth rate of S, i.e. $v_S := \lim_{n\to\infty} |S^n|^{\frac{1}{n}}$. We note at this moment that we would like to think of this and 2. of the previous remark, as the first observations on the rigidity of the spectral radius of a probability measure. In the sequel of this introduction, we shall see more and more intricate rigidity statements/questions on spectral radius, where one end will be the 1 percent question of Breuillard (see below). Continuing our observation, recalling that $r(\mu) \leq 1$, one deduces as a corollary that for any finitely supported symmetric probability measure μ on a group Γ of subexponential growth (i.e. $v_S = 1$), we have $r(\mu) = 1$. Considering amenable groups of exponential growth, 1. of the following striking result of Kesten says in particular that the relation between $r(\mu)$ and the exponential growth is a delicate one :

Theorem 7.3. *(Kesten [76], [77]) Let Γ be a finitely generated group and μ an adapted symmetric probability measure of finite support on Γ.*

1. *Γ is amenable if and only if $r(\mu) = 1$.*
2. *Let H be a normal subgroup of Γ and $\mu_{\Gamma/H}$ denote the push-forward of μ by the projection $\Gamma \twoheadrightarrow \Gamma/H$. Then, $r(\mu) = r(\mu_{\Gamma/H})$ if and only if H is amenable.*

Remark 7.4. *For a recent generalisation of 2. of the previous theorem to invariant (by conjugation) probability measures (sometimes called invariant random subgroups (IRS), in a possible misleading manner) on the set of subgroups of a countable group Γ by Abért-Glasner-Virág, see [1].*

Remark 7.5. *Continuing to speculate about the relation between the exponential growth and the spectral radius, one realises that, at a more profound level, one is in fact interested in understanding the relation between the sequence of probability measures 'of probabilistic origin', i.e. the convolution powers $\mu, \mu^{*2}, \ldots$ and the sequence of probability measures 'of deterministic origin', i.e. the uniform probability measures on S^n, where S denotes the support of μ (we wish to underline that in this setting the difference between the probabilistic and deterministic approaches lies basically in the way of counting, respectively, with and without multiplicities). One important direction in this study is to introduce and analyse other related quantities such as the asymptotic entropy $h(\mu)$ of μ, which in a sense measures the distribution rate of probability mass among the elements of S^n (introduced by Furstenberg [58] and Avez [6])* *and linear drift* l_μ *of* μ *indicating the average word length of the* μ*-random walk* S_n *(first considered by Kesten in [76]), and through these, get a more detailed description of this relation. On this occasion, we mention a fundamental inequality due to Guivarc'h stating that one has* $h(\mu) \leq l_\mu \log v_S$*. For a more detailed discussion of these, we refer the reader to the previous references, as well as to Kaimanovich-Vershik's [75], Vershik's [115] and for a more recent work, to Gouëzel-Mathéus-Maucourant's [62] (see its introduction). However, we shall pursue our study in a different direction.*

Remark 7.6 (Rapid decay)**.** *In this remark, let us mention another direction of study of rigidity of spectral radius, rapid decay property, somewhat closer to our considerations. After Powers' proof in [96] of* C^**-simplicity and unique trace property of the reduced* C^**-algebras of free groups* F_q *based on estimations of spectral radius of some particular probability measures on free groups, there has been an interest in the study of spectral radius, see for example Akemann-Ostrand [2]. This has been continued with Haagerup's use of his estimations of norms of operators of particular type on* l^2F_q *(i.e. norms in* $C^*_rF_q$*) to give a first example of a non-nuclear* C^**-algebra with metric approximation property, namely* $C^*_rF_q$*, in [70]. One aspect of Haagerup's estimations was singled out in a general set-up as the rapid decay proprty (RD) by Jolissaint, who gave a first account of this in [73]. Several equivalent and slightly more general formulations of (RD) is possible ; we content with the following : let* Γ *be a finitely generated group endowed with a word metric coming from a finite generating symmetric set* $S \subset \Gamma$*.* Γ *is said to have (RD) if there exists a polynomial* $P \in \mathbb{R}[X]$ *such that for all* $n \in \mathbb{N}^*$ *and* $f \in \mathbb{R}_+\Gamma$ *whose support* $\{\gamma \in \Gamma \mid f(\gamma) \neq 0\}$ *is contained in the ball of radius* n *in* Γ*, one has* $||\lambda(f)||_{l^2 \to l^2} \leq P(n)||f||_2$*. (RD) turned out to be useful in several areas ; we refer the reader to a recent survey of Sapir [107] and references therein (for earlier accounts, see Bekka-Cowling-de la Harpe [17] and Chatterji-Ruane [40]). We make two immediate observations on the groups with (RD) related to our considerations. First, one sees from the definition that for any* Γ *with (RD) and of exponential growth (in particular,* F_q*), if* S *is a finite symmetric generating set and* $\nu_n := \frac{1}{|S^n|}\sum_{\gamma \in S^n} \delta_\gamma$ *is the uniform probability measure on* S^n*, then* $r(\nu_n)$ *decays exponentially fast as* n *tends to infinity. The second one is in a similar spirit with our considerations in this work : if a uniform probability measure* μ *is supported on* q *elements (large enough depending on* Γ*) has its spectral radius* $r(\mu)$ *larger than some* $\epsilon > 0$*, then the support of* μ *must contain elements of word length larger than* $c_\epsilon q^{\frac{1}{d}}$ *for some* $d \in \mathbb{N}$ *and* $c_\epsilon > 0$ *(although a mere direct one, one notes the weakness of this conclusion under the also somewhat weak assumption* $r(\mu) \geq \epsilon > 0$*. Compare with 1 percent question and our results in Section 7.3.*

After this small digression and before stating the 99 percent result on spectral radius, let us continue with a simple observation and the subsequent uniform spectral gap theorem. Suppose that μ is a finitely supported uniform symmetric probability measure on a group Γ such that there exist $a, b \in S :=$supp(μ) generating (with their inverses) a non-commutative free group in Γ. Set $T := \{a^{\pm 1}, b^{\pm 1}\} \subseteq S$. One can decompose μ as $\mu = \mu(T)\mu_{|T} + \mu(S \setminus T)\mu_{|S\setminus T}$, where for an $R \subseteq S$, the restriction of μ to R is $\mu_{|R} = \mathbb{1}_R \frac{\mu_{|R}}{\mu(R)}$ ($\mu_{|R} := 0$ if $\mu(R) = 0$). Therefore, $\lambda(\mu) = \mu(T)\lambda(\mu_{|T}) + \mu(S \setminus T)\lambda(\mu_{|S\setminus T})$, and as a result, $r(\mu) \leq \mu(T)r(\mu_{|T}) + \mu(S \setminus T)r(\mu_{|S\setminus T})$. Hence, by 2. of Remark 7.2, this gives $r(\mu) \leq \mu(T)\frac{\sqrt{3}}{2} + \mu(S \setminus T) = 1 - \mu(T)(1 - \frac{\sqrt{3}}{2}) < 1$.

In the light of this observation, in the setting of a linear group Γ (i.e. a subgroup of $\mathrm{GL}_d(k)$), it should not be very surprising that the remarkable uniform versions of the Tits alternative obtained by Breuillard-Gelander [34] and (even more uniformly) by Breuillard [35] lead to 'uniform spectral gap' results. More precisely, we have

Theorem 7.7 (Uniform spectral gap, [35]). *Given d, q in $\mathbb{N}$, there exists a constant $0 < \beta < 1$, depending only on d and q, with the property that for all $\gamma_1, \ldots, \gamma_q \in GL_d(k)$ (k any field) such that the group generated by $\{\gamma_1^{\pm 1}, \ldots, \gamma_q^{\pm 1}\}$ in non-amenable, we have*

$$r(\frac{1}{2q}\sum_{i=1}^{2q}(\delta_{\gamma_i} + \delta_{\gamma_i^{-1}})) \leq \beta < 1$$

Similarly (as before the previous theorem), if there exists an amenable subgroup $H < \Gamma$ such that $\mu(H) \geq \alpha$ for some $\alpha \geq 0$, one immediately concludes that $r(\mu) \geq \alpha$. Reading this observation and the one preceding Theorem 7.7 by contrapositive, we see that they yield non-existence (respectively, of amenable subgroup and of free generators of large probability) assertions. 99 percent and 1 percent theorem/question gives/seeks existence results just looking at the spectral radius, in the same spirit as the last observation in Remark 7.6, in the setting of linear groups.

Theorem 7.8 (99 percent result, Breuillard [37]). *For each $d \in \mathbb{N}$ there exists a constant $c(d) > 0$ such that for every subgroup $\Gamma \leq GL_d(k)$ (k any field) and for all symmetric probability measure μ on Γ, if $r(\mu) \geq 1 - \eta$ for some $\eta \geq 0$, then there exists an amenable subgroup H of Γ with $\mu(H) \geq 1 - c(d)\eta$.*

Remark 7.9 (General 99 percent question). *Let us first say a few words about the linear group assumption : in fact, the 99 percent result is false if one omits this assumption and take Γ in an arbitrary non-amenable group. Wilson exhibited (see Theorem 1 in [116]) a first example of a non-amenable group G of non-uniform exponential growth, and moreover, a sequence of generating sets T_n of two elements of G (put $T_n = \{a_n^{\pm 1}, b_n^{\pm 1}\}$) such that the exponential growth rates of these sets satisfy $v_{T_n} \underset{n\to\infty}{\longrightarrow} 1$. As a result, if one considers the uniform probability measures ν_n on T_n's, using the observation following Remark 7.2, one immediately sees that $r(\nu_n) \underset{n\to\infty}{\longrightarrow} 1$ without the conclusion of the 99 percent theorem. As a result, for a '99 percent statement' to be true, one should at least require Γ to be of uniform exponential growth. In that generality, whether the corresponding (to Theorem 7.8) statement holds for Γ, i.e. the answer to the general 99 percent question, is unknown.*

This statement can be proven by combining an instance of the probabilistic method with the uniform spectral gap result (the author would like to thank E. Breuillard for sharing his proof with him). One drawback of this theorem, accounting for the terminology, is that the constant $c(d)$ is in fact strictly larger than one, so that the conclusion becomes void whenever $\eta \geq \frac{1}{c(d)}$, i.e. it is only valid for probability measures of spectral radius close to one. 1 percent question asks whether we have a similar situation for any value of the spectral radius :

1 percent question (Breuillard [37]) : Let Γ be a countable linear group. Given $\epsilon > 0$, can one find $\delta > 0$, depending only on ϵ and Γ, such that if μ is a symmetric probability measure on Γ with $r(\mu) \geq \epsilon$, then there exists an amenable group H in Γ such that a coset of H is charged larger than δ by μ, i.e. there exists $\gamma \in \Gamma$ such that one has $\mu(\gamma H) \geq \delta$.

Remark 7.10. *1. By symmetry of μ, it is indeed sufficient to consider only left cosets since $\mu(\gamma H) = \mu(H\gamma^{-1})$.*

2. One realises that one should indeed include the cosets of amenable groups in the conclusion of this question, since otherwise the answer is trivially negative. Let us give an indication : if for a symmetric probability measure μ, amenable subgroup H of Γ, and an element $\gamma \in \Gamma$, we have $\mu(\gamma H) \geq \delta$, then by symmetry, $\mu(H\gamma^{-1}) \geq \delta$, and hence $\mu^{*2}(H) \geq \delta^2$. As a consequence, by the observation following Theorem 7.7, $r(\mu^{*2}) \geq \delta^2$, and thus, $r(\mu) \geq \delta$ without μ necessarily charging an amenable subgroup significantly (using this indication, it is easy to construct a concrete such μ, for example when Γ is a non-commutative free group). By this observation, one can also formulate the question in a slightly 'easier' but interesting way, by modifying it as "..., then there exists an amenable group H in Γ, such that $\mu(H) \vee \mu^{*2}(H) \geq \delta$".

3. Similar to Theorem 7.8, one can ask uniform, therefore harder, versions of this question. But they seem to be out of reach for the author, for the time being. In Section 7.3 and Remark 7.36, we deal with easier versions of this question, in the setting of a free group (!).

7.2 A geometric view on hyperbolic automorphisms of a tree

In this section, our principal objective is to introduce and study some aspects of a notion (that of defective elements) related to geometric and dynamical properties of the action of a subset of $Aut(T)$, for a tree T, consisting of hyperbolic automorphisms (for these notions, see below). To do this, we start by quickly introducing the classical terminology (following mainly Tits [112] and in parts Serre [108]) and discussing some related results to set the stage. We then introduce some geometrical notions, study their properties and use them in the proofs of some known facts on free groups that are related with our later considerations in the text (see also the first paragraph of Appendix A.1). At the end, we provide a reformulation of 1 percent question in terms of boundary of free groups.

Graphs A graph $G = (V, E) =: (V(G), E(G))$ is the data of a set V, and a set E consisting of subsets with two elements of V. We call the elements of V, the vertices, and those of E, the edges of the graph G. A homomorphism ϕ from a graph $G_1 = (V_1, E_1)$ to a graph $G_2 = (V_2, E_2)$ is the data of a mapping $\phi : V_1 \longrightarrow V_2$ such that for each $\{x, y\} \in E_1$, one has $\phi(\{x, y\}) = \{\phi(x), \phi(y)\} \in E_2$. Call ϕ a monomorphism if the mapping $\phi : V_1 \longrightarrow V_2$ is an injection, and an epimorphism if $\phi : V_1 \longrightarrow V_2$ is a surjection and for each $\{u, v\} \in E_2$, there exists $\{x, y\} \in E_1$ with $\phi(\{x, y\}) = \{u, v\}$. An isomorphism and an automorphism are defined in the obvious manner (by abuse, we will often identify a graph homomorphism with the underlying map on vertex sets). A subgraph H of a graph G is a graph with $V(H) \subseteq V(G)$ and such that the inclusion map on vertex sets yields a monomorphism $\iota : H \longrightarrow G$. The unions and intersections of a family of subgraphs of a graph are defined in the usual manner.

Following Tits [112], let us denote by $Ch(a,b)$ for $a \in \{-\infty\} \cup \mathbb{Z}$ and $b \in \mathbb{Z} \cup \{\infty\}$ with $a \leq b$, the graph, called chain, with vertex set $V = \{i \in \mathbb{Z} \mid a \leq i \leq b\}$ and edge set $E = \{\{i, i+1\} \subset \mathbb{Z} \mid a \leq i \leq b-1\}$ (with corresponding conventions on $\pm\infty$). Given a graph G, we shall call a geodesic, geodesic ray and a geodesic segment, a subgraph of G which is, respectively, monomorphic image in G of the chains $Ch(-\infty, \infty)$, $Ch(a, \infty)$ with $a \in \mathbb{Z}$, and $Ch(a,b)$ with $a, b \in \mathbb{Z}$. If ϕ is the monomorphism in question, for a segment $Ch(a,b)$, denote the corresponding subgraph in G as $[\![\phi(a), \phi(b)]\!]$.

A graph is said to be finite if its set of vertices is finite. The length of a finite chain $Ch(a,b)$ and any monomorphic image of it in a graph G is set to be $b - a \in \mathbb{N}$, we will denote it by using the length function $l(.)$. More generally, a graph G can be given a metric structure by decreeing that for all $x, y \in V(G)$, the distance $d(x,y)$ be the smallest number among the lengths of geodesic segments $[\![\phi(a), \phi(b)]\!]$ with $a, b \in \mathbb{Z}$ such that $\phi(a) = x$ and $\phi(b) = y$. If no such geodesic segment exists, set $d(x,y) = \infty$. A graph G is said to be connected if for all $x, y \in V(G)$, $d(x,y) < \infty$ (this is indeed compatible with the topological realisation of the graph G, see Serre [108] 2.1). Connected components of G are defined in the obvious manner. For two subgraphs A, B of G, set $d(A,B) = \min\{d(a,b) \mid a \in A \text{ and } b \in B\}$. Two vertices $x, y \in V$ are called neighbours (and x is a neighbour of y and vice versa) if $d(x,y) = 1$. For $x \in V$, the degree of x is the number of its neighbours and a graph is called homogeneous if all of its vertices have the same degree. Finally, a vertex is called a terminal vertex or a leaf if its degree is one.

Trees A tree is a graph T such that for all $x, y \in V(T)$, there exists a unique geodesic segment with origin x and extremity y. A subtree S of T is a subgraph of T which is itself a tree. Equivalently, it is a connected subgraph of T. One notes that the intersection of a family of subtrees of a tree is a tree. Let V_0 be a set of vertices of a tree T (and $(S_i)_{i \in I}$ a collection of subgraphs of T), then one can talk about the subtree of T generated by V_0 (respectively, by $(S_i)_{i \in I}$). It is well-defined in a straightforward manner (for example, as the intersection of all subtrees of T containing V_0 in their vertex sets, respectively, admitting S_i's as subtrees).

Two arbitrary subtrees A, B of a tree T such that $\#(V(A) \cap V(B)) \leq 1$, have the property that $d(A,B)$ is realised on a unique pair $(a,b) \in V(A) \times V(B)$, i.e. $d(a,b) = d(A,B)$. For such A and B, and (a,b), the geodesic segment $[\![a,b]\!]$ will be called the bridge between/of A and B and will be denoted $Brid(A,B)$. In case $V(A)$ is a singleton $\{a\}$, the corresponding vertex $b \in V(B)$, will be denoted $p_B(a)$ and reads : the projection of a on B. Consequently, any subtree B of T defines a partition of T as $T = \dot{\bigcup\limits_{b \in B}} p_B^{-1}(b)$ (for empty B or consisting of a single vertex, this is the trivial partition).

A subtree A of a tree T is said to cross a subtree B of T, if there exist $c, d \in V(A) \setminus V(B)$ such that $V([\![c, d]\!]) \cap V(B) \neq \emptyset$. Let $(A_i)_{i \in I}$ be a family of subtrees of a tree T, and let A be the subtree generated by the family $(A_i)_{i \in I}$. For $i \in I$, A_i is called an extremal member of the family $(A_i)_{i \in I}$, if A_i is not crossed by A. In the particular case where each A_i consists of a single vertex of T, an extremal member of their family is a leaf of the subtree generated by the family. Any finite family $A_1, \ldots, A_n$ of subtrees of a tree T has at least one extremal member.

Automorphisms of trees Let $Aut(T)$ denote the group of automorphisms of a tree T and g be a homomorphism from T to T. A preliminary observation is that considering the metric structure of T, g is an automorphism of T if and only if g is an isometry. In particular, an automorphism $g \in Aut(T)$ leaves invariant the set of leaves of T. Moreover, a finite tree always has a leaf (particular case of the above statement for extremal members), and by inducting on the diameter of T, it is not hard to see that T has a vertex or an edge fixed by any automorphism of T (see Serre [108] 2.2). If g fixes an edge of T, it either fixes both vertices of the edge, or it permutes them. In the latter case, it is obvious that g can not fix any vertex in T. As a consequence, for a finite T, an automorphism falls into one of the two mutually exclusive categories according to whether it fixes a vertex or permutes the vertices of an edge.

For an infinite tree T, a third class of automorphisms exists and the three classes are mutually exclusive ; this is expressed in the following result of Tits [112] (3.2 Proposition). To state it, we call an automorphism g of a geodesic T a translation of translation distance $\tau(g) \in \mathbb{N}$, if in an isomorphic identification of T with $Ch(-\infty, \infty)$, it is induced by the vertex permutation defined by $x \mapsto x + \tau(g)$. Finally we note the straightforward observation that an infinite locally finite tree (i.e. each vertex of T has a finite degree) has a (bi-infinite) geodesic.

Theorem 7.11 (Tits, [112])**.** *Let T be a tree and $g \in Aut(T)$. Then, exactly one of the following conditions is satisfied :*

1. *g fixes a vertex of T.*
2. *g permutes two vertices of an edge of T.*
3. *There exists a geodesic A in T, invariant by g (i.e. $g(A) = A$), and g induces a non-trivial translation of translation distance $\tau(g) = \min_{x \in V(T)} d(g.x, x)$ on A. Moreover, in this case, A is the unique such geodesic and it can be characterised as the subtree generated by the set of vertices $\{x \in V(T) \mid d(g.x, x) = \tau(g)\}$.*

The automorphisms of T of type 1,2 and 3 above, are commonly referred to as, respectively, elliptic, parabolic (inversion), and hyperbolic, in analogy with the more standard terminology on, for example, $\mathrm{PSL}(2, \mathbb{R})$ and classification of its elements with respect to their action on the hyperbolic space $\mathbb{H}^2$ and its boundary (see for instance Bekka-Mayer [18] Chapter 2). For hyperbolicity, see also Tits [112] (3.6 Exemple), where a further analogy is exhibited for $\mathrm{SL}(2, \mathrm{k})$, k a local field, action on its Bruhat-Tits building, which is, in this case, a tree. We would like to mention that the study of parabolic automorphisms of a tree T is in a sense contained in the study of elliptic automorphisms by considering the induced action of parabolic elements on the barycentric subdivision of T (see Serre [108]). However, in the rest of this text, we will be only interested in hyperbolic automorphisms.

Hyperbolic automorphisms

Let $g \in Aut(T)$ be an hyperbolic automorphism of a tree T. Before starting, we note that in general, one may even have $Aut(T) = \{e\}$ but, for example, if T is a homogeneous tree, such a g exists. The geodesic A, given by 3. of Theorem 7.11, will be called the (translation) axis of g, and denoted by $\mathrm{axe}(g)$. Considering the projection function defined above, one observes that for $x \in V(T)$, we have $p_{\mathrm{axe}(g)}(g.x) = g.p_{\mathrm{axe}(g)}(x)$. Moreover, for $x \in V(T)$, we have $d(g.x, x) = \tau(g) + 2d(x, \mathrm{axe}(g))$, where, of course, $d(x, \mathrm{axe}(g)) = d(x, p_{\mathrm{axe}(g)}(x))$.

More on trees To proceed further, let us denote by $B(T)$ the set of geodesic rays of T modulo the equivalence relation $\sim$ where, for two geodesic rays ξ, ξ', we have $\xi \sim \xi'$ if and only if $l(\xi \cap \xi') = \infty$. For $x \in V(T)$, let us denote by $B_x(T)$ the set of geodesic rays with origin x. We have a natural bijection $B(T) \xrightarrow{\sim} B_x(T)$ given by $\xi \mapsto [\![x, \infty[\![$, where $[\![x, \infty[\![$ denotes the geodesic ray of equivalence class ξ and of origin x. Furthermore, denote by $b_\xi(.,.)$ the Busemann function on $V(T) \times V(T)$ defined by $b_\xi(x, y) = \lim_{n\to\infty} d(y, z_n) - d(x, z_n)$, where z_n is any sequence of vertices generating a geodesic ray of class ξ. And, say that a segment $[\![x, y]\!]$ in T is directed towards $\xi \in B(T)$, if $b_\xi(x, y) \leq 0$. For an hyperbolic automorphism g of T, denote by $\xi_g^+ \in B(T)$, the class of the geodesic ray containing infinitely many elements of the set $\{g^n x \mid n \geq 1 \text{ and } x \in \mathrm{axe}(g)\}$. This is well-defined and, in particular, does not depend on $x \in \mathrm{axe}(g)$. Set also $\xi_g^- := \xi_{g^{-1}}^+$. The axis of g is then the geodesic $]\!]\xi_g^-, \xi_g^+[\![$, where this latter notation denotes the unique geodesic with one end ξ_g^- and the other ξ_g^+. The uniqueness results easily from the defining property of a tree. By isometric property of an automorphism $g \in Aut(T)$, the image of a geodesic segment in T by g is a geodesic segment. From this, one sees that the action of g on T extends to/induces an action on $B(T)$. Contrary to the action of g on T without fixed points, it is easily observed that g has exactly two fixed points on $B(T)$, these are ξ_g^+, called the 'attracting' fixed point, and ξ_g^- the 'repulsing' fixed point of g.

By an orientation on a tree T, we mean a choice of an element $\xi \in B(T)$ (if T is infinite, we have $B(T) \neq \emptyset$; if T is finite, $B(T)$ is of course finite, and in this situation, an orientation is given by a choice of a leaf). For $\xi_1', \xi_1, \xi_2 \in B(T)$ and $x, y \in B(T)$, a geodesic ray/segment in T denoted as $]\!]\xi_1', \xi_1[\![$, $[\![x, \xi_2[\![$, and $[\![x, y]\!]$ will be considered as belonging to, respectively, the tree T oriented by ξ_1, ξ_2, and $\xi \in B(T)$ where $[\![x, y]\!]$ is directed towards, in the sense of the above paragraph, ξ (thus, they may belong to the same tree with different orientations, this should not cause confusion). Correspondingly, in this situation, we say that x is the origin of $[\![x, y]\!]$ and denote it by $o([\![x, y]\!])$, and y is the end of this segment and denoted it by $e([\![x, y]\!])$. A tree T with a distinguished vertex $x_0 \in V(T)$ will be referred to as a tree based on x_0. In a based tree T_{x_0} (i.e. T based on $x_0 \in V(T)$), a geodesic A will set to be based at $y_0 \in V(A)$, if $p_A(x_0) = y_0$.

We proceed to introduce further notions, with, notably, applications to ping-pong lemma type results (see Remark 7.14 below) in mind : let $S = \{g_i \mid i \in I\}$ be a set of hyperbolic elements in $Aut(T)$. We call the web of S, and denote it by web(S), the subtree of T generated by the translation axes of elements of S. More precisely, web(S) consists of the axes of g_i's and the bridges between the axes of each pair of g_i's. An element $g \in S$ is called an exterior element of S, if axe(g) is an extremal member of the family of geodesics axe(g_i), $i \in I$. We set $ext(S) = \{g \in S \mid g$ is an exterior element of $S\}$, and put $Int(S) = S \setminus Ext(S)$ and call its elements interior elements of S. It is easy to check that for $g \in Int(S)$, one has $\mathrm{web}(S \setminus \{g\}) \cap \mathrm{axe}(g) = \mathrm{web}(S) \cap \mathrm{axe}(g)$, and more generally, for each $g \in S$, $\mathrm{web}(S) \cap \mathrm{axe}(g) = \mathrm{web}(Ext(S)) \cap \mathrm{axe}(g)$.

Definition 7.12. *Given a set S of hyperbolic automorphisms of a tree T, an element $g \in S$ is called a strictly defective, quasi defective, or non-defective element of S, if, respectively, we have*

1. $l(\mathrm{web}(S \setminus \{g\}) \cap \mathrm{axe}(g)) > \tau(g)$
2. $l(\mathrm{web}(S \setminus \{g\}) \cap \mathrm{axe}(g)) = \tau(g)$
3. $l(\mathrm{web}(S \setminus \{g\}) \cap \mathrm{axe}(g)) < \tau(g)$

and simply call g defective, if it is either strictly or quasi defective. We shall also call the identity element $e \in Aut(T)$ as defective.

Remark 7.13. *We wish the underline at this point that an element g in $S \subset Aut(T)$ should rather be called S-defective or S-non-defective, but for notational convenience we omit S- whenever it should not cause any confusion.*

We remind the reader that by definition, $\tau(g)$ is a strictly positive integer. Moreover, to clarify, we note that in the above definition, one can have $\mathrm{web}(S\setminus\{g\})\cap\mathrm{axe}(g) \neq \emptyset$ and $l(\mathrm{web}(S \setminus \{g\}) \cap \mathrm{axe}(g)) = 0$, in which case $\mathrm{web}(S \setminus \{g\}) \cap \mathrm{axe}(g)$ is a single vertex. Let us denote by $D(S)$, the set of defective, and by $ND(S)$, non-defective elements of S. Then, one has a second decomposition of S as $S = D(S) \dot\cup ND(S)$. The straightforward relation between exterior and non-defective elements will be clarified later in a more restricted set-up.

Remark 7.14. *The main motivation in introducing the notion of (non)-defective elements of a given subset of $Aut(T)$ was the application of this notion to the study of spectral radius of a symmetric probability measure on a free group. Then it was remarked by the author that this notion is in fact very closely related to classical ping-pong lemma itself, and in the setting of free groups, to the notion of 'Nielsen reduced' elements. Recall that the classical Nielsen-Schreier theorem states that a subgroup of a free group is itself free. This theorem was first proved by Nielsen in 1921 for a finitely generated subgroup and then generalized to this form by Schreier in 1924 (see [86] and [87] and the references therein). The interest in Nielsen's proof is that he introduced a set of operations, nowadays called Nielsen transformations/reductions, to reduce the elements of a set to another set consisting of Nielsen reduced elements and generating the same subgroup as the initial set (he then concludes by showing that a set of Nielsen reduced elements generates a free subgroup). It turned out that the Nielsen transformations was not restricted to this proof, and to cite Fine-Rosenberger-Stille [53] "Nielsen transformations can be considered as the non-commutative analogs of row reduction* *of matrices and have proved to be indispensable in the theory of free groups." (See also [87]). In a later work, we will exhibit a purely geometric algorithm (based only on the study of translation axes) for Nielsen transformations. As a concrete example, in Appendix A, using the notion of non-defective elements, we show that the fact that a set of Nielsen reduced elements generate a free subgroup is a direct consequence of the Klein's ping-pong lemma (see also below). We also underline that, after Nielsen-Schreier's work, several simpler proofs appeared. The most efficient approach to prove Nielsen-Schreier theorem seems to be the one through topological considerations (see [108] [87] and [86] and the references therein).*

Before ending this part on hyperbolic automorphisms, to give a first flavour of the use of the above Definition 7.12, let us now state a version of the well-known ping-pong lemma :

Lemma 7.15 (Ping-pong lemma). *Let G be a group acting on a set X, and $S = \{g_i \mid i \in I\}$ be a subset of G. Suppose that for each $i \in I$, there exist subsets $A_i, R_i \subset X$ with $A_i \cap R_i = \emptyset$ and such that for all $i \neq j \in I$, we have $A_i \cap A_j = \emptyset = R_i \cap R_j$. Suppose moreover that for each $i \in I$, we have $g_i(X \setminus R_i) \subseteq A_i$ and $g_i^{-1}(X \setminus A_i) \subseteq R_i$. Finally, suppose that the union of $(\bigcup_{i\in I} A_i) \cup (\bigcup_{i\in I} R_i) \neq X$. Then, the elements of S freely generate a subgroup in G.* □

In the setting of the ping-pong lemma, for an element $g \in G$, subsets A_g, R_g of X satisfying $A_g \cap R_g = \emptyset$, $g.(X \setminus R_g) \subseteq A_g$, and $g^{-1}.(X \setminus A_g) \subseteq R_g$ will be called, respectively, an attracting and repulsing set of/for g.

Now, let us come back to our setting of an hyperbolic automorphism g of a tree T. We start by an observation on the attracting and repulsing sets of g :

Lemma 7.16. *Let $x \neq y$ be two vertices on $\mathrm{axe}(g)$ such that the segment $[\![x, y]\!]$ is directed towards ξ_g^+, and satisfies $l([\![x, y]\!]) \leq \tau(g) + 1$. Then, the subsets (i.e. subtrees) $p^{-1}_{[\![x,y]\!]}(y)$ and $p^{-1}_{[\![x,y]\!]}(x)$ are, respectively, attracting and repulsing sets for g.*

Démonstration. The proof is straightforward ; it follows from the definitions of attracting and repulsing sets together with the equivariance of projection mapping, i.e. $g.p_{\mathrm{axe}(g)}(v) = p_{\mathrm{axe}(g)}(g.v)$ for each $v \in V(T)$. □

The following lemma exhibits a first relation of the notion of non-defective elements with the ping-pong type arguments. In view of it, a synonym for S-non-defective elements can be 'in ping-pong position in S'.

Lemma 7.17. *Let S be a subset of $Aut(T)$ consisting of hyperbolic elements. Then, the subgroup of $Aut(T)$ generated by $ND(S) \subseteq S$, non-defective elements of S, is free of rank $|ND(S)|$. In particular, if $D(S) = \emptyset$, then S generates a free group of rank $|S|$.*

Démonstration. The proof follows basically from the previous two lemmata : for each $g \in ND(S)$, fix $x_g \neq y_g \in V(\mathrm{axe}(g))$ with $[\![x_g, y_g]\!]$ is directed towards ξ_g^+, and such that $V(\mathrm{axe}(g) \cap \mathrm{web}(S \setminus \{g\})) \subseteq V([\![x_g, y_g]\!]) \setminus \{x_g, y_g\}$, and $l([\![x_g, y_g]\!]) \leq \tau(g) + 1$. This is indeed possible since, g being non-defective, we have $l(\mathrm{web}(S \setminus \{g\}) \cap$

axe$(g)) < \tau(g)$. Then, in particular, setting $A_g = p^{-1}_{[\![x_g,y_g]\!]}(y_g)$ and $R_g = p^{-1}_{[\![x_g,y_g]\!]}(x_g)$, Lemma 7.16 implies that A_g and R_g are, respectively, attracting and repulsing sets for g. Moreover, it follows by their construction that for $g \neq h \in ND(S)$, we have $A_g \cap R_h = \emptyset = R_g \cap A_h$. Finally, it is easy to see that for our choices of x_g and y_g for a non-defective element g of S, any $v \in V([\![x_g, y_g]\!]) \setminus \{x_g, y_g\}$ has the property that $v \notin (\bigcup_{h\in ND(S)} A_h) \cup (\bigcup_{h\in ND(S)} R_h)$. As a consequence, ping-pong lemma applies and yields the desired result. $\square$

The following simple result is an immediate corollary of the previous lemma. We note that it was already observed by Culler-Morgan (2.6 Lemma [42]). On this occasion, we wish to mention at this point that, the author realizes that Culler-Morgan [42] and (thanks to Frédéric Paulin) Alperin-Bass [4] had a similar (to ours) geometric viewpoint through the considerations of translation axes, but with a different focus.

Corollary 7.18. *Let $g, h \in F \setminus \{e\}$ such that $l(\text{axe}(g) \cap \text{axe}(h)) < \tau(g) \wedge \tau(h)$. Then g, h freely generates a subgroup of F.*

Démonstration. Indeed, the hypothesis implies that g and h are both non-defective elements of $S := \{g, h\}$. Therefore, Lemma 7.17 directly concludes. $\square$

Free groups acting on their Cayley tree

We now restrict to the setting of interest to us : a free group F of finite rank acting on the left on its (right) Cayley graph T with respect to a free generating set. In particular, each element of $F \setminus \{e\}$ (e denoting the identity element) acts by an hyperbolic automorphism of T. Before moving on, let us mention an observation of rigidity for a group Γ acting on a tree T by hyperbolic automorphisms : such an action is of course free and this implies (see Théorème 4 in 3.3 of Serre's [108]) that the group Γ is a free group itself. In the sequel, F_q denotes the free group of rank $q \in \mathbb{N}^*$, when the rank is irrelevant we will often omit q from the notation. In the rest of this part, we recall and prove some properties of F, T and the F-action on T that will be later useful to us, using, in part, our geometric techniques to illustrate their use.

Let $\mathcal{A} = \{a, a^{-1}, b, b^{-1}, \dots\}$ be an alphabet (set of symbols). A word in this alphabet is a finite sequence of letters from $\mathcal{A}$. A word will be called reduced if it contains no consecutive symbols of the form α and α^{-1} from $\mathcal{A}$. The length of a word is the number of symbols (with multiplicities) occurring in this word, it will be denoted by $|.|$. For example, $|aa^{-1}b| = 3$. For a free group F_q, we will denote a fixed free generating set of it as $\{a_1^{\pm1}, \dots, a_q^{\pm1}\}$. F_q is seen to be in natural bijection with the set of reduced words in $\{a_1^{\pm1}, \dots, a_q^{\pm1}\}$ taken as an alphabet. Correspondingly, for an element $g \in F_q$, $|g|$ will denote its word length, i.e. the length of the unique reduced word in $\{a_1^{\pm1}, \dots, a_q^{\pm1}\}$ representing g. Sometimes, we also talk about 'reduced product' when we refer to a concatenation of elements of $\{a_1^{\pm1}, \dots, a_q^{\pm1}\}$ yielding a reduced word. The word length $|.|$ on F_q induces a metric on F_q, by setting, the distance, for $g, h \in F$, $d(g, h) := |g^{-1}h|$. Endowed with this metric, F_q is in isometry with its Cayley graph $T = T_{2q}$ (homogeneous tree of degree $2q$) with respect to the fixed generating set. We shall consider T with its natural labels, its vertices by the elements of F and edges by those of the generating set. A Cayley graph T_q will always be considered as based at the vertex labelled by $e \in F_q$, unless otherwise explicitly mentioned.

An element g of F can be factorised uniquely as a reduced product as $g = c(g)r(g)c(g)^{-1}$ where $c(g), r(g) \in F$, and $c(g)$ is of maximal word length. We will refer to this factorisation as the cyclic factorisation of g, to $c(g)$ as the cyclic part of g, and to $r(g)$ as the cyclically reduced part of g. Correspondingly, if $g = r(g)$, g will be called cyclically reduced. The word length of the cyclically reduced part $r(g)$ of g will be called the translation distance of g. This terminology is compatible with previous ones for the hyperbolic automorphisms of trees : indeed, when considering the action of F on T, for any $g \in F \setminus \{e\}$, its translation axis is seen to be a geodesic based on the vertex labelled by $c(g)$ (see Appendix A.2), and of translation distance $|r(g)|$, i.e. $\tau(g) = |r(g)|$.

To proceed further, we need a more thorough understanding of translation axes of elements of F. More precisely, we seek to understand the translation axis of an element gh, given those of g and h. For this, let us introduce some more terminology : two elements $g, h \in F \setminus \{e\}$ will be called intersecting (non-intersecting), if $l(\mathrm{axe}(g) \cap \mathrm{axe}(h)) \in \mathbb{N}^*$ (respectively, $l(\mathrm{axe}(g) \cap \mathrm{axe}(h)) = 0$). We caution the reader that if the axis of g and h crosses at a single vertex, then the intersection length is zero, and hence g and h are non-intersecting. We also note that for two non-intersecting elements g and h, Corollary 7.18 applies immediately and yields that g, h generate a free subgroup of rank two. Two intersecting elements g, h will said to be intersecting in the same direction/directed similarly, if the orientations induced by ξ_g^+ and ξ_h^+ on $\mathrm{axe}(g) \cap \mathrm{axe}(h)$ is the same ; in other words, for both orientations one has the same origin and extremity for $\mathrm{axe}(g) \cap \mathrm{axe}(h)$. Otherwise, they are said to be intersecting in the opposite directions. For two intersecting elements g, h, $c(g, h)$ will denote the intersection segment of their axes, oriented in the direction of ξ_g^+. Finally, for a vertex $x \in V(T)$, an integer $k \in \mathbb{N}$ and an element $\xi \in B(T)$, we denote by $(x)_{k \to \xi}$, the unique vertex of T belonging to $[\![x, \xi]\!]$ and satisfying $b_\xi(x, (x)_{k \to \xi}) = -k$.

Lemma 7.19. *Let $g, h \in F \setminus \{e\}$. The following statements are equivalent :*

1. *g and h commute.*
2. *g and h satisfy a non-trivial relation (i.e. a reduced word in $\{g, g^{-1}, h, h^{-1}\}$ equals to e in F).*
3. $l(\mathrm{axe}(g) \cap \mathrm{axe}(h)) = \infty$.
4. $\{\xi_g^-, \xi_g^+\} \cap \{\xi_h^+, \xi_h^-\} \neq \emptyset$.
5. *g and h have the same translation axis.*

Démonstration. 1. implies 2. is clear, one has $ghg^{-1}h^{-1} = e$. 2. implies 1. follows from Nielsen-Schreier theorem together with a more general statement (see Corollary 2.13.1 [87]) asserting that any n generators of a free group of rank $n \in \mathbb{N}$ are free generators : indeed, by Nielsen-Schreier theorem, g and h generate a free subgroup of rank at most two, but if the rank is equal to two, then they are free generators, and in particular, they don't satisfy a non-trivial relation. As a result, g and h generate a subgroup isomorphic to $\mathbb{Z}$, and hence they commute. For 1. implies 3. and 5., observe that if g and h commutes, then for each $x \in \text{axe}(g)$, we have $d(x, g.x) = d(x, (h^{-1}gh).x) = d(h.x, g.(h.x))$, thus by Theorem 7.11 $h.x$ is an element of minimal translation length for g, i.e. $h.x \in \text{axe}(g)$. Hence the result follows again by the same Theorem 7.11 (for an alternative way to see that 1. implies 3., we can argue by contrapositive. Suppose 3. is not true and set $d := l(\text{axe}(g) \cap \text{axe}(h)) < \infty$. Let $k \in \mathbb{N}$ such that $k > \frac{d}{\tau(g) \wedge \tau(h)}$. Then, since for all $u \in F \setminus \{e\}$ and $p \in \mathbb{N}^*$, we have $\text{axe}(u^p) = \text{axe}(u)$ and $\tau(u^p) = p\tau(u)$, g^k and h^k satisfy the hypothesis of Corollary 7.18 and generate a free subgroup of rank two in F. In particular, g and h don't commute). The equivalence of 3. and 4. is clear by definition of the space $B(T)$. Let us give a proof of 3. implies 5. by using again translation axes and distances : suppose 3. holds and 5. does not. Up to replacing g, h by their inverses, suppose $\xi_g^- = \xi_h^-$, so that $\xi_g^+ \neq \xi_h^+$. Let x be the unique vertex of T satisfying $\text{axe}(g) \cap \text{axe}(h) =]\!]\xi_g^-, x[\![$, i.e. 'the last common vertex before the axes of g and h separate'. By this property of x, we have $p_{\text{axe}(g)}(hx) = x$, and since $x \in \text{axe}(h)$, $d(h.x, x) = \tau(h) = d(h.x, \text{axe}(g))$. Applying g to $h.x$, by g-equivariance of $p_{\text{axe}(g)}$, get $p_{\text{axe}(g)}(gh.x) = g.x$, and have $d(gh.x, \text{axe}(g)) = d(gh.x, g.x) = d(h.x, x) = \tau(h)$. On the other hand, since $x \in \text{axe}(g)$, $d(g.x, x) = \tau(g)$ and by defining property of x, $p_{\text{axe}(h)}(g.x) = x$, so that by h-equivariance of $p_{\text{axe}(h)}$, $p_{\text{axe}(h)}(hg.x) = h.x$ and hence $p_{\text{axe}(g)}(hg.x) = p_{\text{axe}(g)}(h.x) = x$ As a consequence, $d(hg.x, \text{axe}(g)) = d(hg.x, x) = d(hg.x, h.x) + d(h.x, x) = \tau(g) + \tau(h) \neq \tau(h) = d(gh.x, x)$. In particular, g and h don't commute. Finally, 5. implies 2. is clear since, supposing without loss of generality that g and h intersects in the same direction (otherwise consider g and h^{-1}), for each $x \in \text{axe}(g) = \text{axe}(h)$, one has $g^{-\tau(h)}h^{\tau(g)}.x = x$. Therefore, 2. follows as the action is free. □

Remark 7.20. *Note that by Lemma 7.19, we have a well-defined mapping from the set of abelian subgroups of F to the set of geodesics in T. We shall denote this mapping by the same notation* axe(.).

Using the equivalence of 3. and 5. above, let us see another application of our Lemma 7.17 to the proof of following fact :

Lemma 7.21. *Let $S = \{g_i \mid i \in \mathbb{N}\}$ be a countable set of elements in $F \setminus \{e\}$ with distinct (not necessarily disjoint) axes. Suppose that the lengths of pairwise intersections of axes of elements of S are bounded above, i.e. $M := \sup_{i \neq j \in \mathbb{N}} l(c(g_i, g_j)) < \infty$ (note that this last condition is trivially satisfied if S is finite). Then, for all $k \in \mathbb{N}$ with $k > \frac{M}{\min_{i=1,\dots,n} \tau(g_i)}$, the elements of $\{g_i^k \mid i \in \mathbb{N}\}$ freely generate a subgroup in F.*

Remark 7.22. *Note that by Lemma 7.19 another way to formulate the first hypothesis of Lemma 7.21 is to say that there is no abelian group contaning two elements of S.*

For $n \in \mathbb{N}$, $S \subseteq F$, let us denote by ${}^nS$ the set $\{g^n \mid g \in S\}$.

Proof of Lemma 7.21. Indeed, using the fact that for all $u \in F \setminus \{e\}$ and $k \geq 1$, we have $\text{axe}(u^k) = \text{axe}(u)$ and $\tau(u^k) = k\tau(u)$, one immediately sees from the definitions that for $k \in \mathbb{N}$ as in the statement of the lemma, the set ${}^kS$ consists of non-degenerate elements. In other words, $D({}^kS) = \emptyset$ and the result follows from Lemma 7.17. □

Lemma 7.23. 1. *Let $g \in F \backslash \{e\}$ and $< g >$ be the cyclic subgroup of F generated by g. Then, $< g >$ is contained in a unique maximal abelian subgroup, denoted* $\mathrm{Mab}(g)$, *of F. The cyclic group* $\mathrm{Mab}(g)$ *is generated by one of the two elements u, u^{-1} of the same translation axis $A = \mathrm{axe}(g)$ as g, of minimal translation distance on A, and satisfying $\xi_u^+ = \xi_g^+$ (thus by Lemma 7.19, $\xi_u^- = \xi_g^-$). We shall denote by* $\mathrm{per}(g)$ *and call it period of g, the translation distance of u (for example, for $n \in \mathbb{N}^*$ and two elements $a \neq b^{-1}$ of the generating set of F, we have* $\mathrm{per}(ab) = 2$, $\mathrm{per}((ab)^n) = 2$, $\mathrm{per}(ab^3a^{-1}) = 1$, $\mathrm{per}((ab)^n a) = 2n + 1$*).*

2. *For any segment I in T, there exists an abelian subgroup $< g >$ in F such that* $\mathrm{axe}(< g >) \supset I$.

Démonstration. 1. Let u be as in the lemma, i.e. the element of $\mathrm{axe}^{-1}(\mathrm{axe}(g)) \subset F$ of minimal translation distance with $\xi_u^+ = \xi_g^+$. It follows by Lemma 7.19 that g and u commute, and in fact, we have $u^{\tau(g)} g^{-\tau(u)} = e$. We claim that $\tau(u)|\tau(g)$. Indeed, otherwise, write $\tau(g) = k\tau(u) + r$ with $k \geq 1$ and $\tau(u) > r \geq 1$, and we have that gu^{-k} is an element of F, not equal to identity, commuting to g and u, thus of translation axis A, and of translation distance r. This not being possible, our claim holds and we have $g = u^{\frac{\tau(g)}{\tau(u)}}$. For an element $h \in F \setminus \{e\}$ contained in an abelian group containing also g, using once more Lemma 7.19 and the same argument above, we conclude that $< g >$ is contained in a unique maximal abelian group which is $< u >$.

2. For the second argument, we use the extrinsic label structure of T (with elements of F for vertices). Let the given segment be $[\![x, y]\!]$ with $x, y \in F$ and $d(x, e) < d(y, e)$. Set $c(g) := x$, and put $\tilde{r}(g) := x^{-1}y \neq e$. If the first letter α of $\tilde{r}(g)$ is equal to the inverse of last letter of $\tilde{r}(g)$, set $r(g) := \tilde{r}(g)\beta$ for any $\beta \notin \{\alpha, \alpha^{-1}\}$ where α and β are in the generating set of F; otherwise set $r(g) := \tilde{r}(g)$. Finally, put $g := c(g)r(g)c(g)^{-1}$. Now, it is easy to see that this expression is the cyclic decomposition of g, and the translation axis of g contains the segment $[\![x, y]\!]$.

□

Remark 7.24. *In view of this lemma, in a given symmetric subset $S \subset F \setminus \{e\}$ containing at most two elements (inverses of each other) of maximal abelian subgroups (equivalently, in terms of translation axes : all elements of S have distinct translation axis, except of course for the corresponding inverse element), any external element of S is a non-defective element in S, i.e. $Ext(S) \subseteq ND(S)$.*

We note that at this point, the interested reader can look at Appendix A.1 in a continuation with this section.

A reformulation of 1 percent question for free groups

In this last subsection of Section 7.2, we extend our understanding on the geodesics of a tree T_{2q} in relation to the action of F_q. Using this, we give a reformulation of the 1 percent question for free groups through the 'projections' (axe(.) mappings) of abelian subgroups on 'the boundary', that we mentioned in the previous part (Remark 7.20).

More on geodesics and maximal abelian subgroups Let us start with endowing the boundary of a tree T with a metric ; this is classical : for all $x \in T$, one can define a metric on $B_x(T)$ in the following way : for $\xi_1, \xi_2 \in B_x(T)$, one sets $d_x(\xi_1, \xi_2) := e^{-l([\![x,\xi_1[\![\cap[\![x,\xi_2[\![)}$. This makes $B_x(T)$ into a compact, perfect ultrametric (for all $\xi_1, \xi_2, \xi_3 \in B_x(T)$, $d_x(\xi_1, \xi_3) = d_x(\xi_1, \xi_2) \vee d_x(\xi_2, \xi_3)$) space with diameter one. Let us also denote by d_x the metric induced on $B(T)$ by the bijection $B_x(T) \xrightarrow{\sim} B(T)$. It is not hard to see that for different $x \in V(T)$, d_x's define equivalent metrics, in particular, they define the same topology on $B(T)$. Denote by $\mathrm{ab}(F)$ the set of abelian subgroups of F (which is also the set of amenable subgroups), and by $\mathrm{Mab}(F)$ the set of maximal abelian subgroups of F (we also consider the trivial subgroup consisting of the identity element as an element of $\mathrm{Mab}(F)$). One remarks by Lemma 7.23 that, denoting furthermore by $\mathrm{Geod}(T)$ the set of geodesics in T, the restriction of the mapping $\mathrm{ab}(F) \xrightarrow{\mathrm{axe}} \mathrm{Geod}(T) \cup \{*\}$ given by Remark 7.20 ($\mathrm{ab}(\{e\}) := \{*\}$), to $\mathrm{Mab}(F)$ is injective. In passing, let us also denote by $\mathrm{Geod}_x(T)$ the set of geodesics in T passing through the vertex x in T. We have a natural bijection between $\mathrm{Geod}(T)$ and the set $B(T) \times B(T) \setminus \Delta$, where Δ stands for the diagonal in $B(T) \times B(T)$: each geodesic in T is written uniquely as $]\!]\xi, \xi'[\![$ with $\xi \neq \xi' \in B(T)$. As such, $\mathrm{Geod}(T)$ has a naturally associated topology coming from the product topology on $B(T) \times B(T)$. Now set $\mathrm{Geod}(T)_* := \mathrm{Geod}(T) \cup \{*\}$ as a disjoint union ; $\mathrm{Geod}(F)_*$ an F-action coming from the diagonal action of F on $B(T) \times B(T)$ and accordingly fixing $*$. Let us summarise some more observations in the below remark and lemma :

Remark 7.25. *Let us say a word about this topology on* $\mathrm{Mab}(F)$ *from a different perspective : the 'geometric realisation' of the group F through its action enabled us to find a natural 'augmentation' of its maximal abelian subgroups to translation axes, and using this latter, we put a topology on* $\mathrm{Mab}(F)$ *which is more interesting than its 'unaugmented' version, i.e. the Chabauty topology ; which induces the discrete topology on* $\mathrm{Mab}(F)$.

Lemma 7.26. *The injection* $\mathrm{Mab}(F) \overset{\mathrm{axe}}{\hookrightarrow} \mathrm{Geod}(T)_*$ *is an F-equivariant mapping of dense image, where F acts by conjugation on* $\mathrm{Mab}(F)$ *and on* $\mathrm{Geod}(T)$ *as above.*

Démonstration. The equivariance follows directly by the characterisation of a translation axis of an element as the set of minimally translated vertices of T, together with the fact that F acts by isometries on T. That $\mathrm{axe}(\mathrm{Mab}(F))$ is dense in $\mathrm{Geod}(T)$ follows from 2. of Lemma 7.23. □

A reformulation of 1 percent question We shall rather use the version in 2. of Remark 7.10 of the 1 percent question for the free groups : it asks whether given a free group F_q and $\epsilon > 0$, one can find $\delta > 0$ such that for any symmetric probability measure μ with spectral radius $r(\mu) \geq \epsilon$, there exists an amenable subgroup H in F_q with $\mu(H) \vee \mu^{*2}(H) \geq \delta$. Now recall that we have defined a surjective mapping Mab on F_q onto $\mathrm{Mab}(F)$ associating to a g in F_q the maximal abelian group that contains g, and the axe(.) mapping, injective on $\mathrm{Mab}(F)$, associating to an amenable subgroup its translation axis in T_{2q} ; so that we have $F \overset{\mathrm{Mab}}{\twoheadrightarrow} \mathrm{Mab}(F) \overset{\mathrm{axe}}{\hookrightarrow} \mathrm{Geod}(T)_*$. For a symmetric probability measure μ on F_q, setting $\nu_\mu := (\mathrm{axe} \circ \mathrm{Mab})_* \mu$, the push-forward of μ on $\mathrm{Geod}(T)_*$, we can clearly write the 1 percent question for F_q as follows : Given $\epsilon > 0$, can one find $\delta > 0$ such that if a symmetric probability measure μ on F_q has spectral radius $r(\mu) \geq \epsilon$, then there exists an element x of the F_q-space $\mathrm{Geod}(T)_*$, which is charged more than δ by ν_μ or $\nu_{\mu^{*2}}$, i.e. $\nu_\mu(x) \vee \nu_{\mu^{*2}}(x) \geq \delta$.

7.3 An answer to a (weak) version of the 1 percent question

Contrary to the title, the first and principal goal of this section is to show how the notion of defective/non-defective elements can be related to the spectral radius of a probability measure. More precisely, we show that the proportion of number of defective elements in a set gives an upper bound for the spectral radius of the uniform probability measure on that set. This will be done using our considerations in Section 7.2 and a ping-pong type lemma due to Cowling-Bekka-de la Harpe [17], which we state and prove in the first subsection. Our second aim in this section is to give a first application (Theorem 7.27) of this upper bound. This will answer positively to a weaker version of the 1 percent question. We make this more precise below. At this point, we wish to mention that the use of geometric understanding in Section 7.2 to control the spectral radius is not limited to this application and we obtained several other results in a work in progress (see Remark 7.36).

To discuss one application of Section 7.2 to the 1 percent question, let $l(.) = |.|$ and $d(.,.)$ denote, respectively, the length function and the metric on the free group F_q associated to a free generating set $\{a_1^{\pm 1}, \ldots, a_q^{\pm 1}\}$. Moreover, for each element $g \in F_q$, define the ball of g, denoted $B(g)$, as the ball in F_q of centre g and radius $l(g)$. A straightforward observation on the relation between the balls of elements in F_q and amenable subgroups of F_q (we remind that these are all abelian subgroups), is the following : let S be a finite symmetric subset of an amenable subgroup H of F_q. Let $h^{\pm 1} \in S$ be the elements of maximal translation length in S. Then, $B(h) \cup B(h^{-1}) = S$, and in fact, $\#(B(h) \cap S \setminus \{e\}) = \frac{\#(S \setminus \{e\})}{2}$ (see Lemma 7.32). In particular, if μ is a symmetric probability measure of support S in F_q charging an amenable subgroup H larger than δ (i.e. $\mu(H) \geq \delta$), then there exists $h \in S$, with $\mu(B(h)) \geq \frac{\delta}{2}$. This latter observation is also valid for cosets xH of amenable groups H in F_q. By consequent, an easier version of 1 percent question asks whether for F_q and given $\epsilon \geq 0$, one can find a $\delta > 0$ such that for any symmetric probability measure μ of support $S \subset F_q$, there exists $x \in S$ with $\mu(B(x)) \geq \delta$. The following result clearly answers affirmatively to this question :

Theorem 7.27. *For each $1 \geq \epsilon \geq 0$, there exists $\delta_\epsilon > 0$, such that if μ is a symmetric probability measure on a free group F_q with $r(\mu) \geq \epsilon$, then for all $y \in F$, one can find x in the support of μ such that $\tau_{y*}\mu(B(\tau_y(x))) \geq \delta_\epsilon$.*

Let us now start with proving the ping-pong type lemma, continue with exhibiting the relation of defective elements with spectral radius and finally end with the proof of the previous theorem.

Ping-pong type lemma

Let Γ be a countable group acting on a set X. We shall start with the following result which gives us a way to control norms of some elements of $C_r^*\Gamma$ by looking at combinatorial properties of the actions of these elements on X. This lemma can be traced back to Powers [96], and was used by Bekka-Cowling-de la Harpe in [17] (see Lemma 2.3). See also [37].

Lemma 7.28 (Ping-pong type lemma, [17])**.** *Let $n \in \mathbb{N}$ and $\gamma_1, \ldots, \gamma_n \in \Gamma$ such that there exist subsets $\tilde{A}_1, \ldots, \tilde{A}_r \subseteq X$ and $\tilde{R}_1, \ldots, \tilde{R}_r \subseteq X$ with the properties that for each $i = 1, \ldots, r$, $\gamma_i(X \setminus \tilde{R}_i) \subseteq \tilde{A}_i$, $\sum_{i=1}^n \mathbb{1}_{\tilde{A}_i} \leq M_A$ and $\sum_{i=1}^n \mathbb{1}_{\tilde{R}_i} \leq M_R$ for some constants M_A and M_B in $\mathbb{N}$. Then,*

$$||\lambda(\frac{1}{n}\sum_{i=1}^n \delta_{\gamma_i})||_{l^2\Gamma \to l^2\Gamma} \leq \frac{1}{\sqrt{n}}(M_A^{\frac{1}{2}} + M_B^{\frac{1}{2}})$$

Démonstration. We can suppose that $X = \Gamma$ and that the action is given by left multiplication ; indeed, fix some $x_0 \in X$ and consider the orbit function $\pi_{x_0} : \Gamma \longrightarrow X$ given by $\pi_{x_0}(\gamma) = \gamma.x_0$. Setting, for $i = 1, \ldots, n$, $A_i := \pi_{x_0}^{-1}(\tilde{A}_i)$ and $R_i := \pi_{x_0}^{-1}(\tilde{R}_i)$, one sees that we have $\gamma_i(\Gamma \setminus R_i) \subseteq A_i$, and since $\mathbb{1}_{A_i} = \mathbb{1}_{\tilde{A}_i} \circ \pi_{x_0}$ and $\mathbb{1}_{R_i} = \mathbb{1}_{\tilde{R}_i} \circ \pi_{x_0}$, obviously, $\sum_{i=1}^n \mathbb{1}_{A_i} \leq M_A$ and $\sum_{i=1}^n \mathbb{1}_{R_i} \leq M_R$.

Now, let $f, g \in l^2\Gamma$, $\gamma \in \Gamma$ and $A, R \subset \Gamma$ be such that $\gamma(\Gamma \setminus R) \subseteq A$. Observe that

$$\begin{aligned} | < \lambda(\gamma)f, g > | &= | < \lambda(\gamma)(f\mathbb{1}_R + f\mathbb{1}_{\Gamma\setminus R}), g > | \\ &\leq | < \lambda(\gamma)(f\mathbb{1}_R), g > | + | < \lambda(\gamma)(f\mathbb{1}_{\Gamma\setminus R}), \mathbb{1}_A g > | \\ &\leq ||f\mathbb{1}_R||_2||g||_2 + ||f||_2||g\mathbb{1}_A||_2 \end{aligned} \tag{7.1}$$

where, on the first inequality, we used the fact that $\gamma(\Gamma \setminus R) \subseteq A$ to add the factor $\mathbb{1}_A$ in front of g not modifying the value of the scalar product, and on the second inequality, we used the fact that the Γ action on $l^2\Gamma$ is by unitary transformations and $||f\mathbb{1}_{\Gamma\setminus R}||_2 \leq ||f||_2$. Now, using (7.1) for $\lambda(\frac{1}{n}\sum_{i=1}^n \delta_{\gamma_i})$ with A_i's, R_i's a,d for $f, g \in l^2\Gamma$, we have

$$\begin{aligned} | < \lambda(\frac{1}{n}\sum_{i=1}^n \gamma_i)f, g > | &\leq \sum_{i=1}^n \frac{1}{n}| < \lambda(\gamma_i)f, g > | \\ &\leq ||g||_2\frac{1}{n}\sum_{i=1}^n ||f\mathbb{1}_{R_i}||_2 + ||f||_2\frac{1}{n}\sum_{i=1}^n ||g\mathbb{1}_{A_i}||_2 \\ &\leq ||g||_2\frac{1}{\sqrt{n}}(\sum_{i=1}^n ||f\mathbb{1}_{R_i}||_2^2)^{\frac{1}{2}} + ||f||_2\frac{1}{\sqrt{n}}(\sum_{i=1}^n ||g\mathbb{1}_{A_i}||_2^2)^{\frac{1}{2}} \end{aligned} \tag{7.2}$$

where the last inequality follows for example by Jensen's inequality. As a result, the desired result follow from (7.2), by observing that

$$\sum_{i=1}^n ||f\mathbb{1}_{R_i}||_2^2 = \sum_{i=1}^n\sum_{x\in\Gamma} f^2(x)\mathbb{1}_{R_i}(x) = \sum_{x\in\Gamma} f^2(x)\sum_{i=1}^n \mathbb{1}_{R_i}(x) \leq ||f||_2^2 M_R$$

and similarly for $\sum_{i=1}^n ||g\mathbb{1}_{A_i}||_2^2$. $\square$

Estimating spectral radius with defective elements

Let us now come back to our setting where $\Gamma = F_q$ is a free group and $X = T_{2q}$ is its Cayley tree as before. On the way to put the ping-pong type lemma into action, let us start by singling out a notion and noting some observations. For an element $g \in F$, we have seen in Lemma 7.16, how one can choose attracting and repulsing sets A_g and R_g for g. The choice of these sets can be arbitrary as long as they satisfy the condition in that lemma ; more precisely, in the lemma, one can shift x, y on $\mathrm{axe}(g)$ keeping the distance between them, the same (or decrease the distance, but this is useless for our purposes). Now let S be a subset of F and for each $g \in S$, choose a pair of attracting and repulsing sets A_g, R_g in T by Lemma 7.16 (for $g = e$, we set $A_g = R_g = T$). For a vertex $x \in V(T)$, we call the attracting and repulsing indices of x, respectively, the number of attracting and repulsing sets of elements of S containing x, i.e. these are, respectively, $\sum_{g \in S} \mathbb{1}_{A_g}(x)$ and $\sum_{g \in S} \mathbb{1}_{R_g}(x)$. In view of the ping-pong type lemma, to keep the spectral radius of the uniform probability measure on S small, an optimal choice of A_g's and R_g's will be one that minimises the attracting and repulsing indices of vertices of X. Let us be more precise on this with the following discussion : let S be a finite symmetric set consisting only of non-defective elements ($D(S) = \emptyset$). Let μ_S denote the uniform probability measure on S. In this particular situation, we know by Lemma 7.17 that S freely generates a subgroup and hence by 2. of Remark 7.2, we know the actual value of $r(\mu_S)$; it is $\frac{2\sqrt{|S|-1}}{|S|}$. On the other hand, it is easy to see that one can come up with a 'bad' choice of A_g's and R_g's for $g \in S$, so as to have $x \in V(T)$, for example, with attracting index equal to $\frac{|S|}{2}$ so that ping-pong type lemma gives an upper bound larger than $\frac{1}{2}$. But, as in the proof of Lemma 7.17, in fact, one has a choice of A_g's and R_g's in ping-pong position, i.e. for each vertex in T, the attracting and repelling indices are at most one. As a result, with this 'right' choice, ping-pong type lemma yields the much better upper bound $\frac{2}{|S|^{\frac{1}{2}}}$, which is almost the same as the actual spectral radius, when $|S|$ is large.

Let us now continue with the following simple but main observation of this passage, which was the initial motivation for the author to introduce the notion of defective elements. Before stating the lemma, recall that we have a partition of a subset $S \subseteq F$ into S-defective elements $D(S)$ and S-non-defective elements $ND(S)$, so that $|S| = |D(S)| + |ND(S)|$.

Lemma 7.29. *Let S be a finite symmetric subset of F and μ_S be the uniform probability measure on S. Then, one has*

$$r(\mu_S) \leq \frac{2}{|S|^{\frac{1}{2}}}(2 + \frac{|D(S)|}{2}))^{\frac{1}{2}}$$

Démonstration. Note first that if $g \in S$ is non-defective in S, then it is non-defective in any subset S_0 of S. Now write $S = D(S) \dot\cup ND(S)$ and considering $ND(S)$ alone, as in the proof of Lemma 7.17, get a ping-pong choice of A_g's and R_g's for $g \in ND(S)$. For $h \in D(S)$, choose A_h and R_h arbitrarily using Lemma 7.16. As a result, for all $x \in V(T)$, writing $\sum_{g \in S} \mathbb{1}_{A_g}(x) = \sum_{g \in D(S)} \mathbb{1}_{A_g}(x) + \sum_{g \in ND(S)} \mathbb{1}_{A_g}(x)$, one has $\sum_{g\in ND(S)} \mathbb{1}_{A_g}(x) \leq 1$ and $\sum_{g\in D(S)} \mathbb{1}_{A_g}(x) \leq \left\lceil \frac{|D(S)|}{2} \right\rceil$, and similarly for $\sum_{g\in S} \mathbb{1}_{R_g}(x)$. Hence the result follows by the ping-pong type lemma. □

The previous lemma says that given an $\epsilon > 0$, whenever $r(\mu_S) \geq \epsilon$ for any large enough symmetric subset S, we have a particular information about S ; namely, it contains many (about $\epsilon^2|S|$) defective elements, i.e. those elements whose translation axes are in some sense surrounded by axes of other elements and of relatively small translation distance. In the following subsection, we shall see a concrete application of this result, through a geometric sufficient condition for defectiveness of an element in a free group.

Remark 7.30. *(A better bound) To bound the spectral radius of a probability measure as in Lemma 7.29 using the decomposition into defective and non-defective elements, a more direct approach is to proceed as we did in the paragraph preceding Theorem 7.7 using the particular fact that non-defective elements freely generate a subgroup and Kesten's calculation (2. of Remark 7.2). Accordingly, one gets* $r(\mu_S) \leq \frac{|D(S)|}{|S|} + (1 - \frac{|D(S)|}{|S|})\frac{\sqrt{2|ND(S)|-1}}{|ND(S)|}$. *It turns out that this provides a better bound than in Lemma 7.29 ; we thank J.F. Quint for pointing it out. To comment, this is not very surprising in view of the fact that this latter bound uses a more particular information (that $ND(S)$ is a free subset) than the general ping-pong type Lemma 7.28. We shall stick to the bound of the previous lemma as the particular constants do not play a role in our considerations.*

Charging balls : a proof of a weak version of 1 percent question

Recall that for an element $g \in F$, $B(g)$ denotes the ball in F of centre g and radius $|g| = l(g)$. Note that for all $g \in F$, $e \in B(g)$. We have the following sufficient condition for an element g to be non-defective in a symmetric set S containing g :

Lemma 7.31. *Let S be a symmetric subset of F and let $g \in S \setminus \{e\}$. Suppose that $B(g) \cap (S \setminus \{e, g\}) = B(g^{-1}) \cap (S \setminus \{e, g^{-1}\}) = \emptyset$ and that for each $h \in S \setminus \{g, g^{-1}\}$, $\{g, g^{-1}\} \cap B(h) = \emptyset$. Then, g is a non-defective element of S.*

The proof consists of rather straightforward arguments on the tree T_{2q}, and heavily uses the notation and the terminology introduced in Section 7.2. Before proceeding with the proof, let us first make the following observation on the membership of elements of F to the ball $B(g)$; this clarifies the first assumption on g in Lemma 7.31. To state it, let us introduce a further notation also to be used in the Appendix A.1 : for $g, h \in F$, denote by $c_{\text{in}}(g, h)$ the element of F corresponding to the common initial prefix of g and h seen as (reduced) words in the alphabet $\{a_1, \ldots, a_q, a_1^{-1}, \ldots, a_q^{-1}\}$ of the symmetric generating set of F. For example, $c_{\text{in}}(a_1a_2, a_1^{-1}) = e$, $c_{\text{in}}(a_1a_2, (a_1a_2^{-1})^2) = a_1$, etc.

Lemma 7.32. *For all $g, h \in F$, we have h and h^{-1} belong to $B(g)$ if and only if $h = e$.*

Démonstration. 'If' direction being clear, we only prove the 'only if' part. One starts by observing that for all $g, h \in F$, we have $l(gh) = l(g)+l(h)-2l(c_{\text{in}}(g^{-1}, h))$. Second, recall that by definition, $h \in B(g)$ means $d(h,g) \leq l(g)$ and since $d(h,g) = l(h^{-1}g)$, $h \in B(g)$ implies $l(h) + l(g) - 2l(c_{\text{in}}(g^{-1}, h^{-1})) \leq l(g)$, i.e. $l(h) \leq 2l(c_{\text{in}}(g^{-1}, h^{-1}))$. If, moreover, $h^{-1} \in B(g)$, we similarly have $l(h) = l(h^{-1}) \leq 2l(c_{\text{in}}(g^{-1}, h))$. These two inequalities say that h and h^{-1} have a common initial prefix of length at least the half of $l(h) = l(h^{-1})$. This indeed implies that $h = e$. □

We now give the details of the

Proof of Lemma 7.31. Recall before starting that for each $g \in F \setminus \{e\}$, writing its cyclic decomposition $g = c(g)r(g)c(g)^{-1}$, $\text{axe}(g)$ is based at the vertex $c(g)$ and $\tau(g) = |r(g)|$. In particular, for a cyclically reduced element, $|g| = \tau(g)$. We will only provide the argument for the case g is cyclically reduced ; for a cyclically non-reduced element, the argument goes the same replacing e by $c(g)$ below. Note also that, the notion of a defective element is, of course, invariant by conjugation. We will proceed by a case-by-case analysis. Precisely, we show that the hypothesis of the lemma implies that

$$\text{web}(S \setminus \{g, g^{-1}\}) \cap \text{axe}(g) \subset B(e, \left\lfloor \frac{\tau(g) - 1}{2} \right\rfloor) \tag{7.3}$$

where $B(x,r)$ denotes a ball of centre x and radius r. This yields the result in view of definitions.

For a contradiction, suppose that (7.3) does not hold. Then, two cases are possible : either, up to interchanging g by g^{-1}, there exists a cyclically reduced element $h \in S \setminus \{g, g^{-1}\}$ such that $(\text{axe}(h) \cap \text{axe}(g)) \cap [\![(e)_{(\lfloor \frac{\tau(g)+1}{2} \rfloor) \to \xi_g^+}, \xi_g^+[\![\neq \emptyset$; or, there exists an element $h \in S$, not cyclically-reduced, such that, again up to interchanging h by h^{-1}, $p_{\text{axe}(g)}(\text{axe}(h)) \cap [\![(e)_{(\lfloor \frac{\tau(g)+1}{2} \rfloor) \to \xi_g^+}, \xi_g^+[\![\neq \emptyset$. We claim that in both cases, we have either $g \in B(h)$ or $h \in B(g)$.

For the first case, there are further two possibilities : suppose first $\tau(h) \leq \lfloor \frac{|g|+1}{2} \rfloor$, then, it is easy to see that in this case, in fact, h belongs to $[\![e, g]\!]$, and in particular, $h \in B(g)$. Suppose then $\tau(h) > \lfloor \frac{|g|+1}{2} \rfloor$. But then, either $h \in [\![e, g]\!]$, in which case $h \in B(g)$ and $g \in B(h)$; or $h \notin [\![e, g]\!]$ in which case we have $d(p_{\text{axe}(g)}(h), g) \leq \lfloor \frac{|g|-1}{2} \rfloor$, and therefore, $g \in B(h)$.

For the second case, one proceeds by a similar argument : first, suppose that either $\text{axe}(h)$ is based on a vertex of $[\![(e)_{(\lfloor \frac{\tau(g)+1}{2} \rfloor) \to \xi_g^+}, \xi_g^+[\![$ or its base vertex projects (by $p_{\text{axe}(g)}(.)$) to such a vertex. Then, since we also now that h is based at $c(h)$ in its cyclic decomposition, we deduce that $g \in B(h)$. Second, suppose that we are not in the first case, and thus, h is based on vertex of $[\![(e)_{1 \to \xi_g^+}, (e)_{\lfloor \frac{\tau(g)-1}{2} \rfloor \to \xi_g^+}]\!]$ and $\text{axe}(h)$ intersects $[\![(e)_{(\lfloor \frac{\tau(g)+1}{2} \rfloor) \to \xi_g^+}, g]\!]$. In this case, if $p_{\text{axe}(g)}(c(h)r(h)) \in [\![(e)_{(\lfloor \frac{\tau(g)+1}{2} \rfloor) \to \xi_g^+}, \xi_g^+[\![$, then one sees that $g \in B(h)$. If not, then $p_{\text{axe}(g)}(c(h)r(h)) \in [\![(e)_{1 \to \xi_g^+}, (e)_{\lfloor \frac{\tau(g)-1}{2} \rfloor \to \xi_g^+}]\!]$, and then $h \in B(g)$, concluding the proof of our claim. □

Preparing to prove Theorem 7.27, while bounding the spectral radius of μ by using a probabilistic decomposition argument, we will use the following probabilistic lemma together with Lemma 7.29. Recall that $D(S)$ denotes the defective elements in a subset $S \subset F$.

Lemma 7.33. *Let μ be a symmetric probability measure on $F \setminus \{e\}$, $r \in \mathbb{N}$, and $X_1, \dots, X_r$ be F-valued independent random variables with distribution μ. Suppose that for a $\delta > 0$, for each $g \in supp(\mu)$, we have $\mu(B(g)) \leq \delta$. Denote by* Samp_r *the μ-random subset $\{X_1, \dots, X_r, X_1^{-1}, \dots, X_r^{-1}\}$ of F. Then, for all $k \geq 0$, we have*

$$\mathbb{P}(|D(\mathrm{Samp}_r)| \geq k) \leq \binom{r - \lceil \frac{k}{2} \rceil}{2} 8\delta$$

Remark 7.34. *One can obtain, without substantial effort, a better upper bound in this lemma, but as we will see, we will only need that the right hand side goes to zero as δ goes to zero.*

Démonstration. By Lemma 7.31, we have the following inclusion of events :

$$\{\forall\, i \neq j \in \{1, \dots, r - \lceil \frac{k}{2} \rceil\},\ \{X_i, X_i^{-1}\} \cap (B(X_j) \cup B(X_j^{-1})) = \emptyset\} \subseteq \{|D(\mathrm{Samp}_r)| < k\}$$

so that, by looking at the complements and evaluating the probabilities, we have

$$\begin{aligned}\mathbb{P}(|D(\mathrm{Samp}_r)| \geq k) &\leq \mathbb{P}\Big(\bigcup_{\substack{i \neq j \\ \in \{1, \dots, r - \lceil \frac{k}{2} \rceil\}}} \{X_i \in B(X_j) \text{ or } \dots \text{ or } X_j^{-1} \in B(X_i^{-1})\}\Big) \\ &\leq \binom{r - \lceil \frac{k}{2} \rceil}{2} 8\mathbb{P}(X_1 \in B(X_2))\end{aligned}$$

where in the last inequality we used (twice) the union of events bound and the symmetry of μ. The statement of the lemma follows by remarking that the independence of X_i's and the hypothesis yields $\mathbb{P}(X_1 \in B(X_2)) \leq \delta$. □

We are now in a position to give the

Proof of Theorem 7.27. Note that for a symmetric probability measure μ on F and $y \in F$, the push-forward $\tau_{y*}\mu$ of μ by conjugation τ_y by y, is still a symmetric probability measure with the same mass at identity. It will be clear from our argument that it suffices to prove the statement of the theorem for $y = e$.

Let us first show that, it suffices to prove the statement for finitely supported symmetric probability measures, since one can 'approach' an infinitely supported one by its larger and larger finitely supported restrictions. More precisely, suppose that the result of the theorem is true for finitely supported symmetric probability measures with the constant δ_ϵ^f for each $1 \geq \epsilon \geq 0$. Set $\delta_\epsilon := \frac{\delta_{\epsilon/2}^f}{2}$ and let μ be an infinitely supported symmetric probability measure with $r(\mu) \geq \epsilon$ (it is clear that the constants δ_ϵ also satisfy the conclusion of the theorem for finitely supported symmetric ones). Since all the probability measures on F are σ-finite, there exists a finite subset K of $supp(\mu)$ with $\alpha := \mu(K) \geq (1 - \frac{\epsilon}{2})$. Denoting respectively by $\mu_{|K}$ and $\mu_{|\infty}$ the restrictions of μ to K and to $supp(\mu) \setminus K$, we can write $\mu = \alpha\mu_{|K} + (1 - \alpha)\mu_{|\infty}$. As a result, we have $\epsilon \leq r(\mu) \leq \alpha r(\mu_{|K}) + (1 - \alpha) r(\mu_{|\infty}) \leq \alpha r(\mu_{|K}) + \frac{\epsilon}{2} r(\mu_{|\infty}) \leq$

$\alpha r(\mu_{|K}) + \frac{\epsilon}{2}$, which yields $r(\mu_{|K}) \geq \frac{\epsilon}{2}$. Applying the result for finitely supported symmetric probability measures to $\mu_{|K}$, we get that $\mu_{|K}$ charges a ball of one of its elements larger than $\delta^f_{\epsilon/2}$. In particular, μ charges this ball larger than δ_ϵ. Therefore, we only prove the theorem for finitely supported symmetric probability measures and consider the contrapositive statement.

Start by separating the identity element from the support of μ by decomposing μ as $\mu = (1-\mu(e))\mu_0 + \mu(e)\delta_e$ where $\mu_0 := (1-\delta_e)\frac{\mu}{1-\mu(e)}$ (note that our contrapositive hypothesis implies that $\mu(e) \leq \delta$, $0 < \delta$ to be chosen later). As a result, applying the left-regular representation λ and looking at the operator norm, we get

$$r(\mu) \leq (1-\mu(e))r(\mu_0) + \mu(e) \leq r(\mu_0) + \delta \tag{7.4}$$

so that we are led to bound $r(\mu_0)$ from above.

Now, let $r \geq 1$ and for $\gamma_1, \ldots, \gamma_r \in F$, write $\nu_{\gamma_1,\ldots,\gamma_r} = \frac{\delta_1+\ldots+\delta_{\gamma_r}+\delta_{\gamma_1^{-1}}+\ldots+\delta_{\gamma_r^{-1}}}{2r}$. We readily have the following 'probabilistic decomposition' of μ_0 into uniform probability measures :

$$\mu_0 = \sum_{\gamma_1,\ldots,\gamma_r \in F} \nu_{\gamma_1,\ldots,\gamma_r}\mu_0(\gamma_1)\ldots\mu_0(\gamma_r)$$

To this equality, applying the left-regular representation and taking the operator norm, by triangle inequality, one gets

$$r(\mu_0) \leq \sum_{\gamma_1,\ldots,\gamma_r} r(\nu_{\gamma_1,\ldots,\gamma_r})\mu_0(\gamma_1).\ldots.\mu_0(\gamma_r) \tag{7.5}$$

Now, denoting by $\mathbb{E}_\mu$ the expectation with respect to the distribution of a random subset of F of type $\{X_1^{\pm 1}, \ldots, X_r^{\pm 1}\} =: \mathrm{Samp}_r$ where X_i's are independent F-valued random variables with distribution μ, the following equation is merely a rewriting of its left hand side :

$$\sum_{\gamma_1,\ldots,\gamma_r} r(\nu_{\gamma_1,\ldots,\gamma_r})\mu_0(\gamma_1).\ldots.\mu_0(\gamma_r) = \mathbb{E}_\mu[r(\nu_{\mathrm{Samp}_r})] \tag{7.6}$$

For, $k \leq r$, we can then write

$$\begin{aligned}
\mathbb{E}_\mu[r(\nu_{\mathrm{Samp}_r})] &= \mathbb{E}[r(\nu_{\mathrm{Samp}_r})\mathbb{1}_{|D(\mathrm{Samp}_r)| \geq k}] + \mathbb{E}[r(\nu_{\mathrm{Samp}_r})\mathbb{1}_{|D(\mathrm{Samp}_r)| < k}] \\
&\leq \mathbb{P}(|D(\mathrm{Samp}_r)| \geq k) + \mathbb{E}[r(\nu_{\mathrm{Samp}_r})\mathbb{1}_{|D(\mathrm{Samp}_r)| < k}] \\
&\leq \binom{r - \lceil \frac{k}{2} \rceil}{2} 8\delta + \sqrt{2} r^{-\frac{1}{2}}(2 + \lceil \frac{k}{2} \rceil)^{\frac{1}{2}}
\end{aligned} \tag{7.7}$$

where on the first inequality we used the fact that the spectral radius of a probability measure is at most one, and on the last, Lemma 7.33 for the first term and Lemma 7.29 for the second. Finally, putting (7.4), (7.5) and (7.7) together, we get

$$r(\mu) \leq \delta + \binom{r - \lceil \frac{k}{2} \rceil}{2} 8\delta + \sqrt{2} r^{-\frac{1}{2}}(2 + \lceil \frac{k}{2} \rceil)^{\frac{1}{2}} \tag{7.8}$$

In (7.8) specializing, for instance, to $k = \lfloor r^{\frac{1}{2}} \rfloor$, we get

$$r(\mu) \leq 8r^2\delta + 3r^{-\frac{1}{4}} \tag{7.9}$$

So that choosing any $0 < \delta_\epsilon < \frac{\epsilon^9}{16.6^8}$, one sees that specializing to for example $r = \lceil \frac{6^4}{\epsilon^4} \rceil$ in (7.9), we get $r(\mu) \leq \epsilon$. $\square$

Remark 7.35. *Another idea to use this Lemma 7.29 is to decompose S into relatively smaller components so as to minimise in each component the proportion of defective elements. By restricting the initial (for simplicity say, uniform) probability measure on these components, and working in each component, in some situations, one can get a fairly good control over the spectral radius of the initial probability measure on S. This is the way we obtain other results on rigidity of spectral radius in free group, in a work in progress (see the following remark).*

Remark 7.36. *(Further results) In this remark, we briefly report on our further results/observations around the 1 percent question that we will expose in an upcoming work. The common theme is to restrict to a class of symmetric probability measures of a support with a direct (1. below) or a quantitative (3. below) control over the geometry of its elements.*

1. A first observation is an elementary one ; its proof consists of applying the idea mentioned in Remark 7.35 to the support of the second convolution power of μ. It reads : the answer to the 1 percent question is affirmative, when the question is restricted to the symmetric probability measures whose support consists of non quasi-defective (i.e. strictly defective or non-defective) elements of pairwise disjoint translation axes.

2. Our second observation suggests asking other intermediate questions. It reads : the answer to the 1 percent question is affirmative, when restricted to the symmetric probability measures whose support K satisfies the property that the elements of ${}^2K = \{g^2 \mid g \in K\}$ are all non-defective. In particular, when K_n is a sequence of such finite subsets of F with $|K_n| \underset{n\to\infty}{\longrightarrow} \infty$ and μ_n denote the uniform probability measure on K_n, we have $r(\mu_n) \underset{n\to\infty}{\longrightarrow} 0$. For further intermediate questions, recall (see the proof of Lemma 7.21) that for any finite set K, for all $k \in \mathbb{N}$ large enough, ${}^kK$ consists of non-defective elements.

3. For another result, we introduce a numerical sequence that we associate to a finite subset K of the group F. We call this, the dispersion entropy of K. Intuitively, this sequence expresses the dispersion rate of the translation axes of elements of K. Using 2. above, we obtain partial results on the classes of symmetric probability measures with a support of bounded entropy and consisting of cyclically reduced elements.

Finally, we would do mention that the two type of sets, i.e. consisting of cyclically reduced elements (3. above) or of elements of pairwise disjoint translation axes (1. above) correspond to, in a sense, extreme cases for the geometry of translation axes.

Annexe A

Appendices

A.1 A geometric/ping-pong approach to Nielsen's cancellation theory

In Section 7.2, we introduced the notion of defective/non-defective elements, mainly for their application in Section 7.3 to the study of spectral radius. As it was mentioned in Remark 7.14, it turns out that this geometric-dynamical notion is in close relation to some of the ideas in Nielsen's cancellation theory in the free groups, and therefore, can be used to (re)prove algebraic properties and shed geometric light on the existent results (again, cf. Remark 7.14). We have already seen instances of these, namely in the proofs of Corollary 7.18, Lemma 7.19, Lemma 7.21 and Lemma 7.23. Moreover, in Lemma 7.17, by using defective elements and the ping-pong lemma, we have formulated a criterion for a subset S of a free group F to be free. The aim of this appendix is to exhibit a simple concrete relation between this notion and Nielsen's theory. As already said in the first paragraph of Section 7.1, by this, we only aim at giving a glimpse of an ongoing project, in which we shall give a geometric description of Nielsen's cancellation theory, specifying furthermore a geometric algorithm for the reduction.

Let us recall a definition from Lyndon-Schupp's [86] (Chapter 1) : let F be a free group on a basis S and let $l(.)$ denote the word length function associated to S. A a subset U of F is called N-reduced, if for all $v_1, v_2, v_3 \in U \cup U^{-1}$, one has

$(N0)$ $v_1 \neq e$

$(N1)$ $v_1 v_2 \neq e$ implies $l(v_1 v_2) \geq l(v_1) \vee l(v_2)$

$(N2)$ $v_1 v_2 \neq e$ and $v_2 v_3 \neq e$ implies $l(v_1 v_2 v_3) > l(v_1) - l(v_2) + l(v_3)$

The proof of Nielsen's theorem consists of showing first, that an N-reduced set freely generates a subgroup in F, and second, any finite set V can be carried by Nielsen transformations into an N-reduced set $U \subset F$. We shall show that through our geometric understanding, the first part in this proof reduces to Klein's ping-pong lemma (Lemma 7.15).

We first mention an observation on the translation axes of hyperbolic transformations of a tree T : as it is remarked in its proof by Tits in [112], one aspect of 3.

of Theorem 7.11 is that it comes with an efficient method to determine the translation axis of an element g : one way to determine the axis of an element g is to remark that for two neighbours $x, y \in V(T)$, we have $[\![x, y]\!] \subset \text{axe}(g)$ if and only if $d(g.x, x) = d(g.y, y)$, in which case we also have $\text{axe}(g) = \bigcup_{n\in\mathbb{Z}} g^n.[\![x, g.x]\!]$. Equivalently, one can express this using the notion of orientation : let x, y be two neighbour vertices of T and, up to exchanging x and y, let ξ be an element of $B(T)$ such that $[\![x, \xi[\![$ contains $[\![x, y]\!]$ and $[\![g.x, g.y]\!]$. If the orientation induced by ξ on the segments $[\![x, y]\!]$, $[\![g.x, g.y]\!]$ is the same, then $[\![x, y]\!] \subset \text{axe}(g) = \bigcup_{n\in\mathbb{Z}} g^n.[\![x, g.x]\!]$.

Let us come back to our setting of a free group on the basis S (symmetric free generating set) and recall that for $g, h \in F$, $c_{\text{in}}(g, h)$ the element of F corresponding to the common initial prefix of g and h seen as (reduced) words in the alphabet S, and to clarify the appearing quantities set $c_{\text{end}}(g, h) = c_{\text{in}}(g^{-1}, h^{-1})$. Now, let U be an N-reduced set in F and let us first write conditions $(N1)$ and $(N2)$ in different terms : for $v_1, v_2; v_3 \in F$, we clearly have $l(v_1v_2) = l(v_1) + l(v_2) - 2l(c_{\text{in}}(v_1^{-1}, v_2))$ and in case $l(v_1v_2v_3) > l(v_1) - l(v_2) + l(v_3)$, we have $l(v_1v_2v_3) = l(v_1) + l(v_2) + l(v_3) - 2l(c_{\text{in}}(v_1^{-1}, v_2)) - 2l(c_{\text{in}}(v_2^{-1}, v_3))$. As a result for all $g, h, k \in U$, respectively, $(N1)$ writes and $(N2)$ implies :

$$(D1)\ \ l(g) \geq 2l(c_{\text{in}}(g, h^{-1})) \vee 2l(c_{\text{end}}(g, h^{-1}))$$
$$(D2)\ \ l(g) > l(c_{\text{in}}(g, h^{-1})) + l(c_{\text{end}}(g, k^{-1}))$$

The following lemma shows that a set of N-reduced elements consist of non-defective elements. Its simple proof uses highly technical notation (!), we recommend the interested reader to follow it through a picture.

Lemma A.1. *Let U be a subset of $F \setminus \{e\}$ satisfying $(D1)$ and $(D2)$. Then U consists of non-defective elements, i.e. $D(U) = \emptyset$.*

Démonstration. Up to enlarging U, we can suppose that U is symmetric (recall that, by definition, if an element g of a subset S of F is S-non-defective, then it is also S_0-non-defective for any subset S_0 of S). Recall also that for the labelled structure of the (right) Cayley tree T of F, considering the (left) action of F on T, for each $g \in F \setminus \{e\}$, $\text{axe}(g)$ is based at $c(g)$ (recall that T is marked at e, Section 7.2) where $g = c(g)r(g)c(g)^{-1}$ is the cyclic decomposition of g. Now, let g be in U, and let us show that g is non-defective. since the notion of a non-defective element and $(N1)$ and $(N2')$ are preserved under conjugation, up to conjugating with $c(g)^{-1}$, we can suppose that g is cyclically reduced. One then proceeds with a case-by-case analysis to eliminate, by contradiction, all the cases that renders g defective.

Start by observing that, up to exchanging g by g^{-1}, for a cyclically non-reduced $h \in U$, $\text{axe}(h)$ is necessarily based at or its base points projects to (by $p_{\text{axe}\, g}$) to

$$[\![(e)_{\lfloor \frac{l(g)}{2} \rfloor \to \xi_g^-}, (e)_{\lfloor \frac{l(g)}{2} \rfloor \to \xi_g^+}]\!] \tag{A.1}$$

This follows directly by $(D1)$ and the fact that $\text{axe}(h)$ is based at $c(h)$. Suppose now that there exists a cyclically reduced $h \in U$ such that $\text{axe}(g) \cap \text{axe}(h) \cap [\![(e)_{\lceil \frac{l(g)+1}{2} \rceil \to \xi_g^+}, \xi_g^+[\![\neq$

$\emptyset$. But then, we must have $h \in [\![e, (e)_{\lfloor \frac{l(g)}{2} \rfloor \to \xi_g^+}]\!]$, since otherwise $(D1)$ is violated. But in this case, we have $c_{\text{in}}(g, h) = c_{\text{in}}(g, (h^{-1})^{-1}) = h$ and $(D1)$ is also violated. As a result, reasoning similarly for g^{-1}, we see that the axis of any cyclically reduced element h of U intersects $\text{axe}(g)$ on

$$[\![(e)_{\lfloor \frac{l(g)}{2} \rfloor \to \xi_g^-}, (e)_{\lfloor \frac{l(g)}{2} \rfloor \to \xi_g^+}]\!] \tag{A.2}$$

In view of (A.1) and (A.2), g can only be quasi-defective (hence defective) and this happens only if $l(g)$ is an even integer and there exist $h, k \in U$ such that $\text{axe}(h)$ and $\text{axe}(k)$ projects to (by $p_{\text{axe}(g)}$) or passes through, respectively $(e)_{\frac{l(g)}{2} \to \xi_g^-}$ and $(e)_{\frac{l(g)}{2} \to \xi_g^+}$. It remains thus to see that this is not possible. One observes that if both h and k are cyclically reduced, h or k can not be contained in $[\![(e)_{\frac{l(g)}{2} \to \xi_g^-}, (e)_{\frac{l(g)}{2} \to \xi_g^+}]\!]$ (by (A.2) applied to $\text{axe}(h)$ and $\text{axe}(k)$). But then this situation directly contradicts $(D2)$. Therefore suppose h is not cyclically reduced and k is cyclically reduced (or vice versa). Then, again by (A.2), $l(k) \geq l(g)$ and $[\![e, \xi_g^+[\![\cap \text{axe}(k) = [\![e, (e)_{\frac{l(g)}{2} \to \xi_g^+}]\!]$. In this situation, $\text{axe}(h)$ cannot have $(e)_{\frac{l(g)}{2} \to \xi_g^-}$ as its base point or the projection by $p_{\text{axe}(g)}$ of its base point, since this would again contradict $(D2)$. So, suppose $\text{axe}(h)$ is based on a vertex of $[\![e, (e)_{(\frac{l(g)}{2} - 1) \to \xi_g^-}]\!]$ (note that since $(e)_{\frac{l(g)}{2} \to \xi_g^-} \in \text{axe}(h)$, $\text{axe}(h)$ can not be based elsewhere). But then, for the cyclic decomposition $h = c(h)r(h)c(h)^{-1}$, we must have $c(h)r(h) \in [\![e, (e)_{\frac{l(g)}{2} \to \xi_g^-}]\!]$, since otherwise we again contradiction to $(D2)$. On the other hand, this case gives contradiction to $(D1)$ since $2l(c(h)r(h)) > l(c(h))$ and $l(c_{\text{in}(g,h)}) \geq l(c(h)r(h))$. The remaining case i.e. if h and k are both cyclically non-reduced, is treated very similarly with the last case. we leave out the details. $\square$

Therefore, Lemma 7.17 directly yields

Corollary A.2. *Let U be a symmetric subset of $F \setminus \{e\}$ satisfying $(D1)$ and $(D2)$. Then U freely generates a subgroup in F.* $\square$

We now state a lemma that lists a detailed description of the axis of product of two elements gh in terms of the axes of these elements. This lemma is not used elsewhere in this text (it is related to the announced result in Remark 7.36). It basically follows form the observation, that we mentioned above, on the translation axes of hyperbolic transformations of a tree. Before stating it, note that for two intersecting elements g and h, we have $\text{axe}(g) \cup \text{axe}(h) = \text{web}(\{g, h\})$, and one can not have $l(c(g, h)) \geq \tau(g) = \tau(h)$ unless $g = h$ or $g = h^{-1}$. This last observation follows easily by the fact that the F-action on T is free.

Lemma A.3 (Axe mouvements lemma). *1. (Common axis case) Let $g, h \in F \setminus \{e\}$ with $A := \text{axe}(g) = \text{axe}(h)$.*

(a) If $\tau(g) = \tau(h)$, one has $g = h$ or $g = h^{-1}$ respectively, if $\xi_g^+ = \xi_h^+$ or $\xi_g^+ = \xi_h^-$.

(b) If $g^{-1} \neq h$, one has $\text{axe}(gh) = A$. Moreover, if $\xi_g^+ = \xi_h^+$, we have $\tau(gh) = \tau(g) + \tau(h)$ and $\xi_{gh}^+ = \xi_g^+$; and if $\xi_g^+ = \xi_h^-$ supposing without loss of generality, $\tau(g) > \tau(h)$, one has $\tau(gh) = \tau(g) - \tau(h)$ and $\xi_{gh}^+ = \xi_g^+$.

2. *(Intersecting elements, same direction) Let g, h be two elements intersecting in the same direction. Then, we have $\tau(gh) = \tau(g) + \tau(h)$, $\mathrm{axe}(gh) \cap \mathrm{web}\{g, h\} = [\![(o(c(g,h)))_{\tau(h)\to\xi_h^-}, (e(c(g,h)))_{\tau(g)\to\xi_g^+}]\!]$, and the pairwise intersecting three elements gh, g and h are directed similarly.*
3. *(Intersecting elements, opposite direction) Let g, h be two elements intersecting in the opposite direction.*
   (a) *If $\tau(g), \tau(h) \geq d := l(c(g,h))$, then we have $\tau(gh) = \tau(g) + \tau(h) - 2d$, $\mathrm{axe}(gh) \cap \mathrm{web}(\{g, h\}) = [\![(e(c(g,h)))_{(\tau(h)-d)\to\xi_h^-}, (e(c(g,h)))_{(\tau(g)-d)\to\xi_g^+}]\!]$, and whenever gh is intersecting with g or h, i.e. $\tau(g) > d$ or $\tau(h) > d$, it is oriented similarly with them.*
   (b) *If $\tau(g) \geq d \geq \tau(h)$ or $d \geq \tau(g) \geq \tau(h)$, we have $\tau(gh) = \tau(g) - \tau(h)$, $\mathrm{axe}(gh) \cap \mathrm{web}(\{g, h\}) = [\![(o(c(g,h)))_{\tau(h)\to\xi_h^-}, (e(c(g,h)))_{(\tau(g)-\tau(h))\to\xi_g^+}]\!]$, and gh intersects with g in the same direction.*
   (c) *If $\tau(h) \geq d \geq \tau(g)$ or $d \geq \tau(h) \geq \tau(g)$, we have $\tau(gh) = \tau(h) - \tau(g)$, $\mathrm{axe}(gh) \cap \mathrm{web}(\{g, h\}) = [\![(o(c(g,h)))_{\tau(g)\to\xi_g^+}, (e(c(g,h)))_{(\tau(h)-\tau(g))\to\xi_h^-}]\!]$, and gh intersects with h in the same direction.*

Démonstration. 1.(a) follows by the fact that the F-action on T is free. 1.(b) is clear by the algorithm to determine the translation axis of an element explained previously to the statement of this lemma. 2 and 3 follow by a case-by-case analysis, where each case consists of a straightforward application of the aforementioned algorithm. We omit the details which won't shed more light than a simple picture drawn by the reader. Finally, we note that 3.(c) follows directly from 3.(b) considering $(gh)^{-1}$. □

A.2 Some calculations for the free group

Number of elements of F of given length and translation distance

Fix an alphabet $A := \{a_1, \ldots, a_q, a_1^{-1}, \ldots, a_q^{-1}\}$. For a reduced word g in A, we call its cyclic length, the length of its cyclic component $c(g)$ in its cyclic decomposition $g = c(g)r(g)c(g)^{-1}$ seen as an element of F_q with the free generating set A. In this appendix, we give the exact expression of the number of reduced words in A of length $n \in \mathbb{N}$ and of cyclic length $k \in \mathbb{N}$.

For $k, n \in \mathbb{N}$, denote by w_n the number of reduced words of length n, so that it is equal to $2q(2q-1)^{n-1}$ and by $c_{k,n}$, the number of reduced words of length n and of cyclic length k. Obviously, for each $k \geq 0$ with $2k \geq n$, we have $c_{k,n} = 0$. Furthermore, it is not hard to see that one has the following upper and lower bounds for $c_{k,n}$, which were sufficient our asymptotic considerations in the "L'exemple du groupe libre" of the introduction of the part 1 of this text

$$2q(2q-1)^{n-k-2}(2q-3) \leq c_{k,n} \leq 2q(2q-1)^{n-k-1} \tag{A.3}$$

To calculate the exact expression of $c_{k,n}$, it turns out that one just needs to take a closer look into the set of reduced words of length n and the appearing relation between them. Accordingly, we consider the following decomposition of w_n : we set

$w_{n,1^+}$ = the number of reduced words of length n starting and ending with the same letter
$w_{n,1^-}$ = the number of reduced words of length n starting and ending with opposite letters
$w_{n,2}$ = the number of reduced words of length n starting and ending with $\alpha, \beta \in A$ such that $\alpha \notin \{\beta, \beta^{-1}\}$

Therefore, we can write

$$2q(2q-1)^{2n-1} = w_n = w_{n,1^+} + w_{n,1^-} + w_{n,2} \tag{A.4}$$

Moreover, we have the following relations between these terms :

Lemma A.4. *For $n \geq 2$, we have*

$$w_{n+1,1^+} = w_{n,2} + w_{n,1^+} \tag{1}$$
$$w_{n+1,1^-} = w_{n,2} + w_{n,1^-} \tag{2}$$
$$w_{n+1,2} = (2q-2)w_{n,1^+} + (2q-2)w_{n,1^-} + (2q-3)w_{n,2} \tag{3}$$

Démonstration. The equations follow in a straightforward manner, for example, by fixing the last letter of the word of length $n+1$ on the left hand side of the equations, and considering the word type (in the sense of $w_{n,1^+}, w_{n,1^-}, w_{n,2}$) and the possibilities of the prefix of length n. □

Looking at (1) and (2) of the previous lemma, one sees that the difference $w_{n,1^+} - w_{n,1^-}$ is constant for $n \geq 2$. Calculating for $n = 2$, we have $w_{2,1^-} = 0$ and $w_{2,1^+} = 2q$, so that for all $n \geq 2$, one has

$$w_{n,1^+} = w_{n,1^-} + 2q \tag{A.5}$$

Plugging (A.5) into (A.4), for all $n \geq 2$, we get

$$w_{n,2} = 2q(2q-1)^{n-1} - 2q - 2w_{n,1^-} \tag{A.6}$$

Now, plugging (A.6) into (2) of Lemma A.4, for all $n \geq 1$, we get the following simple recursive relation for $w_{n,1^-}$:

$$w_{n+1,1^-} + w_{n,1^-} = 2q(2q-1)^{n-1} - 2q \tag{A.7}$$

Since $w_{1,1^-} = 0$, solving (A.7), one gets its unique solution as :

$$w_{n,1^-} = \begin{cases} (2q-1)^{n-1} - (2q-1) & \text{if } n \text{ is even} \\ (2q-1)^{n-1} - 1 & \text{if } n \text{ is odd} \end{cases} \tag{A.8}$$

In view of (A.5) and (A.6), this gives the exact expression for $w_{n,1^-}$, $w_{n,1^+}$ and $w_{n,2}$. From this, one concludes by the following observation relating these quantities to $c_{k,n}$ (this lemma was indeed the reason to introduce these quantities) :

Lemma A.5. *For $0 \leq k < \lceil \frac{n}{2} \rceil \in \mathbb{N}$, one has*

$$c_{0,n} = w_n - w_{n,1^-} \qquad (1)$$
$$c_{k,n} = w_k \frac{q-1}{q} c_{0,n-2k} \qquad (2)$$

and $c_{k,n} = 0$ for all other $k \in \mathbb{N}$.

Démonstration. (1) is precisely the definition of $c_{0,n}$ and (2) follows by considering the possible constructions of these elements : w_k stands for the choice of the cyclic part, $c_{0,n-2k}$ stands for the cyclically reduced part which is necessarily of length $n-2k$, and the factor $\frac{q-1}{q}$ appears by considering the possibilities of the first letter of the cyclically reduced part of these elements. □

As a result, explicitly, we have the number of reduced and cyclically reduced words of length n as

$$c_{0,n} = \begin{cases} (2q-1)^n + (2q-1) & \text{if } n \text{ is even} \\ (2q-1)^n + 1 & \text{if } n \text{ is odd} \end{cases}$$

i.e., in a more compressed expression, $c_{0,n} = (2q-1)^n + (-1)^n(q-1) + q$, and thus

$$c_{k,n} = 2(2q-1)^{k-1}(q-1)\left((2q-1)^{n-2k} + (-1)^n(q-1) + q\right)$$

Number of paths in a homogeneous tree

In this part we give the exact expression of some quantities that we have used in the "L'exemple du groupe libre" part of the introduction of Part 1. One of these quantities was related to $c_{k,n}$'s that we expressed in the first part of this appendix. Here we give the number of paths of a fixed origin in a finite degree (q) homogeneous tree of length n and whose two extremities are of distance k (in particular, the number of words of length n and of reduced length k in an alphabet $\{a_1^{\pm 1}, \ldots, a_q^{\pm 1}\}$). Remark that in view of (1.3), (1.8), (1.12) and (1.13) of Part 1 and the previous subsection of this Appendix A.2, denoting by S_n the random walk on a free group F_q with respect to a uniform probability measure on a free generating set, this gives us the exact expressions of $\mathbb{P}(l(S_{2n}) = 2k)$, $\mathbb{P}(\tau(S_n) = 2k)$, $\mathbb{P}(b_\xi(S_{2n}) = 2k)$ for a $\xi \in \partial F_q$ (for this quantities see "l'exemple du groupe libre" in Chapter 1 or Section 7.2).

Let now T_q be a homogeneous tree of degree $q \in \mathbb{N}$ that we consider with its graph metric structure and $e \in V(T_q)$ a fixed vertex. For $n, k \in \mathbb{N}$, denote by $N_q(n,k)$ the number of paths of origin e of length n and of extremity at distance k to e : for clarity, a path of length n of origin e means n successive choices $x_1, \ldots, x_n$ of neighbouring vertices in T_q where x_1 is a neighbour of e (so that we set $x_0 := e$) and of extremity of distance k to e means $d(e, x_n) = k$. By evident relations, it suffices to calculate only $N_q(2n, 2k)$'s for $0 \leq k \leq n$. By using the $d(e, .)$ function along all such paths, we see that we can obtain a mapping from the set of these paths $((x_i)_{0 \leq i \leq 2n})$ into the set of paths of origin $(0,0)$ in $\mathbb{Z}^2$ in the following way (the author was informed by E. Breuillard that this argument had already been used by Kesten in [76], who is, in that work, only interested in an exponential equivalent of

$N_{2q}(2n,0))$: we set $(a_1,b_1)=(1,0)$ and for each $i\geq 1$ if $d(e,x_{i+1})=d(e,x_i)+1$, we set $(a_{i+1},b_{i+1})=(a_i,b_i)+(1,0)$ and if $d(e,x_{i+1})=d(e,x_i)-1$ we set $(a_{i+1},b_{i+1})=(a_i,b_i)+(0,1)$. It is easy to see that the map obtained in this way is surjective on the set of paths $((a_i,b_i))$ of origin $(0,0)\in\mathbb{Z}^2$, of extremity $(n+k,n-k)$ and which stays below the diagonal in $\mathbb{Z}^2$ (i.e. $a_i\geq b_i$ for each $i=0,\ldots,2n$) and such that $(a_{i+1}-a_i,b_{i+1}-b_i)\in\{(0,1),(1,0)\}$. The subtle but simple point is that this map is $q^{j+1}(q-1)^{n+k-j-1}$-to-1 on the paths which touch the diagonal exactly j times except at the origin and the end (i.e. $j=\#\{1\leq i<2n \mid a_i=b_i\}$). Accordingly, for $a,b\in\mathbb{N}$, set $\mathrm{pathz}(a,b)$ the number of paths in as above (image set of our mapping) of origin $(0,0)$ and extremity (a,b). For $0\leq j\leq b$, set $\mathrm{pathz}_j(a,b)$ the number of those touching the diagonal in $\mathbb{Z}^2$ exactly j times except for the origin and the end. For $0\leq j\leq b$ and $i_1\leq\ldots\text{ß}_j$, set $\mathrm{pathz}_j^{2i_1,\ldots,2i_j}(a,b)$ the number of those touching the diagonal at the steps $2i_1,\ldots,2i_j$ (i.e. $a_{i_t}=b_{i_t}$ for $1\leq t\leq j$, and note that a such path can touch the diagonal only on even steps). Then we clearly have

$$\begin{aligned}\mathrm{pathz}(a,b)&=\sum_{j=0}^{b}\mathrm{pathz}_j(a,b)\\ \mathrm{pathz}_j(a,b)&=\sum_{1\leq i_1<\ldots<i_j\leq b}\mathrm{pathz}_j^{2i_1,\ldots,2i_j}(a,b)\end{aligned}\tag{A.9}$$

And decomposing a paths that touches the diagonal j times, and setting $i_0=1$, we have

$$\mathrm{pathz}_j^{2i_1,\ldots,2i_j}(a,b)=\left(\prod_{r=1}^{j}\mathrm{pathz}_0(i_r-i_{r-1},i_r-i_{r-1})\right).\mathrm{pathz}_0(a-i_j,b-i_j)\tag{A.10}$$

As a result of (A.9) and (A.10), for our aims, one just needs to calculate $\mathrm{pathz}_0(a,b)$ for $a\geq b$ in $\mathbb{N}$. Let us first recall André's reflection trick (see [52], the ballot problem of Bertrand) to calculate $\mathrm{pathz}(a,b)$: it consists of decomposing all paths in $\mathbb{Z}^2$ of origin $(0,0)$ and end (a,b) and moving at each step to one unit 'right' $(+(1,0))$ or one unit 'up' $(+(0,1))$ into bad and good paths, where good ones stays below the diagonal $(a_k\geq b_k)$ and bad ones crosses the diagonal (i.e. there exists k with $a_k<b_k$). For each bad path, one considers the first step k such that $a_k<b_k$ and considers the path obtained by reflecting the first one from the step k on, with respect to the line $y=x+\frac{1}{2}$ (in other terms, one exchanges the up's and right's after the step k). This maps the set of bad paths bijectively onto the set of paths (of origin $(0,0)$ moving to the right and up) reaching $(b-1,a+1)$. As a result, the number of good paths, which is what we note by $\mathrm{pathz}(a,b)$, is equal to $\binom{a+b}{a}-\binom{a+b}{a+1}$, which is also $\binom{a+b}{a}\frac{a-b+1}{a+1}$. Now, to calculate $\mathrm{pathz}_0(a,b)$, one realises that these are just the good paths, who, after the first necessary right step, stay below the diagonal passing through $(1,0)$ except possibly for the very end step. By this observation, it easily follows that, for $a>b$, we have $\mathrm{pathz}_0(a,b)=\mathrm{pathz}(a-1,b)=\binom{a+b-1}{a-1}\frac{a-b}{a}$ and for $a=b$, $\mathrm{pathz}_0(a,a)=\mathrm{pathz}(a-1,a-1)=\frac{1}{a}\binom{2a-2}{a-1}$. Then, it follows by the above mentioned property of the mapping-that we constructed using the $d(e,.)$ function, from the paths on the tree of origin e and into $\mathbb{Z}^2$ paths-, (A.9) and (A.10) that, putting again $i_0 := 0$ and $\binom{0}{0} := 1$ for, for $n \geq k > 0$, we have,

$$N_q(2n, 2k) = \sum_{j=1}^{n-k}[q^{j+1}(q-1)^{n+k-1-j}(\sum_{1\leq i_1<\ldots<i_j\leq n-k}\{\frac{2k}{n+k-i_j}\binom{2n-2i_j-1}{n+k-i_j-1}\cdot \prod_{r=1}^{j}\frac{1}{i_r-i_{r-1}}\binom{2(i_r-i_{r-1}-1)}{i_r-i_{r-1}-1}\})] + q(q-1)^{n+k-1}\frac{2k}{n+k}\binom{2n-1}{n+k-1} \tag{A.11}$$

and for $k = 0$, putting $i_0 := 0$ and $i_{j+1} := n$, we have

$$N_q(2n, 0) = \sum_{j=1}^{n-1}[q^{j+1}(q-1)^{n-1-j}(\sum_{1\leq i_1<\ldots<i_j\leq n-1}\{\prod_{r=1}^{j+1}\frac{1}{i_r-i_{r-1}}\binom{2(i_r-i_{r-1}-1)}{i_r-i_{r-1}-1}\})] + q(q-1)^{n-1}\frac{1}{n}\binom{2n-2}{n-1} \tag{A.12}$$

Bibliographie

[1] Abért, M., Glasner, Y., and Virág, B., 2014. Kesten's theorem for invariant random subgroups. Duke Mathematical Journal, 163(3), pp.465-488.

[2] Akemann, C. A., and Ostrand, P. A., 1976. Computing norms in group C^*-algebras. American Journal of Mathematics, 98(4), pp.1015-1047.

[3] Albuquerque, P., 1999. Patterson-Sullivan theory in higher rank symmetric spaces. Geometric & Functional Analysis GAFA, 9(1), pp.1-28.

[4] Alperin, R., and Bass, H., 1987. Length functions of group actions on Λ-trees. Combinatorial group theory and topology, pp.265-378.

[5] Aoun, R., 2011. Random subgroups of linear groups are free. Duke Mathematical Journal, 160(1), pp.117-173.

[6] Avez, A., 1972. Entropie des groupes de type fini. CR Acad. Sci. Paris Sér. AB, 275(A1363-A1366), p.22.

[7] Avez, A., 1976. Croissance des groupes de type fini et fonctions harmoniques. In Théorie Ergodique (pp.35-49). Springer Berlin Heidelberg.

[8] Bahadur, R.R., 1971. Some limit theorems in statistics (No. 04 ; QA276. A1, B3 1971.). Philadelphia : Society for industrial and applied mathematics.

[9] Bahadur, R.R. and Zabell, S.L., 1979. Large deviations of the sample mean in general vector spaces. The Annals of Probability, pp.587-621.

[10] Benoist, Y., 1996. Actions propres sur les espaces homogènes réductifs. Annals of mathematics, pp.315-347.

[11] Benoist, Y., 1997. Propriétés asymptotiques des groupes linéaires. Geometric & Functional Analysis GAFA, 7(1), pp.1-47.

[12] Benoist, Y., 2000, Propriétés asymptotiques des groupes linéaires II. Advanced Studies Pure Math. 26 (2000) pp.33-48

[13] Benoist, Y., 1997, Lectures notes during École Éuropéenne de théorie des groupes, pp.1-70

[14] Benoist, Y. and Quint, J.F., 2013. Random walks on reductive groups. to appear.

[15] Benoist, Y. and Quint, J.F., 2016. Central limit theorem for linear groups. Ann. Probab, 44(2), pp.1308-1340.

[16] Abels, H., Margulis, G.A. and Soifer, G.A., 1995. Semigroups containing proximal linear maps. Israel journal of mathematics, 91(1-3), pp.1-30.

[17] Bekka, M., Cowling, M., and La Harpe, P. D., 1994. Some groups whose reduced C^*-algebra is simple. Publications Mathématiques de l'IHÉS, 80, pp.117-134.

[18] Bekka, M. B., & Mayer, M., 2000. Ergodic theory and topological dynamics of group actions on homogeneous spaces (Vol. 269). Cambridge University Press.

[19] Bellman, R., 1954. Limit theorems for non-commutative operations. I. Duke Mathematical Journal, 21(3), pp.491-500.

[20] Benoist, Y., and Labourie, F., 1993. Sur les difféomorphismes d'Anosov affines à feuilletages stable et instable différentiables. Inventiones mathematicae, 111(1), pp.285-308.

[21] Benoist, Y. and Quint, J.F., 2011 Mesures stationnaires et fermés invariants des espaces homogènes. Annals of mathematics 174(2), pp.1111-1162.

[22] Benoist, Y. and Quint, J.F., 2013. Stationary measures and invariant subsets of homogeneous spaces (II). Journal of the American Mathematical Society, 26(3), pp.659-734.

[23] Benoist, Y. and Quint, J.F., 2013. Stationary measures and invariant subsets of homogeneous spaces (III). Annals of Mathematics, 178(3), pp.1017-1059.

[24] Berger, M.A. and Wang, Y., 1992. Bounded semigroups of matrices. Linear Algebra and its Applications, 166, pp.21-27.

[25] Bochi, J., 2003. Inequalities for numerical invariants of sets of matrices. Linear algebra and its applications, 368, pp.71-81.

[26] Bochi, J. and Morris, I.D., 2014. Continuity properties of the lower spectral radius. Proceedings of the London Mathematical Society, p.pdu058.

[27] Borel, A., 2012. Linear algebraic groups (Vol. 126). Springer Science & Business Media.

[28] Bougerol, P., 1987. Tightness of products of random matrices and stability of linear stochastic systems. The Annals of Probability, pp.40-74.

[29] Bougerol, P., 1981. Théorème central limite local sur certains groupes de Lie. In Annales scientifiques de l'École Normale Supérieure, Vol. 14, No. 4, pp.403-432.

[30] Bougerol, P. and Lacroix J., 1985. Products of random matrices with applications to Schrödinger operators. Progress in Probability and Statistics, 8. Birkhäuser Boston, Inc., Boston, MA, xii+283 pp.

[31] Bousch, T. and Mairesse, J., 2002. Asymptotic height optimization for topical IFS, Tetris heaps, and the finiteness conjecture. Journal of the American Mathematical Society, 15(1), pp.77-111.

[32] Bourgain, J., Furman, A., Lindenstrauss, E. and Mozes, S., 2011. Stationary measures and equidistribution for orbits of nonabelian semigroups on the torus. Journal of the American Mathematical Society, 24(1), pp.231-280.

[33] Breuillard, E., 2005. Effective estimates for the spectral radius of a bounded set of matrices, Unpublished note, pp.1-13.

[34] Breuillard, E. and Gelander, T., 2007. A topological Tits alternative. Annals of mathematics, pp.427-474.

[35] Breuillard, E., 2008. A strong Tits alternative. arXiv preprint arXiv :0804.1395.

[36] Breuillard, E., 2011. A Height Gap Theorem for finite subsets of $GL_d(\overline{\mathbb{Q}})$ and non amenable subgroups, Annals of Math, volume 174-2, pp.1057-1110

[37] Breuillard, E., 2015. Lecture notes in “Masterclass on Groups, boundary actions and group C^*-algebras, Copenhagen".

[38] Breuillard, E., Personal communication

[39] Brezis, H., Ciarlet, P.G. and Lions, J.L., 1999. Analyse fonctionnelle : théorie et applications (Vol. 91). Paris : Dunod.

[40] Chatterji, I., and Ruane, K., 2005. Some geometric groups with rapid decay. Geometric & Functional Analysis GAFA, 15(2), pp.311-339.

[41] Cramér, H., 1938. Sur un nouveau théorème limite de la théorie des probabilités. In Actualités Scientifiques et Industrielles, number 736 in Colloque consacré à la théorie des probabilités, pp.5-23. Hermann, Paris.

[42] Culler, M., and Morgan, J. W., 1987. GROUP ACTIONS ON U-TREES. In Proc. London Math. Soc.(3) (Vol. 55, pp.571-604).

[43] Daubechies, I. and Lagarias, J.C., 1992. Sets of matrices all infinite products of which converge. Linear algebra and its applications, 161, pp.227-263.

[44] Dembo, A. and Zeitouni, O., 2009. Large deviations techniques and applications (Vol. 38). Springer Science & Business Media.

[45] Deuschel, J.D. and Stroock, D.W., 1989. Large deviations. American Mathematical Soc..

[46] Donsker, M.D. and Varadhan, S.S., 1975. Asymptotic evaluation of certain Markov process expectations for large time, I. Communications on Pure and Applied Mathematics, 28(1), pp.1-47.

[47] Donsker, M.D. and Varadhan, S.R.S., 1975. Asymptotic evaluation of certain Markov process expectations for large time, II. Communications on Pure and Applied Mathematics, 28(2), pp.279-301.

[48] Donsker, M.D. and Varadhan, S.R.S., 1976. Asymptotic evaluation of certain Markov process expectations for large time-III. Communications on pure and applied Mathematics, 29(4), pp.389-461.

[49] Ekeland, I. and Temam, R., 1976. Convex analysis and 9 variational problems. Siam.

[50] Ellis, R.S., 1988. Large deviations for the empirical measure of a Markov chain with an application to the multivariate empirical measure. The Annals of Probability, 16(4), pp.1496-1508.

[51] Eskin, A., Mozes, S. and Oh, H., 2005. On uniform exponential growth for linear groups. Inventiones mathematicae, 160(1), pp.1-30.

[52] Feller, W., 1968. An introduction to probability theory and its applications : volume I (Vol. 3). London-New York-Sydney-Toronto : John Wiley & Sons.

[53] Fine, B., Rosenberger, G., and Stille, M., 1995. Nielsen transformations and applications : a survey. In Groups-Korea (Vol. 94, pp.69-105).

[54] Freidlin, M.I. and Wentzell A.D. Random perturbations of dynamical systems. Translated from the 1979 Russian original by Joseph Szücs. Third edition. Grundlehren der Mathematischen Wissenschaften, 260. Springer, Heidelberg, 2012. xxviii+458 pp.

[55] Fulton, W. and Harris, J., 2013. Representation theory : a first course (Vol. 129). Springer Science & Business Media.

[56] Furstenberg, H., 1973. Boundary theory and stochastic processes on homogeneous spaces. Harmonic analysis on homogeneous spaces, 26, pp.193-229.

[57] Furstenberg, H., 1963. Noncommuting random products. Transactions of the American Mathematical Society, 108(3), pp.377-428.

[58] Furstenberg, H., 1963. A Poisson formula for semi-simple Lie groups. Annals of Mathematics, pp.335-386.

[59] Furstenberg, H. and Kesten, H., 1960. Products of random matrices. The Annals of Mathematical Statistics, 31(2), pp.457-469.

[60] Furstenberg, H. and Kifer, Y., 1983. Random matrix products and measures on projective spaces. Israel Journal of Mathematics, 46(1-2), pp.12-32.

[61] Goldsheid, I. and Margulis, G.,1989. Lyapunov Indices of a Product of Random Matrices, Russian Math. Surveys 44 pp.11-81.

[62] Gouëzel, S., Mathéus, F., and Maucourant, F., 2015. Entropy and drift in word hyperbolic groups. arXiv preprint arXiv :1501.05082.

[63] Guivarc'h, Y., 1980. Exposants de Liapunoff des marches aléatoires à pas markovien. Publications mathématiques et informatique de Rennes, (1), pp.1-16.

[64] Goldsheid, I.Y. and Guivarc'h, Y., 1996. Zariski closure and the dimension of the Gaussian low of the product of random matrices. I. Probability theory and related fields, 105(1), pp.109-142.

[65] Guivarc'h, Y., 2008. On the spectrum of a large subgroup of a semisimple group. Journal of modern dynamics, 2(1), p.15.

[66] Guivarc'h, Y., and Le Page, É., 2016. Spectral gap properties for linear random walks and Pareto's asymptotics for affine stochastic recursions. In Annales de l'Institut Henri Poincaré, Probabilités et Statistiques (Vol. 52, No. 2, pp.503-574). Institut Henri Poincaré.

[67] Guivarc'h, Y. and Raugi, A., 1985. Frontière de Furstenberg, propriétés de contraction et théorèmes de convergence. Probability Theory and Related Fields, 69(2), pp.187-242.

[68] Guivarc'h, Y., 1990. Produits de matrices aléatoires et applications aux propriétés géométriques des sous-groupes du groupe linéaire. Ergodic theory and dynamical systems, 10(03), pp.483-512.

[69] Guivarc'h, Y., 2006. Renewal theorems, products of random matrices, and toral endomorphisms. Potential Theory in Matsue, Adv. Stud. Pure Math. Math. Soc. Japan, 44, pp.53-66.

[70] Haagerup, U., 1978. An example of a non nuclear C^*-algebra, which has the metric approximation property. Inventiones Mathematicae, 50(3), pp.279-293.

[71] Hennion, H., and Hervé, L., 2001. Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness (Vol. 1766). Springer Science & Business Media.

[72] Humphreys, J.E., 2012. Linear algebraic groups (Vol. 21). Springer Science & Business Media.

[73] Jolissaint, P., 1990. Rapidly decreasing functions in reduced C^*-algebras of groups. Transactions of the American Mathematical Society, 317(1), pp.167-196.

[74] Kaijser, T., 1978. A limit theorem for Markov chains in compact metric spaces with applications to products of random matrices. Duke Math. J, 45(2), pp.311-349.

[75] Kaimanovich, V. A., & Vershik, A. M., 1983. Random walks on discrete groups : boundary and entropy. The annals of probability, pp.457-490.

[76] Kesten, H., 1959. Symmetric random walks on groups. Transactions of the American Mathematical Society, 92(2), pp. 336-354.

[77] Kesten, H., 1959. Full Banach mean values on countable groups. Mathematica Scandinavica, pp.146-156.

[78] Kostant, B., 1973. On convexity, the Weyl group and the Iwasawa decomposition. In Annales scientifiques de l'Ecole Normale Supérieure (Vol. 6, No. 4, pp.413-455).

[79] Kotani, M. and Sunada, T., 2006. Large deviation and the tangent cone at infinity of a crystal lattice. Mathematische Zeitschrift, 254(4), pp.837-870.

[80] Ledrappier, F., 1984. Quelques propriétés des exposants caractéristiques. In Ecole d'Été de Probabilités de Saint-Flour XII-1982 (pp. 305-396). Springer Berlin Heidelberg.

[81] Ledrappier, F. and Royer, G., 1980. Croissance exponentielle de certains produits aléatoires de matrices. In Acad. Sci (Vol. 280, pp.513-514).

[82] Ledrappier, F., 2001. Some asymptotic properties of random walks on free groups. Topics in probability and Lie groups : boundary theory, 28, pp.117-152.

[83] Lanford, O.E., 1973. Entropy and equilibrium states in classical statistical mechanics. In Statistical mechanics and mathematical problems (pp. 1-113). Springer Berlin Heidelberg.

[84] Le Page, E., 1982. Théorèmes limites pour les produits de matrices aléatoires. In Probability measures on groups (pp. 258-303). Springer Berlin Heidelberg.

[85] Liao, M., 2004. Lévy processes in Lie groups (Vol. 162). Cambridge university press.

[86] Lyndon, R. C., and Schupp, P. E., 2015. Combinatorial group theory. Springer.

[87] Magnus, W., Karrass, A., & Solitar, D., 2004. Combinatorial group theory : Presentations of groups in terms of generators and relations. Courier Corporation.

[88] Mostow, G.D., 1955. Self-adjoint groups. Annals of Mathematics, pp.44-55.

[89] Mostow, G.D., 2016. Strong Rigidity of Locally Symmetric Spaces.(AM-78) (Vol. 78). Princeton University Press.

[90] Nagaev, S.V., 1957. Some limit theorems for stationary Markov chains. Theory of Probability & Its Applications, 2(4), pp.378-406.

[91] Patterson, S.J., 1976. The limit set of a Fuchsian group. Acta mathematica, 136(1), pp.241-273.

[92] Paulin, F., 1988. Topologie de Gromov équivariante, structures hyperboliques et arbres réels. Inventiones mathematicae, 94(1), pp.53-80.

[93] Pays, I. and Valette, A., 1991. Sous-groupes libres dans les groupes d'automorphismes d'arbres. L'enseignement mathématique, 37, pp.151-174.

[94] Pete, G., (Book in preparation) Probability and geometry on groups.

[95] Pittet, C., & Saloff-Coste, L., 2001. A survey on the relationships between volume growth, isoperimetry, and the behavior of simple random walk on Cayley graphs, with examples.

[96] Powers, R. T., 1975. Simplicity of the C^*-algebra associated with the free group on two generators. Duke Math. J, 42(1), pp.151-156.

[97] Prasad, G., 1994. $\mathbb{R}$-regular elements in Zariski dense subgroups. The Quarterly Journal of Mathematics, 45(4), pp.542-545.

[98] Prasad, G. and Rapinchuk, A., 2003. Existence of irreducible $\mathbb{R}$-regular elements in Zariski-dense subgroups. Math. Res. Lett, 10, pp.21-32.

[99] Protasov, V. and Voynov, A., 2014. Matrix semigroups with constant spectral radius. arXiv preprint arXiv :1407.6568.

[100] Quint, J.F., 2002. Divergence exponentielle des sous-groupes discrets en rang supérieur. Commentarii Mathematici Helvetici, 77(3), pp.563-608.

[101] Quint, J.F., 2002. Cônes limites des sous-groupes discrets des groupes réductifs sur un corps local. Transformation groups, 7(3), pp.247-266.

[102] Quint, J.F., 2002. Mesures de Patterson-Sullivan en rang supérieur. Geometric & Functional Analysis GAFA, 12(4), pp.776-809.

[103] Quint, J.F., 2003. L'indicateur de croissance des groupes de Schottky. Ergodic theory and dynamical systems, 23(01), pp.249-272.

[104] Quint, J. F., 2005. Groupes de Schottky et comptage. In Annales de l'institut Fourier, Vol. 55, No. 2, pp.373-429.

[105] Quint, J. F., 2009. Lecture notes "Croissance des groupes d'isométries des espaces symétriques ", GDR Platon, Institut Henri Poincaré, Paris, Octobre 2009.

[106] Ruelle, D., 1967. A variational formulation of equilibrium statistical mechanics and the Gibbs phase rule. Communications in Mathematical Physics, 5(5), pp.324-329.

[107] Sapir, M., 2015. The rapid decay property and centroids in groups. Journal of Topology and Analysis, 7(03), pp.513-541.

[108] Serre, J. P., 1977. Arbres, amalgames, SL2. Astérisque 46.

[109] Stroock, D.W., 2012. An introduction to the theory of large deviations. Springer Science & Business Media.

[110] Tutubalin, V.N., 1977. A central limit theorem for products of random matrices and some of its applications. In Symposia Mathematica (Vol. 21, pp.101-116).

[111] Tits, J., 1972. Free subgroups in linear groups. Journal of Algebra, 20(2), pp.250-270.

[112] Tits, J., 1970. Sur le groupe des automorphismes d'un arbre. In Essays on topology and related topics (pp. 188-211). Springer Berlin Heidelberg.

[113] Tits, J., 1971. Représentations linéaires irréductibles d'un groupe réductif sur un corps quelconque. Journal für die reine und angewandte Mathematik, 247, pp.196-220.

[114] Rota, G.C. and Strang, W., 1960. A note on the joint spectral radius.

[115] Vershik, A. M., 2000. Dynamic theory of growth in groups : entropy, boundaries, examples. Russian Mathematical Surveys, 55(4), pp.667-733.

[116] Wilson, J. S., 2004. On exponential growth and uniformly exponential growth for groups. Inventiones mathematicae, 155(2), pp.287-303.

[117] Woess, W., 2000. Random walks on infinite graphs and groups (Vol. 138). Cambridge university press.