Topics in Random Walks on Lie Groups

Constantin Kogler
Merton College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

2025

Abstract

In Part I, we establish the first examples of finitely supported measures on semisimple Lie groups that satisfy a local limit theorem on the associated symmetric space. We reduce the problem at hand to a spectral gap question for a natural operator associated with the measure. When the given measure satisfies strong Diophantine properties and is supported close to the identity, the latter spectral gap problem is proven. Moreover, quantitative error rates for the local limit theorem are shown under additional assumptions, and C^{ℓ}-smoothness of the Furstenberg measure is discussed.

Part II, which is joint work with Samuel Kittle, is concerned with absolutely continuous self-similar measures. A condition for absolute continuity in arbitrary dimensions is shown. We thereby construct the first explicit absolutely continuous examples of genuinely inhomogeneous self-similar measures in dimensions one and two. Varjú's result for Bernoulli convolutions is strengthened, and in dimension ≥ 3 we improve the condition on absolute continuity by Lindenstrauss-Varjú.

Acknowledgements

I am deeply grateful to my supervisor, Emmanuel Breuillard, for his patient guidance throughout my studies. My time at Oxford and Cambridge was an enriching experience. I thank Emmanuel for the opportunities offered and his kind invitations to Worcester College. It was a privilege to be Emmanuel's student and to benefit from his vast knowledge of mathematics. I thank him for his generosity with ideas and for his unquestioned support in all my pursuits.

I warmly thank Péter Varjú, who acted as an informal co-supervisor to me. Not only was I supported by his ERC grant during the time I spent at Cambridge, but also the first two papers I wrote relied on crucial insights from him. I am deeply grateful for his selfless support as well as for his concise, constructive advice and mathematical comments. I also thank Péter for examining my thesis and for his highly detailed and insightful comments.

I am thankful to Ben Hambly for kindly agreeing to examine my thesis.

I am indebted to my outstanding collaborators and friends, Samuel Kittle and Wooyeon Kim, for sharing their knowledge and insights. The collaboration with Sam is the highlight of my academic journey to date and I admire his remarkable intellectual strength. I am also grateful to Wooyeon for sharing profound ideas and for our many illuminating conversations.

I thank Simon Machado, who has been a friend and mentor ever since we first met in front of King's College Cambridge in February 2020. I am grateful to Timothée Bénard and Amitay Kamber for their friendship and for being positive about my work.

I thank Manfred Einsiedler and Menny Akka for patiently guiding my studies at ETH Zürich and for introducing me to homogeneous dynamics. I am also grateful to Thomas Wurms, my mathematics teacher at the Freies Gymnasium Zürich, for his memorable classes and for suggesting to classify simple closed geodesics on platonic solids for my Maturaarbeit.

Finally, I warmly thank my family and close friends from Zürich, Cambridge and Oxford for the wonderful times spent together and all that I have learned from them.

Statement of Originality

Part I closely follows the article [Kog22], which has been accepted for publication in the Journal d'Analyse Mathématique.

Part II is a joint work with Samuel Kittle and is based on the articles [KK25c] and [KK25b], which are submitted for publication.

Contents

Chapter 1

Introduction

The study of random walks on Lie groups is a broad topic. This thesis addresses two subjects in the area: local limit theorems and the study of stationary measures, with an emphasis on self-similar measures. Both topics have a rich historical background and there are analogies and connections in the recent developments. Crucially, several results from the past 20 years as well as the primary theorems from Part I and Part II of this thesis require that the entries of the matrices in the support of our given distribution are algebraic. More precisely, these advancements rely on distinct problems in Diophantine approximation on Lie groups, which are well understood when the entries are algebraic but remain mostly unresolved in the transcendental case. We will direct our attention towards these facets in the introduction.

To set the stage, let G be a topological group acting on a space X, and let μ be a Borel probability measure on G. Consider independent random variables $X_1, X_2, \ldots$ each distributed according to μ, and fix a starting point $x_0 \in X$. Denote

$$Z_{n,x_0} = X_1 \cdots X_n . x_0. \tag{1.0.1}$$

The main goal in the subject is to study the behavior of Z_{n,x_0} as $n \to \infty$. We further note that Z_{n,x_0} is distributed as $\mu^{*n} * \delta_{x_0}$, where μ^{*n} denotes the n-fold convolution of μ.

In the introduction, we first present some notation, then discuss local limit theorems and state some of the results from Part I. Thereafter, we give an overview on self-similar measures and finally discuss the main result from Part II.

1.1 Notation

We review notation that will be used throughout this thesis. More specific notation that will be used in Part I will be reviewed in section 3.1 and for Part II in section 8.3.

The following asymptotic notation is used. We write $A \ll B$ or $A = O(B)$ to denote that $|A| \leq CB$ for a constant $C > 0$ and for sequences X_n and Y_n we write $X_n = o(Y_n)$ to symbolize $|\frac{X_n}{Y_n}| \to 0$ as $n \to \infty$. If the constant C or the speed of convergence depends on additional parameters, we add subscripts. Moreover, $A \asymp B$ denotes $A \ll B$ and $B \ll A$.

For a topological group G we denote by m_G a fixed choice of a left Haar measure, i.e. a measure satisfying $m_G(hA) = m_G(A)$ for all Borel measurable subsets $A \subset G$ and $h \in G$. Recall that the (left) Haar measure exists and is unique up to scaling when G is a locally compact Hausdorff group. When G is compact, we denote by m_G the Haar probability measure. For $G = \mathbb{R}$ we write $m_{\mathbb{R}}$ for the standard Lebesgue measure.

When G acts on a space X, μ is a Borel measure on G and ν one on X, we define the convolution $\mu * \nu$ as the measure uniquely determined by satisfying

$$(\mu * \nu)(f) = \int \int f(g.x)\, d\mu(g) d\nu(x)$$

for all continuous compactly supported functions $f : X \to \mathbb{R}$.

1.2 Local limit theorems

To review some classical results, we first consider the case where $G = X = \mathbb{R}$, and let $X_1, X_2, \ldots$ be independent random variables distributed according to μ. Assume that the mean $\mathbb{E}[X_i]$ is zero and that the variance $\sigma^2 = \mathbb{E}[X_i^2] < \infty$ is finite. The central limit theorem states that

$$\frac{Z_{n,0}}{\sqrt{n}} \to \mathcal{N}(0, \sigma^2),$$

where the convergence is in distribution and $\mathcal{N}(0, \sigma^2)$ is the distribution of a centered Gaussian with variance σ^2. Thus, the central limit theorem provides information about $Z_{n,0}$ at scale $\sqrt{n}$. On the other hand, it is of interest to describe $Z_{n,0}$ at scale 1, which is addressed by local limit theorems. Indeed, as long as μ is non-lattice, meaning μ is not supported on $\beta\mathbb{Z}$ for some $\beta \in \mathbb{R}$, we have, as $n \to \infty$,

$$\sqrt{n}\mu^{*n} \to \frac{m_{\mathbb{R}}}{\sqrt{2\pi\sigma^2}}, \tag{1.2.1}$$

where $m_{\mathbb{R}}$ denotes the Lebesgue measure on $\mathbb{R}$ and the measures converge vaguely, that is (1.2.1) holds for functions $f \in C_c(\mathbb{R})$.

Our interest in this section lies in similar questions for arbitrary transitive actions of a topological group G on a space X. While a central limit theorem in this general context is difficult to formulate, it is expected that a local limit theorem usually holds, provided the measure in question is non-degenerate and symmetric. Specifically, we give the following definition.

Definition 1.2.1. *A probability measure μ on a locally compact group G is called **aperiodic** if the support of μ is not contained in a coset gH for some $g \in G$ and a proper closed subgroup $H < G$.*

*The probability measure μ is called **symmetric** if for all continuous compactly supported functions $f : G \to \mathbb{R}$ we have that*

$$\int f(g)\, d\mu(g) = \int f(g^{-1})\, d\mu(g).$$

We offer the following general question on the behaviour of random walks for transitive actions.

Question 1.2.2. *(Does a Local Limit Theorem hold?) Let G be a locally compact, second countable and Hausdorff group acting continuously and transitively on a space X. Let μ be an aperiodic, symmetric and compactly supported Borel probability measure on G.*

Does there then exists a sequence of real numbers $(a_n)_{n\geq 1}$ with $a_n > 0$ such that for every $x_0 \in X$ there is a limiting non-zero Borel measure m_{μ,x_0} on X of full support such that

$$a_n \mu^{*n} * \delta_{x_0} \to m_{\mu,x_0} \tag{1.2.2}$$

vaguely? More precisely, we ask whether for all continuous compactly supported functions $f : X \to \mathbb{R}$ it holds that

$$\lim_{n\to\infty} a_n \int f(g.x_0)\, d\mu^{*n}(g) = \int f(x)\, dm_{\mu,x_0}(x).$$

If in addition X is endowed with a Borel measure m_X and the G action on X is measure preserving, is then the limiting measure m_{μ,x_0} absolutely continuous with respect to m_X for every $x_0 \in X$?

We have made the above strong assumptions on the measure μ in order to formulate a question that could be valid for all groups. However, when working with concrete groups, some of the assumptions can be weakened. For example, when $G = \mathbb{R}$, it suffices to assume that μ is centered, or when G is compact, we can drop the assumption that μ is symmetric, as in Theorem 1.2.3.

The author does not know of a group or group action where (1.2.2) is false, yet as the setting is vast, we have not dared to formulate the question as a conjecture. As we discuss later, it is known that Question 1.2.2 can be answered in the affirmative when $G = X$ is a discrete ameanable group or a free group. In the remainder of this section we discuss the case when G is a Lie group and Part I of this thesis addresses G being a semisimple Lie group such as $\mathrm{SL}_n(\mathbb{R})$.

When (1.2.2) is known, it is of further interest to describe the limiting measure m_{μ,x_0} or to prove error rates for the convergence. When G is an ameanable group and $X = G$, then in all known cases the limit measure m_{μ,x_0} is a multiple of the Haar measure. As discussed below, the latter does not hold for the free group (1.2.7) or for semisimple Lie groups (Theorem 1.3.1).

Compact Groups

When G is compact, the Haar probability measure m_G is the natural candidate for the limit measure in (1.2.2). Indeed, for compact groups the local limit theorem was proven by Ito-Kawada in 1940. This is actually the first paper of Kyoshi Ito, the pioneer of stochastic integration.

Theorem 1.2.3. *(*[IK40]*) Let G be a compact topological group and let μ be an aperiodic probability measure on G. Then μ^{*n} converges to m_G vaguely as $n \to \infty$.*

The latter theorem affirms Question 1.2.2 for compact topological groups. In fact, every continuous function on X can be lifted to a continuous function on G and (1.2.2) follows from Theorem 1.2.3.

Having established that μ^{*n} converges to m_G, one may wonder about the speed of convergence. When $G = \mathbb{T} = \mathbb{R}/\mathbb{Z}$ and $\mu = \frac{1}{2}(\delta_\alpha + \delta_{-\alpha})$ for some irrational α, then the speed of equidistribution of μ^{*n} is equivalent to Diophantine properties of α. Indeed, this relies on the Fourier inversion formula and the observation that for $n \in \mathbb{Z}$,

$$\widehat{\mu}(n) = \int e^{2\pi i n x}\, d\mu(x) = \cos(2\pi n\alpha).$$

Similar phenomena occur whenever G is a compact abelian Lie group.

We turn our attention for the remainder of this section to compact simple Lie groups such as, for example, the orthogonal group $SO(d)$ for $d \geq 3$ or the special unitary group $SU(r)$ for $r \geq 2$. When μ is supported on finitely many elements generating a dense subgroup of G, then it follows from the Tits alternative [Tit72] that the cardinality of the support of μ^{*n} grows exponentially. It is therefore natural to conjecture that μ^{*n} converges to m_G with exponential speed, in other words that for every $f \in C^\infty(G)$,

$$\mu^{*n}(f) = m_G(f) + O_{\mu,f}(e^{-\theta n}). \tag{1.2.3}$$

for a constant $\theta = \theta(\mu)$ depending on μ.

The following two important results were proved by Benoist-de Saxcé, generalising pioneering work by Bourgain-Gamburd [BG08], [BG12]. We first define the weak Diophantine property of probability measures.

Definition 1.2.4. *Let G be a compact group. Then μ is called **weakly Diophantine** if there are constants $c_1, c_2 > 0$ such that for sufficiently large n,*

$$\sup_{H<G} \mu^{*n}(B_{e^{-c_1 n}}(H)) \leq e^{-c_2 n},$$

where the supremum is taken over all proper closed subgroups H of G and $B_{e^{-c_1 n}}(H) = \{g \in G \,:\, d(g,H) \leq e^{-c_1 n}\}$.

Theorem 1.2.5. *(*[BdS16, Theorem 1.1]*) Let G be a connected compact simple Lie group and let μ be a symmetric Borel probability measure on G. Then μ is weakly Diophantine if and only if it has a spectral gap on $L^2(G)$, that is there is $\gamma > 0$ such that*

$$||\mu * f||_2 \leq (1-\gamma)||f||_2 \tag{1.2.4}$$

for all $f \in L^2(G)$ with $m_G(f) = 0$ and where $(\mu * f)(x) = \int f(gx)\, d\mu(g)$ for $x \in G$.

Theorem 1.2.6. *(*[BdS16, Theorem 1.2]*) Let $G \subset \mathrm{GL}_d(\mathbb{R})$ be a connected compact simple Lie group. Let μ be a probability measure supported on finitely many matrices in $\mathrm{GL}_d(\overline{Q})$ and generating a dense subgroup. Then μ is weakly Diophantine, and therefore has a spectral gap on $L^2(G)$.*

When μ has a spectral gap on $L^2(G)$, then (1.2.3) holds as it follows for example from [KK24, Corollary 3.11] that there is a constant $\theta = \theta(\mu)$ such that for any bounded Lipschitz function $f : G \to \mathbb{R}$,

$$\mu^{*n}(f) = m_G(f) + O_\mu(\max(||f||_\infty, \mathrm{Lip}(f))e^{-\theta n}),$$

where $||f||_\infty = \sup_{g\in G} |f(g)|$ and $\mathrm{Lip}(f)$ is the Lipschitz constant of f.

Analogously to (1.2.3), it is believed that every probability measure on a connected compact simple Lie group whose support generates a dense subgroup has a spectral gap on $L^2(G)$. It is a major open problem in the field to remove the assumption of Theorem 1.2.6 that the entries of the matrices in the support of μ are algebraic.

Throughout this thesis we will encounter numerous results that are understood in the case when the entries of the matrices in the support are algebraic, yet open in the general case.

$\mathrm{Isom}(\mathbb{R}^d)$ **action on** $\mathbb{R}^d$

Denote by $\mathrm{Isom}(\mathbb{R}^d)$ the group of isometries of $\mathbb{R}^d$, i.e. maps of the form $x \mapsto Ux + b$ for $U \in O(d)$ and $b \in \mathbb{R}^d$ with $x \in \mathbb{R}^d$. Instead of studying the sum of random variables on $\mathbb{R}^d$, it is an interesting question to study random walks for the $\mathrm{Isom}(\mathbb{R}^d)$ action on $\mathbb{R}^d$.

The problem of establishing a central limit theorem or a local limit theorem in this setting can be traced back to the paper by Arnold-Krylov [AK63] from 1963. After the work of several people (we refer to [Var15] for a historical account), Varjú finally proved a local limit theorem in 2015.

Theorem 1.2.7. *(*[Var15]*, Local Limit Theorem with error rates) Let μ be an aperiodic, compactly supported probability measure on* $\mathrm{Isom}(\mathbb{R}^d)$*. Then there exists $y_0 \in \mathbb{R}^d$, a quadratic form Δ and constants $C_\Delta > 0$ and $c > 0$, all depending only on μ, such that the following holds.*

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a smooth function with compact support. Then for Z_{n,x_0} as defined in (1.0.1)*, it holds that*

$$n^{d/2}\mathbb{E}[f(Z_{n,x_0})] = C_\Delta \int f(x) e^{-\Delta(x-y_0,x-y_0)/n}\, dm_{\mathbb{R}^d}(x) \\ + O_\mu(n^{-1/2} + |x_0|^2 n^{-1})||f||_1 + O_\mu(e^{-cn^{1/4}})||f||_{W^{2,(d+1)/2}},$$

where $||\cdot||_1$ is the L^1-norm and $||\cdot||_{W^{2,(d+1)/2}}$ is the L^2 Sobolev norm defined by $||f||^2_{W^{2,(d+1)/2}} = \int |\hat{f}(\xi)|^2(1+|\xi|)^{d+1}\, dm_{\mathbb{R}^d}(\xi)$.

In [Var15], Theorem 1.2.7 is proved under weaker assumptions on μ, yet we have stated the result in the above form for simplicity. We denote by $U(\mu)$ the push forward of μ under the map that sends an isomtery to the orthogonal part, i.e. $g \in \mathrm{Isom}(\mathbb{R}^d) \mapsto U(g) \in O(d)$, where $g(x) = U(g)x + b(g)$ for all $x \in \mathbb{R}^d$. The proof of Theorem 1.2.7 relies on establishing a weakening of the $L^2(O(d))$ spectral gap discussed in the last subsection. In the case when $U(\mu)$ has a spectral gap on $L^2(SO(d))$, the following strengthening of Theorem 1.2.7 was established by Lindenstrauss-Varjú in 2016.

Theorem 1.2.8. *(*[LV16]*, Local Limit Theorem with strong error rates) Let $d \geq 3$ and let μ be a compactly supported probability measure on* $\mathrm{Isom}(\mathbb{R}^d)$*. Assume that $U(\mu)$ is supported on $SO(d)$ and has a spectral gap on $L^2(SO(d))$. Moreover, assume* *that there is no point $x \in \mathbb{R}^d$ such that $X_1(x) = x$ almost surely. Then there exists $y_0 \in \mathbb{R}^d$, a quadratic form Δ and constants $C_\Delta > 0$ and $c > 0$, all depending only on μ, such that the following holds.*

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a smooth function with compact support. Then

$$n^{d/2}\mathbb{E}[f(Z_{n,x_0})] = C_\Delta \int f(x) e^{-\Delta(x-y_0,x-y_0)/n} \, dm_{\mathbb{R}^d}(x)$$
$$+ O_\mu(n^{-1/2} + |x_0|^2 n^{-1})||f||_1 + O_\mu(e^{-cn})||f||_{W^{2,(d+1)/2}}.$$

The difference between Theorem 1.2.8 and Theorem 1.2.7 is that the decay of the term in front of the Sobolev norm $||f||_{W^{2,(d+1)/2}}$ is exponential and not only $O_\mu(e^{-cn^{1/4}})$. The former is optimal because the number of points in the support of μ grows exponentially.

The proof of Theorem 1.2.7 and Theorem 1.2.8 relies on spectral estimates of natural operators arising from the theory of unitary representations on $\mathrm{Isom}(\mathbb{R}^d)$. Indeed, as above, for $g \in \mathrm{Isom}(\mathbb{R}^d)$ we denote by $U(g)$ the rotation part and by $b(g)$ the translation part. Then for $r \in \mathbb{R}$ consider the unitary representation $\rho_r : G \to \mathcal{U}(L^2(\mathbb{S}^{d-1}))$ defined by

$$(\rho_r(g)\varphi)(\xi) = e^{-2\pi i r\langle \xi, b(g)\rangle}\varphi(U(g)^{-1}\xi)$$

for $\varphi \in L^2(\mathbb{S}^{d-1})$, $g \in \mathrm{Isom}(\mathbb{R}^d)$ and $\xi \in \mathbb{S}^{d-1}$.

For a given probability measure μ on $\mathrm{Isom}(\mathbb{R}^d)$ one considers the operator

$$S_r = \rho_r(\mu) = \int \rho_r(g) \, d\mu(g). \tag{1.2.5}$$

We observe that for $n \geq 1$,

$$S_r^n = \rho_r(\mu^{*n}) = \int \rho_r(g) \, d\mu^{*n}(g)$$

Then by the Fourier inversion formula and Fubini's theorem the following holds, which we state for simplicity in the case when $x_0 = 0$ (note that $g.0 = b(g)$),

$$\mathbb{E}[f(Z_{n,0})] = \int f(b(g)) \, d\mu^{*n}(g) = \int\int \widehat{f}(\xi) e^{-2\pi i\langle \xi, b(g)\rangle} \, dm_{\mathbb{R}^d}(\xi) d\mu^{*n}(g)$$
$$= \int \widehat{f}(\xi)(S_{|\xi|}^n 1)(\xi/|\xi|) \, dm_{\mathbb{R}^d}(\xi). \tag{1.2.6}$$

The proof of Theorem 1.2.7 and Theorem 1.2.8 proceeds using distinct bounds for $S_{|\xi|}^n 1$ in the ranges when the frequency ξ is low and when it is high. For the low frequency range one applies a Taylor expansion to deduce satisfactory bounds. For Theorem 1.2.7 the high frequency range is dealt with by subtle estimates relying on the weakened spectral gap for $U(\mu)$. The following important theorem is proven in [LV16] to deal with the high frequency range in Theorem 1.2.8.

Theorem 1.2.9. *(*[LV16, Corollary 8.1]*) Let $d \geq 3$ and suppose that μ is a compactly supported probability measure on* $\mathrm{Isom}(\mathbb{R}^d)$*. Assume that $U(\mu)$ is supported on $SO(d)$ and has a spectral gap on $L^2(SO(d))$. Then there is a constant $c > 0$ such that*

$$\sup_{r \geq 1} ||S_r|| \leq 1 - c,$$

where $||\cdot||$ is the operator norm.

Remarkably, Theorem 1.2.9 has important applications in establishing absolute continuity of self-similar measures in dimension $d \geq 3$, the topic addressed in Part II of this thesis. The paper [LV16] is indeed a central source of inspiration for all the results of this thesis.

Further results

We briefly mention a few further papers on local limit theorems in various contexts.

Nilpotent Lie groups were studied by Breuillard [Bre05b], Hough [Hou19], Diaconis-Hough [DH21] and Bénard-Breuillard [BB23].

The case of discrete groups is also an active area of research. When G is a discrete amenable group, it follows for a symmetric aperiodic probability measure μ on G by [Ave73] (c.f. furthermore [Ger80]) that

$$\lim_{n\to\infty} \frac{\mu^{*n}(g)}{\mu^{*n}(e)} = 1$$

for all $g \in G$, which affirms Question 1.2.2 when $G = X$ is amenable. On the other hand, it was proven by Gouëzel [Gou14], generalising work by Lalley for free groups [Lal93], that if G is a discrete hyperbolic group and μ is an aperiodic probability measure on G, then for every $x, y \in G$ there exists a constant $C(x, y)$ such that

$$\lim_{n\to\infty} \frac{n^{3/2}}{\sigma^n} \mathbb{P}[Z_{n,x} = y] = C(x, y) \tag{1.2.7}$$

for $\sigma \in (0, 1)$ a constant depending on μ.

Another topic of interest is to study volume preserving actions of discrete subgroups on finite-volume homogeneous spaces. We mention Bourgain-Furman-Lindenstrauss-Mozes [BFLM11] establishing quantitative results for $\mathrm{SL}_n(\mathbb{Z})$ acting on $\mathbb{T}^d$, Benoist-Quint [BQ11] classifying stationary measures for the action on finite-volume homogeneous spaces of semi-simple Lie groups G by Zariski dense subgroups of G, and the recent work of Bénard-He [BH24] proving a quantitative local limit theorem in the latter setting for arithmetic quotients of for example $G = \mathrm{SL}_2(\mathbb{R})$ and measures supported on algebraic elements.

1.3 Results of Part I

In contrast to the above, the understanding of local limit theorems for non-compact semisimple Lie groups such as $\mathrm{SL}_n(\mathbb{R})$ is much less complete. The only known case where a local limit theorem is proven assumes that μ is spread out, meaning a convolution power μ^{*n} for some $n \geq 1$ is not singular with respect to the Haar measure. Indeed, the following theorem was proven by Bougerol in 1981.

Theorem 1.3.1. *(*[Bou81]*) Let G be a non-compact connected semisimple Lie group with finite center and let μ be a compactly supported spread out probability measure on G whose support generates a dense subgroup of G. Then there exists a constant $\sigma = \sigma(\mu) \in (0, 1)$ and a continuous function ψ_0 on G depending on μ such that*

$$\lim_{n\to\infty} \frac{n^{\ell/2}}{\sigma^n} \int f(g)\, d\mu^{*n}(g) = \int f(g)\psi_0(g)\, dm_G(g) \tag{1.3.1}$$

for all $f \in C_c^\infty(G)$ and where $\ell = \ell(G)$ is an integer depending only on G. The function ψ_0 satisfies $\mu * \psi_0 = \psi_0 * \mu = \sigma\psi_0$.

The assumption that μ is compactly supported is not necessary. Indeed, one only requires that μ has a finite second moment as defined in (2.0.2). The reader may observe the analogy between (1.3.1) and (1.2.7).

Ever since Bougerol's theorem, it has been an open problem to extend (1.3.1) to finitely supported measures whose support generates a dense subgroup. This question motivates the first part of this thesis. Although we cannot solve this problem, we give the first examples of finitely supported probability measures that satisfy a local limit theorem on the associated symmetric space.

Theorem 1.3.2. *(Follows from Theorem 2.0.1 and Theorem 2.0.7) Let G be a non-compact connected semisimple Lie group with finite center, let $K \subset G$ be a maximal compact subgroup and denote by $X = G/K$ the associated symmetric space. Then there exists finitely supported probability measures μ on G such that* (1.3.1) *holds on X.*

Indeed, there exists a constant $\sigma = \sigma(\mu) \in (0, 1)$ and a continuous function ψ_0 on G depending on μ such that for all $x_0 \in X$,

$$\lim_{n\to\infty} \frac{n^{\ell/2}}{\sigma^n} \int f(g.x_0)\, d\mu^{*n}(g) = \int f(g.x_0)\psi_0(g)\, dm_G(g) \tag{1.3.2}$$

for all $f \in C_c^\infty(X)$ and where $\ell = \ell(G)$ is an integer depending only on G.

The proof of Theorem 1.3.1 and Theorem 1.3.2 is based on studying an analogous class of operators to S_r as defined in (1.2.5) for $\mathrm{Isom}(\mathbb{R}^d)$. For symmetry, we also denote this class of operators as S_r. In Theorem 2.0.1, we prove (1.3.2) under the assumption that S_0 is quasicompact, that is, the essential spectral radius (see (3.1.1)) is strictly less than the spectral radius. We also show quantitative bounds for (1.3.2), inspired by Theorem 1.2.7 and Theorem 1.2.8, in Theorem 2.0.2 and Theorem 2.0.3.

To deduce Theorem 1.3.2 from Theorem 2.0.1, we need to show that there exist finitely supported measures with S_0 quasicompact. When $G = \mathrm{SL}_2(\mathbb{R})$, such examples were constructed by Bourgain [Bou12]. In fact, as we discuss next section, Bourgain showed that there exist finitely supported measures with absolutely continuous Furstenberg measure. The method of [Bou12] was generalised in [BISG17] to arbitrary connected simple Lie groups and we use the results there to conclude that S_0 is quasicompact for numerous examples as well as to generalise Bourgain's result for Furstenberg measures. All this is stated in Theorem 2.0.7 and Theorem 2.0.8. Connecting to the weak Diophantine property from Definition 1.2.4, the measures we work with have strong Diophantine properties while being supported close to the identity. This leads to the definition of (c_1, c_2, ε)-Diophantine measures as given in Definition 2.0.5.

1.4 Self-similar measures

Returning to the setting of (1.0.1), when most of the elements of the measure μ exhibit a contracting behavior, it can be shown that Z_{n,x_0} converges to a limiting distribution. More concretely, let us assume that X is a complete metric space and the measure μ is contracting, that is, for every $g \in \mathrm{supp}(\mu)$ it holds that $\mathrm{Lip}(g) < 1$, for $\mathrm{Lip}(g)$ the Lipschitz constant of g. According to Hutchinson's theorem [Hut81], there exists a unique probability measure ν on X such that for any starting point $x_0 \in \mathbb{R}^d$, as $n \to \infty$,

$$\mu^{*n} * \delta_{x_0} \to \nu. \tag{1.4.1}$$

This measure ν is μ-stationary, satisfying $\mu * \nu = \nu$, and the convergence stated above is exponentially fast. The study of properties of the limiting distributions ν in various settings is an active area of research. We also note that (1.4.1) is an instance of (1.2.2).

From now on, consider the case where $X = \mathbb{R}^d$ and that μ is supported on similarities. A similarity on $\mathbb{R}^d$ is a map $g : \mathbb{R}^d \to \mathbb{R}^d$ that, for some $\rho > 0$, satisfies $d(g(x), g(y)) = \rho \cdot d(x, y)$ for all $x, y \in \mathbb{R}^d$. We denote the group of similarities as

$\mathrm{Sim}(\mathbb{R}^d)$. For each similarity $g \in \mathrm{Sim}(\mathbb{R}^d)$, there exist a scalar $\rho(g) > 0$, an orthogonal matrix $U(g) \in O(d)$, and a vector $b(g) \in \mathbb{R}^d$ such that

$$g(x) = \rho(g)U(g)x + b(g) \tag{1.4.2}$$

for all $x \in \mathbb{R}^d$.

For a contracting measure μ on $\mathrm{Sim}(\mathbb{R}^d)$, the limiting measure ν from (1.4.1) is referred to as the self-similar measure of μ. Self-similar measures are well-known to be exact-dimensional ([FH09]), that is, there exists a constant $\alpha \in [0, d]$ such that

$$\nu(B_r(x)) = r^{\alpha + o_{\mu,x}(1)} \tag{1.4.3}$$

as $r \to 0$ for ν-almost every $x \in \mathbb{R}^d$, where $B_r(x)$ denotes the r-ball around x. The quantity α is known as the dimension of ν. Key questions in the study of self-similar measures are:

1. What is the dimension of ν?
2. Is ν absolutely continuous on $\mathbb{R}^d$? That is, does there exist a function $f_\nu \in L^1(\mathbb{R}^d)$ such that $\nu = f_\nu \cdot \mathrm{vol}_{\mathbb{R}^d}$?

Hochman's Theorem

Recent decades have seen significant advances on these problems. We first recall Hochman's pivotal contributions ([Hoc14], [Hoc17]). Let μ be a finitely supported probability measure. The random walk entropy of μ is defined as

$$h_\mu = \lim_{n\to\infty} \frac{1}{n} H(\mu^{*n}) = \inf_{n\geq 1} \frac{1}{n} H(\mu^{*n}),$$

where $H(\cdot)$ is the Shannon entropy. The Lyapunov exponent is given by

$$\chi_\mu = \mathbb{E}_{g\sim\mu}[\log \rho(g)]. \tag{1.4.4}$$

Furthermore, we say that μ is irreducible if the subgroup generated by $\{U(g) : g \in \mathrm{supp}(\mu)\}$ acts irreducibly on $\mathbb{R}^d$, meaning it has no invariant subspaces other than the trivial ones $\{0\}$ and $\mathbb{R}^d$. When all of the maps $g \in \mathrm{supp}(\mu)$ have the same fixed point $x \in \mathbb{R}^d$, then the Dirac measure δ_x is the self-similar measure of μ. When the latter is not the case, we say that μ does not have a common fixed point.

It is well known (cf. for example [FH09]) that

$$\dim \nu \leq \min\left\{d, \frac{h_\mu}{|\chi_\mu|}\right\}. \tag{1.4.5}$$

Moreover the following conjecture is expected to hold.

Conjecture 1.4.1. *(Generalised Exact Overlaps Conjecture) Let μ be a finitely supported, irreducible and contracting probability measure on* $\mathrm{Sim}(\mathbb{R}^d)$ *without a common fixed point. Then*

$$\dim \nu = \min \left\{ d, \frac{h_\mu}{|\chi_\mu|} \right\}.$$

Much of the literature on self-similar measures assumes that the support of μ generates a free semigroup, which is referred to as the support of μ having no exact overlaps. In this case, if $\mu = \sum_{i=1}^{k} p_i \delta_{g_i}$, then the random walk entropy can be computed as

$$h_\mu = H(p_1, \ldots, p_k) = -\sum_{i=1}^{k} p_i \log p_i.$$

Hochman proved Conjecture 1.4.1 under a mild separation assumption. Indeed, denote by

$$\Delta_n = \min\{d(g,h) \, : \, g, h \in \mathrm{supp}(\mu^{*n}) \text{ with } g \neq h\}, \tag{1.4.6}$$

where $d(\cdot,\cdot)$ is the metric on $\mathrm{Sim}(\mathbb{R}^d)$ given in (7.0.3).

Theorem 1.4.2. *(*[Hoc14], [Hoc17]*) Let μ be a finitely supported, irreducible and contracting probability measure on* $\mathrm{Sim}(\mathbb{R}^d)$ *without a common fixed point. Suppose that there is $c > 0$ and infinitely many $n \geq 1$ such that $\Delta_n \geq e^{-cn}$, then*

$$\dim \nu = \min \left\{ d, \frac{h_\mu}{|\chi_\mu|} \right\}. \tag{1.4.7}$$

We also mention the following proposition, which implies, as we can embed $\mathrm{Sim}(\mathbb{R}^d)$ in $\mathrm{GL}_{d+1}(\mathbb{R})$, that (1.4.7) holds for all μ supported on finitely many similarities with algebraic coefficients.

Proposition 1.4.3. *(follows as Proposition 15.2.3) Let μ be a probability measure on* $\mathrm{GL}_d(\mathbb{R})$ *for some $d \geq 1$ supported on finitely many matrices with algebraic coefficients. Then there exists $c > 0$ such that $\Delta_n \geq e^{-cn}$ for all $n \geq 1$.*

We invite the reader to compare Theorem 1.4.2 with Theorem 1.2.5. In both cases, noticeable results concerning random walks on Lie groups are reduced to Diophantine problems. There is an important difference: The weak Diophantine property of Theorem 1.2.5 requires information on all scales whereas for the condition from Theorem 1.4.2 only infinitely many scales are necessary.

In the following discussion, for simplicity we consider $G = SO(3)$. In the following argument, that was communicated to me by Emmanuel Breuillard, it is shown that the conclusion of Proposition 1.4.3 does not hold for every measure supported on two elements generating a free group.

Denote by

$$T_n = \{(g,h) \in SO(3) \,:\, g^n = h^n = I\}$$

the n-torsion elements. It is simple to verify that $\bigcup_{n\geq k} T_n$ is dense in $SO(3)\times SO(3)$ for every $k \geq 1$. When $n,k \geq 1$ there exists a small number $r_{n,k} > 0$ such that if $d((x,y),(a,b)) < r_{n,k}$ for some tuple $(a,b) \in T_n$ then it holds that

$$d(x^n, I) < e^{-kn} \quad \text{and} \quad d(y^n, I) < e^{-kn}.$$

We then write

$$E_k = \bigcup_{n\geq k} \bigcup_{(a,b)\in T_n} B_{r_{n,k}}((a,b))\backslash V_n,$$

where $B_{r_{n,k}}((a,b))$ denotes the open $r_{n,k}$ ball of (a,b) in $G\times G$ and V_n is the set of all $(g,h) \in SO(3)\times SO(3)$ such that $w(g,h) = I$ for some word w of length $\leq n$. Since V_n is a proper algebraic subset of $SO(3)\times SO(3)$, it follows that E_k is open and dense. We finally denote

$$E = \bigcap_{k\geq 1} E_k,$$

which by the Baire category theorem is a dense subset of $SO(3)\times SO(3)$. Also, every element $(x,y) \in E$ generates a free group and for every k there is some n such that

$$0 < d(x^n, y^n) < 2e^{-kn}.$$

Therefore for a probability measure supported on x and y the conclusion of Proposition 1.4.3 that $\Delta_n \geq e^{-cn}$ for some $c > 0$ and all $n \geq 1$ does not hold.

In contrast to the above discussion, Hochman's theorem (Theorem 1.4.2) only requires separation on infinitely many scales. To the author's knowledge, there is no counterexample to the latter condition and one could therefore conjecture that it always holds.

Conjecture 1.4.4. *For every finitely supported probability measure on* $\mathrm{GL}_d(\mathbb{R})$*, for any* $d \geq 1$*, there exists* $c > 0$ *such that* $\Delta_n \geq e^{-cn}$ *for infinitely many* $n \geq 1$*.*

Bernoulli Convolutions

The simplest interesting example of self-similar measures are Bernoulli convolutions. For $\lambda \in (0,1)$ we denote by

$$\nu_\lambda \qquad \text{the distribution of} \qquad \sum_{i=0}^{\infty} X_i \lambda^i$$

where X_i are independent random variables satisfying

$$\mathbb{P}[X_i = -1] = \mathbb{P}[X_i = 1] = \frac{1}{2}.$$

The measure ν_λ is referred to as the Bernoulli convolution of parameter λ. It is the self-similar measure associated to the unbiased measure supported on the two maps $x \mapsto \lambda x + 1$ and $x \mapsto \lambda x - 1$. We make the following basic observations:

1. When $\lambda \in (0, 1/2)$, then ν_λ is supported on a Cantor set and ν_λ is not absolutely continuous.

2. For $\lambda = 1/2$, ν_λ is the Lebesgue measure on $[-2, 2]$.

3. For $\lambda \in (1/2, 1)$, $\operatorname{supp}(\nu_\lambda) = [\frac{-1}{1-\lambda}, \frac{1}{1-\lambda}]$ and ν_λ may be absolutely continuous or not.

The study of Bernoulli convolution is at least 90 years old and we expose the following selection of known results:

1. (Jessen-Wintner 1935 [JW35, Theorem 11]) ν_λ is of pure type, i.e. is either singular or absolutely continuous to the Lebesgue measure.

2. (Erdős 1939 [Erd39]) When λ^{-1} is a Pisot number, then ν_λ is not absolutely continuous. A Pisot number is an algebraic number with all of its Galois conjugates having modulus < 1, as for example the golden ratio $\lambda = \frac{\sqrt{5}+1}{2}$. Erdös result exploits that powers of Pisot numbers are well approximated by integers.

3. (Erdős 1940 [Erd40]) There exists some $c > 1/2$ such that ν_λ is absolutely continuous for almost all $\lambda \in [c, 1]$.

4. (Garsia 1962 [Gar62]) When λ^{-1} is a Pisot number, then $\dim \nu_\lambda < 1$.

To continue our exposition, we recall the definition of the Mahler measure M_λ.

Definition 1.4.5. *The Mahler measure M_λ of an algebraic number λ is defined as*

$$M_\lambda = |a| \prod_{|z_j|>1} |z_j|$$

where $a(x - z_1) \cdots (x - z_\ell)$ is the minimal polynomial of λ over $\mathbb{Z}$.

4. (Garsia 1962 [Gar62]) ν_λ is absolutely continuous if λ^{-1} is an algebraic integer and $M_\lambda = 2$. Examples include $\lambda = 2^{-1/k}$ and the real roots of $x^{p+n} - x^n - 2$ for any $\max\{p, n\} \geq 2$.

5. (Solomyak 1995 [Sol95]) ν_λ is absolutely continuous for almost all $\lambda \in (1/2, 1)$.

6. (Hochman 2014 [Hoc14]) It follows from Theorem 1.4.2 that $\dim \nu_\lambda = 1$ if λ is algebraic, but not the zero of a polynomial with coefficients in $-1, 0, 1$. Also, it can be deduced from Theorem 1.4.2 that the set of $\lambda \in (1/2, 1)$ such that $\dim \nu_\lambda < 1$ has Hausdorff dimension zero.

7. (Feng-Zhou 2022 [FF22]) Demonstrated that $\dim \nu_\lambda \geq 0.98040856$ for all $\lambda \in (1/2, 1)$.

We give special emphasis to the following two recent results by Varjú, both published in 2019.

Theorem 1.4.6. *(*[Var19a]*) For every $\varepsilon > 0$ there is a constant $c_\varepsilon > 0$ such that ν_λ is absolutely continuous if $\lambda \in (1/2, 1)$ is algebraic and*

$$\lambda > 1 - c_\varepsilon \min\{\log M_\lambda, (\log M_\lambda)^{-1-\varepsilon}\}.$$

Theorem 1.4.6 established the first explicit examples of absolutely continuous Bernoulli convolutions since the work of Garsia [Gar62] in 1962. For instance, Theorem 1.4.6 implies that ν_λ is absolutely continuous for $\lambda = 1 - \frac{p}{q}$ with coprime positive integers p, q and

$$\frac{p(\log q)^{1+\varepsilon}}{q} \leq c_\varepsilon. \tag{1.4.8}$$

Theorem 1.4.7. *(*[Var19b]*) For every transcendental $\lambda \in (1/2, 1)$,*

$$\dim \nu_\lambda = 1.$$

Together with Hochman's Theorem 1.4.2, Theorem 1.4.7 solves Conjecture 1.4.1 for Bernoulli convolutions. Varjú's proof strategy does not establish Conjecture 1.4.4 in the case of Bernoulli convolutions. Rather, the proof, roughly speaking, relies on weakening the separation condition from Theorem 1.4.2 for Bernoulli convolutions, as was achieved by [BV19], as well as on exploiting subtle number-theoretic properties. It is a remarkable achievement as it is the first time in the area where something can be said about genuinely transcendental measures.

Similarly to the work of Breuillard-Varjú [BV19], in my joint paper with Samuel Kittle [KK25a], we show that a weaker separation condition than exponential separation on all scales is sufficient for arbitrary self-similar measures to establish the conclusion of Theorem 1.4.2. The latter leads to an analogue of Theorem 1.4.7 for complex Bernoulli convolutions.

Spectral Gap and Dimension $d \geq 3$

As mentioned above, the work of Lindenstrauss-Varjú [LV16] on the local limit theorem for $\mathrm{Isom}(\mathbb{R}^d)$ has applications to self-similar measures. In fact, the following result is a consequence of Theorem 1.2.9.

Theorem 1.4.8. *(*[LV16, Theorem 1.3]*) Let $d \geq 3$ and μ_U be a probability measure on $SO(d)$ with a spectral gap on $L^2(SO(d))$. Then for every $\ell \geq 1$ and $\varepsilon > 0$ there exists a constant $\rho_\ell = \rho_\ell(d, \varepsilon, \mu_U) \in (0,1)$ depending on d, ℓ, ε and the spectral gap of μ_U such that the following holds.*

Let $\mu = \sum_{i=1}^k p_i \delta_{g_i}$ be a finitely supported probability measure on $\mathrm{Sim}(\mathbb{R}^d)$ without a common fixed point. Then assume that

$$U(\mu) = \mu_U$$

as well as for all $1 \leq i \leq k$,

$$\rho(g_i) \in (\rho_\ell, 1) \qquad \text{and} \qquad p_i \geq \varepsilon.$$

Then the self-similar measure of μ is absolutely continuous with a density in C^ℓ.

Before Theorem 1.4.8 there was essentially nothing known on explicit examples of absolutely continuous self-similar measures beyond the case proved by Garsia [Gar62] in 1962 on Bernoulli convolutions with Mahler measure $M_\lambda = 2$. In addition, C^ℓ densities are established, about which nothing explicit is known for Bernoulli convolutions beyond the case when $\lambda = 2^{-1/k}$ for $k \geq 2$, where Wintner [Win35] proved that ν_λ has a density in $C^{k-2}(\mathbb{R})$.

Furstenberg Measure

Another case of interest are Furstenberg measures. We only discuss $\mathrm{SL}_2(\mathbb{R})$ as much of the recent work on Furstenberg measures is only established for this case.

Consider the $\mathrm{SL}_2(\mathbb{R})$ action on one-dimensional projective space $\mathbb{P}^1(\mathbb{R}) = \mathbb{R}^2/\sim$ given for $g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ and $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in \mathbb{P}^1(\mathbb{R})$ by

$$g.\left[\begin{pmatrix} x \\ y \end{pmatrix}\right] = \left[\begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}\right].$$

We recall the following definitions.

Definition 1.4.9. *A probability measure μ on $\mathrm{SL}_2(\mathbb{R})$ is called **strongly irreducible** if there is no finite subset $S \subset \mathbb{P}^1(\mathbb{R})$ that is invariant under the action of elements of the support of μ.*

*The measure μ is called **unbounded** if its support is not contained in a compact subgroup.*

It is a well-known theorem of Furstenberg (cf. for example [BL85]) that if μ is a strongly irreducible, unbounded probability measure on $\mathrm{SL}_2(\mathbb{R})$, then there exists a unique probability measure ν on $\mathbb{P}^1(\mathbb{R})$ that is μ-stationary, i.e. $\mu * \nu = \nu$. Moreover, the Lyapunov exponent χ_μ, given in this setting by

$$\chi_\mu = \lim_{n\to\infty} \frac{\mathbb{E}_{g\sim\mu^{*n}}[\log ||g||]}{n},$$

is positive, for $||\cdot||$ the operator norm. The measure ν is called the Furstenberg measure of μ. Furstenberg measures and self-similar measures are analogous and there are numerous recent results in both settings.

Kaimanovich-Le Prince [KLP11] initially conjectured that the Furstenberg measure of a finitely supported measure is always singular with respect to the volume measure on $\mathbb{P}^1(\mathbb{R})$. Although there is evidence for the latter conjecture when the support of μ generates a discrete group, Bourgain [Bou12] disproved it.

Theorem 1.4.10. *([Bou12]) For every $\ell \geq 1$, there exists a finitely supported measure on $\mathrm{SL}_2(\mathbb{R})$ whose Furstenberg measure is absolutely continuous with a density in C^ℓ.*

As discussed previously, I have proved in Part I of this thesis that the measures Bourgain constructed satisfy the local limit theorem as in (1.3.2). Moreover, I have generalised Theorem 1.4.10 to simple Lie groups in Theorem 2.0.8. We also point out that Theorem 1.4.10 establishes C^ℓ-densities as in Theorem 1.4.8.

We next mention the following result on the dimension of Furstenberg measures. Hochman and Solomyak proved an analogue of Theorem 1.4.2 for Furstenberg measures.

Theorem 1.4.11. *([HS17]) Let μ be a finitely supported, strongly irreducible and unbounded probability measure on G. Then the Furstenberg measure ν is exact-dimensional (as in (1.4.3)). Moreover, if there is $c > 0$ such that $\Delta_n \geq e^{-cn}$ for all $n \geq 1$, then*

$$\dim \nu = \min\left\{1, \frac{h_\mu}{2\chi_\mu}\right\}.$$

Similarly to (1.4.5), for Furstenberg measures $\dim \nu \leq \min\{1, \frac{h_\mu}{2\chi_\mu}\}$. Therefore, as an absolutely continuous Furstenberg measure has dimension 1, for ν to be absolutely continuous, it must hold $h_\mu \geq 2\chi_\mu$. It is therefore natural to conjecture that

$$\frac{h_\mu}{2\chi_\mu} > 1$$

implies absolute continuity for ν. Recently, my collaborator Samuel Kittle [Kit23] proved a weakening of the latter conjecture and thereby provided numerous new examples of absolutely continuous Furstenberg measures.

To state Kittle's result, let $d(\cdot, \cdot)$ be a left-invariant metric on $\mathrm{SL}_2(\mathbb{R})$, write

$$M_n = \min \left\{ d(g,h) \text{ for } g,h \in \bigcup_{i=0}^{n} \mathrm{supp}(\mu^{*n}) \text{ with } g \neq h \right\},$$

and set

$$S_n = -\frac{1}{n} \log M_n \qquad \text{and} \qquad S_\mu = \limsup_{n\to\infty} S_n. \tag{1.4.9}$$

It is proven in Proposition 15.2.3 that if μ is supported on matrices with coefficients in a number field K of logarithmic height at most L, then

$$S_\mu \ll L[K:\mathbb{Q}]. \tag{1.4.10}$$

Kittle worked with the following definition, for which we endow $\mathbb{P}^1(\mathbb{R})$ with a metric such that it is isometric to $\mathbb{S}^1$.

Definition 1.4.12. *A measure ν on $\mathbb{P}^1(\mathbb{R})$ is said to be (α,t)-**non-degenerate** whenever*

$$\nu(B_t(x)) \leq \alpha$$

for all $x \in \mathbb{P}^1(\mathbb{R})$, there $B_t(x)$ is the open t-ball around x in $\mathbb{P}^1(\mathbb{R})$.

Theorem 1.4.13. *(*[Kit23]*) For every $R > 1$, $\alpha \in (0, \frac{1}{3})$ and $t > 0$ there is a constant $C > 0$ depending on R, α and t such that the following holds.*

Let μ be a finitely supported, strongly irreducible and unbounded probability measure measure on G such that $||g|| \leq R$ for all $g \in \mathrm{supp}(\mu)$ and the Furstenberg measure is (α,t)-non-degenerate. Then ν is absolutely continuous if

$$\frac{h_\mu}{\chi_\mu} > C\left(\max\left\{1, \log\frac{S_\mu}{h_\mu}\right\}\right)^2.$$

We refer to Kittle's paper for a discussion on how to apply Theorem 1.4.13 to construct absolutely continuous Furstenberg measures. As the parameter α needs to be less than $\frac{1}{3}$, the constructed measures are intricate. To state a concrete example, denote for $n \geq 1$ by

$$g = \begin{pmatrix} \frac{n^2-1}{n^2+1} & -\frac{2n}{n^2+1} \\ \frac{2n}{n^2+1} & \frac{n^2-1}{n^2+1} \end{pmatrix} \qquad \text{and} \qquad h = \begin{pmatrix} \frac{n^3+1}{n^3} & 0 \\ 0 & \frac{n^3}{n^3+1} \end{pmatrix}.$$

Then the Furstenberg measure of $\frac{1}{2}(\delta_g + \delta_h)$ is absolutely continuous for sufficiently large n.

1.5 Results of Part II

I will state the main result of Part II (joint work with Samuel Kittle) and mention a few consequences. The main discussion of the result presented will be given in the introduction to Part II. We vastly generalise and strengthen Varjú's result on absolutely continuous Bernoulli convolutions (Theorem 1.4.6) and Lindenstrauss-Varjú's condition on absolute continuity of self-similar measures in dimension $d \geq 3$ (Theorem 1.4.8). The result can be viewed as a strengthening of Theorem 1.4.13 in the context of self-similar measures in arbitrary dimensions.

Similarly to the case of Furstenberg measures, it is conjectured that

$$\frac{h_\mu}{|\chi_\mu|} > d$$

implies that ν is absolutely continuous. The main result of Part II establishes a weakening of this conjecture in the case where the rotation part of the self-similar measure is fixed and the term d is replaced by a constant depending on the rotation part and mildly on the separation rate.

We define S_μ as for Theorem 1.4.13 and note that (1.4.10) holds too.

Theorem 1.5.1. *(Theorem 7.0.4) Let $d \geq 1$ and $\varepsilon \in (0,1)$. Given an irreducible probability measure μ_U on $O(d)$ there exist constants $C \geq 1$ and $\tilde{\rho} \in (0,1)$ depending on d, ε and μ_U such that the following holds. Let $\mu = \sum_{i=1}^k p_i \delta_{g_i}$ be a contracting probability measure on G without a common fixed point satisfying $U(\mu) = \mu_U$ and $p_i \geq \varepsilon$ as well as $\rho(g_i) \in (\tilde{\rho}, 1)$ for all $1 \leq i \leq k$. Then the self-similar measure ν is absolutely continuous if*

$$\frac{h_\mu}{|\chi_\mu|} > C \left(\max \left\{ 1, \log \frac{S_\mu}{h_\mu} \right\} \right)^2.$$

While leaving an extensive discussion of Theorem 1.5.1 to the introduction of Part II, we make the following remarks:

1. Theorem 1.5.1 is significantly simpler to apply than Theorem 1.4.13. However, the proof method of Theorem 1.5.1 is closely related to that of Theorem 1.4.13. The reason why these methods lead to a more applicable result for self-similar measures is because the dynamics of the $\mathrm{Sim}(\mathbb{R}^d)$ action on $\mathbb{R}^d$ is easier to control than the one of the $\mathrm{SL}_2(\mathbb{R})$ action on $\mathbb{P}^1(\mathbb{R})$.

2. Theorem 1.5.1 strengthens Theorem 1.4.6 (see Corollary 7.0.11). For example, when $\lambda = 1 - \frac{p}{q} \in (0,1)$ for coprime natural numbers p, q, then ν_λ is absolutely continuous if
$$\frac{p(\log \log q)^2}{q} \leq c,$$
for c an absolute constant. We deduce similar results for complex Bernoulli convolutions (Corollary 7.0.12).

3. Although our methods cannot conclude results on C^ℓ densities, we can recover the same condition as Theorem 1.4.8 for absolute continuity. Indeed, as discussed in Theorem 8.1.5, our methods establish more explicit dependencies on some of the parameters.

4. An inhomogeneous version of Theorem 1.4.6 (or rather Corollary 7.0.11) is given in Theorem 7.0.7. Indeed, denote for $\lambda_1, \lambda_2 \in (0,1)$ the similarities
$$g_1(x) = \lambda_1 x \qquad \text{and} \qquad g_2(x) = \lambda_2 x + 1.$$
Suppose that $\lambda_i = 1 - \frac{p_i}{q_i}$ for $i = 1,2$ and coprime natural numbers p_i, q_i. Then the self-similar measure of $\mu = \frac{1}{2}\delta_{g_1} + \frac{1}{2}\delta_{g_2}$ is absolutely continuous if
$$\frac{p_i(\log \log q_i)^2}{q_i} \leq c$$
for $i = 1,2$ and c an absolute constant. This established the first absolutely continuous examples of this form when $\lambda_1 \neq \lambda_2$.

5. We construct examples of absolutely continuous self-similar measures for any given collection of irreducible orthogonal matrices $U_1, \ldots, U_k$ and translations $b_1, \ldots, b_k$, provided they all have algebraic entries (Corollary 7.0.8). We note that our result also applies in the case where $U_1, \ldots, U_k$ generates a finite irreducible subgroup of $O(d)$.

6. We also give examples of absolutely continuous measures that arise from generating measures that may have expanding similarities in their support, yet satisfy $\chi_\mu < 0$ (Corollary 7.0.9). We call such measures contracting on average and they will be further discussed in the introduction to Part II.

Part I

Local Limit Theorem on Symmetric Spaces

Chapter 2

Introduction to Part I

Let G be a group and μ a probability measure on G. A fundamental problem in the theory of random walks is to describe the distribution of the product of independent μ-distributed random elements, in other words, to study the measures μ^{*n}. Local limit theorems, which establish the existence of a sequence $a_n \in \mathbb{R}$ such that $a_n\mu^{*n}$ converges to a limit measure, were, as discussed in section 1.2, studied by many authors. The case where G is commutative or compact is classical (cf. for instance [Sto65], [IK40]). Breuillard [Bre05b] and Diaconis-Hough [DH21] considered the Heisenberg group and a local limit theorem for the $\mathrm{Isom}(\mathbb{R}^d)$ action on $\mathbb{R}^d$ was proved by Varjú [Var15] (Theorem 1.2.7). For the latter case, under further assumptions on μ, results with strong error terms were shown by Lindenstrauss-Varjú [LV16] (Theorem 1.2.8). The reader interested in discrete groups may consult Lalley's local limit theorem for the free group [Lal93], which was extended by Gouëzel [Gou14] to hyperbolic groups.

The above results establish local limit theorems for the various mentioned settings under weak assumptions on μ. In contrast, the understanding for non-compact semisimple Lie groups is less developed. The only case where a local limit theorem is known is by assuming that μ is spread out, i.e. a convolution power μ^{*n} for some $n \geq 1$ is not singular to the Haar measure. For spread out measures Bougerol [Bou81] (Theorem 1.3.1) proved in 1981 a local limit theorem.

For a finitely supported measure whose support generates a dense subgroup, the convolutions μ^{*n} become increasingly well-distributed, more and more resembling a continuous measure. Therefore Bougerol's theorem is expected to hold. In this part of the thesis we give the first examples of finitely supported measures on semisimple Lie groups that satisfy Bougerol's theorem for the Lie group acting on the associated symmetric space. Indeed, we reduce the question at hand to understanding spectral properties of a natural operator $S_0 = S_0(\mu)$ associated to μ.

The operator S_0 may be viewed as the Fourier transform of the measure μ at 0 and was studied by Bourgain [Bou12] in his construction of a finitely supported measure on $\mathrm{SL}_2(\mathbb{R})$ with absolutely continuous Furstenberg measure. Further results on S_0 are due to [BISG17], generalizing [Bou12], as well as [BQ18]. These results imply the necessary spectral properties for S_0 in order to establish local limit theorems and will be discussed after stating Theorem 2.0.3. In certain cases, the necessary results for S_0 will also be proved following closely Bourgain's [Bou12] original ideas.

In addition, we deduce quantitative error rates for the local limit theorem (Theorem 2.0.2 and Theorem 2.0.3).

We proceed with stating Bougerol's theorem more precisely than in Theorem 1.3.1. Recall that a measure μ on G is said to be non-degenerate whenever the semigroup generated by its support is dense in G. Let G be a non-compact connected semisimple Lie group with finite center. For a probability measure μ on G, denote $\sigma = ||\lambda_G(\mu)||$, where λ_G is the left regular representation and $\lambda_G(\mu) = \int \lambda_G(g)\, d\mu(g)$. Furthermore denote by p the number of positive indivisible roots of G and by d the rank of G (these notions are further discussed in section 3.1) and write $\ell = 2p + d$. For a non-degenerate and spread out probability measure μ with finite second moment (defined in (2.0.2)), Bougerol [Bou81] showed that there is a continuous function ψ_0 on G (depending on μ) such that

$$\lim_{n\to\infty} \frac{n^{\ell/2}}{\sigma^n} \int f(g)\, d\mu^{*n}(g) = \int f(g)\psi_0(g)\, dm_G(g) \tag{2.0.1}$$

for all $f \in C_c^\infty(G)$. The function ψ_0 satisfies $\mu * \psi_0 = \psi_0 * \mu = \sigma\psi_0$.

To introduce further notation, let K be a maximal compact subgroup of G and denote by $X = G/K$ the associated symmetric space. We recall the definition of the Furstenberg boundary. Let $G = KAN$ be an Iwasawa decomposition of G as introduced in section 3.1. Let M be the centralizer of A in K and write $P = MAN$. The Furstenberg boundary of G is defined as $\Omega = G/P = K/M$. The measure m_Ω is the pushforward of the Haar probability measure m_K onto Ω.

Denote by ρ_0 the Koopman unitary representation of the G action on the measure space (Ω, m_Ω), which is also called the 0-principal series representation (see section 3.1). For a probability measure μ on G, consider the operator $S_0 = \rho_0(\mu) = \int \rho_0(g)\, d\mu(g)$. In order to state the first theorem, recall that a bounded operator is called quasicompact if the essential spectral radius $\rho_{\mathrm{ess}}(A)$ (defined in (3.1.1)) is strictly less than the spectral radius.

Let $\mathfrak{a} = \mathrm{Lie}(A)$ and choose a closed Weyl chamber $\mathfrak{a}^+$. Then for every $g \in G$ denote by $\kappa(g) \in \mathfrak{a}^+$ the unique element such that $g \in K\exp(\kappa(g))K$. We say that μ has finite k-th moment for some $k \geq 1$ if

$$\int |\kappa(g)|^k \, d\mu(g) < \infty. \tag{2.0.2}$$

Theorem 2.0.1. *(Local limit theorem) Let G be a non-compact connected semisimple Lie group with finite center. Choose a maximal compact subgroup K and denote $X = G/K$. Let μ be a non-degenerate probability measure on G with finite second moment and assume that $S_0 = \rho_0(\mu)$ is quasicompact. Write $\sigma = ||\lambda_G(\mu)|| = ||S_0||$ and $\ell = 2p + d$ for p the number of indivisible positive roots of G and d the rank of G.*

Then there is a continuous real-valued function ψ_0 on G satisfying $\mu * \psi_0 = \psi_0 * \mu = \sigma\psi_0$ such that for $x_0 \in X$ and $f \in C_c^\infty(X)$,

$$\lim_{n\to\infty} \frac{n^{\ell/2}}{\sigma^n} \int f(g.x_0)\, d\mu^{*n}(g) = \int f(g.x_0)\psi_0(g)\, dm_G(g). \tag{2.0.3}$$

Moreover, the operator S_0 has a unique σ-eigenfunction $\eta_0 \in L^2(\Omega)$ of unit norm and there exists a unique σ-eigenfunction η_0' of S_0^* satisfying $\langle \eta_0, \eta_0' \rangle = 1$. Then η_0 and η_0' are positive almost surely and ψ_0 is given as $\psi_0(g) = c_\mu \cdot \langle \eta_0, \rho_0(g)\eta_0' \rangle$ for $c_\mu > 0$ a constant depending on μ.

The only difference between (2.0.1) and (2.0.3) is that the latter is only proved on X. Indeed, the limit function of Bougerol's theorem arises as in Theorem 2.0.1 and since a non-degenerate, spread out measure μ satisfies that S_0 is quasicompact (cf. Proposition 2.2.1 of [Bou81]), Theorem 2.0.1 is a generalization of Bougerol's theorem on X. We furthermore mention that it is conjectured that (2.0.1) and therefore also (2.0.3) holds for every non-degenerate probability measure (with finite second moment) on G.

Having stated Theorem 2.0.1, the question arises to give quantitative error rates for (2.0.3). Towards this aim and in order to motivate Theorem 2.0.2, we discuss $G = \mathbb{R}$. Let μ be a non-degenerate measure on $\mathbb{R}$ with mean zero and variance $\sigma^2 < \infty$. The local limit theorem on $\mathbb{R}$ (cf. [Bre92] section 7.4) states that $\sqrt{n}\mu^{*n} \to \frac{m_\mathbb{R}}{\sqrt{2\pi\sigma^2}}$.

Denote

$$\eta_n(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2n\sigma^2}\right).$$

Using that $|\widehat{\mu}(r)| < 1$ for $r \neq 0$ and $\widehat{\mu}(r) = \int e^{irx}\, d\mu(x)$ the Fourier transform of μ, one can show for $f \in C^\infty(\mathbb{R})$ a smooth function whose Fourier transform is compactly supported that there is a constant $c_f = c_f(\mu)$ depending on μ and the support of $\widehat{f}$ such that

$$\sqrt{n}\mu^{*n}(f) = \int f(x)\eta_n(x)\, dm_{\mathbb{R}}(x) + \big(O_\mu(n^{-1}) + O_{\mu,f}(e^{-c_f n})\big)\, ||f||_1, \tag{2.0.4}$$

where the first implied constant depends on μ and the second on μ and the support of $\widehat{f}$. The result (2.0.4) may be referred to as the local central limit theorem as it implies the local limit theorem as well as the central limit theorem. Using that $|\frac{1}{\sqrt{2\pi\sigma^2}} - \eta_n(x)| \ll_\sigma n^{-1}x^2$, it follows that

$$\sqrt{n}\mu^{*n}(f) = \frac{1}{\sqrt{2\pi\sigma^2}} \int f(x)\, dm_{\mathbb{R}}(x) + O_\mu(n^{-1}||f||_*) + O_{\mu,f}(e^{-c_f n}||f||_1) \tag{2.0.5}$$

for

$$||f||_* = \int |f(x)|(1+x^2)\, dm_{\mathbb{R}}(x).$$

We deduce the same behaviour as (2.0.5) even with matching error terms for the G action on its symmetric space under the assumption that S_0 is quasicompact. Choosing a maximal compact subgroup K corresponds to fixing the origin $o = eK \in X$ of X. Denote by $d_X(\cdot,\cdot)$ the distance function induced by a Riemannian metric on X (for which X is a symmetric space, see (3.1.10)). In the theorem below we refer to the Fourier transform of a function $f \in C^\infty(X)$ as discussed in section 3.1. For the asymptotic notation used see also section 3.1.

Theorem 2.0.2. *(Local limit theorem with weak quantitative error rates) With the notation and assumptions from Theorem 2.0.1, assume further that μ has finite fourth moment. Then for $f \in C^\infty(X)$ with compactly supported Fourier transform, there is a constant $c_f = c_f(\mu) > 0$ depending on μ and the support of $\widehat{f}$ such that for $n \geq 1$ and all $x_0 \in X$,*

$$\begin{aligned}\frac{n^{\ell/2}}{\sigma^n}\int f(g.x_0)\, d\mu^{*n}(g) = &\int f(g.x_0)\psi_0(g)\, dm_G(g) \\ &+ O_\mu(n^{-1}||f||_* + n^{-1}d_X(x_0,o)^2||f||_1) + O_{\mu,f}(e^{-c_f n}||f||_1),\end{aligned} \tag{2.0.6}$$

where the first implied constant depends on μ, the second on μ and the support of $\widehat{f}$ and

$$||f||_* = \int |f(x)|(1+d_X(x,o)^2)\, dm_X(x). \tag{2.0.7}$$

For $G = \mathbb{R}$, it is only possible to give strong error rates for (2.0.5) if one gains control over the behaviour of the function $|\widehat{\mu}(r)|$ as $r \to \infty$, which as is shown in

[Bre05a] is equivalent to assuming certain Diophantine properties on the support of μ.

In similar vein, we give strong error rates for (2.0.6) under a suitable Fourier decay assumption. The Schwartz space $\mathscr{S}(X)$ of the theorem below is defined in section 3.1. For $r \in \mathfrak{a}^*$ denote by ρ_r the r-principal series representation defined in (3.1.14) and write

$$S_r = \rho_r(\mu).$$

Theorem 2.0.3. *(Local limit theorem with strong quantitative error rates) With the notation and assumptions from Theorem 2.0.1, assume further that μ has finite fourth moment and that*

$$\sup_{|r|\geq 1} ||S_r|| < ||S_0||. \tag{2.0.8}$$

Then for $f \in \mathscr{S}(X)$, $x_0 \in X$ and $n \geq 1$,

$$\frac{n^{\ell/2}}{\sigma^n} \int f(g.x_0)\, d\mu^{*n}(g) = \int f(g.x_0)\psi_0(g)\, dm_G(g) \\ + O_\mu(n^{-1}||f||_* + n^{-1}d_X(x_0,o)^2||f||_1 + e^{-cn}||f||_{H^s}), \tag{2.0.9}$$

where $c = c(\mu)$ is a constant depending on μ, $s = \frac{1}{2}(\dim X + 1)$, $||\cdot||_{H^s}$ is the Sobolev norm (3.1.18) *of degree s and the implied constant depends only on μ. Moreover, the assumption* (2.0.8) *holds whenever μ is spread out or bi-K-invariant (i.e. $\mu = m_K * \mu * m_K$).*

We proceed with discussing spectral properties of the operator S_0 and also related results on absolute continuity of the Furstenberg measure. In order to introduce convenient notation, the definition of weakly Diophantine measures introduced in [BdS16] and stated in section 1.2 is recalled. We will need to be quantitative about the constants, so we use the following terminology.

Definition 2.0.4. *Let G be a connected Lie group, μ a probability measure on G and let $c_1, c_2 > 0$. The measure μ is called (c_1,c_2)-**weakly Diophantine** or simply (c_1,c_2)-**Diophantine** if*

$$\sup_{H<G} \mu^{*n}(B_{e^{-c_1 n}}(H)) \leq e^{-c_2 n}$$

for sufficiently large n, where $B_{e^{-c_1 n}}(H) = \{g \in G : d(g,H) < e^{-c_1 n}\}$ and the supremum is taken over all closed connected subgroups H of G.

As mentioned in section 1.2, weakly Diophantine measures are useful in understanding random walks on compact groups. Generalizing the Bourgain-Gamburd method developed for SU(2) by [BG08] and for SU(d) in [BG12], it was shown in [BdS16] for K a compact connected simple Lie group, that a symmetric measure μ is (c_1, c_2)-Diophantine for some $c_1, c_2 > 0$ if and only if $\lambda_K(\mu)$ has strong spectral gap on $L^2(K)$ (Definition 4.2.1), in this setting being equivalent to $||\lambda_K(\mu)|_{L^2_0(K)}||_{\text{op}} < 1$ for $L^2_0(K) = \{f \in L^2(K) : m_K(f) = 0\}$. Indeed, the essential spectral radius of $\lambda_K(\mu)$ can be bounded in terms of K, c_1 and c_2. As discussed in section 1.2, strong spectral gap on $L^2_0(K)$ can be used to deduce that μ^{*n} equidistributes towards the Haar measure with exponential speed.

For finitely supported measures, most known spectral results for S_0 also rely on the Bourgain-Gamburd method. However, one requires stronger Diophantine conditions. Indeed, as in contrast to compact groups it is necessary to control the exponential norm growth of the μ-random walk on G, we have to demand that the measure is (c_1, c_2)-Diophantine while being close to the identity in terms of c_1 and c_2. We therefore introduce the following definition.

Definition 2.0.5. *Let G be a connected Lie group, μ a probability measure on G and let $c_1, c_2, \varepsilon > 0$. The measure μ is called (c_1, c_2, ε)-**Diophantine** if*

(i) μ is $(c_1 \log \frac{1}{\varepsilon}, c_2 \log \frac{1}{\varepsilon})$-Diophantine, i.e. for n large enough,

$$\sup_{H<G} \mu^{*n}(B_{\varepsilon^{c_1 n}}(H)) \leq \varepsilon^{c_2 n}.$$

(ii) $\text{supp}(\mu) \subset B_\varepsilon(e)$.

We state a result of [BISG17] showing that there is an abundant collection of examples of (c_1, c_2, ε)- Diophantine measures for arbitrarily small ε.

Theorem 2.0.6. *(Theorem 3.1 of* [BISG17]*) Let G be a connected simple Lie group with finite center and adjoint representation $\text{Ad} : G \to \text{GL}(\mathfrak{g})$. Let $\Gamma < G$ be a countable dense subgroup and assume that there is a basis of $\mathfrak{g}$ such that $\text{Ad}(\gamma)$ is algebraic with respect to that basis for every $\gamma \in \Gamma$.*

Then there exist $c_1, c_2 > 0$ such that for every $\varepsilon_0 > 0$ there is $0 < \varepsilon < \varepsilon_0$ and a finitely supported symmetric (c_1, c_2, ε)-Diophantine probability measure μ satisfying $\text{supp}(\mu) \subset \Gamma \cap B_\varepsilon$.

Using the above defined notion of Diophantine measures, one can establish the following result on quasicompactness of S_0. Together with Theorem 2.0.6, numerous examples of finitely supported measures satisfying (2.0.3) are provided.

Theorem 2.0.7. *Let G be a non-compact connected simple Lie group with finite center. Let $c_1, c_2 > 0$. Then there is $\varepsilon_0 = \varepsilon_0(G, c_1, c_2) > 0$ depending on G and $c_1, c_2 > 0$, such that every symmetric and (c_1, c_2, ε)-Diophantine probability measure μ with $\varepsilon \leq \varepsilon_0$ satisfies that $S_0 = \rho_0(\mu)$ is quasicompact. In particular, Theorem 2.0.1 and Theorem 2.0.2 holds for μ.*

Theorem 2.0.7 is a straightforward consequence of the techniques and results developed in [BISG17] and will be deduced in section 6.0.1. Under the additional assumption that the maximal compact subgroup is semisimple, we offer an alternative proof following more closely the method by Bourgain [Bou12], leading to marginally stronger results (Theorem 6.0.2). Indeed, using an idea from [LV16], we simplify Bourgain's original approach by exploiting that the irreducible representations of K have high dimension.

We proceed with discussing the Furstenberg measure. Let μ be a measure on G whose support generates a Zariski dense subgroup. Then the Furstenberg measure of μ is the unique μ-stationary Borel probability measure ν_{F} on the boundary Ω (cf. for example [GdM89]). It was initially conjectured by Kaimanovich-Le Prince [KLP11] that the Furstenberg measure of a finitely supported measure is singular to the Haar measure m_Ω. However Bourgain [Bou12] and Bárány-Pollicott-Simon [BPS12] disproved the latter conjecture, with Bourgain [Bou12] giving an explicit construction while [BPS12] exploiting probabilistic methods.

[BQ18] also provide examples of finitely supported measures with absolutely continuous Furstenberg measure, yet their construction does not lead to results as versatile as Theorem 2.0.6. It is apparent from their proof, that S_0 is also quasicompact for these examples.

A further result of [Bou12] is the construction of finitely supported measures on $\mathrm{SL}_2(\mathbb{R})$ satisfying $\frac{d\nu_{\mathrm{F}}}{dm_\Omega} \in C^k(\Omega)$ for any $k \in \mathbb{Z}_{\geq 1}$. Following Bourgain's technique, we also deduce smoothness results for the Furstenberg measure for arbitrary simple Lie groups.

Theorem 2.0.8. *Let G be a non-compact connected simple Lie group with finite center. Let $c_1, c_2 > 0$ and $m \in \mathbb{Z}_{\geq 1}$. Then there is $\varepsilon_m = \varepsilon_m(G, c_1, c_2) > 0$ depending on G, c_1, c_2 and m such that every symmetric and (c_1, c_2, ε)-Diophantine probability measure μ with $\varepsilon \leq \varepsilon_m$ has absolutely continuous Furstenberg measure with density in $C^m(\Omega)$.*

After publishing [Kog22], the author became aware of [Leq22] who establishes a similar yet less general result to Theorem 2.0.8. Since our proof is short and differs from [Leq22] for instance in introducing Agmon's inequality (Lemma 6.4.2) for compact Lie groups, it is included in this thesis.

We comment on the organization of Part I. After reviewing the necessary notation and giving an outline of proofs in section 3, we discuss some preliminary results in section 4. Then the local limit theorems Theorem 2.0.1, Theorem 2.0.2 and Theorem 2.0.3 are proved in section 5. Finally, quasicompactness of S_0 and the Furstenberg measure are discussed in section 6, establishing Theorem 2.0.7 and Theorem 2.0.8.

Chapter 3

Notation and Outline

3.1 Notation for Part I

In this section we collect the notations used in Part I. The reader may also recall the notation stated in section 1.1.

Throughout Part I of this thesis, we denote by G a non-compact connected semisimple Lie group with finite center, by K a maximal compact subgroup of G and write $X = G/K$ for the associated symmetric space.

Let $\mathscr{B}$ be a Banach space and let $A : \mathscr{B} \to \mathscr{B}$ be a bounded operator. Recall that A is called a Fredholm operator if there exists a bounded operator T such that $TA - \mathrm{Id}$ and $AT - \mathrm{Id}$ are compact operators. Denote by $\mathrm{spec}(A)$ the spectrum of A. The essential spectrum $\mathrm{spec}_{\mathrm{ess}}(A)$ is defined as the set of complex numbers λ such that $A - \lambda \cdot \mathrm{Id}$ is not Fredholm. The spectral radius is defined as $\rho(A) = \max_{\lambda \in \mathrm{spec}(A)} |\lambda|$ and the essential spectral radius as

$$\rho_{\mathrm{ess}}(A) = \max_{\lambda \in \mathrm{spec}_{\mathrm{ess}}(A)} |\lambda|, \tag{3.1.1}$$

if $\rho_{\mathrm{ess}}(A) \neq \emptyset$ and otherwise $\rho_{\mathrm{ess}}(A) = 0$.

For a locally compact Hausdorff group H, write m_H for a fixed choice of Haar measure. Whenever H is compact, m_H is the Haar probability measure. The left-regular representation is denoted λ_H while we write ρ_H for the right regular representation.

If μ is a finite measure on H and $\pi : H \to \mathscr{U}(\mathscr{H})$ is a unitary representation, where $\mathscr{H}$ is a Hilbert space and $\mathscr{U}(\mathscr{H})$ the space of unitary operators $\mathscr{H} \to \mathscr{H}$, then

$$\pi(\mu) = \int \pi_g \, d\mu(g) \tag{3.1.2}$$

is the operator uniquely characterized by $\langle \pi(\mu)v, w\rangle = \int \langle \pi_g v, w\rangle \, d\mu(g)$ for $v, w \in \mathscr{H}$.

For a group H with metric d_H, for $R > 0$ and $x \in H$ we will denote by $B_R(x) = \{y \in H : d_H(y,x) < R\}$ and abbreviate $B_R = B_R(e)$ for $e \in H$ the identity element. On G we fix a left invariant metric such that $B_R(g) = gB_R(e)$. For a closed subset $H' \subset H$ we define $B_R(H') = \{h \in H : d(h,H') < R\}$, where $d(h,H') = \sup_{h' \in H'} d(h,h')$.

We first fix notation for structure theory on K. Write T for a maximal torus in K with Lie algebra $\mathfrak{t}$ and real dual Lie algebra $\mathfrak{t}^*$. Let W_K be the Weyl group and we fix a W_K-invariant inner product on $\mathfrak{t}$, inducing a W_K-invariant inner product on $\mathfrak{t}^*$. The set of real roots is denoted as R and we choose a fundamental Weyl chamber C which we consider as a subset of $\mathfrak{t}^*$. The fundamental Weyl chamber determines a basis S of the real roots and the set of positive roots R^+. We denote by $I^* \subset \mathfrak{t}^*$ the set of integral forms. Then (cf. [BtD85] section 6) the set $\overline{C} \cap I^*$ parametrizes the irreducible representations of K.

For $\gamma \in \overline{C} \cap I^*$ denote by π_γ the associated irreducible unitary representation of K and by M_γ the span of matrix coefficients of π_γ. By the Peter-Weyl Theorem it holds that

$$L^2(K) = \bigoplus_{\gamma \in \overline{C} \cap I^*} M_\gamma, \tag{3.1.3}$$

where we used the convention applied throughout this paper that by a direct sum we denote the closure of the algebraic direct sum of the involved vector spaces. For any $\gamma \in \overline{C} \cap I^*$ and an orthonormal basis $v_1, \ldots, v_{d_\gamma}$ of π_γ, we set $\chi_{ij}^\gamma(k) = \langle \pi_\gamma(k)v_i, v_j \rangle$. Then the set of functions $d_\gamma^{1/2}\chi_{ij}^\gamma$ forms an orthonormal basis of $L^2(K)$. For $\varphi \in L^2(K)$, we set $\widehat{\varphi}_{ij}^\gamma = a_{ij}^\gamma = \langle \varphi, d_\gamma^{1/2}\chi_{ij}^\gamma \rangle$. For $\varphi \in C^\infty(K)$ and all $k \in K$,

$$\varphi(k) = \sum_{\gamma \in \overline{C} \cap I^*} \sum_{i,j=1}^{d_\gamma} d_\gamma^{1/2} a_{ij}^\gamma \chi_{ij}^\gamma(k). \tag{3.1.4}$$

We want to group together functions on K that oscillate at roughly the same rate. Therefore, one defines

$$V_0 = \bigoplus_{\substack{\gamma \in \overline{C} \cap I^* \\ 0 \le ||\gamma|| < 1}} M_\gamma \qquad \text{and} \qquad V_\ell = \bigoplus_{\substack{\gamma \in \overline{C} \cap I^* \\ 2^{\ell-1} \le ||\gamma|| < 2^\ell}} M_\gamma \tag{3.1.5}$$

for $\ell \ge 1$. The decomposition

$$L^2(K) = \bigoplus_{\ell \ge 0} V_\ell \tag{3.1.6}$$

is referred to as the Littlewood-Paley decomposition of $L^2(K)$. For $\ell \ge 0$ we denote by π_ℓ the orthogonal projection from $L^2(K)$ to V_ℓ. Therefore any $\varphi \in L^2(K)$ can be decomposed as $\varphi = \sum_{\ell \geq 0} \pi_\ell \varphi$. For Littlewood-Paley decompositions on groups in more general contexts we refer the reader to [MKMSG22].

We finally define Sobolev spaces and Sobolev norms on K. Denote by $\mathfrak{k}$ the Lie algebra of K and fix an orthonormal basis $X_1, \ldots, X_n$ of $\mathfrak{k}$. Then the Casimir operator given by $\triangle = -\sum_{i=1}^n X_i \circ X_i$ is a central element of the universal enveloping algebra $U(\mathfrak{k})$. For $\gamma \in \overline{C} \cap I^*$ denote by λ_γ the eigenvalue of $\triangle$ acting on π_γ. For $s \in \mathbb{Z}_{\geq 0}$, we define

$$\begin{aligned} H^s(K) &= \{\varphi \in L^2(K) \,:\, \lambda_K(\triangle)^{s/2}\varphi \in L^2(K)\} \\ &= \left\{ \varphi = \sum_{\gamma \in \overline{C} \cap I^*} \varphi_\gamma \in \bigoplus_{\gamma \in \overline{C} \cap I^*} M_\gamma \,:\, ||\varphi||_{H^s}^2 = \sum_{\gamma \in \overline{C} \cap I^*} \lambda_\gamma^s ||\varphi_\gamma||_2^2 < \infty \right\}. \end{aligned} \tag{3.1.7}$$

We also need structure theory for G. We take care not to confuse the notation introduced for the structure theory of K. The Lie algebra of G is denoted as $\mathfrak{g}$ and we choose a Cartan decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n}$ for $\mathfrak{k}$ the Lie algebra of K. Denote by $\mathfrak{a}^*$ the real dual of $\mathfrak{a}$. Let Σ be the sets of roots, choose a closed Weyl chamber $\mathfrak{a}^+$ and let $\Sigma^+ = \{r_1, \ldots, r_k\} \subset \mathfrak{a}^*$ be the system of positive roots. For a root $r \in \Sigma$ write $m(r)$ for the multiplicity of r and denote by $\delta = \frac{1}{2}\sum_{r \in \Sigma^+} m(r) r$ the half sum of the positive roots counted with multiplicities. We fix a norm $|\cdot|$ on $\mathfrak{g}$ arising from an Ad-invariant inner product. The latter norm restricts to $\mathfrak{a}$ and induces the operator norm on $\mathfrak{a}^*$.

Denote $A = \exp(\mathfrak{a})$, $N = \exp(\mathfrak{n})$ and $P^+ = AN$. Then (cf. [Kna02] chapter VI) the multiplication map $K \times A \times N \to G$ is a diffeomorphism, giving rise to the Iwasawa decomposition $G = KAN$. Write further $K : G \to K$, $A : G \to A$ and $N : G \to N$ for the maps induced from the Iwasawa decomposition and the map $H : G \to \mathfrak{a}$ is defined for $g \in G$ as

$$H(g) = \log A(g). \tag{3.1.8}$$

Set $A^+ = \exp(\mathfrak{a}^+)$. Then the Cartan decomposition $G = KA^+K$ holds and denote by $\kappa : G \to \mathfrak{a}^+$ the map uniquely characterized by $g \in K \exp(\kappa(g))K$. We furthermore define

$$||g|| = |\kappa(g)|. \tag{3.1.9}$$

On the symmetric space $X = G/K$, one defines the metric d_X as

$$d_X(g.o, o) = |\kappa(g)| \tag{3.1.10}$$

for the origin $o = K \in X$ and all $g \in G$. Then for $g \in KA$ it holds that $|H(g)| = d_X(g.o, o) = |\kappa(g)|$. Recall Exercise B2 (iv) from Chapter VI of [Hel78] stating that $d(a.o, o) \leq d(an.o, o)$ for all $a \in A$ and $n \in N$, which follows by applying suitably that the manifolds $A.o$ and $N.o$ are perpendicular at their unique intersection point $o \in X$. It therefore holds for all $g \in G$ that

$$|H(g)| \leq |\kappa(g)| = ||g||. \tag{3.1.11}$$

For each $g \in G$ consider the diffeomorphism

$$\alpha_g : K \to K, \qquad k \mapsto \alpha_g(k) = K(gk).$$

The map $G \to \mathrm{Diff}(K), g \mapsto \alpha_g$ defines an action of G on K. Denote by α'_g the Radon-Nikodym derivative of $(\alpha_g)_* m_K$ with respect to m_K. Then by I Lemma 5.19 of [Hel84],

$$\alpha'_g(k) = \frac{d(\alpha_g)_* m_K}{dm_K}(k) = e^{-2\delta H(g^{-1}k)}. \tag{3.1.12}$$

For $r \in \mathfrak{a}^*$, we consider the unitary representation $\rho_r^+ : G \to L^2(K)$ defined for $g \in G$ and $\varphi \in L^2(K)$ as

$$(\rho_r^+(g)\varphi)(k) = e^{-(\delta+ir)H(g^{-1}k)}\varphi(K(g^{-1}k)) \tag{3.1.13}$$

with $k \in K$.

The representation (3.1.13) is not irreducible in general. In order to make it irreducible, denote by M the centralizer of A in K and write $P = MAN$ for the associated minimal parabolic subgroup. The Furstenberg boundary $\Omega = G/P$ can be identified with K/M and we therefore view functions on Ω as M-invariant functions on K. The probability measure m_Ω is the pushforward of m_K under the projection map. For $r \in \mathfrak{a}^*$ we consider the r-principal series $\rho_r : G \to \mathscr{U}(L^2(\Omega))$ defined for $g \in G$ and $\varphi \in L^2(\Omega)$,

$$(\rho_r(g)\varphi)(\omega) = e^{-(\delta+ir)H(g^{-1}\omega)}\varphi(g^{-1}\omega) \tag{3.1.14}$$

for $\omega \in \Omega$ where we denote by $g^{-1}\omega$ the element $K(g^{-1}k)M$ for any representative $\omega = kM$ with $k \in K$ and note that $H(g^{-1}\omega)$ does not depend on the representative of ω (cf. [War72] section 5.5). The principal series is irreducible.

The Weyl group W_G of G is defined as the group quotient $N_K(\mathfrak{a})/Z_K(\mathfrak{a})$, where $N_K(\mathfrak{a}) = \{k \in K : \mathrm{Ad}(k)\mathfrak{a} \subset \mathfrak{a}\}$ and $Z_K(\mathfrak{a}) = M = \{k \in K : ka = ak \text{ for all } a \in A\}$.

We call a root $r \in \Sigma$ indivisible if $\frac{1}{2}r$ is not a root and we order the positive roots in such a way that $r_1, \ldots, r_p$ are the indivisible roots. For any complex linear form r on $\mathfrak{a}$ denote

$$I(r) = \left(\prod_{\ell=1}^{p} B\left(\frac{m(r_\ell)}{2}, \frac{\langle r, r_\ell \rangle}{\langle r_\ell, r_\ell \rangle} \right) \right) \cdot \left(\prod_{\ell=p+1}^{k} B\left(\frac{m(r_\ell)}{2}, \frac{m(r_\ell/2)}{4} + \frac{\langle r, r_\ell \rangle}{\langle r_\ell, r_\ell \rangle} \right) \right),$$

where $B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}\, dt$ is the Beta function. We further set for $r \in \mathfrak{a}^*$,

$$c(r) = \frac{I(ir)}{I(\delta)}.$$

The spherical function of parameter $r \in \mathfrak{a}^*$ is defined as $\phi_r(g) = \langle \rho_r(g)1, 1 \rangle$. Denote by $\mathscr{D}(G)$ the set of differential operators on G (see [Hel84] chapter 2). The Harish-Chandra Schwartz space introduced in [HC58] (see further page 230 of [Wal88]) is defined as

$$\mathscr{S}(G) = \{ f \in C^\infty(G) \,:\, (1 + |H(g)|)^\ell |Df|(g) \ll_{f,D,\ell} \phi_0(g) \text{ for all } D \in \mathscr{D}(G), \ell \geq 0 \}. \tag{3.1.15}$$

The Schwartz space on X, denoted $\mathscr{S}(X)$, is defined as the set of right K-invariant functions in $\mathscr{S}(G)$.

Recall that a function f on G is called bi-K-invariant or radial if $f(k_1 g k_2) = f(g)$ for all $g \in G$ and $k_1, k_2 \in K$. For a radial function $f \in \mathscr{S}(G)$ we denote by $\rho_r(f)$ as in (3.1.2) the operator $\int f(g)\rho_r(g)\, dm_G(g)$. We then define the spherical Fourier transform as

$$\widehat{f}(r) = \langle 1, \rho_r(f)1 \rangle = \langle \rho_{-r}(f)1, 1 \rangle = \int_G f(g)\phi_{-r}(g)\, dm_G(g).$$

Note that by using that f is bi-K-invariant, it follows that for all $\omega \in \Omega$ we have $\widehat{f}(r) = (\rho_{-r}(f)1)(\omega)$. For all $g \in G$, the spherical Fourier inversion formula holds

$$f(g) = \int_{\mathfrak{a}^*} \widehat{f}(r)\phi_r(g)\, d\nu_{\mathrm{sph}}(r), \tag{3.1.16}$$

where $d\nu_{\mathrm{sph}}(r) = |c(r)|^{-2} dm_{\mathfrak{a}^*}(r)$ is the spherical Plancharel measure.

We furthermore define for $f \in \mathscr{S}(X)$, $r \in \mathfrak{a}^*$ and $\omega \in \Omega$,

$$\widehat{f}(r, \omega) = (\rho_{-r}(f)1)(\omega) = \int_G f(g)(\rho_{-r}(g)1)(\omega)\, dm_G(g).$$

Then it follows by a brief calculation from (3.1.16), for $f \in \mathscr{S}(X)$ and $g \in G$,

$$f(g) = \int_{\mathfrak{a}^*} \int_\Omega \widehat{f}(r, \omega)(\rho_r(g)1)(\omega)\, dm_\Omega(\omega) d\nu_{\mathrm{sph}}(r). \tag{3.1.17}$$

We say that $f \in \mathscr{S}(X)$ has compactly supported Fourier transform if there is a constant $R > 0$ such that $\widehat{f}(r,\omega) = 0$ for $|r| \geq R$ and $\omega \in \Omega$.

We will further need Sobolev spaces and Sobolev norms on X, defined for $s \geq 0$ as

$$H^s(X) = \left\{ f \in L^2(X) \,:\, ||f||_{H^s}^2 = \int_{\mathfrak{a}^*} ||\widehat{f}(r,\cdot)||_{L^2(\Omega)}^2 (1+|r|^2)^s \, d\nu_{\mathrm{sph}}(r) < \infty \right\}. \tag{3.1.18}$$

It holds that $C_c^\infty(X) \subset \mathscr{S}(X) \subset H^s(X)$ for all $s \geq 0$ (c.f. [Hel84] chapter IV).

For a probability measure μ on G, we write for $r \in \mathfrak{a}^*$

$$S_r^+ = \rho_r^+(\mu) \quad \text{ and } \quad S_r = \rho_r(\mu), \tag{3.1.19}$$

using the definition (3.1.2) for the unitary representations ρ_r^+ and ρ_r.

We further use the notation $\sigma = ||S_0||$. Since MAN is an amenable group, it holds by section D of [Gui80] that $\sigma = ||\lambda_G(\mu)||$. If $\lambda(r) \in \mathbb{C}$ satisfying $|\lambda(r)| = \rho(S_r)$ is in the discrete spectrum of S_r, has geometric multiplicity one and is the unique element of $\mathrm{spec}(S_r)$ on the circle of radius $\rho(S_r)$, then we denote by $\eta_r \in L^2(\Omega)$ the $\lambda(r)$-eigenfunction of S_r with unit norm. Furthermore, if the same properties hold for S_r^* and $\overline{\lambda(r)}$, choose η_r' the S_r^*-eigenfunction with eigenvalue $\overline{\lambda(r)}$ satisfying $\langle \eta_r', \eta_r \rangle = 1$, provided there exists such an η_r'. Then we denote

$$\psi_{\mu,r}(g) = \langle \eta_r, \rho_r(g)\eta_r' \rangle \tag{3.1.20}$$

for $g \in G$.

The operator $T_0 : L^2(\Omega) \to L^2(\Omega)$ is defined as

$$T_0\varphi = \int \varphi \circ \alpha_g \, d\mu(g)$$

for $\varphi \in L^2(\Omega)$, where we equally denote by $\alpha_g : \Omega \to \Omega$ the map on Ω induced by $\alpha_g : K \to K$, and

$$T_0^+ : L^2(K) \to L^2(K) \quad \text{ defined as } \quad T_0^+\varphi = \int \varphi \circ \alpha_g \, d\mu(g)$$

for $\varphi \in L^2(K)$.

3.2 Outline of Proofs

For the proof of Theorem 2.0.1, Theorem 2.0.2 and Theorem 2.0.3 one uses the Fourier inversion formula on X to reduce the question at hand to spectral problems about the operators S_r. Indeed, by (3.1.17) it holds for $x_0 = h_0K \in X$ with $h_0 \in G$ and $f \in \mathscr{S}(X)$,

$$\frac{n^{\ell/2}}{\sigma^n}\int f(g.x_0)\,d\mu^{*n}(g) = \frac{n^{\ell/2}}{\sigma^n}\int_{\mathfrak{a}^*}\int_{\Omega}\widehat{f}(r,\omega)(S_r^n\rho_r(h_0)1)(\omega)\,dm_\Omega(\omega)d\nu_{\mathrm{sph}}(r). \quad (3.2.1)$$

One then decomposes (3.2.1) into high and low frequencies. Namely for $\delta_0 \in (0,1)$ small enough depending on μ and for $f \in \mathscr{S}(X)$,

$$\frac{n^{\ell/2}}{\sigma^n}\int f(g.x_0)\,d\mu^{*n}(g) = \frac{n^{\ell/2}}{\sigma^n}\int_{|r|>\delta_0}\int_{\Omega}\widehat{f}(r,\omega)(S_r^n\rho_r(h_0)1)(\omega)\,dm_\Omega(\omega)d\nu_{\mathrm{sph}}(r) \quad (3.2.2)$$

$$+\frac{n^{\ell/2}}{\sigma^n}\int_{|r|\leq\delta_0}\int_{\Omega}\widehat{f}(r,\omega)(S_r^n\rho_r(h_0)1)(\omega)\,dm_\Omega(\omega)d\nu_{\mathrm{sph}}(r). \quad (3.2.3)$$

The following spectral properties of S_r are used to deal with the arising terms:

(1) There are operators E_0 and D_0 such that

$$S_0 = \sigma E_0 + D_0, \quad (3.2.4)$$

where E_0 is a projection to a one-dimensional subspace, $E_0 \circ D_0 = D_0 \circ E_0 = 0$ and D_0 satisfies $\rho(D_0) < \sigma = ||S_0||$. In section 4.2 we refer to the property (3.2.4) as strong spectral gap.

(2) For $|r| \leq \delta_0$, the operator S_r has a decomposition as (3.2.4), i.e.

$$S_r = \lambda(r)E_r + D_r, \quad (3.2.5)$$

for E_r and D_r as in (3.2.4).

(3) For any $r \neq 0$, $\rho(S_r) < \sigma = ||S_0||$.

One deduces (1) from quasicompactness of S_0 and by using that S_0 is a positive operator in the sense of Banach lattices (c.f. section 4.2). (2) will follow as quasicompactness is an open property under certain assumptions (Corollary 4.1.2) and (3) by a convexity argument similar to an argument of Conze-Guivarc'h [CG13]. The necessary spectral properties are proved in section 5.1.

Properties (1) and (2) will be necessary to deal with low frequencies (3.2.3), whereas (3) is used for high frequencies (3.2.2). However, (3) only allows to prove a decay for (3.2.2) either by assuming that f has compactly supported Fourier transform or by imposing the stronger assumption $(\sup_{|r|\geq 1} ||S_r||) < ||S_0||$ of Theorem 2.0.3. One then deduces Theorem 2.0.1 and Theorem 2.0.2 by approximating a given function $f \in \mathscr{S}(X)$ with functions whose Fourier transform is compactly supported.

A novel contribution is the observation that the functions $\psi_{\mu,r}$ as defined in (3.1.20), where $|r| \leq \delta_0$ such that (3.2.5) holds, satisfy

$$\int f \cdot \psi_{\mu,r} \, dm_G = \int_\Omega \widehat{f}(r,\omega)(E_r 1)(\omega) \, dm_\Omega(\omega) \tag{3.2.6}$$

for $f \in \mathscr{S}(X)$ (see Lemma 5.2.1). We further mention that (3.2.6) may be viewed as an analogue of the formula

$$\int f(x) e^{-\sigma^2 \frac{x^2}{2}} \, dm_{\mathbb{R}}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int \widehat{f}(r) e^{-\frac{r^2}{2\sigma^2}} \, dm_{\mathbb{R}}(r) \tag{3.2.7}$$

on $\mathbb{R}$, where $f \in \mathscr{S}(\mathbb{R})$ and $\sigma > 0$, which is used in the proof of the local limit theorem on $\mathbb{R}$.

The outline of the proof of the local limit theorem is concluded. We next discuss quasicompactness of S_0. As in [Bou12] and [BISG17], the main tool are flattening statements for μ. These results, which will be recalled in section 4.4, have as a consequence that for any $\gamma > 0$ and $x \in G$,

$$\mu^{*n}(B_\delta(x)) \ll \delta^{\dim G - \gamma} \tag{3.2.8}$$

for δ small enough depending on μ and γ and $n \asymp_{\mu,\gamma} \log \frac{1}{\delta}$. A measure with property (3.2.8) can be referred to as having high dimension, since an absolutely continuous measure ν satisfies $\nu(B_\delta(x)) \asymp_\nu \delta^{\dim G}$.

The proof of quasicompactness of S_0 comprises two steps. First we will show that the restricted operator $S_0|_{V_\ell}$ has small norm for all ℓ large enough, where V_ℓ is the Littlewood-Paley space introduced in (3.1.5). The second step is to use the latter to deduce that S_0 restricted to $\bigoplus_{\ell \geq L} V_\ell$ has small norm for a suitable $L > 0$ and therefore is quasicompact. This exploits the first step and that the spaces V_ℓ are mutually orthogonal. Indeed, since the measure μ in question is supported close to the identity, the spaces $S_0 V_\ell$ and $V_{\ell'}$ for $\ell \neq \ell'$ are almost orthogonal too.

For the first step, one uses that for $\varphi \in V_\ell$ the matrix coefficients $|\langle \rho_0(g)\varphi, \varphi\rangle|$ are small on average. Indeed, it is shown in section 4.5, following [LV16], that

$$\frac{1}{m_G(B_R)} \int_{B_R} |\langle \rho_0(g)\varphi, \varphi\rangle| \, dm_G(g) \ll 2^{-\ell/2} ||\varphi||_2. \tag{3.2.9}$$

Since μ has high dimension, we are able to use (3.2.9) to give strong estimates for $\langle S_0\varphi, \varphi\rangle$ and therefore conclude a bound on the operator norm of $S_0|_{V_\ell}$.

In order to use (3.2.9), we ought to control the size of the support of μ^{*n} while ensuring that μ^{*n} has high dimension (3.2.8) quickly enough. Analogous to [Bou12] and [BISG17], this is where the (c_1, c_2, ε)-Diophantine property comes into play. Indeed, as ε becomes smaller, a (c_1, c_2, ε)-Diophantine measure is increasingly rapidly non-concentrated on subgroups and therefore a strong flattening lemma applies (Lemma 4.4.2). The latter holds while the measure is still close to the identity, which will allow us to conclude the claimed properties for S_0.

3.3 Relation to Other Work

As mentioned in the introduction, the necessary results for S_0 are also proved in [BISG17]. The main difference between [BISG17] and our proof is in the use of a different Littlewood-Paley decomposition. [BISG17] develop a Littlewood-Paley decomposition on G, which leads to more general results as they are able to deal with all possible quotients of G, while we work with the Littlewood-Paley decomposition on K, leading to marginally stronger results.

For the $\mathrm{Isom}(\mathbb{R}^d)$ action on $\mathbb{R}^d$, a similar representation theoretic decomposition to (3.1.17) holds for a suitable family of unitary representations $\rho_r : \mathrm{Isom}(\mathbb{R}^d) \to \mathscr{U}(L^2(\mathbb{S}^{d-1}))$ for $r \in \mathbb{R}$. In [LV16], a local limit theorem with strong error terms as in Theorem 2.0.3 is proved by just assuming that $S_0 = \rho_0(\mu)$ is quasicompact. Indeed they establish (2.0.8) for their setting by solely assuming that S_0 is quasicompact. It seems reasonable to believe that the same result may hold for a semisimple Lie group acting on its symmetric space, yet the proof of [LV16] is not transferable as several properties only applicable to $\mathrm{Isom}(\mathbb{R}^d)$ are used.

We further mention that in [Tol00] a Berry-Essen result is shown on G for a probability measure with a smooth density of compact support.

Chapter 4

Preliminary Results

4.1 Quasicompact Operators

Throughout this section we denote by $\mathscr{B}$ a separable Banach space and the reader may recall the notations introduced in section 3.1. A bounded operator $A : \mathscr{B} \to \mathscr{B}$ is called quasicompact if $\rho_{\text{ess}}(A) < \rho(A)$. In this secton we show that being quasicompact is an open property. We first state a useful lemma.

Lemma 4.1.1. *For any bounded operator $A : \mathscr{B} \to \mathscr{B}$ the following properties hold:*

(i)
$$\rho_{\text{ess}}(A) = \inf_{U \ \textit{compact}} \rho(A - U).$$

(ii) A is quasicompact whenever A^* is. Moreover,
$$\rho_{\text{ess}}(A^*) = \rho_{\text{ess}}(A).$$

(ii) The set of spectral values of A with modulus $> \rho_{\text{ess}}(A)$ is at most countable and all of its accumulation points have modulus $\rho_{\text{ess}}(A)$.

Proof. (i) follows as the essential spectral radius is the spectral radius of the image of A in the Calkin algebra (c.f section 2.4 in the appendix of [BQ16]) and (ii) as a bounded operator is Fredholm whenever its adjoint is (Corollary 2.12 of appendix B in [BQ16]). Finally (iii) is contained in Proposition 2.14 of appendix B in [BQ16]. □

Corollary 4.1.2. *Let $A_n : \mathscr{H} \to \mathscr{H}$ be a sequence of bounded operators on a Hilbert space $\mathscr{H}$ converging in operator norm to a bounded operator $A : \mathscr{H} \to \mathscr{H}$. If A is quasicompact then so is A_n for n large enough and there is $\varepsilon > 0$ such that for n large enough $\rho_{\text{ess}}(A_n) < \rho_{\text{ess}}(A) + \varepsilon < \rho(A) - \varepsilon < \rho(A_n)$.*

Proof. By Lemma 4.1.1 (i) for any $\varepsilon > 0$ there is a compact operator U (depending on ε) such that $\rho(A-U) < \rho_{\text{ess}}(A)+\varepsilon$. We choose a small $\varepsilon > 0$ such that $\rho_{\text{ess}}(A)+2\varepsilon < \rho(A) - 2\varepsilon$. Recall that the spectral radius is upper semi-continuous and since A is quasicompact, A is a continuity point for the spectral radius (cf. [New51]). Thus for large enough n it holds that $\rho(A_n - U) < \rho(A-U)+\varepsilon$ and $\rho(A)-2\varepsilon < \rho(A_n)$. Then for the above compact operator U,

$$\rho_{\text{ess}}(A_n) \leq \rho(A_n - U) < \rho(A-U) + \varepsilon < \rho_{\text{ess}}(A) + 2\varepsilon < \rho(A) - 2\varepsilon < \rho(A_n),$$

showing the claim upon replacing 2ε by ε. □

4.2 Strong Spectral Gap and Quasicompact Positive Operators

We introduce the following definition of strong spectral gap.

Definition 4.2.1. *Let $S : \mathscr{B} \to \mathscr{B}$ be a bounded operator on a Banach space $\mathscr{B}$. We say that S has **strong spectral gap** if there are two operators $E, D : \mathscr{B} \to \mathscr{B}$ and a decomposition $S = \lambda E + D$ with $\lambda \in \mathbb{C}$ satisfying $|\lambda| = ||S||$ such that the following properties are satisfied:*

(i) The operator E is a projection onto its image and $\dim(\mathrm{Im}(E)) = 1$.

(ii) $E \circ D = D \circ E = 0$.

(iii) $\rho(D) < ||S||$.

In the literature, an operator is referred to as having a spectral gap if there is an isolated eigenvalue λ of maximal modulus and the rest of the spectrum lies in a ball of radius $|\lambda| - \varepsilon$ for some $\varepsilon > 0$. The definition of a strong spectral gap implies the latter while requiring the above further conditions, which explains this choice of terminology.

The aim of this section is to prove Corollary 4.2.4 below on quasicompact operators which are positive in the sense of Banach lattices. We refer to the book [Sch74] as a reference on Banach lattices. For the convenience of the reader, we recall the definition of a Banach lattice from [Sch74] and a few further definitions. In this thesis, we will work with the Banach lattice $L^2(\Omega)$ (or $L^2(K)$) endowed with the partial order defined for $f, g \in L^2(\Omega)$ as $f \leq g$ if and only if $f(\omega) \leq g(\omega)$ for all $\omega \in \Omega$.

Definition 4.2.2. *A Banach space $\mathscr{B}$ with norm $||\circ||$ is called a* ***Banach lattice*** *if it is equipped with a partial order $\leq$ satisfying the following properties:*

(i) (Ordered vector space) For two elements $x, y \in \mathscr{B}$ with $x \leq y$ we have for all $z \in \mathscr{B}$ and $\lambda \in \mathbb{R}_{>0}$ that

$$x + z \leq y + z \qquad \text{and} \qquad \lambda x \leq \lambda y.$$

(ii) (Vector lattice) For two elements $x, y \in \mathscr{B}$ the supremum $\sup\{x, y\}$ (resp. infimum $\inf\{x, y\}$) is the, if it exists, unique element $z \in \mathscr{B}$ with $z \geq x$ and $z \geq y$ (resp. $z \leq x$ and $z \leq y$) that satisfies for every further element $z' \in \mathscr{B}$ with $z' \geq x$ and $z' \geq y$ (resp. $z' \leq x$ and $z' \leq y$) that $z' \geq z$ (resp. $z' \leq z$). We assume that $\mathscr{B}$ is closed under taking suprema and infima of two elements, that is, if $x, y \in \mathscr{B}$ then $\sup\{x, y\}$ and $\inf\{x, y\}$ exist in $\mathscr{B}$.

(iii) (Monotonicity) For $x \in \mathscr{B}$ denote $|x| = \sup\{x, -x\}$. If two elements $x, y \in \mathscr{B}$ satisfy $|x| \leq |y|$, then we have $||x|| \leq ||y||$.

For a Banach lattice $\mathscr{B}$ denote by $\mathscr{B}_+ = \{x \in \mathscr{B} : x \geq 0\}$ the set of positive elements. We note $x \geq y$ whenever $x - y \in \mathscr{B}_+$ and further write $x > y$ if and only if $x \geq y$ and $x \neq y$. We say that a bounded operator $A : \mathscr{B} \to \mathscr{B}$ is **positive** if $A(\mathscr{B}_+) \subset \mathscr{B}_+$, in notation $A \geq 0$. We write $A > 0$ if $Ax > 0$ for $x > 0$.

We furthermore say that the operator $A : \mathscr{B} \to \mathscr{B}$ has a **strictly positive invariant form** if there is a linear form $\eta : \mathscr{B} \to \mathbb{R}$ that maps vectors > 0 to real numbers > 0 and that is invariant under A, i.e. $\eta \circ A = \eta$.

For an element $u \in \mathscr{B}_+$ we denote by

$$I_u = \{x \in \mathscr{B} : 0 \leq |x| \leq \lambda u \text{ for some } \lambda > 0\}$$

the principal ideal generated by u, where as above we write $|x| = \sup\{x, -x\}$. The element u is called **quasi-interior** if I_u is dense in $\mathscr{B}$.

A subspace I of $\mathscr{B}$ is called an **ideal** if $I_u \subset I$ for all $u \in I$. An operator $A : \mathscr{B} \to \mathscr{B}$ is referred to as **irreducible** if the only A-invariant ideals are the trivial ideals $\{0\}$ and $\mathscr{B}$.

The **resolvent** of a bounded operator A is defined as

$$R(\lambda, A) = (\lambda I - A)^{-1},$$

which by [DS58, VII Lemma 3.2] is an analytic collection of operators well-defined on the complement of the spectrum of A. A complex number λ_0 in the spectrum of

A is called a **pole** of the resolvent of A if there is $k \in \mathbb{Z}_{>0}$ and $\varepsilon > 0$ such that for all $\lambda \in \mathbb{C}$ with $|\lambda - \lambda_0| < \varepsilon$ we have that

$$R(\lambda, A) = \sum_{n=-k}^{\infty} B_n (\lambda - \lambda_0)^n$$

for $\{B_n\}_{n=-k}^{\infty}$ a collection of bounded operators from $\mathscr{B} \to \mathscr{B}$. The number k is the **order** of the pole.

The following is the main result on Banach lattices that we use in this thesis, which is a generalisation of the Perron-Frobenius theorem to Banach spaces.

Theorem 4.2.3. *(Corollary to Theorem V 5.2 of* [Sch74]*) Let $\mathscr{B}$ be a Banach lattice and let $A : \mathscr{B} \to \mathscr{B}$ be a positive irreducible bounded operator > 0 satisfying $\rho(A) = 1$ and such that 1 is a pole of the resolvent of A. Then the following properties hold:*

(i) 1 is an eigenvalue. The eigenspace of 1 is one-dimensional and spanned by a quasi-interior element of $\mathscr{B}_+$.

(ii) Every eigenvalue λ of A with $|\lambda| = 1$ is a root of unity and has a one dimensional eigenspace. Moreover, the latter set of eigenvalues form a group.

(iii) 1 is the unique eigenvalue of A with a positive eigenvector.

Using Theorem 4.2.3, combinded with basic properties of quasicompact operators, we can draw the following corollary.

Corollary 4.2.4. *Let $\mathscr{B}$ be a Banach lattice and let $A : \mathscr{B} \to \mathscr{B}$ be a positive quasicompact bounded operator > 0 and assume that A^n is irreducible for every $n \geq 1$. Then A has strong spectral gap.*

Proof. Without loss of generality, we may replace A by $A/\rho(A)$ and assume that $\rho(A) = 1$. Therefore, since A is positive, by [Sch74, Proposition V 4.1] it follows that 1 is an eigenvalue of A. Moreover, by [DS58, VII 8.2], since A is quasicompact, the resolvent $R(\lambda, A)$ has a pole at 1. Therefore, Theorem 4.2.3 applies and as is shown in the proof of the Corollary to Theorem V 5.2 of [Sch74] we moreover have that

$$E = \lim_{\lambda \to 1} (\lambda - 1) R(\lambda, A)$$

is a strictly positive projection of rank 1.

Set $D = A - E$. Then $E \circ D = D \circ E = 0$ as A commutes with E and we claim that $\rho(D) < 1$, which follows if we show that 1 is the unique eigenvalue of A on the circle of radius 1. To show the latter, if λ is an eigenvalue of A with $|\lambda| = 1$ and eigenvector v_λ, then by Theorem 4.2.3 (ii) λ is a root of unity and hence $T^n v_\lambda = v_\lambda$ for some $n > 0$. Therefore by Theorem 4.2.3 (i) applied to T^n, it follows that $v_\lambda \geq 0$. Finally, by Theorem 4.2.3 (iii) applied to T, the vector v_λ must be the unique positive eigenvector of T and hence $\lambda = 1$. □

We return to the operators S_0 and S_0^+ defined in (3.1.19).

Lemma 4.2.5. *Let G be a connected semisimple Lie group with finite center and let μ be a non-degenerate probability measure on G. Then S_0 and S_0^+ are positive bounded operators and S_0^n and $(S_0^+)^n$ are irreducible for all $n \geq 1$.*

Proof. We show that S_0 is irreducible and the same argument will apply to S_0^n and $(S_0^+)^n$ for all $n \geq 1$ since G is connected. By III Proposition 8.3 of [Sch74], it suffices to show for any $\varphi_1, \varphi_2 \in L^2(\Omega)$ with $\varphi_1 > 0$ and $\varphi_2 > 0$ that $\langle S_0^\ell \varphi_1, \varphi_2\rangle$ is > 0 for some $\ell \geq 1$. Indeed, we may reduce to the case where $\varphi_1 = 1_{U_1}$ and $\varphi_2 = 1_{U_2}$ for U_1 and U_2 two sets of positive measure. Using that the support of μ generates a dense subgroup, we may choose ℓ large enough such that the support of $S_0^\ell 1_{U_1}$ has measure larger that $1 - m_\Omega(U_2)/2$ and therefore $\langle S_0^\ell 1_{U_1}, 1_{U_2}\rangle > 0$. □

4.3 Preliminaries on Representation Theory of Compact Lie Groups

Recall the notation introduced in section 3.1.

For $\gamma \in \overline{C} \cap I^*$, by Schur's Lemma, the operator $\pi_\gamma(\triangle)$ acts as a scalar. For functions on K, the operator $\lambda_G(\triangle)$ can be understood as the Laplacian. Therefore (3.1.3) is a decomposition into eigenfunctions of the Laplacian and on M_γ the Laplacian has eigenvalue $\lambda_\gamma = \pi_\gamma(\triangle)$.

Lemma 4.3.1. *For $\gamma \in \overline{C} \cap I^*$ denote $d_\gamma := \dim \pi_\gamma$ and $\lambda_\gamma := \pi_\gamma(\triangle)$. Then for γ large enough it holds that $\lambda_\gamma \asymp ||\gamma||^2$ and $d_\gamma \ll ||\gamma||^{|R^+|}$. Moreover, assuming that K is semisimple, $||\gamma|| \ll d_\gamma$.*

Proof. By Lemma 10.6 of [Hal15], $\lambda_\gamma := \pi_\gamma(\triangle) = \langle \gamma + \rho, \gamma + \rho\rangle - \langle \rho, \rho\rangle$, where $\rho = \frac{1}{2}\sum_{\alpha \in R^+} \alpha$ is the sum of positive half roots (notice that the multiplicity of each root is one cf. Theorem 7.23 of [Hal15]). This easily implies $\lambda_\gamma \asymp ||\gamma||^2$. The upper bound on d_γ follows by Weyl's dimension formula:

$$d_\gamma = \prod_{\alpha \in R^+} \frac{\langle \alpha, \gamma + \rho\rangle}{\langle \alpha, \rho\rangle} \leq \left(\prod_{\alpha \in R^+} \frac{||\alpha||}{|\langle \alpha, \rho\rangle|}\right) ||\gamma + \rho||^{|R^+|} \ll_G ||\gamma||^{|R^+|}$$

for $||\gamma||$ large enough. For the lower bound we recall that in [dS13], also using the Weyl dimension formula, it is proved that $||\gamma||^{|R^+|-p} \ll d_\gamma$, where p is the number of maximal elements of R^+ that are contained in one hyperplane. If K is semisimple, the roots span the vector space $\mathbf{t}^*$ and therefore $(|R^+|-p) \geq 1$. □

Recall the Sobolev spaces defined in (3.1.7). We deduce a condition for a function being in $C^m(K)$ under an assumption on the decay of $||\pi_\ell \varphi||_2$.

Lemma 4.3.2. *Let $m \in \mathbb{Z}_{\geq 0}$, $s > m + \frac{1}{2}\dim K$ and let $\varphi \in L^2(K)$. Assume that for all $\ell \in \mathbb{Z}_{\geq 0}$ large enough,*

$$||\pi_\ell \varphi||_2 \leq 2^{-(s+1)\ell}.$$

Then $\varphi \in H^s(K) \subset C^m(K)$.

Proof. If $\varphi = \sum_{\gamma \in \overline{C} \cap I^*} \varphi_\gamma$, by the assumption for large enough ℓ, $2^{2s\ell}||\pi_\ell \varphi||_2^2 = 2^{2s\ell} \sum_{2^{\ell-1} \leq ||\gamma|| < 2^\ell} ||\varphi_\gamma||_2^2 \leq 2^{-2\ell}$ and hence using Lemma 4.3.1,

$$\sum_{\gamma \in \overline{C} \cap I^*} \lambda_\gamma^s ||\varphi_\gamma||_2^2 \ll \sum_{\ell \geq 0} 2^{2s\ell} \sum_{2^{\ell-1} \leq ||\gamma|| < 2^\ell} ||\varphi_\gamma||_2^2 \ll \sum_{\ell \geq 0} 2^{-2\ell} < \infty,$$

showing that $\varphi \in H^s(K)$. The inclusion $H^s(K) \subset C^m(K)$ follows from the Sobolev embedding theorem (cf. [Aub98] Theorem 2.10). □

4.4 Flattening of μ^{*n}

In this section we state strong flattening results from [BISG17] for (c_1, c_2, ε)-Diophantine measures. To introduce notation, denote

$$P_\delta = \frac{1_{B_\delta}}{m_G(B_\delta)}$$

and for a measure ν and $g \in G$, we note that $(\nu * P_\delta)(g) = \frac{\nu(B_\delta(g))}{m_G(B_\delta)}$. We also use the notation $\nu_\delta = (\nu)_\delta = \nu * P_\delta$.

We first relate the condition that a measure is (c_1, c_2, ε)-Diophantine to the assumptions of several theorems in [BISG17].

Lemma 4.4.1. *Let $c_1, c_2, \varepsilon > 0$ and let μ be a probability measure on G satisfying $\mathrm{supp}(\mu) \subset B_\varepsilon$. Then μ is (c_1, c_2, ε)-Diophantine if and only if for δ small enough and $n = \frac{\log \frac{1}{\delta}}{c_1 \log \frac{1}{\varepsilon}}$,*

$$\sup_{H < G} \mu^{*n}(B_\delta(H)) \leq \delta^{\frac{c_2}{c_1}},$$

where $B_\delta(H) = \{g \in G : d(g, H) < \delta\}$ and the supremum is taken over all closed subgroups of G.

Proof. This follows from the fact that μ is (c_1, c_2)-Diophantine if and only if $\sup_{H<G} \mu^{*n}(B_\delta(H)) \leq \delta^{\frac{c_2}{c_1}}$ for $n = \frac{1}{c_1}\log\frac{1}{\delta}$. □

We state Corollary 4.2 from [BISG17].

Theorem 4.4.2. *(Flattening Lemma, Corollary 4.2 of* [BISG17]*) Let G be a connected simple Lie group with finite center. Let $c_1, c_2 > 0$. Then for every $\gamma > 0$ there is $\varepsilon_0 = \varepsilon_0(c_1, c_2, \gamma) > 0$ and $C_0 = C_0(c_1, c_2, \gamma) > 0$ such that the following holds:*

If $0 < \varepsilon < \varepsilon_0$ and μ is a symmetric and (c_1, c_2, ε)-Diophantine probability measure on G, then for $\delta > 0$ small enough,

$$||(\mu^{*n})_\delta||_2 \leq \delta^{-\gamma} \qquad \textit{for any integer} \qquad n \geq C_0 \frac{\log\frac{1}{\delta}}{\log\frac{1}{\varepsilon}}.$$

4.5 Estimate of Averages of Matrix Coefficients for Oscillating Functions

In this subsection we prove the following proposition, which is inspired by [LV16]. We denote $B_R = \{g \in G \,:\, d(g, e) < R\}$.

Proposition 4.5.1. *Let G be a non-compact semisimple Lie group with finite center and maximal compact subgroup K. Recall the Littlewood-Paley decomposition* (3.1.6) *$L^2(K) = \bigoplus_{\ell \geq 0} V_\ell$ and assume further that K is a semisimle Lie group. Then for any $r \in \mathbb{R}$ and $\ell \in \mathbb{Z}_{\geq 1}$, for $\varphi_1, \varphi_2 \in V_\ell \subset L^2(K)$,*

$$\frac{1}{m_G(B_R)} \int_{B_R} |\langle \rho_r^+(g)\varphi_1, \varphi_2\rangle| \, dm_G(g) \ll 2^{-\ell/2} ||\varphi_1||_2 ||\varphi_2||_2,$$

where the representation ρ_r^+ is defined in (3.1.13)*.*

We recall the following lemma from [LV16].

Lemma 4.5.2. *(Proposition 5.1 of* [LV16]*) Let $(\pi, \mathscr{H})$ be a unitary representation of a compact group K and let D be the minimum of the dimension of all irreducible representations contained in π. Then for any vectors $u, v \in \mathscr{H}$,*

$$\left(\int |\langle \pi(g)u, v\rangle|^2 \, dm_K(k)\right)^{1/2} \leq \frac{||u|| \, ||v||}{D^{1/2}}.$$

If π is irreducible, then Lemma 4.5.2 follows from Schur's Lemma (see [Kna02] section I.5). For the general case one decomposes π as a direct sum of irreducible representations.

Proof. (of Proposition 4.5.1) Let $B'_R = B_R \cdot K$. By left invariance of the metric, it follows that $B_{R'} \subset B_{R+C}$ for C an absolute constant and therefore $m_G(B_{R'}) \ll m_G(B_R)$. Using Cauchy-Schwarz and that for $k \in K$ the operator $\rho_r^+(k)$ acts as the regular representation, it follows by Lemma 4.5.2,

$$\begin{aligned}
\int_{B_R} |\langle \rho_r^+(g)\varphi_1, \varphi_2\rangle| \, dm_G(g) &\leq \int_{B'_R} |\langle \rho_r^+(g)\varphi_1, \varphi_2\rangle| \, dm_G(g) \\
&= \int_{B_{R'}} \left(\int_K |\langle \rho_r^+(k)\varphi_1, \rho_r^+(g^{-1})\varphi_2\rangle| \, dm_K(k) \right) dm_G(g) \\
&\leq \int_{B_{R'}} \left(\int_K |\langle \rho_r^+(k)\varphi_1, \rho_r^+(g^{-1})\varphi_2\rangle|^2 \, dm_K(k) \right)^{1/2} dm_G(g) \\
&\leq m_G(B_{R'}) \left(\min_{2^{\ell-1} \leq ||\gamma|| < 2^\ell} d_\gamma \right)^{-1/2} ||\varphi_1|| \, ||\varphi_2|| \\
&\ll m_G(B_R) 2^{-\ell/2} ||\varphi_1|| \, ||\varphi_2||,
\end{aligned}$$

having used in the last line that $||\gamma|| < d_\gamma$ from Lemma 4.3.1 under the assumption that K is semisimple. □

Chapter 5

Proof of Local Limit Theorem

We fix throughout a non-compact semisimple Lie group G with finite center. In this section we prove Theorem 2.0.1, Theorem 2.0.2 and Theorem 2.0.3. The reader may recall the outline given in section 3.2.

In section 5.1 we prove the necessary spectral properties for the operators S_r. Then in section 5.2 we prove the claimed properties of the limit measure as well as deduce (3.2.6). In section 5.3 we deal with the high frequency term (3.2.2) while in section 5.4 we establish most of the necessary results to deal with the low frequency term (3.2.3). The proof of Theorem 2.0.2 and Theorem 2.0.3 is then completed in section 5.5, while Theorem 2.0.1 is deduced in section 5.6.

5.1 Spectral Properties of S_r

In this section we discuss spectral results for the operators S_0 and S_r and the function $r \mapsto \rho(S_r)$ under the assumption that S_0 is quasicompact and using the results developed in section 4.1 and section 4.2. Notice that if μ is non-degenerate and S_0 is quasicompact, then by Lemma 4.2.5 and Corollary 4.2.4 the operator S_0 has strong spectral gap.

Before stating the first lemma, we mention that $|S_r\eta| \leq S_0|\eta|$ for all $r \in \mathfrak{a}^*$ and $\eta \in L^2(\Omega)$, which implies $\rho(S_r) \leq ||S_0||$. Lemma 5.1.1 is concerned with improving the latter inequality to $\rho(S_r) < ||S_0||$ under suitable assumptions on μ.

Lemma 5.1.1. *Let μ be a non-degenerate probability measure and assume that S_0 is quasicompact. Then for any non-zero $r \in \mathfrak{a}^*$,*

$$\rho(S_r) < \rho(S_0) = ||S_0||. \tag{5.1.1}$$

Moreover, for any $c_2 > c_1 > 0$ and n large enough depending on c_1 and c_2,

$$\sup_{c_1 \leq |r| \leq c_2} ||S_r^n||^{\frac{1}{n}} < ||S_0||. \tag{5.1.2}$$

Proof. To prove (5.1.1), we follow ideas from the proof of Theorem 3.9 of [CG13]. Fix a non-zero $r \in \mathfrak{a}^*$. We assume for a contradiction that $\rho(S_r) = \rho(S_0)$ and therefore there is $\lambda = e^{i\gamma}\rho(S_0) \in \operatorname{spec}(S_r)$ for $\gamma \in \mathbb{R}$. Then (cf. section 12.1 of [EW17]) either λ is in the discrete spectrum or in the approximate spectrum, i.e. there is a sequence $\eta_\ell \in \ker(S_r - \lambda \cdot \mathrm{Id})^\perp$ with $||\eta_\ell|| = 1$ and

$$\lim_{\ell \to \infty} ||S_r \eta_\ell - \lambda \eta_\ell|| = 0. \tag{5.1.3}$$

Note that as S_0 is quasicompact, $\rho(S_0) = ||S_0||$. We first treat the case where λ is in the discrete spectrum, i.e. that there exists $\eta \in L^2(\Omega)$ such that $S_r\eta = \lambda\eta$. Then $||S_0||\,|\eta| = |S_r\eta| \leq S_0|\eta|$ and thus $||\,S_0|\eta|\,|| = ||S_0||\,||\eta||$. Denote by η_0 the $||S_0||$-eigenfunction of S_0 with unit norm. As S_0 has strong spectral gap (by Lemma 4.2.5 and Corollary 4.2.4), it follows that $\eta(\omega) = e^{i\theta(\omega)}\eta_0(\omega)$, for $\theta : \Omega \to \mathbb{R}$ a measurable function and $\omega \in \Omega$.

Then for almost all $\omega \in \Omega$ and $n \geq 1$,

$$\begin{aligned}
\int e^{-(\delta+ir)H(g^{-1}\omega)+i\theta(g^{-1}\omega)}\eta_0(g^{-1}\omega)\, d\mu^{*n}(g) &= (S_r^n\eta)(\omega) \\
&= \lambda^n \eta(\omega) \\
&= e^{in\gamma}||S_0||^n e^{i\theta(\omega)}\eta_0(\omega) \\
&= e^{i(n\gamma+\theta(\omega))}\int e^{-\delta H(g^{-1}\omega)}\eta_0(g^{-1}\omega)\, d\mu^{*n}(g).
\end{aligned}$$

As η_0 is a quasi-interior element by Theorem 4.2.3, it must hold that $\eta_0(\omega) > 0$ for almost all $\omega \in \Omega$. Hence for almost all $\omega \in \Omega$ and $g \in \operatorname{supp}(\mu^{*n})$,

$$e^{-i(rH(g^{-1}\omega)-\theta(g^{-1}\omega)+\theta(\omega)+n\gamma)} = 1.$$

If $r \neq 0$, for a fixed $\omega \in \Omega$ and $n \geq 1$, we can choose $h_n \in G$ such that $e^{-irH(h_n^{-1}\omega)} = e^{i(n\gamma+\pi)}$ yet $e^{i(\theta(h_n^{-1}\omega)-\theta(\omega))} = 1$. Indeed, for a representative $\omega = kM$ for $k \in K$, we may choose $h_n = ka_nk^{-1}$ for an element $a_n \in A$ satisfying $e^{-irH(a_n^{-1})} = e^{i(n\gamma+\pi)}$ as then $H(h_n^{-1}k) = H(a_n^{-1})$ and $\theta(h^{-1}\omega) = \theta(\omega)$. We may choose the h_n within a bounded region of G and therefore upon replacing h_n with a subsequence we may assume that h_n converges to some element $h \in G$. Since μ is non-degenerate we can find some n and $g \in \operatorname{supp}(\mu^{*n})$ such that g becomes arbitrarily close to h and hence for n large enough also to h_n. This is a contradiction.

It remains to assume that λ is in the approximate spectrum. Let η_ℓ as in (5.1.3). Since $\langle S_r\eta_\ell, \lambda\eta_\ell\rangle = \langle S_r\eta_\ell - \lambda\eta_\ell, \lambda\eta_\ell\rangle + ||S_0||^2$, it follows that $\langle S_r\eta_\ell, \lambda\eta_\ell\rangle \xrightarrow{\ell\to\infty} ||S_0||^2$ and furthermore exploiting $|\langle S_r\eta_\ell, \lambda\eta_\ell\rangle| \leq \langle S_0|\eta_\ell|, ||S_0||\,|\eta_\ell|\rangle$ one concludes

$$\lim_{\ell\to\infty}\langle S_0|\eta_\ell|, ||S_0||\,|\eta_\ell|\rangle = ||S_0||^2$$

and hence $||S_0|\eta_\ell| - ||S_0||\,|\eta_\ell|\,||^2 \leq 2||S_0||^2 - 2\langle S_0|\eta_\ell|, ||S_0||\,|\eta_\ell|\rangle \xrightarrow{\ell\to\infty} 0$.

Denote $\psi_\ell = |\eta_\ell| - \langle|\eta_\ell|, \eta_0\rangle\eta_0 \in \langle\eta_0\rangle^\perp \subset L^2(\Omega)$. Then it holds that

$$||(S_0 - ||S_0||)\psi_\ell||_2 = ||S_0|\eta_\ell| - ||S_0||\,|\eta_\ell|\,||_2 \xrightarrow{\ell\to\infty} 0.$$

Since ψ_ℓ in $\langle\eta_0\rangle^\perp$ and $S_0 - ||S_0||$ is invertible on $\langle\eta_0\rangle^\perp$ it follows that $||\psi_\ell||_2 \to 0$. Notice that $||\psi_\ell||_2^2 = 1 - \langle|\eta_\ell|, \eta_0\rangle^2$ and hence $\langle|\eta_\ell|, \eta_0\rangle \to 1$ and further $||\,|\eta_\ell| - \eta_0||_2 \to 0$. Upon replacing ℓ by a subsequence, we can assume that $|\eta_\ell|$ converges pointwise to η_0 almost everywhere.

We further note that for all $n \geq 1$, $\langle S_r^n\eta_\ell, \lambda^n\eta_\ell\rangle \to ||S_0||^{2n}$ as $\ell \to \infty$. Indeed this follows by induction as

$$\begin{aligned}&\langle S_r^n\eta_\ell - S_r^{n-1}\lambda\eta_\ell + S_r^{n-1}\lambda\eta_\ell, \lambda^n\eta_\ell\rangle\\ &= \langle S_r^{n-1}(S_r\eta_\ell - \lambda\eta_\ell), \lambda^n\eta_\ell\rangle + ||S_0||^2\langle S_r^{n-1}\eta_\ell, \lambda^{n-1}\eta_\ell\rangle \to ||S_0||^{2n}.\end{aligned}$$

Write $\lambda = e^{i\gamma}||S_0||$ and $\eta_\ell(\omega) = e^{i\theta_\ell(\omega)}|\eta_\ell|(\omega)$ for $\theta_\ell : \Omega \to \mathbb{R}$ a measurable function and $\omega \in \Omega$. Notice that $\langle S_r^n\eta_\ell, \lambda^n\eta_\ell\rangle$ equals

$$\int\int e^{-(\delta+ir)H(g^{-1}\omega)+i(\theta_\ell(g^{-1}\omega)-\theta_\ell(\omega)-n\gamma)}||S_0||^n\,|\eta_\ell|(g^{-1}\omega)|\eta_\ell|(\omega)\,d\mu^{*n}(g)dm_\Omega(\omega)$$

and on the other hand

$$\langle S_0^n\eta_0, ||S_0||^n\eta_0\rangle = \int\int e^{-\delta H(g^{-1}\omega)}||S_0||^n\eta_0(g^{-1}\omega)\eta_0(\omega)\,d\mu^{*n}(g)dm_\Omega(\omega).$$

As $\langle S_r^n\eta_\ell, \lambda^n\eta_\ell\rangle \xrightarrow{\ell\to\infty} ||S_0||^{2n} = \langle S_0^n\eta_0, ||S_0||^n\eta_0\rangle$ and since almost surely $|\eta_\ell| \to \eta_0$, we conclude that for almost all $g \in \operatorname{supp}(\mu^{*n})$ and $\omega \in \Omega$,

$$\lim_{\ell\to\infty} e^{i(rH(g^{-1}\omega)-\theta_\ell(g^{-1}\omega)+\theta_\ell(\omega)+\gamma)} = 1.$$

This leads to a contradiction by a similar argument to the case of the discrete spectrum.

To prove (5.1.2), we notice that for an operator T on a Hilbert space $\mathscr{H}$ with $||T|| \leq 1$, the value of $||T^n||^{\frac{1}{n}}$ for a given n controls $||T^k||^{\frac{1}{k}}$ for any $k \geq n$. Indeed (cf. [Rem]) if $k = \ell n + j$ for $0 \leq j \leq n-1$ then it holds that

$$||T^k||^{\frac{1}{k}} \leq (||T^{\ell n}||^{\frac{1}{\ell n}})^{\frac{\ell n}{k}}||T||^{\frac{j}{k}} \leq (||T^n||^{\frac{1}{n}})^{1-\frac{j}{k}}||T||^{\frac{j}{k}}. \quad (5.1.4)$$

Therefore for k large enough in terms of n, $||T^k||^{\frac{1}{k}}$ is at most slightly larger than $||T^n||^{\frac{1}{n}}$. Assume now for a contradiction that (5.1.2) does not hold. Then there is a sequence $(n_i)_{i\geq 1}$ with $n_i \to \infty$ and for each i there is r_i with $||S_{r_i}^{n_i}||^{\frac{1}{n_i}} = ||S_0||$. As the set $\{c_1 \leq |r| \leq c_2\}$ is compact, we may choose a subsequence of the i such that r_i converges to $r \in \mathfrak{a}^*$ with $c_1 \leq |r| \leq c_2$. We arrive at a contradiction as by (5.1.4), $||S_{r_i}^{n_i}||^{\frac{1}{n_i}}$ is at most marginally larger than $||S_r^n||^{\frac{1}{n}}$ for r_i close enough to r. Indeed, choose $\varepsilon > 0$ small enough such that $\rho(S_r) + 3\varepsilon < ||S_0||$ and fix n large enough such that $||S_r^n||^{\frac{1}{n}} \leq \rho(S_r)+\varepsilon$. Then for r_i close enough to r, $||S_{r_i}^n||^{\frac{1}{n}} \leq \rho(S_r)+2\varepsilon$ and hence by (5.1.4), choosing i sufficiently large, $||S_{r_i}^{n_i}||^{\frac{1}{n_i}} \leq \rho(S_r)+3\varepsilon < ||S_0||$, a contradiction to the assumption. □

Proposition 5.1.2. *Let μ be a non-degenerate probability measure with finite second moment and assume that S_0 is quasicompact. Then there is $\delta_0 = \delta_0(\mu) > 0$ such that for any $r \in \mathfrak{a}^*$ with $|r| \leq \delta_0$ the operators S_r and S_r^* have strong spectral gap.*

More precisely there is $0 < \delta_0 < 1$ small enough satisfying the following properties. For $|r| \leq \delta_0$ we can write

$$S_r = \lambda(r)E_r + D_r \qquad \text{and} \qquad S_r^* = \overline{\lambda(r)}E_r^* + D_r^* \tag{5.1.5}$$

where $\lambda(r)$, E_r and D_r and equally $\overline{\lambda(r)}, E_r^*$ and D_r^* satisfy the assumptions of Definition 4.2.1, and the following properties hold:

(i) $\sup_{|r|\leq\delta_0} ||D_r|| \leq (1-c)||S_0||$ *for $c = c(\mu) > 0$.*

(ii) $||E_r - E_0|| \ll_\mu |r|^2$ *and* $||E_r^* - E_0^*|| \ll_\mu |r|^2$ *for $|r| \leq \delta_0$.*

(iii) Let η_r be the unique $\lambda(r)$-eigenfunction of S_r with unit norm. Then for small enough r there exists a unique $\overline{\lambda(r)}$-eigenfunction η_r' of S_r^* satisfying $\langle \eta_r', \eta_r\rangle = 1$. Additionally, for $\varphi \in L^2(\Omega)$,

$$E_r\varphi = \langle \varphi, \eta_r'\rangle \eta_r.$$

(iv) Moreover,

$$||\eta_r - \eta_0||_2 \ll_\mu |r|^2, \qquad \text{and} \qquad ||\eta_r' - \eta_0'|| \ll_\mu |r|^2$$

for $|r| \leq \delta_0$.

Proof. As μ has finite second moment, the directional derivatives of second order of the family of operators S_r and S_r^* exist. Therefore the function $r \mapsto ||S_r - S_0||$ is C^2. Since $\overline{S_r\varphi} = S_{-r}\overline{\varphi}$ for $\varphi \in L^2(\Omega)$, it follows by Taylor's theorem that $||S_r - S_0|| \ll_\mu$

$|r|^2$ for small r. By Corollary 4.1.2 and Corollary 4.2.4, S_0 has strong spectral gap and S_r is quasicompact for small r. Equally by Lemma 4.1.1 (ii) and since $S_0^* = \int \rho_0(g^{-1})\, d\mu(g)$ is a positive operator too, it follows that S_0^* has strong spectral gap and S_r^* is quasicompact for small r.

We show that there is $\delta_0, c > 0$ small enough such that for $|r| \leq \delta_0$ and two orthogonal functions of unit norm $\varphi_1, \varphi_2 \in L^2(\Omega)$ it must hold for either $i = 1$ or $i = 2$ that

$$||S_r\varphi_i||_2 \leq (1-c)||S_0||. \tag{5.1.6}$$

Indeed, assume for a contradiction that (5.1.6) does not hold. Then $||S_0\varphi_i||_2 \geq ||S_r\varphi_i||_2 - ||(S_r - S_0)\varphi_i||_2 \geq (1-c)\lambda(0) + O_\mu(|r|^2) \geq (1-2c)||S_0||$ for r small enough. For c small enough, as S_0 has strong spectral gap and $\langle \varphi_1, \varphi_2\rangle = 0$, the latter is a contradiction.

Therefore we have shown for $|r| \leq \delta_0$ that the $\lambda(r)$-eigenspace of S_r is one dimensional and on its complement the norm of S_r is bounded by $(1-c)||S_0||$. Choose $\delta_0 > 0$ in addition small enough such $||S_0||(1-\frac{c}{2}) < \inf_{|r|\leq\delta_0} \lambda(r)$. Denote by $\gamma_1 : \mathbb{S}^1 \to \mathbb{C}$ a smooth parametrization of the closed circle of radius $||S_0||(1-\frac{c}{2})$ around zero and by $\gamma_2 : \mathbb{S}^1 \to \mathbb{C}$ a smooth parametrization of the circle of radius $\frac{||S_0||c}{2}$ around $||S_0||$. Consider the operators

$$P_r = -\frac{1}{2\pi i}\int_{\gamma_1} R(z, S_r)\, dz, \qquad \text{and} \qquad E_r = -\frac{1}{2\pi i}\int_{\gamma_2} R(z, S_r)\, dz, \tag{5.1.7}$$

for $R(z, S_r) = (S_r - z\cdot\mathrm{Id})^{-1}$ the resolvent of S_r at z. Then by Theorem 6.17 of Chapter 3 in [Kat95], the operators E_r and P_r are commuting projections with $\mathrm{Id} = E_r + P_r$ and where $\ker(P_r) = \mathrm{Im}(E_r)$ is the one dimensional eigenspace of S_r with eigenvalue $\lambda(r)$. By setting $D_r = S_rP_r$, we therefore have shown that $S_r = S_r(E_r + P_r) = \lambda(r)E_r + D_r$ has strong spectral gap and that (i) holds.

We claim that the operators E_r and P_r are also C^2. Indeed by Lemma 3 of Chapter VII.6 of [DS58], it holds that whenever $||S_r - S_0|| < ||R(z, S_0)||^{-1}$, then for any z in the resolvent set of S_0 that z is also in the resolvent set for S_r and that

$$R(z, S_r) = R(z, S_0)\sum_{n=0}^{\infty}(S_r - S_0)^n R(z, S_0)^n.$$

Since S_r is C^2 it therefore follows that for r small enough $R(z, S_r)$ is also C^2 on γ_1 and γ_2. Thus $||P_r - P_0|| \ll_\mu |r|^2$ and $||E_r - E_0|| \ll_\mu |r|^2$ and the claim for E_r^* is established similarly.

To show (iii), first assume that such an η_r' exists. Then as $E_r\varphi = \langle\varphi, \psi\rangle\eta_r$ for some $\psi \in L^2(\Omega)$ with $S_rE_r = E_rS_r$ and $E_r^2 = E_r$ it follows that $S_r^*\psi = \overline{\lambda(r)}\psi$ and that $\langle \eta_r, \psi \rangle = 1$, which implies that $\psi = \eta'_r$. By the above, it follows that there is a unique $\overline{\lambda(r)}$-eigenfunction of S_r^* with unit norm for $|r| \leq \delta_0$, yet we need to show that there exists one with $\langle \eta_r, \eta'_r \rangle = 1$. For $r = 0$ this holds as both eigenfunctions are positive almost surely and for small r we apply (iv) (for η'_r with a fixed norm) to show that there is a $\overline{\lambda(r)}$-eigenfunction η'_r of S_r^* satisfying $\langle \eta_r, \eta'_r \rangle \neq 0$ and therefore upon normalizing η'_r the claim follows.

To conclude, we show (iv) for $||\eta_r - \eta_0||_2$ and note that the same argument applies to $||\eta'_r - \eta'_0||_2$. The claim is deduced from (ii) by noticing that for δ_0 small enough, $\eta_r = \frac{E_r \eta_0}{||E_r \eta_0||}$. Indeed, $||E_r \eta_0|| = |||E_r \eta_0 - \eta_0 + \eta_0|| \geq ||\eta_0|| - ||(E_r - E_0)\eta_0|| > \frac{1}{2}$ for δ small enough. To prove (iv), notice

$$\begin{aligned} ||\eta_r - \eta_0||_2 &\leq ||\tfrac{E_r \eta_0}{||E_r \eta_0||} - \tfrac{E_0 \eta_0}{||E_r \eta_0||}||_2 + ||\tfrac{E_0 \eta_0}{||E_r \eta_0||} - \eta_0||_2 \\ &\ll_\mu ||E_r - E_0|| + |\tfrac{1}{||E_r \eta_0||} - 1| \ll_\mu |r|^2, \end{aligned}$$

using that $1 = ||E_0 \eta_0||$ and $|\frac{1}{||E_r \eta_0||} - 1| \leq |\frac{||E_0 \eta_0|| - ||E_r \eta_0||}{||E_r \eta_0||}| \ll_\mu ||E_r - E_0|| \ll_\mu |r|^2$. □

Proposition 5.1.3. *Let μ be a non-degenerate probability measure with finite second moment and assume that S_0 is quasicompact. Then $\lambda(r)$ is a C^2-function and the Hessian $H_{\lambda,0}$ of λ at 0 is a negative definite sesquilinear form.*

Proof. Using the notation of the proof of Proposition 5.1.2, it holds that $\lambda(r) = \frac{\langle S_r E_r \eta_0, \eta'_0 \rangle}{\langle E_r \eta_0, \eta'_0 \rangle}$ and therefore for r small enough it follows that $r \mapsto \lambda(r)$ is a C^2-function.

For the remainder we follow roughly the proof of Proposition 2.2.7 of [Bou81]. To show that $H_{\lambda,0}$ is negative definite, we fix a non-zero element $r \in \mathfrak{a}^*$ and prove that the function $\xi(t) = \lambda(tr)$ has strictly negative second derivative at zero. Consider the function $h_n(t) = \langle D_{tr}^n \eta_0, \eta'_0 \rangle$. As $D_{tr}^n = (\mathrm{Id} - E_{tr}) D_{tr}^n (\mathrm{Id} - E_{tr})$ it holds that

$$\begin{aligned} |h_n(t)| &= |\langle D_{tr}^n (\mathrm{Id} - E_{tr}) \eta_0, (\mathrm{Id} - E_{tr})^* \eta'_0 \rangle| \\ &\leq ||D_{tr}||^n ||(\mathrm{Id} - E_{tr}) \eta_0|| \, ||(\mathrm{Id} - E_{tr})^* \eta'_0|| \\ &\leq ||D_{tr}||^n ||(E_0 - E_{tr}) \eta_0|| \, ||(E_0 - E_{tr})^* \eta'_0|| \\ &\leq ||D_{tr}||^n ||(E_0 - E_{tr})|| \, ||(E_0^* - E_{tr}^*)|| \ll ||D_{rt}^n|| t^2, \end{aligned}$$

using Proposition 5.1.2 (ii). In particular, using Proposition 5.1.2 (i), $\lambda(0)^{-n} |h_n(t)| \ll_{\mu,r} t^2$ for all $n \geq 1$ and small t and therefore $\lambda(0)^{-n} h''_n(0)$ is bounded for all $n \geq 1$ as otherwise Taylor's theorem would yield a contradiction.

As $\xi(0) = \lambda(0)$ and $\xi'(0) = 0$, it follows that

$$\frac{d^2}{d^2t}\bigg|_{t=0} \big(\lambda(0)^{-n}\langle S^n_{tr}\eta_0, \eta_0'\rangle\big) = \frac{d^2}{d^2t}\bigg|_{t=0} \big(\lambda(0)^{-n}\xi(t)^n\langle E_{tr}\eta_0, \eta_0'\rangle + \lambda(0)^{-n}h_n(t)\big)$$
$$= n\lambda(0)^{-1}\xi''(0) + \frac{d^2}{d^2t}\bigg|_{t=0} \langle E_{tr}\eta_0, \eta_0'\rangle + \lambda(0)^{-n}h_n''(0). \tag{5.1.8}$$

Note that $\frac{d^2}{d^2t}|_{t=0}\langle E_{tr}\eta_0, \eta_0'\rangle$ is also bounded as by Proposition 5.1.2, $|\langle E_{tr}\eta_0, \eta_0'\rangle| \ll_{\mu,r} 1 + t^2$.

We finally consider the functions $f_n(t) = \lambda(0)^{-n}\langle S^n_{tr}\eta_0, \eta_0'\rangle$ for $n \geq 1$. We claim that the function $f_n(t)$ is positive definite. Indeed, for $t_1, \dots, t_m \in \mathbb{R}$ and $\alpha_1, \dots, \alpha_m \in \mathbb{C}$,

$$\lambda(0)^n \sum_{k,\ell} \alpha_k\overline{\alpha_\ell} f_n(t_k - t_\ell) = \sum_{k,\ell} \langle S^n_{(t_k-t_\ell)r}\alpha_k\eta_0, \alpha_\ell\eta_0'\rangle$$
$$= \sum_{k,\ell} \int \alpha_k\overline{\alpha_\ell} e^{-i(t_k-t_\ell)rH(g^{-1}k)} e^{-\delta H(g^{-1}k)}\eta_0(g^{-1}.k)\eta_0'(k)\, d\mu(g)dm_\Omega(k)$$
$$= \int \bigg|\sum_k e^{-it_k rH(g^{-1}k)}\alpha_k\bigg|^2 e^{-\delta H(g^{-1}k)}\eta_0(g^{-1}.k)\eta_0'(k)\, d\mu(g)dm_\Omega(k),$$

which is positive as $\eta_0 \geq 0$ and $\eta_0' \geq 0$. Therefore by Bochner's theorem and since $f_n(0) = 1$ one may expresses f_n as the Fourier transform of a real valued random variable X_n, i.e. $f_n(t) = \int e^{itx}\, d\mu_{X_n}(x)$. Denote by $v_n = -if_n'(0)$ the expected value of X_n and by $\sigma_n^2 = -f_n''(0)$ its variance. For any given $c > 0$ we notice that $P[|X_n - v_n| < c] \to 0$ as $n \to \infty$ since by Lemma 5.1.1 it holds that $f_n(t) \to 0$ for $t \neq 0$ as $n \to \infty$ and therefore μ_n weakly converges to the zero measure. Applying Chebyschev's inequality,

$$1 - \frac{\sigma_n^2}{c^2} \leq 1 - P[|X_n - v_n| \geq c] = P[|X_n - v_n| < c] \to 0$$

and hence $\sigma_n^2 \geq c^2/2$ for any large enough n. Thus $f_n''(0) \to -\infty$ which by (5.1.8) can only happen if $\xi''(0) < 0$. This concludes the proof. □

5.2 The Limit Measure

In this section we establish the claimed properties of the functions $\psi_{\mu,r}$ as stated in (3.2.6). A multiple of $\psi_{\mu,0}$ is the limit function of Theorem 2.0.1.

The main lemma of this section may be viewed as a Lie group analogue of (3.2.7).

Lemma 5.2.1. *Let μ and $\delta_0 \in (0,1)$ be as in Proposition 5.1.2. Denote for $|r| \le \delta_0$ by η_r the unique $\lambda(r)$-eigenfunction of S_r with unit norm and by η'_r the S^*_r-eigenfunction with eigenvalue $\overline{\lambda(r)}$ satisfying $\langle \eta'_r, \eta_r\rangle = 1$. Then the continuous function*

$$\psi_{\mu,r}(g) = \langle \eta_r, \rho_r(g)\eta'_r\rangle \tag{5.2.1}$$

satisfies $\mu * \psi_{\mu,r} = \psi_{\mu,r} * \mu = \lambda(r)\psi_{\mu,r}$. Moreover, for any $f \in \mathscr{S}(X)$ and $h \in G$,

$$\int f \cdot \rho_G(h)\psi_{\mu,r}\, dm_G = \int_\Omega \widehat{f}(r,\omega)(E_r\rho_r(h^{-1})1)(\omega)\, dm_\Omega(\omega), \tag{5.2.2}$$

where ρ_G is the right regular representation of G and we view f as a right K-invariant eigenfunction on G.

Proof. The relation (5.2.2) follows as for $f \in \mathscr{S}(X)$ and $h \in G$,

$$\begin{aligned}\int f \cdot \rho_G(h)\psi_{\mu,r}\, dm_G &= \langle \eta_r, \rho_r(f)\rho_r(h)\eta'_r\rangle \\ &= \langle \eta_r, \rho_r(f)\rho_r(m_K)\rho_r(h)\eta'_r\rangle \\ &= \langle \eta_r, \rho_r(f)\langle \eta'_r, \rho_r(h^{-1})1\rangle 1\rangle \\ &= \langle \langle \rho_r(h^{-1})1, \eta'_r\rangle \eta_r, \rho_r(f)1\rangle \\ &= \int_\Omega \widehat{f}(r,\omega)(E_r\rho_r(h^{-1})1)(\omega)\, dm_\Omega(\omega),\end{aligned}$$

having used in the last line that $\widehat{f}(r,k) = \rho_{-r}(f)(1) = \overline{\rho_r(f)(1)}$.

To show that $\mu * \psi_{\mu,r} = \lambda(r)\psi_{\mu,r}$, we calculate for $g \in G$

$$\begin{aligned}(\mu * \psi_{\mu,r})(g) &= \int \psi_{\mu,r}(h^{-1}g)\, d\mu(h) \\ &= \langle \eta_r, S^*_r\rho_r(g)\eta'_r\rangle \\ &= \langle S_r\eta_r, \rho_r(g)\eta'_r\rangle = \lambda(r)\psi_{\mu,r}(g).\end{aligned}$$

A similar argument shows that $\psi_{\mu,r} * \mu = \lambda(r)\psi_{\mu,r}$. □

For later reference we show the following lemma.

Lemma 5.2.2. *Let μ be a non-degenerate probability measure on G with finite second moment and assume that S_0 is quasicompact. Denote by δ_0 the constant obtained from Proposition 5.1.2. Then for $|r| \le \delta_0$ with δ_0 small enough, and $g \in G$,*

$$|\psi_{\mu,r}(g) - \psi_{\mu,0}(g)| \ll |r|(1 + ||g||). \tag{5.2.3}$$

Moreover, for $|r| \le \delta_0$ and $g \in G$,

$$\left|\frac{\psi_{\mu,r}(g) + \psi_{\mu,-r}(g)}{2} - \psi_{\mu,0}(g)\right| \ll |r|^2(1 + ||g||^2). \tag{5.2.4}$$

Proof. Observe that

$$\begin{aligned}
|\psi_{\mu,r}(g) - \psi_{\mu,0}(g)| &= |\langle \eta_r, \rho_r(g)\eta_r' \rangle - \langle \eta_0, \rho_0(g)\eta_0' \rangle| \\
&= |\langle \rho_r(g^{-1})\eta_r, \eta_r' \rangle - \langle \rho_0(g^{-1})\eta_0, \eta_0' \rangle| \\
&\leq |\langle \rho_r(g^{-1})\eta_r, \eta_r' - \eta_0' \rangle| + |\langle \rho_r(g^{-1})\eta_r - \rho_0(g^{-1})\eta_0, \eta_0' \rangle| \\
&\ll_\mu ||\eta_r' - \eta_0'||_2 + ||(\rho_r(g^{-1}) - \rho_0(g^{-1}))\eta_0||_2.
\end{aligned}$$

Thus in order to prove (5.2.3), by using Proposition 5.1.2 (iii) it suffices to deal with $||(\rho_r(g^{-1}) - \rho_0(g^{-1}))\eta_0||_2$. One calculates that for $g \in G$ and $\omega \in \Omega$,

$$\begin{aligned}
|(\rho_r(g^{-1}) - \rho_0(g^{-1}))\eta_0(\omega)| &= |(e^{-irH(g\omega)} - 1)||e^{-\delta H(g\omega)}\eta_0(g\omega)| \\
&\ll |r| \, ||g|| \, |e^{-\delta H(g\omega)}\eta_0(g\omega)|. \qquad (5.2.5)
\end{aligned}$$

Equation (5.2.3) therefore follows by squaring the latter term, integrating over Ω and using that $||\rho_0(g)\eta_0||_2 = ||\eta_0||_2 = 1$. For (5.2.4) one performs the same calculation and notices that

$$\left| \left(\frac{\rho_r(g) + \rho_{-r}(g)}{2} - \rho_0(g) \right) \eta_0(\omega) \right| = |(\cos(rH(g^{-1}\omega)) - 1)||e^{-\delta H(g^{-1}\omega)}\eta_0(g^{-1}\omega)|.$$

Then (5.2.4) follows by using that $|(\cos(rH(g^{-1}\omega)) - 1)| \ll |r|^2||g||^2$. □

5.3 High Frequency Estimate

For a Schwartz function $f \in \mathscr{S}(X)$, we say that the Fourier transform $\widehat{f} : \mathfrak{a} \times K \to \mathbb{C}$ has compact support if there is $R \geq 0$ such that $\widehat{f}(r, \omega) = 0$ for $r \geq R$ and all $\omega \in \Omega$. In this section with make no notational difference between a function $f \in \mathscr{S}(X)$ and its G-lift. We first prove a preliminary lemma on the Fourier transform.

Lemma 5.3.1. *For $f \in \mathscr{S}(X)$,*

$$||\widehat{f}(r, \cdot)||_{L^2(\Omega)} \leq ||f||_1$$

Proof. We calculate for $r \in \mathfrak{a}$ and $\omega \in \Omega$ that

$$\begin{aligned}
|\widehat{f}(r, \omega)|^2 &= \left| \int_G f(g)(\rho_{-r}(g)1)(\omega) \, dm_G(g) \right|^2 \\
&\leq \left| \int_G |f(g)| \, |(\rho_{-r}(g)1)(\omega)| \, dm_G(g) \right|^2 \\
&\leq \left| \int_G |f(g)| \, |(\rho_0(g)1)(\omega)| \, dm_G(g) \right|^2.
\end{aligned}$$

Set $f_1 = \frac{|f|}{||f||_1}$ so that it follows that

$$|\widehat{f}(r,\omega)|^2 \le ||f||_1^2 \cdot \left| \int_G (\rho_0(g)1)(\omega)\, f_1(g)\, dm_G(g) \right|^2.$$

Recall that if X is a random variable on a probability space then by Jensen's inequality $E[X]^2 \le E[X^2]$. By construction $f_1 dm_G$ is a probability measure and hence it follows that

$$\begin{aligned} |\widehat{f}(r,\omega)|^2 &\le ||f||_1^2 \int_G (\rho_0(g)1)(\omega)^2\, f_1(g)\, dm_G(g) \\ &\le ||f||_1^2 \int_G \frac{d(\alpha_g)_* m_\Omega}{dm_\Omega}(\omega)\, f_1(g)\, dm_G(g) \end{aligned}$$

Thus we conclude that

$$\begin{aligned} ||\widehat{f}(r,\cdot)||_2^2 &\le ||f||_1^2 \int_G \left(\int_\Omega \frac{d(\alpha_g)_* m_\Omega}{dm_\Omega}(\omega)\, dm_\Omega(\omega) \right) f_1(g)\, dm_G(g) \\ &\le ||f||_1^2. \end{aligned}$$

□

Lemma 5.3.2. *Let μ be a non-degenerate probability measure on G assume that S_0 is quasicompact and let $\delta_0 \in (0,1)$ be the constant from Proposition 5.1.2. Let $R \ge 1$ and let $f \in \mathscr{S}(X)$ be a Schwartz function whose Fourier transform satisfies $\widehat{f}(r,\omega) = 0$ for all $|r| \ge R$ and $\omega \in \Omega$. Then there is $c_R = c_R(\mu) > 0$ depending on μ and R such that for $n \ge 1$,*

$$\left| \frac{n^{\ell/2}}{\sigma^n} \int_{|r|\ge\delta_0} \int_\Omega \widehat{f}(r,\omega)\, (S_r^n \rho_r(h_0)1)(\omega)\, dm_\Omega(\omega) d\nu_{\mathrm{sph}}(r) \right| \ll_\mu R^{\dim X} e^{-c_R n} ||f||_1.$$

Proof. Choose R such that $\widehat{f}(r,\omega) = 0$ for $r \ge R$ and $\omega \in \Omega$. Then using Cauchy-Schwarz and Lemma 5.1.1,

$$\begin{aligned} &\left| \frac{n^{\ell/2}}{\sigma^n} \int_{\delta_0\le|r|\le R} \int_\Omega \widehat{f}(r,\omega)\, (S_r^n \rho_r(h_0)1)(\omega)\, dm_\Omega(\omega) d\nu_{\mathrm{sph}}(r) \right| \\ &\le \frac{n^{\ell/2}}{\sigma^n} \int_{\delta_0\le|r|\le R} ||\widehat{f}(r,\cdot)||_{L^2(\Omega)} ||S_r^n \rho_r(h_0)1||_2\, d\nu_{\mathrm{sph}}(r) \\ &\le \frac{n^{\ell/2}}{\sigma^n} \sup_{\delta_0\le|r|\le R} ||S_r^n|| \int_{1\le|r|\le R} ||\widehat{f}(r,\cdot)||_{L^2(\Omega)}\, d\nu_{\mathrm{sph}}(r) \\ &\le e^{-c_R n} \int_{\delta_0\le|r|\le R} ||\widehat{f}(r,\cdot)||_{L^2(\Omega)}\, d\nu_{\mathrm{sph}}(r) \\ &\ll_\mu e^{-c_R n} ||f||_1 \int_{|r|\le R} |c(r)|^{-2}\, dm_{\mathfrak{a}^*}(r) \\ &\ll_\mu e^{-c_R n} ||f||_1 \int_{|r|\le R} (1 + |r|^{\dim N})\, dm_{\mathfrak{a}^*}(r) \ll_\mu R^{\dim X} e^{-c_R n} ||f||_1, \end{aligned}$$

using (5.1.2) in order to choose a constant $c_R > 0$ depending on μ and R such that $\left(\frac{n^{\ell/2}}{\sigma^n}\sup_{\delta_0\leq|r|\leq R}||S_r^n||\right) \leq e^{-c_R n}$ for n large enough and Proposition 7.2 of chapter IV in [Hel84], asserting that $|c(r)|^{-2} \ll 1+|r|^{\dim N}$ for any $r \in \mathfrak{a}^*$ □

Towards proving Theorem 2.0.3, we strengthen Lemma 5.3.2 under strong assumptions on $||S_r||$.

Lemma 5.3.3. *Let μ be a non-degenerate probability measure on G. Assume that S_0 is quasicompact and that $\left(\sup_{|r|\geq 1}||S_r||\right) < ||S_0||$. Let δ_0 be the constant from Proposition 5.1.2. Then for $f \in \mathscr{S}(X)$, $s > \frac{1}{2}\dim X$ and $n \geq 1$,*

$$\left|\frac{n^{\ell/2}}{\sigma^n}\int_{|r|\geq\delta_0}\int_\Omega \widehat{f}(r,\omega)\,(S_r^n\rho_r(h_0)1)(\omega)\,dm_\Omega(\omega)d\nu_{\mathrm{sph}}(r)\right| \ll_{\mu,s} e^{-cn}||f||_{H^s}.$$

Proof. The left hand side of the claimed equation is bounded by

$$\begin{aligned}
&\leq \frac{n^{\ell/2}}{\sigma^n}\int_{|r|\geq\delta_0}||\widehat{f}(r,\cdot)||_{L^2(\Omega)}||S_r^n\rho_r(h_0)1||_2\,d\nu_{\mathrm{sph}}(r)\\
&\leq e^{-cn}\int_{|r|\geq\delta_0}||\widehat{f}(r,\cdot)||_{L^2(\Omega)}|r|^s|r|^{-s}\,d\nu_{\mathrm{sph}}(r)\\
&\leq e^{-cn}\sqrt{\int_{|r|\geq\delta_0}|r|^{-2s}\,d\nu_{\mathrm{sph}}(r)}\sqrt{\int_{|r|\geq\delta_0}||\widehat{f}(r,\cdot)||^2_{L^2(\Omega)}|r|^{2s}\,d\nu_{\mathrm{sph}}(r)}\\
&\ll_{\delta_0,s} e^{-cn}||f||_{H^s},
\end{aligned}$$

for n large enough and choosing s sufficiently large such that $\int_{|r|\geq 1}|r|^{-2s}\,d\nu_{\mathrm{sph}}(r)$ is bounded. Indeed, by Proposition 7.2 of chapter IV in [Hel84], it holds that $|c(r)|^{-2} \ll 1+|r|^{\dim N}$ for any $r \in \mathfrak{a}^*$ and therefore $|c(r)|^{-2} \ll_{\delta_0} |r|^{\dim N}$ for $|r| \geq \delta_0$. Thus $\int_{|r|\geq\delta_0}|r|^{-2s}\,d\nu_{\mathrm{sph}}(r) \ll_{\delta_0} \int_{|r|\geq\delta_0}|r|^{\dim N-2s}\,dm_{\mathfrak{a}^*}(r)$ and the latter term is $< \infty$ whenever $\dim N - 2s < -\dim A$. □

5.4 Low Frequency Estimate

Throughout this section we assume that S_0 is quasicompact and denote by $\delta_0 \in (0,1)$ the constant from Proposition 5.1.2. In this section we deal with the some preliminary estimates for the frequency range $|r| \leq \delta_0$. We recall that by Proposition 5.1.2 for $|r| \leq \delta_0$ we have a decomposition

$$S_r = \lambda(r)E_r + D_r,$$

where E_r and D_r satisfy the properties of Definition 4.2.1. We first show that we can ignore the contribution of D_r.

Lemma 5.4.1. *Let μ be a non-degenerate probability measure on G and assume that S_0 is quasicompact. There exists a constant $c > 0$ depending on μ such that for all $f \in \mathscr{S}(X)$ and $h_0 \in G$,*

$$\left| \frac{n^{\ell/2}}{\sigma^n} \int_{|r|\leq\delta_0} \int_\Omega \widehat{f}(r,\omega)(D_r^n \rho_r(h_0)1)(\omega)\, dm_\Omega(\omega) d\nu_{\mathrm{sph}}(r) \right| \ll ||f||_1 e^{-cn}.$$

Proof. Using Proposition 5.1.2, we deduce $\frac{n^{\ell/2}}{\sigma^n} \sup_{|r|\leq\delta_0} ||D_r^n \rho_r(h_0)1|| \ll e^{-cn}$ for $c > 0$ a constant depending on μ. Using Cauchy-Schwarz the term in question is bounded by

$$\frac{n^{\ell/2}}{\sigma^n} \int_{|r|\leq\delta_0} ||\widehat{f}(r,\cdot)||_{L^2(\Omega)}\, ||D_r^n \rho_r(h_0)1||_2 \, d\nu_{\mathrm{sph}}(r).$$

The lemma follows as $||\widehat{f}(r,\cdot)||_{L^2(\Omega)} \leq ||f||_1$ by Lemma 5.3.1 and by estimating $\int_{|r|\leq\delta_0} 1\, d\nu_{\mathrm{sph}}(r) \ll 1$ since $\delta_0 \leq 1$. □

Therefore, up to an exponential error term, we only need to deal with

$$\frac{n^{\ell/2}}{\sigma^n} \int_{|r|\leq\delta_0} \lambda(r)^n \int_\Omega \widehat{f}(r,\omega)(E_r \rho_r(h_0)1)(\omega)\, dm_\Omega(\omega) d\nu_{\mathrm{sph}}(r). \tag{5.4.1}$$

Recall that $\ell = 2p + d$ for d the rank of G, where the rank is defined as the real dimension of $\mathfrak{a}$. We therefore may rewrite (5.4.1) by replacing r by $\frac{r}{\sqrt{n}}$ as

$$\frac{n^p}{\sigma^n} \int_{|r|\leq\delta_0\sqrt{n}} \lambda(\tfrac{r}{\sqrt{n}})^n |c(\tfrac{r}{\sqrt{n}})|^{-2} \int_\Omega \widehat{f}(\tfrac{r}{\sqrt{n}},\omega)(E_{\frac{r}{\sqrt{n}}} \rho_{\frac{r}{\sqrt{n}}}(h_0)1)(\omega)\, dm_\Omega(\omega) dm_{\mathfrak{a}^*}(r). \tag{5.4.2}$$

Towards proving the local limit theorem, we first replace $\frac{\lambda(r/\sqrt{n})^n}{\sigma^n}$ by a suitable function. Before doing so we give some elementary calculative results.

Lemma 5.4.2. *The following inequalities hold:*

(i) For any $A, B \in \mathbb{R}$,

$$|e^A - e^B| \leq |A - B| \max\{e^A, e^B\}$$

(ii) For any $c > 0, r \neq 0$ and $n \geq 1$,

$$ne^{-cnr^2} \leq \frac{2}{c} e^{-cnr^2/2} r^{-2}.$$

Proof. For the first inequality by assuming without loss of generality that $A \geq B$ we deduce that $|e^A - e^B| \leq e^A|1 - e^{B-A}|$ and hence reduce to showing that $|1 - e^{B-A}| \leq |A - B|$. For this we use that $e^x \geq 1 + x$ and hence as $B - A$ is negative, $|1 - e^{B-A}| = 1 - e^{B-A} \leq -(B - A) = |A - B|$.

For the second inequality we apply the observation that $e^{-x} \leq \frac{1}{x}$ to deduce that $ne^{-cnr^2/2} \leq n\frac{2}{cnr^2} = \frac{2}{cr^2}$ which implies the claim by multiplication with $e^{-cnr^2/2}$. □

As in the proof of Proposition 5.1.3 one shows that $\lambda(r)$ is C^4 if μ has finite fourth moment. Indeed, by conducting a Taylor expansion of λ, for small r,

$$\lambda(r) = \lambda(0) - Q(r, r) + O_G(|r|^4),$$

where $Q(r, r) = -H_{\lambda,0}(r, r)/2$ for $H_{\lambda,0}$ the Hessian of λ at 0. By Proposition 5.1.3 the sesquilinear form Q is positive definite.

Lemma 5.4.3. *Assume that μ has finite fourth moment. There are constants $c_2, c^* > 0$ such that for Q the above positive definite sesquilinear on $\mathfrak{a}$ we have for $|r| \leq \delta_0$,*

$$\left|\frac{\lambda(r)^n}{\sigma^n} - e^{-c_2 n Q(r,r)}\right| \ll_\mu e^{-c^* n|r|^2}|r|^2.$$

In particular, for $|r| \leq \delta_0\sqrt{n}$,

$$\left|\frac{\lambda(r/\sqrt{n})^n}{\sigma^n} - e^{-c_2 Q(r,r)}\right| \ll_\mu n^{-1}e^{-c^*|r|^2}|r|^2.$$

Proof. We may choose for small enough r a constant $c_* > 0$ such that $|\lambda(r)| \leq \lambda(0)(1 - c^*|r|^2)$. Using that $\ln(1 + x) \leq x$, it therefore follows that

$$n\ln(\tfrac{\lambda(r)}{\lambda(0)}) \leq -c^* n|r|^2.$$

Throughout set $c_2 = \frac{1}{\lambda(0)}$ and choose $c^* \leq c_2$. Then

$$\max\{e^{-c_2 nQ(r,r)}, e^{n\ln(\frac{\lambda(r)}{\lambda(0)})}\} \leq e^{-c^* n|r|^2}.$$

Using Lemma 5.4.2 (i) it follows that

$$\begin{aligned}
|\tfrac{\lambda(r)^n}{\lambda(0)^n} - e^{-c_2 nQ(r,r)}| &= |e^{n\ln(\frac{\lambda(r)}{\lambda(0)})} - e^{-c_2 nQ(r,r)}| \\
&\leq \max\{e^{-c_2 nQ(r,r)}, e^{n\ln(\frac{\lambda(r)}{\lambda(0)})}\}|n\ln(\tfrac{\lambda(r)}{\lambda(0)}) + c_2 nQ(r,r)| \\
&\ll e^{-c^* nQ(r,r)}n|r|^4 \\
&\ll e^{-c^* nQ(r,r)}|r|^2,
\end{aligned}$$

by using Lemma 5.4.2 (ii) in the last line by changing the constant c^*. □

Recall by the definition of the c-function:

$$
\begin{aligned}
|c(r)|^{-2} &= \frac{1}{I(\delta)}\left(\prod_{\ell=1}^{p}\left|B\left(\frac{m(r_i)}{2},\frac{i\langle r,r_\ell\rangle}{\langle r_\ell,r_\ell\rangle}\right)\right|\right)^{-2}\cdot\left(\prod_{\ell=p+1}^{k}\left|B\left(\frac{m(r_\ell)}{2},\frac{m(r_\ell/2)}{4}+\frac{i\langle r,r_\ell\rangle}{\langle r_\ell,r_\ell\rangle}\right)\right|\right)^{-2}\\
&= \frac{1}{I(\delta)}\left(\prod_{\ell=1}^{p}\frac{|\Gamma(\frac{m(r_\ell)}{2}+\frac{i\langle r,r_\ell\rangle}{\langle r_\ell,r_\ell\rangle})|^2}{|\Gamma(\frac{m(r_\ell)}{2})|^2|\Gamma(\frac{i\langle r,r_\ell\rangle}{\langle r_\ell,r_\ell\rangle})|^2}\right)\cdot\left(\prod_{\ell=p+1}^{k}\frac{|\Gamma(\frac{m(r_\ell)}{2}+\frac{m(r_\ell/2)}{4}+\frac{i\langle r,r_\ell\rangle}{\langle r_\ell,r_\ell\rangle})|^2}{|\Gamma(\frac{m(r_\ell)}{2})|^2|\Gamma(\frac{m(r_\ell/2)}{4}+\frac{i\langle r,r_\ell\rangle}{\langle r_\ell,r_\ell\rangle})|^2}\right),
\end{aligned}
\tag{5.4.3}
$$

where $B(x,y)=\int_0^1 t^{x-1}(1-t)^{y-1}\,dt$ is the Beta function satisfying $B(x,y)=\frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$.

Lemma 5.4.4. *There is a constant c_G depending only on G such that for $|r|\le\delta_0$,*

$$|c(r)|^{-2}=c_G\prod_{\ell=1}^{p}|\langle r,r_\ell\rangle|^2+O\left(|r|^{2p+2}\right).$$

In particular, for $|r|\le\delta_0\sqrt{n}$,

$$\left|n^p|c(\tfrac{r}{\sqrt{n}})|^{-2}-c_G\prod_{\ell=1}^{p}|\langle r,r_\ell\rangle|^2\right|\ll n^{-1}|r|^{2p+2}.$$

Proof. As the singularities of the Γ function are at $0,-1,-2,\dots$ and $\Gamma(z)$ behaves around 0 like $\frac{1}{z}$, it holds that $|\frac{1}{|\Gamma(ix)|^2}-x^2|\ll x^4$ and $|\,|\Gamma(\frac{n}{2}+ix)|^2-\Gamma(\frac{n}{2})^2|\ll x^2$. Therefore,

$$\left|\frac{|\Gamma(\frac{m(r_\ell)}{2}+\frac{i\langle r,r_\ell\rangle}{\langle r_\ell,r_\ell\rangle})|^2}{|\Gamma(\frac{m(r_\ell)}{2})|^2|\Gamma(\frac{i\langle r,r_\ell\rangle}{\langle r_\ell,r_\ell\rangle})|^2}-\frac{|\Gamma(\frac{m(r_\ell)}{2})|\,|\langle r,r_\ell\rangle|^2}{|\Gamma(\frac{m(r_\ell)}{2})|^2\,|\langle r_\ell,r_\ell\rangle|^2}\right|\ll|r|^4$$

and similarly

$$\left|\frac{|\Gamma(\frac{m(r_\ell)}{2}+\frac{m(r_\ell/2)}{4}+\frac{i\langle r,r_\ell\rangle}{\langle r_\ell,r_\ell\rangle})|^2}{|\Gamma(\frac{m(r_\ell)}{2})|^2|\Gamma(\frac{m(r_\ell/2)}{4}+\frac{i\langle r,r_\ell\rangle}{\langle r_\ell,r_\ell\rangle})|^2}-\frac{|\Gamma(\frac{m(r_\ell)}{2}+\frac{m(r_\ell/2)}{4})|^2}{|\Gamma(\frac{m(r_\ell)}{2})|^2|\Gamma(\frac{m(r_\ell/2)}{4})|^2}\right|\ll|r|^4.$$

Using these two estimates in (5.4.3) the lemma follows for a suitable constant c_G. □

Denote by

$$\gamma(r)=c_G e^{-c_2Q(r,r)}\prod_{\ell=1}^{p}|\langle r,r_\ell\rangle|^2$$

for c_G the constant from Lemma 5.4.4. We then may draw the following corollary.

Corollary 5.4.5. *Assume that μ has finite fourth moment. For $|r|\le\delta_0\sqrt{n}$ and $c'>0$ a constant depending on μ,*

$$\left|\frac{n^p}{\sigma^n}\lambda(\tfrac{r}{\sqrt{n}})^n|c(\tfrac{r}{\sqrt{n}})|^{-2}-\gamma(r)\right|\ll_\mu n^{-1}e^{-c'|r|^2}.$$

Proof. Combining Lemma 5.4.3 and Lemma 5.4.4,

$$\begin{aligned}\left|\frac{n^p}{\sigma^n}\lambda(\tfrac{r}{\sqrt{n}})^n|c(\tfrac{r}{\sqrt{n}})|^{-2}-\gamma(r)\right| &\leq \left|\frac{\lambda(\frac{r}{\sqrt{n}})^n}{\sigma^n}-e^{-c_2Q(r,r)}\right||n^p c(\tfrac{r}{\sqrt{n}})^{-2}| \\ &\quad+\left|n^p|c(\tfrac{r}{\sqrt{n}})|^{-2}-c_G\prod_{\ell=1}^p|\langle r,r_\ell\rangle|^2\right|e^{-c_2Q(r,r)} \\ &\ll_\mu n^{-1}e^{-c'|r|^2},\end{aligned}$$

by using in the last line and that $|r|^{2p+2}e^{-c_2Q(r,r)} \ll_\mu e^{-c'|r|^2}$ for a suitable constant $c' > 0$. □

5.5 Proof of Theorem 2.0.2 and Theorem 2.0.3

Throughout this section assume that μ has finite fourth moment. We are now in a suitable position to prove Theorem 2.0.2 and Theorem 2.0.3. Let $f \in \mathscr{S}(X)$. Recall that we expressed in (3.2.1) the term in question $\frac{n^{\ell/2}}{\sigma^n}\int f(g.x_0)\,d\mu^{*n}(g)$ for $x_0 = h_0K$ by using the Fourier inversion formula as

$$\frac{n^{\ell/2}}{\sigma^n}\int_{\mathfrak{a}^*}\int_\Omega \widehat{f}(r,\omega)(S_r^n\rho_r(h_0)1)(\omega)\,dm_\Omega(\omega)d\nu_{\mathrm{sph}}(r).$$

The latter term is decomposed into the high frequency (3.2.2) and low frequency (3.2.3) component for $\delta_0 \in (0,1)$ small enough such that Lemma 5.1.2 holds. Under the assumption $\sup_{|r|\geq 1}||S_r|| < 1$, the high frequency term (3.2.2) is dealt with by Lemma 5.3.3 collecting an error term of size $O_\mu(e^{-cn}||f||_{H^s})$ for $s = \frac{1}{2}(\dim X + 1)$. Without this assumption, one requires that the Fourier transform of f is compactly supported yielding by Lemma 5.3.2 an error term of size $O_{\mu,f}(e^{-c_fn})$.

For the low frequency term, one applies Lemma 5.4.1, thereby collecting an error term of size $O_\mu(e^{-cn}||f||_1)$. It remains to deal with (5.4.1), which after the substitution r to $\frac{n}{\sqrt{r}}$ is of the form (5.4.2). Using Lemma 5.2.1 and Corollary 5.4.5, we arrive at the term

$$\begin{aligned}&\int_{|r|\leq\delta_0\sqrt{n}}\gamma(r)\int_\Omega\widehat{f}(\tfrac{r}{\sqrt{n}},\omega)(E_{\frac{r}{\sqrt{n}}}\rho_{\frac{r}{\sqrt{n}}}(h)1)(\omega)\,dm_\Omega(\omega)dm_{\mathfrak{a}^*}(r)\\ &=\int f(g)\rho_G(h^{-1})\left(\int_{|r|\leq\delta_0\sqrt{n}}\gamma(r)\psi_{\mu,\frac{r}{\sqrt{n}}}(g)\,dm_{\mathfrak{a}^*}(r)\right)dm_G(g)\end{aligned}$$

admitting an additional error term of size

$$\ll_\mu n^{-1}||f||_1\int_{|r|\leq\delta_0\sqrt{n}}e^{-c'|r|^2}\,dm_{\mathfrak{a}^*}(r) \ll_\mu n^{-1}||f||_1,$$

using that the latter integral converges.

We define for $n \geq 1$ the continuous real-valued functions on G,

$$\psi_n(g) = \int_{|r|\leq \delta_0\sqrt{n}} \gamma(r)\psi_{\mu,\frac{r}{\sqrt{n}}}(g)\, dm_{\mathfrak{a}^*}(r) \qquad \text{and} \tag{5.5.1}$$

$$\psi_0(g) = c_\mu \cdot \psi_{\mu,0}(g) \qquad \text{for} \qquad c_\mu = \int_{r\in\mathfrak{a}^*} \gamma(r)\, dm_{\mathfrak{a}^*}(r).$$

While $\psi_{\mu,\frac{r}{\sqrt{n}}}$ is not necessarily real-valued, the function ψ_n is as $\overline{\psi_{\mu,\frac{r}{\sqrt{n}}}} = \psi_{\mu,-\frac{r}{\sqrt{n}}}$ and the definition of ψ_n is invariant under $r \mapsto -r$.

We have so far collected a total error of size

$$O_\mu(n^{-1}||f||_1 + e^{-cn}||f||_{H^s})$$

under the assumption $\sup_{|r|\geq 1} ||S_r|| < ||S_0||$ and for $f \in \mathscr{S}(X)$ and

$$O_\mu(n^{-1}||f||_1) + O_{\mu,f}(e^{-c_f n}||f||_1)$$

without the latter assumption yet requiring that the Fourier transform of f has compact support. To conclude the proof, we show the following lemma.

Lemma 5.5.1. *For $g \in G$ and $n \geq 1$,*

$$|\psi_n(g) - \psi_0(g)| \ll_\mu n^{-1}(1 + ||g||^2).$$

Proof. Since $\gamma(r) \ll_\mu e^{-c'|r|^2}$ for a suitable constant c' it follows that

$$|\psi_{\mu,0}(g)| \int_{|r|>\delta_0\sqrt{n}} \gamma(r)\, dm_{\mathfrak{a}^*}(r)$$

decays exponentially fast in n (using that $|\psi_{\mu,0}(g)| = |\langle \eta_0, \rho_0(g)\eta_0'\rangle| \ll_\mu 1$) and therefore we need to deal with

$$\left|\psi_n(g) - \int_{|r|\leq\delta_0\sqrt{n}} \gamma(r)\psi_{\mu,0}(g)\, dm_{\mathfrak{a}^*}(r)\right|. \tag{5.5.2}$$

By Lemma 5.2.2 it holds that

$$\left|\frac{\psi_{\mu,\frac{r}{\sqrt{n}}}(g) + \psi_{\mu,-\frac{r}{\sqrt{n}}}(g)}{2} - \psi_{\mu,0}(g)\right| \ll_\mu n^{-1}|r|^2(1 + ||g||^2)$$

and therefore using again that $\gamma(r) \ll_\mu e^{-c^*|r|^2}$ and as the defining integral of ψ_n is invariant under replacing r by $-r$,

$$\begin{aligned}
(5.5.2) &\ll \int_{|r|\leq\delta_0\sqrt{n}} \gamma(r)|\psi_{\mu,\frac{r}{\sqrt{n}}}(g) - \psi_{\mu,0}(g)|\, dm_{\mathfrak{a}^*}(r) \\
&\ll_\mu n^{-1}(1 + ||g||^2) \int_{|r|\leq\delta_0\sqrt{n}} \gamma(r)|r|^2\, dm_{\mathfrak{a}^*}(r) \\
&\ll_\mu n^{-1}(1 + ||g||^2).
\end{aligned}$$

□

Recall that we have defined

$$||f||_* = \int |f(x)|(1 + d_X(x,o)^2)\, dm_X(x) = \int |f(g)|(1 + ||g||^2)\, dm_G(g),$$

where we make no notational difference between f and its lift to G. To conclude the proof of (2.0.6) and (2.0.9) we estimate

$$\begin{aligned}
&\left| \int f(g.x_0)\psi_n(g)\, dm_G(g) - \int f(g.x_0)\psi_0(g)\, dm_G(g) \right| \\
&\leq \int |f(g)||\psi_n(gh_0^{-1}) - \psi_0(gh_0^{-1})|\, dm_G(g) \\
&\ll_\mu n^{-1} \int |f(g)|(1 + ||gh_0^{-1}||^2)\, dm_G(g) \\
&\ll_\mu n^{-1} \int |f(g)|(1 + ||g||^2 + ||h_0||^2)\, dm_G(g) \\
&\ll_\mu n^{-1}||f||_* + n^{-1} d_X(x_0, o)^2 ||f||_1,
\end{aligned}$$

having used in the penultimate line that $||gh_0^{-1}|| \leq ||g|| + ||h_0^{-1}||$ by Corollary 7.20 of [BQ16] as G is connected. This concludes the proof of Theorem 2.0.2 and of (2.0.9). The final claim of Theorem 2.0.3 is proved in the following lemma.

Lemma 5.5.2. *Let G be a non-compact connected semisimple Lie group with finite center and let μ be a non-degenerate probability measure on G with finite second moment. Assume that μ satisfies one of the following properties:*

(i) μ is spread out.

(ii) μ is bi-K-invariant, i.e. $m_K * \mu * m_K$.

Then S_0 is quasicompact and $\left(\sup_{|r|\geq 1} ||S_r||\right) < ||S_0||$.

Proof. The claim of the lemma was established for spread out measures in section 2.2 of [Bou81]. It remains to treat the case where μ is bi-K-invariant. Note that as $S_r = \rho_r(m_K) * S_r * \rho_r(m_K)$ it holds that $S_r 1_\Omega = \lambda(r) 1_\Omega$ and $S_r \langle 1_\Omega \rangle^\perp = \{0\}$ and therefore $\lambda(r) = \int \phi_r(g)\, d\mu(g)$. The claim now follows as $\phi_r(g) \to 0$ (cf. for example appendix A of [FM21]) for fixed $g \in G \backslash K$ and $r \to \infty$ and using that $\mu(G\backslash K) > 0$ as μ is non-degenerate. □

5.6 Proof of Theorem 2.0.1

Lemma 5.6.1. *Let G and μ be as in Theorem 2.0.1. Let $f \in \mathscr{S}(X)$ be a Schwartz function whose Fourier transform is compactly supported. Then*

$$\lim_{n\to\infty} \frac{n^{\ell/2}}{\sigma^n} \int f(g.x_0)\, d\mu^{*n}(g) = \int f(g.x_0)\psi_0(g)\, dm_G(g).$$

Proof. The proof is as the one of Theorem 2.0.3 expect that we cannot use Lemma 5.4.3. Revising the argument of Lemma 5.4.3, it follows that for the positive definite quadratic form Q from Lemma 5.4.3, under the assumption that μ has finite second moment, it holds that $\lambda(r) = \lambda(0) - Q(r,r) + o(|r|^2)$ and therefore for $|r| \le \delta_0\sqrt{n}$,

$$\lim_{n\to\infty} \frac{\lambda(r/\sqrt{n})^n}{\sigma^n} = e^{-c_2 Q(r,r)} \qquad \text{and} \qquad \frac{\lambda(r/\sqrt{n})^n}{\sigma^n} \ll e^{c'|r|^2}$$

for a suitable constant $c' > 0$. Similarly to Lemma 5.4.5,

$$\lim_{n\to\infty} \frac{n^p}{\sigma^n} \lambda(\tfrac{r}{\sqrt{n}})^n |c(\tfrac{r}{\sqrt{n}})|^{-2} = \gamma(r).$$

Arguing as in the proof of Theorem 2.0.2, it therefore follows by dominated convergence,

$$\begin{aligned}
&\lim_{n\to\infty} \frac{n^{\ell/2}}{\sigma^n} \int f(g.x_0)\, d\mu^{*n}(g) = \lim_{n\to\infty} (5.4.1) \\
&= \int_{r\in\mathfrak{a}^*} \gamma(r) \int_\Omega \widehat{f}(0,\omega)(E_0\rho_0(h_0)1)(\omega)\, dm_\Omega(\omega) dm_{\mathfrak{a}^*}(r) \\
&= \int f(g.x_0)\psi_0(g)\, dm_G(g).
\end{aligned}$$

□

Lemma 5.6.2. *Let $f \in \mathscr{S}(X)$. Then*

$$\limsup_{n\to\infty} \left| \frac{n^{\ell/2}}{\sigma^n} \int f(g.x_0)\, d\mu^{*n}(g) \right| \ll ||f||_1,$$

where the implied constant depends only on G.

Proof. One may reduce to functions $f \ge 0$. By covering the latter function suitably by a linear combination of characteristic functions, it suffices to show the claim for $f = 1_{B_\varepsilon(x)}$ with $\varepsilon > 0$ small and $x \in X$. By Theorem 5.7 of [And04] there is a positive function $h \in \mathscr{S}(X)$, whose Fourier transform has compact support, satisfying $1_{B_\varepsilon(x)} \le h$ and $||h||_1 \ll \mathrm{vol}_X(B_\varepsilon)$. The lemma follows by applying Lemma 5.6.1 to h. □

Proof. (of Theorem 2.0.1) Let $\delta_\ell \in \mathscr{S}(X)$ be an approximation to the identity on G that is bi-K-invariant and whose Fourier transform has compact support. Such functions exists by choosing a sequence ω_ℓ of smooth bi-K-invariant approximations to the identity that are supported on smaller and smaller balls around $e \in G$. As a Schwartz function is characterized by its Fourier transform, it suffices to determine $\widehat{\delta_\ell}$. Indeed one may choose $\widehat{\delta_\ell}$ to be equal to $\widehat{\omega_\ell}$ in a sufficiently large ball around the identity and to decay to zero rapdily outside of it. One then readily checks that δ_ℓ satisfies the required properties.

Then for $f \in \mathscr{S}(X)$, it holds for $r \in \mathfrak{a}^*$ and $k \in \Omega$,

$$\widehat{f * \delta_\ell}(r,k) = (\rho_{-r}(f * \delta_\ell)1)(k) = (\rho_{-r}(f)\rho_{-r}(\delta_\ell)1)(k) = \widehat{f}(r,k)\widehat{\delta_\ell}(r).$$

Therefore the Fourier transform of $f * \delta_\ell$ has compact support.

Combining Corollary 5.6.1 and Lemma 5.6.2, for $f \in \mathscr{S}(X)$,

$$\begin{aligned} &\frac{n^{\ell/2}}{\sigma^n}\int f(g.x_0)\, d\mu^{*n}(g) \\ &= \frac{n^{\ell/2}}{\sigma^n}\int (f * \delta_\ell)(g.x_0)\, d\mu^{*n}(g) + \frac{n^{\ell/2}}{\sigma^n}\int (f - f * \delta_\ell)(g.x_0)\, d\mu^{*n}(g) \\ &= \int f(g.x_0)\psi_0(g)\, dm_G(g) + O_\mu(||f - f * \delta_\ell||_1) + o_{f,\ell}(1) \end{aligned}$$

having used Lemma 5.6.2 and that $|\int (f - f*\delta_\ell)(g)\psi_0(gh_0^{-1})\, dm_G(g)| \ll_\mu ||f - f*\delta_\ell||_1$ as ψ_0 is bounded. The claim follows by choosing ℓ sufficiently slowly increasing in n. □

Chapter 6

Quasicompactness of S_0

In this section we discuss how to establish quasicompactness of S_0 under strong Diophantine assumption. The reader may recall the Littlewood-Paley decomposition $L^2(K) = \bigoplus_{\ell \geq 0} V_\ell$ (see (3.1.6)), where the space of functions V_ℓ can be pictured as oscillating with frequency 2^ℓ. The main result of this section states that under suitable assumptions, the operator S_0 has small norm on the space of functions with high enough oscillations.

Recall that we denoted by ρ_0^+ the Koopman representation induced by the G action on K, which contains the zero principal series ρ_0 as a subrepresentation and write $S_0^+ = \rho_0^+(\mu)$. Instead of considering S_0, we study S_0^+, which leads to stronger statements.

Theorem 6.0.1. *Let G be a non-compact connected simple Lie group with finite center. For $c_1, c_2 > 0$ there exists $\varepsilon_0 = \varepsilon_0(c_1, c_2) > 0$ such that the following holds. For any $0 < \varepsilon < \varepsilon_0$ and any symmetric and (c_1, c_2, ε)-Diophantine probability measure μ there is $L = L(c_1, c_2) \in \mathbb{Z}_{\geq 1}$ such that for $\varphi \in \bigoplus_{\ell \geq L} V_\ell$,*

$$||S_0^+ \varphi||_2 \leq \frac{1}{4} ||\varphi||_2. \tag{6.0.1}$$

Theorem 6.0.1 will be deduced in section 6.1 using results and ideas from [BISG17], thereby exploiting that the measure μ has high dimension (3.2.8) as well as a Littlewood-Paley decomposition and a mixing inequality on G. Under the additional assumption that K is semisimple, one may instead follow Bourgain's [Bou12] original ideas and improve (6.0.1).

Theorem 6.0.2. *Let G be a non-compact connected simple Lie group with finite center and maximal compact subgroup K. Assume that K is semisimple. For $c_1, c_2 > 0$ there exists $\varepsilon_0 = \varepsilon_0(c_1, c_2) > 0$ such that the following holds. For any $0 < \varepsilon < \varepsilon_0$*

and any symmetric and (c_1, c_2, ε)*-Diophantine probability measure* μ *there is* $L = L(c_1, c_2) \in \mathbb{Z}_{\geq 1}$ *such that for* $\varphi \in \bigoplus_{\ell \geq L} V_\ell$,

$$||S_0^+\varphi||_2 \leq \varepsilon^{O_{c_1,c_2}(1)}||\varphi||_2. \tag{6.0.2}$$

The proof of Theorem 6.0.2 was exposed in section 3.2. As in [BISG17] we exploit that μ has high dimension, yet we work with the Littlewood-Paley decomposition on K and use that the averages of matrix coefficients of V_ℓ are small (Proposition 4.5.1). From these results, one may easily deduce that S_0 and S_0^+ are quasicompact, therefore also implying Theorem 2.0.7.

Corollary 6.0.3. *Let* G *be a non-compact connected simple Lie group with finite center. For* $c_1, c_2 > 0$ *there exists* $\varepsilon_0 = \varepsilon_0(c_1, c_2) > 0$ *such that the following holds. For any* $0 < \varepsilon < \varepsilon_0$ *and any symmetric and* (c_1, c_2, ε)*-Diophantine probability measure* μ*, the operators* S_0 *and* S_0^+ *are quasicompact.*

Proof. As $||S_0|| = ||S_0^+||$ (by section D of [Gui80]) and since ρ_0^+ is a subrepresentation of ρ_0, it suffices to show that S_0^+ is quasicompact. By Lemma 4.1.1, the estimate (6.0.1) implies that $\rho_{\text{ess}}(S_0^+) \leq \frac{1}{4}$. As for $\varepsilon > 0$ small enough, $||\sqrt{\alpha_g'} - 1||_\infty \ll |\delta|\,||g|| \ll \varepsilon^{O(1)}$ for $g \in B_\varepsilon$, it holds that $||S_0^+|| \geq 1 - \varepsilon^{O(1)}$ and hence the claim follows. □

We next explain how to deduce from (6.0.1) that the Furstenberg measure is absolutely continuous. Given a non-degenerate probability measure, we study the operator

$$T_0 : L^2(\Omega) \to L^2(\Omega), \qquad \varphi \mapsto T_0\varphi = \int \varphi \circ \alpha_g \, d\mu(g).$$

As we discuss in the proof of Corollary 6.0.4, it is shown in [BQ18] that if $\rho_{\text{ess}}(T_0) < 1$, then the Furstenberg measure of μ is absolutely continuous. The following corollary is also necessary to establish Theorem 2.0.8.

Corollary 6.0.4. *Let* G *be a non-compact connected simple Lie group with finite center. For* $c_1, c_2 > 0$ *there exists* $\varepsilon_0 = \varepsilon_0(c_1, c_2) > 0$ *such that the following holds. For any* $0 < \varepsilon < \varepsilon_0$ *and any symmetric and* (c_1, c_2, ε)*-Diophantine probability measure* μ *there is* $L = L(c_1, c_2) \in \mathbb{Z}_{\geq 1}$ *such that*

$$||T_0\varphi||_2 \leq \frac{1}{2}||\varphi||_2 \qquad \textit{for} \qquad \varphi \in \left(L^2(\Omega) \cap \bigoplus_{\ell \geq L} V_\ell \right). \tag{6.0.3}$$

Moreover, $\rho_{\text{ess}}(T_0) < 1$ *and the Furstenberg measure of* μ *is absolutely continuous.*

Proof. Using as in the proof of Corollary 6.0.3 that $||\sqrt{\alpha'_g} - 1||_\infty \ll |\delta|\,||g|| \ll \varepsilon^{O(1)}$ for $g \in B_\varepsilon$ and $\varepsilon > 0$ small enough, it follows that $||S_0 - T_0|| \leq \varepsilon^{O(1)}$. Therefore (6.0.3) is implied by (6.0.1). By Lemma 4.1.1 we hence conclude $\rho_{\mathrm{ess}}(T_0) < 1$.

We finally review the argument from [BQ18] to show that the Furstenberg measure of μ is absolutely continuous under the assumption that $\rho_{\mathrm{ess}}(T_0) < 1$. Indeed as $T_0 1 = 1$, it follows that 1 is in the discrete spectrum of T_0. If $\rho_{\mathrm{ess}}(T_0) < 1$, one furthermore concludes (cf. Fact 2.3 of [BQ18]) that 1 is in the discrete spectrum of the adjoint operator T_0^* and therefore there is a function $\psi_{\mathrm{F}} \in L^2(\Omega)$ satisfying $T_0^* \psi_{\mathrm{F}} = \psi_{\mathrm{F}}$. One then readily checks that $\psi_{\mathrm{F}} dm_\Omega$ is a μ-stationary measure and thus by uniqueness of the Furstenberg measure it holds $d\nu_{\mathrm{F}} = \psi_{\mathrm{F}} dm_K$. □

We comment on the organization of this section. Theorem 6.0.1 is proved in section 6.1. The proof of Theorem 6.0.2 comprises two steps. In section 6.2 we first establish using the flattening results from Theorem 4.4.2 that $S_0^+|_{V_\ell}$ has small operator norm. In section 6.3 we complete the proof of Theorem 6.0.2 by using that $S_0^+ V_\ell$ and $V_{\ell'}$ are almost orthogonal. Finally in section 6.4 we show how to deduce that the Furstenberg measure has a $C^m(K)$ density.

6.1 Proof of Theorem 6.0.1

Write $T_0^+ \varphi = \int \varphi \circ \alpha_g \, d\mu(g)$ for $\varphi \in L^2(K)$. Since $||S_0^+ - T_0^+|| \leq \varepsilon^{O(1)}$, as argued in the proof of Corollary 6.0.4, in order to prove Theorem 6.0.1 it suffices to show that

$$||T_0^+ \varphi||_2 \leq \frac{1}{8} ||\varphi||_2 \tag{6.1.1}$$

for $\varphi \in \bigoplus_{\ell \geq L} V_\ell$ and $L = L(c_1, c_2)$.

We proceed similarly to the proof of Corollary C of [BISG17]. Indeed, we reduce the problem at hand to studying the regular representation on $L^2(G)$. One then uses the following result of [BISG17], which may be considered as their core technical contribution, which uses that μ has high dimension as well as a novel Littlewood-Paley decomposition and a mixing inequality on G. We rephrase their result using the notion of (c_1, c_2, ε)-Diophantine measures.

To introduce notation, for a measurable subset $B \subset G$ we consider the norm

$$||f||^2_{L^2(B)} = \int_B |f(g)|^2 \, dm_G.$$

Theorem 6.1.1. *(Theorem 6.7 of* [BISG17]*) Let G be a connected simple Lie group with finite center and $B \subset G$ a measurable set with compact closure. Let $c_1, c_2 > 0$.*

Then there is $\varepsilon_0 = \varepsilon_0(B, c_1, c_2) > 0$ such that the following holds. For any $0 < \varepsilon < \varepsilon_0$ and any symmetric and (c_1, c_2, ε)-Diophantine probability measure μ there is a finite dimensional subspace $V_B \subset L^2(B)$ such that

$$||\lambda_G(\mu)|_{(V_B)^\perp}||_{\mathrm{op},L^2(B)} \leq \varepsilon^{O_{B,c_1,c_2}(1)}.$$

In order to apply Theorem 6.1.1, we use the following lemma, which is inspired by the proof of Corollary C of [BISG17]. Denote by $\pi_K : G \to K = G/P^+$ the natural projection.

Lemma 6.1.2. *Denote $B = \{g \in G : |\kappa(g)| \leq c\}$ for $c > 0$. For small enough $c > 0$ there is a constant $D > 1$ depending on G and $c > 0$ such for all $\varphi \in L^2(K)$,*

$$D^{-1}||\varphi||_{L^2(K)} \leq ||\varphi \circ \pi_K||_{L^2(B)} \leq D||\varphi||_{L^2(K)}. \tag{6.1.2}$$

Proof. Recall that we denote $P^+ = AN$. By [BdlHV08] Theorem B.1.4 there is a continuous function $\rho : G \to \mathbb{R}_{>0}$ such that

$$\int_G f(g)\rho(g)\, dm_G(g) = \int_K \int_{P^+} f(kp)\, dm_P(p) dm_K(k) \tag{6.1.3}$$

for all $f \in L^1(G)$ with compact support. It moreover holds that $\alpha'_g(xP^+) = \frac{\rho(gx)}{\rho(x)}$ for all $x, g \in G$. We then calculate for $\varphi_1, \varphi_2 \in L^2(K)$ using (6.1.3),

$$\begin{aligned}
&|m_G(B)\langle\varphi_1, \varphi_2\rangle_{L^2(K)} - \langle\varphi_1 \circ \pi_K, \varphi_2 \circ \pi_K\rangle_{L^2(B)}| \\
&= \left| \int_K \int_{P^+} \varphi_1(k)\overline{\varphi_2(k)} 1_B(p)\, dm_P(p) dm_K(k) - \int 1_B(g)\varphi_1(\pi_K(g))\overline{\varphi_2(\pi_K(g))}\, dm_G(g) \right| \\
&= \left| \int_B \varphi_1(\pi_K(g))\overline{\varphi_2(\pi_K(g))}(1 - \rho(g))\, dm_G(g) \right| \\
&\leq ||\varphi_1 \circ \pi_K||_{L^2(B)} \sqrt{\int_B |\varphi_2(\pi_K(g))|^2\, |1 - \rho(g)|^2\, dm_G(g)} \\
&\leq D'||\varphi_1 \circ \pi_K||_{L^2(B)}||\varphi_2 \circ \pi_K||_{L^2(B)},
\end{aligned}$$

for a suitable constant D' using that $|1 - \rho(g)|$ is bounded on the compact set B. By a similar argument and possibly enlarging the constant D', we may also estimate the latter term by

$$m_G(B) D'||\varphi_1||_{L^2(K)}||\varphi_2||_{L^2(K)}.$$

Setting $\varphi = \varphi_1 = \varphi_2$ the claim is readily implied by choosing D suitably in terms of D' and $m_G(B)$. □

Throughout the following denote by $B = \{g \in G : |\kappa(g)| \leq c\}$ a set from Lemma 6.1.2 such that (6.1.2) holds. We are now in a suitable position to apply Theorem 6.1.1. Indeed for $\varphi \in L^2(K)$ it holds by (6.1.2) that

$$||T_0\varphi||_{L^2(K)} \leq D||(T_0\varphi) \circ \pi_K||_{L^2(B)} = D||\lambda_G(\mu)(\varphi \circ \pi_K)||_{L^2(B)}. \tag{6.1.4}$$

Let $V_B \subset L^2(B)$ the finite dimensional subspace of Theorem 6.1.1. We then may choose L large enough such that if $\varphi \in \bigoplus_{\ell \geq L} V_\ell$ then

$$||\varphi \circ \pi_K - (\varphi \circ \pi_K)_{(V_B)^\perp}||_{L^2(B)} \leq \frac{1}{16D^2}||\varphi \circ \pi_K||_{L^2(B)}, \tag{6.1.5}$$

where $(\varphi \circ \pi_K)_{(V_B)^\perp}$ is the projection of $\varphi \circ \pi_K$ onto $(V_B)^\perp$. Indeed this follows using (6.1.3) and that V_B is finite dimensional.

We conclude using Theorem 6.1.1, (6.1.2),(6.1.4) and (6.1.5),

$$\begin{aligned}
||T_0\varphi||_{L^2(K)} &\leq D||\lambda_G(\mu)(\varphi \circ \pi_K)||_{L^2(B)} \\
&\leq D||\lambda_G(\mu)\left(\varphi \circ \pi_K - (\varphi \circ \pi_K)_{(V_B)^\perp}\right)||_{L^2(B)} + D||\lambda_G(\mu)(\varphi \circ \pi_K)_{(V_B)^\perp}||_{L^2(B)} \\
&\leq \frac{1}{16D}||\varphi \circ \pi_K||_{L^2(B)} + D\varepsilon^{O_{c_1,c_2}(1)}||\varphi \circ \pi_K||_{L^2(B)} \\
&\leq \left(\frac{1}{16} + D^2\varepsilon^{O_{c_1,c_2}(1)}\right)||\varphi||_{L^2(K)},
\end{aligned}$$

showing (6.1.1) by choosing ε small enough in terms of c_1 and c_2. The proof of Theorem 6.0.1 is complete.

6.2 Operator Norm Estimate for S_0^+ on V_ℓ

In this section we prove the following proposition.

Proposition 6.2.1. *For $c_1, c_2 > 0$ there exists $\varepsilon_0 = \varepsilon_0(G, c_1, c_2) > 0$ such that the following holds. For any $0 < \varepsilon < \varepsilon_0$ and any symmetric and (c_1, c_2, ε)-Diophantine probability measure μ, there is $L = L(G, c_1, c_2) \in \mathbb{Z}_{\geq 1}$ such that $||S_0^+|_{V_\ell}||_{\mathrm{op}} \leq \varepsilon^{O_{c_1,c_2}(1)}$ for $\ell \geq L$.*

Recall that as introduced in section 4.4,

$$P_\delta = \frac{1_{B_\delta}}{m_G(B_\delta)},$$

where B_δ is the open δ-ball around $e \in G$. For the proof of Proposition 6.2.1, one estimates by the triangle inequality for $n \geq 1$ and $\varphi \in V_\ell$,

$$||(S_0^+)^n\varphi||_2 \leq ||(S_0^+)^n\varphi - \rho_0^+(\mu^{*n} * P_\delta)\varphi||_2 + ||\rho_0^+(\mu^{*n} * P_\delta)\varphi||_2. \tag{6.2.1}$$

We aim to show that (6.2.1) is very small for a suitably chosen n and δ. For the first term of (6.2.1), we use that the Lipschitz constant of φ is $\asymp ||\gamma||^{O(1)}$. Therefore, a δ-perturbation of $(S_0^+)^n\varphi = \rho_0^+(\mu^{*n})\varphi$ is small provided we choose δ miniscule in terms of ℓ.

The second term of (6.2.1) is dealt with by using that μ has high dimension. Indeed by Lemma 4.4.2 it will follow that $\mu^{*n} * P_\delta$ has small $||\cdot||_\infty$-norm for n chosen in terms of δ. This will allow us to compare $||\rho_0^+(\mu^{*n} * P_\delta)\varphi||_2$ to the average estimate of matrix coefficients

$$\frac{1}{m_G(B_R)} \int_{B_R} |\langle \rho_0^+(g)\varphi, \varphi\rangle| \, dm_G(g) \ll 2^{-\ell/2}||\varphi||_2$$

that was discussed in section 4.5.

We proceed with some preliminary lemmas used in the proof of Proposition 6.2.1. First, we estimate how much $\rho_0^+(g)\varphi$ differs from φ, given that $\varphi \in V_\ell$ and $g \in B_\delta$.

Lemma 6.2.2. *Fix $\ell \geq 0$. Then for $\varphi \in V_\ell$ and $0 < \delta \ll 2^{-\ell}$, it holds for $g \in B_\delta$,*

$$||\rho_0^+(g)\varphi - \varphi||_2 \ll e^{O(1)\ell}\delta^{O(1)}||\varphi||_2.$$

Proof. We first fix $\gamma \in \overline{C} \cap I^*$ and denote as usual by π_γ the associated irreducible representation and let $v_1, \ldots, v_n \in \pi_\gamma$ be an orthonormal basis of the representation space of π_γ. For $k \in B_\delta$ in K for δ small enough, it holds by Lemma 3.1 of [dS13] that $||\pi_\gamma(k) - \mathrm{Id}_{\pi_\gamma}||_{\mathrm{op}} \ll d_K(k, e)||\gamma||$. Indeed, upon conjugation, we can assume that k is inside the maximal torus T of K and hence we can write $k = e^X$ for $X \in \mathfrak{t} = \mathrm{Lie}(T)$ with $||X|| \ll d_K(k, e)$. With these assumptions, the eigenvalues of $\pi_\gamma(k) - \mathrm{Id}_{\pi_\gamma}$ can be calculated as $e^{\gamma'(X)} - 1$ for γ' the weights of π_γ. Choosing $\delta \ll 2^{-\ell}$, and therefore having $|\gamma'(X)| \ll 1$, we can bound $\max_{\gamma'} |e^{\gamma'(X)} - 1| \ll \max_{\gamma'} |\gamma'(X)| \ll d_K(g, e)||\gamma||$, showing the claim.

Denote by ψ the matrix coefficient $k \mapsto \sqrt{d_\pi}\langle \pi_\gamma(k)v_i, v_j\rangle$, satifying $||\psi||_2 = 1$. We first show that $||\rho_0^+(g)\psi - \psi||_2 \ll \delta^{O(1)}||\gamma||^{O(1)}$ for $g \in B_\delta$. Indeed, using as in the proof of Corollary 6.0.3 that $||\sqrt{\alpha_g'}(k) - 1||_\infty \ll \delta^{O(1)}$ and Lemma 4.3.1,

$$\begin{aligned} |(\rho_0^+(g)\psi)(k) - \psi(k)| &\leq \left| \left(\sqrt{\alpha_g'}(k) - 1\right) \psi(g^{-1}.k)\right| + \left|\psi(g^{-1}.k) - \psi(k)\right| \\ &\ll \delta^{O(1)}|\psi(g^{-1}.k)| + \sqrt{d_\pi}||\pi_\gamma(g^{-1}.k) - \pi_\gamma(k)||_{\mathrm{op}} \\ &\ll \delta^{O(1)}|\psi(g^{-1}.k)| + \delta^{O(1)}||\gamma||^{O(1)}, \end{aligned}$$

which implies the claim using $|\psi(g^{-1}.k)| \leq |\psi(g^{-1}.k) - \psi(k)| + |\psi(k)|$.

To prove the lemma, denote by $(\psi_i)_{i\in I}$ an orthonormal basis of V_ℓ with functions as in the previous paragraph. Then $|I| \ll e^{O(1)\ell}$ and for $\varphi \in V_\ell$ we decompose $\varphi = \sum_{i\in I} a_i\psi_i$, implying using Cauchy-Schwarz,

$$||\rho_0^+(g)\varphi - \varphi||_2 \leq \sum_{i\in I} |a_i|\, ||\rho_0^+(g)\psi - \psi||_2 \ll e^{O(1)\ell}\delta^{O(1)}||\varphi||_2.$$

□

We next show how to compare $\pi(\nu)\varphi$ with $\pi(\nu * P_\delta)\varphi$ for a suitable vector φ and a unitary representation π and probability measure ν.

Lemma 6.2.3. *Let $(\pi, \mathscr{H})$ be a unitary representation of G and let $\delta > 0$. Fix $\varphi \in \mathscr{H}$. Assume that $||\pi(g)\varphi - \varphi|| \leq C_\delta||\varphi||$ for all $g \in B_\delta$ and $C_\delta > 0$ a constant. Then for any probability measure ν,*

$$||\pi(\nu)\varphi - \pi(\nu * P_\delta)\varphi|| \leq C_\delta||\varphi||.$$

Proof. Using Fubini's theorem and that $1_{B_\delta(g)}(h) = 1_{B_\delta(h)}(g)$,

$$\begin{aligned}\pi(\nu)\varphi &= \int \frac{1}{m_G(B_\delta(e))}\left(\int 1_{B_\delta(g)}(h)\pi(g)\varphi\, dm_G(h)\right) d\nu(g)\\ &= \int \frac{1}{m_G(B_\delta(e))}\left(\int_{B_\delta(h)} \pi(g)\varphi\, d\nu(g)\right) dm_G(h).\end{aligned}$$

Furthermore, by the assumption and using that $B_\delta(h) = hB_\delta(e)$ (the metric on G is left invariant),

$$\begin{aligned}\left|\left|\int_{B_\delta(h)} \pi(g)\varphi\, d\nu(g) - \nu(B_\delta(h))\cdot\pi(h)\varphi\right|\right| &\leq \int_{B_\delta(h)} ||(\pi(g) - \pi(h))\varphi||\, d\nu(g)\\ &\leq \int_{B_\delta(h)} ||\pi(h)(\pi(h^{-1}g) - \mathrm{Id})\varphi||\, d\nu(g)\\ &\leq \nu(B_\delta(h))C_\delta||\varphi||.\end{aligned}$$

Finally, as $(\nu * P_\delta)(h) = \frac{\nu(B_\delta(h))}{m_G(B_\delta(e))}$,

$$\begin{aligned}&||\pi(\nu)\varphi - \pi(\nu * P_\delta)\varphi||\\ &= \left|\left|\int \frac{1}{m_G(B_\delta(e))}\left(\int_{B_\delta(h)} \pi(g)\varphi\, d\nu(g)\right) dm_G(h) - \int \pi(h)\varphi\,(\nu * P_\delta)(h)\, dm_G(h)\right|\right|\\ &\leq \int \frac{1}{m_G(B_\delta(e))}\left|\left|\int_{B_\delta(h)} \pi(g)\varphi\, d\nu(g) - \nu(B_\delta(h))\cdot\pi(h)\varphi\right|\right| dm_G(h)\\ &\leq C_\delta||\varphi||\cdot\int \frac{\nu(B_\delta(h))}{m_G(B_\delta(e))}\, dm_G(h) = C_\delta||\varphi||,\end{aligned}$$

using in the last line that by Fubini's theorem $\int \frac{\nu(B_\delta(h))}{m_G(B_\delta(e))}\, dm_G(h) = 1$ as ν is a probability measure. □

Proof. (of Proposition 6.2.1) Let $\gamma > 0$ be a fixed constant to be determined later. Then by Proposition 4.4.2 there is $\varepsilon_0 = \varepsilon_0(c_1, c_2) > 0$ and $C_0 = C_0(c_1, c_2) > 0$ such that for $\delta > 0$ small enough it holds that $||(\mu^{*n})_\delta||_2 \leq \delta^{-\gamma}$ for any $n \geq C_0 \frac{\log \frac{1}{\delta}}{\log \frac{1}{\varepsilon}}$ and

$$(\mu^{*n})_\delta = \mu^{*n} * P_\delta.$$

Let $\varphi \in V_\ell$ with $||\varphi||_2 = 1$. Then by the triangle inequality,

$$||(S_0^+)^n \varphi||_2 \leq ||(S_0^+)^n \varphi - \rho_0^+(\mu^{*n} * P_\delta)\varphi||_2 + ||\rho_0^+(\mu^{*n} * P_\delta)\varphi||_2.$$

The first term can be estimated using Lemma 6.2.2 and Lemma 6.2.3 as $||(S_0^+)^n \varphi - \rho_0^+(\mu^{*n} * P_\delta)\varphi||_2 \ll e^{O(1)\ell}\delta^{O(1)}$ assuming that $\delta \ll 2^{-\ell}$. For the second term, first notice that by applying Cauchy-Schwarz it follows that $||(\mu^{*n})_\delta * (\mu^{*n})_\delta||_\infty \leq ||(\mu^{*n})_\delta||_2^2$. Then with Theorem 4.4.2 and Proposition 4.5.1,

$$\begin{aligned}
||\rho_0^+(\mu^{*n} * P_\delta)\varphi||_2^2 &= \langle \rho_0^+(\mu^{*n} * P_\delta * \mu^{*n} * P_\delta)\varphi, \varphi\rangle \\
&\leq \int |\langle \rho_0^+(g)\varphi, \varphi\rangle| \left((\mu^{*n})_\delta * (\mu^{*n})_\delta\right)(g)\, dm_G(g) \\
&\leq \delta^{-2\gamma} \int_{B_{4n\varepsilon}} |\langle \rho_0^+(g)\varphi, \varphi\rangle|\, dm_G(g) \\
&\ll \delta^{-2\gamma} m_G(B_{4n\varepsilon}) e^{-O(1)\ell} \leq \delta^{-2\gamma} e^{O(1)n\varepsilon} e^{-O(1)\ell}.
\end{aligned}$$

Let n be a power of 2 satisfying $n \asymp C_0 \frac{\log \frac{1}{\delta}}{\log \frac{1}{\varepsilon}}$. Then by using that S_0^+ is self-adjoint and n a power of 2, it follows by induction on k with $2^k = n$ that $||(S_0^+)\varphi||_2^n \leq ||(S_0^+)^n \varphi||_2$. Therefore it follows for $\delta \ll 2^{-\ell}$ that

$$||S_0^+|_{V_\ell}||_{\mathrm{op}} \leq D^{\frac{1}{n}} \max\{e^{\frac{\sigma_1 \ell}{n}} \delta^{\frac{\sigma_2}{n}}, \delta^{-\frac{\gamma}{n}} e^{-\frac{\sigma_3 \ell}{n}}\},$$

for $D, \sigma_1, \sigma_2, \sigma_3 > 0$ absolute constants. We choose $\delta = e^{-\max\{1, \frac{2\sigma_1}{\sigma_2}\}\ell}$ so that $\delta \ll 2^{-\ell}$ and $e^{\frac{\sigma_1 \ell}{n}} \delta^{\frac{\sigma_2}{n}} \leq e^{-\frac{\sigma_1 \ell}{n}}$. We furthermore set $\gamma = \frac{\sigma_3}{2\max\{1, 2\frac{\sigma_1}{\sigma_2}\}}$ and therefore $\delta^{-\frac{\gamma}{n}} e^{-\frac{\sigma_3 \ell}{n}} = e^{-\frac{\sigma_3 \ell}{2n}}$. With these choices, $||S_0^+|_{V_\ell}||_{\mathrm{op}} \leq D^{\frac{1}{n}} e^{-O(1)\frac{\ell}{n}}$. In addition we make ℓ large enough in terms of c_1 and c_2 such that δ becomes small enough for Proposition 4.4.2 to hold. To conclude, it holds by construction that $\frac{\ell}{n} \asymp_{c_1, c_2} \log \frac{1}{\varepsilon}$ and therefore $e^{-O(1)\frac{\ell}{n}} = \varepsilon^{O_{c_1, c_2}(1)}$ and similarly $D^{\frac{1}{n}} = \varepsilon^{-\frac{O_{c_1, c_2}(1)}{\ell}}$, so choosing ℓ additionally larger than a further constant depending on c_1 and c_2, the claim follows. □

6.3 Proof of Theorem 6.0.2

Having established that $||S_0^+|_{V_\ell}||_{\mathrm{op}}$ is small for $\ell \geq L(c_1, c_2)$, we aim to convert this to an estimate that $||S_0^+|_{\bigoplus_{\ell \geq L} V_\ell}||_{\mathrm{op}}$ is also small. We use that the spaces $S_0^+ V_\ell$ and $V_{\ell'}$ are almost orthogonal for $\ell \neq \ell'$ as shown in Lemma 6.3.2.

The Lie algebra of K is denoted $\mathfrak{k}$ and we also write λ_K for the Lie algebra representation induced by the regular representation λ_K on K. Indeed, for a smooth function φ on K the function $(\lambda_K(X)\varphi)(k) = \lim_{t\to 0} \frac{1}{t}(\varphi(e^{-tX}k) - \varphi(k))$ with $X \in \mathfrak{k}$ and $k \in K$ is the directional derivative of φ in the direction $-X$.

As in [Bou12], we use an argument based on partial integration to show that $S_0^+ V_\ell$ and $V_{\ell'}$ are almost orthogonal. For a general manifold there is no suitable partial integration formula. However, for compact Lie groups we overcome this issue by exploiting that the Laplacian acts as a scalar on functions on $L^2(K)$ induced by the representation π_γ. Indeed, for a fixed orthonormal basis $X_1, \ldots, X_{\dim K}$ of $\mathfrak{k}$ recall that the Casimir element is defined as $\triangle = -\sum_i X_i \circ X_i$. We then use as replacement to partial integration that

$$\langle \varphi_1, \lambda_K(\triangle)\varphi_2\rangle = \sum_i \langle \lambda_K(-X_i)\varphi_1, \lambda_K(X_i)\varphi_2\rangle. \tag{6.3.1}$$

In order to give a suitable estimate for (6.3.1), we first analyse $||\lambda_K(X)\varphi||_2$ for $X \in \mathfrak{k}$.

Lemma 6.3.1. *Let $\ell \geq 0$ and $\varepsilon > 0$. Then for $\varphi \in V_\ell, g \in B_\varepsilon$ and $X \in \mathfrak{k}$ of unit norm,*

$$||\lambda_K(X)\varphi||_2 \ll 2^\ell ||\varphi|| \qquad \text{and} \qquad ||\lambda_K(X)(\rho_0^+(g)\varphi)||_2 \ll (1 + O(\varepsilon^{O(1)}))2^\ell ||\varphi||_2.$$

Proof. Without loss of generality we assume that $X \in \mathfrak{t}$. Fix $\gamma \in \overline{C} \cap I^*$. The eigenvalues of the operator $\pi_\gamma(e^{tX}) - \mathrm{Id}$ can be calculated as $e^{t\gamma'(X)} - 1$ for γ' the various weights of the representation π_γ. Therefore the operator $\pi_\gamma(X) = \lim_{t\to 0} \frac{1}{t}(\pi_\gamma(e^{tX}) - \mathrm{Id})$ has eigenvalues $\gamma'(X)$. Let $v_1, \ldots, v_n$ be an orthonormal basis of eigenvectors of $\pi_\gamma(X)$. Then the functions $\psi(k) = \sqrt{d_\gamma}\langle \pi_\gamma(k)v_i, v_j\rangle$ for $k \in K$ satisfy $(\lambda_K(X)\psi)(k) = \sqrt{d_\gamma}\langle \pi_\gamma(k)v_i, \pi_\gamma(X)v_j\rangle = (\gamma'(X)\psi)(k)$. The first claim follows as $||\gamma'(X)|| \ll ||\gamma|| \leq 2^\ell$ and by decomposing the function φ as a sum of functions of the form ψ.

For the second claim recall that $\rho_0^+(g)\varphi = \sqrt{\alpha_g'} \cdot (\varphi \circ \alpha_g)$ and therefore

$$\lambda_K(X)(\rho_0^+(g)\varphi) = \left(\lambda_K(X)\sqrt{\alpha_g'}\right) \cdot (\varphi \circ \alpha_g) + \sqrt{\alpha_g'} \cdot \lambda_K(X)(\varphi \circ \alpha_g). \tag{6.3.2}$$

To deal with the first term of (6.3.2), since α_g' is a smooth polynomial perturbation of the identity, it follows that $||\lambda_K(X)\sqrt{\alpha_g'}||_\infty \leq (1 + O(\varepsilon^{O(1)}))$ and furthermore using integration by substitution, $||\varphi \circ \alpha_g||_2 \ll (1 + O(\varepsilon^{O(1)}))||\varphi||_2$. For the second term of (6.3.2), we use the chain rule and the first step to conclude that $||\lambda_K(X)(\varphi \circ \alpha_g)||_2 \ll (1 + O(\varepsilon^{O(1)}))2^\ell ||\varphi||_2$, concluding the lemma. □

We now apply (6.3.1) to prove the following lemma.

Lemma 6.3.2. *For $\varphi_{\ell_1} \in V_{\ell_1}$ and $\varphi_{\ell_2} \in V_{\ell_2}$ with $\ell_1 \neq \ell_2$ and $g \in B_\varepsilon$,*

$$|\langle \rho_0^+(g)\varphi_{\ell_1}, \varphi_{\ell_2}\rangle| \ll (1 + O(\varepsilon^{O(1)}))2^{-|\ell_1-\ell_2|}||\varphi_{\ell_1}||_2||\varphi_{\ell_2}||_2.$$

Proof. Without loss of generality we assume that $\ell_2 > \ell_1$. Denote by $\psi \in V_{\ell_2}$ the function such that $\lambda_K(\triangle)\psi = \varphi_{\ell_2}$. Then by Lemma 4.3.1, $||\psi||_2 \ll 2^{-2\ell_2}||\varphi_{\ell_2}||_2$. Using then (6.3.1) and Lemma 6.3.1,

$$\begin{aligned}
|\langle \rho_0^+(g)\varphi_{\ell_1}, \varphi_{\ell_2}\rangle| &= |\langle \rho_0^+(g)\varphi_{\ell_1}, \lambda_K(\triangle)\psi\rangle| \\
&= \left| \sum_i \langle \lambda_K(-X_i)\rho_0^+(g)\varphi_{\ell_1}, \lambda_K(X_i)\psi\rangle \right| \\
&\leq \sum_i ||\lambda_K(-X_i)(\rho_0^+(g)\varphi_{\ell_1})||\,||\lambda_K(X_i)\psi|| \\
&\ll (1 + O(\varepsilon^{O(1)}))2^{\ell_1+\ell_2}||\varphi_{\ell_1}||_2||\psi||_2 \\
&\ll (1 + O(\varepsilon^{O(1)}))2^{\ell_1-\ell_2}||\varphi_{\ell_1}||_2||\varphi_{\ell_2}||_2.
\end{aligned}$$

□

We conclude this section by proving Theorem 6.0.2 by combining Proposition 6.2.1 and Lemma 6.3.2.

Proof. (of Theorem 6.0.2) By Proposition 6.2.1, there is $\varepsilon_0 = \varepsilon_0(c_1, c_2) > 0$ and $L = L(c_1, c_2) \in \mathbb{Z}_{\geq 1}$ such that $||S_0^+|_{V_\ell}||_{\mathrm{op}} \leq \varepsilon^{O_{c_1,c_2}(1)}$ for $\ell \geq L$. Let $\varphi \in \bigoplus_{\ell \geq L} V_\ell$ and let $N \geq 1$ to be determined later. Then

$$\begin{aligned}
||S_0^+\varphi||_2^2 &\leq \sum_{\ell,\ell' \geq L} |\langle S_0^+\pi_\ell\varphi, S_0^+\pi_{\ell'}\varphi\rangle| \\
&= \sum_{|\ell-\ell'| \leq N} |\langle S_0^+\pi_\ell\varphi, S_0^+\pi_{\ell'}\varphi\rangle| + \sum_{|\ell-\ell'|>N} |\langle S_0^+\pi_\ell\varphi, S_0^+\pi_{\ell'}\varphi\rangle|,
\end{aligned}$$

where both of the sums are with $\ell, \ell' \geq L$. For the first of these two terms one uses the conclusion of Proposition 6.2.1,

$$\sum_{|\ell-\ell'| \leq N} ||S_0^+\pi_\ell\varphi||\,||S_0^+\pi_{\ell'}\varphi|| \leq N \sum_{\ell \geq L} ||S_0^+\pi_\ell\varphi||_2^2 \leq N\varepsilon^{O_{c_1,c_2}(1)}||\varphi||_2^2.$$

Lemma 6.3.2 is used to bound the second term:

$$\begin{aligned}
\sum_{|\ell-\ell'|>N} |\langle S_0^+\pi_\ell\varphi, S_0^+\pi_{\ell'}\varphi\rangle| &\ll \sum_{|\ell-\ell'|>N} 2^{-|\ell-\ell'|}||\pi_\ell\varphi||\,||\pi_{\ell'}\varphi|| \\
&\ll \sum_{|\ell-\ell'|>N} 2^{-|\ell-\ell'|}||\pi_\ell\varphi||_2^2 \\
&\ll 2^{-N}\sum_{\ell \geq L} ||\pi_\ell\varphi||_2^2 = 2^{-N}||\varphi||_2^2.
\end{aligned}$$

Therefore it follows that $||S_0^+\varphi||_2 \leq \sqrt{N\varepsilon^{O_{c_1,c_2}(1)} + 2^{-N}}||\varphi||_2$. Setting $N = \log\frac{1}{\varepsilon}$ implies the claim of the theorem. $\square$

6.4 Smoothness of the Furstenberg Measure

In this section we prove Theorem 2.0.8, which we restate here for convenience of the reader.

Theorem 6.4.1. *(Theorem 2.0.8) Let G be a non-compact connected simple Lie group with finite center. Let $c_1, c_2 > 0$ and $m \in \mathbb{Z}_{\geq 1}$. Then there is $\varepsilon_m = \varepsilon_m(G, c_1, c_2) > 0$ depending on G, c_1, c_2 and m such that every symmetric and (c_1, c_2, ε)-Diophantine probability measure μ with $\varepsilon \leq \varepsilon_m$ has absolutely continuous Furstenberg measure with density in $C^m(\Omega)$.*

By Corollary 6.0.4, we know that the Furtstenberg measure is absolutely continuous if we choose ε_m small enough, i.e. there is $\psi_{\mathrm{F}} \in L^2(\Omega)$ such that $d\nu_{\mathrm{F}} = \psi_{\mathrm{F}} dm_\Omega$. In order to prove Theorem 2.0.8, we use the smoothness condition from Lemma 4.3.2 for ψ_F. Indeed, for π_ℓ the projection from $L^2(K)$ to V_ℓ, it suffices to show

$$||\pi_\ell \psi_F||_2 \leq 2^{-(s+1)\ell}$$

for $s > m + \frac{1}{2}\dim K$ and ℓ large enough.

By the characterization of the Furstenberg measure, for any $n \geq 1$ it holds that $\nu_{\mathrm{F}} = \mu^{*n} * \nu_{\mathrm{F}}$ and therefore for $\varphi \in L^2(K)$,

$$\begin{aligned} |\langle \psi_{\mathrm{F}}, \varphi \rangle| &= \left| \int \varphi \, d\nu_{\mathrm{F}} \right| \\ &= \left| \int \int \varphi(g.k) \, d\mu^{*n}(g) d\nu_{\mathrm{F}}(k) \right| \\ &\leq \left|\left| \int \varphi \circ \alpha_g \, d\mu^{*n}(g) \right|\right|_\infty . \end{aligned} \tag{6.4.1}$$

We thus study the L^∞-norm of the function

$$\Phi_n = T_0^n \varphi = \int \varphi \circ \alpha_g \, d\mu^{*n}(g).$$

We will use Corollary 6.0.4 to give L^2-estimates of Φ_n. In order to convert these estimates to an L^∞-bound, we use Agmon's inequality (cf. [Agm65] chapter 13), which we introduce for compact Lie groups.

Lemma 6.4.2. *(Agmon's Inequality for Compact Lie Groups). Let K be a compact Lie group. Then there is $t \in \mathbb{Z}_{\geq 2}$ depending on K such that for any $\varphi \in C^\infty(K)$,*

$$||\varphi||_\infty \ll ||\varphi||_2^{1/2} ||\varphi||_{H^t}^{1/2}.$$

Proof. For $M \in \mathbb{R}_{>0}$ to be determined, we group together the contribution of the representations with $||\gamma|| \leq M$ and $||\gamma|| > 0$. Indeed, by (3.1.4), for $k \in K$,

$$\begin{aligned}
\varphi(k) &= \sum_{\gamma \in \overline{C} \cap I^*} \sum_{i,j=1}^{d_\gamma} d_\gamma^{1/2} a_{ij}^\gamma \chi_{ij}^\gamma(k) \\
&= \sum_{||\gamma|| \leq M} \sum_{i,j=1}^{d_\gamma} d_\gamma^{1/2} a_{ij}^\gamma \chi_{ij}^\gamma(k) + \sum_{||\gamma|| > M} \sum_{i,j=1}^{d_\gamma} d_\gamma^{1/2} a_{ij}^\gamma \chi_{ij}^\gamma(k) \\
&= \sum_{||\gamma|| \leq M} \sum_{i,j=1}^{d_\gamma} d_\gamma^{1/2} a_{ij}^\gamma \chi_{ij}^\gamma(k) + \sum_{||\gamma|| > M} \sum_{i,j=1}^{d_\gamma} \lambda_\gamma^t \lambda_\gamma^{-t} d_\gamma^{1/2} a_{ij}^\gamma \chi_{ij}^\gamma(k),
\end{aligned}$$

where in the last line we multiplied the second term by $1 = \lambda_\gamma^t \lambda_\gamma^{-t}$ for some $t \in \mathbb{Z}_{\geq 0}$. By Cauchy-Schwarz and using Lemma 4.3.1, the first term can be bounded by $||\varphi||_2 \sqrt{\sum_{||\gamma|| \leq M, i,j} d_\gamma} \ll M^C ||\varphi||_2$, where C is a constant depending on K. For the second term, we choose t large enough such that $\sqrt{\sum_{||\gamma|| > M, i,j} \lambda_\gamma^{-2t} d_\gamma} \ll M^{-C}$. Again using Cauchy-Schwarz, the second term is bounded by $M^{-C} ||\varphi||_{H^t}$. The claim is implied by setting $M = (\frac{||\varphi||_{H^t}}{||\varphi||_2})^{1/2C}$. □

Lemma 6.4.3. *For $\varphi \in V_\ell$ set $\Phi_n = T_0^n \varphi$. Let $\gamma \in \overline{C} \cap I^*$ and $r \in \mathbb{Z}_{\geq 1}$. Then it holds for $\widehat{\Phi_n}(\gamma) = \pi_\gamma(\Phi_n)$,*

$$||\widehat{\Phi_n}(\gamma)||_{\text{op}} \ll_r 2^{O(1)\ell - r\ell} (1+\varepsilon)^{O(1)nr} ||\gamma||^{O(1)r} ||\varphi||_2.$$

Proof. Let $v_1, \ldots, v_{d_\gamma}$ be an orthonormal basis of π_γ. Then

$$\begin{aligned}
||\widehat{\Phi_n}(\gamma)||_{\text{op}} &\leq d_\gamma \sup_{1 \leq i \leq d_\gamma} ||\widehat{\Phi}(\gamma) v_i|| \\
&\leq d_\gamma \sup_{1 \leq i,j \leq d_\gamma} |\langle \widehat{\Phi}(\gamma) v_i, v_j \rangle| \\
&= d_\gamma \sup_{1 \leq i,j \leq d_\gamma} |\langle \Phi_n, \chi_{ij}^\gamma \rangle| \\
&\leq d_\gamma \sup_{\substack{g \in \text{supp}(\mu^{*n}) \\ 1 \leq i,j \leq d_\gamma}} |\langle \varphi \circ \alpha_g, \chi_{ij}^\gamma \rangle|. \qquad (6.4.2)
\end{aligned}$$

Notice further that for $g \in B_\varepsilon$ and a further $\gamma' \in \overline{C} \cap I^*$ and $1 \leq i', j' \leq d_\gamma$,

$$\begin{aligned}
|\langle \chi^{\gamma'}_{i'j'} \circ \alpha_g, \chi^{\gamma}_{ij} \rangle| &= \frac{\lambda^r_\gamma}{\lambda^r_\gamma} |\langle \chi^{\gamma'}_{i'j'} \circ \alpha_g, \chi^{\gamma}_{ij} \rangle| \\
&= \frac{1}{\lambda^r_\gamma} |\langle \chi^{\gamma'}_{i'j'} \circ \alpha_g, \lambda_K(\triangle)^r \chi^{\gamma}_{ij} \rangle| \\
&\leq \frac{1}{\lambda^r_\gamma} \sum_{i_1,\ldots,i_r} |\langle \lambda(-X_{i_1}) \cdots \lambda(-X_{i_r}) \chi^{\gamma'}_{i'j'} \circ \alpha_g, \lambda(X_{i_1}) \cdots \lambda(X_{i_r}) \chi^{\gamma}_{ij} \rangle| \\
&\ll_r \lambda^{-r}_\gamma (1+\varepsilon)^{O(1)nr} ||\gamma'||^r \, ||\gamma||^r \\
&\ll_r (1+\varepsilon)^{O(1)nr} ||\gamma'||^r \, ||\gamma||^{-r}
\end{aligned}$$

where for the penultimate line one argues as in Lemma 6.3.1 and in the last line we use Lemma 4.3.1. Similarly, it holds that $|\langle \chi^{\gamma'}_{i'j'} \circ \alpha_g, \chi^{\gamma}_{ij} \rangle| \ll_r (1+\varepsilon)^{O(1)nr} ||\gamma'||^{-r} \, ||\gamma||^r$. Then using the decomposition

$$\varphi = \sum_{2^{\ell-1} \leq ||\gamma'|| < 2^\ell} \sum_{i',j'=1}^{d_\gamma} d^{1/2}_{\gamma'} a^{\gamma'}_{i'j'} \chi^{\gamma'}_{i'j'}$$

we conclude

$$\begin{aligned}
|\langle \varphi \circ \alpha_g, \chi^{\gamma}_{ij} \rangle| &\leq \sum_{\gamma',i',j'} d^{1/2}_{\gamma'} |a^{\gamma'}_{i'j'}| \, |\langle \chi^{\gamma'}_{i'j'} \circ \alpha_g, \chi^{\gamma}_{ij} \rangle| \\
&\leq 2^{O(1)\ell} ||\varphi||_2 \sup_{\gamma',i',j'} |\langle \chi^{\gamma'}_{i'j'} \circ \alpha_g, \chi^{\gamma}_{ij} \rangle| \\
&\ll_r 2^{O(1)\ell - r\ell} (1+\varepsilon)^{O(1)nr} ||\gamma||^r ||\varphi||_2.
\end{aligned}$$

This implies the claim by (6.4.2) and using Lemma 4.3.1. □

Proof. (of Theorem 2.0.8) Let $\varphi \in V_\ell$ be of unit norm and write $\Phi_n = T^n_0 \varphi$. It suffices to prove for $\varepsilon < \varepsilon_m$ and some some $n \geq 1$ that

$$||\Phi_n||_\infty \leq 2^{-(s+1)\ell}, \tag{6.4.3}$$

where s is a constant depending on G and m. Indeed, if (6.4.3) holds, then by (6.4.1),

$$||\pi_\ell \psi_F||_2 \ll 2^{O(1)\ell} 2^{-(s+1)\ell},$$

which satisfies the smoothness condition from Lemma 4.3.2 for s large enough depending on G and m.

We will use Agmon's inequality to prove (6.4.3). Notice first that for the fixed $t \in \mathbb{Z}_{\geq 2}$ from Lemma 6.4.2,

$$\begin{aligned} ||\Phi_n||_{H^t} &= ||\lambda_K(\triangle)^{t/2}\Phi_n||_2 \\ &\leq \sup_{g \in \mathrm{supp}\mu^{*n}} ||\lambda_K(\triangle)^{t/2}(\varphi \circ \alpha_g)||_\infty \\ &\leq ||\lambda_K(\triangle)^{t/2}\varphi||_\infty (1+\varepsilon)^{O(1)n} \leq 2^{A\ell}, \end{aligned}$$

for a constant A depending only on t and where we choose $n = \frac{1}{10E_2\varepsilon}\ell$ for $E_2 \geq 1$ a fixed constant to be determined later.

We next bound $||\Phi_n||_2$. In order to do so, we decompose Φ_n into a low and high frequency part:

$$\Phi_n = \Phi_n^{(1)} + \Phi_n^{(2)} \quad \text{where} \quad \Phi_n^{(1)} = \sum_{||\gamma|| \leq L(c_1,c_1)} \sum_{i,j}^{d_\gamma} d_\gamma^{1/2} \widehat{(\Phi_n)}_{ij}^{\gamma} \chi_{ij}^{\gamma}.$$

Then for $n \geq 1$, exploiting Corollary 6.0.4

$$\begin{aligned} ||\Phi_n||_2 &\leq \left|\left| \int \Phi_{n-1}^{(1)} \circ \alpha_g \, d\mu(g) \right|\right|_\infty + \left|\left| \int \Phi_{n-1}^{(2)} \circ \alpha_g \, d\mu(g) \right|\right|_2 \\ &\leq ||\Phi_{n-1}^{(1)}||_\infty + \frac{1}{2}||\Phi_{n-1}^{(2)}||_2 \end{aligned} \tag{6.4.4}$$

Using Lemma 6.4.3, it follows for all $m \leq n$ and $r \geq 1$,

$$||\Phi_m^{(1)}||_\infty \ll_r 2^{O(1)\ell - r\ell}(1+\varepsilon)^{O(1)nr} L(c_1,c_2)^{O(1)r}||\varphi||_2.$$

Iterating (6.4.4), there are absolute constants $E_1, E_2, E_3 \geq 1$ such that

$$||\Phi_n||_2 \ll_r (n 2^{E_1\ell - r\ell}(1+\varepsilon)^{E_2 nr} L(c_1,c_2)^{E_3 r} + 2^{-n})||\varphi||_2.$$

By Lemma 6.4.2, it therefore follows that

$$||\Phi_n||_\infty \ll_r (n 2^{(E_1+A)\ell - r\ell}(1+\varepsilon)^{E_2 nr} L(c_1,c_2)^{E_3 r} + 2^{-n})||\varphi||_2.$$

Setting the parameters suitably, the proof is concluded. Indeed, choose for instance

$$r = 2(s+1) + E_1 + A + 100$$

and $n = \frac{1}{10E_2\varepsilon}\ell$. For s large enough and choosing ε small enough in terms of r and s the claim (6.4.3) holds for large ℓ (depending on s and ε). □

Part II

Absolute Continuity of Self-Similar Measures

Chapter 7

Introduction to Part II

In the study of self-similar measures it is fundamental to determine their dimension and to find conditions for absolute continuity. For the former problem progress was made by Hochman [Hoc14], [Hoc17] (Theorem 1.4.2), relating the dimension of a self-similar measure to the entropy and Lyapunov exponent provided the generating measure satisfies a mild separation condition. While it was shown by Saglietti-Shmerkin-Solomyak [SSS18] that, under suitable assumptions, generic one-dimensional self-similar measures are absolutely continuous, finding explicit examples remains challenging. It was shown by Varjú [Var19a] (Theorem 1.4.6) that Bernoulli convolutions are absolutely continuous if their defining parameter is sufficiently close to 1 in terms of the Mahler measure. In dimension $d \geq 3$, assuming that the rotation part of the self-similar measure is fixed and has a spectral gap on $L^2(O(d))$, Lindenstrauss-Varjú [LV16] (Theorem 1.4.8) showed absolute continuity if all of the contraction rates are sufficiently close to 1. In this thesis we strengthen and vastly generalise these two results. Moreover, we give the first explicit examples of absolutely continuous self-similar measures in dimension one and two with non-uniform contraction rates. For instance consider for $x \in \mathbb{R}$ the similarities

$$g_1(x) = \frac{n}{n+1}x \quad \text{and} \quad g_2(x) = \frac{n}{n+2}x + 1. \tag{7.0.1}$$

We then show that the self-similar measure of $\frac{1}{2}\delta_{g_1} + \frac{1}{2}\delta_{g_2}$ is absolutely continuous on $\mathbb{R}$ for any sufficiently large integer $n \geq 1$. Furthermore, our methods allow us to construct several classes of explicit absolutely continuous examples for $g_i(x) = \rho_i U_i x + b_i$ for $x \in \mathbb{R}^d$ in any dimension $d \geq 1$ as well as for every collection of orthogonal matrices U_i acting irreducibly on $\mathbb{R}^d$ and distinct vectors $b_i \in \mathbb{R}^d$, provided they all have algebraic entries.

Let $G = \mathrm{Sim}(\mathbb{R}^d)$ be the group of similarities on $\mathbb{R}^d$ and let $O(d)$ be the group of orthogonal $d \times d$ matrices. For each $g \in G$ there exists a scalar $\rho(g) > 0$, an orthogonal matrix $U(g) \in O(d)$ and a vector $b(g) \in \mathbb{R}^d$ such that $g(x) = \rho(g)U(g)x + b(g)$ for all $x \in \mathbb{R}^d$. A similarity is called contracting if $\rho(g) < 1$ and expanding when $\rho(g) > 1$.

The Lyapunov exponent of a probability measure μ on G is defined, whenever it exists, as

$$\chi_\mu = \mathbb{E}_{g\sim\mu}[\log \rho(g)].$$

Throughout Part II we use the following terminology.

Definition 7.0.1. *If $\chi_\mu < 0$, we call μ* ***contracting on average****. Moreover, if every $g \in \mathrm{supp}(\mu)$ is contracting, we say that μ is* ***contracting****. When $\chi_\mu < 0$ and there is $g \in \mathrm{supp}(\mu)$ such that $\rho(g) > 1$, then we call μ* ***contracting only on average****.*

It is well-known ([Hut81], [BE88], [BP92]) that when μ is a finitely supported contracting on average probability measure on G, then there exists a unique probability measure ν on $\mathbb{R}^d$ that is μ-stationary (i.e. ν satisfies $\mu * \nu = \nu$) and referred to as the self-similar measure of μ. Under these assumptions, it follows from the moment estimates of [GP16, Proposition 5.1] that ν has a polynomial tail decay in the sense that there exists some $\alpha = \alpha(\mu) > 0$ such that as $R \to \infty$,

$$\nu(x \in \mathbb{R}^d \,:\, |x| \geq R) \ll_\mu R^{-\alpha} \tag{7.0.2}$$

for an implied constant depending only on μ. The authors have given in [KK25d] an independent proof of (7.0.2) for contracting on average measures on arbitrary metric spaces.

Throughout we denote by ν the self-similar measure associated to μ. If μ is (only) contracting on average, we say that ν is a (only) contracting on average self-similar measure. Moreover, μ or respectively ν is called homogeneous if there are $r \in \mathbb{R}_{>0}$ and $U \in O(d)$ such that $r = \rho(g)$ and $U = U(g)$ for all $g \in \mathrm{supp}(\mu)$. When this is not the case, we say that μ and ν are inhomogeneous. A particular goal of this thesis is to give explicit examples of inhomogeneous as well as contracting only on average self-similar measures which are absolutely continuous.

To state our main result, we first discuss the Hausdorff dimension of ν, which is defined as

$$\dim \nu = \inf\{\dim E \,:\, E \subset \mathbb{R}^d \text{ measurable and } \nu(E) > 0\}$$

where $\dim E$ is the Hausdorff dimension of E. In order to state the landmark results by Hochman [Hoc14], [Hoc17], recall that the random walk entropy of a finitely supported measure μ is defined as

$$h_\mu = \lim_{n\to\infty} \frac{1}{n} H(\mu^{*n}) = \inf_{n\geq 1} \frac{1}{n} H(\mu^{*n}),$$

where $H(\cdot)$ is the Shannon entropy. Observe that if $\mathrm{supp}(\mu)$ has no exact overlaps, meaning that $\mathrm{supp}(\mu)$ generates a free semigroup, then $h_\mu = H(\mu) = -\sum_i p_i \log p_i$.

Moreover, as in [Hoc17], denote by $d(\cdot,\cdot)$ the metric on G defined for $g = \rho_1 U_1 + b_1$ and $h = \rho_2 U_2 + b_2$ as

$$d(g,h) = |\log \rho_1 - \log \rho_2| + ||U_1 - U_2|| + |b_1 - b_2| \tag{7.0.3}$$

for $||\cdot||$ the operator norm and $|\cdot|$ the euclidean norm.

To distinguish between the results for dimension and absolute continuity, denote

$$\Delta_n = \min\{d(g,h) \text{ for } g,h \in \mathrm{supp}(\mu^{*n}) \text{ with } g \neq h\}$$

and

$$M_n = \min\left\{ d(g,h) \text{ for } g,h \in \bigcup_{i=0}^{n} \mathrm{supp}(\mu^{*i}) \text{ with } g \neq h \right\}.$$

Furthermore we set

$$S_n = -\frac{1}{n}\log M_n \qquad \text{and} \qquad S_\mu = \limsup_{n\to\infty} S_n,$$

where S_μ is referred to as the splitting rate.

We call a subgroup H of $O(d)$ irreducible if H acts irreducibly on $\mathbb{R}^d$, i.e. the only H-invariant subspaces of $\mathbb{R}^d$ are $\{0\}$ and $\mathbb{R}^d$. Moreover, we say that a measure $\mu = \sum_{i=1}^n p_i \delta_{g_i}$ on G or $O(d) \subset G$ is irreducible if the group generated by $\{U(g_1), \ldots, U(g_n)\}$ is irreducible. When the elements in the support of μ have a common fixed point $x \in \mathbb{R}^d$, then δ_x is the self-similar measure of μ. To avoid the latter case, we say that μ has no common fixed point if the similarities in $\mathrm{supp}(\mu)$ do not.

It follows by Hochman [Hoc17], generalising [Hoc14], that if μ is a finitely supported, contracting and irreducible probability measure on G without a common fixed point such that $\Delta_n \geq e^{-cn}$ for some $c > 0$ and infinitely many $n \geq 1$, then $\dim \nu = \min\{d, \frac{h_\mu}{|\chi_\mu|}\}$.

In the paper [KK25a] we use the techniques of Part II of this thesis to generalise Hochman's result to contracting on average measures. Moreover, we show that a weaker requirement than exponential separation at all scales is sufficient (see [KK25a] for a discussion). We work with M_n instead of Δ_n for convenience only and in order to apply the general entropy gap results from [KK25b].

Theorem 7.0.2. *(*[KK25a, Theorem 1.2 and Theorem 1.3]*) Let μ be a finitely supported, contracting on average and irreducible probability measure on G without a common fixed point. Assume that either of the following two properties holds:*

(i) For some $c > 0$, $M_n \geq e^{-cn}$ *for infinitely many* $n \geq 1$,

(ii) For some $\varepsilon > 0$, $\log M_n \geq -n \exp((\log n)^{1/3-\varepsilon})$ *for all sufficiently large* $n \geq 1$.

Then

$$\dim \nu = \min \left\{ d, \frac{h_\mu}{|\chi_\mu|} \right\}.$$

It is well-established that $\dim \nu \leq \{d, \frac{h_\mu}{|\chi_\mu|}\}$. Therefore ν can only be absolutely continuous if $h_\mu \geq d\,|\chi_\mu|$. The following general conjecture is expected to hold.

Conjecture 7.0.3. *Let* μ *be a finitely supported, contracting on average and irreducible probability measure on* G *without a common fixed point. Then* ν *is absolutely continuous if*

$$\frac{h_\mu}{|\chi_\mu|} > d.$$

We observe that the latter conjecture is completely open and is not known for any class of self-similar measures. Our main result establishes a weakening of the latter conjecture. Indeed, when the $O(d)$-part of our measure μ is fixed, we show Conjecture 7.0.3 with the d being replaced by a constant depending on the $O(d)$-part as well as the logarithmic separation rate $\log S_\mu$. Given a measure μ on G we denote by $U(\mu)$ the pushforward of μ under the map $g \mapsto U(g)$. We first state a version of our main theorem for contracting measures.

Theorem 7.0.4. *Let* $d \geq 1$ *and* $\varepsilon \in (0, 1)$. *Given an irreducible probability measure* μ_U *on* $O(d)$ *there exist constants* $C \geq 1$ *and* $\tilde{\rho} \in (0, 1)$ *depending on* d, ε *and* μ_U *such that the following holds. Let* $\mu = \sum_{i=1}^{k} p_i \delta_{g_i}$ *be a contracting probability measure on* G *without a common fixed point satisfying* $U(\mu) = \mu_U$ *and* $p_i \geq \varepsilon$ *as well as* $\rho(g_i) \in (\tilde{\rho}, 1)$ *for all* $1 \leq i \leq k$. *Then the self-similar measure* ν *is absolutely continuous if*

$$\frac{h_\mu}{|\chi_\mu|} > C \left(\max \left\{ 1, \log \frac{S_\mu}{h_\mu} \right\} \right)^2.$$

Theorem 7.0.4 is a special case of the more general Theorem 8.1.4, which requires a few new definitions we state in Section 8.1. When $d = 1$ we note that every probability measure on $O(1)$ is irreducible. We further observe that while Theorem 7.0.4 applies in the case when the spectral gap of μ_U is zero, the dependence of C and $\tilde{\rho}$ can be made more explicit in the presence of a spectral gap. To introduce notation, given a closed subgroup $H \subset G$ and assuming that μ_U is a probability measure on $O(d)$ with $\mathrm{supp}(\mu_U) \subset H$, we denote by $\mathrm{gap}_H(\mu_U)$ the L^2-spectral gap of μ_U in H as defined in (8.3.4).

Theorem 7.0.5. *Let d, ε, μ_U and μ be as in Theorem 7.0.4. Assume further that $\mathrm{gap}_H(\mu_U) \geq \varepsilon > 0$ for H the closure of the subgroup generated by the support of μ_U. Then there exists $C \geq 1$ and $\tilde{\rho} \in (0,1)$ only depending on d and ε such that the conclusion of Theorem 7.0.4 holds.*

We point out that in Theorem 7.0.5 the constants are independent of the subgroup H and the statement applies when H is a finite irreducible subgroup of $O(d)$ as well as when H is a positive dimensional irreducible Lie subgroup of $O(d)$. As is shown in section 14, this observation relies on uniform convergence of μ_U^{*n} towards the Haar probability measure m_H and on Schur's lemma implying that $\mathbb{E}_{h\sim m_H}[|x \cdot hy|^2] = d^{-1}$ for any unit vectors $x, y \in \mathbb{R}^d$ and any irreducible subgroup $H \subset O(d)$.

To construct explicit examples of absolutely continuous self-similar measures on $\mathbb{R}^d$, Theorem 7.0.4 requires us to estimate $h_\mu, |\chi_\mu|$ and S_μ. It is straightforward to deal with $|\chi_\mu|$ as it can be explicitly computed. Lower bounds on the random walk entropy follow in many cases (see Section 15.1) by the ping-pong lemma or Breuillard's strong Tits alternative [Bre08]. It also holds that $h_{U(\mu)} \leq h_\mu$, so when $h_{U(\mu)} > 0$, we only need to control $|\chi_\mu|$ and S_μ. With current methods we can usually only bound S_μ if all of the coefficients of the elements in the support of μ are algebraic. In the latter case, as shown in Section 15.2, when all of the coefficients of elements in the support of μ lie in a number field K and have logarithmic height at most L (see (7.0.5)), then $S_\mu \ll_d L \cdot [K : \mathbb{Q}]$. We observe that $\log S_\mu$ is usually very small as it is double logarithmic in the arithmetic complexity of the coefficients. All this information makes it straightforward to find explicit examples of absolutely continuous self-similar measures. The constants C and $\tilde{\rho}$ in Theorem 7.0.4 can be computed from the involved terms, yet we do not make the dependence explicit in this work.

The proof of Theorem 7.0.4 and Theorem 8.1.4 builds on new techniques initiated by Samuel Kittle in [Kit23] and further developed in this part of the thesis, while being inspired by ideas from [Hoc14], [Hoc17], [Var19a] and [Kit21]. We give an outline of our proof in Section 8.2 and note that the main novelties exploited are strong product bounds for detail at scale r (a notion introduced in [Kit21]) and a decomposition theory for stopped random walks to capture the amount of variance we can gain at a given scale, a technique we call the variance summation method. [Kit23] is concerned with constructing absolutely continuous Furstenberg measures of $\mathrm{SL}_2(\mathbb{R})$ on 1-dimensional projective space $\mathbb{P}^1(\mathbb{R}) = \mathbb{R}^2/\sim$ and an analogue of Theorem 8.1.4 is shown. However, we currently can't deduce a result similar to Theorem 7.0.4 for Furstenberg measures of $\mathrm{SL}_2(\mathbb{R})$ as the dynamics of the $\mathrm{SL}_2(\mathbb{R})$ action on $\mathbb{P}^1(\mathbb{R})$

are more difficult to control than the one of the $\mathrm{Sim}(\mathbb{R}^d)$ action on $\mathbb{R}^d$. Indeed, we exploit that one can rescale and translate self-similar measures without changing the Lyapunov exponent, the separation rate, the random walk entropy or the spectral gap of the generating measure. Moreover, an analogue of Theorem 8.1.4 as well as Theorem 7.0.2 for Furstenberg measures of arbitrary dimensions is presently out of reach since the current methods cannot deal with non-conformal measures.

To also treat contracting on average measures, we state the following version of Theorem 7.0.4. We require some control on the scaling rate of the expanding similarities.

Theorem 7.0.6. *Let d and μ_U be as in Theorem 7.0.4 and let $R > 1$ and $\varepsilon > 0$. Let $\mu = \sum_{i=1}^k p_i \delta_{g_i}$ be a contracting on average probability measure on G without a common fixed point satisfying $U(\mu) = \mu_U$ and $p_i \geq \varepsilon$ as well as $\rho(g_i) \in [R^{-1}, R]$ for all $1 \leq i \leq k$. Then there is some $\tilde{\rho} \in (0,1)$ and $C > 1$ depending on d, R, ε and μ_U such that the conclusion of Theorem 7.0.4 holds provided that for some $\hat{\rho} \in (\tilde{\rho}, 1)$ we have*

$$\frac{\mathbb{E}_{\gamma\sim\mu}[|\hat{\rho} - \rho(\gamma)|]}{1 - \mathbb{E}_{\gamma\sim\mu}[\rho(\gamma)]} < 1 - \varepsilon.$$

In the presence of a spectral gap, the analogue of Theorem 7.0.5 also holds for Theorem 7.0.6. Using Theorem 7.0.4, Theorem 7.0.6 and Theorem 8.1.4 one can construct a versatile collection of explicit absolutely continuous self-similar measures. We give a few cases below and encourage the reader to find further examples. Indeed, as shown in Corollary 7.0.8 and Corollary 7.0.9, for any given irreducible probability measure μ_U on $O(d)$ supported on matrices with algebraic entries and algebraic vectors $b_1, \ldots, b_k$ with $b_1 \neq b_2$, we can find explicit contracting as well as contracting only on average measures $\mu = \sum_{i=1}^k p_i \delta_{g_i}$ on G with $U(\mu) = \mu_U$ and $b(g_i) = b_i$ for $1 \leq i \leq k$ and having absolutely continuous self-similar measure.

Inhomogeneous Self-Similar Measures in Dimension 1

As a first example, we present results for self-similar measures supported on two similarities in dimension one. Upon conjugating, we can assume without loss of generality that our generating measure is supported on $x \mapsto \lambda_1 x$ and $x \mapsto \lambda_2 x + 1$ for $\lambda_1, \lambda_2 \in (0,1)$.

We recall the definition of the height of algebraic numbers, which measures the arithmetic complexity. For a number field K and an algebraic number $\alpha \in K$ one defines the absolute height as

$$\mathcal{H}(\alpha) = \left(\prod_{v \in M_K} \max(1, |\alpha|_v)^{n_v} \right)^{1/[K:\mathbb{Q}]} \tag{7.0.4}$$

where M_K is the set of places of K, $n_v = [K_v : \mathbb{Q}_v]$ is the local degree at v and $|\cdot|_v$ is the absolute value associated with the place v. We refer to [Mas16] for basic properties of heights and note that the height of α is independent of the number field K. We will also work with the logarithmic height

$$h(\alpha) = \log \mathcal{H}(\alpha). \tag{7.0.5}$$

Corollary 7.0.7. *For every $\varepsilon > 0$ there exists a small constant $c = c(\varepsilon) > 0$ such that the following holds. Let K be a number field and $\lambda_1, \lambda_2 \in K \cap (0,1)$ and write $h(\lambda_1, \lambda_2) = \max\{h(\lambda_1), h(\lambda_2)\}$. Consider the similarities given for $x \in \mathbb{R}$ as*

$$g_1(x) = \lambda_1 x \qquad \text{and} \qquad g_2(x) = \lambda_2 x + 1.$$

Then the self-similar measure of $\frac{1}{2}\delta_{g_1} + \frac{1}{2}\delta_{g_2}$ is absolutely continuous if

$$h(\lambda_1, \lambda_2) \geq \varepsilon \qquad \text{and} \qquad |\chi_\mu| \max\{1, \log([K:\mathbb{Q}]h(\lambda_1, \lambda_2))\}^2 < c.$$

Concretely, generalising the example discussed in (7.0.1), if $\lambda_i = 1 - p_i/q_i$ is rational for $i \in \{1, 2\}$ with coprime integers $p_i, q_i \geq 1$ then the self-similar measure of $\frac{1}{2}\delta_{g_1} + \frac{1}{2}\delta_{g_2}$ is absolutely continuous if for $i \in \{1, 2\}$,

$$p_i \frac{(\log \log q_i)^2}{q_i} \leq c.$$

Corollary 7.0.7 can be viewed as an inhomogeneous version of our strengthening of Varjú's result for Bernoulli convolutions (Corolarry 7.0.11), yet with an additional dependence on the number field K and on the lower bound of $\max\{h(\lambda_1), h(\lambda_2)\}$. We further note that Lehmer's conjecture states the existence of an absolute $\varepsilon_0 > 0$ such that $\max\{h(\lambda_1), h(\lambda_2)\} \geq \varepsilon_0/[K:\mathbb{Q}]$ for all $\lambda_1, \lambda \in K$ for any number field K.

It is straightforward to adapt Corollary 7.0.7 to multiple maps and also to include contracting on average measures. We next discuss such examples in arbitrary dimensions.

Self-similar measures on $\mathbb{R}^d$

With Theorem 7.0.4 and Theorem 7.0.6 numerous explicit classes of absolutely continuous self-similar measures in $\mathbb{R}^d$ can be constructed. In order to apply these results we need to estimate h_μ. In the following examples we have used the ping-pong lemma (see section 15) in two ways in order to establish lower bounds on h_μ. For the first class of examples we have applied p-adic ping-pong as in Lemma 15.1.4.

Corollary 7.0.8. *Let $d \geq 1$ and $\varepsilon > 0$, let $\mu_U = \sum_{i=1}^k p_i \delta_{U_i}$ be an irreducible probability measure on $O(d)$ with $p_i \geq \varepsilon$ and let $b_1, \dots, b_k \in \mathbb{R}^d$ with $b_1 \neq b_2$. Assume that $U_1, \dots, U_k$ and $b_1, \dots, b_k$ have algebraic coefficients. Let q be a prime number and for $1 \leq i \leq k$ consider*

$$g_i(x) = \frac{q}{q + a_{i,q}} U_i x + b_i \qquad \textit{for any integer} \qquad a_{i,q} \in [1, q^{1-\varepsilon}].$$

Assume that $g_1, \dots, g_k$ do not have a common fixed point and consider $\mu = \sum_{i=1}^k p_i \delta_{g_i}$. Then the self-similar measure of μ is absolutely continuous for q a sufficiently large prime depending on $d, \varepsilon, U_1, \dots, U_k$ and $b_1, \dots, b_k$.

We point out that any choice of integers $a_{i,q}$ works and that the necessary size of q to derive absolute continuity does not depend on this choice, leading to a vast number of examples. Moreover, we can adapt Corollary 7.0.8 to give contracting only on average examples. In order to satisfy the assumption from Theorem 7.0.6, we require that $\mu = \sum_{i=1}^k p_i \delta_{g_i}$ satisfies that $p_k \leq \frac{1}{3}$. This nonetheless leads to absolutely continuous examples with $U(\mu) = \mu_U$ for any given irreducible probability measure $\mu_U = \sum_{i=1}^k p_i \delta_{U_i}$ on $O(d)$ as we do not require that the U_i are distinct.

Corollary 7.0.9. *Let d, ε and $\mu_U = \sum_{i=1}^k p_i \delta_{U_i}$ as well as $b_1, \dots, b_k$ be as in Proposition 7.0.8. Let q be a prime number and consider for $1 \leq i \leq k-1$*

$$g_i(x) = \frac{q}{q+3} U_i x + b_i \qquad \textit{and} \qquad g_k(x) = \frac{q}{q-1} U_k x + b_k.$$

Assume that $g_1, \dots, g_k$ do not have a common fixed point and further that

$$p_k \leq \frac{1}{3}.$$

Then the self-similar measure of $\mu = \sum_{i=1}^k p_i \delta_{g_i}$ is absolutely continuous for q a sufficiently large prime depending on $d, \varepsilon, U_1, \dots, U_k$ and $b_1, \dots, b_k$.

We give a second class of examples that rely on Galois ping-pong in as Lemma 15.1.4.

Corollary 7.0.10. *Let $d \geq 1$ and $\varepsilon \in (0,1)$ and $\mu_U = \sum_{i=1}^{k} p_i \delta_{U_i}$ an irreducible probability measure on $O(d)$ with $p_i \geq \varepsilon$ for all $1 \leq i \leq k$. Assume furthermore that $U_1, \ldots, U_k$ have algebraic entries. Let $\tilde{\rho} \in (0,1)$ be sufficiently close to 1 in terms of d, ε and μ_U and let $C > 1$ be sufficiently large depending on the same parameters.*

Suppose that $g_i(x) = \frac{a_i + b_i\sqrt{q}}{c_i} U_i x + d_i$ with $a_i, b_i, c_i \in \mathbb{Z}$ and $d_i \in \mathbb{Z}^d$ for $1 \leq i \leq k$ and a prime number q do not have a common fixed point. Then the self-similar measure associated to $\mu = \sum_{i=1}^{k} p_i \delta_{g_i}$ is absolutely continuous if the following properties are satisfied:

(i) $\frac{a_i + b_i\sqrt{q}}{c_i} \in (\tilde{\rho}, 1)$ *for* $1 \leq i \leq k$,

(ii) for $j = 1$ and for $j = 2$ we have

$$\left| \frac{a_j - b_j\sqrt{q}}{c_j} \right| < \frac{1}{3},$$

(iii) For $L = \max(\sqrt{q}, |a_i|, |b_i|, |c_i|, |d_i|_\infty)$ we have

$$C|\chi_\mu| \leq \frac{1}{(\log(\log L))^2}.$$

As a particular case of Corollary 7.0.10, we can consider as shown in Lemma 15.4.2 the maps

$$g_i(x) = \frac{\lceil \sqrt{q} \rceil - m_{i,q} + 2\sqrt{q}}{3\lceil \sqrt{q} \rceil} U_i x + d_i$$

for any $m_{i,q} \in \mathbb{Z}$ and $d_i \in \mathbb{Z}^d$ satisfying for some $\varepsilon > 0$ that

$$m_{i,q} \in [0, q^{1/2-\varepsilon}] \qquad \text{and} \qquad |d_i|_\infty \leq \exp(\exp(q^{\varepsilon/3})).$$

Then the self-similar measure of $\mu = \sum_{i=1}^{n} p_i \delta_{g_i}$ is absolutely continuous for sufficiently large primes q depending on d, μ_U and ε, provided that $g_1, \ldots, g_k$ do not have a common fixed point. We note that since we have a double exponential range for d_i, we get abundantly many examples.

Real and Complex Bernoulli Convolutions

While Theorem 7.0.4 applies to arbitrary self-similar measures, it gives new results for Bernoulli convolutions. Let $\lambda \in (1/2, 1)$ and denote by ν_λ the unbiased Bernoulli convolution of parameter λ, i.e. the law of the random variable $\sum_{n=0}^{\infty} \xi_n \lambda^n$ with $\xi_0, \xi_1, \ldots$ independent Bernoulli random variables with $\mathbb{P}[\xi_i = 1] = \mathbb{P}[\xi_i = -1] = 1/2$. It was shown by Solomyak [Sol95] that for almost all $\lambda \in (1/2, 1)$ the Bernoulli convolution ν_λ has a density in $L^2(\mathbb{R})$, while Erdős [Erd39] proved that ν_λ is singular whenever λ^{-1} is a Pisot number.

The Mahler measure of an algebraic number λ is defined as

$$M_\lambda = |a| \prod_{|z_j|>1} |z_j|$$

with $a(x-z_1)\cdots(x-z_\ell)$ the minimal polynomial of λ over $\mathbb{Z}$. We note that as in Corollary 5.9 of [Kit23] it holds that

$$S_{\nu_\lambda} \le \log M_\lambda. \tag{7.0.6}$$

Garsia [Gar62, Theorem 1.8] showed that ν_λ is absolutely continuous for algebraic λ with $M_\lambda = 2$, while Samuel Kittle [Kit21] established that ν_λ is absolutely continuous if $M_\lambda \approx 2$. In landmark work, Varjú [Var19a] (Theorem 1.4.6) proved for every $\varepsilon > 0$ there is a constant $C_\varepsilon > 1$ such that that ν_λ is absolutely continuous if

$$\lambda > 1 - C_\varepsilon^{-1} \min\{\log M_\lambda, (\log M_\lambda)^{-1-\varepsilon}\}. \tag{7.0.7}$$

When applying Theorem 7.0.4 to Bernoulli convolutions we deduce the following strengthening of (7.0.7), exploiting the comparison between the entropy and the Mahler measure for Bernoulli convolution due to [BV20].

Corollary 7.0.11. *There is an absolute constant $C > 1$ such that the following holds. Let $\lambda \in (1/2, 1)$ be a real algebraic number. Then the Bernoulli convolution ν_λ is absolutely continuous on $\mathbb{R}$ if*

$$\lambda > 1 - C^{-1} \min\{\log M_\lambda, (\log\log M_\lambda)^{-2}\}. \tag{7.0.8}$$

We estimate that a direct application of our method would lead to $C \approx 10^{10}$ in Corollary 7.0.11. It would be an interesting further direction to try to optimise C for Bernoulli convolutions and in particular for the case $\lambda = 1 - \frac{1}{n}$.

Our most general result, Theorem 8.1.4, also applies to complex Bernoulli convolutions, which are defined analogously for $\lambda \in \mathbb{D} = \{\lambda \in \mathbb{C} : |\lambda| < 1\}$. When $|\lambda| \in (0, 2^{-1/2})$, then $\dim \nu_\lambda \le \frac{\log 2}{|\log \lambda|} < 2$ and ν_λ is singular to the Lebesgue measure on $\mathbb{C}$. It was shown by Shmerkin-Solomyak [SS16a] that the set of $\lambda \in \mathbb{C}$ with $|\lambda| \in (2^{-1/2}, 1)$ and ν_λ is singular has Hausdorff dimension zero, whereas Solomyak-Xu [SX03] showed that ν_λ is absolutely continuous on $\mathbb{C}$ for a non-real algebraic $\lambda \in \mathbb{D}$ with $M_\lambda = 2$. We extend Corollary 7.0.11 to complex parameters while assuming (15.6.1) in order to ensure that the rotation part of λ mixes fast enough and so that our measure is sufficiently non-degenerate (see section 8.1).

Corollary 7.0.12. *For every $\varepsilon > 0$ there is a constant $C \geq 1$ such that the following holds. Let $\lambda \in \mathbb{C}$ be a complex algebraic number such that $|\lambda| \in (2^{-1/2}, 1)$ and*

$$|\mathrm{Im}(\lambda)| \geq \varepsilon. \tag{7.0.9}$$

Then the Bernoulli convolution ν_λ is absolutely continuous on $\mathbb{C}$ if

$$|\lambda| > 1 - C^{-1} \min\{\log M_\lambda, (\log\log M_\lambda)^{-2}\}.$$

Dimension $d \geq 3$

Finally we discuss the case when $d \geq 3$. Under this assumption, $O(d)$ is a simple non-abelian Lie group and therefore instead of using the entropy and separation rate on G we can use the same quantities on $O(d)$.

We recall that Lindenstrauss-Varjú [LV16] (Theorem 1.4.8) proved the following. Given $d \geq 3$, $\varepsilon \in (0, 1)$ and a finitely supported probability measure μ_U on $SO(d)$, whose support generates a dense subgroup of $SO(d)$ and with $\mathrm{gap}_{SO(d)}(\mu_U) \geq \varepsilon$. Then there exists a constant $\widetilde{\rho} \in (0, 1)$ depending on d and ε such that every finitely supported contracting probability measure $\mu = \sum_{i=1}^k p_i \delta_{g_i}$ on G with $U(\mu) = \mu_U$ and

$$p_i \geq \varepsilon \quad \text{as well as} \quad \rho(g_i) \in (\tilde{\rho}, 1) \qquad \text{for all} \quad 1 \leq i \leq k \tag{7.0.10}$$

has absolutely continuous self-similar measure ν. Moreover, [LV16] show that ν has a C^k-density if the constant $\widetilde{\rho}$ is in addition sufficiently close to 1 in terms of k. As discussed in section 1.2, by current methods ([BG08], [BdS16]) spectral gap of $U(\mu)$ is only known when $\mathrm{supp}(U(\mu))$ generates a dense subgroup and all of the entries of elements in $\mathrm{supp}(U(\mu))$ are algebraic.

We note that $h_{U(\mu)} \leq h_\mu$ yet we do not have in general that $S_{U(\mu)} \geq S_\mu$. In the case when $S_{U(\mu)} \geq S_\mu$, which for example holds when the support of $U(\mu)$ generates a free group, (7.0.10) follows from Theorem 7.0.4. Moreover, our method can be adapted to work with $S_{U(\mu)}$ instead of S_μ and we establish a generalisation of (7.0.10) (in the case when $\mathrm{supp}(\mu_U)$ consists of matrices with algebraic coefficients) that we state in Theorem 8.1.5. We note that our method does not require that $\mathrm{supp}(\mu_U)$ generates a dense subgroup of $O(d)$ or $SO(d)$ and we can also treat contracting on average self-similar measures. Moreover, as shown in Corollary 7.0.8 and Corollary 7.0.10, we can also give examples when $\mathrm{supp}(\mu_U)$ generates a finite irreducible subgroup of $O(d)$.

Discussion of other work

In addition to the discussed above [Gar62], [SX03], [LV16], [Var19a] and [Kit21] there is little known about explicit examples of absolutely continuous self-similar measures. To the authors knowledge, the only other papers addressing this topic are [DFW07] and [Str24], which are concerned with homogeneous self-similar measures on $\mathbb{R}$ whose contraction rate λ satisfies that all of its Galois conjugates have absolute value < 1.

As exposed in section 1.4, a related problem is to study the Furstenberg measure of $\mathrm{SL}_2(\mathbb{R})$ or of arbitrary simple non-compact Lie groups. The first examples of explicit absolutely continuous Furstenberg measures arising from finitely supported generating measures were established by [Bou12] (Theorem 1.4.10), giving an intricate number theoretic construction and also providing examples with a C^k-density for any $k \geq 1$. Bourgain's methods were generalised and further used by [BISG17], [Leq22] and in Part I of this thesis. Moreover, numerous new examples were recently given by [Kit23] (Theorem 1.4.13).

Returning to self-similar measures, we observe that the behavior of generic self-similar measures on $\mathbb{R}$ or $\mathbb{C}$ is better understood. [Shm14] showed, thereby improving [Sol95], that the set of $\lambda \in (1/2, 1)$ such that the Bernoulli convolution ν_λ is singular has Hausdorff dimension zero. In [SSS18] it was shown that when the translation part (with distinct translations) and the probability vector is fixed, then generic one-dimensional self-similar measures on $\mathbb{R}$ are almost surely absolutely continuous in the range where the similarity dimension > 1. This was generalised to $\mathbb{C}$ by [SS23]. A further line of research is to show that certain parametrized families of self-similar measures or other types of invariant function systems are generically absolutely continuous, see for example [Hoc14], [Hoc17], [SS16b] and [BSSŚ22].

We finally mention that Fourier decay of self-similar measures was studied by numerous authors recently. The interested reader is referred to [LS20], [Bré21], [LS22], [Rap22], [Sol22], [VY22] and [BKS24] and as well as [ARHW21] and [BS23] for self-conformal measures.

Chapter 8

Main Result and Outline

In this section we first state our main results and give an outline of the proof of the main theorem in section 8.2. Then we collect for the convenience of the reader some notation used throughout Part II in section 8.3 and comment on the organisation in section 8.4.

8.1 Main Result

Let μ be a probability measure on $G = \mathrm{Sim}(\mathbb{R}^d)$. To state our main results in full generality we introduce notions that capture how well $U(\mu)$ mixes on $O(d)$ and how degenerate ν is.

Denote by $\gamma_1, \gamma_2, \ldots$ independent samples from μ, write $q_n := \gamma_1\gamma_2 \ldots \gamma_n$ and given $\kappa > 0$ let τ_κ be the stopping time defined by

$$\tau_\kappa := \inf\{n \geq 1 : \rho(q_n) \leq \kappa\}.$$

We then have the following definitions.

Definition 8.1.1. *Let μ be a probability measure on G generating a self-similar measure ν.*

*(i) We say that μ is (α_0, θ, A)-**non-degenerate** for $\alpha_0 \in (0,1)$ and $\theta, A > 0$ if for any proper subspace $W \subset \mathbb{R}^d$ and $y \in \mathbb{R}^d$,*

$$\nu(\{x \in \mathbb{R}^d \,:\, |x - (y + W)| < \theta \text{ or } |x| \geq A\}) \leq \alpha_0.$$

*(ii) We say that μ is (c, T)-**well-mixing** for $c \in (0,1)$ and $T \geq 0$ if there is some κ_0 such that for any $\kappa < \kappa_0$ and any unit vectors $x, y \in \mathbb{R}^d$ we have*

$$\mathbb{E}[|x \cdot U(q_{\tau_\kappa + F})y|^2] \geq c,$$

where F is a uniform random variable on $[0, T]$ which is independent of the γ_i.

For $d = 1$ our measure μ will always be $(1,1)$-well-mixing. As we show in section 14.1, when $U(\mu)$ is fixed and irreducible, there exists (c, T) depending only on $U(\mu)$ such that μ is (c, T)-well-mixing. This follows as $U(q_F) \to m_H$ in distribution as $T \to \infty$, where H is the closure of the subgroup generated by $\mathrm{supp}(U(\mu))$ and m_H the Haar probability measure on H. The latter would not be true if we fix F to be a deterministic random variable and therefore we have introduced the above definition.

Dealing with non-degeneracy is more involved and uniform results for many classes of self-similar measures do not hold. However, instead of our given measure we can consider a conjugated measure to establish uniform non-degeneracy results. Indeed, for $\mu = \sum_{i=1}^k p_i \delta_{g_i}$ a measure on G and $h \in G$ we denote

$$\mu_h = \sum_{i=1}^k p_i \delta_{h g_i h^{-1}} \quad \text{and} \quad \mu'_h = \frac{1}{2}\delta_e + \frac{1}{2}\sum_{i=1}^k p_i \delta_{h g_i h^{-1}}.$$

Then as we show in Lemma 14.2.1, absolute continuity of any of the self-similar measures of μ, μ_h or μ'_h is equivalent and all relevant quantities such as h_μ, S_μ and $|\chi_\mu|$ are the same or comparable.

Towards Theorem 7.0.4, Theorem 7.0.5 and Theorem 7.0.6, as we state in Proposition 8.1.2 and Proposition 8.1.3 we have essentially uniform (c, T)-mixing and uniform (α_0, θ, A)-non-degeneracy as long as we fix $U(\mu)$. We first state a uniform mixing result adapted for Theorem 7.0.4 and Theorem 7.0.5 in the contracting case.

Proposition 8.1.2. *Let $d \geq 1$, $\varepsilon \in (0,1)$ and let μ_U be an irreducible probability measure on $O(d)$. Then there exists $\tilde{\rho} \in (0,1)$, (c, T) and (α_0, θ, A) depending on d, ε and μ_U such that the following holds. Let $\mu = \sum_{i=1}^k p_i \delta_{g_i}$ be a contracting probability measure on G without a common fixed point and with $U(\mu) = \mu_U$ and*

$$p_i \geq \varepsilon \quad \text{as well as} \quad \rho(g_i) \in (\tilde{\rho}, 1) \qquad \text{for all} \quad 1 \leq i \leq k.$$

Then there is $h \in G$ such that $\mu'_h = \frac{1}{2}\delta_e + \frac{1}{2}\sum_{i=1}^k p_i \delta_{h g_i h^{-1}}$ is (c, T)-well-mixing and (α_0, θ, A)-non-degenerate.

Moreover, if $\mathrm{gap}_H(\mu_U) \geq \varepsilon > 0$ for H the closure of the subgroup generated by the support of μ_U, then there exist (c, T) and (α_0, θ, A) depending only on d and ε such that the above conclusion holds.

For Theorem 7.0.6 we state a similar result for contracting on average measures.

Proposition 8.1.3. *Let d and μ_U be as in Theorem 8.1.2 and let $\varepsilon > 0$. Let $\mu = \sum_{i=1}^k p_i \delta_{g_i}$ be a contracting on average probability measure on G without a common* *fixed point satisfying* $U(\mu) = \mu_U$ *and* $p_i \geq \varepsilon$ *for* $1 \leq i \leq k$. *Then there is some* $\tilde{\rho} \in (0,1)$ *and* $C > 1$ *depending on* d, ε *and* μ_U *such that the following holds.*

The conclusion of Proposition 8.1.2 holds provided that for some $\hat{\rho} \in (\tilde{\rho}, 1)$ *we have*

$$\frac{\sum_{i=1}^{k} |\hat{\rho} - \rho(g_i)|}{k - \sum_{i=1}^{k} \rho(g_i)} < 1 - \varepsilon.$$

Proposition 8.1.2 and Proposition 8.1.3 are proved in section 14. We are now in a suitable position to state our main result. Theorem 7.0.4,Theorem 7.0.5 and Theorem 7.0.6 follow from the main result Theorem 8.1.4 by applying Proposition 8.1.2 and Proposition 8.1.3 as well as Lemma 14.2.1.

Theorem 8.1.4. *For every* $d \in \mathbb{Z}_{\geq 1}$ *and* $R, c, T, \alpha_0, \theta, A > 0$ *with* $c, \alpha_0 \in (0,1)$ *and* $T \geq 0$ *there is a constant* $C = C(d, R, c, T, \alpha_0, \theta, A)$ *depending on* $d, R, c, T, \alpha_0, \theta$ *and* A *such that the following holds. Let* μ *be a finitely supported, contracting on average, exponentially separated,* (c,T)*-well-mixing and* (α_0, θ, A)*-non-degenerate probability measure on* G *with* $\mathrm{supp}(\mu) \subset \{g \in G : \rho(g) \in [R^{-1}, R]\}$ *and satisfying*

$$\frac{h_\mu}{|\chi_\mu|} > C \left(\max\left\{1, \log \frac{S_\mu}{h_\mu}\right\}\right)^2.$$

Then the associated self-similar measure ν *is absolutely continuous.*

A similar result for Furstenberg measures of $\mathrm{SL}_2(\mathbb{R})$ was established by Samuel Kittle [Kit23]. However in [Kit23] it is necessary to assume that $\alpha_0 \in (0, 1/3)$ and we currently can't prove an analogue of Proposition 8.1.2 for Furstenberg measures. Therefore Theorem 7.0.4 can be deduced in the case of self-similar measures and we also note that the examples of absolutely continuous Furstenberg measures in [Kit23] are more intricate.

We next state a version of our main theorem for $d \geq 3$ that implies (7.0.10) by Proposition 8.1.2, provided that μ_U is supported on matrices with algebraic coefficients.

Theorem 8.1.5. *Let* $d \geq 3$ *and* $R, c, T, \alpha_0, \theta, A > 0$ *with* $c, \alpha_0 \in (0,1)$ *and* $T \geq 1$. *Then there is a constant* $C = C(d, R, c, T, \alpha_0, \theta, A)$ *such that the following holds. Let* μ *be a finitely supported, contracting on average,* (c,T)*-well-mixing and* (α_0, θ, A)*-non-degenerate probability measure on* G *with* $\mathrm{supp}(\mu) \subset \{g \in G : \rho(g) \in [R^{-1}, R]\}$. *Moreover assume that all of the coefficients of the matrices in* $\mathrm{supp}(U(\mu))$ *lie in the number field* K *and have logarithmic height at most* $L \geq 1$. *Then* ν *is absolutely continuous if*

$$\frac{h_{U(\mu)}}{|\chi_\mu|} \geq C \max\left\{1, \log\left(\frac{L[K:\mathbb{Q}]}{h_{U(\mu)}}\right)\right\}^2.$$

As in (7.0.10) we do not assume in Theorem 8.1.5 that all the entries of elements in $\mathrm{supp}(\mu)$ are algebraic and only require the latter for $U(\mu)$. By Breuillard's uniform Tits alternative [Bre08], there is a constant $c_d > 0$ only depending on d such that $h_{U(\mu)} > c_d$ as long as the group generated by $\mathrm{supp}(U(\mu))$ in not virtually solvable. The advantage of Theorem 8.1.5 over (7.0.10) is that our result is particularly effective when $U(\mu)$ has high entropy (for example when $\mathrm{supp}(U(\mu))$ generates a free semi-group) and is explicit in terms of the dependence of the heights of the coefficients of $\mathrm{supp}(U(\mu))$. In addition, Theorem 8.1.5 applies to contracting only on average measures and does not require $\mathrm{supp}(U(\mu))$ to generate a dense subgroup of $SO(d)$.

8.2 Outline

We give a sketch for the proof of Theorem 8.1.4. Our proof extends the strategy of [Kit23] to self-similar measures and generalises it to higher dimensions, which in turn is inspired by ideas and techniques developed in [Hoc14], [Hoc17], [Var19a] and [Kit21]. Proposition 8.1.2 will be discussed and proved in section 14. An entropy theory for random walks on general Lie groups was developed in [KK25b] and will be used throughout Part II.

Let μ be a measure on $G = \mathrm{Sim}(\mathbb{R}^d)$ and let $\gamma_1, \gamma_2, \dots$ be independent μ-distributed random variables. For a stopping time τ write $q_\tau = \gamma_1 \cdots \gamma_\tau$. Note that if x is a sample of ν then so is $q_\tau x$. The basic idea of our proof is to decompose $q_\tau x$ as a sum

$$q_\tau x = X_1 + \dots + X_n \tag{8.2.1}$$

with $X_1, \dots, X_n$ independent random variables. We aim to show that for each scale $r > 0$ and a suitable stopping time τ that we can find a decomposition (8.2.1) such that for all $i \in [n]$,

$$|X_i| \leq C^{-1} r \quad \text{and} \quad \sum_{i=1}^{n} \mathrm{Var}\, X_i \geq C(\log\log r^{-1}) r^2 I \tag{8.2.2}$$

for a sufficiently large fixed constant $C = C(d) > 0$ only depending on d, where $\mathrm{Var}\, X_j$ is the covariance matrix of X_j and we denote by $\geq$ the partial order defined in (8.3.1). The proof of Theorem 8.1.4 comprises to establish (8.2.2) and to deduce from (8.2.2) that ν is absolutely continuous. For the former we use adequate entropy results and for the latter we work with the detail of a measure. The constant C is be closely related to the one from Theorem 8.1.4.

From Decomposition to Absolute Continuity

The notion of detail $s_r(\nu)$ at scale $r > 0$ of a measure ν is a tool introduced in [Kit21] measuring how smooth ν is at scale r. Detail is an analogue of the entropy between scales $1 - H(\nu; r|2r)$ used by [Var19a], yet with better properties. Our goal is to deduce from (8.2.2) that our self-similar measure ν satisfies for r sufficiently small,

$$s_r(\nu) \leq (\log r^{-1})^{-2}, \tag{8.2.3}$$

which implies that ν is absolutely continuous, as shown in [Kit21].

A novelty introduced in [Kit23] is a strong product bound for detail on $\mathbb{R}$, which we prove for $\mathbb{R}^d$ in section 10. Indeed, if $\lambda_1, \ldots, \lambda_k$ are measures on $\mathbb{R}^d$, $a < b$ and $r > 0$ with $s_r(\lambda_i) \leq \alpha$ for some $\alpha > 0$ and all $r \in [a, b]$ and $1 \leq i \leq k$, then, as shown in Corollary 10.2.4,

$$s_{a\sqrt{k}}(\lambda_1 * \cdots * \lambda_k) \leq Q'(d)^{k-1}(\alpha^k + k!ka^2b^{-2}) \tag{8.2.4}$$

for some constant $Q'(d)$ depending only on d. To prove (8.2.4), [Kit23] introduced k order detail, which we generalise to $\mathbb{R}^d$. We note that (8.2.4) is stronger than the product bounds [Kit21, Theorem 1.17] and [Var19a, Theorem 3] and is required in our proof.

To convert (8.2.2) into (8.2.3), we first partition $[n]$ as $J_1 \sqcup \ldots \sqcup J_k$ for $k \asymp \log\log r^{-1}$ such that the random variables $Y_j = \sum_{i \in J_j} X_i$ satisfy $\operatorname{Var} Y_j \gg_d C$. Then we apply a Berry-Essen type result to deduce that Y is well-approximated by a Gaussian random variable and therefore that $s_r(Y_j) \leq \alpha$ for some constant α depending on C, with α tending to zero as C tends to ∞. Finally we conclude by (8.2.4) that we roughly get $s_r(\nu) \leq Q'(d)^k \alpha^k = e^{k(\log Q'(d) + \log \alpha)}$. We choose $k \asymp \log\log r^{-1}$ and therefore deduce (8.2.3) provided that α is sufficiently small in terms of d or equivalently C is sufficiently large. This proves that ν is absolutely continuous.

From Decomposition on $\mathbb{R}^d$ to Decomposition on G

It remains to explain how to establish (8.2.2), which we first translate into an analogous question on G. Indeed, we will make a decomposition of q_τ into

$$q_\tau = g_1 \exp(U_1) g_2 \exp(U_2) \cdots g_n \exp(U_n) \tag{8.2.5}$$

for random variables $g_1, \ldots, g_n$ on G and $U_1, \ldots, U_n$ on the Lie algebra $\mathfrak{g}$ of G. In order to express $q_\tau v$ as a sum of random variables using (8.2.5), we apply Taylor's theorem in Proposition 9.1.4 to deduce

$$q_\tau v \approx g_1 \cdots g_n v + \sum_{i=1}^{n} \zeta_i(U_i), \tag{8.2.6}$$

where

$$\zeta_i = D_u(g_1 g_2 \cdots g_i \exp(u) g_{i+1} g_{i+2} \cdots g_n v)|_{u=0}.$$

For notational convenience we write in this outline of proofs

$$g_i' = g_1 \cdots g_i \quad \text{and} \quad g_i'' = g_{i+1} \cdots g_n$$

and denote

$$\rho_x = D_u(\exp(u)x)|_{u=0}.$$

Then by the chain rule, as shown in Lemma 9.1.3,

$$\mathrm{Var}(\zeta_i(U_i)) = \rho(g_i')^2\, U(g_i')\, \mathrm{Var}(\rho_{g_i''x}(U_i)) U(g_i')^T.$$

We will use the (c, T)-well-mixing and (α_0, θ, A)-non-degeneracy condition to ensure that

$$\mathrm{Var}(\zeta_i(U_i)) \geq c_1 \rho(g_i')^2 \mathrm{tr}(U_i) I = c_1 \mathrm{tr}(\rho(g_i') U_i) I \tag{8.2.7}$$

for some constant $c_1 > 0$ depending on $d, c, T, \alpha_0, \theta$ and A and where $\mathrm{tr}(U_i)$ is the trace of the covariance matrix of U_i. This will be shown in Proposition 13.2.1 by ensuring that each of the g_i is a product of sufficiently many γ_j such that we can apply well-mixing and non-degeneracy as $g_i x$ is close in distribution to ν. In fact, we exploit suitable properties of the derivative of ρ_x and use a principal component decomposition.

So in order to achieve (8.2.2), we require that

$$|U_i| \leq \rho(g_i')^{-1} r \qquad \text{and} \qquad \sum_{i=1}^{n} \mathrm{tr}(\rho(g_i') U_i) \geq C^3 c_1^{-1} (\log\log r^{-1}) r^2 \tag{8.2.8}$$

for the constant C from (8.2.2). Note that to arrive at (8.2.2) we replace U_i by $C^{-1}U_i$ and use (8.2.7).

Entropy Gap and Trace Bounds for Stopped Random Walk

We prove (8.2.8) by establishing suitable entropy bounds on G and then translate them to the necessary trace bounds. We use the following notation. For a random variable g on G and $s > 0$, we define $\mathrm{tr}(g; s)$ to be the supremum of all $t \geq 0$ such that we can find some σ-algebra $\mathscr{A}$ and some $\mathscr{A}$-measurable random variable h taking values in G such that

$$|\log(h^{-1}g)| \leq s \quad \text{and} \quad \mathbb{E}[\mathrm{tr}(\log(h^{-1}g)|\mathscr{A})] \geq ts^2,$$

where $\log : G \to \mathfrak{g}$ is the Lie group logarithm and we assume that $h^{-1}g$ is supported on a small ball around the identity. The reason we need to work with the conditional trace is to use (8.2.12).

To establish (8.2.8) we therefore need to find a collection of scales $s_i = \rho(g_i')^{-1}r$ such that

$$\sum_{i=1}^{n} \mathrm{tr}(q_\tau; s_i) \geq Cc_1^{-1} \log\log r^{-1} \tag{8.2.9}$$

for C an absolute constant depending only on d.

To show (8.2.9) one converts entropy estimates for q_τ into trace estimates, using in essence that for an absolutely continuous random variable Z on $\mathbb{R}^\ell$ we have

$$H(Z) \leq \frac{\ell}{2} \log\left(\frac{2\pi e}{\ell} \cdot \mathrm{tr}(Z)\right), \tag{8.2.10}$$

where H is the differential entropy and $\mathrm{tr}(Z)$ is the trace of the covariance matrix of Z. Equality holds in (8.2.10) if and only if Z is a spherical Gaussian.

We will work with entropy between scales on G. Precise definitions are given in section 11. For the purposes of this outline consider the entropy between scales defined for a random variable g taking values in G, two scales $r_1, r_2 > 0$ and a parameter $a > 0$ as

$$H_a(g; r_1|r_2) = (H(gs_{r_1,a}) - H(s_{r_1,a})) - (H(gs_{r_2,a}) - H(s_{r_2,a})),$$

where $H(\cdot)$ is the differential entropy and $s_{r,a}$ is a smoothing function supported on a ball of radius ar and satisfying for $\ell = \dim \mathfrak{g}$ that

$$\mathrm{tr}(\log(s_{r,a})) \asymp \ell r^2 \quad \text{and} \quad H(s_{r,a}) = \frac{\ell}{2}\log 2\pi e r^2 + O_d(e^{-a^2/4}) - O_{d,a}(r). \tag{8.2.11}$$

The function $s_{r,a}$ is chosen such that $H(s_{r,a})$ is essentially maximal while being compactly supported, which is necessary towards establishing (8.2.9). The parameter $a > 0$ is useful as it gives us a uniform error bound in (8.2.11). By using moreover (8.2.10), we relate in Theorem 11.5.1 entropy between scales and the trace by

$$\mathrm{tr}(g; 2ar) \gg a^{-2}(H_a(g; r|2r) - O_d(e^{-a^2/4}) - O_{d,a}(r)). \tag{8.2.12}$$

For $\kappa > 0$ denote by

$$\tau_\kappa = \inf\{n \geq 1 \,:\, \rho(\gamma_1 \cdots \gamma_n) \leq \kappa\}.$$

It is then shown in Proposition 12.1.1 for $r_1 < r_2$ and with $r_1 \leq \kappa^{\frac{S_\mu}{|\chi_\mu|}}$ that as $\kappa \to 0$ the following entropy gap holds:

$$H_a(q_{\tau_\kappa}; r_1|r_2) \geq \left(\frac{h_\mu}{|\chi_\mu|} - d\right) \log \kappa^{-1} + \ell \cdot \log r_2 + o_{\mu,d,a}(\log \kappa^{-1}). \tag{8.2.13}$$

We will give a sketch of the proof of (8.2.13) in the beginning of section 12 and just note that the main point of (8.2.13) is that most of the elements in the support of q_{τ_κ} are separated by $\kappa^{\frac{S_\mu}{|\chi_\mu|}}$, which by standard properties of entropy implies that $H(q_{\tau_\kappa} s_{r_1,a}) \approx H(q_{\tau_\kappa}) + H(s_{r_1,a})$. As we have to use a stopping time in (8.2.13), we will need to work with q_τ instead of a deterministic time throughout our proof.

By (8.2.13) it follows, assuming $h_\mu/|\chi_\mu|$ is sufficiently large and κ is sufficiently small, that

$$H_a(q_{\tau_\kappa}; \kappa^{\frac{S_\mu}{|\chi_\mu|}} | \kappa^{\frac{h_\mu}{2\ell|\chi_\mu|}}) \gg_d \frac{h_\mu}{|\chi_\mu|} \log \kappa^{-1}. \tag{8.2.14}$$

Using (8.2.14) and (8.2.12), we show in Proposition 12.2.2 with setting $S = 2\max\{S_\mu, h_\mu\}$ that for a collection of scales

$$s_i \in (\kappa^{\frac{S}{|\chi_\mu|}}, \kappa^{\frac{h_\mu}{2\ell|\chi_\mu|}}) \qquad \text{with} \qquad 1 \leq i \leq \hat{m}$$

and $\hat{m}$ being a fixed constant depending on S_μ and χ_μ that

$$\sum_{i=1}^{\hat{m}} \operatorname{tr}(q_{\tau_\kappa}; s_i) \gg_d \frac{h_\mu}{|\chi_\mu|} \max\left\{1, \log \frac{S_\mu}{h_\mu}\right\}^{-1}. \tag{8.2.15}$$

As we explain at the beginning of section 12, the error term $\max\left\{1, \log \frac{S_\mu}{h_\mu}\right\}^{-1}$ arises from the error $O_d(e^{-a^2/4})$ in (8.2.12).

Conclusion of Proof

The trace bound (8.2.15) is not sufficient to establish (8.2.9) as we require a lower bound depending on $\log \log r^{-1}$. To achieve such a bound and to conclude the proof, we concatenate several decompositions arising from (8.2.15) and therefore develop a suitable theory of such decompositions in section 13.

It therefore remains to find sufficiently many parameters $\kappa_1, \ldots, \kappa_m$ such that the resulting intervals

$$(\kappa_1^{\frac{S}{|\chi_\mu|}}, \kappa_1^{\frac{h_\mu}{2\ell|\chi_\mu|}}), \quad (\kappa_2^{\frac{S}{|\chi_\mu|}}, \kappa_2^{\frac{h_\mu}{2\ell|\chi_\mu|}}), \quad \ldots \quad (\kappa_m^{\frac{S}{|\chi_\mu|}}, \kappa_m^{\frac{h_\mu}{2\ell|\chi_\mu|}})$$

are disjoint. As we require that all of the scales are $\geq r$, we set $\kappa_1 = r^{\frac{|\chi_\mu|}{S}}$. On the other hand, we want all scales to be sufficiently small. We, for example, therefore require that $\kappa_m^{\frac{h_\mu}{2\ell|\chi_\mu|}} < e^{-10}$. Thus setting $\kappa_{i+1} = \kappa_i^{\frac{h_\mu}{3\ell S}}$, thereby ensuring that the resulting intervals are disjoint (provided h_μ/χ_μ is sufficiently large), a calculation shows that the maximal m we can take is

$$m \asymp \max\left\{1, \log\frac{S_\mu}{h_\mu}\right\}^{-1} \log\log r^{-1}.$$

Combining all of the above, it follows that when summing over all the scales

$$\sum_i \operatorname{tr}(q_{\tau_{\kappa_1}}; s_i) \gg_d \frac{h_\mu}{|\chi_\mu|} \max\left\{1, \log\frac{S_\mu}{h_\mu}\right\}^{-2} \log\log r^{-1}.$$

We therefore require in order to satisfy (8.2.9) that

$$\frac{h_\mu}{|\chi_\mu|} \max\left\{1, \log\frac{S_\mu}{h_\mu}\right\}^{-2} \geq C^3 c_1^{-1},$$

which leads us to the condition from Theorem 8.1.4 and concludes our sketch of the proof.

8.3 Notation for Part II

The reader may recall the notation stated in section 1.1. For an integer $n \geq 1$ we abbreviate $[n] = \{1, 2, \dots, n\}$. On $\mathbb{R}^d$ the euclidean norm is denoted $|\cdot|$.

Given two positive semi-definite symmetric real $d \times d$ matrices M_1 and M_2 we write

$$M_1 \geq M_2 \qquad \text{if and only if} \qquad x^T M_1 x \geq x^T M_2 x \quad \text{for all } x \in \mathbb{R}^d. \tag{8.3.1}$$

For a random variable X on $\mathbb{R}^d$ we denote by $\operatorname{Var}(X)$ the covariance matrix of X and by $\operatorname{tr}(X) = \operatorname{tr}\operatorname{Var}(X)$ the trace of the covariance matrix.

Given a metric space (M, d), $p \in [1, \infty)$ and two probability measures λ_1 and λ_2 on M, we define the L^p-Wasserstein metric as

$$\mathcal{W}_p(\lambda_1, \lambda_2) = \inf_{\gamma \in \Gamma(\lambda_1, \lambda_2)} \left(\int_{M \times M} d(x, y)^p \, d\gamma(x, y) \right)^{\frac{1}{p}}, \tag{8.3.2}$$

where $\Gamma(\lambda_1, \lambda_2)$ is the set of couplings of λ_1 and λ_2, i.e. of probability measures γ on $M \times M$ whose projections to the first coordinate is λ_1 and to the second is λ_2.

Throughout Part II we fix $d \geq 1$ and write $G = \mathrm{Sim}(\mathbb{R}^d)$, except for section 11, where G will be an arbitrary Lie group. The Lie algebra of G will be denoted $\mathfrak{g}$ and $\ell = \dim \mathfrak{g}$. We usually consider a fixed probability measure μ on G and independent samples $\gamma_1, \gamma_2, \ldots$ of μ. We write for $\kappa > 0$

$$q_n = \gamma_1 \cdots \gamma_n \qquad \text{and} \qquad \tau_\kappa = \inf\{n \geq 1\,;\, \rho(\gamma_n) \leq \kappa\}.$$

When μ is a probability measure on $G = \mathrm{Sim}(\mathbb{R}^d)$ and ν is a probability measure $\mathbb{R}^d$ we denote by $\mu * \nu$ the probability measure uniquely characterized by

$$(\mu * \nu)(f) = \int \int f(gx)\, d\mu(g) d\nu(x)$$

for $f \in C_c(\mathbb{R}^d)$. When $\mu = \sum_i p_i \delta_{g_i}$ is finitely supported, then

$$\mu * \nu = \sum_i p_i g_i \nu, \tag{8.3.3}$$

where $g_i\nu$ is the pushforward of ν by g_i defined by $(g_i\nu)(B) = \nu(g_i^{-1}B)$ for all Borel sets $B \subset \mathbb{R}^d$.

The various notions of entropy between scales as well as $\mathrm{tr}(g, r)$ will be given in section 11.

We will denote by m_G a normalised Haar measure on $\mathrm{Sim}(\mathbb{R}^d)$. Moreover if $H \subset O(d)$ is a closed subgroup, we will denote by m_H the Haar probability measure on H. For a probability measure μ_U on H, the L^2-spectral gap of μ_U in H is defined as

$$\mathrm{gap}_H(\mu_U) = 1 - ||T_{\mu_U}|_{L^2_0(G)}||, \tag{8.3.4}$$

where $(T_{\mu_U} f)(k) = \int f(hk)\, d\mu_U(h)$ for $f \in L^2(H)$ and $L^2_0(H) = \{f \in L^2(H) : m_H(f) = 0\}$ for $|| \circ ||$ the operator norm.

8.4 Organisation

In section 9 the Taylor expansion bound (8.2.6) is proved and we establish several probabilistic preliminaries. We discuss order k detail in section 10, establish (10.0.2) as well as show how to convert (8.2.2) into suitable detail bounds. Entropy results for general Lie groups are discussed in section 11. In section 12 we prove (8.2.13) and (8.2.15). Finally, we deduce Theorem 8.1.4 as well as Theorem 8.1.5 in section 13 by developing a decomposition theory for stopped random walks. We study (c, T)-well-mixing and (α_0, θ, A)-non-degeneracy in section 14 and prove Proposition 8.1.2 and Proposition 8.1.3. In section 15 we establish explicit examples and in particular we prove Corollary 7.0.11, Corollary 7.0.12, Corollary 7.0.8, Corollary 7.0.9 and Corollary 7.0.10.

Chapter 9

Preliminaries

In this section we first study the derivatives of the G action on $\mathbb{R}^d$ in section 9.1, then regular conditional distributions in section 9.2 and finally versions of the large deviation principle in section 9.3.

9.1 Derivative Bounds

9.1.1 Basic Properties

Let $G = \mathrm{Sim}(\mathbb{R}^d)$ with Lie algebra $\mathfrak{g} = \mathrm{Lie}(G)$. For $x \in \mathbb{R}^d$ consider the map

$$w_x : \mathfrak{g} \to \mathbb{R}^d, \qquad u \mapsto \exp(u)x.$$

Denote by $\psi_x = D_0 w_x : \mathfrak{g} \to \mathbb{R}^d$ the differential at zero of w_x.

Note that we can embed $G = \mathrm{Sim}(\mathbb{R}^d)$ into $\mathrm{GL}_{d+1}(\mathbb{R})$ via the map

$$g \mapsto \begin{pmatrix} r(g)U(g) & b(g) \\ 0 & 1 \end{pmatrix}.$$

We can therefore identify $\mathfrak{g}$ as a matrix Lie algebra and so can write

$$\mathfrak{g} = \left\{ \begin{pmatrix} \alpha & \beta \\ 0 & 0 \end{pmatrix} : \alpha \in \mathbb{R}I + \mathfrak{so}_d(\mathbb{R}), \beta \in \mathbb{R}^d \right\} \subset \mathfrak{gl}_{n+1}(\mathbb{R})$$

Thus for $u = \left(\begin{smallmatrix} \alpha & \beta \\ 0 & 0 \end{smallmatrix}\right)$ it follows that $\psi_x(u) = u\left(\begin{smallmatrix} x \\ 1 \end{smallmatrix}\right) = \alpha x + \beta$. With this viewpoint we also use the following convenient notation

$$ux = \psi_x(u) = \alpha x + \beta \tag{9.1.1}$$

We fix an inner product on $\mathfrak{g}$ and denote by $|\circ|$ the associated norm. Moreover, we choose an ordered orthonormal basis of $\mathfrak{g}$, endowing $\mathfrak{g}$ with a coordinate system. So every element $u \in \mathfrak{g}$ can be written as a sum $u = \sum_{i=1}^{\ell} u_i$, where u_i is the projection of u to the i-th basis vector. On numerous occasions we will consider derivatives with respect to u_i

In the following lemma, some properties about the derivatives of w_x, ψ_x and the map g are collected. For notational convenience, we denote throughout this subsection by $\frac{\partial f}{\partial x}$ the derivative $D_x f$ of a function $f : \mathbb{R}^{d_1} \to \mathbb{R}^{d_2}$ at a vector $x \in \mathbb{R}^{d_1}$. We furthermore write $\ell = \dim \mathfrak{g}$.

Lemma 9.1.1. *The following properties hold:*

(i) Let $g = \rho U + b \in G$. Then for all $x \in \mathbb{R}^d$, it holds that $\frac{\partial g}{\partial x} = \rho U$ and all of the second derivatives of g are zero.

(ii) Whenever $x \in \mathbb{R}^d$ and $|u| \le 1$ and $1 \le i, j \le \ell$,

$$\left|\frac{\partial w_x}{\partial u_i}\right| \ll_d \max(|x|, 1) \quad \text{and} \quad \left|\frac{\partial w_x}{\partial u_i \partial u_j}\right| \ll_d \max(|x|, 1).$$

(iii) For any $x_1, x_2 \in \mathbb{R}^d$ we have that

$$||\psi_{x_1} - \psi_{x_2}|| \ll_d |x_1 - x_2|.$$

(iv) Let $u \in \mathfrak{g} \backslash \{0\}$. Then there is a proper subspace $W_u \subset \mathbb{R}^d$ and a vector $u_0 \in \mathbb{R}^d$ such that if $\psi_x(u) = 0$ then $x \in u_0 + W_u$ for $x \in \mathbb{R}^d$.

(v) For all $\theta, A > 0$ there is $\delta > 0$ such that the following is true. Let $v \in \mathfrak{g}$ be a unit vector. Then there is a proper subspace $W_v \subset \mathbb{R}^d$ and a vector $v_0 \in \mathbb{R}^d$ such that if

$$x \in \mathbb{R}^d \backslash B_\theta(v_0 + W_v) \quad \text{and} \quad |x| \le A$$

for $B_\theta(v_0 + W_v)$ the θ-ball around $v_0 + W_v$ then

$$|\psi_x(v)| \ge \delta.$$

Proof. (i) follows by definition and (ii) by compactness of $\{u \in \mathfrak{g} : |u| \le 1\}$ and using that a pure translation by a small vector has norm $O_d(1)$. For (iii) using notation (9.1.1) it holds for $u = \left(\begin{smallmatrix} \alpha & \beta \\ 0 & 0 \end{smallmatrix}\right) \in \mathfrak{g}$ with $|u| \le 1$ that

$$|\psi_{x_1}(u) - \psi_{x_2}(u)| = |\alpha x_1 - \alpha x_2| \le ||\alpha|| \cdot |x_1 - x_2| \\ \ll_d |\alpha| \cdot |x_1 - x_2| \le |u| \cdot |x_1 - x_2|$$

using that the operator norm $|| \circ ||$ is equivalent to the inner product norm on $\mathfrak{g}$. To show (iv), we may assume that $\beta \in \mathrm{Im}(\alpha)$ as otherwise there is nothing to show. Then set $W_u = \ker(\alpha)$ and $u_0 \in \mathbb{R}^d$ such that $\alpha u_0 = -\beta$, implying the claim. (v) follows from (iv) by continuity. $\square$

For $u \in \mathfrak{g}\backslash\{0\}$ we define

$$E_\theta(u) = \mathbb{R}^d \backslash B_\theta(u_0 + W_u).$$

Given a random variable U taking values in $\mathfrak{g}$, we say that $u \in \mathfrak{g}$ is a first principal component if it is an eigenvector of its covariance matrix with maximal eigenvalue. Set

$$E_\theta(U) = \bigcup_{v \in P} E_\theta(v),$$

where P is the set of first principal components of U. Similarly if μ is a probability measure which is the law of a random variable U then we define $E_\theta(\mu) = E_\theta(U)$. Recall that given a random variable U in $\mathbb{R}^\ell$, we denote by $\mathrm{tr}(U)$ the trace of the covariance matrix of U.

Proposition 9.1.2. *For all $\theta, A > 0$ there is some $\delta = \delta(d, \theta, A) > 0$ such that the following is true. Suppose that U is a random variable taking values in $\mathfrak{g}$ and that $x \in \mathbb{R}^d$ with $|x| \leq A$. Suppose that $x \in E_\theta(U)$. Then*

$$\mathrm{tr}(Ux) \geq \delta \cdot \mathrm{tr}(U).$$

Proof. We applied here the notation (9.1.1) that $\psi_x(U) = Ux$. Indeed, we do not identify $\mathfrak{g}$ as a column vector here, but simply use the latter convenient notation.

Write $\ell = \dim \mathfrak{g}$ and let $w_1, \ldots, w_\ell$ be an orthonormal basis of eigenvectors of the covariance matrix $\mathrm{Var}(U)$. We may assume that U has mean zero. Denote by $U_i = \langle U, w_i \rangle = U^T w_i$ for $1 \leq i \leq \ell$ and assume without loss of generality that $\mathrm{Var}(U_1) \geq \ldots \geq \mathrm{Var}(U_\ell)$ so that w_1 is a principal component. Then the $(U_i)_{1 \leq i \leq \ell}$ are uncorrelated since for $i \neq j$

$$\begin{aligned}\mathrm{cov}(U_i, U_j) = \mathbb{E}[U_i U_j] &= \mathbb{E}[\langle U^T w_i, U^T w_j \rangle] \\ &= \mathbb{E}[\langle U U^T w_i, w_j \rangle] = \langle \mathrm{Var}(U) w_i, w_j \rangle = 0\end{aligned}$$

and it holds that $U = \sum_{i=1}^\ell U_i w_i$ and that $\mathrm{Var}(U_1) \geq \frac{1}{\ell}\mathrm{tr}(U)$. Also by Proposition 9.1.1 (v) it holds that $|\psi_x(w_1)| \geq \delta$. We then compute

$$\mathrm{tr}(\rho_x(U)) = \mathbb{E}[|\rho_x(U)|^2] = \mathbb{E}\left[\sum_{i=1}^\ell U_i^2 |\rho_x(w_i)|^2\right] \geq \mathbb{E}[U_1^2 |\rho_x(w_1)|^2] \geq \frac{\delta}{\ell}\mathrm{tr}(U).$$

□

Lemma 9.1.3. *Let U be a random variable on $\mathfrak{g}$ and let $g \in G$ and $x \in \mathbb{R}^d$. Denote*

$$\zeta = D_u g \exp(u) x|_{u=0}.$$

Then

$$\mathrm{Var}(\zeta(U)) = \rho(g)^2 \cdot U(g)\psi_x \circ \mathrm{Var}(U) \circ \psi_x^T U(g)^T.$$

Proof. Note that by the chain rule $\zeta(U) = \rho(g)U(g)\psi_x(U)$ and therefore

$$\mathrm{Var}\,\zeta(U) = \rho(g)^2 U(g)\,\mathrm{Var}(\psi_x(U))U(g)^T$$

Viewing $\psi_x : \mathfrak{g} \to \mathbb{R}^d$ as a matrix with our choice of coordinate system we write $\psi_x(U) = \psi_x \circ U$ and the claim follows. □

9.1.2 Taylor Expansion Bound

The aim of this subsection is to prove the following proposition, which crucially relies on the G action on $\mathbb{R}^d$ having vanishing second derivatives.

Proposition 9.1.4. *For every $A > 0$ there exists $C = C(d, A) > 1$ such that the following holds. Let $n \geq 1$, $r \in (0, 1)$ and let $u^{(1)}, \ldots, u^{(n)} \in \mathfrak{g}$. Let $g_1, \ldots, g_n \in G$ with*

$$\rho(g_i) < 1, \qquad |b(g_i)| \leq A \quad \text{and} \quad |u^{(i)}| \leq \rho(g_1 \cdots g_i)^{-1} r < 1.$$

Let $v \in \mathbb{R}^d$ with $|v| \leq A$ and write

$$x = g_1 \exp(u^{(1)}) \cdots g_n \exp(u^{(n)}) v$$

and

$$\zeta_i = D_0(g_1 g_2 \cdots g_i \exp(u) g_{i+1} \cdots g_{n-1} g_n v)$$

and let

$$S = g_1 \cdots g_n v + \sum_{i=1}^{n} \zeta_i(u^{(i)}).$$

Then it holds that

$$|x - S| \leq C^n \rho(g_1 \cdots g_n)^{-1} r^2.$$

To prove Proposition 9.1.4 we use the following version of Taylor's theorem.

Theorem 9.1.5. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^2-function, let $R_1, \ldots, R_n > 0$ and write $B = [-R_1, R_1] \times \ldots \times [-R_n, R_n]$. For integers $i, j \in [1, n]$ let $K_{ij} = \sup_B |\frac{\partial^2 f}{\partial x_i \partial x_j}|$ and let $x \in B$. Then we have that*

$$\left| f(x) - f(0) - \sum_{i=1}^{n} x_i \frac{\partial f}{\partial x_i}\bigg|_{x=0} \right| \leq \frac{1}{2} \sum_{i,j=1}^{n} K_{i,j} \, |x_i| \, |x_j|.$$

Lemma 9.1.6. *Let*

$$w : \mathfrak{g} \times \mathfrak{g} \longrightarrow \mathbb{R}^d, \qquad (x, y) \longmapsto \exp(x) g \exp(y) v$$

for fixed g, v*. Then if* $|x|, |y| \leq 1$ *it holds that*

$$\left| \frac{\partial w(x,y)}{\partial x_i \partial y_i} \right| \ll_d \rho(g) \max(|v|, 1).$$

Proof. Let $\hat{v} = \exp(y)v$ and by Lemma 9.1.1 (ii), $|\frac{\partial \hat{v}}{\partial y_i}| \ll_d \max(|v|, 1)$. Now let $\tilde{v} = g\hat{v}$. Therefore, by Lemma 9.1.1 (i), $||\frac{\partial \tilde{v}}{\partial \hat{v}}|| \leq \rho(g)$ and moreover, since $w = \exp(x)\tilde{v}$ and $|x| \leq 1$, it is readily shown that $||\frac{\partial^2 w}{\partial x_i \partial \tilde{v}}|| \ll_d 1$. We conclude therefore by the chain rule

$$\left| \frac{\partial w}{\partial x_i \partial y_i} \right| = \left| \left| \frac{\partial w}{\partial x_i \partial \tilde{v}} \right| \right| \cdot \left| \left| \frac{\partial \tilde{v}}{\partial \hat{v}} \right| \right| \cdot \left| \frac{\partial \hat{v}}{\partial y_i} \right| \ll_d \rho(g) \max(|v|, 1).$$

□

Proposition 9.1.7. *There exists a constants* $C = C(d) > 1$ *such that the following holds. Suppose that* $n \in \mathbb{Z}_{>0}$*,* $g_1, g_2, \ldots, g_n \in G$ *and let* $u^{(1)}, \ldots, u^{(n)} \in \mathfrak{g}$ *be such that* $|u^{(i)}| \leq 1$*.*

Let $v \in \mathbb{R}^d$ *and*

$$x = g_1 \exp(u^{(1)}) g_2 \exp(u^{(2)}) \cdots g_n \exp(u^{(n)}) v.$$

Then for any $1 \leq i, j \leq \ell$ *and any integers* $k, m \in [1, n]$ *with* $k \leq m$ *we have*

$$\left| \frac{\partial^2 x}{\partial u_i^{(k)} \partial u_j^{(m)}} \right| \leq C^n \rho(g_1 \cdots g_m) \max(|g_{m+1} \exp(u^{(m+1)}) \cdots g_n \exp(u^{(n)}) v|, 1).$$

Proof. First, we deal with the case $k = m$. Let

$$a = g_1 \exp(u^{(1)}) g_2 \exp(u^{(2)}) \cdots g_{k-1} \exp(u^{(k-1)}) g_k$$

and

$$b = g_{k+1} \exp(u^{(k+1)}) g_{k+2} \exp(u^{(k+2)}) \cdots g_n \exp(u^{(n)}) v$$

and let $\tilde{b} = \exp(u^{(k)}) b$. We have

$$\frac{\partial x}{\partial u_i^{(k)}} = \frac{\partial x}{\partial \tilde{b}} \frac{\partial \tilde{b}}{\partial u_i^{(k)}}.$$

Note that by Lemma 9.1.1 (i) all of the second derivatives of x with respect to $\tilde{b}$ are zero and therefore

$$\left| \frac{\partial^2 x}{\partial u_i^{(k)} \partial u_j^{(k)}} \right| \leq \left| \left| \frac{\partial x}{\partial \tilde{b}} \right| \right| \cdot \left| \frac{\partial^2 \tilde{b}}{\partial u_i^{(k)} \partial u_j^{(k)}} \right|. \tag{9.1.2}$$

Thus by Lemma 9.1.1 (i) and (ii) we conclude that

$$\left|\frac{\partial^2 x}{\partial u_i^{(k)}\partial u_j^{(k)}}\right| \ll_d \rho(a)\max(|b|,1) \leq C^n \rho(g_1\cdots g_\ell)\max(|b|,1)$$

for a suitable constant $C > 1$ using that $\rho(\exp(u^{(i)}))$ is bounded.

For the case $k < m$ we consider

$$\begin{aligned} a_1 &= g_1 \exp(u^{(1)}) g_2 \exp(u^{(2)}) \cdots g_{k-1} \exp(u^{(k-1)}) g_k \\ a_2 &= g_{k+1} \exp(u^{(k+1)}) g_{k+2} \exp(u^{(k+2)}) \cdots g_m \\ b &= g_{m+1} \exp(u^{(m+1)}) g_{m+2} \exp(u^{(m+2)}) \cdots g_n \exp(u^{(n)}) v. \end{aligned}$$

Then we consider $\tilde{b} = \exp(u^{(k)}) a_2 \exp(u^{(m)}) b$ and as before we conclude

$$\frac{\partial^2 x}{\partial u_i^{(k)}\partial u_j^{(m)}} = \frac{\partial x}{\partial \tilde{b}} \frac{\partial^2 \tilde{b}}{\partial u_i^{(k)}\partial u_j^{(m)}}.$$

We again arrive at (9.1.2) and deduce the claim as in the case $k = m$ using Lemma 9.1.6 instead of Lemma 9.1.1 (i). □

Proof. (of Proposition 9.1.4) We first show that there is a constant $C_1 = C_1(A, d)$ depending on A such that for all $1 \leq i \leq n$ we have that

$$|g_i \exp(u^{(i)}) \cdots g_n \exp(u^{(n)}) v| \leq C_1^{n-i+1}. \tag{9.1.3}$$

Indeed, we note that for any $u \in \mathfrak{g}$ with $|u| \leq 1$ and $v_0 \in \mathbb{R}^d$ it holds that $|\exp(u)v_0 - v_0| \leq C_2(|v_0| + 1)$ for an absolute constant $C_2 = C_2(d)$. Without loss of generality we assume that $C_2(d) > 1$. Therefore $|\exp(u^{(n)})v| \leq C_2(2|v| + 1)$. Next note that as $\rho(g_n) < 1$,

$$\begin{aligned} |g_n \exp(u^{(n)}) v| &\leq |g_n \exp(u^{(n)}) v - g_n(0)| + |g_n(0)| \\ &\leq \rho(g_n) |\exp(u^{(n)}) v| + |b(g_n)| \\ &\leq C_2(2|v| + |b(g_n)| + 1) \leq 4C_2(A+1), \end{aligned}$$

using that $\rho(g_n) < 1$ and that $|v| \leq A$ and $|b(g_n)| \leq A$. Continuing this argument inductively, we may conclude that

$$|g_i \exp(u^{(i)}) \cdots g_n \exp(u^{(n)}) v| \leq 4^{n-i+1} C_2^{n-i+1} (A + (n-i) + 1),$$

which implies (9.1.3).

Note that by the assumptions

$$\rho(g_1 \cdots g_\ell)|u^{(\ell)}|^2 \leq \rho(g_1 \cdots g_\ell)^{-1} r^2 \leq \rho(g_1 \cdots g_n)^{-1} r^2.$$

Therefore, by applying Theorem 9.1.5 together with Proposition 9.1.7 and (9.1.3) for a sufficiently large constant C depending on A and d in each of the coordinates of $\mathbb{R}^d$,

$$|x - S| \leq dn^2 C^n \rho(g_1 \cdots g_n)^{-1} r^2,$$

which implies the claim upon enlarging the constant C. □

9.2 Regular Conditional Distributions

In this section we review the definition of regular conditional distributions that will be used in section 11. On a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, we denote the conditional expectation by $\mathbb{E}[f|\mathscr{A}]$ for $f \in L^1(\Omega, \mathscr{F}, \mathbb{P})$ and a σ-algebra $\mathscr{A} \subset \mathscr{F}$. Given two measurable spaces $(\Omega_1, \mathscr{A}_1)$ and $(\Omega_2, \mathscr{A}_2)$, recall that a Markov kernel on $(\Omega_1, \mathscr{A}_1)$ and $(\Omega_2, \mathscr{A}_2)$ is a map $\kappa : \Omega_1 \times \mathscr{A}_2 \to [0, 1]$ if for any $A_2 \in \mathscr{A}_2$, the map $\kappa(\cdot, A_2)$ is $\mathscr{A}_1$-measurable and for any ω_1 the map $A_2 \to \kappa(\omega_1, A_2)$ is a probability measure.

Definition 9.2.1. *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space and let $\mathscr{A} \subset \mathscr{F}$ be a σ-algebra. Let (E, ξ) be a measurable space and let $Y : (\Omega, \mathscr{F}) \to (E, \xi)$ be a random variable. Then we say that a Markov kernel*

$$(Y|\mathscr{A}) : \Omega \times \xi \to [0, 1]$$

on $(\Omega, \mathscr{A})$ and (E, ξ) is a ***regular conditional distribution*** *if for all $B \in \xi$,*

$$(Y|\mathscr{A})(\omega, B) = \mathbb{P}[Y \in B \,|\, \mathscr{A}](\omega) = \mathbb{E}[1_{Y^{-1}(B)} \,|\, \mathscr{A}](\omega).$$

In other words,

$$\mathbb{E}[(Y|\mathscr{A})(\cdot, B) 1_A] = \mathbb{P}[A \cap \{Y \in B\}]$$

for all $A \in \mathscr{A}$.

Regular conditional distributions exists whenever $(\Omega, \mathscr{F}, \mathbb{P})$ is a standard probability space. To give a construction, recall (c.f. section 3 of [EW11]) that there exist conditional measures $\mathbb{P}_\omega^{\mathscr{A}}$ uniquely characterized by

$$\mathbb{E}[f|\mathscr{A}](\omega) = \int f \, d\mathbb{P}_\omega^{\mathscr{A}}.$$

Then

$$(Y|\mathscr{A})(\omega,\cdot) = Y_*\mathbb{P}_\omega^{\mathscr{A}}$$

Indeed, for $B \in \xi$,

$$(Y|\mathscr{A})(\omega, B) = E[1_{Y^{-1}(B)}|\mathscr{A}](\omega) = \int 1_{Y^{-1}(B)}\, d\mathbb{P}_\omega^{\mathscr{A}} = \mathbb{P}_\omega^{\mathscr{A}}(Y^{-1}(B)) = Y_*\mathbb{P}_\omega^{\mathscr{A}}(B).$$

We denote by $[Y|\mathscr{A}]$ a random variable defined on a separate probability space with law $(Y|\mathscr{A})$.

We recall that given two further σ-algebras $\mathscr{G}_1, \mathscr{G}_2 \subset \mathscr{F}$, we say that they are independent given $\mathscr{A}$ if for all $U \in \mathscr{G}_1$ and $V \in \mathscr{G}_2$

$$\mathbb{P}[U \cap V|\mathscr{A}] = \mathbb{P}[U|\mathscr{A}]\mathbb{P}[V|\mathscr{A}]$$

almost surely. Similarly, two random variables Y_1 and Y_2 are independent given $\mathscr{A}$ if the σ-algebra they generate are. Note that if Y_1 is $\mathscr{A}$-measurable, then it is independent given $\mathscr{A}$ to every random variable Y_2.

Given a topological group G and two measures μ_1 and μ_2 we recall that the convolution $\mu_1 * \mu_2$ is defined as

$$(\mu_1 * \mu_2)(B) = \int\int 1_B(gh)\, d\mu_1(g) d\mu_2(g)$$

for any measurable set $B \subset G$.

Lemma 9.2.2. *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, G be a topological group and g, h be G-valued random variables. Let $\mathscr{A} \subset \mathscr{F}$ be a σ-algebra and assume that g and h are independent given $\mathscr{A}$. Then the following properties hold:*

(i) $(gh|\mathscr{A}) = (g|\mathscr{A}) * (h|\mathscr{A})$ almost surely.

(ii) $[gh|\mathscr{A}] = [g|\mathscr{A}] \cdot [h|\mathscr{A}]$ almost surely.

Proof. To show (i) we note that by assumption g and h are independent with respect to $\mathbb{P}_\omega^{\mathscr{A}}$ for almost all $\omega \in \Omega$. This implies that

$$\mathbb{E}_{\mathbb{P}_\omega^{\mathscr{A}}}[f(gh)] = \mathbb{E}_{\mathbb{P}_\omega^{\mathscr{A}}}[\mathbb{E}_{\mathbb{P}_\omega^{\mathscr{A}}}[f(gh)|h]] = \mathbb{E}_{(z_1,z_2)\sim\mathbb{P}_\omega^{\mathscr{A}}\times\mathbb{P}_\omega^{\mathscr{A}}}[f(g(z_1)h(z_1))],$$

proving (i). (ii) follows from (i) on a suitable separate probability space. □

9.3 Large Deviation Principle

In this subsection we review various versions of the large deviation principle. Throughout this section we denote by μ a measure on G and by $\gamma_1, \gamma_2, \dots$ independent samples from μ. Applying the classical large deviation principle to ρ, we can state the following.

Lemma 9.3.1. *Let μ be a compactly supported, contracting on average probability measure on G. Then for every $\varepsilon > 0$ there is $\delta = \delta(\mu, \varepsilon) > 0$ such that for all sufficiently large n,*

$$\mathbb{P}\Big[|n\chi_\mu - \log \rho(\gamma_1) \cdots \rho(\gamma_n)| > \varepsilon n \Big] \leq e^{-\delta n}.$$

We generalise Lemma 9.3.1 to stopping times.

Lemma 9.3.2. *Let μ be a compactly supported contracting on average probability measure on G and let $\kappa > 0$ and denote*

$$\tau_\kappa = \inf\{n \geq 1 \, : \, \rho(\gamma_1 \dots \gamma_n) \leq \kappa\}.$$

Then for every $\varepsilon > 0$ there is $\delta > 0$ such that for sufficiently small κ

$$\mathbb{P}\Big[\Big|\tau_\kappa - \frac{\log \kappa^{-1}}{|\chi_\mu|}\Big| > \varepsilon \log \kappa^{-1}\Big] \leq e^{-\delta \log \kappa^{-1}}$$

Proof. If $\tau_\kappa > \frac{\log \kappa^{-1}}{|\chi_\mu|} + \varepsilon \log \kappa^{-1}$ then

$$\rho(\gamma_1 \cdots \gamma_{\lfloor \frac{\log \kappa^{-1}}{|\chi_\mu|} + \varepsilon \log \kappa^{-1} \rfloor}) \geq \kappa,$$

which by Lemma 9.3.1 has probability at most $e^{-\delta \log \kappa^{-1}}$ for some $\delta > 0$ and sufficiently small κ.

Write $R = \inf\{\rho(g) \, : \, g \in \mathrm{supp}(\mu)\} \in (0,1)$, which is non-zero since μ is compactly supported. Therefore when $\tau_\kappa < \frac{\log \kappa^{-1}}{|\chi_\mu|} - \varepsilon \log \kappa^{-1}$ happens there must be some integer

$$k \in \left[\frac{\log \kappa^{-1}}{|\log R|}, \frac{\log \kappa^{-1}}{|\chi_\mu|} - \varepsilon \log \kappa^{-1}\right]$$

such that

$$\log \rho(\gamma_1 \cdots \gamma_k) \leq \log \kappa.$$

Note that for sufficiently small κ we have $k|\chi_\mu| \leq \log \kappa^{-1} - \varepsilon |\chi_\mu| \, |\log R|$ and therefore

$$\log \rho(\gamma_1 \cdots \gamma_k) \leq \log \kappa \leq k(\chi_\mu + \varepsilon |\log R| \chi_\mu). \tag{9.3.1}$$

By Lemma 9.3.1 the probability that (9.3.1) happens is $\leq e^{-\delta' k} = e^{-\delta' O_\mu(\log \kappa^{-1})}$ for some $\delta' > 0$. Since there are at most $O_\mu(\log \kappa^{-1})$ many possibilities for k, the claim follows by the union bound. □

From Lemma 9.3.1 and (7.0.2) we can deduce the following corollary.

Corollary 9.3.3. *Let μ be a contracting on average probability measure on G. Then for every $\varepsilon > 0$ there is $\delta = \delta(\mu, \varepsilon) > 0$ such that for all sufficiently large N*

$$\mathbb{P}\Big[\exists n \geq N : \rho(\gamma_1 \cdots \gamma_n) \geq \exp((\chi_\mu + \varepsilon)n)\Big] \leq e^{-\delta N} \tag{9.3.2}$$

and

$$\mathbb{P}\Big[\exists n, m \geq N : |b(\gamma_1 \cdots \gamma_n) - b(\gamma_1 \cdots \gamma_m)| \geq \exp((\chi_\mu + \varepsilon)\min(m, n))\Big] \leq e^{-\delta N}.$$

Proof. Equation (9.3.2) follows from Lemma 9.3.1 and Borel-Cantelli. For (9.3.3) note that when $m \geq n + 1$,

$$|b(\gamma_1 \cdots \gamma_n) - b(\gamma_1 \cdots \gamma_m)| \leq \rho(\gamma_1 \cdots \gamma_n)|b(\gamma_{n+1} \cdots \gamma_m)|.$$

Therefore by (9.3.2) it suffices to show that for sufficiently large N we have that

$$\mathbb{P}[\exists k \geq 1 \,:\, |b(\gamma_1 \cdots \gamma_k)| \geq e^{\varepsilon N}] \leq e^{-\delta N},$$

which readily follows from (7.0.2) and Borel-Cantelli as $b(\gamma_1 \cdots \gamma_k)$ converges exponentially fast in distribution to ν. □

The next lemma was proved in [Kit23].

Lemma 9.3.4. *(Corollary 7.9 of* [Kit23]*) There is a constant $c > 0$ such that the following is true for all $a \in [0, 1)$ and $n \geq 1$. Let $X_1, \ldots, X_n$ be random variables taking values in $[0, 1]$ and let $m_1, \ldots, m_n \geq 0$ be such that we have almost surely $\mathbb{E}[X_i | X_1, \ldots, X_{i-1}] \geq m_i$ for $1 \leq i \leq n$. Suppose that $\sum_{i=1}^n m_i = an$. Then*

$$\log \mathbb{P}\left[X_1 + \ldots + X_n \leq \frac{1}{2}na\right] \leq -cna.$$

We generalise Lemma 9.3.4 to higher dimensions.

Lemma 9.3.5. *There is some absolute constant $c > 0$ such that the following is true. Suppose that $X_1, \ldots, X_n$ are random $d \times d$ symmetric positive semi-definite matrices such that $X_i \leq bI$ for some $b > 0$ and*

$$\mathbb{E}[X_i | X_1, \ldots, X_{i-1}] \geq m_i I.$$

Suppose that $\sum_{i=1}^n m_i = an$. Then there is some constant $C = C(a/b, d)$ depending only on a/b and d such that

$$\mathbb{P}\left[X_1 + \cdots + X_n > \frac{na}{4} I\right] \geq 1 - Ce^{-can}$$

Here we are using the partial ordering (8.3.1).

Proof. For convenience write $Y_n = X_1 + \ldots + X_n$ and choose a set S of unit vectors in $\mathbb{R}^d$ such that if y is any unit vector in $\mathbb{R}^d$ then there exists $x \in S$ with $\|x - y\| \leq \frac{a}{8b}$. Note that the size of S depends only on d and a/b.

By Lemma 9.3.4 we know that for any $x \in S$,

$$\log \mathbb{P}\left[x^T Y_n x \leq \frac{na}{2}\right] \leq -can.$$

Let A be the event that there exists some $x \in S$ with $x^T Y_n x \leq \frac{na}{2}$. We have that $\log \mathbb{P}[A]$ is at most $-can + \log |S|$. It suffices therefore to show that on A^C we have $Y_n > \frac{na}{4} I$.

Indeed let $y \in \mathbb{R}^d$ be a unit vector. Choose some $x \in \mathbb{R}^d$ with $\|x - y\| \leq a/8b$. Suppose that A^C occurs. Note that we must have $Y_n \leq bnI$ and therefore $||Y_n|| \leq bn$. This means

$$\begin{aligned} y^T Y_n y &= x^T Y_n x + x^T Y_n (y - x) + (y - x)^T Y_n y \\ &> \frac{an}{2} - 2bn \cdot \frac{a}{8b} = \frac{an}{4}. \end{aligned}$$

and result follows. □

Chapter 10

Order k Detail

The goal of this section is to prove the product bound (8.2.4) and to show how to convert (8.2.2) into suitable estimates for detail. We first recall in section 10.1 the definition of the detail $s_r(\lambda)$ of a measure λ on $\mathbb{R}^d$ at scale $r > 0$ that was first introduced by [Kit21]. We then expand the definition and results of order k detail $s_r^{(k)}(\lambda)$ of a measure from [Kit23] to measures on $\mathbb{R}^d$.

As mentioned in the outline of proofs, the advantage of using k-order detail over detail is that it leads to stronger product bounds. Indeed, we will show in Lemma 10.2.1 that

$$s_r^{(k)}(\lambda_1 * \cdots * \lambda_k) \leq s_r(\lambda_1) \cdots s_r(\lambda_k) \tag{10.0.1}$$

for measures $\lambda_1, \ldots, \lambda_k$ on $\mathbb{R}^d$ and $r > 0$. Moreover, if $s_r^{(k)}(\lambda) \leq \alpha$ for all $r \in [a, b]$ and some $k \geq 1$ then we show in Proposition 10.2.3 for a constant $Q'(d)$ depending only on d that

$$s_{a\sqrt{k}}(\lambda) \leq Q'(d)^{k-1}(\alpha + k!ka^2b^{-2}). \tag{10.0.2}$$

Combining (10.0.1) and (10.0.2), we deduce the strong product bound (Corollary 10.2.4) mentioned at (8.2.4) in the outline of proofs.

In section 10.3, we show that the difference in the detail of two measures is bounded in term of their Wasserstein distance. Finally, in section 10.4 we show how to convert the conditions from (8.2.2) into good estimates for detail. The latter requires Berry-Essen type results, the Wasserstein distance bounds from section 10.3, (10.0.1) and a suitable partition of $\sum_i X_i$.

All of these results will be used in section 13.

10.1 Definitions

Denote by η_y the standard Gaussian density on $\mathbb{R}^d$ with covariance matrix $y \cdot \mathrm{I}_d$, i.e.

$$\eta_y(x) = \frac{1}{(2\pi y)^{d/2}} \exp\left(-\frac{||x||^2}{2y}\right).$$

Moreover, we write

$$\eta_y^{(1)} = \frac{\partial}{\partial y}\eta_y.$$

Given a probability measure λ on $\mathbb{R}^d$ the detail of λ at scale $r > 0$ is defined as

$$s_r(\lambda) = r^2\, Q(d)\, ||\lambda * \eta_{r^2}^{(1)}||_1,$$

where $Q(d) = ||\eta_1^{(1)}||^{-1} = \frac{1}{2}\Gamma(\frac{d}{2})(\frac{d}{2e})^{-d/2}$ and note that by Stirling's approximation $d^{-1/2} \leq Q(d) \leq ed^{-1/2}$ for all $d \geq 1$. Moreover, $r^2 Q(d) = ||\eta_{r^2}^{(1)}||^{-1}$ and therefore $s_r(\lambda) \leq 1$ for every probability measure λ.

Proposition 10.1.1. [Kit21, section 2] *Let λ and μ be probability measures on $\mathbb{R}^d$. Then the following properties hold:*

(i) Suppose that there is $\beta > 1$ such that $s_r(\lambda) < (\log r^{-1})^{-\beta}$ for sufficiently small r. Then λ is absolutely continuous.

(ii) $s_r(\lambda * \mu) \leq s_r(\lambda)$.

Definition 10.1.2. *Given a probability measure λ on $\mathbb{R}^d$ and some $k \geq 1$ we define the **order** k **detail of** λ **at scale** r as*

$$s_r^{(k)}(\lambda) = r^{2k}\, Q(d)^k\, ||\lambda * \eta_{kr^2}^{(k)}||_1,$$

where $\eta_y^{(k)} = \frac{\partial^k}{\partial y^k}\eta_y$.

10.2 Bounding Detail

We have the following properties:

Lemma 10.2.1. *Let $k \geq 1$ and let $\lambda_1, \lambda_2, \dots, \lambda_k$ be probability measures on $\mathbb{R}^d$. Then*

$$s_r^{(k)}(\lambda_1 * \lambda_2 * \dots * \lambda_k) \leq s_r(\lambda_1) s_2(\lambda_2) \cdots s_r(\lambda_k). \tag{10.2.1}$$

In particular, for any probability measure λ on $\mathbb{R}^d$ and $k \geq 1$,

$$s_r^{(k)}(\lambda) \leq 1. \tag{10.2.2}$$

Proof. Recall that by the Heat equation $\frac{\partial}{\partial y}\eta_y(x) = \frac{1}{2}\sum_{i=1}^d \frac{\partial^2}{\partial x_i^2}\eta_y(x)$ and therefore by standard properties of convolution

$$\begin{aligned}
\eta_{kr^2}^{(k)} &= \frac{1}{2^k}\sum_{i_1,\ldots,i_k=1}^{d} \frac{\partial^2}{\partial x_{i_1}^2}\cdots\frac{\partial^2}{\partial x_{i_k}^2}\eta_{kr^2} \\
&= \underbrace{\left(\frac{1}{2}\sum_{i=1}^{d}\frac{\partial^2}{\partial x_i^2}\eta_{r^2}\right) * \left(\frac{1}{2}\sum_{i=1}^{d}\frac{\partial^2}{\partial x_i^2}\eta_{r^2}\right) * \cdots * \left(\frac{1}{2}\sum_{i=1}^{d}\frac{\partial^2}{\partial x_i^2}\eta_{r^2}\right)}_{k \text{ times}} \\
&= \underbrace{\eta_{r^2}^{(1)} * \eta_{r^2}^{(1)} * \cdots * \eta_{r^2}^{(1)}}_{k \text{ times}}.
\end{aligned}$$

This concludes the proof of (10.2.1) as

$$\begin{aligned}
||\lambda_1 * \ldots * \lambda_k * \eta_{kr^2}^{(k)}||_1 &= ||\lambda_1 * \eta_{r^2}^{(1)} * \lambda_2 * \eta_{r^2}^{(1)} * \cdots * \lambda_k * \eta_{r^2}^{(1)}||_1 \\
&\leq ||\lambda_1 * \eta_{r^2}^{(1)}||_1 \cdot ||\lambda_2 * \eta_{r^2}^{(1)}||_1 \cdots ||\lambda_k * \eta_{r^2}^{(1)}||_1.
\end{aligned}$$

To show (10.2.2) we set $\lambda_1 = \lambda$ and $\lambda_2 = \ldots = \lambda_k = \delta_e$ and use that $s_r(\lambda_i) \leq 1$. □

Lemma 10.2.2. *Let k be an integer greater than 1 and suppose that λ is a probability measure on $\mathbb{R}^d$. Suppose that $a, b, c > 0$ and $\alpha \in (0,1)$. Assume that $a < b$ and that for all $r \in [a,b]$ it holds that*

$$s_r^{(k)}(\lambda) \leq \alpha + cr^{2k}.$$

Then for all $r \in \left[a\sqrt{\frac{k}{k-1}}, b\sqrt{\frac{k}{k-1}}\right]$ we have

$$s_r^{(k-1)}(\lambda) \leq 2eQ(d)^{-1}\left(\alpha + (b^{-2(k-1)} + ckb^2)r^{2(k-1)}\right).$$

Proof. By the assumption and the definition of detail for $y \in [ka^2, kb^2]$ and writing $y = kr^2$,

$$||\lambda * \eta_y^{(k)}||_1 \leq r^{-2k}Q(d)^{-k}(\alpha + cr^{2k}) = \alpha y^{-k}k^kQ(d)^{-k} + cQ(d)^{-k}.$$

Therefore with $y \in [ka^2, kb^2]$,

$$\begin{aligned}
||\lambda * \eta_y^{(k-1)}||_1 &\leq ||\lambda * \eta_{kb^2}^{(k-1)}||_1 + \int_y^{kb^2} ||\lambda * \eta_u^{(k)}||_1 \, du \\
&\leq ||\eta_{kb^2}^{(k-1)}||_1 + \int_y^{kb^2} \alpha u^{-k}k^kQ(d)^{-k} + cQ(d)^{-k}\, du \\
&\leq (\tfrac{kb^2}{k-1})^{-(k-1)}Q(d)^{-(k-1)} + \alpha k^kQ(d)^{-k}\tfrac{y^{-(k-1)}}{k-1} + Q(d)^{-k}ckb^2,
\end{aligned}$$

where we bounded in the last inequality $||\eta_{kb^2}^{(k-1)}||_1$ by using that order $(k-1)$-detail is at most one, $\int_y^{kb^2} \alpha u^{-k} k^k Q(d)^{-k}\, du$ by $\int_y^{\infty} \alpha u^{-k} k^k Q(d)^{-k}\, du$ and $\int_y^{kb^2} cQ(d)^{-k}\, du$ by $\int_0^{kb^2} cQ(d)^{-k}\, du$. Using that $(\frac{k}{k-1})^{-(k-1)} < 1$ we therefore get

$$||\lambda * \eta_y^{(k-1)}||_1 \leq \alpha k^k Q(d)^{-k} \frac{y^{-(k-1)}}{k-1} + (b^{-(2k-2)} + Q(d)^{-1} ckb^2) Q(d)^{-(k-1)}.$$

Substituting the definition of order k detail gives for $y = (k-1)r^2 \in [ka^2, kb^2]$ or equivalently $r \in \left[a\sqrt{\frac{k}{k-1}}, b\sqrt{\frac{k}{k-1}}\right]$,

$$\begin{aligned}
s_r^{(k-1)}(\lambda) &= r^{2(k-1)} Q(d)^{k-1} ||\lambda * \eta_{(k-1)r^2}^{(k-1)}||_1 \\
&\leq \alpha r^{2(k-1)} k^k Q(d)^{-1} \frac{((k-1)r^2)^{-(k-1)}}{k-1} + r^{2(k-1)}(b^{-2(k-1)} + Q(d)^{-1} ckb^2) \\
&\leq \alpha Q(d)^{-1} \left(1 + \frac{1}{k-1}\right)^k + (b^{-2(k-1)} + Q(d)^{-1} ckb^2) r^{2(k-1)}.
\end{aligned}$$

Finally using that $\left(1 + \frac{1}{k-1}\right)^k \leq 2e$ and that $2eQ(d)^{-1} \geq 1$ the proof is concluded. □

Proposition 10.2.3. *Let k be an integer greater than 1 and suppose that λ is a probability measure on $\mathbb{R}^d$. Suppose that $a, b > 0$ and $\alpha \in (0,1)$. Assume that $a < b$ and that for all $r \in [a,b]$ we have*

$$s_r^{(k)}(\lambda) \leq \alpha.$$

Then we have that

$$s_{a\sqrt{k}}(\lambda) \leq Q'(d)^{k-1}(\alpha + k! \cdot ka^2 b^{-2})$$

for $Q'(d) = 4eQ(d)^{-1} \geq 1$.

Proof. We will show by induction for $j = k, k-1, \dots, 1$ that for all $r \in \left[a\sqrt{\frac{k}{j}}, b\sqrt{\frac{k}{j}}\right]$ we have

$$s_r^{(j)}(\lambda) \leq Q'(d)^{k-j} \left(\alpha + \frac{k!}{j!} b^{-2j} r^{2j}\right), \tag{10.2.3}$$

which implies the claim by setting $j = 1$ and $r = a\sqrt{k}$. The case $j = k$ follows from the conditions of the lemma. For the inductive step assume now that for all $r \in \left[a\sqrt{\frac{k}{j}}, b\sqrt{\frac{k}{j}}\right]$ we have that (10.2.3) holds. Then by Lemma 10.2.2 we have for all $r \in \left[a\sqrt{\frac{k}{j-1}}, b\sqrt{\frac{k}{j-1}}\right]$

$$\begin{aligned}
s_r^{(j-1)}(\lambda) &\leq Q'(d)^{k-j} 2eQ(d)^{-1} \left(\alpha + \left(b^{-2(j-1)} + \frac{k!}{j!} b^{-2j} jb^2\right) r^{2(j-1)}\right) \\
&\leq Q'(d)^{k-j} 2eQ(d)^{-1} \left(\alpha + \left(1 + \frac{k!}{(j-1)!}\right) b^{-2(j-1)} r^{2(j-1)}\right) \\
&\leq Q'(d)^{k-(j-1)} \left(\alpha + \frac{k!}{(j-1)!} b^{-2(j-1)} r^{2(j-1)}\right).
\end{aligned}$$

□

Combining Lemma 10.2.1 and Proposition 10.2.3, we arrive at the following corollary.

Corollary 10.2.4. *Let $k \geq 1$ and let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be probability measures on $\mathbb{R}^d$. Suppose that $a, b > 0$ and $\alpha \in (0,1)$. Assume that $a < b$ and that for all $r \in [a,b]$ and $i \in [k]$ we have*

$$s_r(\lambda_i) \leq \alpha.$$

Then it holds that

$$s_{a\sqrt{k}}(\lambda) \leq Q'(d)^{k-1}(\alpha^k + k! \cdot ka^2b^{-2}).$$

10.3 Wasserstein Distance

Recall as in (8.3.2) that the Wasserstein 1-distance on $\mathbb{R}^d$ between λ_1 and λ_2 is defined as

$$\mathcal{W}_1(\lambda_1, \lambda_2) = \inf_{\gamma \in \Gamma(\lambda_1, \lambda_2)} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y| \, d\gamma(x, y),$$

where $\Gamma(\lambda_1, \lambda_2)$ is the set of couplings between λ_1 and λ_2. We show that detail is bounded up to a constant by the Wasserstein distance.

Lemma 10.3.1. *Let λ_1 and λ_2 be probability measures on $\mathbb{R}^d$. Then for $k \geq 1$ and $r > 0$,*

$$|s_r^{(k)}(\lambda_1) - s_r^{(k)}(\lambda_2)| \leq edr^{-1}\mathcal{W}_1(\lambda_1, \lambda_2),$$

where e is Euler's number.

Proof. Let X and Y be random variables with laws λ_1 and λ_2 respectively. Then

$$(\lambda_1 - \lambda_2) * \eta_{kr}^{(k)}(v) = \mathbb{E}\left[\eta_{kr}^{(k)}(v - X) - \eta_{kr}^{(k)}(v - Y)\right]$$

and therefore

$$\left|(\lambda_1 - \lambda_2) * \eta_{kr}^{(k)}(v)\right| \leq \mathbb{E}\left[\left|\eta_{kr}^{(k)}(v - X) - \eta_{kr}^{(k)}(v - Y)\right|\right].$$

Note that

$$\left|\eta_{kr}^{(k)}(v - X) - \eta_{kr}^{(k)}(v - Y)\right| \leq \int_X^Y \left|\nabla \eta_{kr}^{(k)}(v - u)\right| |du|,$$

where $\int_x^y \cdot |du|$ is understood to be the integral along the shortest path between x and y and ∇ is the gradient. Thus

$$\begin{aligned}||(\lambda_1 - \lambda_2) * \eta_{kr}^{(k)}||_1 &\leq \int_{\mathbb{R}^d} \mathbb{E}\left[\int_X^Y \left|\nabla \eta_{kr}^{(k)}(v-u)\right| |du|\right] dv \\ &= \mathbb{E}\left[\int_X^Y \int_{\mathbb{R}^d} \left|\nabla \eta_{kr}^{(k)}(v-u)\right| dv\, |du|\right] \\ &= ||\nabla \eta_{kr}^{(k)}||_1 \mathbb{E}[|X-Y|] \\ &\leq \left(\sum_{i=1}^d \left|\left|\frac{\partial}{\partial x_i}\eta_{kr}^{(k)}\right|\right|_1\right) \mathbb{E}[|X-Y|]\end{aligned}$$

We next bound $||\frac{\partial}{\partial x_i}\eta_{kr}^{(k)}||_1$. As in the proof of Lemma 10.2.1, it follows that

$$\frac{\partial}{\partial x_i}\eta_{kr^2}^{(k)} = \left(\frac{\partial}{\partial x_i}\eta_{\frac{k}{k+1}r^2}\right) * \underbrace{\eta_{\frac{k}{k+1}r^2}^{(1)} * \dots * \eta_{\frac{k}{k+1}r^2}^{(1)}}_{k \text{ times}}.$$

Using standard properties of Gaussian integrals,

$$\left|\left|\frac{\partial}{\partial x_i}\eta_{\frac{k}{k+1}r^2}\right|\right|_1 = \sqrt{\frac{2(k+1)}{k\pi}}r^{-1} \leq \sqrt{\frac{k+1}{k}}r^{-1}$$

and therefore

$$\begin{aligned}\left|\left|\frac{\partial}{\partial x_i}\eta_{kr}^{(k)}\right|\right|_1 &\leq \left|\left|\frac{\partial}{\partial x_i}\eta_{\frac{k}{k+1}r^2}\right|\right|_1 \cdot ||\eta_{\frac{k}{k+1}r^2}^{(1)}||_1^k \\ &\leq \left(\frac{k+1}{k}\right)^{(k+1)/2} Q(d)^{-k}r^{-2k-1}.\end{aligned}$$

Using that $\left(\frac{k+1}{k}\right)^{(k+1)/2} \leq e$, we conclude

$$\begin{aligned}|s_r^{(k)}(\lambda_1) - s_r^{(k)}(\lambda_2)| &\leq r^{2k}Q(d)^k||(\lambda_1-\lambda_2) * \eta_{kr}^{(k)}||_1 \\ &\leq der^{-1}\mathbb{E}[|X-Y|].\end{aligned}$$

Choosing a coupling for X and Y which minimizes $\mathbb{E}[|X-Y|]$ gives the required result. □

10.4 Small Random Variables Bound in $\mathbb{R}^d$

The aim of this subsection is to show that the sum of independent random variables in $\mathbb{R}^d$ have small detail whenever they are supported close to 0 and have a sufficiently large variance. To state our result, we use the partial order (8.3.1) for positive semi-definite symmetric matrices.

Proposition 10.4.1. *For every positive integer $d \geq 1$ and every $\alpha > 0$ there exists some $C = C(\alpha, d) > 0$ such that the following is true for all $r > 0$ and positive integers k. Let $X_1, X_2, \ldots, X_n$ be independent random variables taking values in $\mathbb{R}^d$ such that almost surely*

$$|X_i| \leq C^{-1} r \quad \text{and} \quad \sum_{i=1}^{n} \operatorname{Var} X_i \geq Ckr^2 I.$$

Then

$$s_r^{(k)}(X_1 + \ldots + X_n) \leq \alpha^k.$$

Proposition 10.4.1 relies on a higher dimensional Berry-Essen type result, which implies Proposition 10.4.1 for $k = 1$, as deduced in Lemma 10.4.4. To prove the higher dimensional Berry-Essen type result we first need the following.

Theorem 10.4.2. *Let $X_1, X_2, \ldots, X_n$ be independent random variables taking values in $\mathbb{R}$ with mean 0 and for each $i \in [n]$ let $\mathbb{E}[X_i^2] = \omega_i^2$ and $\mathbb{E}[|X_i|^3] = \gamma_i^3 < \infty$. Let $\omega^2 = \sum_{i=1}^n \omega_i^2$ and let $S = X_1 + \cdots + X_n$. Let N be a normal distribution with mean 0 and variance ω^2. Then for an absolute implied constant*

$$\mathcal{W}_1(S, N) \ll \frac{\sum_{i=1}^n \gamma_i^3}{\sum_{i=1}^n \omega_i^2}.$$

Proof. A proof of this result may be found in [Eri73]. □

From this we may deduce the following higher dimensional Berry-Essen type result by using projections onto one-dimensional subspaces.

Lemma 10.4.3. *Let $X_1, X_2, \ldots, X_n$ be independent random variables taking values in $\mathbb{R}^d$ with mean 0 and for each $i \in [n]$ write*

$$\Sigma_i = \operatorname{Var} X_i.$$

Suppose that $\delta > 0$ is such that for each $i \in [n]$ we have $|X_i| \leq \delta$ almost surely. Let $\Sigma = \sum_{i=1}^n \Sigma_i$ and $S = X_1 + \ldots + X_n$. Let N be a multivariate normal distribution with mean 0 and covariance matrix Σ. Then

$$\mathcal{W}_1(S, N) \ll_d \delta.$$

Proof. First, we will deduce this from Theorem 10.4.2 when $d = 1$. In this case simply note that

$$\sum_{i=1}^{n} \gamma_i^3 = \sum_{i=1}^{n} \mathbb{E}[|X_i|^3] \leq \sum_{i=1}^{n} \mathbb{E}[\delta |X_i|^2] = \delta \sum_{i=1}^{n} \omega_i^2,$$

showing the claim.

Now in the case $d \geq 1$ the lemma follows by using, as shown in [BG21, Theorem 2.1], that

$$\mathcal{W}_1(S, N) \ll_d \sup_p \mathcal{W}_1(pS, pN),$$

where the supremum is taken over all one dimensional projections p. The result is therefore deduced as in the one dimensional case by using that $\mathbb{E}[|pX_i|^3] \leq \delta \mathbb{E}[|pX_i|^2]$. □

Lemma 10.4.4. *For every positive integer $d \geq 1$ and every $\alpha > 0$ there exists some $C = C(\alpha, d) > 0$ such that the following is true. Let $r > 0$ and let $X_1, X_2, \dots, X_n$ be independent random variables taking values in $\mathbb{R}^d$ such that*

$$|X_i| \leq C^{-1} r \quad \textit{and} \quad \sum_{i=1}^{n} \operatorname{Var} X_i \geq C r^2 I.$$

Then

$$s_r(X_1 + \dots + X_n) \leq \alpha.$$

Proof. Denote for $1 \leq i \leq n$ by $X_i' = X_i - \mathbb{E}[X_i]$ and let $S' = \sum_{i=1}^n X_i'$. Note that $s_r(\sum_{i=1}^n X_i) = s_r(S')$. Write $\Sigma_i = \operatorname{Var} X_i$ and let $\Sigma = \sum_{i=1}^n \Sigma_i$. Let N be a multivariate normal distribution with mean 0 and covariance matrix Σ. Note that $|X_i'| \leq 2C^{-1} r$ almost surely. Therefore by Lemma 10.4.3,

$$\mathcal{W}_1(S', N) \ll_d C^{-1} r.$$

Also

$$s_r(N) \leq s_r(\eta_{C^2 r^2}) = \frac{\|\eta^{(1)}_{C^2 r^2 + r^2}\|}{\|\eta^{(1)}_{r^2}\|} = \frac{1}{C^2 + 1}.$$

Thus by Lemma 10.3.1,

$$s_r(X_1 + \dots + X_n) = s_r(S') \ll_d C^{-1} + \frac{1}{1 + C^2},$$

implying the claim. □

The proof of Proposition 10.4.1 in the case $k \geq 2$ is more involved than the proof in the case $k = 1$. In order to prove this proposition we also need the following lemma and a corollary of it.

Lemma 10.4.5. *Let V be a Euclidean vector space, let $v_1, \ldots, v_n \in V$ and write $S = v_1 + \cdots + v_n$. Let $c_1, c_2 > 0$ be such that for all $i \in [n]$ we have*

$$|v_i| \leq c_1 \quad \text{and} \quad v_i \cdot S \geq c_2 |v_i||S|.$$

Let k be a positive integer. Then we can partition $[n]$ as $J_1 \sqcup J_2 \sqcup \cdots \sqcup J_k$ such that for each $j \in [k]$ we have

$$|S_j - \tfrac{1}{k}S| < c_2^{-1}\sqrt{\tfrac{2c_1}{k}|S|} + 2c_2^{-2}c_1$$

where $S_j = \sum_{i \in J_j} v_i$.

Proof. Choose the J_j such that

$$\sum_{j=1}^{k} |S_j|^2 \tag{10.4.1}$$

is minimized. For each $i \in [n]$ let $j(i)$ denote the unique $j \in [k]$ such that $i \in J_j$. For each $i \in [n]$ and $j' \in [k]$ we know that moving i from $J_{j(i)}$ to $J_{j'}$ cannot decrease the sum in (10.4.1). Therefore

$$|S_{j(i)} - v_i|^2 + |S_{j'} + v_i|^2 \geq |S_{j(i)}|^2 + |S_{j'}|^2.$$

Expanding this out and cancelling gives

$$S_{j(i)} \cdot v_i - |v_i|^2 \leq S_{j'} \cdot v_i$$

and summing over all $i \in J_j$, we get

$$S_j \cdot S_j \leq S_j \cdot S_{j'} + \sum_{i \in J_j} |v_i|^2.$$

Let A_j denote $\sum_{i \in J_j} |v_i|^2$. Note that the above equation gives $|S_j - S_{j'}|^2 \leq A_j + A_{j'}$ and so

$$|S_j - \tfrac{1}{k}S| \leq \max_{j' \in [k]} |S_j - S_{j'}| \leq \sqrt{2 \max_{j' \in [k]} A_{j'}}. \tag{10.4.2}$$

Now let $\Lambda^2 = \max_{j' \in [k]} A_{j'}$. We compute

$$\begin{aligned}
\sum_{i \in J_j} |v_i|^2 &\leq c_2^{-2}|S|^{-2} \sum_{i \in J_j} (v_i \cdot S)^2 \\
&\leq c_2^{-2}|S|^{-2} \sum_{i \in J_j} (v_i \cdot S) c_1 |S| \\
&= c_2^{-2}c_1|S|^{-1} S \cdot S_j \leq c_2^{-2}c_1|S_j| \leq c_2^{-2}c_1(\tfrac{1}{k}|S| + \sqrt{2}\Lambda).
\end{aligned}$$

Therefore $\Lambda^2 \leq c_2^{-2} c_1(|S|/k + \sqrt{2}\Lambda)$, which gives

$$\left(\Lambda - c_2^{-2} c_1/\sqrt{2}\right)^2 \leq c_2^{-2} c_1 |S|/k + c_2^{-4} c_1^2/2$$

and so

$$\begin{aligned}\Lambda &\leq \sqrt{\frac{c_2^{-2} c_1 |S|}{k} + \frac{c_2^{-4} c_1^2}{2}} + \frac{c_2^{-2} c_1}{\sqrt{2}} \\ &\leq c_2^{-1} \sqrt{\tfrac{c_1}{k} |S|} + c_2^{-2} c_1 \sqrt{2},\end{aligned}$$

showing the required result by (10.4.2). □

Corollary 10.4.6. *Let $A_1, \ldots, A_n$ be symmetric positive semi-definite $d \times d$ matrices. Suppose that $\sum_{i=1}^n A_i \geq CkI$ and that for each $i \in [n]$ we have $\|A_i\| \leq c$. Then we can partition $[n]$ as $J_1 \sqcup J_2 \sqcup \cdots \sqcup J_k$ such that for each $j \in [k]$ we have*

$$\sum_{i \in J_j} A_i \geq \left(C - d\sqrt{2cC} - 2d^{3/2} c\right) I.$$

Proof. Let $M = \sum_{i=1}^n A_i$. We know that M is symmetric positive semi-definite and so it may be diagonalised as $M = P^{-1} D P$ for some orthogonal matrix P and a diagonal matrix D with non-zero real entries. Since $M \geq CkI$ all of the diagonal entries of D are at least Ck. Let $D' = \sqrt{CkD^{-1}}$ be a diagonal matrix and for each $i \in [n]$ let $A_i' = QA_iQ$ where $Q = P^{-1} D' P$. Note that A_i' is symmetric positive semi-definite, $\|A_i'\| \leq c$ as $\|Q\| \leq 1$ and that $\sum_{i=1}^n A_i' = CkI$ since

$$QMQ = (P^{-1} D' P)(P^{-1} D P)(P^{-1} D' P) = P^{-1} D' D D' P = CkI.$$

We now apply Lemma 10.4.5 with V being the space of symmetric $d \times d$ matrices with inner product given by $A \cdot B = \sum_{x=1}^n \sum_{y=1}^n A_{xy} B_{xy} = \operatorname{tr} AB$ and with $v_1, \ldots, v_n$ being $A_1', \ldots, A_n'$. We will denote the norm induced by this inner product by $|\cdot|$. Note that given a symmetric matrix A we have that $|A|^2$ is equal to the sum of the squares of the eigenvalues of A and so in particular $\|\cdot\| \leq |\cdot| \leq \sqrt{d}\|\cdot\|$. This means that we can take $c_1 = \sqrt{d}c$ so that $|A_1'| \leq c_1$.

All that we need to do is find some lower bound on $A_i' \cdot CkI$ in terms of $|A_i'| \cdot |CkI|$. Note that $\operatorname{tr} A_i'$ is equal to the sum of the eigenvalues of A_i' and that $|A_i'|^2$ is equal to the sum of the squares of these eigenvalues. In particular since the eigenvalues are non-negative $\operatorname{tr} A_i' \geq |A_i'|$ and so

$$A_i' \cdot CkI = Ck \operatorname{tr} A_i' \geq Ck|A_i'| = |A_i'| \cdot |CkI|/\sqrt{d}.$$

This means that we can take $c_2 = 1/\sqrt{d}$.

We now apply Lemma 10.4.5 with $S = \sum_{i=1}^n A_i' = CkI$ to construct our partition $[n] = J_1 \sqcup J_2 \sqcup \cdots \sqcup J_k$ such that for all $j \in [k]$,

$$\left\| \sum_{i \in J_j} A_i' - CI \right\| \leq \left| \sum_{i \in J_j} A_i' - CI \right| \leq d\sqrt{2cC} + 2d^{3/2}c.$$

Therefore

$$\left\| \sum_{i \in J_j} A_i - CQ^{-2} \right\| \leq (d\sqrt{2cC} + 2d^{3/2}c)||Q^{-2}||$$

and hence,

$$\begin{aligned}\sum_{i \in J_j} A_i &\geq CQ^{-2} - \left(d\sqrt{2cC} + 2d^{3/2}c\right)||Q^{-2}||I \\ &\geq CI - \left(d\sqrt{2cC} - 2d^{3/2}c\right)||Q^{-2}||I \\ &\geq \left(C - d\sqrt{2cC} - 2d^{3/2}c\right) I\end{aligned}$$

using in the penultimate line that $Q^{-2} = P^{-1}(D')^{-2}P$ is symmetric and all eigenvalues are ≥ 1 and in the last line that $||Q^{-1}|| \geq 1$. □

Finally we can prove Proposition 10.4.1.

Proof of Proposition 10.4.1. Note that since $|X_i| \leq C^{-1}r$ almost surely we have $\| \operatorname{Var} X_i \| \leq C^{-2}r^2$. By Corollary 10.4.6 we can partition $[n]$ as $J_1 \sqcup J_2 \sqcup \cdots \sqcup J_k$ such that for each $j \in [k]$ we have

$$\sum_{i \in J_j} \operatorname{Var} X_i \geq \left(C - d\sqrt{2C^{-1}} - 2d^{3/2}C^{-2}\right) r^2 I.$$

This means that by Lemma 10.4.4, provided that C is sufficiently large in terms of d, we know that

$$s_r\left(\sum_{i \in J_j} X_i\right) \leq \alpha.$$

The result now follows from Proposition 10.2.1. □

Chapter 11

Entropy and Variance on General Lie groups

Throughout this section let G be an arbitrary Lie group of dimension ℓ with a fixed choice of Haar measure m_G and let $\mathfrak{g}$ be the Lie algebra of G. We fix an inner product on $\mathfrak{g}$, inducing an associated norm $|\circ|$. Also denote by

$$\log : G \to \mathfrak{g}$$

the logarithm on G, which is defined in a small neighbourhood around the identity.

We study entropy on arbitrary Lie groups. As exposed in the outline of proofs, we shall convert entropy estimates of a random variable Z to estimates of the variance of Z. Indeed, recall that if Z is an absolutely continuous random variable on $\mathbb{R}$ with variance σ^2 then

$$H(Z) \leq \frac{1}{2}\log(2\pi e\sigma^2), \tag{11.0.1}$$

where $H(Z)$ is the differential entropy of Z and equality holds in (11.0.1) if and only if Z is distributed like a Gaussian with variance σ^2. We will prove an analogue of this fact on Lie groups. To do so, for random variables g that are supported within small balls of a given point g_0 we consider the covariance matrix of the Lie group logarithm applied to $g_0^{-1}g$. This viewpoint allows us to apply a higher dimensional analogue of (11.0.1) to deduce an analogous result on G.

Indeed, we recall that for an ℓ-dimensional random variable X we denote by $\mathrm{tr}(X)$ the trace of the covariance matrix of X. In particular, we use the following definition. Given $g_0 \in G$ and a random variable g on G we define

$$\mathrm{tr}_{g_0}(g) = \mathrm{tr}(\log(g_0^{-1}g)),$$

whenever $\log(g_0^{-1}g)$ is defined. The analogue of (11.0.1), which will be proved in Proposition 11.3.1, then amounts to

$$H(g) \leq \frac{\ell}{2} \log\left(\frac{2\pi e}{\ell} \cdot \mathrm{tr}_{g_0}(g)\right) + O_G(\varepsilon) \tag{11.0.2}$$

for random variables supported on $B_\varepsilon(g_0)$ and $\varepsilon > 0$ sufficiently small.

A further goal of this section is to study entropy between scales on G. Indeed, we will define in section 11.5 an explicit family of smoothing distributions $s_{r,a}$ on G, which satisfy

$$\mathrm{tr}_e(s_{r,a}) \approx \ell r^2 \quad \text{and} \quad H(s_{r,a}) = \frac{\ell}{2} \log 2\pi e r^2 + O_\ell(e^{-a^2/4}) + O_{G,a}(r), \tag{11.0.3}$$

while being supported on $B_{ar}(e)$. The error $O_\ell(e^{-a^2/4})$ arises since $s_{r,a}$ is compactly supported while equality holds in (11.0.1) for Gaussians, which are non-compactly supported.

We then define the entropy at a scale $r > 0$ of a random variable as

$$H_a(g; r) = H(gs_{r,a}) - H(s_{r,a})$$

and the entropy between scales between two scales $r_1, r_2 > 0$ as

$$H_a(g; r_1|r_2) = H_a(g; r_1) - H_a(g; r_2).$$

Roughly speaking, $H_a(g; r_1|r_2)$ measures how much more information g has on scale ar_1 than it has on scale ar_2. We work with the parameter a as the uniform bounds (11.0.3) are useful for our purposes.

We next aim to relate the entropy between scales to the trace of a random variable. To do so we introduce the trace $\mathrm{tr}(g; r)$ for a random variable g at scale r, which we define as the supremum of all $t \geq 0$ such that we can find some σ-algebra $\mathscr{A}$ and some $\mathscr{A}$-measurable random variable h taking values in G such that

$$|\log(h^{-1}g)| \leq r \quad \text{and} \quad \mathbb{E}[\mathrm{tr}_h(g|\mathscr{A})] \geq tr^2.$$

Then we show in Proposition 11.5.1 that

$$\mathrm{tr}(g; 2ar) \gg a^{-2}(H_a(g; r|2r) - O_\ell(e^{-a^2/4}) - O_{G,a}(r)). \tag{11.0.4}$$

In section 11.1 we give definitions and discuss basic properties of entropy on G, after which we discuss the Kullback-Leibler divergence on G in section 11.2. In section 11.3 we prove (11.0.2), after which we study conditional entropy in section 11.4. Finally we prove (11.0.4) in section 11.5.

11.1 Entropy and Basic Properties

For notational convenience, we denote

$$h(x) = -x\log(x)$$

for $x \in (0,\infty)$ and recall that h is concave. If $\lambda = \sum_i p_i\delta_{g_i}$ is a discrete probability measure on G, we define the Shannon entropy of λ as

$$H(\lambda) = \sum_i h(p_i).$$

On the other hand, given an absolutely continuous probability measure λ on G with density f_λ we define

$$H(\lambda) = \int h(f_\lambda)\, dm_G.$$

We extend the definition to finite positive measures λ that are either absolutely continuous or discrete by setting

$$H(\lambda) = ||\lambda||_1 H(\lambda/||\lambda||_1).$$

In this subsection we collect some useful basic properties of entropy.

Lemma 11.1.1. *Let $\lambda_1,\dots,\lambda_n$ be absolutely continuous finite measures on G. Then*

$$H(\lambda_1+\ldots+\lambda_n) \geq H(\lambda_1)+\ldots+H(\lambda_n).$$

Proof. It suffices to prove the claim for $n=2$. Let f_1 and f_2 be the densities of λ_1 and λ_2. Then since h is concave

$$\begin{aligned} H(\lambda_1+\lambda_2) &= (||\lambda_1||_1+||\lambda_2||_1)\int h\left(\frac{f_1+f_2}{||\lambda_1||_1+||\lambda_2||_1}\right) dm_G \\ &\geq (||\lambda_1||_1+||\lambda_2||_1)\int \frac{||\lambda_1||_1}{||\lambda_1||_1+||\lambda_2||_1} h\left(\frac{f_1}{||\lambda_1||_1}\right) dm_G \\ &+ (||\lambda_1||_1+||\lambda_2||_1)\int \frac{||\lambda_2||_1}{||\lambda_1||_1+||\lambda_2||_1} h\left(\frac{f_2}{||\lambda_2||_1}\right) dm_G \\ &= H(\lambda_1)+H(\lambda_2). \end{aligned}$$

□

Lemma 11.1.2. *Let $p=(p_1,p_2,\dots)$ be a probability vector and let $\lambda_1,\lambda_2,\dots$ be probability measures on G either all absolutely continuous measures or all discrete measures with finite entropy such that $||\lambda_i|| = p_i$. Then*

$$H\left(\sum_{i=1}^{\infty}\lambda_i\right) \leq H(p)+\sum_{i=1}^{\infty} H(\lambda_i).$$

In particular, if $p_i = 0$ for all $i > k$ for some integer $k \geq 1$ then

$$H\left(\sum_{i=1}^{k} \lambda_i\right) \leq \log k + \sum_{i=1}^{k} H(\lambda_i).$$

Proof. Upon taking limits it suffices to prove the claim for n-dimensional probability vectors $p = (p_1, \ldots, p_n)$ and we only consider the case of absolutely continuous measures as the proof is analogous in the discrete case. We prove the first line in the case when the λ_i are absolutely continuous and denote their densities by f_i. Note that $h(\sum_{i=1}^n a_i) \leq \sum_{i=1}^n h(a_i)$ for any $a_1, \ldots, a_n \geq 0$. Therefore

$$\begin{aligned}
H(\lambda_1 + \ldots + \lambda_n) &= \int h\left(\sum_{i=1}^{n} f_i\right) dm_G \\
&\leq \sum_{i=1}^{n} \int h(f_i)\, dm_G \\
&= \sum_{i=1}^{n} \int (-f_i(x)\log(p_i^{-1} f_i) - f_i(x)\log(p_i))\, dm_G \\
&= \sum_{i=1}^{n} \int p_i h(p_i^{-1} f_i) dm_G + h(p_i) \\
&= \sum_{i=1}^{n} H(\lambda_i) + H(p).
\end{aligned}$$

□

Lemma 11.1.3. *Let λ_1 be a discrete and and λ_2 be continuous probability measures on G. Then*

$$H(\lambda_1 * \lambda_2) \leq H(\lambda_1) + H(\lambda_2)$$

Suppose further that λ_1 is supported on finitely many points with separation at least $2r$ and that the support of λ_2 is contained in a ball of radius r. Then

$$H(\lambda_1 * \lambda_2) = H(\lambda_1) + H(\lambda_2).$$

Proof. Write $\lambda_1 = \sum_{i=1}^n p_i \delta_{g_i}$ and let f be the density of λ_2. Then the density of

$\lambda_1 * \lambda_2$ is given by $\sum_{i=1}^n p_i\, f \circ g_i^{-1}$. As $h(\sum_{i=1}^n a_i) \leq \sum_{i=1}^n h(a_i)$ for any $a_1, \ldots, a_n \geq 0$,

$$\begin{aligned} H(\lambda_1 * \lambda_2) &= \int h\left(\sum_{i=1}^n p_i\, f \circ g_i^{-1}\right) dm_G \\ &\leq \sum_{i=1}^n \int h(p_i\, f \circ g_i^{-1})\, dm_G \\ &= \sum_{i=1}^n \int (p_i\, f \circ g_i^{-1})(\log(p_i) + \log(f \circ g_i^{-1}))\, dm_G \\ &= H(\lambda_1) + H(\lambda_2). \end{aligned}$$

If λ_1 is supported on finitely many points with separation at least $2r$ and that the support of λ_2 is contained in a ball of radius r, then the support of the functions $f \circ g_i^{-1}$ is disjoint and the inequality in the second line is an equality. □

11.2 Kullback-Leibler Divergence

If $\nu \ll \mu$ are measures on G, then we define the Kullback-Leibler divergence as

$$D_{\mathrm{KL}}(\nu \,||\, \mu) = -\int \log \frac{d\nu}{d\mu}\, d\nu.$$

Observe that if ν is absolutely continuous, then $H(\nu) = D_{\mathrm{KL}}(\nu \,||\, m_G)$. We collect some basic results on the Kullback-Leibler divergence on G.

Lemma 11.2.1. *Let $\nu \ll \mu$ be measures on G and assume that ν is a probability measure supported on a set A of positive μ measure. Then*

$$D_{\mathrm{KL}}(\nu \,||\, \mu) \leq \log(\mu(A)).$$

Proof. For convenience write $\nu = f_\nu\, d\mu$. Then by Jensen's inequality,

$$D_{\mathrm{KL}}(\nu \,||\, \mu) = \int_A h\left(f_\nu \frac{\mu(A)}{\mu(A)}\right) d\mu = \int h(f_\nu \mu(A)) \frac{1_A}{\mu(A)}\, d\mu + \log(\mu(A)) \leq \log(\mu(A)).$$

□

Lemma 11.2.2. *Assume that we can write $X = X_1 \times \ldots \times X_m$ as a product of sub-manifolds $X_i \subset X$ and assume that $m_X = m_{X_1} \times \ldots \times m_{X_m}$ for a measure m_X on X and measures m_{X_i} on X_i. Let μ be a probability measure on X with $\mu \ll m_G$. Denote by π_i the projection from X to X_i and by $\pi_i \mu$ the pushforward of μ under π_i. Then*

$$D_{\mathrm{KL}}(\mu \,||\, m_X) \leq D_{\mathrm{KL}}(\pi_1 \mu \,||\, m_{X_1}) + \ldots + D_{\mathrm{KL}}(\pi_m \mu \,||\, m_{X_m}).$$

Proof. It suffices to prove the claim for $m = 2$. Denote by f_μ the density of μ with respect to m_G and write

$$f_\mu^1(x_2) = \int f_\mu(x_1, x_2)\, dm_{X_1}(x_1) \quad \text{and} \quad f_\mu^2(x_1) = \int f_\mu(x_1, x_2)\, dm_{X_2}(x_2).$$

Therefore,

$$\begin{aligned}
D_{\mathrm{KL}}(\mu \,||\, m_G) &= \int\int h(f_\mu(x_1, x_2))\, dm_{X_1}(x_1) dm_{X_2}(x_2) \\
&= \int\int h\left(\frac{f_\mu(x_1, x_2)}{f_\mu^2(x_1)} f_\mu^2(x_1)\right) dm_{X_1}(x_1) dm_{X_2}(x_2) \\
&= \int\int h\left(\frac{f_\mu(x_1, x_2)}{f_\mu^2(x_1)}\right) f_\mu^2(x_1)\, dm_{X_1}(x_1) dm_{X_2}(x_2) \\
&+ \int\int -\log(f_\mu^2(x_1)) f_\mu(x_1, x_2)\, dm_{X_1}(x_1) dm_{X_2}(x_2) \\
&\leq \int h(f_\mu^1(x_2))\, dm_{X_2}(x_2) + \int h(f_\mu^2(x_1))\, dm_{X_1}(x_1) \\
&= D_{\mathrm{KL}}(\pi_1 \mu \,||\, m_{X_1}) + D_{\mathrm{KL}}(\pi_2 \mu \,||\, m_{X_2}),
\end{aligned}$$

having used that h is concave and Jensen's inequality in the penultimate line. □

Lemma 11.2.3. *Let (X, m_X) and (Y, m_Y) be manifolds endowed with Radon measures of full support, and let $\Phi : X \to Y$ be a diffeomorphism with $\Phi_* m_X = m_Y$. Then for a measure $\nu \ll m_X$ on X it holds that*

$$D_{\mathrm{KL}}(\Phi_* \nu || m_Y) = D_{\mathrm{KL}}(\nu || m_X).$$

Proof. Let $f : Y \to \mathbb{R}$ be a continuous compactly supported function. Then

$$\int f\, d\Phi_* \nu = \int (f \circ \Phi)\, d\nu = \int (f \circ \Phi) \frac{d\nu}{dm_X}\, dm_X$$

as well as

$$\begin{aligned}
\int f\, d\Phi_* \nu &= \int f\, \frac{d\Phi_* \nu}{dm_Y}(y) dm_Y \\
&= \int f\, \frac{d\Phi_* \nu}{dm_Y}(y) d\Phi_* m_X = \int (f \circ \Phi) \left(\frac{d\Phi_* \nu}{dm_Y} \circ \Phi\right) dm_X.
\end{aligned}$$

Therefore, as Φ is a diffeomorphism and since m_X has full support,

$$\frac{d\Phi_* \nu}{dm_Y} \circ \Phi = \frac{d\nu}{dm_X}$$

and thus

$$\begin{aligned}
D_{\mathrm{KL}}(\Phi_*\nu||m_Y) &= -\int \log\frac{d\Phi_*\nu}{dm_Y}\,d\Phi_*\nu \\
&= -\int \log\left(\frac{d\Phi_*\nu}{dm_Y}\circ\Phi\right)\,d\nu \\
&= -\int \log\left(\frac{d\nu}{dm_X}\right)\,d\nu = D_{\mathrm{KL}}(\nu||m_X).
\end{aligned}$$

□

Lemma 11.2.4. *Let λ_1 be a probability measure on G and let λ_2 and λ_3 be measures on G such that $\lambda_1 \ll \lambda_2$, $\lambda_1 \ll \lambda_3$ and $\lambda_2 \ll \lambda_3$. Let $U \subset E$ and suppose that the support of λ_1 is contained in U. Then*

$$|D_{\mathrm{KL}}(\lambda_1\,||\,\lambda_2) - D_{\mathrm{KL}}(\lambda_1\,||\,\lambda_3)| \leq \sup_{x\in U}\left|\log\frac{d\lambda_2}{d\lambda_3}\right|.$$

Proof. We calculate

$$\begin{aligned}
|D_{\mathrm{KL}}(\lambda_1\,||\,\lambda_2) - D_{\mathrm{KL}}(\lambda_1\,||\,\lambda_3)| &= \left|\int_U \log\frac{d\lambda_1}{d\lambda_2}\,d\lambda_1 - \int_U \log\frac{d\lambda_1}{d\lambda_3}\,d\lambda_1\right| \\
&\leq \int_U \left|\log\frac{d\lambda_1}{d\lambda_2} - \log\frac{d\lambda_1}{d\lambda_3}\right| d\lambda_1 \\
&= \int_U \left|\log\frac{d\lambda_2}{d\lambda_3}\right| d\lambda_1 \\
&\leq \sup_{x\in U}\left|\log\frac{d\lambda_2}{d\lambda_3}\right|.
\end{aligned}$$

□

11.3 Entropy and Trace

In this subsection we prove (11.0.2). Recall that given $g_0 \in G$ and a random variable g on G we define

$$\mathrm{tr}_{g_0}(g) = \mathrm{tr}(\log(g_0^{-1}g)),$$

whenever $\log(g_0^{-1}g)$ is defined.

Proposition 11.3.1. *Let G be a Lie group of dimension ℓ. Let $\varepsilon > 0$ and suppose that g is a continuous random variable taking values in $B_\varepsilon(g_0)$ for some $g_0 \in G$. If ε is sufficiently small depending on G,*

$$H(g) \leq \frac{\ell}{2}\log\left(\frac{2\pi e}{\ell}\cdot\mathrm{tr}_{g_0}(g)\right) + O_G(\varepsilon).$$

Proof. We first note that if X is an ℓ-dimensional random vector, then

$$H(X) \leq \frac{\ell}{2} \log \left(\frac{2\pi e}{\ell} \cdot \operatorname{tr}(X) \right) \tag{11.3.1}$$

Indeed, it follows from the 1-dimensional case (11.0.1) that $H(X) \leq \frac{1}{2} \log((2\pi e)^{\ell} \cdot |\operatorname{Var}(X)|)$, where $|\operatorname{Var}(X)|$ is the determinant of the covariance matrix. Note that by the AM-GM inequality $|\operatorname{Var}(X)| \leq \operatorname{tr}(X)^{\ell} \ell^{-\ell}$, which implies (11.3.1).

Since $H(g_0^{-1} g) = H(g)$ and $\operatorname{tr}_{g_0}(g) = \operatorname{tr}_e(g_0^{-1} g)$, we may assume without loss of generality that $g_0 = e$. The density $\frac{dm_G}{d(m_{\mathfrak{g}} \circ \log)}$ is smooth and for $\varepsilon > 0$ sufficiently small is $1 + O_G(\varepsilon)$ on $B_\varepsilon(e)$ and therefore $\sup_{B_\varepsilon(e)} \left| \log \frac{dm_G}{d(m_{\mathfrak{g}} \circ \log)} \right| \ll_G \varepsilon$. Thus by Lemma 11.2.4,

$$|D_{\mathrm{KL}}(g \,||\, m_G) - D_{\mathrm{KL}}(g \,||\, m_{\mathfrak{g}} \circ \log)| \ll_G \varepsilon.$$

The claim follows since by (11.3.1)

$$D_{\mathrm{KL}}(g \,||\, m_{\mathfrak{g}} \circ \log) = D_{\mathrm{KL}}(\log(g) \,||\, m_{\mathfrak{g}}) = H(\log(g)) \leq \frac{\ell}{2} \log \left(\frac{2\pi e}{\ell} \operatorname{tr}_e(g) \right).$$

□

11.4 Conditional Entropy and Conditional Trace

The aim of this subsection is to prove an abstract result relating entropy between scales and the trace. To do so, we first discuss conditional entropy and conditional trace. Let Y be a random variable on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and $\mathscr{A} \subset \mathscr{F}$ be a σ-algebra. Denote by $(Y|\mathscr{A})$ the regular conditional distribution as defined in section 9.2. Assuming that $(Y|\mathscr{A})$ is almost surely absolutely continuous, we define

$$H((Y \mid \mathscr{A}))(\omega) = H((Y|\mathscr{A})(\omega)).$$

Recall that if X_1 and X_2 are two random variables then entropy of X_1 given X_2 is $H(X_1|X_2) = H(X_1, X_2) - H(X_2)$. If X_1 and X_2 have finite entropy and finite joint entropy, then by [Vig21],

$$H(X_1|X_2) = \mathbb{E}[H((X_1|X_2))]. \tag{11.4.1}$$

We next give an abstract definition of the entropy at a scale and for a smoothing functions s. Indeed, let g and s be random variables on G and assume that s is absolutely continuous. Then the entropy at scale s is defined as

$$H(g; s_1) = H(g s_1) - H(s_1)$$

Moreover, if s_1 and s_2 are absolutely continuous smoothing functions we define the entropy between scales s_1 and s_2 as

$$H(g; s_1|s_2) = H(g; s_1) - H(g; s_2).$$

The following basic result on the growth of conditional entropy holds.

Lemma 11.4.1. *Let g, s_1, s_2 be independent random variables taking values in G. Assume that s_1 and s_2 are absolutely continuous with finite differential entropy and assume that gs_1 and gs_2 also have finite differential entropy. Then*

$$H(gs_1|gs_2) \geq H(g; s_1|s_2) + H(s_1).$$

Proof. Note that

$$H(gs_2|gs_1) \geq H(gs_2|g, s_1) = H(gs_2|g) = H(s_2)$$

and so

$$H(gs_2, gs_1) = H(gs_2|gs_1) + H(gs_1) \geq H(gs_1) + H(s_2).$$

Therefore

$$\begin{aligned} H(gs_1|gs_2) &= H(gs_2, gs_1) - H(gs_2) \\ &\geq H(gs_1) - H(gs_2) + H(s_2) \\ &\geq H(g; s_1|s_2) + H(s_1). \end{aligned}$$

□

We next define the conditional trace of a random variable on G and relate it to the entropy between scales.

Definition 11.4.2. *Let g be a random variable defined on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and taking values in G. Let $\mathscr{A} \subset \mathscr{F}$ be a σ-algebra let g_0 be a $\mathscr{A}$-measurable random variable taking values on G. Then we denote by $\mathrm{tr}_{g_0}(g \,|\, \mathscr{A})$ the $\mathscr{A}$-measurable function given for $\omega \in \Omega$ by*

$$\mathrm{tr}_{g_0}(g \,|\, \mathscr{A})(\omega) = \mathrm{tr}_{g_0(\omega)}((g \,|\, \mathscr{A})(\omega)).$$

We note here that the variance of a measure μ is defined as the variance of a random variable with law μ. It follows from Proposition 11.3.1 that

$$H((g|\mathscr{A})) \leq \frac{\ell}{2} \log\left(\frac{2\pi e}{\ell} \cdot \mathrm{tr}_{g_0}(g|\mathscr{A})\right) + O_G(\varepsilon). \tag{11.4.2}$$

Theorem 11.4.3. *Let g, s_1 and s_2 be independent absolutely continuous random variables taking values in G and suppose that that s_1 and s_2 are supported on B_ε for some sufficiently small $\varepsilon > 0$ and have finite differential entropy. Write $c = \frac{\ell}{2}\log\frac{2\pi e}{\ell}\mathrm{tr}_e(s_1) - H(s_1)$ and suppose that $\mathrm{tr}_e(s_1) \geq A\varepsilon^2$ for some positive constant A. Then*

$$\mathbb{E}[\mathrm{tr}_{gs_2}(g|gs_2)] \geq \frac{2}{\ell}(H(g; s_1|s_2) - c - C\varepsilon)\mathrm{tr}_e(s_1),$$

where C is some positive constant depending only on A and ℓ.

We first prove some basic result on the trace of the product of two random variables.

Lemma 11.4.4. *Let $\varepsilon > 0$ be sufficiently small and let a, b be random variables and $\mathscr{A}$ a σ-algebra. Suppose that b is independent from a and $\mathscr{A}$ and let g_0 be an $\mathscr{A}$-measurable random variable. Suppose that $g_0^{-1}a$ and b are almost surely contained in B_ε. Then*

$$\mathrm{tr}_{g_0}(ab|\mathscr{A}) = \mathrm{tr}_{g_0}(a|\mathscr{A}) + \mathrm{tr}_e(b) + O(\varepsilon^3).$$

Note that under the assumptions of Lemma 11.4.4 it holds by Lemma 9.2.2 that

$$[ab|\mathscr{A}] = [a|\mathscr{A}][b|\mathscr{A}] = [a|\mathscr{A}]b.$$

Therefore the claim follows from the following unconditional version.

Lemma 11.4.5. *Let $\varepsilon > 0$ be sufficiently small and let g and h be independent random variables taking values in G. Suppose that the image of g is contained in B_ε and the image of h is contained in $B_\varepsilon(h_0)$ for some $h_0 \in G$. Then*

$$\mathrm{tr}_{h_0}(hg) = \mathrm{tr}_{h_0}(h) + \mathrm{tr}_e(g) + O(\varepsilon^3).$$

Proof. Let $X = \log(h_0^{-1}h)$ and let $Y = \log(g)$. Then $|X|, |Y| \leq \varepsilon$ almost surely and by Taylor's theorem there is a random variable E with $|E| \ll \varepsilon^2$ almost surely such that

$$\log(\exp(X)\exp(Y)) = X + Y + E.$$

Therefore

$$\begin{aligned}\mathrm{tr}_{h_0}(hg) &= \mathbb{E}[|X+Y+E|^2] - |\mathbb{E}[X+Y+E]|^2 \\ &= \mathbb{E}[|X+Y|^2] - |\mathbb{E}[X+Y]|^2 \\ &+ 2\mathbb{E}[(X+Y)\cdot E] + \mathbb{E}[|E|^2] - 2\mathbb{E}[X+Y]\mathbb{E}[E] - |\mathbb{E}[E]|^2 \\ &= \mathrm{Var}[X+Y] + O(\varepsilon^3) = \mathrm{Var}[X] + \mathrm{Var}[Y] + O(\varepsilon^3).\end{aligned}$$

□

Proof. (of Theorem 11.4.3) We note that by 11.4.1 and Lemma 11.4.1, it holds that

$$\mathbb{E}[H((gs_1|gs_2))] \geq H(g; s_1|s_2) + H(s_1)$$

and so by (11.4.2),

$$\mathbb{E}\left[\frac{\ell}{2}\log\frac{2\pi e}{\ell}\mathrm{tr}_{gs_2}(gs_1|gs_2)\right] + O(\varepsilon) \geq H(g; s_1|s_2) + H(s_1).$$

Note that $(gs_2)^{-1}g = s_2^{-1}$, which is contained in $B_\varepsilon(e)$. Therefore by Lemma 11.4.4,

$$\mathrm{tr}_{gs_2}(gs_1|gs_2) \leq \mathrm{tr}_{gs_2}(g|gs_2) + \mathrm{tr}_e(s_1) + O(\varepsilon^3)$$

and so

$$H(g; s_1|s_2) + H(s_1) \leq \mathbb{E}\left[\frac{\ell}{2}\log\frac{2\pi e}{\ell}\left(\mathrm{tr}_{gs_2}(g|gs_2) + \mathrm{tr}_e(s_1) + O(\varepsilon^3)\right)\right] + O(\varepsilon).$$

Thus

$$\frac{2}{\ell}\left(H(g; s_1|s_2) - c\right) \leq \mathbb{E}\left[\log\left(1 + \frac{\mathrm{tr}_{gs_2}(g|gs_2)}{\mathrm{tr}_e(s_1)} + O_A(\varepsilon)\right)\right].$$

Using that $\log(1+x) \leq x$ for $x \geq 0$, we conclude the claim. □

11.5 Entropy Between Scales

In this subsection we prove an explicit result relating the entropy between scales and $\mathrm{tr}(g)$. To do so, we construct a suitable family of smoothing functions. Indeed for given $r > 0$ and $a \geq 1$, denote by $\eta_{r,a}$ a random variable on $\mathfrak{g}$ with density function $f_{r,a} : \mathfrak{g} \to \mathbb{R}$ given by

$$f_{r,a}(x) = \begin{cases} C_{r,a}e^{-\frac{|x|^2}{2r^2}} & \text{if } |x| \leq ar, \\ 0 & \text{otherwise,} \end{cases}$$

where $C_{r,a}$ is a normalizing constant to ensure that $f_{r,a}$ integrates to 1. We furthermore define

$$s_{r,a} = \exp(\eta_{r,a}).$$

We then define the entropy at scale r as

$$H_a(g; r) = H(g; s_{r,a}) = H(gs_{r,a}) - H(s_{r,a})$$

and the entropy between scales $r_1, r_2 > 0$ as

$$\begin{aligned} H_a(g; r_1|r_2) &= H(g; s_{r_1,a}|s_{r_2,a}) = H_a(g; r_1) - H_a(g; r_2) \\ &= (H(gs_{r_1,a}) - H(s_{r_1,a})) - (H(gs_{r_2,a}) - H(s_{r_2,a})). \end{aligned}$$

Recall that $\mathrm{tr}(g;r)$ is defined to be the supremum of all $t \geq 0$ such that we can find some σ-algebra $\mathscr{A}$ and some $\mathscr{A}$-measurable random variable h taking values in G such that

$$|\log(h^{-1}g)| \leq r \quad \text{and} \quad \mathbb{E}[\mathrm{tr}_h(g|\mathscr{A})] \geq tr^2.$$

Proposition 11.5.1. *Let g be a random variable taking values in G, let $a \geq 1$ and $r > 0$ be such that ar is sufficiently small in terms of G and assume that $g, s_{r,a}$ and $s_{2r,a}$ are independent random variables. Then*

$$\mathrm{tr}(g; 2ar) \gg a^{-2}(H_a(g; r|2r) - O_\ell(e^{-a^2/4}) - O_{G,a}(r)),$$

for the implied constants depending on G.

Proposition 11.5.1 relies on the following lemma.

Lemma 11.5.2. *The following properties hold for $r > 0$ and $a \geq 1$:*

(i) $\ell r^2 \ll \mathrm{tr}(\eta_{r,a}) \leq \ell r^2$ *and*

$$H(\eta_{r,a}) = \frac{\ell}{2}\log 2\pi e r^2 + O_\ell(e^{-a^2/4}).$$

(ii) If ar is sufficiently small, $\ell r^2 \ll \mathrm{tr}_e(s_{r,a}) \leq \ell r^2$ *and*

$$H(s_{r,a}) = \frac{\ell}{2}\log 2\pi e r^2 + O_\ell(e^{-a^2/4}) + O_{G,a}(r).$$

Proof. We note that (ii) follows from (i) and the claim $\ell r^2 \ll \mathrm{tr}(\eta_{r,a}) \leq \ell r^2$ is obvious. To complete the proof of (i), we deal with $r = 1$ case first. Note first that

$$\int_{x\in\mathbb{R}^\ell, |x|\leq a} e^{-|x|^2/2}\, dx \leq \int_{x\in\mathbb{R}^\ell} e^{-|x|^2/2}\, dx = \prod_{i=1}^{\ell} \int_{\mathbb{R}} e^{-x_i^2/2}\, dx_i = (2\pi)^{\ell/2}$$

and by using spherical coordinates

$$\begin{aligned}\int_{x\in\mathbb{R}^\ell, |x|\geq a} e^{-|x|^2/2}\, dx &= c_\ell \int_a^\infty u^{\ell-1} e^{-u^2/2}\, du \\ &\ll_\ell \int_a^\infty e^{-u^2/3}\, du \leq \int_a^\infty e^{-au/3}\, du = \frac{3}{a} e^{-a^2/3} \ll_\ell e^{-a^2/4}.\end{aligned}$$

Thus we conclude

$$\int_{x\in\mathbb{R}^\ell, |x|\leq a} e^{-|x|^2/2}\, dx = (2\pi)^{\ell/2} - \int_{x\in\mathbb{R}^\ell, |x|\geq a} e^{-||x||^2/2}\, dx \geq (2\pi)^{\ell/2} - O_\ell(e^{-a^2/4})$$

and therefore $C_{1,a} = (2\pi)^{-\ell/2} + O_\ell(e^{-a^2/4})$. We are now in a suitable position to calculate $H(\eta_{1,a})$. Indeed,

$$\begin{aligned} H(\eta_{1,a}) &= \int_{|x|\le a} -C_{1,a} e^{-|x|^2/2} \log\left(C_{1,a} e^{-|x|^2/2}\right) dx \\ &= \int_{|x|\le a} C_{1,a}\left(\frac{|x|^2}{2} - \log C_{1,a}\right) e^{-|x|^2/2}\, dx \end{aligned}$$

We calculate

$$\begin{aligned} &\int_{x\in\mathbb{R}^\ell} C_{1,a}\left(\frac{|x|^2}{2} - \log C_{1,a}\right) e^{-|x|^2/2}\, dx \\ &= (2\pi)^{\ell/2} C_{1,a}\left(\frac{\ell}{2} - \log C_{1,a}\right) \\ &= \left(1 + O_\ell(e^{-a^2/4})\right)\left(\frac{\ell}{2}\log e + \frac{\ell}{2}\log 2\pi + O_\ell(e^{-a^2/4})\right) \\ &= \frac{\ell}{2}\log 2\pi e + O_\ell(e^{-a^2/4}). \end{aligned}$$

and again using spherical coordinates,

$$\begin{aligned} &\int_{|x|\ge a} C_{1,a}\left(\frac{|x|^2}{2} - \log C_{1,a}\right) e^{-|x|^2/2}\, dx \\ &= c_\ell \int_a^\infty C_{1,a}\left(\frac{u^2}{2} - \log C_{1,a}\right) u^{\ell-1} e^{-u^2/2}\, dx \\ &\ll_\ell O_\ell(e^{-a^2/4}). \end{aligned}$$

Thus the claimed bound on $H(\eta_{1,a})$ follows. Since $f_{r,a}(x) = r^\ell C_{1,a} f_{1,a}(x/r)$ it follows that $H(\eta_{r,a}) = \log(r^\ell) + H(\eta_{1,a})$ and hence the proof is complete. □

Proof. (of Proposition 11.5.1) We apply Theorem 11.4.3 to $s_1 = s_{r,a}$ and $s_2 = s_{2r,a}$ and we set $\varepsilon = \ell a r$. By Lemma 11.5.2 (ii) we have that $\mathrm{tr}_e(s_1) \gg \ell r^2 \gg_{a,\ell} \varepsilon^2$ and $c = \frac{\ell}{2}\log\frac{2\pi e}{\ell}\mathrm{tr}_e(s_1) - H(s_1) \le O_\ell(e^{-a^2/4}) + O_{a,\ell}(r)$. Applying Theorem 11.4.3,

$$\mathbb{E}[\mathrm{tr}_{gs_2}(g|gs_2)] \ge cr^2(H(g;r|2r) - O_\ell(e^{-a^2/4}) - O_{G,a}(r))$$

for some absolute constant c depending on G. On the other hand, we have that $|\log((gs_2)^{-1}g)| = |\log s_2| \le 2ar$ and therefore

$$\mathrm{tr}(g;2ar) \ge (2ar)^{-2}\mathbb{E}[\mathrm{tr}_{gs_2}(g|gs_2)] \gg a^{-2}(H(g;r|2r) - O_\ell(e^{-a^2/4}) - O_{G,a}(r)).$$

□

Chapter 12

Entropy Gap and Variance Growth on $\mathrm{Sim}(\mathbb{R}^d)$

In this section we return to $G = \mathrm{Sim}(\mathbb{R}^d)$ with dimension $\ell = \frac{d(d+1)}{2} + 1$. For μ a probability measure on G we denote by $\gamma_1, \gamma_2, \ldots$ independent μ-distributed samples of μ and write

$$q_n = \gamma_1 \cdots \gamma_n.$$

For $\kappa > 0$ be denote by τ_κ the stopping time

$$\tau_\kappa = \inf\{n \,:\, \rho(q_n) \leq \kappa\}.$$

The goal of this section is to give bounds for $\sum_{i=1}^N \mathrm{tr}(q_{\tau_\kappa}, s_i)$ for suitable scales s_i. Towards the proof of our main theorem as discussed in section 8.2, it would be ideal to give a bound roughly of the form

$$\sum_{i=1}^N \mathrm{tr}(q_{\tau_\kappa}, 2^i ar) \gg \frac{h_\mu}{|\chi_\mu|} \log \kappa^{-1} \qquad \text{with} \quad r \approx \kappa^{\frac{S_\mu}{|\chi_\mu|}} \quad \text{and} \quad 2^N r \approx \kappa^{\frac{h_\mu}{2\ell|\chi_\mu|}} \tag{12.0.1}$$

for sufficiently small κ. As we explain below, we can't quite achieve (12.0.1) and the bound we arrive at will also depend on the separation rate S_μ. To estimate the left hand side of (12.0.1) we apply Proposition 11.5.1 to each of the terms $\mathrm{tr}(q_{\tau_\kappa}, 2^i ar)$ which gives

$$\sum_{i=1}^N \mathrm{tr}(q_{\tau_\kappa}, 2^i ar) \gg a^{-2}(H_a(q_{\tau_\kappa}; r|2^N r) + O_d(Ne^{-a^2/4}) + O_d(r)) \tag{12.0.2}$$

having used that by a telescoping sum

$$H_a(q_{\tau_\kappa}; r|2^N r) = \sum_{i=1}^N H_a(q_{\tau_\kappa}; 2^{i-1} r|2^i r).$$

The main contribution from (12.0.1) comes from suitable estimates for $H_a(q_\tau; r|2^N r)$. Indeed, we will show in Proposition 12.1.1 that, up to negligible error terms,

$$H_a(q_{\tau_\kappa}; r|2^N r) \gg \frac{h_\mu}{|\chi_\mu|} \log \kappa^{-1}. \tag{12.0.3}$$

To show this, we recall that

$$H_a(q_{\tau_\kappa}; r|2^N r) = H_a(q_{\tau_\kappa}; r) - H_a(q_{\tau_\kappa}; 2^N r)$$

and therefore we need to estimate the terms $H_a(q_{\tau_\kappa}; r)$ and $H_a(q_{\tau_\kappa}; 2^N r)$. To bound the first term, as we explain after the statement of Lemma 12.1.2, we use that with high probability $\tau_\kappa \approx \log(\kappa^{-1})/|\chi_\mu|$ and so the points in the support of q_{τ_κ} are separated by distance $r \approx \kappa^{\frac{S_\mu}{|\chi_\mu|}} \approx \exp(-S_\mu \tau_\kappa)$. For the second term we use the large deviation principle and the polynomial decay of our self-similar measure.

Combining (12.0.2) with (12.0.3) would lead to (12.0.1) would it not be for the error term $O_d(Ne^{-a^2/\ell})$. Indeed, to not cancel out the lower bound from (12.0.3) we require that

$$Ne^{-a^2/\ell} \leq c\frac{h_\mu}{|\chi_\mu|} \log \kappa^{-1}$$

for a sufficiently small constant c. By our choice of N it holds that $N \approx \frac{S_\mu}{|\chi_\mu|} \log \kappa^{-1}$ and therefore

$$e^{-a^2/\ell} \leq c\frac{h_\mu}{S_\mu}.$$

So we have to set

$$a^2 = c \max\left\{1, \log \frac{S_\mu}{h_\mu}\right\}.$$

Applying then (12.0.2), since the error term $O_d(r)$ is negligible, we conclude that

$$\sum_{i=1}^{N} \operatorname{tr}(q_{\tau_\kappa}, 2^i ar) \gg \frac{h_\mu}{|\chi_\mu|} \log \kappa^{-1} \max\left\{1, \log \frac{S_\mu}{h_\mu}\right\}^{-1}. \tag{12.0.4}$$

We will give a precise proof of the latter bound in Proposition 12.2.2.

12.1 Entropy Gap of Stopped Random Walk

In this subsection we show that the entropy between scales is large for a suitable stopped random walk on $G = \operatorname{Sim}(\mathbb{R}^d)$. Indeed, we establish the following more precise version of (12.0.3).

Proposition 12.1.1. *Let μ be a finitely supported, contracting on average probability measure on G. Suppose that $S_\mu < \infty$ and that $h_\mu/|\chi_\mu|$ is sufficiently large. Let $S > S_\mu$, $\kappa > 0$ and $a \geq 1$ and suppose that $0 < r_1 < r_2 < a^{-1}$ with $r_1 < \exp(-S\log(\kappa^{-1})/|\chi_\mu|)$. Then as $\kappa \to 0$,*

$$H_a(q_{\tau_\kappa}; r_1|r_2) \geq \left(\frac{h_\mu}{|\chi_\mu|} - d\right) \log \kappa^{-1} + H(s_{r_2,a}) + o_{\mu,d,S,a}(\log \kappa^{-1}).$$

Proposition 12.1.1 directly follows from Lemma 12.1.2 and Lemma 12.1.3.

Lemma 12.1.2. *Under the assumptions of Proposition 12.1.1, as $\kappa \to 0$,*

$$H_a(q_{\tau_\kappa}; r_1) \geq \frac{h_\mu}{|\chi_\mu|} \log \kappa^{-1} + o_{\mu,d,S,a}(\log \kappa^{-1}).$$

Recall that $H_a(q_{\tau_\kappa}; r_1) = H(q_{\tau_\kappa} s_{r_1,a}) - H(s_{r_1,a})$. To give the proof idea, note that with high probability $\tau_\kappa \approx \log(\kappa^{-1})/|\chi_\mu|$. Also, by definition of h_μ, we have that $H(q_{\log(\kappa^{-1})/|\chi_\mu|}) \geq h_\mu \log(\kappa^{-1})/|\chi_\mu|$. On the other hand, $s_{r_1,a}$ is mostly contained in a ball around the identity with radius $O(\exp(-S\log(\kappa^{-1})/|\chi_\mu|))$, and therefore by Lemma 11.1.3 we have $H(q_{\log(\kappa^{-1})/|\chi_\mu|} \cdot s_{r_1,a}) = H(q_{\log(\kappa^{-1})/|\chi_\mu|}) + H(s_{r_1,a})$, which implies the claim. We proceed with a more rigorous proof.

Proof. For ease of notation we write in this proof $\tau = \tau_\kappa$. Fix some $\varepsilon > 0$ which is sufficiently small in terms of S and μ. Let $m = \lfloor \log(\kappa^{-1})/|\chi_\mu| \rfloor$ and define τ' as

$$\tau' = \begin{cases} \lceil (1+\varepsilon)m \rceil & \text{if } \tau > \lceil (1+\varepsilon)m \rceil, \\ \lfloor (1-\varepsilon)m \rfloor & \text{if } \tau < \lfloor (1-\varepsilon)m \rfloor, \\ \tau & \text{otherwise.} \end{cases}$$

For a random variable X denote by $\mathcal{L}(X)$ its law. Furthermore, given an event A, we will denote by $\mathcal{L}(X)|_A$ the measure given by the push forward of the restriction of $\mathbb{P}$ to A under the random variable X. Note that $||\mathcal{L}(X)|_A|| = \mathbb{P}[A]$.

By applying Lemma 11.1.1,

$$\begin{aligned} H(q_\tau s_{r_1,a}) &= H(\mathcal{L}(q_\tau) * \mathcal{L}(s_{r_1,a})) \\ &\geq H(\mathcal{L}(q_\tau)|_{\tau=\tau'} * \mathcal{L}(s_{r_1,a})) + H(\mathcal{L}(q_\tau)|_{\tau\neq\tau'} * \mathcal{L}(s_{r_1,a})) \\ &\geq H(\mathcal{L}(q_\tau)|_{\tau=\tau'} * \mathcal{L}(s_{r_1,a})) + \mathbb{P}[\tau \neq \tau'] H(\mathcal{L}(s_{r_1,a})), \end{aligned} \tag{12.1.1}$$

having used that

$$\begin{aligned} H(\mathcal{L}(q_\tau)|_{\tau\neq\tau'} * \mathcal{L}(s_{r_1,a})) &\geq H(\mathcal{L}(q_\tau)|_{\tau\neq\tau'} * \mathcal{L}(s_{r_1,a})|q_\tau) \\ &\geq \mathbb{P}[\tau \neq \tau'] H(\mathcal{L}(s_{r_1,a})). \end{aligned}$$

We next apply that $s_{r_1,a}$ has small support. Set $\delta = \frac{1}{4}(S - S_\mu)$. Write $D_m = \bigcup_{n=1}^m \operatorname{supp}(\mu^{*i})$ for all $m \geq 1$. Then for every N sufficiently large, $\exp(-(S_\mu+\delta)N) < d(x,y)$ for all $x, y \in D_N$. Therefore for ε and κ sufficiently small, $\exp(-(S_\mu+2\delta)m) < d(x,y)$ for all $x, y \in D_{\lceil (1+\varepsilon)m\rceil}$. As $d(s_{r_1,a}, e) \ll_G r_1 a$ it follows that if κ is sufficiently small in terms of μ, a and S,

$$d(s_{r_1,a}, \mathrm{Id}) < O(a\exp(-Sm)) < \frac{1}{2} \min_{x,y\in\operatorname{supp}(q_{\tau'}), x\neq y} d(x,y).$$

In particular, by Lemma 11.1.3,

$$H(\mathcal{L}(q_\tau)|_{\tau=\tau'} * \mathcal{L}(s_{r_1,a})) = H(\mathcal{L}(q_\tau)|_{\tau=\tau'}) + \mathbb{P}[\tau = \tau']H(\mathcal{L}(s_{r_1,a})). \tag{12.1.2}$$

Combining (12.1.2) with (12.1.1),

$$H(q_\tau s_{r_1,a}) \geq H(\mathcal{L}(q_\tau)|_{\tau=\tau}) + H(s_{r_1,a}).$$

It remains to estimate $H(\mathcal{L}(q_\tau)|_{\tau=\tau'})$. Consider the random variable

$$X' = (q_{\lfloor (1-\varepsilon)m\rfloor}, \gamma_{\lfloor (1-\varepsilon)m\rfloor+1}, \gamma_{\lfloor (1-\varepsilon)m\rfloor+2}, \dots, \gamma_{\lceil (1+\varepsilon)m\rceil+1}).$$

As $q_{\tau'}$ is completely determined by X', we have $H(X'|q_{\tau'}) = H(X') - H(q_{\tau'})$.

Let K be the number of points in the support of μ. Note that if

$$\gamma_{\lfloor (1-\varepsilon)m\rfloor+1}, \gamma_{\lfloor (1-\varepsilon)m\rfloor+2}, \dots, \gamma_{\lceil (1+\varepsilon)m\rceil}$$

and τ' are fixed, then for any possible value of $q_{\tau'}$ there is at most one choice of $q_{\lfloor (1-\varepsilon)m\rfloor}$ which would lead to this value of $q_{\tau'}$. Therefore for each y in the image of $q_{\tau'}$ there are at most $(2\varepsilon m+2)K^{2\varepsilon m+2}$ elements x in the image of X' such that $\mathbb{P}[X' = x \cap q_{\tau'} = y] > 0$. Therefore $(X'|q_{\tau'})$ is almost surely supported on less than $(2\varepsilon m+2)K^{2\varepsilon m+2}$ points and hence by (11.4.1),

$$H(X'|q_{\tau'}) \leq \log\left((2\varepsilon m+2)K^{2\varepsilon m+2}\right) \leq \frac{2\varepsilon\log K}{|\chi_\mu|}\log\kappa^{-1} + o_{\mu,\varepsilon}(\log\kappa^{-1}).$$

On the other hand,

$$H(X') \geq H(q_m) \geq h_{\mathrm{RW}}\cdot m \geq \frac{h_{RW}}{|\chi_\mu|}\log\kappa^{-1} - o_\mu(\log\kappa^{-1}) \tag{12.1.3}$$

and therefore

$$H(q_{\tau'}) \geq \frac{h_{RW} - 2\varepsilon\log K}{|\chi_\mu|}\log\kappa^{-1} - o_{\mu,\varepsilon}(\log\kappa^{-1}).$$

To continue, we note that by Lemma 11.1.2,

$$H(q_{\tau'}) \le H(\mathcal{L}(q_{\tau'})|_{\tau=\tau'}) + H(\mathcal{L}(q_{\tau'})|_{\tau\neq\tau'}) + \log 2. \tag{12.1.4}$$

We wish to bound $H(\mathcal{L}(q_{\tau'})|_{\tau=\tau'})$ from below. By the large deviation principle, $\mathbb{P}[\tau \neq \tau'] \le \alpha^m$ for $\alpha \in (0,1)$ only depending on ε and μ. We also know that conditional on $\tau \neq \tau'$, there are at most $2K^{\lceil (1+\varepsilon)m \rceil}$ possible values for $q_{\tau'}$ and therefore

$$H(\mathcal{L}(q_{\tau'})|_{\tau\neq\tau'}) \le \alpha^m \log\left(2K^{\lceil (1+\varepsilon)m \rceil}\right) = o_{\mu,\varepsilon}(\log \kappa^{-1}).$$

This implies

$$H(\mathcal{L}(q_{\tau'})|_{\tau=\tau'}) \ge \frac{h_{RW} - 2\varepsilon \log K}{|\chi_\mu|} \log \kappa^{-1} - o_{\mu,\varepsilon}(\log \kappa^{-1}).$$

Since ε can be made arbitrarily small, the claim follows. □

Lemma 12.1.3. *Under the assumptions of Proposition 12.1.1, as $\kappa \to 0$,*

$$H(q_{\tau_\kappa} s_{r_2,a}) \le d \log \kappa^{-1} + o_{\mu,d,a}(\log \kappa^{-1}).$$

Proof. As in the proof of Lemma 12.1.2, write $\tau = \tau_\kappa$ and $K = |\mathrm{supp}(\mu)|$. We use the product structure on G combined with Lemma 11.2.2. Indeed, note that a choice of Haar measure on G is given as

$$\int f \, dm_G = \int f(\rho U + b)\, \rho^{-(d+1)} d\rho dU db,$$

for dr, db the Lebesgue measure and dU the Haar probability measure on $O(d)$. Therefore by Lemma 11.2.2, $H(q_\tau s_{r_2,a}) \le$

$$D_{\mathrm{KL}}(\rho(q_\tau s_{r_2,a}) \,||\, \rho^{-(d+1)} d\rho) + D_{\mathrm{KL}}(U(q_\tau s_{r_2,a}) \,||\, dU) + D_{\mathrm{KL}}(b(q_\tau s_{r_2,a}) \,||\, db).$$

We give suitable bounds for each these terms. As dU is a probability measure $D_{\mathrm{KL}}(U(q_\tau s_{r_2,a}) \,||\, dU) \le 0$ by Lemma 11.2.1.

We next deal with $D_{\mathrm{KL}}(b(q_\tau s_{r_2,a}) \,||\, db)$. Denote by ν_τ the distribution of $b(q_\tau s_{r_2,a})$. We claim that there is $\alpha = \alpha(\mu, d, a)$ such that

$$\nu_\tau(B_R^c) \le R^{-\alpha} \tag{12.1.5}$$

for all sufficiently small κ and sufficiently large R. Note that

$$|b(q_\tau s_{r_2,a})| = |\rho(q_\tau) U(q_\tau) b(s_{r_2,a}) + b(q_\tau)| \le \kappa |b(s_{r_2,a})| + |b(q_\tau)|$$

and therefore it suffices to show (12.1.5) for the distribution of $b(q_\tau)$, which we denote by ν'_τ. For $x \in \mathbb{R}^d$,

$$|b(q_\tau) - q_\tau(x)| \le |q_\tau(0) - q_\tau(x)| \le \rho(q_\tau)|x| \le \kappa|x|$$

and so $|b(q_\tau)| \le |q_\tau(x)| + \kappa|x|$. Therefore if $R \le |b(q_\tau)|$ then either $R/2 \le |q_\tau(x)|$ or $R/2 \le \kappa|x|$. Also note that if x is sampled from ν independently from $\gamma_1, \gamma_2, \ldots$, so is $q_\tau(x)$. By (7.0.2) this implies that

$$\nu'_\tau(B_R^c) \le \nu(B_{R/2}^c) + \nu(B_{R/2\kappa}^c) \le R^{-\alpha_2} 2^{\alpha_2}(1+\kappa^{-1})^{-\alpha_2},$$

showing (12.1.5).

To conclude we deduce from (12.1.5) that $D_{\mathrm{KL}}(\nu_\tau \,||\, db)$ is bounded by a constant depending on μ, d and a and therefore is $\le o_{\mu,d,a}(\log \kappa^{-1})$. Indeed denote by f_τ the density of ν_τ such that

$$D_{\mathrm{KL}}(\nu_\tau \,||\, db) = \int -f_\tau \log f_\tau \, dm_{\mathbb{R}^d}.$$

Also let $L > 1$ be a constant and for $i = 0, 1, 2, \ldots$ write $p_i = \nu_\tau(B_{L^{i+1}} \backslash B_{L^i})$ such that $p_i \le \nu_\tau(B_{L^i}^c) \le L^{-i\alpha}$. Thus it holds by Jensen's inequality for $h(x) = -x \log x$,

$$\begin{aligned} D_{\mathrm{KL}}(\nu_\tau \,||\, db) &= \sum_{i \ge 0} \int_{B_{L^{i+1}} \backslash B_{L^i}} -f_\tau \log f_\tau \, dm_{\mathbb{R}^d} \\ &= \sum_{i \ge 0} \int_{B_{L^{i+1}} \backslash B_{L^i}} -f_\tau \log \left(\frac{f_\tau p_i}{p_i} \right) dm_{\mathbb{R}^d} \\ &= \sum_{i \ge 0} \left(\int h(f_\tau p_i) \frac{1_{B_{L^{i+1}} \backslash B_{L^i}}}{p_i} \, dm_{\mathbb{R}^d} + p_i \log(p_i) \right) \\ &\le \sum_{i \ge 0} h(p_i) \le \sum_{0 \le i \le I} h(p_i) + \sum_{i \ge I} h(L^{-i\alpha}) < \infty, \end{aligned}$$

having used in the last line that $\log(p_i) \le 0$ and that $h(x)$ is monotonically decreasing for small x and therefore $h(p_i) \le h(L^{-i\alpha})$ for $i \ge I$ with I sufficiently large.

Finally, we estimate $D_{\mathrm{KL}}(\rho(q_{\tau_\kappa} s_{r_2,a}) \,||\, \rho^{-(d+1)} d\rho)$. Fix $\varepsilon > 0$ and let A be the event that $\rho(q_\tau) \ge \kappa^{(1+\varepsilon)}$. By Lemma 9.3.2 there is $\delta > 0$ only depending on μ and ε such that $\mathbb{P}[A^c] \le \kappa^\delta$. By Lemma 11.2.1,

$$\begin{aligned} D_{\mathrm{KL}}(\mathcal{L}(\rho(q_{\tau_\kappa} s_{r_2,a}))|_A \,||\, \rho^{-(d+1)} d\rho) &\le \log \left(\int_{\kappa^{1+\varepsilon}}^\infty \rho^{-(d+1)} \, d\rho \right) \\ &= \log \left(d^{-1} \kappa^{-d(1+\varepsilon)} \right) \le d(1+\varepsilon) \log \kappa^{-1}. \end{aligned}$$

To bound $H(\mathcal{L}(q_\tau s_{r_2,a})|_{A^c})$, we note that as in Lemma 11.1.3 it suffices to bound the Shannon entropy of $H(\mathcal{L}(q_\tau)|_{A^c})$. If $\tau \leq 2\frac{\log\kappa^{-1}}{|\chi_\mu|}$, the contribution can be bounded by $\kappa^{\delta\frac{2\log\kappa^{-1}}{|\chi_\mu|}}\log K$. By the large deviation principle, when $n \geq 2\frac{\log\kappa^{-1}}{|\chi_\mu|}$ it holds that $\mathbb{P}[\tau = n] \leq \alpha^n$ for some $\alpha \in (0,1)$. Therefore the contribution in this case is $\leq \alpha^n n \log K$ where $\alpha \in (0,1)$ is some constant depending on μ. Summing over all $n \geq 2\frac{\log\kappa^{-1}}{|\chi_\mu|}$ and using Lemma 11.1.2, we conclude that $H(\mathcal{L}(q_\tau s_{r_2,a})|_{A^c})$ is bounded and therefore $o_{\mu,\varepsilon}(\log\kappa^{-1})$. As $\varepsilon > 0$ was arbitrary the claim follows. □

12.2 Trace Bounds for Stopped Random Walk

In this subsection we give a precise proof of (12.0.4) following the sketch given at the beginning of this section. We first convert Proposition 12.1.1 into an integral bound.

Proposition 12.2.1. *Let μ be a finitely supported, contracting on average probability measure on $G = \mathrm{Sim}(\mathbb{R}^d)$ and write $\ell = \dim G = \frac{d(d+1)}{2}+1$. Suppose that $S_\mu < \infty$ and that $h_\mu/|\chi_\mu|$ is sufficiently large. Let $S > S_\mu$ and suppose that S is chosen sufficiently large such that $h_\mu \leq S$. Then for sufficiently small κ,*

$$\int_{\kappa^{\frac{S}{|\chi_\mu|}}}^{\kappa^{\frac{h_\mu}{2\ell|\chi_\mu|}}} \frac{1}{u}\mathrm{tr}(q_{\tau_\kappa}; u)\, du \gg \left(\frac{h_\mu}{|\chi_\mu|}\right) \max\left\{1, \log\frac{S}{|\chi_\mu|}\right\}^{-1} \log\kappa^{-1}.$$

Proof. Let $\tau = \tau_\kappa$ and let $a \geq 1$ to be determined. Let

$$r_1 = a^{-1}\kappa^{\frac{S}{|\chi_\mu|}} = a^{-1}\exp\left(-\frac{S}{|\chi_\mu|}\log\kappa^{-1}\right)$$

and

$$N = \left\lfloor \left(\frac{S}{|\chi_\mu|} - \frac{h_\mu}{2\ell|\chi_\mu|}\right)\frac{\log\kappa^{-1}}{\log 2}\right\rfloor - 1.$$

Note that

$$\frac{1}{4}\frac{\kappa^{\frac{h_\mu}{2\ell|\chi_\mu|}}}{ar_1} = \frac{1}{4}\frac{\kappa^{-\frac{S}{|\chi_\mu|}}}{\kappa^{-\frac{h_\mu}{2\ell|\chi_\mu|}}} \leq 2^N \leq \frac{1}{2}\frac{\kappa^{-\frac{S}{|\chi_\mu|}}}{\kappa^{-\frac{h_\mu}{2\ell|\chi_\mu|}}} = \frac{1}{2}\frac{\kappa^{\frac{h_\mu}{2\ell|\chi_\mu|}}}{ar_1}.$$

Given $u \in [1,2)$ and an integer $0 \leq i \leq N-1$ let

$$k_i(u) = H_a(q_\tau; 2^{i-1}ur_1 | 2^i ur_1).$$

Then by Proposition 11.5.1, there is some constant $c = c(d) > 0$ depending only on d such that

$$\mathrm{tr}(q_\tau; a2^i ur_1) \geq ca^{-2}(k_i(u) - O_d(e^{-\frac{a^2}{4}}) - O_{d,a}(2^i r_1)). \tag{12.2.1}$$

Thus

$$\sum_{i=1}^{N} \operatorname{tr}(q_\tau; a2^i ur_1) \geq ca^{-2} \sum_{i=1}^{N} k_i(u) - O_d(Ne^{-\frac{a^2}{4}} a^{-2}) - O_{d,a}(N2^N r_1).$$

Note that for $u \in [1, 2)$ we have $a2^N ur_1 \leq \kappa^{\frac{h_\mu}{2\ell|\chi_\mu|}}$ and $aur_1 \geq \kappa^{\frac{S}{|\chi_\mu|}}$. Therefore,

$$\begin{aligned}
&\int_{\kappa^{\frac{S}{|\chi_\mu|}}}^{\kappa^{\frac{h_\mu}{2\ell|\chi_\mu|}}} \frac{1}{u} \operatorname{tr}(q_\tau; u)\, du \\
\geq& \sum_{i=1}^{N} \int_{a2^i ur_1}^{a2^{i+1} ur_1} \frac{1}{u} \operatorname{tr}(q_\tau; u)\, du \\
\geq& \sum_{i=1}^{N} \int_{1}^{2} \frac{1}{u} \operatorname{tr}(q_\tau; a2^i ur_1)\, du \\
\geq& ca^{-2} \int_{1}^{2} \frac{1}{u} \left(\sum_{i=1}^{N} k_i(u) - O_d(Ne^{-\frac{a^2}{4}} a^{-2}) - O_{d,a}(N2^N r_1) \right) du. \qquad (12.2.2)
\end{aligned}$$

Observe that $\sum_{i=1}^{N} k_i(u) = H_a(q_{\tau_\kappa}; ur_1 | 2^N ur_1)$ and therefore by Proposition 12.1.1 and Lemma 11.5.2,

$$\begin{aligned}
\sum_{i=1}^{N} k_i(u) &\geq \left(\frac{h_\mu}{|\chi_\mu|} - d \right) \log \kappa^{-1} + \ell \cdot \log 2^N ur_1 + o_{\mu,d,S,a}(\log \kappa^{-1}) \\
&\geq \left(\frac{h_\mu}{|\chi_\mu|} - d - \frac{h_\mu}{2|\chi_\mu|} \right) \log \kappa^{-1} + o_{\mu,d,S,a}(\log \kappa^{-1}). \qquad (12.2.3)
\end{aligned}$$

Let $C = C(d)$ be chosen such that the error term $O(Ne^{-\frac{a^2}{4}} a^{-2})$ in (12.2.2) can be bounded above by $CNe^{-\frac{a^2}{4}} a^{-2}$. Note that this is at most $C \frac{S}{|\chi_\mu| \log 2} e^{-\frac{a^2}{4}} a^{-2} \log \kappa^{-1}$. Let c be as in (12.2.1). We take our value of a to be

$$a = 2\sqrt{\log \left(\frac{4C}{c \log 2} \frac{S}{h_\mu} \right)}.$$

Then

$$CNe^{-\frac{a^2}{4}} a^{-2} \leq ca^{-2} \frac{h_\mu}{4|\chi_\mu|} \log \kappa^{-1}.$$

We also note that $N2^N r_1 \leq o_{\mu,d,S}(\log \kappa^{-1})$. Therefore combining (12.2.2) and (12.2.3),

$$\int_{\kappa^{\frac{S}{|\chi_\mu|}}}^{\kappa^{\frac{h_\mu}{2\ell|\chi_\mu|}}} \frac{1}{u} \operatorname{tr}(q_\tau; u)\, du \geq ca^{-2} \left(\frac{h_\mu}{|\chi_\mu|} - d - \frac{h_\mu}{2|\chi_\mu|} - \frac{h_\mu}{4|\chi_\mu|} \right) \log \kappa^{-1} + o_{\mu,d,S}(\log \kappa^{-1}).$$

Note further that $a^2 \ll_d \max\{1, \log \frac{S}{h_\mu}\}$. Thus we have for all sufficiently small κ (depending on μ and M),

$$\int_{\kappa^{\frac{S}{|\chi_\mu|}}}^{\kappa^{\frac{h_\mu}{2\ell|\chi_\mu|}}} \frac{1}{u}\mathrm{tr}(q_\tau; u)\, du \gg_d \left(\frac{h_\mu}{|\chi_\mu|}\right) \max\left\{1, \log \frac{S}{h_\mu}\right\}^{-1} \log \kappa^{-1}.$$

□

Finally we prove the following more precise version of (12.2.2). We show further that $s_{i+1} \geq \kappa^{-3} s_i$ in order to apply Proposition 11.5.2 to concatenate proper decompositions as defined and discussed in section 13.

Proposition 12.2.2. *Let μ be a finitely supported, contracting on average probability measure on $G = \mathrm{Sim}(\mathbb{R}^d)$ and write $\ell = \dim G = \frac{d(d+1)}{2} + 1$. Suppose that $S_\mu < \infty$ and that $h_\mu / |\chi_\mu|$ is sufficiently large. Let $S > S_\mu$ be chosen large enough that $S \geq h_\mu$. Suppose that κ is sufficiently small (depending on μ and S) and let $\widehat{m} = \lfloor \frac{S}{100|\chi_\mu|} \rfloor$.*

Then there exist $s_1, s_2, \dots, s_{\widehat{m}} > 0$ such that for each $i \in [\widehat{m}]$,

$$s_i \in \left(\kappa^{\frac{S}{|\chi_\mu|}}, \kappa^{\frac{h_\mu}{2\ell|\chi_\mu|}}\right)$$

and for each $i \in [\widehat{m} - 1]$ $s_{i+1} \geq \kappa^{-3} s_i$ and

$$\sum_{i=1}^{\widehat{m}} \mathrm{tr}(q_{\tau_\kappa}; s_i) \gg_d \left(\frac{h_\mu}{|\chi_\mu|}\right) \max\left\{1, \log \frac{S}{h_\mu}\right\}^{-1}.$$

Proof. Let $A = \kappa^{\frac{h_\mu}{4\widehat{m}\ell|\chi_\mu|} - \frac{S}{2\widehat{m}|\chi_\mu|}}$. Define $a_1, a_2, \dots, a_{2\widehat{m}+1}$ by $a_i = \kappa^{\frac{S}{|\chi_\mu|}} A^{i-1}$. Therefore $a_1 = \kappa^{\frac{S}{|\chi_\mu|}}$ and $a_{2\widehat{m}+1} = \kappa^{\frac{h_\mu}{2\ell|\chi_\mu|}}$. Furthermore, provided $h_\mu/|\chi_\mu|$ is sufficiently large, we have $\kappa^{-3} \leq A \leq \kappa^{-50}$. In particular $a_{i+1} \geq \kappa^{-3} a_i$.

Let U and V be defined by

$$U = \bigcup_{i=1}^{\widehat{m}} [a_{2i-1}, a_{2i}) \quad \text{and} \quad V = \bigcup_{i=1}^{\widehat{m}} [a_{2i}, a_{2i+1}).$$

Without loss of generality, upon replacing U with V, by Proposition 12.2.1

$$\int_U \frac{1}{u}\mathrm{tr}(q_{\tau_\kappa}; u)\, du \gg_d \left(\frac{h_\mu}{|\chi_\mu|}\right) \max\left\{1, \log \frac{S}{|\chi_\mu|}\right\}^{-1} \log \kappa^{-1}.$$

For $i \in [\widehat{m}]$ let $s_i \in (a_{2i-1}, a_{2i})$ be chosen such that

$$\mathrm{tr}(q_{\tau_\kappa}; s_i) \geq \frac{1}{2} \sup_{u \in (a_{2i-1}, a_{2i})} \mathrm{tr}(q_{\tau_\kappa}; u).$$

In particular,

$$\operatorname{tr}(q_{\tau_\kappa}; s_i) \geq \frac{1}{2\log A} \int_{a_{2i-1}}^{a_{2i}} \frac{1}{u} \operatorname{tr}(q_{\tau_\kappa}; u)\, du.$$

Summing over i gives

$$\begin{aligned} \sum_{i=1}^{\widehat{m}} \operatorname{tr}(q_{\tau_\kappa}; s_i) &\geq \frac{1}{2\log A} \int_U \frac{1}{u} \operatorname{tr}(q_{\tau_\kappa}; u)\, du \\ &\geq \frac{c}{2\log A} \left(\frac{h_\mu}{|\chi_\mu|} \right) \max \left\{ 1, \log \frac{S}{|\chi_\mu|} \right\}^{-1} \log \kappa^{-1}. \end{aligned}$$

As $\log A \asymp \log \kappa^{-1}$ it follows that, provided that κ is sufficiently small depending on μ, d, S,

$$\sum_{i=1}^{\widehat{m}} \operatorname{tr}(q_{\tau_\kappa}; s_i) \gg_d \left(\frac{h_\mu}{|\chi_\mu|} \right) \max \left\{ 1, \log \frac{S}{|\chi_\mu|} \right\}^{-1}.$$

Finally we note that as $A \geq \kappa^{-3}$ we have that $s_{i+1} \geq \kappa^{-3} s_i$. □

Chapter 13

Decomposition of Stopped Random Walk

In this section Theorem 8.1.4 is proved. We construct samples from ν in a suitable way in order to bound the order k detail of ν. Given a probability measure μ on $G = \mathrm{Sim}(\mathbb{R}^d)$ we denote by $\gamma_1, \gamma_2, \ldots$ independent μ-distributed random variables and write $q_n = \gamma_1 \cdots \gamma_n$. Recall that if x is distributed like ν and τ is a stopping time, then by Lemma 2.24 from [Kit23] the random variable $q_\tau x$ is distributed like ν.

As discussed in the outline of proofs, one uses Proposition 12.2.2 to make a decomposition

$$q_{\tau_\kappa} x = g_1 \exp(U_1) g_2 \exp(U_2) \cdots g_n \exp(U_n) x \tag{13.0.1}$$

with a suitable $\kappa > 0$ and integer $n \geq 1$ that satisfies for $1 \leq i \leq n$,

$$|U_i| \leq \rho(g_1 \cdots g_i)^{-1} r \quad \text{and} \quad \sum_{i=1}^{n} \mathrm{tr}(\rho(g_1 \cdots g_i) U_i) \geq C r^2 \tag{13.0.2}$$

for a sufficiently large constant C and a given scale $r > 0$. The definition of $\mathrm{tr}(q_{\tau_\kappa}, s_i)$ requires us to work with a σ-algebra $\mathscr{A}$ and with the conditional trace in (13.0.2). As stated in (8.2.9), we need to have (13.0.2) at $O(\log \log r^{-1})$ many suitable times κ_i.

Indeed, in order to deduce (13.0.2) from Proposition 12.2.2 we need to combine all the information at the scales $s_1, \ldots, s_{\hat{m}}$. One also needs to ensure that the assumptions from the Taylor-approximation result Proposition 9.1.4 are satisfied for each scale s_i and that we can apply our (c, T)-well-mixing and (α_0, θ, A)-non-degeneracy conditions to deduce that

$$\mathrm{Var}(\zeta_i(U_i)) \geq c_1 \mathrm{tr}(\rho(g_1 \cdots g_i) U_i) I$$

for c_1 a constant depending on $d, c, T, \alpha_0, \theta$ and A. We will achieve the latter by ensuring that each g_i is a product of sufficiently many γ_i so that $g_i x$ is in distribution sufficiently close to ν.

To combine the trace bounds at the various scales while ensuring that the above conditions are satisfied, a theory of decompositions of the form (13.0.1) will be developed. We call decompositions (13.0.1) satisfying suitable properties *proper decompositions*. It is important for our purposes to track the amount of variance we can gain from a given proper decomposition, which is a quantity we will call the variance sum and denote by $V(\mu, n, K, \kappa, A; r)$ (see definition 13.1.2 for the various parameters).

In section 13.2 we will show that there exist proper decompositions that allow us to compare the variance sum V and tr. Proper decompositions can be concatenated in such a way that the variance sum is additive, as is shown in section 13.3. We establish how to convert an estimate on the variance sum V into an estimate for detail in section 13.4. The proof of Theorem 8.1.4 culminates in section 13.5 combining the previous results. Finally, we establish Theorem 8.1.5 in section 13.6.

13.1 Proper Decompositions

Definition 13.1.1. *Let μ be a probability measure on G, let $n, K \in \mathbb{Z}_{\geq 0}$ and let $A, r > 0$ and $r \in (0,1)$. Then a **proper decomposition** of (μ, n, K, A) at scale r consists of the following data*

(i) $f = (f_i)_{i=1}^n$ and $h = (h_i)_{i=1}^n$ random variables taking values in G,

(ii) $U = (U_i)_{i=1}^n$ random variables taking values in $\mathfrak{g}$,

(iii) $\mathscr{A}_0 \subset \mathscr{A}_1 \subset \ldots \subset \mathscr{A}_n$ a nested sequence of σ-algebras,

(iv) $\gamma = (\gamma_i)_{i=1}^\infty$ be i.i.d. samples from μ and let $\mathscr{F} = (\mathscr{F}_i)_{i=1}^\infty$ be a filtration for γ with γ_{i+1} being independent from $\mathscr{F}_i$ for $i \geq 1$,

(v) stopping times $S = (S_i)_{i=1}^n$ and $T = (T_i)_{i=1}^n$ for the filtration $\mathscr{F}$,

(vi) $m = (m_i)_{i=1}^n$ non-negative real numbers,

satisfying the following properties:

A1 *The stopping times satisfy*

$$S_1 \leq T_1 \leq S_2 \leq T_2 \leq \ldots \leq S_n \leq T_n,$$

$S_1 \geq K$ as well as $S_i \geq T_{i-1} + K$ and $T_i \geq S_i + K$ for $i \in [n]$,

A2 *We have* $f_1 \exp(U_1) = \gamma_1 \dots \gamma_{S_1}$ *and for* $2 \leq i \leq n$ *we have* $f_i \exp(U_i) = \gamma_{T_{i-1}+1} \cdots \gamma_{S_i}$. *Furthermore for each* i *we have that* f_i *is* $\mathscr{A}_i$*-measurable,*

A3 $h_i = \gamma_{S_i+1} \cdots \gamma_{T_i}$ *and* h_i *is* $\mathscr{A}_i$*-measurable,*

A4 $\rho(f_i) < 1$ *for all* $1 \leq i \leq n$*,*

A5 *Whenever* $|b(h_i)| > A$*, we have* $U_i = 0$*,*

A6 *For each* $1 \leq i \leq n$ *we have*

$$|U_i| \leq \rho(f_1 h_1 f_2 h_2 \cdots h_{i-1} f_i)^{-1} r,$$

A7 *For each* $1 \leq i \leq n$*, we have that* U_i *is conditionally independent of* $\mathscr{A}_n$ *given* $\mathscr{A}_i$*,*

A8 *The* U_i *are conditionally independent given* $\mathscr{A}_n$*,*

A9 *For each* $1 \leq i \leq n$*, it holds*

$$\mathbb{E}\left[\frac{\mathrm{Var}(\rho(f_i)U(f_i)U_i b(h_i)|\mathscr{A}_i)}{\rho(f_1 h_1 f_2 h_2 \cdots f_{i-1} h_{i-1})^{-2} r^2} \,|\, \mathscr{A}_{i-1}\right] \geq m_i I.$$

Note that in **A9** by Var we mean the covariance matrix and we are using the ordering given by positive semi-definiteness (8.3.1) and we denote, as in section 9.1, by $U_i b(h_i) = \psi_{b(h_i)}(U_i)$.

A proper decomposition as above gives us

$$\gamma_1 \cdots \gamma_{T_n} = f_1 \exp(U_1) h_1 f_2 \exp(U_2) h_2 \cdots h_{n-1} f_n \exp(U_n) h_n \tag{13.1.1}$$

We briefly comment on the various properties of proper decompositions. We use parameter K and **A1** to ensure that each of the $f_i x$ and $h_i x$ for $x \in \mathbb{R}^d$ are close in distribution to ν. Properties **A4**, **A5** and **A6** are needed in order to apply Proposition 9.1.4. We require **A7** so that we have $\mathrm{Var}(U_i|\mathscr{A}_n) = \mathrm{Var}(U_i|\mathscr{A}_i)$ and in particular the latter is a $\mathscr{A}_i$-measurable random variable. **A8** is needed so that $[U_1|\mathscr{A}_n], \dots, [U_n|\mathscr{A}_n]$ are independent random variables and therefore we can apply Proposition 10.4.1.

One works with two sequences of random variables f and h instead of one in order to be able to concatenate proper decompositions as in Proposition 13.3.1. Indeed, if we had proper decompositions of the form

$$\gamma_1 \cdots \gamma_{T_n} = g_1 \exp(U_1) g_2 \exp(U_2) g_3 \cdots g_n \exp(U_n) g_{n+1}$$

we could show a variant of (13.3.1) and all other results on proper decompositions. However we could not prove anything like Proposition 13.3.1, whose flexible choice of the parameter M is necessary to apply Proposition 12.2.2.

We next define the V function mentioned above. The additional parameter $\kappa > 0$ is introduced in order to be able to concatenate the decompositions in a suitable way (Proposition 13.3.1).

Definition 13.1.2. *Given (μ, n, K, A) and $\kappa, r > 0$ we denote by*

$$V(\mu, n, K, \kappa, A; r)$$

*the **variance sum** defined as the supremum for $k = 0, 1, 2, \ldots, n$ of all possible values of*

$$\sum_{i=1}^{k} m_i$$

for a proper decomposition of (μ, k, K, A) at scale r with $\rho(f_1 h_1 \cdots f_k h_k) \geq \kappa$ almost surely.

It is clear that for any $\kappa' > 0$ with $\kappa' \leq \kappa$ we have

$$V(\mu, n, K, \kappa', A; r) \geq V(\mu, n, K, \kappa, A; r). \tag{13.1.2}$$

13.2 Existence of Proper Decompositions

We show that for a suitable dependence of the involved parameters, we can construct proper decompositions comparing the variance sum and the trace.

Proposition 13.2.1. *Let $d \in \mathbb{Z}_{\geq 1}$ and $c, T, \alpha_0, \theta, A, R > 0$ with $c, \alpha_0 \in (0, 1)$ and $T \geq 1$. Then there exists $c_1 = c_1(d, R, c, T, \alpha_0, \theta, A) > 0$ such that the following is true. Let μ be a contracting on average, (c, T)-well-mixing and (α_0, θ, A)-non-degenerate probability measure on G such that $\rho(g) \in [R^{-1}, R]$ for all $g \in \mathrm{supp}(\mu)$.*

Let $\kappa, s > 0$ with κ and s sufficiently small (in terms of μ and R). Let K be sufficiently large in terms of μ, R, and T. Then

$$V(\mu, 1, K, R^{-3K}\kappa, A; R^{-K}\kappa s) \geq c_1 \mathrm{tr}(q_{\tau_\kappa}; s).$$

Proof. We construct a proper decomposition with $n = 1$. Let F be uniform on $[0,T]\cap\mathbb{Z}$ and independent of γ. Let $\underline{S}$ be defined as

$$\underline{S} = \inf\{n : \rho(q_n) \le R^{-K-1}\} + F$$

and let

$$S_1 := \inf\{n \ge \underline{S} : \rho(\gamma_{\underline{S}+1}\cdots\gamma_n) \le \kappa\}.$$

Denote

$$\underline{f} = \gamma_1\cdots\gamma_{\underline{S}} \qquad \text{and} \qquad g = \gamma_{\underline{S}+1}\gamma_{\underline{S}+2}\cdots\gamma_{S_1}.$$

By the definition of $\mathrm{tr}(q_{\tau_\kappa}, s)$ there is some σ-algebra $\mathscr{A}$, some random variable V taking values in $\mathfrak{g}$, some $\mathscr{A}$-measurable random variable $\overline{f}$ taking values in G such that $g = \overline{f}\exp(V)$ with $|V| \le s$ and

$$\mathbb{E}[\mathrm{tr}(V|\mathscr{A})] \ge \frac{1}{2}s^2\mathrm{tr}(q_{\tau_\kappa}, s). \tag{13.2.1}$$

We define $T_1 = S_1 + K$ and set

$$h_1 = \gamma_{S_1+1}\gamma_{S_1+2}\cdots\gamma_{T_1}.$$

Denote

$$U_1 = \begin{cases} V & \text{if } |b(h_1)| \le A, \\ 0 & \text{otherwise} \end{cases} \qquad \text{and} \qquad f_1 = \begin{cases} \underline{f}\overline{f} & \text{if } |b(h_1)| \le A, \\ \underline{f}g & \text{otherwise.} \end{cases}$$

Furthermore we set $\mathscr{A}_1 = \sigma(\underline{f}, f_1, h_1, \mathscr{A})$.

We have

$$R^{-K-3}R^{-T}\kappa \le \rho(\underline{f}g) \le R^{-K-1}R^{T}\kappa.$$

In particular, we note that $|U_1| \le s$ and so providing κ and s are sufficiently small in terms of R, we have $R^{-K-4}R^{-T}\kappa \le \rho(f_1) \le R^{T-K}\kappa < 1$ for K sufficiently large in terms of T. This means that $|U_1| \le s \le \rho(f_1)^{-1}R^{T-K}\kappa s$.

Now let $x \in \mathbb{R}^d$ be a unit vector. We wish to show that

$$\mathbb{E}\left[\mathrm{Var}(x\cdot\rho(f_1)U(f_1)U_1b(h_1)|\mathscr{A}_1)\right] \ge c_1\mathrm{tr}(q_{\tau_\kappa}; s)R^{-2K}\kappa^2 s^2.$$

Let $f' = \underline{f}^{-1}f_1$ and let $P_1, \dots, P_d$ be orthogonal eigenvectors of the covariance matrix of $(U_1b(h_1)|\mathscr{A})$ with eigenvalues $\lambda_1 \ge \cdots \ge \lambda_d$. We have

$$\begin{aligned}
&\mathrm{Var}(x\cdot\rho(f_1)U(f_1)U_1b(h_1)|\mathscr{A}_1)\\
&\ge R^{-2K-8}R^{-2T}\kappa^2\,\mathrm{Var}(x\cdot U(\underline{f})U(f')U_1b(h_1)|\mathscr{A}_1)\\
&= R^{-2K-8}R^{-2T}\kappa^2\sum_{i=1}^{d}\left|x\cdot U(\underline{f})U(f')P_i\right|^2\lambda_i\\
&\ge R^{-2K-8}R^{-2T}\kappa^2\left|x\cdot U(\underline{f})U(f')P_1\right|^2\mathrm{tr}(U_1b(h_1)|\mathscr{A}_1)/d.
\end{aligned} \tag{13.2.2}$$

By Proposition 9.1.2 we know that when $b(h_1) \in E_\theta(V)$ and $|b(h_1)| \le A$ we have

$$\mathrm{tr}(U_1 b(h_1)|\mathscr{A}_1) \ge \delta \cdot \mathrm{tr}(U_1|\mathscr{A}_1).$$

By our (α_0, θ, A)-non-degeneracy condition and since $\mu^{*n} * \delta_x$ converges to ν exponentially fast (see for example [KK25d, Lemma 2.2]) we know that providing K is sufficiently large this happens, conditional on $\mathscr{A}$, with probability at least $\frac{1}{2}(1-\alpha)$. Therefore by (13.2.1)

$$\mathbb{E}[\mathrm{tr}(U_1 b(h_1)|\mathscr{A}_1)] \ge \frac{1}{4}(1-\alpha)\delta \mathrm{tr}(q_{\tau_\kappa}; s)s^2.$$

By our (c, T)-well-mixing condition we have that providing K is sufficiently large in terms of μ,

$$\mathbb{E}\left[\left|x \cdot U(\underline{f})U(f')P_1\right|^2 |\sigma(h_1, \mathscr{A})\right] \ge c.$$

Clearly $\mathrm{Var}(U_1 b(h_1)|\mathscr{A}_1)$ is $\sigma(h_1, \mathscr{A})$-measurable. Therefore, by (13.2.2) and conditioning by $\sigma(h_1, \mathscr{A})$,

$$\begin{aligned}
&\mathbb{E}\left[\mathrm{Var}(x \cdot \rho(f_1)U(f_1)U_1 b(h_1)|\mathscr{A}_1)\right] \\
&\ge R^{-2K-8}R^{-2T}\kappa^2 d^{-1}\mathbb{E}\left[\left|x \cdot U(\underline{f})U(f')P_1\right|^2 \mathrm{tr}(U_1 b(h_1)|\mathscr{A}_1)\right] \\
&\ge R^{-2K-8}R^{-2T}\kappa^2 d^{-1}\mathbb{E}\left[\mathbb{E}\left[\left|x \cdot U(\underline{f})U(f')P_1\right|^2 \mathrm{tr}(U_1 b(h_1)|\mathscr{A}_1)\,\middle|\, \sigma(h_1, \mathscr{A})\right]\right] \\
&\ge R^{-2K-8}R^{-2T}\kappa^2 d^{-1}\mathbb{E}\left[\mathbb{E}\left[\left|x \cdot U(\underline{f})U(f')P_1\right|^2 \,\middle|\, \sigma(h_1, \mathscr{A})\right] \mathrm{tr}(U_1 b(h_1)|\mathscr{A}_1)\right] \\
&\ge R^{-2K-8}R^{-2T}\kappa^2 d^{-1} c \cdot \mathbb{E}\left[\mathrm{tr}(U_1 b(h_1)|\mathscr{A}_1)\right] \\
&\ge R^{-2K-8}R^{-2T} d^{-1} c \cdot \frac{1}{4}(1-\alpha)\delta \mathrm{tr}(q_{\tau_\kappa}; s)\kappa^2 s^2 \\
&= c_1 \mathrm{tr}(q_{\tau_\kappa}; s) R^{-2K}\kappa^2 s^2
\end{aligned}$$

where $c_1 = R^{-8}R^{-2T}d^{-1}(1-\alpha)\delta c/4$. Since this is true for any unit vector $x \in \mathbb{R}^d$ we have

$$\mathbb{E}\left[\frac{\mathrm{Var}(\rho(f_1)U(f_1)U_1 b(h_1)|\mathscr{A}_1)}{R^{-2K}\kappa^2 s^2}\right] \ge c_1 \mathrm{tr}(q_{\tau_\kappa}; s) I$$

as required. Finally note that

$$\rho(f_1 h_1) \ge R^{-1}\rho(\underline{f}gh_1) \ge R^{-1}R^{-K-3}R^{-T}\kappa \cdot R^{-K} = \kappa R^{-2K-4-T} \ge R^{-3K}\kappa$$

providing K is sufficiently large in terms of T and R. □

13.3 Concatenating Decompositions

We note that it is straightforward to show that for any measure μ and any admissible choice of coefficients, the variance sum is additive

$$V(\mu, n_1+n_2, K, \kappa_1\kappa_2, A; r) \geq V(\mu, n_1, K, \kappa_1, A; r) + V(\mu, n_2, K, \kappa_2, A; \kappa_1^{-1}r). \tag{13.3.1}$$

However, in order to use Proposition 12.2.2 it is necessary to work with different scales r_1 and r_2 and therefore we show the following proposition.

Proposition 13.3.1. *Let μ be a probability measure on G. Let $R > 1$ be such that $\rho(g) \in [R^{-1}, R]$ for every $g \in \operatorname{supp}(\mu)$. Let $n_1, n_2, K \in \mathbb{Z}_{\geq 0}$ with $n_2, K > 0$ and let $\kappa_1, \kappa_2, r \in (0,1)$. Let $A > 0$ and let $M \geq R$. Then*

$$V(\mu, n_1+n_2, K, R^{-1}M^{-1}\kappa_1\kappa_2, A; r) \geq V(\mu, n_1, K, \kappa_1, A; r) + V(\mu, n_2, K, \kappa_2, A; M\kappa_1^{-1}r).$$

Proof. For $j \in \{1,2\}$ let $\gamma_1^{(j)}, \gamma_2^{(j)}, \dots$ be a sequence of i.i.d. samples from μ defined on the probability space $\left(\Omega_{(j)}, \mathscr{F}_{(j)}, \mathbb{P}_{(j)}\right)$. Let $\hat{\gamma}_1, \hat{\gamma}_2, \dots$ be a sequence of i.i.d. samples from μ defined on the probability space $\left(\hat{\Omega}, \hat{\mathscr{F}}, \hat{\mathbb{P}}\right)$. Consider the product probability space

$$(\Omega, \mathscr{F}, \mathbb{P}) = \left(\Omega_1 \times \hat{\Omega} \times \Omega_2, \mathscr{F}_1 \times \hat{\mathscr{F}} \times \mathscr{F}_2, \mathbb{P}_1 \times \hat{\mathbb{P}} \times \mathbb{P}_2\right).$$

Let $\left(\gamma_i^{(1)}, S_i^{(1)}, T_i^{(1)}, f_i^{(1)}, U_i^{(1)}, h_i^{(1)}, \mathscr{A}_i^{(1)}, m_i^{(1)}\right)$ be a proper decomposition for $(\mu, k_1, K, \kappa_1, A)$ at scale r defined on the probability space $\left(\Omega^{(1)}, \mathscr{F}^{(1)}, \mathbb{P}^{(1)}\right)$ such that $\sum_{i=1}^{k_1} m_i^{(1)}$ approaches $V(\mu, n_1, K, \kappa_1, A; r)$ and

$$\rho(f_1^{(1)}h_1^{(1)} \cdots f_{k_1}^{(1)}h_{k_1}^{(1)}) \geq \kappa_1.$$

Given $\omega_1 \in \Omega_1$ and $\hat{\omega} \in \hat{\Omega}$, let $\tau = \tau(\omega_1, \hat{\omega})$ be given by

$$\tau = \min\{k \in \mathbb{Z}_{\geq 0} : \rho(f_1^{(1)}h_1^{(1)}f_2^{(1)}h_2^{(1)} \dots f_{k_1}^{(1)}h_{k_1}^{(1)}\hat{\gamma}_1\hat{\gamma}_2 \dots \hat{\gamma}_k) < M^{-1}\kappa_1\}$$

and let $\hat{\rho} = \rho(f_1^{(1)}h_1^{(1)}f_2^{(1)}h_2^{(1)} \dots f_{k_1}^{(1)}h_{k_1}^{(1)}\hat{\gamma}_1\hat{\gamma}_2 \dots \hat{\gamma}_\tau)$ such that

$$\hat{\rho} \in [M^{-1}R^{-1}\kappa_1, M^{-1}\kappa_1].$$

Now given $\omega_1 \in \Omega_1$ and $\hat{\omega} \in \hat{\Omega}$, let $\left(\gamma_i^{(2)}, S_i^{(2)}, T_i^{(2)}, f_i^{(2)}, U_i^{(2)}, h_i^{(2)}, \mathscr{A}_i^{(2)}, m_i^{(2)}\right)$ be a proper decomposition for $(\mu, k_2, K, \kappa_2, A)$ at scale $M\kappa_1^{-1}r$ defined on the probability space $\left(\Omega^{(2)}, \mathscr{F}^{(2)}, \mathbb{P}^{(2)}\right)$ such that $\sum_{i=1}^{k_2} m_i^{(2)}$ approaches $V(\mu, n_2, K, \kappa_2, A; M\kappa_1^{-1}r)$ and

$$\rho(f_1^{(1)}h_1^{(1)} \cdots f_{k_2}^{(1)}h_{k_2}^{(1)}) \geq \kappa_2.$$

We now concatenate the two decompositions as follows. Let $\gamma_1, \gamma_2, \dots$ be the sequence of random variables on the probability space $(\Omega, \mathscr{F}, \mathbb{P})$ defined by

$$\gamma_i = \begin{cases} \gamma_i^{(1)} & \text{if } i \le T_{k_1}^{(1)} \\ \hat{\gamma}_{i-T_{k_1}^{(1)}} & \text{if } i > T_{k_1}^{(1)} \text{ and } i \le T_{k_1}^{(1)} + \tau \\ \gamma_{i-T_{k_1}^{(1)}-\tau} & \text{if } i > T_{k_1}^{(1)} + \tau. \end{cases}$$

Clearly these are i.i.d. samples from μ. For $i = 1, 2, \dots, k_1 + k_2$ we define S_i by

$$S_i = \begin{cases} S_i^{(1)} & \text{if } i \le k_1 \\ S_{i-k_1}^{(2)} + T_{k_1}^{(1)} + \tau & \text{if } i > k_1 \end{cases}$$

and we define T_i analogously. We define f_i by

$$f_i = \begin{cases} f_i^{(1)} & \text{if } i \le k_1 \\ \hat{\gamma}_1 \dots \hat{\gamma}_\tau f_1^{(2)} & \text{if } i = k_1 + 1 \\ f_{i-k_1}^{(2)} & \text{if } i > k_1 + 1. \end{cases}$$

We define U_i by

$$U_i = \begin{cases} U_i^{(1)} & \text{if } i \le k_1 \\ U_{i-k_1}^{(2)} & \text{if } i > k_1. \end{cases}$$

and define h_i and m_i analogously. Finally we define $\mathscr{A}_i$ by

$$\mathscr{A}_i = \begin{cases} \mathscr{A}_i^{(1)} \times \hat{\Omega} \times \Omega^{(2)} & \text{if } i \le k_1 \\ \mathscr{A}_{k_1}^{(1)} \times \hat{\mathscr{F}} \times \mathscr{A}_{i-k_1}^{(2)} & \text{if } i > k_1. \end{cases}$$

It is easy to check that $(\gamma_i, S_i, T_i, f_i, U_i, h_i, \mathscr{A}_i, m_i)$ is a proper decomposition for $(\mu, R, k_1 + k_2, K, R^{-1}M^{-1}\kappa_1\kappa_2, A)$ at scale r and it holds that

$$\sum_{i=1}^{k_1+k_2} m_i = \sum_{i=1}^{k_1} m_i^{(1)} + \sum_{i=1}^{k_2} m_i^{(2)}.$$

Indeed, we note that for $i > k_2$ we have that since $M\kappa_1^{-1} \le \hat{\rho}^{-1}$,

$$\begin{aligned} |U_i| = |U_{i-k_1}^{(2)}| &\le \rho(f_1^{(2)} h_1^{(2)} f_2^{(2)} h_2^{(2)} \cdots h_{i-k_1-1}^{(2)} f_{i-k_1}^{(2)})^{-1} M\kappa_1^{-1} r \\ &\le \hat{\rho}^{-1} \rho(f_1^{(2)} h_1^{(2)} f_2^{(2)} h_2^{(2)} \cdots h_{i-k_1-1}^{(2)} f_{i-k_1}^{(2)})^{-1} r \\ &= \rho(f_1 h_1 f_2 h_2 \cdots h_{i-1} f_i)^{-1} r. \end{aligned}$$

Similarly, for $i > k_2 + 1$ and using that $\hat{\rho}^2 M^2 \kappa_1^{-2} \leq 1$,

$$\begin{aligned}
&\mathbb{E}\left[\frac{\mathrm{Var}(\rho(f_i)U(f_i)U_i b(h_i)|\mathscr{A}_i)}{\rho(f_1h_1f_2h_2\cdots f_{i-1}h_{i-1})^{-2}r^2} \mid \mathscr{A}_{i-1}\right] \\
&\quad = \mathbb{E}\left[\frac{\mathrm{Var}(\rho(f_{i-k_1}^{(2)})U(f_{i-k_1}^{(2)})U_{i-k_1}^{(2)} b(h_{i-k_1}^{(2)})|\mathscr{A}_i)}{\hat{\rho}^{-2}\rho(f_1^{(2)}h_1^{(2)}f_2^{(2)}h_2^{(2)}\cdots h_{i-k_1}^{(2)})^{-2}r^2} \mid \mathscr{A}_{i-1}\right] \\
&\quad \geq \mathbb{E}\left[\frac{\mathrm{Var}(\rho(f_{i-k_1}^{(2)})U(f_{i-k_1}^{(2)})U_{i-k_1}^{(2)} b(h_{i-k_1}^{(2)})|\mathscr{A}_i)}{\hat{\rho}^{-2}\rho(f_1^{(2)}h_1^{(2)}f_2^{(2)}h_2^{(2)}\cdots h_{i-k_1}^{(2)})^{-2}\hat{\rho}^2 M^2 \kappa_1^{-2} r^2} \mid \mathscr{A}_{i-1}\right] \\
&\quad \geq m_{i-k_1}^{(2)} I.
\end{aligned}$$

The remainder of the properties are straightforward to check. □

Corollary 13.3.2. *Let μ be a probability measure on G. Let $R > 1$ be such that $\rho(g) \in [R^{-1}, R]$ for every $g \in \mathrm{supp}(\mu)$. Let $n, K \in \mathbb{Z}_{>0}$ and let $\kappa, r \in (0,1)$. Let $C, A > 0$ and let $M \geq R$. Then*

$$V(\mu, n, K, R^{-1}M^{-1}\kappa, A; M^{-1}r) \geq V(\mu, n, K, \kappa, A; r)$$

Proof. By Proposition 13.3.1 we have

$$\begin{aligned}
&V(\mu, n, K, R^{-1}M^{-1}\kappa, A; M^{-1}r) \\
&\qquad \geq V(\mu, 0, K, 1, A; M^{-1}r) + V(\mu, n, K, \kappa, A; r).
\end{aligned}$$

and simply note that $V(\mu, 0, K, 1, A; M^{-1}r) = 0$. □

13.4 From Variance Sum to Bounding Detail

Proposition 13.4.1. *For every $d \geq 1$ and $A, \alpha > 0$ there is a constants $C = C(d, A, \alpha) > 0$ such that the following is true. Suppose that μ is a contracting on average probability measure on G. Then there is some $c = c(\mu) > 0$ such that whenever $\kappa \leq 1$ and $k, K, n \in \mathbb{Z}_{>0}$ with K and n sufficiently large (in terms of A, α and μ) and $r > 0$ is sufficiently small (in terms of A, α and μ) and*

$$V(\mu, n, K, \kappa, A; r) > Ck$$

we have

$$s_r^{(k)}(\nu) < \alpha^k + n\exp(-cK) + C^n \kappa^{-1} r.$$

Proof. Suppose that $(f, h, U, \mathscr{A}, \gamma, \mathscr{F}, S, T, m)$ is a proper decomposition of (μ, n, K, A) at scale r such that $\sum_{i=1}^{n} m_i \geq Ck/2$ and let v be an independent sample from ν. Let

$$I = \{i \in [1, n] \cap \mathbb{Z} : |b(h_i)| \leq A\}$$

and let $m = |I|$. Enumerate I as $i_1 < i_2 < \cdots < i_m$ and define $g_1, \dots, g_m$ by $g_1 = f_1 h_1 \dots f_{i_1}$ and $g_j = h_{i_{j-1}} f_{i_{j-1}+1} \dots f_{i_j}$ for $2 \leq j \leq m$. Define $\overline{v}$ by $\overline{v} = h_{i_m} f_{i_m+1} \dots h_n v$ and let $V_j = U_{i_j}$. Let x be defined by

$$x = g_1 \exp(V_1) \dots g_m \exp(V_m) \overline{v}.$$

Note that x is a sample from ν. Let $\hat{\mathscr{A}}$ be the σ-algebra generated by $\mathscr{A}_n$ and v. Note that the g_j and $\overline{v}$ are $\hat{\mathscr{A}}$-measurable.

We will bound the order k detail of x by showing that with high probability we can apply Proposition 9.1.4 to $g_1, \dots, g_m$, $V_1, \dots, V_m$, and $\overline{v}$ and then bound the order k detail of this using Proposition 10.4.1.

Let E be the event that $|\overline{v}| \leq 2A$ and that for each $j = 1, \dots, m$ we have $|b(g_j)| \leq 2A$, $\rho(g_j) < 1$ and $|V_j| \leq \rho(g_1 \dots g_j)^{-1} r$. By Corollary 9.3.3 we know that $\mathbb{P}[E^C] \leq \exp(-c_1 K)$ for some $c_1 = c_1(\mu, A) > 0$.

For $j = 1, \dots, m$ define ζ_j by

$$\zeta_j = D_u(g_1 \cdots g_j \exp(u) g_{j+1} \cdots g_m \overline{v})|_{u=0}.$$

By Proposition 9.1.4 on E we have

$$\left| x - g_1 \dots g_m \overline{v} - \sum_{j=1}^{m} \zeta_j(V_j) \right| \leq C_1^m \rho(g_1 \dots g_m)^{-1} r^2$$

for some $C_1 = C_1(A) > 0$. Clearly the right hand side is at most $C_1^n \kappa^{-1} r^2$. By Lemma 10.3.1 this means that on E we have

$$s_r^{(k)}(x|\hat{\mathscr{A}}) \leq s_r^{(k)} \left(\sum_{j=1}^{m} \zeta_j(V_j) | \hat{\mathscr{A}} \right) + C_1^n e d \kappa^{-1} r$$

where e is Euler's number.

Let $C_3 = C_3(\alpha, d)$ be the constant C from Proposition 10.4.1 with the same values of α and d and let F be the event that

$$\sum_{j=1}^{m} \operatorname{Var} \zeta_j(V_j | \hat{\mathscr{A}}) \geq k C_3 I.$$

By Proposition 10.4.1, using that by **A8** the $[V_1|\hat{\mathscr{A}}], \ldots, [V_m|\hat{\mathscr{A}}]$ are independent almost surely, we have that on F

$$s_r^{(k)}\left(\sum_{j=1}^{m} \zeta_j(V_j)|\hat{\mathscr{A}}\right) \leq \alpha^k.$$

Therefore

$$s_r^{(k)}(x|\hat{\mathscr{A}}) \leq \alpha^k + C_1^n ed\kappa^{-1}r + \mathbb{I}_{E^C \cup F^C}$$

and so by the convexity of order k detail we have

$$s_r^{(k)}(x) \leq \alpha^k + C_1^n ed\kappa^{-1}r^2 + \mathbb{P}[E^C] + \mathbb{P}[F^C].$$

We already have that $\mathbb{P}[E^C] \leq \exp(-c_1 K)$ so it only remains to bound $\mathbb{P}[F^C]$.

For $i = 1, \ldots, n$ define

$$\hat{\zeta}_i = D_u(f_1 h_1 \cdots h_{i-1} f_i \exp(u) b(h_i))|_{u=0}$$

and let $\underline{F}$ be the event that

$$\left\| \sum_{i=1}^{n} \operatorname{Var} \hat{\zeta}_i(U_i|\hat{\mathscr{A}}) - \sum_{j=1}^{m} \operatorname{Var} \zeta_j(V_j|\hat{\mathscr{A}}) \right\| < 1.$$

Recall that $C_3 = C_3(\alpha, d)$ is the constant C from Proposition 10.4.1 with the same values of α and d and let $\overline{F}$ be the $\hat{\mathscr{A}}$-measurable event that $\sum_{i=1}^{n} \operatorname{Var}(\hat{\zeta}_i(U_i)|\hat{\mathscr{A}}) \geq (C_3+1)kIr^2$. Clearly $\underline{F} \cup \overline{F} \subset F$ so it suffices to bound $\mathbb{P}[\underline{F}^C]$ and $\mathbb{P}[\overline{F}^C]$.

Since $g_1, \ldots, g_m$ and $\overline{v}$ are $\hat{\mathscr{A}}$ measurable, by Lemma 9.1.3 we have for $j = 1, \ldots, m$ that $\operatorname{Var}(\zeta_j(V_j)|\hat{\mathscr{A}})$ is equal to

$$\rho(g_1 \ldots g_j)^2 \cdot U(g_1 \ldots g_j)\psi_{g_{j+1} \cdots g_m \overline{v}} \circ \operatorname{Var}(V_j|\hat{\mathscr{A}}) \circ \psi^T_{g_{j+1} \cdots g_m} U(g_1 \ldots g_j)^T$$

and that

$$\operatorname{Var}(\hat{\zeta}_{i_j}(U_{i_j})|\hat{\mathscr{A}}) = \rho(g_1 \cdots g_j)^2 \cdot U(g_1 \ldots g_j)\psi_{b(h_{i_j})} \circ \operatorname{Var}(V_j|\hat{\mathscr{A}}) \circ \psi^T_{b(h_{i_j})} U(g_1 \ldots g_j)^T.$$

We also have that $|V_j| \leq \rho(g_1 \cdots g_j)^{-1} r$ almost surely and so consequently $\| \operatorname{Var} V_j \| \leq \rho(g_1 \cdots g_j)^{-2} r^2$. Therefore by Lemma 9.1.1 (iii),

$$\| \operatorname{Var} \zeta_j(V_i|\hat{\mathscr{A}}) - \operatorname{Var} \hat{\zeta}_{i_j}(U_{i_j}|\hat{\mathscr{A}}) \| \ll_d |b(h_j) - g_{j+1} \ldots g_m \overline{v}|^2 r^2.$$

Furthermore we have that whenever $i \notin I$ that $\operatorname{Var}(\hat{\zeta}_i(U_i)|\hat{\mathscr{A}}) = 0$. We may assume without loss of generality that $n \exp(-K\chi_\mu/10) < 1$. This means that, providing K

is sufficiently large (in terms of d), in order for $\underline{F}$ to occur it is sufficient that for each $j = 1, \dots, m$ we have

$$|b(h_j) - g_{j+1} \dots g_m \overline{v}| < \exp(-K\chi_\mu/10) < 1/n.$$

By Corollary 9.3.3 this occurs with probability at least $1 - m\exp(-c_2 K)$ for some $c_2 = c_2(\mu) > 0$ and therefore $\mathbb{P}[\underline{F}^C] \leq m\exp(-c_2 K) \leq n\exp(-c_2 K)$.

Finally we wish to bound $\mathbb{P}[\overline{F}^C]$. Let

$$\begin{aligned}\Sigma_i &= r^{-2}\operatorname{Var}(\hat{\zeta}_i(U_i)|\hat{\mathscr{A}}) = r^{-2}\operatorname{Var}(\hat{\zeta}_i(U_i)|\mathscr{A}_i) \\ &= r^{-2}\operatorname{Var}(\rho(f_1 h_1 \cdots h_{i-1} f_i) U(f_1 h_1 \cdots h_{i-1} f_i) U_i b(h_i)|\mathscr{A}_i))\end{aligned}$$

By construction we know that

$$\mathbb{E}[\Sigma_i|\Sigma_1, \dots, \Sigma_{i-1}] \geq m_i I.$$

We also know that $\|\Sigma_i\| \leq A^2$ since $||\psi_{b(h_i)}|| \leq |b(h_i)| \leq A$. This means that we can apply Lemma 9.3.5. By Lemma 9.3.5 we know that providing C is sufficiently large we have

$$\mathbb{P}\left[\sum_{i=1}^{n} \Sigma_i \geq (C_3+1)kI\right] \geq 1 - \exp\left(-c_3 k \sum_{i=1}^{n} m_i\right)$$

for some absolute $c_3 > 0$. Providing we choose C to be sufficiently large, we therefore have $\mathbb{P}[\overline{F}^C] \leq \exp(-c_3 k C) \leq \alpha^k$ this is less than α^k.

Putting everything together we have

$$s_r^{(k)}(x) \leq 2\alpha^k + n\exp(-c_3 K) + edC_1^m \kappa^{-1} r.$$

Replacing α be a slightly smaller value gives the required result. □

13.5 Conclusion of Proof of Theorem 8.1.4

We finally show a decay in detail under the assumption of Theorem 8.1.4. What follows is a rather intricate calculation and we refer the reader to the outline of proofs in section 8.2 for intuition and a sketch of the argument.

Proposition 13.5.1. *Let $d \in \mathbb{Z}_{\geq 1}$ and $c, T, \alpha_0, \theta, A, R > 0$ with $c, \alpha \in (0,1)$ and $T \geq 1$. Then there exists $C = C(d, R, c, T, \alpha_0, \theta, A) > 0$ such that the following is true. Let μ be a contracting on average, (c,T)-well-mixing and (α_0, θ, A)-non-degenerate probability measure on G with $\rho(g) \in [R^{-1}, R]$ for all $g \in \operatorname{supp}(\mu)$ and assume that*

$$\frac{h_\mu}{|\chi_\mu|} > C \max\left\{1, \left(\log \frac{S_\mu}{h_\mu}\right)^2\right\}.$$

Then for all sufficiently small $r > 0$ and all integers $k \in [\log\log r^{-1}, 2\log\log r^{-1}]$ we have that

$$s_r^{(k)}(\nu) < (\log r^{-1})^{-10d}.$$

Proof. We prove this by repeatedly applying Proposition 13.2.1 and Proposition 13.3.1 and then applying Proposition 13.4.1. First let C be as in Proposition 13.4.1 with $\alpha = \exp(-20d)$.

Now let $r > 0$ be sufficiently small and let $K = \exp(\sqrt{\log\log r^{-1}})$. This value of K is chosen so that K grows more slowly than $(\log r^{-1})^{\varepsilon}$ but faster than any polynomial in $\log\log r^{-1}$ as $r \to 0$. Let $S = 2\max\{h_\mu, S_\mu\}$.

Note that $\frac{h_\mu}{2\ell S} < 1$ and for $i = 1, 2, \dots$ let

$$\kappa_i = \exp\left(-\frac{|\chi_\mu|\log r^{-1}}{2S}\left(\frac{h_\mu}{3\ell S}\right)^{i-1}\right) = r^{\frac{|\chi_\mu|}{2S}(\frac{h_\mu}{3\ell S})^{i-1}}$$

with $\ell = \dim G$. Then

$$\kappa_1 = r^{\frac{|\chi_\mu|}{2S}} \quad \text{and} \quad \kappa_{i+1} = \kappa_i^{\frac{h_\mu}{3\ell S}}$$

and let m be chosen as large as possible such that

$$\kappa_m < \min\{R^{-10K}, 2^{-10K}\}.$$

We require $\kappa_m < R^{-10K}$ later in the proof and assume $\kappa_m < 2^{-10K}$ so that κ_m is surely sufficiently small when r is small enough so that we can apply Proposition 12.2.2. Note that this gives

$$\log\log R + \sqrt{\log\log r^{-1}} \ll \log\log r^{-1} + m\log\frac{h_\mu}{2\ell S} + \log\frac{\chi_\mu}{2S}$$

which is equivalent to

$$m\log\left(4\ell\max\left\{1, \frac{S_\mu}{h_\mu}\right\}\right) = m\log\frac{2\ell S}{h_\mu} \ll_d \log\log r^{-1}$$

and therefore we have

$$m \asymp \left(\max\left\{1, \log\frac{S_\mu}{h_\mu}\right\}\right)^{-1}\log\log r^{-1}.$$

Now as in Proposition 12.2.2 let $\hat{m} = \lfloor\frac{S}{100|\chi_\mu|}\rfloor$. For each $i = 1, 2, \dots, m$ let $s_1^{(i)}, s_2^{(i)}, \dots, s_{\hat{m}}^{(i)} > 0$ be the s_i from Proposition 12.2.2 with κ_i in the role of κ. So $s_j^{(i)} \in (\kappa_i^{\frac{S}{|\chi_\mu|}}, \kappa_i^{\frac{h_\mu}{2\ell|\chi_\mu|}})$. By Proposition 13.2.1 we have for each $j \in [\hat{m}]$,

$$V(\mu, 1, K, R^{-3K}\kappa_i, A; R^{-K}\kappa_i s_j^{(i)}) \geq c_1 \mathrm{tr}(q_{\tau_{\kappa_i}}; s_j^{(i)})$$

for some constant $c_1 = c_1(c, T, \alpha_0, \theta, A, R, d) > 0$. Therefore by Proposition 13.3.1 with $M = R^{-1_{\{\geq 2\}}(j)} R^{-3K} \kappa_i s_{j+1}^{(i)} / s_j^{(i)}$, where we denote $1_{\{\geq 2\}}(j) = 1$ whenever $j \geq 2$, we can prove inductively for $j = 2, 3, \ldots, \hat{m}$ that

$$V(\mu, j, K, R^{-1} R^{-3K} \kappa_i s_1^{(i)} / s_j^{(i)}, A; R^{-K} \kappa_i s_1^{(i)}) \geq c_1 \sum_{a=1}^{j} \operatorname{tr}(q_{\tau_{\kappa_i}}; s_j^{(i)}).$$

We have used here that $s_{j+1}^{(i)} / s_j^{(i)} \geq \kappa_i^{-3}$ and so $M \geq R^{-6K} \kappa_i^{-2} \geq R^{10K} \geq R$ since $\kappa_i < R^{-10K}$. By Proposition 12.2.2 and (13.1.2) we conclude that

$$V(\mu, \hat{m}, K, R^{-4K} \kappa_i s_1^{(i)} / s_{\hat{m}}^{(i)}, A; R^{-K} \kappa_i s_1^{(i)}) \geq c_2 \frac{h_\mu}{|\chi_\mu|} \max \left\{ 1, \log \frac{S_\mu}{h_\mu} \right\}^{-1}$$

for some constant $c_2 > 0$ depending on all of the parameters.

Note that for $i = 1, 2, \ldots, m-1$ when $h_\mu / |\chi_\mu|$ is sufficiently large we have

$$\begin{aligned} R^{-4K} \kappa_{i+1} s_1^{(i+1)} / s_{\hat{m}}^{(i)} &\geq R^{-4K} \kappa_{i+1}^{\frac{S}{|\chi_\mu|} + 1} \kappa_i^{-\frac{h_\mu}{2\ell |\chi_\mu|}} \\ &\geq R^{-4K} \kappa_i^{\frac{h_\mu}{3\ell |\chi_\mu|} - \frac{h_\mu}{2\ell |\chi_\mu|} + \frac{h_\mu}{3\ell S}} \\ &\geq R^{-4K} \kappa_i^{-1} \geq R^{6K} \geq R. \end{aligned}$$

as $\kappa_{i+1} = \kappa_i^{\frac{h_\mu}{3\ell S}}$ and $\kappa_i < R^{-10K}$ and so we may repeatedly apply Proposition 13.3.1 with

$$M = R^{-1_{\{\geq 2\}}(i)} R^{-4K} \kappa_{i+1} s_1^{(i+1)} / s_{\hat{m}}^{(i)},$$

where we denote $1_{\{\geq 2\}}(i) = 1$ whenever $i \geq 2$, to inductively show for $i = 2, 3, \ldots, m$ that

$$\begin{aligned} &V(\mu, i\hat{m}, K, R^{-1} R^{-4K} \kappa_1 s_1^{(1)} / s_{\hat{m}}^{(i)}, A; R^{-K} \kappa_1 s_1^{(1)}) \\ &\geq c_2 i \frac{h_\mu}{|\chi_\mu|} \max \left\{ 1, \log \frac{S_\mu}{h_\mu} \right\}^{-1}. \end{aligned}$$

This means using (13.1.2)

$$\begin{aligned} &V(\mu, m\hat{m}, K, R^{-5K} \kappa_1 s_1^{(1)} / s_{\hat{m}}^{(m)}, A; R^{-K} \kappa_1 s_1^{(1)}) \\ \geq &c_3 \frac{h_\mu}{|\chi_\mu|} \max \left\{ 1, \log \frac{S_\mu}{h_\mu} \right\}^{-2} \log \log r^{-1} \end{aligned}$$

for some constant $c_3 > 0$ depending on all of the parameters. Since

$$R^{-K} \kappa_1 s_1^{(1)} \geq R^{-K} \kappa_1^{\frac{S}{|\chi_\mu|} + 1} = R^{-K} r^{\frac{1}{2} + \frac{|\chi_\mu|}{2S}} \geq R^{-K} r^{\frac{1}{2} + \frac{1}{4d}} \geq r$$

for r sufficiently small by Corollary 13.3.2 with $M = R^{-K}\kappa_1 s_1^{(1)} r^{-1} \geq R$

$$V(\mu, m\hat{m}, K, R^{-5K}r/s_{\hat{m}}^{(m)}, A; r) \geq c_3 \frac{h_\mu}{|\chi_\mu|} \max\left\{1, \log \frac{S_\mu}{h_\mu}\right\}^{-2} \log\log r^{-1}.$$

Note that $1/s_{\hat{m}}^{(m)} \geq \kappa_m^{-\frac{h_\mu}{2\ell|\chi_\mu|}}$ and so in particular providing $h_\mu/|\chi_\mu|$ is sufficiently large we have $R^{-5K}r/s_{\hat{m}}^{(m)} \geq R^K r$. By Proposition 13.4.1 provided

$$\frac{h_\mu}{|\chi_\mu|} \max\left\{1, \log \frac{S_\mu}{h_\mu}\right\}^{-2} \geq 2c_3^{-1}C$$

we deduce

$$s_r^{(k)}(\nu) \leq \exp(-20dk) + m\hat{m}\exp(-c_4 K) + R^{-K}C^{m\hat{m}}$$

for some constant $c_4 = c_4(\mu) > 0$ and $k \in [\log\log r^{-1}, 2\log\log r^{-1}]$. Since $m\hat{m} \ll_\mu \log\log r^{-1}$ it is easy to see that

$$m\hat{m}\exp(-c_4 K) + R^{-K}C^{m\hat{m}} < \left(\log r^{-1}\right)^{-20d}$$

whenever $r > 0$ is sufficiently small (in terms of μ). Since $k \geq \log\log r^{-1}$ we have that $\exp(-20dk) \leq (\log r^{-1})^{-20d}$. Overall this means that provided $r > 0$ is sufficiently small (in terms of μ) we have

$$s_r^{(k)}(\nu) < \left(\log r^{-1}\right)^{-10d}. \qquad \square$$

We deduce the main theorem from Proposition 13.5.1.

Proof. (of Theorem 8.1.4) We combine Proposition 13.5.1 with Lemma 10.2.3. Given $r > 0$ sufficiently small, let $k = \frac{3}{2}\log\log r^{-1}$, $a = r/\sqrt{k}$ and $b = rk^k$.

Suppose that $s \in [a, b]$ and note that then $k \in [\log\log s^{-1}, 2\log\log s^{-1}]$ and $\frac{1}{2}\log r^{-1} < \log s^{-1}$ for r sufficiently small and therefore by Proposition 13.5.1

$$s_s^{(k)}(\nu) < (\log s^{-1})^{-10d} < 2^{10d}(\log r^{-1})^{-10d}.$$

By Lemma 10.2.3 it follows that

$$s_r(\nu) \leq Q'(d)^{k-1}(2^{10d}(\log r^{-1})^{-10d} + k^{-k}),$$

which is easily shown to be $\leq (\log r^{-1})^{-2}$ for r sufficiently small. Indeed, recall that $Q'(d) \leq ed^{-1/2} \leq e$ for all $d \geq 1$ and therefore $Q'(d)^k \leq (\log(r^{-1}))^e$.

This concludes the proof of the main theorem of Part II of this thesis. $\square$

13.6 Proof of Theorem 8.1.5

In this section we show how to work with the entropy and separation rate on $O(d)$ instead of the one on G. Recall that for a measure μ on G the measure $U(\mu)$ on $O(d)$ is the pushforward of μ under the map $g \mapsto U(g)$. We then denote for a finitely supported μ by $h_{U(\mu)}$ and $S_{U(\mu)}$ the analogously defined Shannon entropy and separation rate of $U(\mu)$. As we show in section 15.2, when all of the coefficients of the matrices in $\mathrm{supp}(U(\mu))$ lie in the number field K and have logarithmic height at most $L \geq 1$, then

$$S_{U(\mu)} \ll_d L[K : \mathbb{Q}].$$

Therefore Theorem 8.1.5 follows from Theorem 13.6.1.

Theorem 13.6.1. *Let $d \geq 3$ and $R, c, T, \alpha_0, \theta, A > 0$ with $c, \alpha_0 \in (0,1)$ and $T \geq 1$. Then there is a constant $C = C(d, R, c, T, \alpha_0, \theta, A)$ such that the following holds. Let μ be a finitely supported, contracting on average, (c,T)-well-mixing and (α_0, θ, A)-non-degenerate probability measure on G with $\mathrm{supp}(\mu) \subset \{g \in G : \rho(g) \in [R^{-1}, R]\}$. Then ν is absolutely continuous if*

$$\frac{h_{U(\mu)}}{|\chi_\mu|} \geq C \max\left\{1, \log\left(\frac{S_{U(\mu)}}{h_{U(\mu)}}\right)\right\}^2.$$

The proof of Theorem 13.6.1 is analogous to the proof of Theorem 8.1.4. The only point where a slightly different argument is needed is the following version of Proposition 12.1.1. The remainder of the proof is verbatim to the proof of Theorem 8.1.4 with only changing the notation of h_μ to $h_{U(\mu)}$ and S_μ to $S_{U(\mu)}$.

Proposition 13.6.2. *Let μ be a finitely supported, contracting on average probability measure on G. Suppose that $S_{U(\mu)} < \infty$ and that $h_{U(\mu)}/|\chi_\mu|$ is sufficiently large. Let $S > S_{U(\mu)}$, $\kappa > 0$ and $a \geq 1$ and suppose that $0 < r_1 < r_2 < a^{-1}$ with $r_1 < \exp(-S\log(\kappa^{-1})/|\chi_\mu|)$. Then as $\kappa \to 0$,*

$$H_a(q_{\tau_\kappa}; r_1|r_2) \geq \left(\frac{h_{U(\mu)}}{|\chi_\mu|} - d - 1\right)\log\kappa^{-1} + H(s_{r_2,a}) + o_{\mu,d,S,a}(\log\kappa^{-1}).$$

Proof. The proof is similar to the one of Proposition 12.1.1 thus we only provide a sketch. Lemma 12.1.3 still holds and therefore we only need to show that

$$H_a(q_{\tau_\kappa}; r_1) \geq \left(\frac{h_{U(\mu)}}{|\chi_\mu|} - 1\right)\log\kappa^{-1} + o_{\mu,d,S,a}(\log\kappa^{-1}), \tag{13.6.1}$$

where $H_a(q_{\tau_\kappa}; r_1) = H(q_{\tau_\kappa} s_{r_1,a}) - H(s_{r_1,a})$. To show (13.6.1) we apply Lemma 11.2.3 with $X = \mathbb{R}_{>0} \times O(d) \times \mathbb{R}^d$ and $\Phi : G \to X, g \mapsto (\rho(g), U(g), b(g))$ and m_X the product measure on X as used in Lemma 12.1.3. Indeed, as $\Phi_* m_G = m_X$, it follows by Lemma 11.2.3 that $H(q_{\tau_\kappa} s_{r_1,a}) = D_{\mathrm{KL}}(q_{\tau_\kappa} s_{r_1,a} \,||\, m_G) = D_{\mathrm{KL}}(\Phi_* q_{\tau_\kappa} s_{r_1,a} \,||\, m_X)$ and therefore, since m_X is a product measure,

$$\begin{aligned} H(q_{\tau_\kappa} s_{r_1,a}) = {} & D_{\mathrm{KL}}(U(q_{\tau_\kappa} s_{r_1,a}) \,||\, dU) + D_{\mathrm{KL}}(\rho(q_{\tau_\kappa} s_{r_1,a}) \,||\, \rho^{-(d+1)} d\rho) \\ & + D_{\mathrm{KL}}(b(q_{\tau_\kappa} s_{r_1,a}) \,||\, db). \end{aligned}$$

As in Proposition 12.1.1 one shows that

$$D_{\mathrm{KL}}(U(q_{\tau_\kappa} s_{r_1,a}) \,||\, dU) \geq \frac{h_{U(\mu)}}{|\chi_\mu|} \log \kappa^{-1} + D_{\mathrm{KL}}(U(s_{r_1,a}) \,||\, dU) + o_{\mu,d,S,a}(\log \kappa^{-1}).$$

On the other hand,

$$D_{\mathrm{KL}}(\rho(q_{\tau_\kappa} s_{r_1,a}) \,||\, \rho^{-(d+1)} d\rho) \gg D_{\mathrm{KL}}(\rho(s_{r_1,a}) \,||\, \rho^{-(d+1)} d\rho)$$

and

$$D_{\mathrm{KL}}(b(q_{\tau_\kappa} s_{r_1,a}) \,||\, db) \gg D_{\mathrm{KL}}(b(s_{r_1,a}) \,||\, db)$$

and note that by [KK25b, Lemma 2.5],

$$D_{\mathrm{KL}}(U(s_{r_1,a}) \,||\, dU) + D_{\mathrm{KL}}(\rho(s_{r_1,a}) \,||\, \rho^{-(d+1)} d\rho) + D_{\mathrm{KL}}(b(s_{r_1,a}) \,||\, db) \geq H(s_{r_1,a}).$$

All these estimates combined imply the claim. □

Chapter 14

Well-Mixing and Non-Degeneracy

In this section we study (c, T)-well mixing as well as (α_0, θ, A)-non-degeneracy. The goal of this section is prove Proposition 8.1.2 and Proposition 8.1.3. We treat (c, T)-well-mixing in section 14.1 and show that we have uniform results as long as $U(\mu)$ is fixed. In section 14.2 we conclude the proofs of Proposition 8.1.2 and Proposition 8.1.3 by proving strong results on non-degeneracy.

14.1 (c, T)-well-mixing

In this subsection we establish in Lemma 14.1.2 that we have uniform (c, T)-well-mixing whenever $U(\mu)$ is fixed and show that (c, T) can taken to be uniform when we know a lower bound on the spectral gap of $U(\mu)$. We start with a preliminary lemma that will also be used in section 14.2. Throughout this section and next we denote by m_H the Haar probability measure on H and by $I \in O(d)$ the identity matrix.

Lemma 14.1.1. *(Schur-type Lemma) Suppose that $d \geq 1$ and that H is an irreducible subgroup of $O(d)$ and let V be a uniform random variable on H. Let B be a random variable independent from V taking values in $\mathbb{R}^d$. Then VB has mean zero and covariance matrix of the form λI for some $\lambda \geq 0$.*

Proof. For $h \in H$ the random variables hVB and VB have the same law. This means that the mean of VB is invariant under H and so since H is irreducible it must be zero. Moreover the covariance matrix M of VB is invariant under conjugation by elements of H. Since M is symmetric positive definite, it has an eigenvector v and therefore $Mv = \lambda v$ and $Mhv = hMv = \lambda hv$ for some $\lambda \geq 0$ and all $h \in H$. Since H is irreducible it therefore follows that $M = \lambda I$ as claimed. □

Lemma 14.1.2. *Let μ_U be a finitely supported probability measure on $O(d)$ such that* $\operatorname{supp}(\mu_U)$ *acts irreducibly on $\mathbb{R}^d$. Then there exists $T = T(\mu_U)$ only depending* *on μ_U such that every finitely supported probability measure μ on G with $U(\mu)$ is $(\frac{1}{2d}, T)$-well-mixing.*

Proof. Let $H \subset O(d)$ be the closure of the group generated by $\mathrm{supp}(\mu_U)$. Then H is compact and let m_H be the Haar probablility measure on G and denote by V a uniform random variable on H. We first claim that for all unit vectors x and y in $\mathbb{R}^d$ we have

$$\mathbb{E}[|x \cdot Vy|^2] = \frac{1}{d}. \tag{14.1.1}$$

Indeed, we can view y as a random variable independent from V and therefore, by Lemma 14.1.1, the random variable Vy has mean zero and covariance matrix λI. Moreover, since $\mathbb{E}[|Vy|^2] = d\lambda = 1$ it follows that $\lambda = \frac{1}{d}$ and therefore (14.1.1) holds.

Let F be a uniform random variable on $[0, T]$. Then F is distributed as

$$\frac{1}{T+1}\sum_{i=0}^{T} \mu^{*i}. \tag{14.1.2}$$

We claim that (14.1.2) converges as $T \to \infty$ to m_H in the weak*-topology. Indeed, we note that any weak*-limit m of (14.1.2) is μ_U-stationary and, upon performing an ergodic decomposition, we may assume without loss of generality that m is in addition ergodic. As this is equivalent to the measure being extremal, we conclude that m is invariant under the group generated by $\mathrm{supp}(\mu_U)$ and therefore also by H, implying that $m = m_H$.

Finally, we just choose $c = \frac{1}{2d}$ and T sufficiently large depending on μ_U such that (14.1.2) is sufficiently close in distribution to m_H and therefore $\mathbb{E}[|x \cdot U(q_F)y|^2] \geq \frac{1}{2d}$ for all unit vectors $x, y \in \mathbb{R}^d$, implying the claim. □

For a closed subgroup $H \subset O(d)$ and a probability measure μ_U supported on H we denote, as defined in (8.3.4), by $\mathrm{gap}_H(\mu_U)$ the L^2-spectral gap of μ_U on $L^2(H)$. We aim to show uniform well-mixing as long as $\mathrm{gap}_H(\mu_U) \geq \varepsilon$ independent of the subgroup H. To do so, we first show that we have uniform convergence in the Wasserstein distance with a rate only depending on ε and d.

Lemma 14.1.3. *Let $d \geq 1, \varepsilon \in (0,1)$ and let μ_U be a probability measure on $O(d)$. Assume that $\mathrm{gap}_H(\mu_U) \geq \varepsilon$ for H the subgroup generated by the support of μ_U. Then for $n \geq 1$*

$$\mathcal{W}_1(\mu_U^{*n}, m_H) \ll_d (1-\varepsilon)^{\alpha n}$$

for $\alpha = (1 + \frac{1}{2}\dim O(d))^{-1}$.

Proof. We consider the metric $d(g_1, g_2) = ||g_1 - g_2||$ on $O(d)$ for $|| \circ ||$ the operator norm and note that it is bi-invariant and restricts to H. Denote by $B_\delta^H(h)$ for $h \in H$ and $\delta > 0$ the δ-ball around $h \in H$ and denote

$$P_\delta = \frac{1_{B_\delta^H(e)}}{m_H(B_\delta^H(e))}.$$

For $\delta \in (0, 1)$ we note that $m_H(B_\delta^H(e)) \gg_d \delta^{\dim O(d)}$ for an implied constant depending only on d and therefore $||P_\delta||_2 \ll_d \delta^{-(\dim O(d))/2}$. Also we note that for $h \in H$ we have $(\mu^{*n} * P_\delta)(h) = \frac{\mu^{*n}(B_\delta^H(h))}{m_H(B_\delta^H(e))}$. By the triangle inequality,

$$\mathcal{W}_1(\mu^{*n}, m_H) \leq \mathcal{W}_1(\mu^{*n}, \mu^{*n} * P_\delta) + \mathcal{W}_1(\mu^{*n} * P_\delta, m_H).$$

Note $\mathcal{W}_1(\mu^{*n}, \mu^{*n} * P_\delta) \ll_d \delta$ and since H is compact,

$$\begin{aligned}
\mathcal{W}_1(\mu^{*n} * P_\delta, m_H) &\ll_d ||\mu^{*n} * P_\delta - 1||_1 \\
&\leq ||\mu^{*n} * P_\delta - 1||_2 \\
&\leq (1 - \varepsilon)^n ||P_\delta||_2 \ll_d (1 - \varepsilon)^n \delta^{-(\dim O(d))/2}.
\end{aligned}$$

To conclude, if follows

$$\mathcal{W}_1(\mu^{*n}, m_H) \ll_d \delta + (1 - \varepsilon)^n \delta^{-(\dim O(d))/2}.$$

Therefore setting $\delta = (1 - \varepsilon)^{\alpha n}$ for $\alpha = (1 + \frac{1}{2} \dim O(d))^{-1}$ implies the claim. □

Lemma 14.1.4. *Let $d \geq 1, \varepsilon \in (0, 1)$ and let μ_U be a probability measure on $O(d)$. Assume that $\mathrm{gap}_H(\mu_U) \geq \varepsilon$ for H the subgroup generated by the support of μ_U. Then there exists $T = T(d, \varepsilon)$ only depending on d and ε such every probability measure μ on G with $U(\mu) = \mu_U$ is $(\frac{1}{2d}, T)$-well-mixing.*

Proof. The proof is similar to the one of Lemma 14.1.2 and recall the notation used in it. Consider a list of tuples of unit vectors $(x_1, y_1), \ldots, (x_m, y_m)$ such that for every two unit vectors x and y in $\mathbb{R}^d$ there is some $i \in [m]$ such that

$$\sup_{U \in O(d)} \left| |x \cdot Uy|^2 - |x_i \cdot Uy_i|^2 \right| < \frac{1}{4d}.$$

Such a list of tuples exists as the action of $O(d)$ on $\mathbb{S}^{d-1} \subset \mathbb{R}^d$ is uniformly continuous. We claim that for T large enough depending only on ε we have for all $i \in [m]$ that

$$\mathbb{E}[|x_i \cdot U(q_F) y_i|^2] \geq \frac{3}{4d}.$$

Indeed, we note that for $h_1, h_2 \in H$ we have

$$\big|\, |x_i \cdot h_1 y_i|^2 - |x_i \cdot h_2 y_i|^2 \,\big| \leq \big|\, |x_i \cdot h_1 y_i| + |x_i \cdot h_2 y_i| \,\big| \cdot \big|\, |x_i \cdot h_1 y_i| - |x_i \cdot h_2 y_i| \,\big| \leq 2||h_1 - h_2||.$$

Thus it follows that

$$\mathbb{E}[|x_i \cdot V y_i|^2 - |x_i \cdot U(q_n) y_i|^2] \leq 2\mathcal{W}_1(\mu^{*n}, m_H)$$

and the claim follows by Lemma 14.1.3. This concludes the proof as for all x and y we have

$$\mathbb{E}[|x \cdot U(q_F) y|^2] \geq \sup_{i \in [m]} \mathbb{E}[|x_i \cdot U(q_F) y_i|^2] - \frac{1}{4d} \geq \frac{1}{2d}.$$

□

Another direction to show uniform well-mixing would be to study the stopped random walk $U(q_{\tau_\kappa})$ and to show that $U(q_{\tau_\kappa}) \to m_H$. We do not pursue this direction further and just note that the results by Kesten [Kes74] can be applied to this problem.

14.2 (α_0, θ, A)-non-degeneracy

In order to state our results on (α_0, θ, A)-non-degeneracy it is useful to understand that we can translate and rescale our generating measures, without changing any of the fundamental properties. It is also beneficial to replace μ by $\frac{1}{2}\delta_e + \frac{1}{2}\mu$ and we show in the following lemma that these changes do not change our self-similar measure or any of the relevant constants in a fundamental way.

Lemma 14.2.1. *Let $\mu = \sum_i p_i \delta_{g_i}$ be a contracting on average probability measure on G with self-measure ν. Let $h \in G$ and consider the measures*

$$\mu_h = \sum_i p_i \delta_{h g_i h^{-1}} \quad \text{and} \quad \mu_h' = \frac{1}{2}\delta_e + \frac{1}{2}\mu_h.$$

Then the following properties hold:

(i) $h_\mu = h_{\mu_h} = 2h_{\mu_h'}$,

(ii) $\chi_\mu = \chi_{\mu_h} = 2\chi_{\mu_h'}$,

(iii) $S_\mu = S_{\mu_h} = S_{\mu_h'}$,

(iv) $\mathrm{gap}_H(\mu) = \mathrm{gap}_{U(h)HU(h)^{-1}}(\mu_h) = 2\mathrm{gap}_{U(h)HU(h)^{-1}}(\mu_h')$,

(v) *μ_h and $\mu_{h'}$ have $h\nu$ as self-similar measure.*

Proof. As conjugation is a bijection on G and by using [HS17, Lemma 6.8], (i) follows. Moreover, (ii) follows since $\rho(hg_ih^{-1}) = \rho(g_i)$ and (iv) follows similarly. To show (iii) note $S_{\mu_h} = S_{\mu'_h}$ since by the triangle inequality $d(g,h) \leq d(g,e) + d(e,h)$ for all $g, h \in G$. To show that $S_\mu = S_{\mu_h}$, set

$$A = \min_{g_1, g_2 \in \mathrm{supp}(\mu), g_1 \neq g_2} d(g_1, g_2)$$

and note that there is a constant C_h depending on h such that $d(hg_1h^{-1}, hg_2h^{-1}) \leq C_h d(g_1, g_2)$ for $d(g_1, g_2) \leq A$. Thus it holds that

$$\begin{aligned} S_{\mu_h} &= \limsup_{g_1, g_2 \in S_n, g_1 \neq g_2} -\frac{1}{n} \log d(hg_1h^{-1}, hg_2h^{-1}) \\ &\leq \limsup_{g_1, g_2 \in S_n, g_1 \neq g_2} -\frac{1}{n} \log C_h d(g_1, g_2) = S_\mu \end{aligned}$$

Applying the same argument to conjugation by h^{-1} implies the claim. Finally, we note that μ_h and μ'_h have the same self-similar measure and it holds that

$$h\nu = h \sum_i p_i g_i \nu = \sum_i p_i hg_ih^{-1} h\nu$$

and therefore $h\nu$ is the self-similar measure of μ_h and μ'_h. □

In particular, it follows that the self-similar measure of μ is absolutely continuous if and only if the one of μ_h or $\mu_{h'}$ is and all of the relevant quantities are the same up to a factor of 2.

To give an idea of the proof of the main results in this subsection, we first discuss how to show that real Bernoulli convolutions ν_λ are uniformly non-degenerate. Indeed, we distinguish between $\lambda \geq \lambda_0$ and $\lambda \leq \lambda_0$ for some λ_0 sufficiently close to 1. Note that ν_λ is supported on $[-(1-\lambda)^{-1}, (1-\lambda)^{-1}]$ and thus when $\lambda \leq \lambda_0$ one easily shows uniform non-degeneracy depending only on λ_0 by compactness of the support. In the case $\lambda \geq \lambda_0$ it follows from the Berry-Essen Theorem 10.4.2 that $\mathcal{W}_1(\nu_\lambda, \mathcal{N}(0, \frac{1}{\sqrt{1-\lambda^2}})) \approx 2/3$. The latter then implies then the claim by Lemma 14.2.5 and by rescaling ν_λ to have variance 1.

Our results will be deduced from suitable results in the case when μ has a uniform contraction ratio and then in the general case from comparing our given measure with a self-similar measure with uniform contraction ratio. We now state the main proposition of this section.

Proposition 14.2.2. *Let $d \geq 1$, $\varepsilon > 0$ and let μ_U be an irreducible probability measure on $O(d)$. Then there is $\tilde{\rho} \in (0,1)$ and some (α_0, θ, A) depending on d, ε and μ_U such that the following is true. Let $\mu = \sum_{i=1}^{k} p_i \delta_{g_i}$ be a contracting on average probability measure on G satisfying $U(\mu) = \mu_U$ and $p_i \geq \varepsilon$ for $1 \leq i \leq k$. Suppose further that there is some $\hat{\rho} \in (\tilde{\rho}, 1)$ such that*

$$\frac{\mathbb{E}_{\gamma \sim \mu} |\hat{\rho} - \rho(\gamma)|}{1 - \mathbb{E}_{\gamma \sim \mu}[\rho(\gamma)]} < 1 - \varepsilon.$$

Then there is some $h \in G$ with $U(h) = I$ such that the conjugate measure $\mu'_h = \frac{1}{2}\delta_e + \frac{1}{2}\sum_i p_i \delta_{h g_i h^{-1}}$ is (α_0, θ, A)-non-degenerate.

Moreover, if in addition $\operatorname{gap}_H(\mu) \geq \varepsilon$, for H the closure of the subgroup generated by $\operatorname{supp}(\mu)$, then $\tilde{\rho}$ and (α_0, θ, A) can be made uniform in d and ε.

We first show how to deduce from Proposition 14.2.2 the two propositions 8.1.2 and 8.1.3 from section 8.1. To do so we first state the following lemma.

Lemma 14.2.3. *Suppose $x_1 < x_2$ and let X be a real-valued random variable such that $X \leq x_2$ almost surely and $\mathbb{P}[X \leq x_1] \geq 1/2 + p$ for some $p > 0$. Then*

$$\mathbb{E}[|X - x_1|] \leq \mathbb{E}[|X - x_2|] - 2p(x_2 - x_1).$$

Proof. Let X_1 and X_2 have the same law as X and be coupled such that at least one of them is at most x_1 almost surely. Let A be the event that both X_1 and X_2 are at most x_1. Noting that A has probability at least $2p$ we compute

$$\begin{aligned}
\mathbb{E}[|X_1 - x_1| + |X_2 - x_1|] &= \mathbb{E}[(|X_1 - x_1| + |X_2 - x_1|)\mathbb{I}_{A^C}] \\
&\quad + \mathbb{E}[(|X_1 - x_1| + |X_2 - x_1|)\mathbb{I}_{A}] \\
&\leq \mathbb{E}[(|X_1 - x_2| + |X_2 - x_2|)\mathbb{I}_{A^C}] \\
&\quad + \mathbb{E}[(|X_1 - x_2| + |X_2 - x_2| - 2(x_2 - x_1))\mathbb{I}_{A}] \\
&\leq \mathbb{E}[|X_1 - x_2| + |X_2 - x_2|] - 4p(x_2 - x_1).
\end{aligned}$$

The result follows. □

We now prove Proposition 8.1.2 and Proposition 8.1.3.

Proof of Proposition 8.1.2. Let $\gamma_1, \gamma_2, \ldots$ be i.i.d. samples from μ. Let $p_{\min}$ be the smallest of the $p_1, \ldots, p_k$ and let $\rho_{\min}$ be the smallest of the $\rho(g_1), \ldots, \rho(g_k)$. Clearly

$$\mathbb{P}[\rho(\gamma_1 \ldots \gamma_n) \leq \rho_{\min}] \geq 1 - (1 - p_{\min})^n.$$

In particular there is some n depending only on ε such that this is at least 3/4. Note that by Lemma 14.2.3 with $x_1 = \rho_{\min}$ and $x_2 = 1$ and $p = \frac{1}{4}$ we have

$$\begin{aligned}\frac{\mathbb{E}[|\rho(\gamma_1 \cdots \gamma_n) - \rho_{\min}|]}{1 - \mathbb{E}[\rho(\gamma_1 \dots \gamma_n)]} &\leq \frac{1 - \mathbb{E}[\rho(\gamma_1 \dots \gamma_n)] - (1 - \rho_{\min})/2}{1 - \mathbb{E}[\rho(\gamma_1 \dots \gamma_n)]} \\ &= 1 - \frac{1 - \rho_{\min}}{2(1 - \mathbb{E}[\rho(\gamma_1 \dots \gamma_n)])} \\ &\leq 1 - \frac{1 - \rho_{\min}}{2(1 - \rho_{\min}^n)} \\ &\leq 1 - \frac{1}{2n}.\end{aligned}$$

The result now follows by applying Proposition 14.2.2, Lemma 14.1.2 and Lemma 14.1.4 to μ^{*n}. □

Proof of Proposition 8.1.3. This follows directly by Proposition 14.2.2, Lemma 14.1.2 and Lemma 14.1.4. □

Now we prove Proposition 14.2.2. We use the following definition.

Definition 14.2.4. *Given two measures λ_1, λ_2 on $\mathbb{R}^d$ we define*

$$\mathcal{PW}_1(\lambda_1, \lambda_2) := \inf_{\gamma \in \Gamma(\lambda_1, \lambda_2)} \sup_{p \in P(d)} \int |px - py| \, d\gamma(x, y)$$

where $P(d)$ is the set of orthogonal projections onto one dimensional subspaces of $\mathbb{R}^d$ and $\Gamma(\lambda_1, \lambda_2)$ is the set of couplings between λ_1 and λ_2.

We use this to show that if a measure is sufficiently close to a spherical normal distribution then it is (α_0, θ, A)-non-degenerate.

Lemma 14.2.5. *Let I be the $d \times d$ identity matrix. Then given any $p \in P(d)$ we have*

$$\mathbb{E}_{x \sim N(0,I)}[|px|] = \sqrt{\frac{2}{\pi}}.$$

Moreover, for any $\varepsilon > 0$ there exists $\alpha_0 \in (0, 1)$ and $\theta, A > 0$ such that if ν is a measure on $\mathbb{R}^d$ and

$$\mathcal{PW}_1(\nu, N(0, I)) < \sqrt{\frac{2}{\pi}} - \varepsilon$$

then ν is (α_0, θ, A)-non-degenerate.

Proof. The first part follows since if $X \sim \mathcal{N}(0, I)$ and $u \in \mathbb{R}^d$ is a unit vector, then $\langle X, u \rangle$ is distributed as $\mathcal{N}(0, 1)$. The second part follows from the first part, the fact that the $y \in \mathbb{R}$ such that $\mathbb{E}_{x \sim N(0,1)}|x - y|$ is minimal is $y = 0$ and Markov's inequality.

More precisely, we aim to estimate for all $y_0 \in \mathbb{R}^d$ and all proper subspaces $W \subset \mathbb{R}^d$

$$\nu(\{x \in \mathbb{R}^d : |x - (y_0 + W)| < \theta \text{ or } |x| \geq A\}),$$

which is bounded by $\nu(\{x \in \mathbb{R}^d : |x - (y_0 + W)| < \theta\}) + \nu(\{x \in \mathbb{R}^d : |x| \geq A\})$. To deal with the second term we note that by Markov's inequality for a coupling γ between ν and $\mathcal{N}(0, 1)$ we have

$$\begin{aligned}\nu(\{x \in \mathbb{R}^d : |x| \geq A\}) &\leq A^{-1} \int |x| \, d\nu(x) \\ &\leq A^{-1} \left(\int |y| \, d\mathcal{N}(0, I)(y) + \int |x - y| \, d\gamma(x, y) \right).\end{aligned}$$

In order to apply our bound for $\mathcal{PW}_1(\nu, N(0, I))$ we consider the projections $p_1, \ldots, p_d$ to the coordinate axes. Then $|x - y| \leq \sum_{i=1}^d |p_i x - p_i y|$ and therefore by choosing a suitable coupling, it follows that for A sufficiently large only depending on d and ε we have that $\nu(\{x \in \mathbb{R}^d : |x| \geq A\}) \leq \varepsilon/16$.

To deal with the first term $\nu(\{x \in \mathbb{R}^d : |x - (y_0 + W)| < \theta\})$, we assume without loss of generality that W has dimension $d - 1$ and we let p be the orthogonal projection to the orthogonal complement of W. Then it holds that $|x - (y_0 + W)| = |px - py_0|$ and therefore

$$\nu(\{x \in \mathbb{R}^d : |x - (y_0 + W)| < \theta\}) = \nu(\{x \in \mathbb{R}^d : |px - py_0| < \theta\}).$$

In the following we identify $p\mathbb{R}^d$ as the real line. Let γ be any coupling between ν and $\mathcal{N}(0, I)$. Then it holds that

$$\begin{aligned}\int |px - py| \, d\gamma(x, y) &\geq \int |px - py| 1_{|px - py_0| < \theta}(x, y) \, d\gamma(x, y) \\ &\geq \nu(\{x \in \mathbb{R}^d : (|px - py_0| < \theta\}) \int |py - py_0| - \theta) \, d\mathcal{N}(0, I)(y) \\ &\geq \nu(\{x \in \mathbb{R}^d : (|px - py_0| < \theta\}) \left(\sqrt{\frac{2}{\pi}} - \theta \right),\end{aligned}$$

having used in the last line that $y \in \mathbb{R}$ such that $\mathbb{E}_{x \sim N(0,1)}|x - y|$ is minimal is $y = 0$. By choosing a suitable coupling and setting $\theta = \varepsilon/4$ it therefore follows for ε sufficiently small that

$$\nu(\{x \in \mathbb{R}^d : (|px - py_0| < \theta\}) \leq \frac{\sqrt{\frac{2}{\pi}} - \varepsilon/2}{\sqrt{\frac{2}{\pi}} - \varepsilon/4} \leq 1 - \varepsilon/8.$$

The claim follows by combining the above two estimates. □

To make this useful we need to show that our self-similar measures are close to spherical normal distributions. We prove this in the case where all of the ρ_i are equal with the following proposition.

Proposition 14.2.6. *Given any $\varepsilon > 0$ and any irreducible probability measure $\mu_U = \sum_{i=1}^k p_i \delta_{U_i}$ on $O(d)$ there is some $\tilde{\rho} \in (0,1)$ depending on ε and μ_U such that the following is true. Let $\mu = \sum_{i=1}^k p_i \delta_{g_i}$ be a probability measure on G without a common fixed point and with $U(\mu) = \mu_U$ as well as $p_i \geq \varepsilon$ for all $1 \leq i \leq k$. Assume there is $\rho \in (\tilde{\rho}, 1)$ such that $\rho(g_i) = \rho$ for all $1 \leq i \leq k$. Then there exists some $h \in G$ with $U(h) = I$ such that the self similar measure ν'_h generated by the conjugate measure $\mu'_h = \frac{1}{2}\delta_e + \frac{1}{2}\sum_i p_i \delta_{hg_ih^{-1}}$ satisfies*

$$\mathcal{W}_3(\nu'_h, N(0,I)) < \varepsilon.$$

If moreover $\operatorname{gap}_H(\mu_U) \geq \varepsilon$ then $\tilde{\rho}$ is uniform in d and ε.

We then extend to the general case using the following lemma.

Lemma 14.2.7. *Let γ and $\tilde{\gamma}$ be contracting on average random variables taking values in G such that $U(\gamma) = U(\tilde{\gamma})$ and $z(\gamma) = z(\tilde{\gamma})$ almost surely. Let ν and $\tilde{\nu}$ be the self similar measures generated by the laws of γ and $\tilde{\gamma}$ respectively. Then*

$$\mathcal{PW}_1(\nu, \tilde{\nu}) \leq \frac{\mathbb{E}[|\rho(\gamma) - \rho(\tilde{\gamma})|]}{1 - \mathbb{E}[\rho(\gamma)]} \sup_{p \in P(d)} \mathbb{E}_{x \sim \tilde{\nu}} |px|.$$

We now have all the ingredients needed to prove Proposition 14.2.2.

Proof of Proposition 14.2.2. Without loss of generality we replace μ by $\frac{1}{2}\delta_e + \frac{1}{2}\mu$. Let $\tilde{g}_i : x \mapsto \hat{\rho} U_i x + b_i$ and let $\tilde{\mu} = \sum_{i=1}^n p_i \delta_{\tilde{g}_i}$ with self-similar measure $\tilde{\nu}$. Then by Proposition 14.2.6 there is some $h \in G$ with $U(h) = I$ such that

$$\mathcal{W}_3(\tilde{\nu}_h, N(0,I)) < \varepsilon/10.$$

Clearly this implies $\mathcal{W}_1(\tilde{\nu}_h, N(0,I)) < \varepsilon/10$ and therefore $\mathcal{PW}_1(\tilde{\nu}_h, N(0,I)) < \varepsilon/10$ and so by Lemma 14.2.7 if we define $\mu_h = \sum_{i=1}^k p_i \delta_{hg_ih^{-1}}$ and let ν_h be the self similar measure generated by μ_h we have $\mathcal{PW}_1(\nu_h, N(0,I)) < \sqrt{\frac{\pi}{2}} - \varepsilon/2$. The result follows by Lemma 14.2.5. □

Now we just need to prove Lemma 14.2.7 and Proposition 14.2.6. We start with Lemma 14.2.7.

Proof of Lemma 14.2.7. Let x be a sample from ν and $\tilde{x}$ be a sample from $\tilde{\nu}$ such that $(x, \tilde{x})$ is independent from $(\gamma, \tilde{\gamma})$. Note that this means that γx is a sample from ν and $\tilde{\gamma}\tilde{x}$ is a sample from $\tilde{\nu}$. Let $p \in P(d)$. We have

$$\begin{aligned}\mathbb{E}[|p\gamma x - p\tilde{\gamma}\tilde{x}|] &\leq \mathbb{E}[|p\gamma(x - \tilde{x})|] + \mathbb{E}[|p(\gamma - \tilde{\gamma})\tilde{x}|] \\ &= \mathbb{E}[\rho(\gamma)]\mathbb{E}[|pU(\gamma)(x - \tilde{x})|] + \mathbb{E}[|\rho(\gamma) - \rho(\tilde{\gamma})|]\mathbb{E}[|pU(\gamma)(\tilde{x})|].\end{aligned}$$

Therefore by taking a series of couplings such that $\sup_{p \in P(d)} \mathbb{E}[|px - p\tilde{x}|] \to \mathcal{PW}_1(\nu, \tilde{\nu})$ we get

$$\mathcal{PW}_1(\nu, \tilde{\nu}) \leq \mathbb{E}[\rho(\gamma)]\mathcal{PW}_1(\nu, \tilde{\nu}) + \mathbb{E}[|\rho(\gamma) - \rho(\tilde{\gamma})|]\mathbb{E}_{x \sim \tilde{\nu}}[|p(x)|].$$

□

Now we wish to prove Proposition 14.2.6. First we need the following result.

Lemma 14.2.8. *Let μ_U be a probability measure on $O(d)$ and let H be the closure of the group generated by the support of μ_U and let V be a uniform random variable on H. Let $\gamma_1, \gamma_2, \dots$ be independent samples from $\frac{1}{2}\delta_e + \frac{1}{2}\mu_U$. Then for every $\varepsilon > 0$ there exists $N \in \mathbb{Z}_{>0}$ depending on d, H and ε such that whenever $n \geq N$ we have*

$$\mathcal{W}_3(\gamma_1 \dots \gamma_n, V) < \varepsilon.$$

Furthermore, if in addition $\operatorname{gap}_H(\mu_U) \geq \varepsilon$, then N can be made uniform d and ε.

Proof. This follows similarly to the arguments given in section 14.1 since the measure $\mu'_U = \frac{1}{2}\delta_e + \frac{1}{2}\mu_U$ satisfies that $(\mu'_U)^{*n} \to m_H$ as $n \to \infty$. In the presence of a spectral gap we apply Lemma 14.1.3 and use that by compactness of H the L^3-Wasserstein distance is comparable with the L^1-Wasserstein distance. □

It is convenient to work with measures which are appropriately translated.

Definition 14.2.9. *We say that a probability measure μ on G is* centred at zero *if $\mathbb{E}_{\gamma \sim \mu}[\gamma(0)] = 0$.*

Lemma 14.2.10. *Suppose that μ is a probability measure on G which is centred at zero and has uniform contraction ratio $\rho \in (0, 1)$. Then if $\gamma_1, \gamma_2, \dots$ are i.i.d. samples from μ and $n \in \mathbb{Z}_{>0}$ we have*

$$\mathbb{E}[\gamma_1 \dots \gamma_n(0)] = 0$$

and

$$\mathbb{E}[|\gamma_1 \dots \gamma_n(0)|^2] = \frac{1 - \rho^{2n}}{1 - \rho^2}\mathbb{E}[|\gamma_1(0)|^2].$$

Proof. Both of these follow by an induction argument left to the reader. □

In order to prove Proposition 14.2.6, we need the following theorem of Sakhanenko from [Sak85].

Theorem 14.2.11. *For every $p, d \geq 1$ there is some constant $C = C(p, d) > 0$ such that the following holds. Suppose that $X_1, \dots, X_n$ are independent random variables taking values in $\mathbb{R}^d$ with mean 0. Let $\Sigma_i = \operatorname{Var} X_i$, suppose that $\sum_{i=1}^n \Sigma_i = I$ and let $L_p = (\sum_{i=1}^n \mathbb{E}[|X_i|^p])^{1/p}$. Then*

$$\mathcal{W}_p\left(\sum_{i=1}^n X_i, N(0, I)\right) \leq CL_p.$$

This is enough to deduce the following estimate. We note that we work with $\mathcal{W}_3$ norm in order to establish the decaying $(n')^{-1/6}$ term in (14.2.2).

Lemma 14.2.12. *Let $(p_1, \dots, p_k)$ be a probability vector, $U_1, \dots, U_k \in O(d)$ generate an irreducible subgroup, $b_1, \dots, b_k \in \mathbb{R}^d$ and let $\rho \in (0, 1)$. Let μ be the probability measure on G given by $\mu = \sum_{i=1}^k p_i \delta_{g_i}$ where $g_i : x \mapsto \rho U_i x + b_i$. Suppose that μ is centred at zero and that all of the b_i have modulus at most 1. Let $\gamma_1, \gamma_2, \dots$ be i.i.d. samples from μ. Let $\varepsilon \in (0, 1)$.*

Given $\ell \in \mathbb{Z}_{>0}$ we define $S_\ell := \mathbb{E}[|\gamma_1 \dots \gamma_\ell(0)|^2]$ and

$$W_\ell := \mathcal{W}_3\left(d^{1/2} S_\ell^{-1/2} \gamma_1 \dots \gamma_\ell(0), N(0, I)\right).$$

Suppose that there exist $m, n \in \mathbb{Z}_{>0}$ such that for V a uniform random variable on the closure of the subgroup generated by the $U_1, \dots, U_k$ we have

$$\mathcal{W}_3(U(\gamma_1 \dots \gamma_m), V) < \varepsilon \qquad \text{and} \qquad \frac{m}{S_n^{1/2}} < \varepsilon.$$

Then for $n' \in \mathbb{Z}_{>0}$,

$$W_{(m+n)n'} \ll_d (T^{-1/6} + T^{1/2}\varepsilon)(W_n + 1) \tag{14.2.1}$$

where $T := \sum_{i=0}^{n'-1} \rho^{(m+n)i}$. In particular if $\rho^{(m+n)n'} > 1/2$ then $n'/2 \leq T \leq n'$ and therefore

$$W_{(m+n)n'} \ll_d ((n')^{-1/6} + (n')^{1/2}\varepsilon)(W_n + 1) \tag{14.2.2}$$

Proof. For $i = 1, \dots, n'$ let

$$X_i := \gamma_{(i-1)(n+m)+1} \cdots \gamma_{(i-1)(n+m)+m}$$

and

$$Y_i := \gamma_{(i-1)(n+m)+m+1} \cdots \gamma_{i(n+m)}$$

such that

$$Z_i = X_i Y_i = \gamma_{(i-1)(n+m)+1} \cdots \gamma_{i(n+m)}.$$

Furthermore consider $V_1, \ldots, V_k$ independent random variables which are uniform on H (the closure of the subgroup generated by the U_i), independent of the Y_i and are such that

$$\mathbb{E}[\|U(X_i) - V_i\|^3] < \varepsilon^3.$$

Note that

$$\begin{aligned} Z_1 \ldots Z_{n'}(0) = Z_1(0) + \rho^{(m+n)} U(Z_1) Z_2(0) + \\ \cdots + \rho^{(m+n)(n'-1)} U(Z_1 \ldots Z_{n'-1}) Z_{n'}(0). \end{aligned}$$

Also note that

$$\begin{aligned} &\mathcal{W}_3 \left(\rho^{(m+n)(i-1)} U(Z_1 \ldots Z_{i-1}) Z_i(0), \rho^{(m+n)(i-1)+m} V_i Y_i(0)\right) \\ &\quad = \rho^{(m+n)(i-1)} \mathcal{W}_3 \left(U(Z_1 \ldots Z_{i-1})(\rho^m U(X_i) Y_i(0) + X_i(0)), \rho^m V_i Y_i(0)\right) \\ &\quad \leq \rho^{(m+n)(i-1)} (m + \varepsilon \rho^m (\mathbb{E}\left[|Y_i(0)|^3\right])^{1/3}) \\ &\quad \ll_d \varepsilon \rho^{(m+n)(i-1)} S_n^{1/2} (W_n + 1), \end{aligned}$$

having used the triangle inequality in the second line and that $|X_i(0)| \leq m$ as $\sup_i |b_i| \leq 1$ as well as that

$$\begin{aligned} &\mathcal{W}_3 \left(U(Z_1 \ldots Z_{i-1}) U(X_i) Y_i(0), V_i Y_i(0)\right) \\ &\quad = \mathcal{W}_3 \left(U(Z_1 \ldots Z_{i-1}) U(X_i) Y_i(0), U(Z_1 \ldots Z_{i-1}) V_i Y_i(0)\right) \end{aligned}$$

as V_i is distributed like the Haar measure on H.

Note that by Lemma 14.1.1 the covariance matrix of $V_i Y_i(0)$ is $d^{-1} S_n I$. Therefore by Theorem 14.2.11 letting $A = d^{-1/2} \left(\frac{1-\rho^{2n'(m+n)}}{1-\rho^{2(m+n)}}\right)^{1/2} S_n^{1/2}$ we have that

$$\begin{aligned} &\mathcal{W}_3 \left(A^{-1} \left(\sum_{i=1}^{n'} \rho^{(m+n)(i-1)} V_i Y_i(0)\right), N(0, I)\right) \\ &\quad \ll \left(\sum_{i=1}^{n'} \mathbb{E}\left[|A^{-1} \rho^{(m+n)(i-1)} Y_i(0)|^3\right]\right)^{1/3} \\ &\quad \ll_d A^{-1} \left(\frac{1 - \rho^{3(m+n)n'}}{1 - \rho^{3(m+n)}}\right)^{1/3} (W_n + 1) \\ &\quad \ll_d T^{-1/6} (W_n + 1), \end{aligned}$$

where we exploited that

$$\frac{1-\rho^{2n'(m+n)}}{1-\rho^{2(m+n)}} = \frac{1-\rho^{n'(m+n)}}{1-\rho^{(m+n)}} \frac{1+\rho^{n'(m+n)}}{1+\rho^{(m+n)}} \in [T/2, T]$$

and a similar estimate for $\left(\frac{1-\rho^{3(m+n)n'}}{1-\rho^{3(m+n)}}\right)^{1/3}$.

Therefore we may deduce that

$$\mathcal{W}_3\left(A^{-1}\gamma_1 \dots \gamma_{(m+n)n'}(0), N(0,I)\right) \ll_d T^{-1/6}(W_n+1) + \varepsilon T^{1/2}(W_n+1)$$

By Lemma 14.2.10 we have that

$$\frac{d^{-1/2}S_{n'}^{-1/2}}{A} = 1 + O(\frac{m}{n}) = 1 + O(\varepsilon).$$

We conclude

$$\begin{aligned} W_{(m+n)n'} &\ll_d T^{-1/6}(W_n+1) + \varepsilon T^{1/2}(W_n+1) + \varepsilon \\ &\ll_d T^{-1/6}(W_n+1) + \varepsilon T^{1/2}(W_n+1) \end{aligned}$$

as required. □

From this we can deduce the following.

Corollary 14.2.13. *For every $\varepsilon > 0$ and every irreducible probability measure μ_U on $O(d)$ there is $C > 0$ and $\tilde{\rho} \in (0,1)$ such that the following is true. Let $\mu = \sum_{i=1}^{k} p_i \delta_{g_i}$ be a probability measure on G such that $U(\mu) = \mu_U$ and $p_i \geq \varepsilon$ for all $1 \leq i \leq k$. Assume further that $\max_{1 \leq i \leq k} |b(g_i)| = 1$ and for some $\rho \in (\tilde{\rho}, 1)$ we have $\rho(g_i) = \rho$ for all $1 \leq i \leq k$. Suppose that μ is centred at zero and let $\gamma_1, \gamma_2, \dots$ be i.i.d. samples from μ. Then for every $k \in \mathbb{Z}_{>0}$ such that $C^{k+1} < \frac{\rho^C}{1-\rho^C}$ there is some $n \in \mathbb{Z}_{>0}$ such that*

$$\frac{1}{1-\rho^n} \in [C^k, C^{k+1}]$$

and

$$\mathcal{W}_3(d^{1/2} S_n^{-1/2} \gamma_1 \dots \gamma_n(0), N(0,I)) < C.$$

Moreover, if $\mathrm{gap}_H(\mu) \geq \varepsilon$, then C and $\tilde{\rho}$ can be made uniform d and ε.

Proof. Let $\varepsilon' > 0$ be sufficiently small. Choose $m = m(\mu_U, \varepsilon')$ such that

$$\mathcal{W}_3(U(\gamma_1 \dots \gamma_m), V) < \varepsilon'$$

and choose $n_0 = n_0(\varepsilon, \varepsilon', \tilde{\rho})$ such that

$$\frac{m}{S_{n_0}^{1/2}} < \varepsilon'.$$

Note that this is possible by Lemma 14.2.10 as $\varepsilon \leq \mathbb{E}[|\gamma_1(0)|^3] \leq 1$ and providing we choose $\tilde{\rho}$ to be sufficiently close to 1 in terms of ε'. Now inductively chose n'_k such that $\sum_{i=0}^{n'_k-1} \rho^{(m+n_k)i} \in [\varepsilon'^{-3/2}, 2\varepsilon'^{-3/2}]$ and define $n_{k+1} := n'_k(n_k+m)$. Repeat this process until we find some k such that $\sum_{i=0}^{\infty} \rho^{(m+n_k)i} < \varepsilon'^{-3/2}$ and let k^* denote this value of k. By Lemma 14.2.12 this means that for $i = 1, \ldots, k^*$ we have

$$W_i \ll_d \varepsilon'^{1/4}(W_{i-1}+1).$$

Providing we take $\tilde{\rho}$ to be sufficiently close to 1 we can bound n_0 and W_{n_0} from above purely in terms of ε and ε'. This means that, providing we choose ε' to be sufficiently small, there is some $C_1 = C_1(\varepsilon, \varepsilon')$ such that for each $i = 1, \ldots, k^*$ we have

$$W_{n_i} < C_1.$$

We also have that

$$\frac{1-\rho^{n_{i+1}}}{1-\rho^{m+n_i}} \in [\varepsilon'^{-3/2}, 2\varepsilon^{-3/2}]$$

and so providing we choose $\tilde{\rho}$ to be sufficiently large we have

$$\frac{1-\rho^{n_{i+1}}}{1-\rho^{n_i}} \leq 4\varepsilon'^{-3/2}.$$

The result follows. When we have a spectral gap, all of these constants can be chosen to be uniform. $\square$

Now we have enough to prove Proposition 14.2.6.

Proof of Proposition 14.2.6. Without loss of generality we may assume that μ is centred at zero and that $\max_{i=1}^{k} |b_i| = 1$.

Let $\varepsilon' > 0$. By Lemma 14.2.8 there is some $m \in \mathbb{Z}_{>0}$ depending only on ε and ε' such that

$$\mathcal{W}_3(U(\gamma_1 \ldots \gamma_m(0)), V) < \varepsilon'.$$

By Lemma 14.2.10 there is some N depending only on μ_U and ε' such that for any $n \geq N$ we have

$$\frac{m}{S_n^{1/2}} < \varepsilon'.$$

Let C be as in Corollary 14.2.13 and choose n such that

$$\frac{1}{1-\rho^{m+n}} \in [C^{-1}\varepsilon'^{-3/2}, C\varepsilon'^{-3/2}].$$

Providing we choose $\tilde{\rho}$ sufficiently close to 1 we will also have $n \geq N$. By letting $n' \to \infty$ in Lemma 14.2.12 we deduce that

$$\mathcal{W}_3(A^{-1}\nu, N(0,I)) \ll_d C\varepsilon'^{1/4}$$

where $A = d^{1/2}(1-\rho^2)^{1/2} = \lim_{\ell\to\infty} d^{1/2} S_\ell^{-1/2}$. In the presence of a spectral gap, all of these bounds are easily seen to be uniform. □

Chapter 15

Construction of Examples

Throughout this section we denote as usual by $G = \mathrm{Sim}(\mathbb{R}^d)$. We first study random walk entropy in section 15.1 and then the separation rate in section 15.2. We prove Corollary 7.0.11 on real Bernoulli convolutions in section 15.5 as well as treat complex Bernoulli convolutions in section 15.6 proving Corollary 7.0.12. Finally, we discuss examples in $\mathbb{R}^d$ in section 15.4 and show Corollary 7.0.8, Corollary 7.0.9 and Corollary 7.0.10.

15.1 Bounding Random Walk Entropy

The techniques from [HS17, Section 6.3] or [Kit23, Section 9.2] follow through to our setting. In particular we have the following using Breuillard's strong Tits alternative.

Proposition 15.1.1. *([HS17, Section 6.3]) Let $d \geq 1$. Then for every $p_0 > 0$ there exists $\rho = \rho(p_0, d)$ such that if $\mu = \sum_{i=1}^k p_i \delta_{g_i}$ is a finitely supported probability measure on G with $p_i \geq p_0$ and $\mathrm{supp}(\mu)$ generates a non-virtually solvable subgroup, then $h_\mu \geq \rho$.*

We will also use the following version of the ping-pong lemma for which we provide a full proof for the convenience of the reader.

Lemma 15.1.2. *(Ping-Pong) Let G be a group acting on a set X and let $g_1, g_2 \in G$. Assume there exist disjoint non-empty sets $A_1, A_2 \subset X$ such*

$$g_1(A_1 \cup A_2) \subset A_1 \quad \text{and} \quad g_2(A_1 \cup A_2) \subset A_2.$$

Then g_1 and g_2 generate a free semigroup.

When this happens we say that g_1 and g_2 *play ping pong.*

Proof. Let $w_1 = h_1h_2\cdots h_{\ell_1}$ and $w_2 = f_1f_2\cdots f_{\ell_2}$ with distinct sequences $h_i, f_j \in \{g_1, g_2\}$. Assume without loss of generality that $\ell_1 \leq \ell_2$. First assume that there is some $1 \leq k \leq \ell_1$ such that $h_k \neq f_k$. Choose the smallest such k and note that it suffices to show that $h_k\cdots h_{\ell_1} \neq f_k\cdots f_{\ell_2}$, which follows by applying the resulting maps to any $x \in A_1 \cup A_2$ and noting that $h_k\cdots h_{\ell_1}x \neq f_k\cdots f_{\ell_2}x$. On the other hand assume that $h_i = f_i$ for all $1 \leq i \leq \ell_1$, in which case we need to show that $w' = f_{\ell_1+1}\cdots f_{\ell_2}$ is not the identity. Without loss of generality assume that $f_{\ell_1+1} = g_1$. Then for $x \in A_2$ we have that $w'x \in A_1$ and thus w' is not the identity. We note that in particular it follows by the assumptions that g_1 and g_2 have infinite order. □

Lemma 15.1.3. *Let μ be a finitely supported probability measure on G such that $g_1, g_2 \in \mathrm{supp}(\mu)$ generate a free semigroup. Then*

$$h_\mu \gg \min\{\mu(g_1), \mu(g_2)\}.$$

Proof. Denote $\mu' = \frac{1}{2}\delta_e + \frac{1}{2}\mu$. Then by [HS17, Lemma 6.8] we have $h_{\mu'} = h_\mu/2$. Thus the claim follows from [Kit23, Proposition 9.7] (generalised to G and applied to $K = \min\{\mu(g_1), \mu(g_2)\}/2$). □

15.1.1 p-adic Ping-Pong

We first use ping-pong in a p-adic setting. For a number field K with ring of integers O_K. Let $\mathfrak{p} \subset O_K$ be a prime ideal and we denote by $R_\mathfrak{p}$ the localization of O_K at P defined as

$$R_\mathfrak{p} = \left\{\frac{a}{b} : a \in O_K, b \in O_K\backslash\mathfrak{p}\right\}.$$

Lemma 15.1.4. *(p-adic Ping-Pong) Let K be a number field and let O_K be its ring of integers. Let $\mathfrak{p} \subset O_K$ be a prime ideal and let $M_\mathfrak{p}$ be the ideal of $R_\mathfrak{p}$ defined by*

$$M_\mathfrak{p} = \left\{\frac{a}{b} : a \in \mathfrak{p}, b \in O_K\backslash\mathfrak{p}\right\}.$$

Let $g_1, g_2 \in G$ be such that all of the entries of $\rho(g_1)U(g_1)$ and $\rho(g_2)U(g_2)$ are in $M_\mathfrak{p}$ and all components of b_1 and b_2 are in $R_\mathfrak{p}$. Suppose that

$$M_\mathfrak{p} \times \cdots \times M_\mathfrak{p} + b_1 \neq M_\mathfrak{p} \times \cdots \times M_\mathfrak{p} + b_2.$$

Then g_1 and g_2 generate a free semigroup.

Proof. This follows immediately from Lemma 15.1.2 with $X = R_\mathfrak{p} \times \cdots \times R_\mathfrak{p}$ and $A_i = M_\mathfrak{p} \times \cdots \times M_\mathfrak{p} + b_i$ for $i = 1, 2$. □

15.1.2 Ping-Pong under a Galois transform

We can also apply the ping-pong lemma using field automorphisms. Recall that given a number field K, the automorphism group $\mathrm{Aut}(K/\mathbb{Q})$ consists of field automorphisms that fix $\mathbb{Q}$.

Lemma 15.1.5. *(Galois Ping-Pong) Let g_1 and g_2 be two elements in G whose coefficients lie in a real number field K and without a common fixed point. Let $\Phi \in \mathrm{Aut}(K/\mathbb{Q})$ be such that for $i = 1, 2$ we have*

$$|\rho(\Phi(g_i))| < 1/3.$$

Then g_1 and g_2 generate a free semigroup.

Proof. For $i = 1, 2$ write $h_i = \Phi(g_i)$ and let p_i be the fixed point of h_i, which has coefficients in K since it arises from a linear equation over K. Then $h_1 \neq h_2$ as g_1 and g_2 have no common fixed point. Consider $A_i = B_{d(h_1,h_2)/2}(h_i)$ (the open ball around h_i of radius $d(h_1, h_2)/2$) and note further that $h_1(A_1 \cup A_2) \subset A_1$ and $h_2(A_1 \cup A_2) \subset A_2$. So the claim follows by Lemma 15.1.2. □

15.1.3 Height Entropy Bound in Dimension One

In dimension one we also have the following tool for bounding the random walk entropy. We use the absolute height $\mathcal{H}(\alpha)$ and the logarithmic heigh $h(\alpha)$ of an algebraic number α as defined in (7.0.4) and (7.0.5).

Proposition 15.1.6. *Suppose that μ is a finitely supported probability measure on G and that there exist $f, g \in \mathrm{supp}(\mu)$ which are of the form $f : x \mapsto \lambda_1 x + 1$ and $g : x \mapsto \lambda_2 x$ with λ_1 and λ_2 real algebraic and $\lambda_2 \neq 1$. Let $n = \lceil \frac{\log 3}{\max\{h(\lambda_1), h(\lambda_2)\}} \rceil + 2$. Then*

$$h_\mu \gg \frac{1}{n} \min\{\mu(f), \mu(g)\}^n.$$

This is a simple consequence of the following lemma.

Lemma 15.1.7. *Suppose that λ is algebraic and in some number field K. Let $f, g \in G$ be defined by $f : x \mapsto \lambda(x - a) + a$ and $g : x \mapsto \lambda(x - b) + b$ for some $a, b \in K$ with $a \neq b$. Suppose that $\mathcal{H}(\lambda) > 3$. Then f and g freely generate a free semi-group.*

Proof. First note that

$$\mathcal{H}(\lambda) = \mathcal{H}(\lambda^{-1}) = \prod_{v \in M_K} \min(1, |\lambda|_v)^{-\frac{n_v}{[K:\mathbb{Q}]}}.$$

This means that either there is some Archimedean place v such that $|\lambda|_v < 1/3$ or there is some non-Archimedean place v such that $|\lambda|_v < 1$.

In the Archimedean case there is some Galois transform ρ such that $|\rho(\lambda)| < 1/3$ and the result follows from Lemma 15.1.5. In the non-Archimedean case there is some prime ideal $\mathfrak{p} \subset \mathcal{O}_K$ with $\lambda \in \mathfrak{p}$ and Lemma 15.1.4 applies. □

We now deduce Proposition 15.1.6.

Proof of Proposition 15.1.6. For $n = \lceil \frac{\log 3}{\max\{h(\lambda_1), h(\lambda_2)\}} \rceil + 2$, using that $h(\alpha^n) = |n| h(\alpha)$ and $h(\alpha\beta) \geq h(\alpha) - h(\beta)$ for all α, β algebraic and $n \in \mathbb{Z}$, there exists $f', g' \in \{f, g\}^n$ satisfying the conditions of Lemma 15.1.7. Therefore by Lemma 15.1.3 we deduce that

$$h_{\mu^{*n}} \gg \min\{\mu(f), \mu(g)\}^n \qquad \text{and so} \qquad h_\mu \gg \frac{1}{n} \min\{\mu(f), \mu(g)\}^n$$

as required. □

15.2 Heights and Separation

In this subsection we will review some techniques for bounding S_μ using heights as defined in (7.0.4) and (7.0.5). We wish to bound the size of polynomials of algebraic numbers. To do this we need the following way of measuring the complexity of a polynomial.

Definition 15.2.1. *Given some polynomial $P \in \mathbb{Z}[X_1, X_2, \dots, X_n]$ we define the* length *of P, which we denote by $\mathcal{L}(P)$, to be the sum of the absolute values of the coefficients of P.*

We recall the following basic facts about heights.

Lemma 15.2.2. *The following properties hold:*

(i) $\mathcal{H}(\alpha^{-1}) = \mathcal{H}(\alpha)$ *for any non-zero algebraic number α.*

(ii) If α is a non-zero algebraic number of degree d,

$$\mathcal{H}(\alpha)^{-d} \leq |\alpha| \leq \mathcal{H}(\alpha)^d.$$

(iii) Given $P \in \mathbb{Z}[X_1, X_2, \dots, X_n]$ of degree at most $L_1 \geq 0$ in X_1, ..., $L_n \geq 0$ in X_n and algebraic numbers $\xi_1, \xi_2, \dots, \xi_n$ we have

$$\mathcal{H}(P(\xi_1, \xi_2, \dots, \xi_n)) \leq \mathcal{L}(P) \mathcal{H}(\xi_1)^{L_1} \dots \mathcal{H}(\xi_n)^{L_n}$$

Proof. (i) and (ii) are well-known and (iii) is [Mas16, Proposition 14.7]. □

Proposition 15.2.3. *Suppose that μ is a finitely supported measure on $G = \mathrm{Sim}(\mathbb{R}^d)$. Let S be the set of coefficients of $\rho(g), U(g)$ and $b(g)$ with $g \in \mathrm{supp}(\mu)$ supported on a finite set of points. Suppose that all of the elements of S are algebraic and let K be the number field generated by S. Then*

$$S_\mu \ll_d [K : \mathbb{Q}] \max(h(y) : y \in S\} \cup \{1\}).$$

Proof. We let $m, n \in \mathbb{Z}_{>0}$ and we consider an expression of the from

$$a_1^{-1} a_2^{-1} \dots a_n^{-1} b_1 b_2 \dots b_m$$

for $a_1, \dots, a_n$ and $b_1, \dots, b_m$ elements in the support of μ. We wish to show that this is either the identity or at least some distance away from the identity. Let $C := \max\{\mathcal{H}(y) : y \in S\}$. First note that

$$\rho(a_1^{-1} a_2^{-1} \dots a_n^{-1} b_1 b_2 \dots b_m) - 1$$

is a polynomial in elements of S and their inverses with length 2 and total degree at most $n + m$. Therefore by Lemma 15.2.2

$$H(\rho(a_1^{-1} a_2^{-1} \dots a_n^{-1} b_1 b_2 \dots b_m) - 1) \leq 2C^{m+n}$$

and so either $\rho(a_1^{-1} a_2^{-1} \dots a_n^{-1} b_1 b_2 \dots b_m) = 1$ or

$$|\rho(a_1^{-1} a_2^{-1} \dots a_n^{-1} b_1 b_2 \dots b_m) - 1| \geq 2^{-[K:\mathbb{Q}]} C^{-(m+n)[K:\mathbb{Q}]}.$$

By a similar argument, using that the coefficients of the inverse matrix of a matrix are polynomial in the coefficients of the given matrix, we see that either

$$U(a_1^{-1} a_2^{-1} \dots a_n^{-1} b_1 b_2 \dots b_m) = I$$

or

$$||U(a_1^{-1} a_2^{-1} \dots a_n^{-1} b_1 b_2 \dots b_m) - I|| \geq (d^{m+n} + 1)^{-[K:\mathbb{Q}]} C^{-O_d(m+n)[K:\mathbb{Q}]}$$

and that either $b(a_1^{-1} a_2^{-1} \dots a_n^{-1} b_1 b_2 \dots b_m) = 0$ or

$$|b(a_1^{-1} a_2^{-1} \dots a_n^{-1} b_1 b_2 \dots b_m)| \geq (d^{m+n} + 1)^{-[K:\mathbb{Q}]} C^{-O_d(m+n)[K:\mathbb{Q}]}.$$

Overall this means that either $a_1^{-1} a_2^{-1} \dots a_n^{-1} b_1 b_2 \dots b_m = \mathrm{Id}$ or

$$\log d(a_1^{-1} a_2^{-1} \dots a_n^{-1} b_1 b_2 \dots b_m, \mathrm{Id}) \gg_d -(m+n)(\log C + 1)[K : \mathbb{Q}].$$

The result follows. □

15.3 Inhomogeneous examples in $\mathbb{R}$

In this section we prove Corollary 7.0.7, which we recall for convenience of the reader.

Corollary (Restatement of Corollary 7.0.7). *For every $\varepsilon > 0$ there exists a small constant $c = c(\varepsilon) > 0$ such that the following holds. Let K be a number field and $\lambda_1, \lambda_2 \in K \cap (0,1)$ and write $h(\lambda_1, \lambda_2) = \max\{h(\lambda_1), h(\lambda_2)\}$. Consider the similarities given for $x \in \mathbb{R}$ as*

$$g_1(x) = \lambda_1 x \qquad \textit{and} \qquad g_2(x) = \lambda_2 x + 1.$$

Then the self-similar measure of $\frac{1}{2}\delta_{g_1} + \frac{1}{2}\delta_{g_2}$ is absolutely continuous if

$$h(\lambda_1, \lambda_2) \geq \varepsilon \qquad \textit{and} \qquad |\chi_\mu| \max\{1, \log([K:\mathbb{Q}]h(\lambda_1, \lambda_2))\}^2 < c.$$

Proof. (of Corollary 7.0.7) Write $\mu = \frac{1}{2}\delta_{g_1} + \frac{1}{2}\delta_{g_2}$ By Proposition 15.1.6 for every $\varepsilon > 0$ there exists a $\delta > 0$ such that if $h(\lambda_1, \lambda_2) \geq \varepsilon$ then it follows that $h_\mu \geq \delta$. Therefore by Theorem 7.0.4 and using that $S_\mu \ll h(\lambda_1, \lambda_2)[K:\mathbb{Q}]$ it follows that μ is absolutely continuous if for absolute constants C_1, C_2 it holds that

$$\frac{\delta}{|\chi_\mu|} \geq C_1 \max\{1, \log(C_2 \delta^{-1} h(\lambda_1, \lambda_2)[K:\mathbb{Q}])\}^2,$$

which easily implies the claim. □

15.4 Examples in $\mathbb{R}^d$

In this section we prove Corollary 7.0.8, Corollary 7.0.9 and Corollary 7.0.10 on general examples with absolutely continuous self-similar measures, which we all again recall for convenience of the reader.

Corollary (Restatement of Corollary 7.0.8). *Let $d \geq 1$ and $\varepsilon > 0$, let $\mu_U = \sum_{i=1}^k p_i \delta_{U_i}$ be an irreducible probability measure on $O(d)$ with $p_i \geq \varepsilon$ and let $b_1, \ldots, b_k \in \mathbb{R}^d$ with $b_1 \neq b_2$. Assume that $U_1, \ldots, U_k$ and $b_1, \ldots, b_k$ have algebraic coefficients. Let q be a prime number and for $1 \leq i \leq k$ consider*

$$g_i(x) = \frac{q}{q + a_{i,q}} U_i x + b_i \qquad \textit{for any integer} \qquad a_{i,q} \in [1, q^{1-\varepsilon}].$$

Assume that $g_1, \ldots, g_k$ do not have a common fixed point and consider $\mu = \sum_{i=1}^k p_i \delta_{g_i}$. Then the self-similar measure of μ is absolutely continuous for q a sufficiently large prime depending on $d, \varepsilon, U_1, \ldots, U_k$ and $b_1, \ldots, b_k$.

Proof of Corollary 7.0.8. We first show that g_1 and g_2 generate a free semigroup for sufficiently large q by using Lemma 15.1.4. For simplicity we first treat the case when all of the entries are rational. Then consider the q-adic numbers $\mathbb{Q}_q$ and the q-adic integers $\mathbb{Z}_q$. As the $U_1, \ldots, U_k$ and the $b_1, \ldots, b_k$ are fixed, for a sufficiently large prime q all of their entries are in $\mathbb{Z}_q \backslash q\mathbb{Z}_q$. On the other hand, by construction $\rho(g_i) \in q\mathbb{Z}_q$ for $1 \leq i \leq k$ and as $q\mathbb{Z}_q$ is an ideal therefore also all of the entries of $\rho(g_i)U_i$ are in $q\mathbb{Z}_q$. By Lemma 15.1.4 it therefore suffices to check that $(q\mathbb{Z}_q)^d + b_1 \neq (q\mathbb{Z}_q)^d + b_2$ or equivalently $b_1 - b_2 \notin (q\mathbb{Z}_q)^d$, which is clearly the case for sufficiently large q. Thus g_1 and g_2 generate a free semigroup. The same argument applies in the general case for K the number field generated by the coefficients of the entries of g_i and by choosing any prime ideal that factors (q).

Thus it follows by Lemma 15.1.3 that $h_\mu \gg \varepsilon$ and note that by Lemma 15.2.2 it holds that $S_\mu \ll_{K,d} \log q$. Hence there exists a constant C depending on all the relevant parameters such that the self-similar measure of μ is absolutely continuous if

$$C|\chi_\mu| \leq \frac{1}{(\log \log q)^2}.$$

Therefore it remains to estimate the Lyapunov exponent. Indeed, note that

$$\log\left(\frac{q}{q + a_{i,q}}\right) = \log\left(1 - \frac{a_{i,q}}{q + a_{i,q}}\right) \geq \log\left(1 - \frac{q^{1-\varepsilon}}{q}\right) \gg -q^{-\varepsilon}.$$

Therefore $|\chi_\mu| \ll q^{-\varepsilon}$ and the claim follows for sufficiently large q. □

Corollary (Restatement of Corollary 7.0.9). *Let d, ε and $\mu_U = \sum_{i=1}^k p_i \delta_{U_i}$ as well as $b_1, \ldots, b_k$ be as in Proposition 7.0.8. Let q be a prime number and consider for $1 \leq i \leq k-1$*

$$g_i(x) = \frac{q}{q+3} U_i x + b_i \qquad \text{and} \qquad g_k(x) = \frac{q}{q-1} U_k x + b_k.$$

Assume that $g_1, \ldots, g_k$ do not have a common fixed point and further that

$$p_k \leq \frac{1}{3}.$$

Then the self-similar measure of $\mu = \sum_{i=1}^k p_i \delta_{g_i}$ is absolutely continuous for q a sufficiently large prime depending on $d, \varepsilon, U_1, \ldots, U_k$ and $b_1, \ldots, b_k$.

Proof of Corollary 7.0.9. As in the proof of Corollary 7.0.8, g_1 and g_2 generate a free semigroup for sufficiently large q and therefore $h_\mu \gg \varepsilon$. Write $\alpha_1 = p_1 + \ldots + p_{k-1}$ and $\alpha_2 = p_k$. Then we have as $\alpha_1 + \alpha_2 = 1$,

$$\mathbb{E}_{\gamma \sim \mu}[\rho(\gamma)] = \alpha_1 \frac{q}{q+3} + \alpha_2 \frac{q}{q-1} = \frac{q^2 + (4\alpha_2 - 1)q}{(q+3)(q-1)}$$

and thus

$$1 - \mathbb{E}_{\gamma\sim\mu}[\rho(\gamma)] = \frac{(q+3)(q-1) - (q^2 + (4\alpha_2 - 1)q)}{(q+3)(q-1)} = \frac{(3-4\alpha_2)q - 3}{(q+3)(q-1)}.$$

On the other hand, choosing $\hat{\rho} = \frac{q}{q+3}$ we have

$$\mathbb{E}_{\gamma\sim\mu}[|\hat{\rho} - \rho(\gamma)|] = \alpha_2 \left(\frac{q}{q-1} - \frac{q}{q+3} \right) = \frac{4q\alpha_2}{k(q+3)(q-1)}.$$

Thus it follows that

$$\lim_{q\to\infty} \frac{\mathbb{E}_{\gamma\sim\mu}[|\hat{\rho} - \rho(\gamma)|]}{1 - \mathbb{E}_{\gamma\sim\mu}[\rho(\gamma)]} = \frac{4\alpha_2}{3 - 4\alpha_2} < 1 \tag{15.4.1}$$

provided that $\alpha_2 = p_k < \frac{3}{8}$. If we assume that $p_k \leq \frac{1}{3}$ then we have that the limit in (15.4.1) is uniformly away from 1. As in Corollary 7.0.8, we have that $S_\mu \ll_{K,d} \log q$. Therefore by Theorem 7.0.6 there exists a constant C depending on all of the parameters such that μ is absolutely continuous if

$$C|\chi_\mu| \leq \frac{1}{(\log\log q)^2}.$$

As in Corollary 7.0.8 it follows that $|\chi_\mu| \ll q^{-1}$ and hence the claim follows. □

We next prove Corollary 7.0.10 and first show the following basic lemma.

Lemma 15.4.1. *Let K be a real algebraic number field satisfying $\mathbb{Q}(\sqrt{q}) \subset K$ for a prime q. Then there exists a field automorphism $\Phi \in \mathrm{Aut}(K/\mathbb{Q})$ such that $\Phi(\sqrt{q}) = -\sqrt{q}$.*

Proof. Write $K_0 = \mathbb{Q}(\sqrt{q})$ and assume that $K = K_0(\alpha_1, \ldots, \alpha_\ell)$ for some $\alpha_1, \ldots, \alpha_\ell \in K$. Denote by $\Theta \in \mathrm{Aut}(K_0/\mathbb{Q})$ the automorphism with $\Theta(\sqrt{q}) = -\sqrt{q}$. When $\ell = 1$ we consider the surjective map $K_0[X] \to K_0(\alpha)$ with $P \mapsto \Theta(P)(\alpha_1)$ for $\Theta(P)$ the polynomial to which all coefficients we have applied Θ. This map induces a field automorphism of $K_0(\alpha)$ with the required properties and our proof is concluded by an induction on ℓ with the same argument. □

Corollary (Restatement of Corollary 7.0.10)**.** *Let $d \geq 1$ and $\varepsilon \in (0,1)$ and $\mu_U = \sum_{i=1}^k p_i \delta_{U_i}$ an irreducible probability measure on $O(d)$ with $p_i \geq \varepsilon$ for all $1 \leq i \leq k$. Assume furthermore that $U_1, \ldots, U_k$ have algebraic entries. Let $\tilde{\rho} \in (0,1)$ be sufficiently close to 1 in terms of d, ε and μ_U and let $C > 1$ be sufficiently large depending on the same parameters.*

Suppose that $g_i(x) = \frac{a_i + b_i\sqrt{q}}{c_i} U_i x + d_i$ with $a_i, b_i, c_i \in \mathbb{Z}$ and $d_i \in \mathbb{Z}^d$ for $1 \leq i \leq k$ and a prime number q do not have a common fixed point. Then the self-similar measure associated to $\mu = \sum_{i=1}^k p_i \delta_{g_i}$ is absolutely continuous if the following properties are satisfied:

(i) $\frac{a_i+b_i\sqrt{q}}{c_i} \in (\tilde{\rho}, 1)$ *for* $1 \leq i \leq k$,

(ii) for $j = 1$ *and for* $j = 2$ *we have*

$$\left|\frac{a_j - b_j\sqrt{q}}{c_j}\right| < \frac{1}{3},$$

(iii) For $L = \max(\sqrt{q}, |a_i|, |b_i|, |c_i|, |d_i|_\infty)$ *we have*

$$C|\chi_\mu| \leq \frac{1}{(\log(\log L))^2}.$$

Proof of Corollary 7.0.10. By Theorem 7.0.4 there exists $\tilde{\rho} \in (0,1)$ and $C \geq 1$ depending on d, ε and μ_U such that μ is absolutely continuous if $p_i \geq \varepsilon$ as well as $\frac{a_i+b_i\sqrt{q}}{c_i} \in (\tilde{\rho}, 1)$ for all $1 \leq i \leq k$ as well as

$$\frac{h_\mu}{|\chi_\mu|} \geq C\left(\max\left\{1, \log\frac{S_\mu}{h_\mu}\right\}\right)^2.$$

Let K be the number field generated by all the coefficients of elements in $\mathrm{supp}(\mu)$. Then by Lemma 15.4.1 there is a field automorphism $\Phi \in \mathrm{Aut}(K/\mathbb{Q})$ such that $\Phi(\sqrt{q}) = -\sqrt{q}$ and therefore we have that $|\rho(\Phi(g_j))| < \frac{1}{3}$ for $j = 1, 2$. Thus by Lemma 15.1.5 and Lemma 15.1.3 we have that $h_\mu \gg \varepsilon$. We also have $h_\mu \leq \log\varepsilon^{-1}$. On the other hand, it follows by Lemma 15.2.2 (iii) and Proposition 15.2.3 that $S_\mu \ll_{d,\mu_U} \log L$, which readily implied the claim upon changing the constant C. □

Lemma 15.4.2. *In the setting of Corollary 7.0.10, for* $\varepsilon > 0$ *choose*

$$a_i = \lceil\sqrt{q}\rceil - m_{i,q}, \qquad b_i = 2 \qquad c_i = 3\lceil\sqrt{q}\rceil$$

for $m_{i,q}$ *an integer satisfying* $m_{i,q} \in [0, q^{1/2-\varepsilon}]$ *and any* $d_i \in \mathbb{Z}^d$ *with* $|d_i|_\infty \leq \exp(\exp(q^{\varepsilon/3}))$. *Then* μ *is absolutely continuous for sufficiently large* q *depending on* d, p_0, ε *and* $U_1, \ldots, U_k$, *provided* $g_1, \ldots, g_k$ *does not have a common fixed point.*

Proof. It holds that $\frac{a_i+b_i\sqrt{q}}{c_i} \in (0,1)$ converges to 1 as $q \to \infty$ and that $|\frac{a_i-b_i\sqrt{q}}{c_i}| < \frac{1}{3}$. We next estimate the Lyapunov exponent of μ. Indeed, note that for q large enough,

$$\begin{aligned}\log\left(\frac{a_i+b_i\sqrt{q}}{c_i}\right) &\geq \log\left(\frac{\lceil\sqrt{q}\rceil - q^{1/2-\varepsilon} + 2\sqrt{q}}{3\lceil\sqrt{q}\rceil}\right)\\ &\geq \log\left(1 - \frac{2(\lceil\sqrt{q}\rceil - \sqrt{q}) + q^{1/2-\varepsilon}}{3\lceil\sqrt{q}\rceil}\right) \gg -q^{-\varepsilon}\end{aligned}$$

and therefore $|\chi_\mu| \ll q^{-\varepsilon}$. In our case, for large q we have $L = |d_i|_\infty = \exp(\exp(q^{\varepsilon/3}))$ and therefore $\log(\log L) = q^{\varepsilon/3}$. Thus for sufficiently large q we have that $C|\chi_\mu| \leq (\log\log L)^{-2} = q^{-2\varepsilon/3}$ and the claim follows. □

15.5 Real Bernoulli Convolutions

In this section we prove Corollary 7.0.11.

Corollary (Restatement of Corollary 7.0.11)**.** *There is an absolute constant $C > 1$ such that the following holds. Let $\lambda \in (1/2, 1)$ be a real algebraic number. Then the Bernoulli convolution ν_λ is absolutely continuous on $\mathbb{R}$ if*

$$\lambda > 1 - C^{-1} \min\{\log M_\lambda, (\log\log M_\lambda)^{-2}\}. \tag{15.5.1}$$

Proof of Corollary 7.0.11. As in the paragraph before Proposition 14.2.2, Bernoulli convolutions are uniformly non-degenerate. Since we are in $d = 1$ they are $(1, 0)$-well-mixing and therefore Theorem 8.1.4 applies. For convenience write $\eta = \log M_\lambda$ and $h_\lambda = h_{\nu_\lambda}$. We don't keep track of possible enlargements of C. That Bernoulli convolutions are uniformly non-degenerate follows from Proposition 8.1.2. Then Theorem 8.1.4 implies that if

$$(1-\lambda)^{-1} h_\lambda > C \left(\max\left\{1, \log \eta / h_\lambda\right\}\right)^2, \tag{15.5.2}$$

then ν_λ is absolutely continuous. Recall that by [BV20, Theorem 5] (which is stated with logarithms base 2) there is an absolute $c_0 \in (0, 1)$ such that $c_0 \min(\log 2, \eta) \leq h_\lambda \leq \min(\log 2, \eta)$.

We proceed with a case distinction. First assume that $\eta \leq \log 2$. Then $c_0^{-1} \geq \eta / h_\lambda \geq 1$ and therefore by (15.5.2) the condition $(1-\lambda)^{-1} c_0 \eta > C$ is sufficient for absolute continuity, which is equivalent to

$$\lambda > 1 - C^{-1} \eta. \tag{15.5.3}$$

Next assume that $\eta \geq \log 2$. Then $c_0 \log 2 \leq h_\lambda \leq \log 2$ and so (15.5.2) gives

$$(1-\lambda) \max\{1, \log \eta + \log(c_0 \log 2)^{-1}\}^2 < C^{-1}.$$

Note that $\max\{1, \log \eta + \log(c_0 \log 2)^{-1}\} \leq 2 \log(c_0 \log 2)^{-1} \max\{1, \log \eta\}$. Therefore we get the condition

$$\lambda > 1 - C^{-1} \max\{1, \log \eta\}^{-2} = 1 - C^{-1} \min\{1, (\log \eta)^{-2}\}. \tag{15.5.4}$$

To deduce (15.5.1), we note that there is a unique $\eta' > 0$ with $\eta' = (\log \eta')^{-2}$ and this η' satisfies $2 \leq \eta' \leq 5/2$. Moreover $\log \eta < (\log \eta)^{-2}$ for $0 < \eta < \eta'$ and $\log \eta > (\log \eta)^{-2}$ for $\eta > \eta'$. Therefore (15.5.1) holds for $\eta < \log(2)$ and $\eta > 2\eta'$ by (15.5.3) and (15.5.4). In the range $\log(2) < \eta < 2\eta'$, we enlarge C to ensure that (15.5.1) holds. □

We note that if λ is algebraic and not the root of any non-zero polynomial with coefficients $0, \pm 1$, then $h_\lambda = 2$ and also as mentioned in Remark 5.10 of [Kit21], $M_\lambda \geq 2$. Therefore for such a λ, ν_λ is absolutely continuous if

$$\lambda > 1 - C^{-1} \min\{1, (\log \log M_\lambda)^{-2}\}. \tag{15.5.5}$$

15.6 Complex Bernoulli Convolutions

Corollary (Restatement of Corollary 7.0.12)**.** *For every $\varepsilon > 0$ there is a constant $C \geq 1$ such that the following holds. Let $\lambda \in \mathbb{C}$ be a complex algebraic number such that $|\lambda| \in (2^{-1/2}, 1)$ and*

$$|\mathrm{Im}(\lambda)| \geq \varepsilon. \tag{15.6.1}$$

Then the Bernoulli convolution ν_λ is absolutely continuous on $\mathbb{C}$ if

$$|\lambda| > 1 - C^{-1} \min\{\log M_\lambda, (\log \log M_\lambda)^{-2}\}.$$

Proof of Corollary 7.0.12. We can't directly apply Proposition 8.1.2 so we give a direct proof of mixing and non-degeneracy. First note that (15.6.1) ensures that there is some $c > 0$ and $T \geq 1$ depending only on ε such that the (c, T)-well-mixing property is satisfied.

To deal with non-degeneracy, we distinguish the case when $|\lambda| \leq \lambda_0$ and $|\lambda| \geq \lambda_0$ for some λ_0 sufficiently close to 1. As in the case of real Bernoulli convolution, for any given λ_0, the family of Bernoulli convolutions with $|\lambda| \leq \lambda_0$ are easily seen to be uniformly non-degenerate depending on λ_0. To deal with the case $\lambda \geq \lambda_0$, we rescale our measure to the one given by the law of $B_\lambda = \sqrt{1 - |\lambda|^2} \sum_{i=0}^{\infty} \pm \lambda^i$ and denote the resulting measure by ν'_λ. Now let Σ be the covariance matrix of ν'_λ under the natural identification of $\mathbb{C}$ with $\mathbb{R}^2$. Note that the trace of Σ is 1 and we claim that the smallest eigenvalue of Σ is $\gg_\varepsilon 1$. Indeed, for a unit vector $x \in \mathbb{R}^2$ we want to estimate $x^T \Sigma x$, which is by identifying $\mathbb{C}$ with $\mathbb{R}^2$ equal to

$$\mathbb{E}[|B_\lambda \cdot x|^2] = (1 - |\lambda|^2) \sum_{i=0}^{\infty} |\lambda^i \cdot x|^2 \gg_\varepsilon 1,$$

which follows as $|\lambda^i \cdot x|^2 \gg |\lambda|^2$ unless λ^i and x are almost colinear, which is only the case for a very small proportion of i's. It follows that

$$\inf_{p \in P(2)} \mathbb{E}_{x \sim \mathcal{N}(0, \Sigma)}[|px|] \gg_\varepsilon 1$$

for p ranging in the orthogonal projections of $\mathbb{R}^2$ as in section 14.2. By for example Lemma 10.4.3 we know that $\mathcal{W}_1(\nu'_\lambda, N(0,\Sigma)) \ll \sqrt{1-|\lambda|^2}$. Therefore for λ_0 sufficiently close to 1 in terms of ε, uniform non-degeneracy follows as in Lemma 14.2.5. Having establish uniform well-mixing and non-degeneracy, Corollary 7.0.12 is established by the same argument as the proof of Corollary 7.0.11. $\square$

Bibliography

[Agm65] S. Agmon, *Lectures on elliptic boundary value problems*, Van Nostrand Mathematical Studies, No. 2, D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto-London, 1965. Prepared for publication by B. Frank Jones, Jr. with the assistance of George W. Batten, Jr.

[AK63] V. I. Arnold and A. L. Krylov, *Uniform distribution of points on a sphere and certain ergodic properties of solutions of linear ordinary differential equations in a complex domain*, Dokl. Akad. Nauk SSSR **148** (1963), 9–12.

[And04] N. Andersen, *Real Paley-Wiener theorems for the inverse Fourier transform on a Riemannian symmetric space*, Pacific Journal of Mathematics **213** (2004), 1–13.

[ARHW21] A. Algom, F. Rodriguez Hertz, and Z. Wang, *Pointwise normality and Fourier decay for self-conformal measures*, Adv. Math. **393** (2021).

[Aub98] T. Aubin, *Some nonlinear problems in Riemannian geometry*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 1998.

[Ave73] A. Avez, *Limite de quotients pour des marches aléatoires sur des groupes*, C. R. Acad. Sci. Paris Sér. A-B **276** (1973).

[BB23] T. Bénard and E. Breuillard, *Local limit theorems for random walks on nilpotent lie groups* (2023). https://arxiv.org/abs/2103.12684.

[BdlHV08] B. Bekka, P. de la Harpe, and A. Valette, *Kazhdan's property (T)*, New Mathematical Monographs, vol. 11, Cambridge University Press, Cambridge, 2008.

[BdS16] Y. Benoist and N. de Saxcé, *A spectral gap theorem in simple Lie groups*, Invent. Math. **202** (2016), no. 2, 337–361.

[BE88] M. F. Barnsley and J. H. Elton, *A new class of markov processes for image encoding*, Adv. in Appl. Probab **20** (1988), no. 1, 14–32.

[BFLM11] J. Bourgain, A. Furman, E. Lindenstrauss, and S. Mozes, *Stationary measures and equidistribution for orbits of nonabelian semigroups on the torus*, J. Amer. Math. Soc. **24** (2011), no. 1, 231–280.

[BG08] J. Bourgain and A. Gamburd, *On the spectral gap for finitely-generated subgroups of* SU(2), Invent. Math. **171** (2008), no. 1, 83–121.

[BG12] J. Bourgain and A. Gamburd, *A spectral gap theorem in* $\mathrm{SU}(d)$, J. Eur. Math. Soc. (JEMS) **14** (2012), no. 5, 1455–1511.

[BG21] E. Bayraktar and G. Guo, *Strong equivalence between metrics of Wasserstein type*, Electron. Commun. Probab. **26** (2021).

[BH24] T. Bénard and W. He, *Multislicing and effective equidistribution for random walks on some homogeneous spaces* (2024). https://www.arxiv.org/pdf/2409.03300.

[BISG17] R. Boutonnet, A. Ioana, and A. Salehi Golsefidy, *Local spectral gap in simple Lie groups and applications*, Invent. Math. **208** (2017), no. 3, 715–802.

[BKS24] S. Baker, O. Khalil, and T. Sahlsten, *Fourier decay from l2-flattening* (2024). https://arxiv.org/abs/2407.16699.

[BL85] P. Bougerol and J. Lacroix, *Products of random matrices with applications to Schrödinger operators*, Progress in Probability and Statistics, vol. 8, Birkhäuser Boston, Inc., Boston, MA, 1985.

[Bou12] J. Bourgain, *Finitely supported measures on $SL_2(\mathbb{R})$ which are absolutely continuous at infinity*, Geometric aspects of functional analysis, 2012, pp. 133–141.

[Bou81] P. Bougerol, *Théorème central limite local sur certains groupes de Lie*, Ann. Sci. École Norm. Sup. (4) **14** (1981), no. 4, 403–432 (1982).

[BP92] P. Bougerol and N. Picard, *Strict stationarity of generalized autoregressive processes*, Ann. Probab. **20** (1992), no. 4, 1714–1730.

[BPS12] B. Bárány, M. Pollicott, and K. Simon, *Stationary measures for projective transformations: the Blackwell and Furstenberg measures*, J. Stat. Phys. **148** (2012), no. 3, 393–421.

[BQ11] Y. Benoist and J.-F. Quint, *Mesures stationnaires et fermés invariants des espaces homogènes*, Ann. of Math. (2) **174** (2011), no. 2, 1111–1162.

[BQ16] ———, *Random walks on reductive groups*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics], vol. 62, Springer, Cham, 2016.

[BQ18] ———, *On the regularity of stationary measures*, Israel J. Math. **226** (2018), no. 1, 1–14.

[Bré21] J. Brémont, *Self-similar measures and the Rajchman property*, Ann. H. Lebesgue **4** (2021), 973–1004.

[Bre05a] E. Breuillard, *Distributions diophantiennes et théorème limite local sur $\mathbb{R}^d$*, Probab. Theory Related Fields **132** (2005), no. 1, 39–73.

[Bre05b] ———, *Local limit theorems and equidistribution of random walks on the heisenberg group*, Geom. Funct. Anal. **25** (2005), 35–82.

[Bre08] E. Breuillard, *A strong tits alternative* (2008). https://arxiv.org/abs/0804.1395.

[Bre92] L. Breiman, *Probability*, Classics in Applied Mathematics, vol. 7, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992. Corrected reprint of the 1968 original.

[BS23] S. Baker and T. Sahlsten, *Spectral gaps and fourier dimension for self-conformal sets with overlaps* (2023). https://arxiv.org/abs/2306.01389.

[BSSŚ22] B. Bárány, K. Simon, B. Solomyak, and A. Śpiewak, *Typical absolute continuity for classes of dynamically defined measures*, Adv. Math. **399** (2022).

[BtD85] T. Bröcker and T. tom Dieck, *Representations of compact Lie groups*, Graduate Texts in Mathematics, vol. 98, Springer-Verlag, New York, 1985.

[BV19] E. Breuillard and P. P. Varjú, *On the dimension of Bernoulli convolutions*, Ann. Probab. **47** (2019), no. 4, 2582–2617.

[BV20] E. Breuillard and P. P. Varjú, *Entropy of Bernoulli convolutions and uniform exponential growth for linear groups*, J. Anal. Math. **140** (2020), no. 2, 443–481.

[CG13] J.-P. Conze and Y. Guivarc'h, *Ergodicity of group actions and spectral gap, applications to random walks and Markov shifts*, Discrete Contin. Dyn. Syst. **33** (2013), no. 9, 4239–4269.

[DFW07] X.-R. Dai, D.-J. Feng, and Y. Wang, *Refinable functions with non-integer dilations*, J. Funct. Anal. **250** (2007), no. 1, 1–20.

[DH21] P. Diaconis and R. Hough, *Random walk on unipotent matrix groups*, Ann. Sci. Éc. Norm. Supér. (4) **54** (2021), no. 3, 587–625.

[dS13] N. de Saxcé, *Trou dimensionnel dans les groupes de Lie compacts semisimples via les séries de Fourier*, J. Anal. Math. **120** (2013), 311–331.

[DS58] N. Dunford and J. T. Schwartz, *Linear Operators. I. General Theory*, Pure and Applied Mathematics, Vol. 7, Interscience Publishers, Inc., New York; Interscience Publishers, Ltd., London, 1958. With the assistance of W. G. Bade and R. G. Bartle.

[Erd39] P. Erdös, *On a family of symmetric Bernoulli convolutions*, Amer. J. Math. **61** (1939), 974–976.

[Erd40] P. Erdös, *On the smoothness properties of a family of Bernoulli convolutions*, Amer. J. Math. **62** (1940), 180–186.

[Eri73] R. V. Erickson, *On an L_p version of the Berry-Esseen theorem for independent and m-dependent variables*, Ann. Probability **1** (1973), 497–503. MR383502

[EW11] M. Einsiedler and T. Ward, *Ergodic theory with a view towards number theory*, Graduate Texts in Mathematics, vol. 259, Springer-Verlag London, Ltd., London, 2011. MR2723325

[EW17] ———, *Functional analysis, spectral theory, and applications*, Graduate Texts in Mathematics, vol. 276, Springer, Cham, 2017.

[FF22] D.-J. Feng and Z. Feng, *Estimates on the dimension of self-similar measures with overlaps*, J. Lond. Math. Soc. (2) **105** (2022), no. 4, 2104–2135.

[FH09] D.-J. Feng and H. Hu, *Dimension theory of iterated function systems*, Comm. Pure Appl. Math. **62** (2009), no. 11, 1435–1500.

[FM21] T. Finis and J. Matz, *On the asymptotics of Hecke operators for reductive groups*, Math. Ann. **380** (2021), no. 3-4, 1037–1104.

[Gar62] A. M. Garsia, *Arithmetic properties of Bernoulli convolutions*, Trans. Amer. Math. Soc. **102** (1962), 409–432.

[GdM89] I. Ya. Gol′ dsheĭd and G. A. Margulis, *Lyapunov exponents of a product of random matrices*, Uspekhi Mat. Nauk **44** (1989), no. 5(269), 13–60.

[Ger80] P. Gerl, *A ratio limit theorem*, Conference on Random Walks (Kleebach, 1979) (French), 1980, pp. 7–14.

[Gou14] S. Gouëzel, *Local limit theorem for symmetric random walks in gromov-hyperbolic groups*, J. Amer. Math. Soc **27** (2014), no. 4, 893–928.

[GP16] Y. Guivarc'h and E. L. Page, *Spectral gap properties for linear random walks and pareto's asymptotics for affine stochastic recursions*, Ann. Inst. H. Poincaré Probab. Statist. **52** (2016), no. 2, 503–574.

[Gui80] Y. Guivarc'h, *Sur la loi des grands nombres et le rayon spectral d'une marche aléatoire*, Conference on Random Walks (Kleebach, 1979) (French), 1980, pp. 47–98, 3.

[Hal15] B. C. Hall, *Lie groups, Lie algebras, and representations*, Second, Graduate Texts in Mathematics, vol. 222, Springer-Verlag, Cham, 2015. An elementary introduction.

[HC58] Harish-Chandra, *Spherical functions on a semisimple Lie group. II*, Amer. J. Math. **80** (1958), 553–613.

[Hel78] S. Helgason, *Differential geometry, Lie groups, and symmetric spaces*, Pure and Applied Mathematics, vol. 80, Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1978.

[Hel84] ______, *Groups and geometric analysis*, Pure and Applied Mathematics, vol. 113, Academic Press, Inc., Orlando, FL, 1984. Integral geometry, invariant differential operators, and spherical functions.

[Hoc14] M. Hochman, *On self-similar sets with overlaps and inverse theorems for entropy*, Ann. of Math. (2) **180** (2014), no. 2, 773–822.

[Hoc17] ______, *On self-similar sets with overlaps and inverse theorems for entropy in R^d* (2017). https://arxiv.org/pdf/1503.09043.pdf.

[Hou19] R. Hough, *The local limit theorem on nilpotent Lie groups*, Probab. Theory Related Fields **174** (2019), no. 3-4, 761–786.

[HS17] M. Hochman and B. Solomyak, *On the dimension of Furstenberg measure for $SL_2(\mathbb{R})$ random matrix products*, Invent. Math. **210** (2017), no. 3, 815–875.

[Hut81] J. E. Hutchinson, *Fractals and self-similarity*, Indiana Univ. Math. J. **30** (1981), no. 5, 713–747.

[IK40] K. Ito and Y. Kawada, *On the probability distribution on a compact group*, Proc. Phys.-Math. Soc. Japan **22** (1940), 977–998.

[JW35] B. r. Jessen and A. Wintner, *Distribution functions and the Riemann zeta function*, Trans. Amer. Math. Soc. **38** (1935).

[Kat95] T. Kato, *Perturbation theory for linear operators*, Classics in Mathematics, Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition.

[Kes74] H. Kesten, *Renewal theory for functionals of a Markov chain with general state space*, Ann. Probability **2** (1974), 355–386.

[Kit21] S. Kittle, *Absolutely continuous self-similar measures with exponential separation* (2021). https://arxiv.org/abs/2103.12684, to appear at Annales scientifiques de l'Ecole normale supérieure.

[Kit23] ______, *Absolutely continuous Furstenberg measures* (2023). https://arxiv.org/pdf/2305.05757.pdf.

[KK24] W. Kim and C. Kogler, *Effective density of non-degenerate random walks on homogeneous spaces*, Int. Math. Res. Not. **2024** (2024), no. 11, 9218–9236.

[KK25a] S. Kittle and C. Kogler, *Dimension of contracting on average self-similar measures* (2025). https://arxiv.org/abs/2501.17795.

[KK25b] ______, *Entropy theory for random walks on lie groups* (2025). https://arxiv.org/abs/2501.05395.

[KK25c] ______, *On absolute continuity of inhomogeneous and contracting on average self-similar measures* (2025). https://arxiv.org/abs/2409.18936.

[KK25d] ______, *Polynomial tail decay for stationary measures* (2025). https://arxiv.org/abs/2502.00086.

[KLP11] V. A. Kaimanovich and V. Le Prince, *Matrix random products with singular harmonic measure*, Geom. Dedicata **150** (2011), 257–279.

[Kna02] A. W. Knapp, *Lie groups beyond an introduction*, Second, Progress in Mathematics, vol. 140, Birkhäuser Boston, Inc., Boston, MA, 2002.

[Kog22] C. Kogler, *Local limit theorem for random walks on symmetric spaces* (2022). https://arxiv.org/abs/2211.11128, To appear at J. Anal. Math.

[Lal93] S. Lalley, *Finite range random walk on free groups and homogeneous trees*, Ann. Probab. **21** (1993), no. 4, 2087–2130.

[Leq22] F. Lequen, *Absolutely continuous Furstenberg measures for finitely-supported random walks* (2022). `https://arxiv.org/abs/2205.11138.pdf`.

[LS20] J. Li and T. Sahlsten, *Fourier transform of self-affine measures*, Adv. Math. **374** (2020).

[LS22] ———, *Trigonometric series and self-similar sets*, J. Eur. Math. Soc. (JEMS) **24** (2022), no. 1, 341–368.

[LV16] E. Lindenstrauss and P. P. Varjú, *Random walks in the group of Euclidean isometries and self-similar measures*, Duke Math. J. **165** (2016), no. 6, 1061–1127.

[Mas16] D. Masser, *Auxiliary polynomials in number theory*, Cambridge Tracts in Mathematics, vol. 207, Cambridge University Press, Cambridge, 2016.

[MKMSG22] K. Mallahi-Karai, A. Mohammadi, and A. Salehi Golsefidy, *Locally random groups*, Michigan Math. J. **72** (2022), 479–527.

[New51] J. D. Newburgh, *The variation of spectra*, Duke Math. J. **18** (1951), 165–176.

[Rap22] A. Rapaport, *On the Rajchman property for self-similar measures on* $\mathbb{R}^d$, Adv. Math. **403** (2022), Paper No. 108375, 53. MR4405371

[Rem] C. Remling, *Uniform decay of operator norm for smooth family of operators*. URL:https://mathoverflow.net/q/434679 (version: 2022-11-15).

[Sak85] A. I. Sakhanenko, *Estimates in an invariance principle*, Matematicheskie Trudy [In Russian] **5** (1985), 27–44.

[Sch74] H. H. Schaefer, *Banach lattices and positive operators*, Die Grundlehren der mathematischen Wissenschaften, Band 215, Springer-Verlag, New York-Heidelberg, 1974.

[Shm14] P. Shmerkin, *On the exceptional set for absolute continuity of Bernoulli convolutions*, Geom. Funct. Anal. **24** (2014), no. 3, 946–958.

[Sol22] B. Solomyak, *Fourier decay for homogeneous self-affine measures*, J. Fractal Geom. **9** (2022), no. 1-2, 193–206.

[Sol95] ———, *On the random series* $\sum \pm\lambda^n$ *(an Erdos problem)*, Ann. of Math. (2) **142** (1995), no. 3, 611–625.

[SS16a] P. Shmerkin and B. Solomyak, *Absolute continuity of complex bernoulli convolutions*, Math. Proc. Cambridge Philos. Soc. **161** (2016), 435–453.

[SS16b] ———, *Absolute continuity of self-similar measures, their projections and convolutions*, Trans. Amer. Math. Soc. **368** (2016), no. 7, 5125–5151.

[SS23] B. Solomyak and A. Spiewak, *Absolute continuity of self-similar measures on the plane* (2023). `https://arxiv.org/abs/2301.10620`, to appear at Indiana Univ. Math. J.

[SSS18] S. Saglietti, P. Shmerkin, and B. Solomyak, *Absolute continuity of non-homogeneous self-similar measures*, Adv. Math. **335** (2018), 60–110. MR3836658

[Sto65] C. Stone, *A local limit theorem for nonlattice multi-dimensional distribution functions*, Ann. Math. Stat. **36** (1965), 546–551.

[Str24] L. Streck, *On absolute continuity and maximal garsia entropy for self-similar measures with algebraic contraction ratio* (2024). `https://arxiv.org/abs/2303.07785`.

[SX03] B. Solomyak and H. Xu, *On the 'mandelbrot set' for a pair of linear maps and complex bernoulli convolutions*, Nonlinearity **16** (2003), no. 5, 1733–1749.

[Tit72] J. Tits, *Free subgroups in linear groups*, J. Algebra **20** (1972), 250–270.

[Tol00] F. Tolli, *A Berry-Esseen theorem on semisimple Lie groups*, Ann. Inst. H. Poincaré Probab. Statist. **36** (2000), no. 3, 275–290.

[Var15] P. P. Varjú, *Random walks in euclidean space*, Ann. of Math. (2) **181** (2015), 243–301.

[Var19a] P. P. Varjú, *Absolute continuity of Bernoulli convolutions for algebraic parameters*, J. Amer. Math. Soc. **32** (2019), no. 2, 351–397.

[Var19b] P. P. Varjú, *On the dimension of Bernoulli convolutions for all transcendental parameters*, Ann. of Math. (2) **189** (2019), no. 3, 1001–1011.

[Vig21] J. P. Vigneaux, *Entropy under disintegrations*, Geometric science of information, 2021, pp. 340–349.

[VY22] P. P. Varjú and H. Yu, *Fourier decay of self-similar measures and self-similar sets of uniqueness*, Anal. PDE **15** (2022), no. 3, 843–858.

[Wal88] N. R. Wallach, *Real reductive groups. I*, Pure and Applied Mathematics, vol. 132, Academic Press, Inc., Boston, MA, 1988.

[War72] G. Warner, *Harmonic analysis on semi-simple Lie groups. I*, Die Grundlehren der mathematischen Wissenschaften, Band 188, Springer-Verlag, New York-Heidelberg, 1972.

[Win35] A. Wintner, *On convergent poisson convolutions.*, Amer. J. Math. **57** (1935), 827—838.