RANDOM WALKS, SPECTRAL GAPS, AND KHINTCHINE'S THEOREM ON FRACTALS

OSAMA KHALIL AND MANUEL LUETHI

ABSTRACT. This work addresses problems on simultaneous Diophantine approximation on fractals, motivated by a long standing problem of Mahler regarding Cantor's middle 1/3 set. We obtain the first instances where a complete analogue of Khintchine's Theorem holds for fractal measures. Our results apply to fractals which are self-similar by a system of rational similarities of $\mathbb{R}^d$ (for any $d \geq 1$) and have sufficiently small Hausdorff co-dimension. A concrete example of such measures in the context of Mahler's problem is the Hausdorff measure on the "middle 1/5 Cantor set"; i.e. the set of numbers whose base 5 expansions miss a single digit.

The key new ingredient is an effective equidistribution theorem for certain fractal measures on the homogeneous space $\mathcal{L}_{d+1}$ of unimodular lattices; a result of independent interest. The latter is established via a new technique involving the construction of S-arithmetic operators possessing a spectral gap and encoding the arithmetic structure of the maps generating the fractal. As a consequence of our methods, we show that spherical averages of certain random walks naturally associated to the fractal measures effectively equidistribute on $\mathcal{L}_{d+1}$.

1. INTRODUCTION

Given a function $\psi : \mathbb{N} \to \mathbb{R}_+$, we say that $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$ is ψ-**approximable** if for infinitely many $q \in \mathbb{N}$, we have

$$\max_{1 \leqslant i \leqslant d} |qx_i - p_i| < \psi(q), \quad \text{for some } \mathbf{p} = (p_1, \ldots, p_d) \in \mathbb{Z}^d. \tag{1.1}$$

We denote by $W(\psi) \subseteq \mathbb{R}^d$ the set of ψ-approximable vectors. A much studied example is the function $\psi_\tau(q) := q^{-\tau-1/d}$. The union over $\tau > 0$ of $W(\psi_\tau)$ comprises the set of Very Well Approximable (VWA) vectors. Khintchine's Theorem [Kh26] in its modern formulation asserts that if ψ is non-increasing and Leb is the Lebesgue measure on $\mathbb{R}^d$, then

$$\text{Leb}(W(\psi)) = \begin{cases} 0 & \text{if } \sum_{q \geqslant 1} \psi^d(q) < \infty, \\ \text{FULL} & \text{if } \sum_{q \geqslant 1} \psi^d(q) = \infty. \end{cases} \tag{1.2}$$

Motivated by the study of approximation of real numbers by algebraic numbers, Mahler conjectured in 1932 that the Veronese curve $\mathcal{V}$ parametrized by $x \mapsto (x, x^2, \ldots, x^d)$ is *extremal*, i.e., the Lebesgue measure on $\mathcal{V}$ assigns zero mass to the set of VWA vectors. This conjecture set forth a field of study aiming to understand the prevalence of ψ-approximable vectors with respect to measures which may be singular with respect to Lebesgue on $\mathbb{R}^d$, e.g., volume measures on manifolds and affine subspaces as well as fractal measures. In particular, in 1984, Mahler asked:

Question 1.1 (Section 2, [Ma84])**.** *How close can irrational elements of Cantor's set be approximated by rational numbers not in Cantor's set?*

Mahler's conjecture on the extremality of $\mathcal{V}$ was settled by Sprindžuk who conjectured that every non-degenerate submanifold (e.g., analytic submanifolds which are not contained in a proper affine subspace) of $\mathbb{R}^d$ is extremal. This latter conjecture was resolved by Kleinbock and Margulis in [KM98]. Subsequently, Weiss proved in [We01] that a large class of measures on $\mathbb{R}$, which includes the canonical measure on Cantor's set, is extremal. This result, along with the work of [KM98], was generalized in [KLW04] to show the extremality of a wide class of measures on $\mathbb{R}^d$ which the

authors called *friendly measures*. This class includes volume measures on non-degenerate manifolds as well as measures which are self-similar by an irreducible iterated function system (IFS for short); cf. Section 2 for the corresponding definitions. Naturally, the authors posed the following problem.

Question 1.2 (Question 10.1, [KLW04])**.** *Suppose μ is a friendly measure on $\mathbb{R}^d$. Does the analogue of* (1.2) *hold with* Leb *replaced by μ?*

Questions 1.1 and 1.2 have generated intense activity in recent years. We refer the reader to [BDV07, Be12, VV06, BVVZ17, Hu20] and references therein for recent breakthroughs on this problem for non-degenerate manifolds. The question for fractals remains wide open however.

Finally, we refer the reader to Question 1.5 and Theorem D below for another motivation for our work beyond Diophantine approximation, regarding the equidistribution of "spherical averages" of random walks on homogeneous spaces.

1.1. **Statement of the results.**
The goal of this article is to obtain a complete analogue of Khintchine's Theorem for certain self-similar fractal measures, Theorem A. This answers Question 1.2 in the affirmative for those measures. The class of measures for which our results hold includes Hausdorff measures on missing digit Cantor sets of sufficiently small Hausdorff co-dimension, Theorem C. This provides the first evidence that a similar result is to be expected for Cantor's set in the setting of Question 1.1. The main ingredient in our proof is a new effective equidistribution result for fractal measures on the space of unimodular lattices, Theorem B.

We introduce notation to be used throughout the rest of the introduction. We refer the reader to Section 2 for detailed definitions.

Let $\mathcal{F} = \{f_i : i \in \Lambda\}$ be an IFS consisting of a finite collection of contractive similarities of $\mathbb{R}^d$ with respect to some inner product. We say that $\mathcal{F}$ is **rational** if $f_i = \rho_i O_i + b_i$ with

$$0 < \rho_i < 1, \qquad \rho_i \in \mathbb{Q}, \qquad O_i \in \mathrm{SO}_d(\mathbb{R}) \cap \mathrm{SL}_d(\mathbb{Q}), \qquad b_i \in \mathbb{Q}^d, \qquad \forall i \in \Lambda, \qquad (1.3)$$

and $\mathrm{SO}_d(\mathbb{R})$ is the special orthogonal group of some inner product on $\mathbb{R}^d$.

Let $(\lambda_i)_{i \in \Lambda}$ be a probability vector, i.e., $\lambda_i > 0$ for all i and $\sum \lambda_i = 1$. Denote by μ the unique self-similar probability measure on $\mathbb{R}^d$ determined by $\mathcal{F}$ and λ and by s the Hausdorff dimension (denoted $\dim_H$) of the attractor of $\mathcal{F}$. Set

$$\rho_{\min} := \min\{\rho_i : i \in \Lambda\}, \qquad \rho_{\max} := \max\{\rho_i : i \in \Lambda\}.$$

We define $\lambda_{\min}$ and $\lambda_{\max}$ similarly. The following is the main result of this article.

Theorem A. *There exists an explicit $\epsilon_0 > 0$, depending only on d, such that the following holds. Suppose $\mathcal{F}$ is a rational IFS satisfying the open set condition. Assume further that*

$$\left(\frac{d \log \rho_{\min}}{\log \lambda_{\max}} - 1\right) \frac{\log \lambda_{\min}}{s \log \rho_{\max}} < \epsilon_0. \qquad (1.4)$$

Let $\psi : \mathbb{N} \to \mathbb{R}_+$ be any non-increasing function. Then,

$$\mu(W(\psi)) = \begin{cases} 0 & \text{if } \sum_{q \geqslant 1} \psi^d(q) < \infty, \\ 1 & \text{if } \sum_{q \geqslant 1} \psi^d(q) = \infty. \end{cases} \qquad (1.5)$$

An explicit choice of ϵ_0 is stated in (6.8). We note that even the convergence part of Theorem A is new.

Remark 1.3. In the special case of equal contraction ratios and λ being the uniform probability vector, Condition (1.4) amounts to requiring that the Hausdorff dimension of the fractal is sufficiently close to that of the ambient Euclidean space.

The key ingredient in the proof of Theorem A is the following dynamical theorem. Let $G = \mathrm{SL}_{d+1}(\mathbb{R})$ and $\Gamma = \mathrm{SL}_{d+1}(\mathbb{Z})$. For $t > 0$ and $\mathbf{x} \in \mathbb{R}^d$, define the following elements of G:

$$g_t = \begin{pmatrix} e^{t/d}\mathrm{Id}_d & \mathbf{0} \\ \mathbf{0} & e^{-t} \end{pmatrix}, \qquad u(\mathbf{x}) = \begin{pmatrix} \mathrm{Id}_d & \mathbf{x} \\ \mathbf{0} & 1 \end{pmatrix}, \tag{1.6}$$

where Id_d denotes the $d \times d$ identity matrix.

Theorem B. *Under the hypotheses of Theorem A, there exists $\delta > 0$ and $\ell \in \mathbb{N}$ such that for every $\varphi \in \mathrm{C}_c^\ell(G/\Gamma)$, and $t > 0$, the following holds:*

$$\int \varphi\left(g_t u(\mathbf{x})\Gamma\right)\, d\mu(\mathbf{x}) = \int \varphi\, dm_{G/\Gamma} + O\big(\mathcal{S}_{\infty,\ell}(\varphi)e^{-\delta t}\big),$$

where $m_{G/\Gamma}$ is the unique G-invariant measure on G/Γ and the Sobolev norm $\mathcal{S}_{\infty,\ell}$ is defined in (3.2).

It is worth noting that Theorem B is new even in its qualitative form. The reader is referred to Theorem 6.1 for a more precise statement. In the special case of missing digit Cantor sets, we obtain the following sharper statement.

Theorem C. *Theorems A and B hold when μ is the Hausdorff measure on a missing digit Cantor set $\mathcal{K}$ in a prime base, cf. Definition A.1, satisfying*

$$\dim_H(\mathcal{K}) > 0.839. \tag{1.7}$$

In particular, these results hold for $\mathcal{K}$ the set of numbers whose base 5 expansions miss a single digit.

Remark 1.4. For comparison, we note that Theorem A implies that Khintchine's Theorem holds for the Hausdorff measure on a missing digit set $\mathcal{K}$ whenever $\dim_H(\mathcal{K}) \geq 0.9992$.

1.2. **Random walks.** To demonstrate the scope of the methods introduced in this article, we establish the equidistribution of certain random walks on G/Γ, motivated by the breakthroughs of Benoist-Quint and Bourgain-Furman-Lindenstrauss-Mozes. Given $\mathcal{F}$ and λ as above, let

$$c_i = \rho_i^{-1/(d+1)}, \qquad s_i = \begin{pmatrix} c_iO_i & c_ib_i \\ \mathbf{0} & c_i^{-d} \end{pmatrix}, \qquad \nu = \sum_{i \in \Lambda} \lambda_i \delta_{s_i}, \tag{1.8}$$

where we regard ν as a probability measure on G. Then, the IFS induces a random walk on G/Γ with law ν. The methods used to establish Theorem B yield the following:

Theorem D. *There exists $\varrho_0 > 0$ such that the following holds. Suppose $\mathcal{F}$ is a missing digit IFS on $\mathbb{R}$, cf. Definition A.1, with attractor $\mathcal{K}$ and λ is the uniform probability measure on Λ. Assume that $\dim_H(\mathcal{K}) > 1 - \varrho_0$. Let ν be as in (1.8). Then,*

$$\nu^{*n} * \delta_e \to m_{G/\Gamma},$$

where $e \in G/\Gamma$ is the coset of the identity element in G. The speed of convergence is exponential in n for sufficiently smooth functions on G/Γ.

In fact, our methods apply to certain more general basepoints and IFS; cf. Eq. (7.1) and Remark 7.2. This result is motivated by the following well-known open problem.

Question 1.5 (Question 3, [BQ12])**.** *Suppose ν is a compactly supported measure on G and let Γ_ν denote the subsemigroup generated by its support. Assume that the Zariski closure of Γ_ν is semisimple without compact factors. Let $x \in G/\Gamma$. As $n \to \infty$, do the measures $\nu^{*n} * \delta_x$ converge towards the unique homogeneous probability measure on $\overline{\Gamma_\nu \cdot x}$?*

Note that the measures ν in (1.8) do not fall under Question 1.5. In the setting of Question 1.5, Benoist and Quint showed that $\frac{1}{N}\sum_1^N \nu^{*n} * \delta_x$ converge to the expected limit [BQ11, BQ13, BQ13]. Question 1.5 was previously resolved in [BFLM11] in the setting of random walks on the torus (under certain additional hypotheses). In that result, a rate of equidistribution was also provided.

Generalizing the work of Benoist-Quint, Simmons and Weiss [SW19] studied random walks with law ν as in (1.8) arising from IFS and proved that for all $x \in G/\Gamma$,

$$\frac{1}{N} \sum_{n=1}^{N} \nu^{*n} * \delta_x \to m_{G/\Gamma}. \tag{1.9}$$

Motivated by Question 1.5, it is natural to ask whether the Cesaro averaging can be removed in (1.9). Theorem D solves this problem in the cases considered.

1.3. **Generalizations.** It is worth noting that we do not require the probability vector to be rational nor the contraction ratios be equal. Our results also apply to fractals in all dimensions.

In order to keep the article to a manageable length, we have not included the most general statements that can be obtained with our methods. We describe below several generalizations of our results we hope to address in forthcoming work.

(1) Jarník-Besicovitch Theorem: H. Yu recently proved, using Fourier analytic techniques, that the set of VWA numbers have full dimension inside missing digit Cantor sets $\mathcal{K} \subset \mathbb{R}$ whose Hausdorff dimensions are close to 1 [Yu20]. Recalling the notation at the beginning of the introduction, his methods also yield the exact value of the dimension of $W(\psi_\tau) \cap \mathcal{K}$ when τ is sufficiently small. We expect our results can be used to provide an alternative proof of those facts. We hope to provide a more complete Hausdorff measure theory of the intersections $W(\psi) \cap \mathcal{K}$ and to address more general fractals $\mathcal{K}$ in future work.

(2) Gallagher's Theorem: Our proof of Theorem B extends with minor modifications to more general diagonal flows, which commute with the IFS in a suitable sense. In ongoing work, we are studying the application of such extensions to obtain generalizations of Gallagher's Theorem in multiplicative Diophantine approximation [Ga62] for fractal measures. The reader is referred to [CY19] for related recent developments.

(3) Khintchine-Groshev Theorem: Our proof of Theorem B also extends to cover rational self-similar measures on the space of systems of linear forms under a suitable analogue of Hypothesis (1.4). In particular, the convergence case of Theorem A holds for those measures as well. We leave the divergence case for those measures to future work.

1.4. **Related work.** The best known result towards Question 1.2 for fractal measures was obtained by Pollington and Velani in [PV05] (cf. [We02] for the case of Cantor sets on the line). They show that for an absolutely friendly measure μ which is (C, α)-absolutely decaying for some constants $C, \alpha > 0$ (cf. [PV05] and (2.11) for definitions), the following holds for non-increasing functions ψ:

$$\sum_{q \geq 1} q^{\frac{\alpha}{d}-1} \psi^\alpha(q) < \infty \implies \mu(W(\psi)) = 0. \tag{1.10}$$

Simmons-Weiss [SW19] recently proved that for μ-almost every $\mathbf{x}$ the measures

$$\frac{1}{T} \int_0^T \delta_{g_t u(\mathbf{x})\Gamma} \, \mathrm{d}t$$

converge towards the Haar measure under the minimal necessary hypotheses, i.e., that the IFS is irreducible. For comparison with Theorem B, this implies that the averaged measures

$$\frac{1}{T} \int_0^T \int \delta_{g_t u(\mathbf{x})\Gamma} \, \mathrm{d}\mu(\mathbf{x}) \, \mathrm{d}t \tag{1.11}$$

converge to the Haar measure. At the heart of their proof is a generalization of the measure classification results of Benoist and Quint [BQ11]. A weaker equidistribution result for the measures in (1.11) was obtained earlier in [EFS11] in the case where μ is a $\times n$-invariant measure on the circle or an invariant measure for a hyperbolic toral automorphism. Their method relies on the

measure classification results of Lindenstrauss [Li06]. In particular, the methods in both instances are inherently non-effective. Moreover, the additional averaging in t is necessary in both cases.

In [CY19], Chow and Yang showed that the translates by certain diagonal flows (in the interior of the standard positive Weyl chamber) of the Lebesgue measure on a straight line (with Diophantine parameters) become effectively equidistributed on $\mathrm{SL}_3(\mathbb{R})/\mathrm{SL}_3(\mathbb{Z})$. They applied this result to obtain refinements of Gallagher's Theorem [Ga62]. Their methods are of a completely different nature to ours and build on an effective equidistribution theorem by Strömbergsson on the space of affine lattices in $\mathbb{R}^2$.

The reader is also referred to [LSV07, AB19] for related results on intrinsic Diophantine approximation on fractals and to [ACY20] for results on the approximation of points on Cantor's middle thirds set by dyadic rationals.

1.5. **Outline of the proof.** We first describe the deduction of Theorem A from the equidistribution theorem[1]. In view of the connection between ψ-approximability and cusp excursions, we show in Section 9 that the convergence part of Theorem A holds for any (not necessarily self-similar) measure μ satisfying the conclusion of Theorem B. In Section 12, we show that the divergence part holds for any (not necessarily rational) self-similar measure with the open set condition satisfying a stronger form of Theorem B; namely Corollary 6.4.

The main difficulties in deducing the divergence part from Corollary 6.4 arise from the fact that our error terms are in terms of Sobolev norms of L^∞-type and are not uniform over certain basepoints associated to the fractal, even as they vary in a fixed compact set in G/Γ. This complicates the independence arguments, especially when the approximation function ψ has slowly diverging partial sums. We remark that these issues do not arise in [KM99]; for instance the error terms in *loc. cit.* are in terms of L^2-Sobolev norms. In particular, for smooth approximations of shrinking cusp neighborhoods, these L^2-Sobolev norms provide additional decay in the error terms due to the decay of the measure of the support of such functions.

To overcome these issues, we prove a converse to the classical Borel-Cantelli lemma, Proposition 11.1, which is adapted to our problem. This result requires as input two quasi-independence estimates of different nature. To explain the idea, suppose E_n is a sequence of events in a probability space (Ω, μ) such that $\sum \mu(E_n) = \infty$. The first such independence estimate roughly takes the form

$$\mu(E_n \cap E_m) \leq C\mu(E_n)\mu(E_m) + O(e^{-\delta|n-m|}), \tag{1.12}$$

whenever

$$n \geq C_* m \qquad \text{or} \qquad m \leq n \leq (1 + \varepsilon_*)m \tag{1.13}$$

for some constants $C, C_* \geq 1$ and $0 < \delta, \varepsilon_* < 1$ and for all $n, m \in \mathbb{N}$. This estimate is most useful when $n \geq m$ are sufficiently separated; namely when $n \gtrsim m - \log \mu(E_m)$. To account for close-by pairs of n and m, we use an estimate roughly of the form

$$\mu(E_n \cap E_m) \ll \mu(E_m)\mu(E_n)^\varepsilon, \tag{1.14}$$

for some $\varepsilon > 0$ and for all $n > m$. Additionally, this result requires control over the failure of monotonicity of the measures $\mu(E_n)$; cf. Proposition 11.1(4).

The first estimate (1.12) is deduced in Proposition 10.1 from a stronger version of the equidistribution theorem, Corollary 6.4, which holds for more general basepoints besides the identity cosets. Proposition 10.1 can be viewed as a substitute for mixing of the flow g_t. The reason we cannot establish (1.12) for all pairs n and m is explained below.

The second estimate (1.14) is proved in Proposition 12.11. The proof relies on the simplex lemma and self-similarity and is similar in spirit to some proofs of the classical Khintchine Theorem. It also

---

[1]Here and throughout, we refer to Theorem 6.1, which is the more precise form of Theorem B, as the equidistribution theorem.

requires our equidistribution theorem. The proof of the divergence part of Theorem A is completed in Section 12.

The key ingredient in verifying all the above estimates is the equidistribution theorem. To explain its proof, define

$$\mu_t = \int \delta_{g_t u(\mathbf{x})\Gamma} \, d\mu(\mathbf{x}), \qquad t \in \mathbb{R}. \tag{1.15}$$

The key idea is to construct a random walk which *commutes* with the flow g_t and which leaves each μ_t stationary. To do this, we lift the problem to a suitable S-arithmetic cover G_S/Γ_S of G/Γ. The set of finite primes used to define the cover comes from the rational parameters of the IFS. The random walk is supported on a finite set $\{\gamma_i : i \in \Lambda\} \subset G_S$, defined in (4.2), and satisfying the following key identity

$$\gamma_i \cdot g_t u(\mathbf{x})\Gamma_S = g_t u(f_i(\mathbf{x}))\Gamma_S, \qquad (i \in \Lambda),$$

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

Denote by $\mathcal{P}$ the averaging operator associated to this random walk and the probability vector λ. A key step in the proof is to show that $\mathcal{P}$ has a spectral gap, in a suitable sense, as an operator on $L^2(G_S/\Gamma_S)$. This is Proposition 4.3 where we give an explicit estimate on the size of the spectral gap. The essential observation used in the proof is that the subsemigroup generated by the support of the random walk remains at a uniformly bounded distance from the lattice Γ_S. This allows us to use the fact that the matrix coefficients of Γ_S, acting on $L^2(G_S/\Gamma_S)$, belong to $\ell^p(\Gamma_S)$ to deduce that $\mathcal{P}$ has a spectral gap. At some stage in the proof, we use the fact that this subsemigroup is free and hence our proof is valid for all IFS without exact overlaps, cf. (2.9).

In Section 4.3, we find an explicit choice of p so that the matrix coefficients belong to $\ell^p(\Gamma_S)$. In Appendix A, we give sharper estimates on the spectral gap of $\mathcal{P}$ in the special case of missing digit Cantor sets, using more elementary techniques; cf. Proposition A.3. It is desirable to generalize these methods to more general fractals.

The proof of Theorem B is carried out in Section 6. Using the fact that all the maps in the IFS are contractions, we approximate μ—with an explicit bound on the approximation error—by an absolutely continuous probability measure on $\mathbb{R}^d$ (Theorem 2.3). As the approximation happens along the unstable manifold of g_t, the approximation errors blow up with t. A crucial Cauchy-Schwarz step allows us to bring the spectral properties of $\mathcal{P}$ into the argument, cf. (6.22). When our assumption in (1.4) holds, the spectral gap of $\mathcal{P}$ is stronger than the approximation error allowing us to obtain the result in this case.

Over the course of the proof, we apply effective equidistribution of translates of absolutely continuous measures by g_t to functions of the form $\mathcal{P}^n(\varphi)$, where φ is the lift of a smooth function from G/Γ to G_S/Γ_S. As $n \to \infty$, the functions $\mathcal{P}^n(\varphi)$ become less smooth in the S-arithmetic sense, i.e., they correspond to functions which live on suitable (congruence) covers of G/Γ. In Proposition 5.1, we verify the needed equidistribution statements, with uniform error rates and *uniform implied constants* over the family of covers in question.

Using suitable conjugation of the operator $\mathcal{P}$, along with the above arguments, allows us to prove an equidistribution statement of translates of fractal measures anchored at certain rational basepoints (cf. (6.6)) in G/Γ which are naturally associated with the IFS. This more general statement is crucial for the independence result in Proposition (10.1) which is a key ingredient in the divergence part of Theorem A. However, the index of the congruence cover on which we apply Proposition 5.1 depends on the conjugation of $\mathcal{P}$; i.e. on the basepoint. As the error terms in Proposition 5.1 depend on the index of the congruence cover, this causes non-uniformity of our error terms for Corollary 6.4 over basepoints in a compact set in G/Γ. This is the reason we are not able to prove the estimate (1.12) for all pairs n and m.

Theorem D is proved in Section 7 using a similar strategy to the proof of the equidistribution theorem. In this case, we appeal to the equidistribution of rational points instead of absolutely continuous measures (Proposition 7.1); cf. Remark 7.2 for a discussion of the reason for this difference. In Appendix A, we provide the needed modifications on the proofs to obtain Theorem C.

2. Preliminaries

In this section, we recall several facts regarding self similar measures.

2.1. **Iterated Function Systems.** We fix an arbitrary inner product on $\mathbb{R}^d$ and denote by $\mathrm{SO}_d(\mathbb{R})$ the subgroup of $\mathrm{SL}_d(\mathbb{R})$ which preserves it. A finite collection of maps $\mathcal{F} = \{f_i : i \in \Lambda\}$ on $\mathbb{R}^d$ is said to be an **iterated function system** (IFS for short) if each f_i is a contractive similarity of $\mathbb{R}^d$ relative to our chosen inner product, i.e., f_i has the form

$$f_i = \rho_i O_i + b_i,$$

where $0 < \rho_i < 1$, $O_i \in \mathrm{SO}_d(\mathbb{R})$, and $b_i \in \mathbb{R}^d$. Let

$$\Lambda^* = \bigcup_{k \geq 0} \Lambda^k,$$

where we use the convention $\Lambda^0 = \{\emptyset\}$ and $f_\emptyset$ is the identity mapping. In particular, $\rho_\emptyset = 1$, $b_\emptyset = \mathbf{0}$, and $O_\emptyset$ is the identity element of $\mathrm{SO}_d(\mathbb{R})$.

Using a fixed point theorem, it is shown in [Hu81] that there exists a unique compact set $\mathcal{K} \subset \mathbb{R}^d$ which is invariant by $\mathcal{F}$ in the sense that

$$\mathcal{K} = \bigcup_{i \in \Lambda} f_i(\mathcal{K}). \tag{2.1}$$

We refer to the set $\mathcal{K}$ as the **attractor** of $\mathcal{F}$. Given $\omega = (\omega_i) \in \Lambda^k$, we let

$$f_\omega = f_{\omega_1} \circ \cdots \circ f_{\omega_k}.$$

Denoting by $\mathbf{0}$ the origin in $\mathbb{R}^d$, the maps f_ω take the form $\rho_\omega O_\omega + b_\omega$, where

$$\rho_\omega = \prod_{i=0}^{k} \rho_{\omega_i}, \quad O_\omega = O_{\omega_1} \cdots O_{\omega_k}, \quad b_\omega = f_\omega(\mathbf{0}). \tag{2.2}$$

Hence, by induction, for all $k \in \mathbb{N}$, we have

$$\mathcal{K} = \bigcup_{\omega \in \Lambda^k} \mathcal{K}_\omega, \qquad \mathcal{K}_\omega := f_\omega(\mathcal{K}).$$

Definition 2.1. Given an IFS $\mathcal{F} = \{f_i : i \in \Lambda\}$ and probability vector λ on Λ, define the operator P_λ on $\mathrm{C}(\mathbb{R}^d)$ as follows. For all $\varphi \in \mathrm{C}(\mathbb{R}^d)$ let

$$P_\lambda(\varphi)(\mathbf{x}) = \sum_{i \in \Lambda} \lambda_i \varphi(f_i(\mathbf{x})) \quad (\mathbf{x} \in \mathbb{R}^d).$$

The dual operator on measures, also denoted P_λ, is defined similarly by

$$P_\lambda(\nu) = \sum_{i \in \Lambda} \lambda_i (f_i)_* \nu,$$

for all Borel measures ν on $\mathbb{R}^d$. We say a compactly supported probability measure μ on $\mathbb{R}^d$ is **self-similar** if

$$\mu = P_\lambda(\mu). \tag{2.3}$$

Given a probability vector λ, induction applied to (2.3) shows that

$$\mu = P_\lambda^k(\mu) = \sum_{\omega \in \Lambda^k} \lambda_\omega (f_\omega)_* \mu, \tag{2.4}$$

whenever μ is a self-similar measure, where $\lambda_\omega = \prod_{i=1}^k \lambda_i$. We also note that given a probability measure ν and $\psi \in \mathrm{C}_c(\mathbb{R}^d)$, we have

$$\int \psi(\mathbf{x}) \, \mathrm{d} P_\lambda^k(\nu)(\mathbf{x}) = \sum_{\omega \in \Lambda^k} \lambda_\omega \psi(f_\omega(\mathbf{x})) \, \mathrm{d}\nu(\mathbf{x}). \tag{2.5}$$

We say a map $P : X \to X$ of a metric space X is a contraction with ratio $r \in (0,1)$ if for every $x_1, x_2 \in X$, $d(P(x_1), P(x_2)) \le r d(x_1, x_2)$. We need the following elementary lemma.

Lemma 2.2. *Suppose $P : X \to X$ is a contraction mapping of a metric space with contraction ratio $r \in (0,1)$. Let $y, x_0 \in X$ be such that $P^{n_k}(y) \to x_0$ along a sequence $n_k \in \mathbb{N}$. Then, $P^n(y) \to x_0$, x_0 is the unique fixed point of P, and for all $n \in \mathbb{N}$ we have*

$$d(x_0, P^n(y)) \le r^n \frac{d(y, P(y))}{1 - r}. \tag{2.6}$$

Proof. For each $m \in \mathbb{N}$, let $y_m = P^m(y)$ and let $y_0 = y$. Then, for every $m > n$,

$$\begin{aligned}
d(y_m, y_n) &\le d(y_m, y_{m-1}) + d(y_{m-1}, y_{m-2}) + \cdots + d(y_{n+1}, y_n) \\
&\le r^{m-1} d(y_1, y_0) + r^{m-2} d(y_1, y_0) + \cdots + r^n d(y_1, y_0) \\
&= r^n d(y_1, y_0) \sum_{k=0}^{m-n-1} r^k \le r^n d(y_1, y_0) \sum_{k=0}^{\infty} r^k = r^n \frac{d(y_1, y_0)}{1 - r}.
\end{aligned}$$

Since $y_{n_k} \to x_0$, then (2.6) follows by continuity of the distance function. It follows that $y_n \to x_0$. In particular, $x_0 = \lim_n y_n = P(\lim_n y_{n-1}) = P(x_0)$ and hence x_0 is fixed by P. Uniqueness follows since P is a contraction. $\qquad\square$

Given a measure space (X, μ) and $\varphi \in \mathrm{L}^1(\mu)$, we use the notation

$$\mu(\varphi) := \int \varphi \, \mathrm{d}\mu.$$

Given a Lipschitz function φ on $\mathbb{R}^d$, we let $\mathrm{Lip}(\varphi)$ denote its Lipschitz constant. We use $\|\varphi\|_\infty$ to denote the sup-norm of φ. Denote by $\mathrm{Prob}_c(\mathbb{R}^d)$ the space of compactly supported probability measures on $\mathbb{R}^d$. Following [Hu81], we define the L-metric on $\mathrm{Prob}_c(\mathbb{R}^d)$ as follows:

$$L(\mu, \nu) = \sup |\mu(\varphi) - \nu(\varphi)|,$$

for all $\mu, \nu \in \mathrm{Prob}_c(\mathbb{R}^d)$, where the supremum is taken over all Lipschitz functions φ with Lipschitz constant at most 1.

Theorem 2.3 (Theorem 4.4.1 (ii), [Hu81]). *Let $\mathcal{F}$ be an IFS and λ be a probability vector. There exists a unique probability measure $\mu \in \mathrm{Prob}_c(\mathbb{R}^d)$ satisfying* (2.3). *Moreover, for every Lipschitz function φ on $\mathbb{R}^d$, $\nu \in \mathrm{Prob}_c(\mathbb{R}^d)$, and all $m \in \mathbb{N}$, we have*

$$|\mu(\varphi) - P_\lambda^m(\nu)(\varphi)| \ll_\nu r^m \mathrm{Lip}(\varphi),$$

where r is given by

$$r = \sum_{i \in \Lambda} \lambda_i \rho_i.$$

Proof. First, we show that P_λ is a contraction in the L-metric on $\mathrm{Prob}_c(\mathbb{R}^d)$ with ratio r. Indeed, we have for all $\nu_1, \nu_2 \in \mathrm{Prob}_c(\mathbb{R}^d)$,

$$|P_\lambda(\nu_1)(\varphi) - P_\lambda(\nu_2)(\varphi)| \leqslant \sum_{i \in \Lambda} \lambda_i \rho_i \left| \nu_1(\rho_i^{-1}\varphi \circ f_i) - \nu_2(\rho_i^{-1}\varphi \circ f_i) \right|$$
$$\leqslant r L(\nu_1, \nu_2) \mathrm{Lip}(\varphi),$$

where on the second line, we used the fact that $\mathrm{Lip}(\varphi \circ f_i) \leq \rho_i \mathrm{Lip}(\varphi)$.

In order to apply Lemma 2.2, it remains to check that $P_\lambda^n(\nu)$ converges along a subsequence to some $\mu \in \mathrm{Prob}_c(\mathbb{R}^d)$ in the L-metric. Since all the maps in $\mathcal{F}$ are contractions, there is a closed ball $B \subset \mathbb{R}^d$ around $\mathbf{0}$ containing the supports of the measures $P_\lambda^n(\nu)$ for all n. By compactness of the weak-$*$ topology on $\mathrm{Prob}(B)$, we can find $\mu \in \mathrm{Prob}(B) \subset \mathrm{Prob}_c(\mathbb{R}^d)$ and a sequence n_k such that $P_\lambda^{n_k}(\nu) \to \mu$ in the weak-$*$ topology.

We note that this implies convergence in the L-metric on $\mathrm{Prob}_c(\mathbb{R}^d)$. Indeed, for every $\varphi \in \mathrm{C}(B)$ with $\mathrm{Lip}(\varphi) \leq 1$, and for $\nu_1, \nu_2 \in \mathrm{Prob}(B)$, we have

$$|\nu_1(\varphi) - \nu_2(\varphi)| = |\nu_1(\varphi - \varphi(\mathbf{0})) - \nu_2(\varphi - \varphi(\mathbf{0}))|.$$

Hence, it suffices to check convergence in the L-metric where the supremum is restricted to functions which vanish at $\mathbf{0}$. For such functions, we have $\|\varphi\|_\infty \leq |\varphi(\mathbf{0})| + R = R$, where R is the radius of B since $\mathrm{Lip}(\varphi) \leq 1$. This set of function is pre-compact in the uniform norm on $\mathrm{C}(B)$ in view of the Arzelà-Ascoli theorem. One then verifies that this implies convergence in the L-metric. Applying Lemma 2.2 completes the proof. □

Finally, we record the following simple lemma concerning averages of multiplicative cocycles.

Lemma 2.4. *Suppose a tuple $(\tau_i)_{i \in \Lambda} \in \mathbb{R}^{|\Lambda|}$ and a probability vector λ on Λ are given. For $\omega = (\omega_1, \dots, \omega_k) \in \Lambda^k$, let $\tau_\omega = \tau_{\omega_1} \cdots \tau_{\omega_k}$. Then, for all $n \in \mathbb{N}$,*

$$\sum_{\omega \in \Lambda^n} \lambda_\omega \tau_\omega = \left(\sum_{i \in \Lambda} \lambda_i \tau_i \right)^n.$$

Proof. Let $a_n = \sum_{\omega \in \Lambda^n} \lambda_\omega \tau_\omega$. Given two words α and ω, let $\alpha\omega$ denote the word obtained by concatenating ω to the end of α. We then note that $\lambda_{\alpha\omega} = \lambda_\alpha \lambda_\omega$ and $\tau_{\alpha\omega} = \tau_\alpha \tau_\omega$. It follows that, for all $m, n \in \mathbb{N}$, we have

$$a_{m+n} = \sum_{\omega \in \Lambda^{m+n}} \lambda_\omega \tau_\omega = \sum_{\alpha \in \Lambda^m} \sum_{\omega \in \Lambda^n} \lambda_{\alpha\omega} \tau_{\alpha\omega} \sum_{\alpha \in \Lambda^m} \sum_{\omega \in \Lambda^n} \lambda_\alpha \lambda_\omega \tau_\alpha \tau_\omega = a_m a_n.$$

□

2.2. **Notational Convention.** For the remainder of this article, we will denote by $(\mathcal{F}, \lambda)$ a tuple consisting of an IFS and a probability vector λ. With such a tuple, we implicitly fix a choice of an inner product relative to which $\mathcal{F}$ consists of similarity maps and denote the resulting norm simply by $\|\cdot\|$. We denote by $\mathrm{SO}_d(\mathbb{R})$ the orthogonal group of this inner product. We extend this norm to $\mathbb{R}^{d+1}$ as follows:
$$\|(x_1, \ldots, x_{d+1})\| = \max\{\|(x_1, \ldots, x_d)\|, |x_{d+1}|\}.$$
We will denote by $\mathcal{K} = \mathcal{K}_{\mathcal{F}}$ the attractor, which is completely determined by $\mathcal{F}$, and by $\mu = \mu_{(\mathcal{F}, \lambda)}$ the unique self-similar measure for the operator P_λ provided by Theorem 2.3.

2.3. **The open set condition and null overlaps.** In general, serious problems in the analysis occur if the images of the fractal by distinct maps of the IFS overlap. We recall several conditions under which such overlap becomes negligible in a precise sense.

We say $\mathcal{F}$ satisfies the **open set condition** if there exists an open set $U \subset \mathbb{R}^d$ such that for all $i \neq j \in \Lambda$,
$$f_i(U) \subseteq U, \qquad f_i(U) \cap f_j(U) = \emptyset. \tag{2.7}$$
We say that a self-similar measure μ has **null overlaps** if for all $i \neq j \in \Lambda$,
$$\mu(\mathcal{K}_i \cap \mathcal{K}_j) = 0. \tag{2.8}$$
Note that by self similarity, the null overlaps property implies that $\mu(\mathcal{K}_\alpha \cap \mathcal{K}_\omega) = 0$ for all $\alpha \neq \omega \in \Lambda^n$ and for all $n \in \mathbb{N}$.

Lemma 2.5 (Lemma 3.3, [Gr95]). *Suppose $\mathcal{F}$ is an IFS satisfying the open set condition, λ is a probability vector and μ is a self-similar measure for $(\mathcal{F}, \lambda)$. Then, μ has null overlaps.*

For a Borel set A and a Borel measure μ, we denote by $\mu|_A$ the restriction of μ to A. That is for every Borel set B, $\mu|_A(B) = \mu(B \cap A)$. The following consequence of null overlaps will be useful for us.

Lemma 2.6. *Suppose $\mathcal{F}$ is an IFS, λ is a probability vector and μ is a self-similar measure for $(\mathcal{F}, \lambda)$ having null overlaps. Then, for every $k \in \mathbb{N}$ and all $\omega \in \Lambda^k$,*
$$\mu|_{f_\omega(\mathcal{K})} = \lambda_\omega (f_\omega)_* \mu$$
and, in particular, $\mu(\mathcal{K}_\omega) = \lambda_\omega$.

Proof. The main assertion follows from the null overlaps property and equation (2.4).   □

The weakest notion of separation for an IFS is having **no exact overlaps**, where $\mathcal{F}$ has no exact overlaps if
$$\alpha \neq \omega \implies f_\alpha \neq f_\omega, \qquad \forall \alpha, \omega \in \Lambda^*. \tag{2.9}$$
Having no exact overlaps turns out to be sufficient for spectral gap arguments, cf. § 4.

2.4. **A Zero-Full law for fractals.** The following lemma will be used in the proof of the divergence part of Theorem A to upgrade positivity of the measure of the set of ψ-approximable vectors to the statement that they have full measure. Note that Cassel's Zero-Full law [Ca50] does not apply directly to fractal measures; cf. Remark 12.4.

Lemma 2.7. *Let $(\mathcal{F}, \lambda)$ as in Section 2.2 and suppose that the associated self-similar measure μ has null overlaps. Let $A \subseteq \mathbb{R}^d$ be a Borel measurable set and suppose that there exists $c > 0$ such that for every $\alpha \in \Lambda^*$, we have that $\mu(A \cap \mathcal{K}_\alpha) \geq c\mu(\mathcal{K}_\alpha)$. Then, $\mu(A) = 1$.*

Proof. Let $\Sigma = \Lambda^{\mathbb{N}}$ and endow it with the product topology induced from the discrete topology on Λ. Denote by $\lambda^{\mathbb{N}}$ the product measure on Σ defined by λ. Let $\pi : \Lambda^{\mathbb{N}} \to \mathcal{K}$ be the coding map defined by $\pi(\alpha) = \lim_{n \to \infty}(f_{\alpha_1} \circ \cdots \circ f_{\alpha_n})(\mathbf{0})$. Then, π is continuous and $\mu = \pi_* \lambda^{\mathbb{N}}$; cf. [Hu81, Theorems 3.1(3) and 4.4(4)]. In particular, $\mu(A) = \lambda^{\mathbb{N}}(\pi^{-1}(A))$. For $\alpha \in \Lambda^*$, let Σ_α denote the cylinder set determined by α. For all x, let $\Sigma(x, k)$ denote the unique cylinder set of length k containing x.

By our null overlaps hypothesis, the symmetric difference between $\pi^{-1}(\mathcal{K}_\alpha)$ and Σ_α has measure 0. Hence, our hypothesis shows that

$$\lambda^{\mathbb{N}}(\pi^{-1}(A) \cap \Sigma_\alpha) \geq c\lambda^{\mathbb{N}}(\Sigma_\alpha). \tag{2.10}$$

On the other hand, if $B = \Sigma \setminus \pi^{-1}(A)$, then by a version of the Lebesgue density theorem for $(\Sigma, \lambda^{\mathbb{N}})$, cf. Lemma A.6, for almost every $x \in B$, we have

$$\lim_{k \to \infty} \frac{\lambda^{\mathbb{N}}(\Sigma(x,k) \cap B)}{\lambda^{\mathbb{N}}(\Sigma(x,k))} = 1.$$

It follows that if B has positive measure, then we can find some cylinder $\Sigma(x,k)$ so that $\lambda^{\mathbb{N}}(\Sigma(x,k) \cap B) > (1-c)\lambda^{\mathbb{N}}(\Sigma(x,k))$. This contradicts (2.10). $\qquad\square$

2.5. **Absolutely decaying measures.** For a subset $\mathcal{L} \subset \mathbb{R}^d$ and $\varepsilon > 0$, we denote

$$\mathcal{L}^{(\varepsilon)} = \left\{ \mathbf{x} \in \mathbb{R}^d : d(\mathbf{x}, \mathcal{L}) < \varepsilon \right\}.$$

The IFS $\mathcal{F}$ is said to be **irreducible** if no proper affine subspace of $\mathbb{R}^d$ is invariant under all the maps in $\mathcal{F}$.

The following absolute decay property was established in [KLW04] for a wide class of natural measures on $\mathbb{R}^d$. We recall their result in our setting.

Proposition 2.8 (Theorem 2.3, [KLW04]). *Let $(\mathcal{F}, \lambda)$ as in Section 2.2 and suppose that $\mathcal{F}$ is irreducible and satisfies the open set condition. Let μ be the associated self-similar measure. Then, there exist constants $C \geq 1$ and $\gamma > 0$ such that for every word $\alpha \in \Lambda^*$, for every proper affine subspace $\mathcal{L}$, and for every $\varepsilon > 0$, we have*

$$\mu(\mathcal{K}_\alpha \cap \mathcal{L}^{(\varepsilon)}) \leqslant C \left(\frac{\varepsilon}{\rho_\alpha} \right)^\gamma \mu(\mathcal{K}_\alpha). \tag{2.11}$$

Proof. Let $s = \dim_H \mathcal{K}$. The proposition follows from the argument establishing the (C, α)-absolutely decaying property for μ in [KLW04, Theorem 2.3], under the open set condition, in the case μ is the Hausdorff measure supported on the fractal, i.e., μ is the self-similar measure for the probability vector $(\rho_i^s)_{i \in \Lambda}$. The proof adapts verbatim to general self-similar measures. Indeed, the only place in the proof in [KLW04, Section 8] where the fact that μ is the Hausdorff measure is used is to show that for any Borel set $A \subset \mathbb{R}^d$,

$$\mu(f_\omega(A \cap \mathcal{K})) = \rho_\omega^s \mu(A), \tag{2.12}$$

for all $\omega \in \Lambda^*$; cf. proof of (8.5) and (8.6) in *loc. cit.* For general probability vectors λ, (2.12) holds with λ_ω in place of ρ_ω^s by Lemmas 2.5 and 2.6. $\qquad\square$

3. Congruence quotients of PGL

The goal of this section is to introduce notation for the S-arithmetic groups and homogeneous spaces we work with in our proofs. We also discuss several technical connectedness issues which arise naturally in equidistribution problems and which stem from the failure of Strong Approximation for PGL. We also recall bounds on matrix coefficients of PGL which we use in later sections.

3.1. **S-arithmetic Setup.** We let PGL_{d+1} denote the automorphism group of the algebra Mat_{d+1}. Throughout this article we fix the $\mathbb{Q}$-structure (and the integral structures) to be the one induced by the faithful representation $\mathrm{PGL}_{d+1} \hookrightarrow \mathrm{GL}_{(d+1)^2}$ induced by the standard basis of Mat_{d+1}.

In what follows, we write $\mathbf{G} = \mathrm{PGL}_{d+1}$. Given a finite set S of places of $\mathbb{Q}$ possibly containing the archimedean place, we denote

$$S_{\mathrm{f}} = S \setminus \{\infty\}, \quad \mathbb{Z}_\infty = \mathbb{R}, \quad \mathbb{Q}_S = \prod_{v \in S} \mathbb{Q}_v, \quad \mathbb{Z}_{\mathrm{f}} = \prod_{p \in S_{\mathrm{f}}} \mathbb{Z}_p, \quad \mathbb{Z}[S^{-1}] = \mathbb{Z}[\tfrac{1}{p}; p \in S_{\mathrm{f}}].$$

Accordingly, we define the following:

$$\mathbf{G}_S = \mathbf{G}(\mathbb{Q}_S), \quad \mathbf{G}_\infty = \mathbf{G}(\mathbb{R}), \quad \mathbf{G}_{\mathrm{f}} = \mathbf{G}(\mathbb{Q}_{S_{\mathrm{f}}}),$$
$$\Gamma_S = \mathbf{G}(\mathbb{Z}[S^{-1}]) = \mathbf{G}(\mathbb{Q}_S) \cap \mathrm{GL}_{(d+1)^2}(\mathbb{Z}[S^{-1}]),$$
$$K_{\mathrm{f}} = \mathbf{G}(\mathbb{Z}_{\mathrm{f}}) = \mathbf{G}(\mathbb{Q}_{S_{\mathrm{f}}}) \cap \mathrm{GL}_{(d+1)^2}(\mathbb{Z}_{\mathrm{f}}).$$

If $\infty \in S$, then Γ_S embedded diagonally in $\mathbf{G}_S$ is a lattice and we denote

$$X_S = \mathbf{G}_S/\Gamma_S.$$

We also use the following notation:

$$M_v = \begin{cases} \mathrm{PO}_{d+1}(\mathbb{R}) & \text{if } v = \infty, \\ \mathbf{G}(\mathbb{Z}_v) & \text{if } v \text{ is finite.} \end{cases}$$

We will denote by $\Phi : \mathrm{GL}_{d+1} \to \mathrm{GL}_{(d+1)^2}$ the rational representation defined by mapping an element $x \in \mathrm{GL}_{d+1}$ to the automorphism $\Phi(x)$ of Mat_{d+1} defined by

$$\Phi(x)w = xwx^{-1} \quad (w \in \mathrm{Mat}_{d+1}). \tag{3.1}$$

By the Skolem-Noether theorem, for every field k we have $\mathbf{G}(k) = \Phi(\mathrm{GL}_{d+1}(k))$. We also note that $\ker \Phi = k^\times$ and therefore $\mathbf{G}(k) \cong \mathrm{GL}_{d+1}(k)/k^\times$.

We will denote by $\|\cdot\|_v$ the operator norm on $\mathrm{GL}(\mathrm{Mat}_{d+1}(\mathbb{Q}_v))$ given by the sup-norm with respect to the standard basis if v is finite, and induced by the inner product

$$\langle L, M \rangle = \mathrm{tr}(^t LM) \quad \left(L, M \in \mathrm{GL}(\mathrm{Mat}_{d+1}(\mathbb{Q}_v))\right)$$

if v is infinite. Abusing notation, we denote

$$\|g\|_v = \|\mathrm{Ad}_g\|_v \quad (g \in \mathbf{G}(\mathbb{Q}_v)). \tag{3.2}$$

We note that for all $g \in \mathbf{G}(\mathbb{Q}_v)$ and for all $k_1, k_2 \in M_v$ we have (cf. Corollary C.2)

$$\|k_1 g k_2\|_v = \|g\|_v. \tag{3.3}$$

Lemma 3.1. *Let* $\mathrm{SL}_{d+1}(\mathbb{R})$ *act on* $\mathbf{G}_\infty/\mathbf{G}(\mathbb{Z})$ *via the representation* Φ. *This action is transitive and the stabilizer of the identity coset is* $\mathrm{SL}_{d+1}(\mathbb{Z})$. *In particular, we have*

$$\mathrm{SL}_{d+1}(\mathbb{R})\big/\mathrm{SL}_{d+1}(\mathbb{Z}) \cong \mathbf{G}_\infty\big/\mathbf{G}(\mathbb{Z})$$

as $\mathrm{SL}_{d+1}(\mathbb{R})$-*spaces.*

Remark 3.2. The above statement is a well-known fact in the theory of lattices in $\mathbb{R}^{d+1}$ once we know that $\mathbf{G}(\mathbb{Z})$ agrees with the image of $\mathrm{GL}_{d+1}(\mathbb{Z})$ under projection with respect to the center, which is not apriori clear.

Proof. We first show that $\mathbf{G}(\mathbb{Z}_p) = \Phi(\mathrm{GL}_{d+1}(\mathbb{Z}_p))$ for every prime p. In fact, for every $x \in \mathrm{GL}_{d+1}(\mathbb{Q}_p)$ satisfying $\Phi(x) \in \mathbf{G}(\mathbb{Z}_p)$ there is some $n_p \in \mathbb{Z}$ such that $p^{n_p}x \in \mathrm{GL}_{d+1}(\mathbb{Z}_p)$. Indeed, choose $n_p \in \mathbb{Z}$ such that $\|p^{n_p}x\|_p = 1$. Using the KAK-decomposition on $\mathrm{GL}_{d+1}(\mathbb{Q}_p)$, let $p^{n_p}x = k_1 a k_2$, where $k_1, k_2 \in \mathrm{GL}_{d+1}(\mathbb{Z}_p)$ and a is a diagonal matrix whose entries are decreasingly ordered with respect to the p-adic valuation. As $\mathrm{GL}_{d+1}(\mathbb{Z}_p)$ preserves $\mathrm{Mat}_{d+1}(\mathbb{Z}_p)$ under left- and right-multiplication, we find that that

$$a\mathrm{Mat}_{d+1}(\mathbb{Z}_p)a^{-1} = \mathrm{Mat}_{d+1}(\mathbb{Z}_p).$$

In particular, the automorphism of $\mathrm{Mat}_{d+1}(\mathbb{Z}_p)$ given by conjugation by a maps the standard basis to a $\mathbb{Z}_p$-basis of $\mathrm{Mat}_{d+1}(\mathbb{Z}_p)$. This implies that $a \in \mathrm{GL}_{d+1}(\mathbb{Z}_p)$ and in particular $p^{n_p}x \in \mathrm{GL}_{d+1}(\mathbb{Z}_p)$. As $\Phi(\lambda x) = \Phi(x)$ for all $\lambda \in \mathbb{Q}_p^\times$, the initial claim then follows.

We next deduce that $\mathbf{G}(\mathbb{Z}) = \Phi(\mathrm{GL}_{d+1}(\mathbb{Z}))$. To this end let $g \in \mathbf{G}(\mathbb{Z})$ and let $x \in \mathrm{GL}_{d+1}(\mathbb{Q})$ such that $g = \Phi(x)$. As of the previous argument, we know that for all primes p there is $n_p \in \mathbb{Z}$ such that $p^{n_p} x \in \mathrm{GL}_{d+1}(\mathbb{Z}_p)$ and we note that $n_p = 0$ for almost all p. Therefore the number

$$c = \prod_p p^{n_p} \in \mathbb{Q}^\times$$

is well-defined and $cx \in \mathrm{GL}_{d+1}(\mathbb{Z}_p)$ for all primes p. As

$$\mathrm{GL}_{d+1}(\mathbb{Z}) = \bigcap_p \big(\mathrm{GL}_{d+1}(\mathbb{Q}) \cap \mathrm{GL}_{d+1}(\mathbb{Z}_p) \big),$$

we get $g = \Phi(cx) \in \Phi(\mathrm{GL}_{d+1}(\mathbb{Z}))$.

Two lattices $\Delta_1, \Delta_2 \subseteq \mathbb{R}^{d+1}$ are homothetic if there exists a scalar $c \in \mathbb{R}$ such that $\Delta_1 = c\Delta_2$. The natural action of $\mathrm{GL}_{d+1}(\mathbb{R})$ on lattices in $\mathbb{R}^{d+1}$ induces a transitive action of $\mathbf{G}_\infty$. Moreover, as every homothety class admits a unique unimodular representative, this induces a transitive action of $\mathbf{G}_\infty$ on $\mathrm{SL}_{d+1}(\mathbb{R})/\mathrm{SL}_{d+1}(\mathbb{Z})$ and the stabilizer of the identity coset is $\mathbf{G}(\mathbb{Z})$. The identification is clearly $\mathrm{SL}_{d+1}(\mathbb{R})$-equivariant and hence the claim follows. $\qquad\square$

Given $N \in \mathbb{N}$ we also denote by $\Gamma(N) \leq \mathbf{G}(\mathbb{Z})$ the principal congruence subgroup of level N, i.e., the kernel of the homomorphism defined by coordinate-wise reduction mod N. For the remainder of the article, we use the notation

$$X_\infty(N) := \mathbf{G}_\infty/\Gamma(N).$$

We abuse notation and let $\Gamma(1) = \mathbf{G}(\mathbb{Z})$. A subgroup $\Delta \leq \Gamma(1)$ is called a congruence subgroup if it contains a principal congruence subgroup.

We note that for the chosen representation of $\mathbf{G}$, the group has class number one (cf. Proposition B.8 and Corollary B.10), i.e.,

$$(\mathbf{G}_\infty \times K_{\mathrm{f}})\Gamma_S = \mathbf{G}_S. \tag{3.4}$$

Denote by $\mathbb{N}_0$ the set of non-negative integers. Given $\mathbf{m} = (m_p)_{p \in S_{\mathrm{f}}} \in \mathbb{N}_0^{S_{\mathrm{f}}}$, let $S_{\mathrm{f}}^{\mathbf{m}} = \prod_{p \in S_{\mathrm{f}}} p^{m_p}$ and denote by $K_{\mathrm{f}}[S_{\mathrm{f}}^{\mathbf{m}}]$ the kernel of the canonical homomorphism $\mathbf{G}(\mathbb{Z}_{\mathrm{f}}) \to \prod_{p \in S_{\mathrm{f}}} \mathbf{G}(\mathbb{Z}/p^{m_p}\mathbb{Z})$. It will be useful to abuse notation and let $K_{\mathrm{f}}[1] = K_{\mathrm{f}}$. Note that $\Gamma(S_{\mathrm{f}}^{\mathbf{m}}) = \Gamma_S \cap K_{\mathrm{f}}[S^{\mathbf{m}}]$.

As $\mathbf{G}$ has class number one, there is $r(\mathbf{m}) \in \mathbb{N}$ such that

$$K_{\mathrm{f}}[S_{\mathrm{f}}^{\mathbf{m}}]\backslash \mathbf{G}_S / \Gamma_S \cong \bigsqcup_{i=1}^{r(\mathbf{m})} X_\infty(S_{\mathrm{f}}^{\mathbf{m}}), \tag{3.5}$$

as $\mathbf{G}_\infty$-spaces; cf. Proposition B.9.

Let $\mathbf{G}_S^+$ denote the image of $\mathrm{SL}_{d+1}(\mathbb{Q}_S)$ in $\mathbf{G}_S$ and similarly $\mathbf{G}_v^+$ denotes the image of $\mathrm{SL}_{d+1}(\mathbb{Q}_v)$, for each $v \in S$. These are normal subgroups of finite index in $\mathbf{G}_S$ and $\mathbf{G}_v$ respectively. The number $r(\mathbf{m})$ of cosets in (3.5) is bounded by the index of $\mathbf{G}_S^+$ inside $\mathbf{G}_S$, cf. Proposition B.13.

In what follows, we will call the copies of $X_\infty(S_{\mathrm{f}}^{\mathbf{m}})$ the components of $K_{\mathrm{f}}[S_{\mathrm{f}}^{\mathbf{m}}]\backslash \mathbf{G}_S/\Gamma_S$. We want to point out that these components are in general not connected. In fact, the connected components are precisely the $\mathbf{G}_\infty^+$ orbits.

3.2. **Sobolev norms.** We introduce a family of Sobolev norms to be used throughout the article. We fix a basis $\mathcal{G}_d$ of the Lie algebra of $\mathbf{G}_\infty$ and, given $\ell \in \mathbb{N}_0$, we denote by $\mathfrak{X}_\ell$ the set of all monomials in the elements of $\mathcal{G}_d$ of degree at most ℓ. The elements of $\mathfrak{X}_\ell$ act as differential operators on $\mathrm{C}^\infty(\mathbf{G}_\infty/\Gamma)$, for any lattice Γ in $\mathbf{G}_\infty$. Given $\varphi \in \mathrm{C}^\infty(\mathbf{G}_\infty/\Gamma)$, we denote by $\mathcal{S}_{p,\ell}$ the L^p-Sobolev norm of degree ℓ defined by

$$\mathcal{S}_{p,\ell}(\varphi) = \sum_{\mathcal{D} \in \mathfrak{X}_\ell} \|\mathcal{D}\varphi\|_p, \tag{3.6}$$

where the L^p-norm is defined with respect to the $\mathbf{G}_\infty$-invariant probability measure. Given $p \in [1, \infty]$, we let

$$B^\infty_{p,\ell}(\mathbf{G}_\infty/\Gamma) = \{\varphi \in C^\infty(\mathbf{G}_\infty/\Gamma) : \mathcal{S}_{p,\ell}(\varphi) < \infty\}. \tag{3.7}$$

Note that in case $\ell = 0$, $B^\infty_{p,\ell}(\mathbf{G}_\infty/\Gamma) = C^\infty(\mathbf{G}_\infty/\Gamma) \cap L^p(\mathbf{G}_\infty/\Gamma)$.

We write $C^\infty(X_S)$ for the space of functions on X_S which are invariant by some compact open subgroup of K_f and which are smooth along the $\mathbf{G}_\infty$ directions. Accordingly, we can extend (3.6) and (3.7) to $C^\infty(X_S)$, where the L^p-norms are defined with respect to the $\mathbf{G}_S$-invariant probability measure. We shall use the following basic estimates on these norms. Throughout the article, for a function φ on X_S and $g \in \mathbf{G}_S$, we use $\varphi \circ g$ to denote the composition of φ with the left multiplication action of g on X_S.

Lemma 3.3. *For all $\ell \in \mathbb{N}$, the following holds with implicit constants depending on ℓ:*

   (1) *For all $\varphi \in B^\infty_{\infty,\ell}(X_S)$ and $g \in \mathbf{G}_S$, $\mathcal{S}_{\infty,\ell}(\varphi \circ g) \ll \|\mathrm{Ad}(g)\|^\ell_\infty \mathcal{S}_{\infty,\ell}(\varphi)$, where $\|\mathrm{Ad}(g)\|_\infty$ denotes the operator norm of the adjoint action of the Archimedean component of g on the Lie algebra of $\mathbf{G}_\infty$.*

   (2) *For any $p \geq 1$ and $\varphi, \psi \in B^\infty_{p,\ell}(X_S)$, $\mathcal{S}_{p,\ell}(\varphi\psi) \ll_\ell \mathcal{S}_{2p,\ell}(\varphi)\mathcal{S}_{2p,\ell}(\psi)$. In particular, we have $\mathcal{S}_{\infty,\ell}(\varphi^2) \ll \mathcal{S}_{\infty,\ell}(\varphi)^2$.*

   (3) *For each non-zero $\mathcal{D}$ in the Lie algebra of $\mathbf{G}_\infty$, $\|\mathcal{D}\varphi\|_\infty \ll \|\mathcal{D}\| \mathcal{S}_{\infty,1}(\varphi)$, where $\|\mathcal{D}\|$ is taken with respect to any fixed choice of norm on the Lie algebra and the implicit constant depends on such choice.*

   (4) *For $\theta \in C^\infty_c(\mathbf{G}_\infty)$ and $\varphi \in L^\infty(X_S)$, $\mathcal{S}_{\infty,\ell}(\theta * \varphi) \ll m_{\mathbf{G}_\infty}(\mathrm{supp}(\theta))\mathcal{S}_{\infty,\ell}(\theta) \|\varphi\|_\infty$, where $\theta * \varphi$ denotes the convolution of the two functions and $m_{\mathbf{G}_\infty}$ is the Haar measure on $\mathbf{G}_\infty$.*

Proof. The independence of the estimate in Item (1) from the non-Archimedean component of g follows from the fact that the differential operators in the definition of our Sobolev norms commute with $\mathbf{G}_f$. Item (2) follows by Cauchy-Schwarz; cf. [Ve10, Lemma 8.1]. Item (3) follows from expressing $\mathcal{D}$ in terms of the basis $\mathcal{G}_d$ with coefficients bounded by $\|\mathcal{D}\|$. To show Item (4), it suffices to note that invariance of the Haar measure implies that $\mathcal{D}(\theta * \varphi) = (\mathcal{D}\theta) * \varphi$ for any differential operator $\mathcal{D}$ on $\mathbf{G}_\infty$. $\square$

Remark 3.4. Throughout the remainder of the article, we omit the dependence of implicit constants in our estimates on the order ℓ of the Sobolev norms in question for the sake of readability.

3.3. **Uniform bounds on matrix coefficients.** The main reference for the material in this section is [GMO08, Section 3]. Much of the general discussion there is a lot simpler in the setting where $\mathbf{G} = \mathrm{PGL}_{d+1}$, and we restrict ourselves to this case.

We let v denote a possibly infinite place of $\mathbb{Q}$, A_v the image of the diagonal subgroup of $\mathrm{GL}_{d+1}(\mathbb{Q}_v)$ in $\mathbf{G}$, and Σ^+_v the system of positive roots of $\mathbf{G}(\mathbb{Q}_v)$ relative to $A^+_v \leq A_v$, where

$$A^+_\infty = \Phi\left\{\begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_{d+1} \end{pmatrix} : a_i \in \mathbb{R}, a_1 \geq \cdots \geq a_{d+1} = 1\right\}$$

and for any finite rational prime p

$$A^+_p = \Phi\left\{\begin{pmatrix} p^{-n_1} & & \\ & \ddots & \\ & & p^{-n_{d+1}} \end{pmatrix} : n_i \in \mathbb{N}, n_1 \geq \cdots \geq n_{d+1} = 0\right\}.$$

In what follows we will usually identify elements in A^+_v with their representative in the sets on the right hand side of the above equations.

Recall that $M_v = \mathrm{PO}_{d+1}(\mathbb{R})$ if $v = \infty$ and $M_v = \mathbf{G}(\mathbb{Z}_v)$ if v is finite. Using the polar decomposition for infinite v or the p-adic analogue for finite v, we have $\mathbf{G}(\mathbb{Q}_v) = M_v A^+_v M_v$; cf. Appendix C.

Choose a maximal strongly orthogonal system $\mathcal{S}_v$ in Σ_v^+; cf. [Oh02]. Define a bi-M_v invariant function ξ_v on $\mathbf{G}(\mathbb{Q}_v)$ as follows: for every $g = m_1 a m_2 \in \mathbf{G}(\mathbb{Q}_v)$,

$$\xi_v(g) = \prod_{\alpha \in \mathcal{S}_v} \Xi_{\mathrm{PGL}_2(\mathbb{Q}_v)} \begin{pmatrix} \alpha(a) & 0 \\ 0 & 1 \end{pmatrix}, \tag{3.8}$$

where $\Xi_{\mathrm{PGL}_2(\mathbb{Q}_v)}$ is the Harish-Chandra function on $\mathrm{PGL}_2(\mathbb{Q}_v)$. We then define $\xi_{\mathbf{G}}$ on $\mathbf{G}_S$ by

$$\xi_{\mathbf{G}}(g) = \prod_{v \in S} \xi_v(g_v) \quad (g \in \mathbf{G}_S). \tag{3.9}$$

Denote by $\mathrm{L}_{00}^2(X_S)$ the closed subspace of $\mathrm{L}^2(X_S)$ which is orthogonal to the subspace spanned by $\mathbf{G}(\mathbb{Q}_v)^+$-invariant functions for all v. We note that the normality of $\mathbf{G}(\mathbb{Q}_v)^+$ implies that $\mathrm{L}_{00}^2(X_S)$ is a $\mathbf{G}_S$-invariant subspace. The following lemma will be of importance in Section 6.

Lemma 3.5. *Assume that $\varphi \in \mathrm{L}^2(X_S)$ is K_{f}-invariant and has mean zero. Then, φ belongs to $\mathrm{L}_{00}^2(X_S)^{K_{\mathrm{f}}}$.*

Proof. As φ is by assumption K_{f}-invariant, we only have to show that the average of φ over every $\mathbf{G}_S^+$-orbit in X_S vanishes. As $\mathbf{G}$ has class number one, cf. (3.4), we know that $K_{\mathrm{f}} \backslash X_S \cong X_\infty(1)$. Hence transitivity of the action of $\mathbf{G}_\infty^+$ on $X_\infty(1)$, cf. Lemma 3.1, implies that the average of φ over any $\mathbf{G}_\infty^+ \times K_{\mathrm{f}}$-orbit vanishes. As shown in the proof of Proposition B.13, the orbits of $\mathbf{G}_\infty^+ \times K_{\mathrm{f}}$ in X_S agree with the orbits of $\mathbf{G}_S^+ K_{\mathrm{f}}$ in X_S. Hence K_{f}-invariance of φ implies that the average of φ on every $\mathbf{G}_S^+$-orbit in X_S vanishes. $\qquad\square$

We recall the following bound on matrix coefficients given in [GMO08].

Theorem 3.6 (Theorems 3.10 and 3.20, [GMO08]). *For any compact open subgroup W_{f} of $\mathbf{G}_{\mathrm{f}}$, there exists $C_{W_{\mathrm{f}}} > 0$ such that the following holds: for all W_{f}-invariant M_∞-finite vectors $v, w \in \mathrm{L}_{00}^2(X_S)$ and for all $g \in \mathbf{G}_S$,*

$$|\langle gv, w \rangle| \leq C_{W_{\mathrm{f}}} \dim\langle M_\infty \cdot v \rangle \dim\langle M_\infty \cdot w \rangle \|v\|_2 \|w\|_2 \xi_{\mathbf{G}}(g)^{\epsilon(d)}, \tag{3.10}$$

where $\langle M_\infty \cdot v \rangle$ is the span of the orbit of v under M_∞ and

$$\epsilon(d) = \begin{cases} 1/2 & \text{if } d = 1, \\ 1 & \text{otherwise.} \end{cases} \tag{3.11}$$

Theorem 3.6 has the following corollary for smooth functions, which are not necessarily M_∞-finite.

Corollary 3.7. *Let $W_{\mathrm{f}} < \mathbf{G}_f$ be a compact open subgroup. For all $g \in \mathbf{G}_S$ and W_{f}-invariant smooth functions $\psi_1, \psi_2 \in \mathrm{L}_{00}^2(X_S)$ such that $\mathcal{S}_{2,\ell}(\psi_i) < \infty$, $i = 1, 2$, we have*

$$|\langle g\psi_1, \psi_2 \rangle| \ll_{W_{\mathrm{f}}} \mathcal{S}_{2,\ell}(\psi_1) \mathcal{S}_{2,\ell}(\psi_2) \xi_{\mathbf{G}}(g)^{\epsilon(d)},$$

where $\ell = \dim(M_\infty)$.

Proof. As ψ_1 and ψ_2 admit an orthogonal decomposition in terms of M_∞-finite vectors, the argument in [EMV09, Section 6.2.1] applies with minimal changes to deduce the corollary where one replaces Eq. (6.7) in *loc. cit.* with (3.10). Note that Eq. (6.7) in [EMV09] involves $\dim\langle M_\infty \cdot \rangle^{1/2}$. The argument goes through however and the resulting Sobolev norm is of order $\ell = \dim M_\infty$ (instead of $\ell = \lceil \dim M_\infty/2 \rceil$ as in [EMV09]).

$\qquad\square$

4. Spectral Gap of S-Arithmetic Random Walks

The goal of this section is to introduce certain S-arithmetic operators which are naturally associated to a rational IFS and which leave fractal measures invariant. Moreover, we prove that these operators possess a spectral gap and we provide an estimate on their operator norms, Proposition 4.3. The results of this section are key ingredients in our equidistribution theorems.

4.1. **The S-arithmetic random walk.** Given $(t, s) \in \mathbb{R}^\times \times \mathbb{Q}_{S_f}^\times$ and $(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^d \times \mathbb{Q}_{S_f}^d$ regarded as a pair of column vectors, we define the following elements of $\mathbf{G}_S$:

$$a(t, s) = \left(\begin{pmatrix} t\mathrm{Id}_d & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix}, \begin{pmatrix} s\mathrm{Id}_d & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix} \right), \qquad u(\mathbf{x}, \mathbf{y}) = \left(\begin{pmatrix} \mathrm{Id}_d & \mathbf{x} \\ \mathbf{0} & 1 \end{pmatrix}, \begin{pmatrix} \mathrm{Id}_d & \mathbf{y} \\ \mathbf{0} & 1 \end{pmatrix} \right),$$

where Id_d is the identity matrix in dimension d. We also use the following notation:

$$a(t) = \begin{pmatrix} t\mathrm{Id}_d & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix}, \qquad u(\mathbf{x}) = \begin{pmatrix} \mathrm{Id}_d & \mathbf{x} \\ \mathbf{0} & 1 \end{pmatrix}.$$

The above matrices are regarded as elements of $\mathbf{G}_\infty$ (resp. $\mathbf{G}_f$) whenever their parameters belong to $\mathbb{R}^\times$ (resp. $\mathbb{Q}_{S_f}^\times$). We will denote by Id the identity element in $\mathbf{G}$.

Throughout the remainder of this section, we fix a tuple $(\mathcal{F}, \lambda)$ as in Section 2.2 such that $\mathcal{F}$ is rational. We will define an analogue of the operator P_λ on the homogeneous space $\mathbf{G}_S / \Gamma_S$; cf. Definition 2.1. Using the rationality of $\mathcal{F}$, let $S(\mathcal{F})$ denote the smallest set of primes such that

$$\rho_i \in \mathbb{Z}[S(\mathcal{F})^{-1}]^\times, \quad b_i \in \mathbb{Z}[S(\mathcal{F})^{-1}]^d, \qquad O_i \in \mathrm{GL}(d, \mathbb{Z}[S(\mathcal{F})^{-1}]),$$

for all $i \in \Lambda$. In addition to the prime factors of the numerators and denominators of ρ_i, $S(\mathcal{F})$ need only contain primes in the denominators of the components of b_i and the entries in O_i, but not the numerators. We let $S = S(\mathcal{F}) \cup \{\infty\}$, omitting the dependence on $\mathcal{F}$.

Given $\alpha \in \Lambda^*$, let $k_\alpha \in \Gamma_S$ be given by

$$k_\alpha = \begin{pmatrix} O_\alpha & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix}, \tag{4.1}$$

where O_α is defined as in (2.2). The following elements of $\mathbf{G}_S$ will be central to the analysis to follow:

$$\gamma_\alpha = u(\mathbf{0}, -b_\alpha) a(\rho_\alpha, \rho_\alpha)(k_\alpha, k_\alpha). \tag{4.2}$$

The crucial property of γ_α is that for all $\mathbf{x} \in \mathbb{R}^d$ we have

$$\gamma_\alpha u(\mathbf{x}, \mathbf{0}) \gamma_\alpha^{-1} u(b_\alpha, \mathbf{0}) = u(f_\alpha(\mathbf{x}), \mathbf{0}). \tag{4.3}$$

We note that if $\alpha = (\alpha_i) \in \Lambda^k$, then the following equality holds:

$$\gamma_\alpha = \gamma_{\alpha_1} \cdots \gamma_{\alpha_k}. \tag{4.4}$$

Given a probability vector λ, we define an operator $\mathcal{P}_\lambda$ by

$$\mathcal{P}_\lambda(\varphi)(x) = \sum_{i \in \Lambda} \lambda_i \varphi(\gamma_i \cdot x), \tag{4.5}$$

for all $\varphi \in \mathrm{C}(X_S)$, and all $x \in X_S$. In fact, we shall need a more general family of operators which we now define. Given $\alpha, \omega \in \Lambda^*$, we set

$$\gamma_\omega^\alpha := (k_\alpha^{-1}, \mathrm{Id}) \gamma_\alpha \gamma_\omega \gamma_\alpha^{-1}(k_\alpha, \mathrm{Id}), \tag{4.6}$$

and define $\alpha \cdot \mathcal{P}_\lambda$ by

$$(\alpha \cdot \mathcal{P}_\lambda)(\varphi)(x) = \sum_{i \in \Lambda} \lambda_i \varphi(\gamma_i^\alpha \cdot x), \tag{4.7}$$

where $\alpha \cdot \mathcal{P}_\lambda = \mathcal{P}_\lambda$ if α is the empty word. For all $n \in \mathbb{N}$, we have

$$(\alpha \cdot \mathcal{P}_\lambda)^n(\varphi)(x) = \sum_{\omega \in \Lambda^n} \lambda_\omega \varphi(\gamma_\omega^\alpha \cdot x).$$

These operators are among the main objects of study in this article.

Remark 4.1. In Appendix A, we introduce a variant of the above operators which can be used to obtain sharper cutoffs in the case when the contraction ratios of the IFS are all equal.

We need the following lemma.

Lemma 4.2. *There exist positive constants A and L such that the following holds. Suppose φ is a function on $X_\infty(1)$. Then, for all $n \in \mathbb{N}$ and $\alpha \in \Lambda^*$, there exists $N \geq 1$ such that $(\alpha \cdot \mathcal{P}_\lambda)^n(\varphi)$ is invariant under $K_\mathrm{f}[N] \lhd K_\mathrm{f}$ and*

$$[\Gamma(1) : \Gamma(N)] \ll_{S,d} \rho_\alpha^{-A} \rho_{\min}^{-nL},$$

where $\rho_{\min} = \min\{\rho_i : i \in \Lambda\}$. When $\mathcal{F}$ is a missing digit IFS, cf. Definition A.1, we may take $A = 6$, $L = 3$.

Proof. For $g \in \mathbf{G}_S$, let $\mathrm{con}(g)(K_\mathrm{f}) = g^{-1}K_\mathrm{f}g \cap K_\mathrm{f}$. Note that $g^{-1}K_\mathrm{f}g \subseteq \mathbf{G}_\mathrm{f}$ is a compact open subgroup for every $g = (g_\infty, g_\mathrm{f}) \in \mathbf{G}_S$, as $g^{-1}K_\mathrm{f}g = g_\mathrm{f}^{-1}K_\mathrm{f}g_\mathrm{f}$. Denote by $W \subseteq K_\mathrm{f}$ the compact open subgroup given by

$$W = \bigcap_{\omega \in \Lambda^n} \mathrm{con}(\gamma_\omega^\alpha)(K_\mathrm{f}).$$

Since φ is K_f-invariant, the function $(\alpha \cdot \mathcal{P}_\lambda)^n(\varphi)$ is invariant under W. We will find N so that W contains $K_\mathrm{f}[N]$. In view of Proposition B.7, to bound the index $[\Gamma(1) : \Gamma(N)]$, it will suffice to bound the index $[K_\mathrm{f} : K_\mathrm{f}[N]]$. We will obtain bounds on the latter by reducing the question to bounds on the index of the kernel of reduction mod p^ν for $\mathrm{SL}_{d+1}(\mathbb{Z}_p)$ for appropriate ν. By [Ha06, Corollary 2.8], we have

$$|\mathrm{SL}_{d+1}(\mathbb{Z}/p^\nu\mathbb{Z})| = p^{(d^2+2d)\nu} \prod_{k=2}^{d+1} \left(1 - \frac{1}{p^k}\right). \tag{4.8}$$

Since $\mathrm{SL}_{d+1}(\mathbb{Z}/p^\nu\mathbb{Z})$ is generated by unipotents [HO89, Theorem 4.3.9], the reduction mod p^ν is surjective and therefore the right hand side of (4.8) gives the desired bound at place p.

Let $p \in S$ be a finite prime and recall that $M_p = \mathbf{G}(\mathbb{Z}_p)$. Let $g_p \in \mathbf{G}(\mathbb{Q}_p)$. By the KAK-decomposition (cf. Appendix C), we can write $g_p = k_1 a k_2$ for $k_1, k_2 \in M_p$ and $a \in A_p^+$. Then,

$$g_p M_p g_p^{-1} \cap M_p = k_1(aM_pa^{-1} \cap M_p)k_1^{-1}.$$

Let $\|g_p\|_p = p^\nu$ denote the operator norm of the adjoint action of g_p. We claim that $aM_pa^{-1} \cap M_p$ (and thus also $g_p M_p g_p^{-1} \cap M_p$) contains $M_p[\|g_p\|_p]$, i.e. the kernel of reduction mod $p^\nu = \|g_p\|_p$. Denote $D = (d + 1)^2$ and let $n_1 \geq \cdots \geq n_D = 0$ be so that $a = \mathrm{diag}(p^{-n_1}, \ldots, p^{-n_D})$. Note that $n_1 = \nu$ by Corollary C.2. Let $h \in M_p[p^\nu]$. The (i,j)-coordinate of $a^{-1}ha$ is given by multiplying the corresponding coordinate of h by $p^{n_i - n_j} \geq p^{-\nu}$. As the diagonal entries of h are preserved and the off-diagonal entries of h are multiples of p^ν, we obtain that $a^{-1}ha \in M_p$, i.e. $h \in aM_pa^{-1} \cap M_p$. It follows that

$$W = \prod_p (g_p M_p g_p^{-1} \cap M_p)$$

contains the subgroup $K_\mathrm{f}[N] = \prod_p M_p[|N|_p^{-1}]$, where

$$N = \prod_p \max\{\|\gamma_\alpha \gamma_\omega \gamma_\alpha^{-1}\|_p : \omega \in \Lambda^n\}$$

where $\|\gamma_\alpha \gamma_\omega \gamma_\alpha^{-1}\|_p$ denotes the operator norm for the adjoint action of the p-adic component. It thus remains to bound $[M_p : M_p[p^\nu]]$.

As shown in the proof of Lemma 3.1, we have $M_p = \Phi(\mathrm{GL}_{d+1}(\mathbb{Z}_p))$. Hence $M_p = FU_p$, where $U_p = \Phi(\mathrm{SL}_{d+1}(\mathbb{Z}_p))$ and F is the image of

$$T_p = \left\{\begin{pmatrix} x & 0 \\ 0 & \mathrm{Id}_d \end{pmatrix} : x \in \mathbb{Z}_p^\times\right\} < \mathrm{GL}_{d+1}(\mathbb{Z}_p), \tag{4.9}$$

under Φ. Given $r \in \mathbb{N}$, let $\mathbb{Z}_p^{\times r} \subseteq \mathbb{Z}_p^\times$ denote the set of elements admitting an r-th root in $\mathbb{Z}_p$. We have that $F/(F \cap U_p) \cong \mathbb{Z}_p^\times/\mathbb{Z}_p^{\times(d+1)}$ is finite with cardinality depending only on d and p. Letting

$U_p[p^\nu] = U_p \cap M_p[p^\nu]$, we obtain

$$[M_p : M_p[p^\nu]] \ll_{p,d} [U_p : U_p[p^\nu]].$$

As the kernel of the reduction mod p^ν in $\mathrm{SL}_{d+1}(\mathbb{Z}_p)$ is mapped into $U_p[p^\nu]$ under Φ, (4.8) yields

$$[U_p : U_p[p^\nu]] \leq p^{(d^2+2d)\nu} \prod_{k=2}^{d+1} \left(1 - \frac{1}{p^k}\right). \tag{4.10}$$

Hence, we conclude that

$$[K_{\mathrm{f}} : W] \leq [K_{\mathrm{f}} : K_{\mathrm{f}}[N]] \ll_{S,d} \prod_p \max\{\|\gamma_\alpha \gamma_\omega \gamma_\alpha^{-1}\|_p : \omega \in \Lambda^n\}^{d^2+2d}.$$

By Lemmas C.1 and C.5, the norm $\|\cdot\|_p$ is submultiplicative and satisfies $\|g_p\|_p = \|g_p^{-1}\|_p$. Hence, we find that

$$[K_{\mathrm{f}} : K_{\mathrm{f}}[N]] \ll_{S,d} \prod_p \|\gamma_\alpha\|_p^{2(d^2+2d)} \max\{\|\gamma_\omega\|_p : \omega \in \Lambda^n\}^{d^2+2d}.$$

Given $p \in S_{\mathrm{f}}$, we let $c_p = \max\{\|\gamma_i\|_p : i \in \Lambda\}$. Note that $c_p \geq 1$, and hence there are $L_p, A_p > 0$ such that

$$\rho_{\min}^{-L_p} = \rho_{\max}^{-A_p} = c_p.$$

We define

$$L = (d^2 + 2d) \sum_{p \in S_{\mathrm{f}}} L_p, \quad A = 2(d^2 + 2d) \sum_{p \in S_{\mathrm{f}}} A_p.$$

Using submultiplicativity and (4.4) and denoting by $|\alpha|$ the length of α, we get

$$[K_{\mathrm{f}} : K_{\mathrm{f}}[N]] \ll_{S,d} \prod_p c_p^{(d^2+2d)(2|\alpha|+n)} = \rho_{\max}^{-|\alpha|A} \rho_{\min}^{-nL} \leq \rho_\alpha^{-A} \rho_{\min}^{-nL}. \tag{4.11}$$

Recall that $\Gamma(N) = \Gamma_S \cap K_{\mathrm{f}}[N]$; cf. Proposition B.7. Applying the second isomorphism theorem with ambient group K_{f} and subgroups $\Gamma_S \cap K_{\mathrm{f}}$ and $K_{\mathrm{f}}[N] \lhd K_{\mathrm{f}}$, we get

$$[\Gamma(1) : \Gamma(N)] = [\Gamma_S \cap K_{\mathrm{f}} : \Gamma_S \cap K_{\mathrm{f}}[N]] \leq [K_{\mathrm{f}} : K_{\mathrm{f}}[N]].$$

Combining all of the above, one obtains

$$[\Gamma(1) : \Gamma(N)] \ll_{S,d} \rho_\alpha^{-2A} \rho_{\min}^{-nL}.$$

This completes the proof in the general case. For missing digit Cantor sets, assume that $q \in \mathbb{N}$ is at least two and $\Lambda \subseteq \{0, \ldots, q-1\}$ such that $\#\Lambda \geq 2$. Then $\rho_i = 1/q$, $k_i = \mathrm{Id}$, and $b_i = i/q$ for all $i \in \Lambda$. For $\alpha \in \Lambda^*$, using (4.4) and denoting by $\gamma_{\alpha,p}$ the p-adic component of γ_α, we have

$$\gamma_{\alpha,p} = \begin{pmatrix} \rho_\alpha & -b_\alpha \\ 0 & 1 \end{pmatrix}.$$

Let $x = \rho_\alpha$, $t = -b_\alpha$, and fix a prime divisor p of q. If $|t|_p \leq |x|_p$, we have

$$\gamma_{\alpha,p} = \begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & x^{-1}t \\ 0 & 1 \end{pmatrix}$$

and hence $\|\gamma_\alpha\|_p = |x|_p$. If $|x|_p < |t|_p$, then we note that

$$M_p \gamma_{\alpha,p} M_p = M_p \gamma_{\alpha,p} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} M_p$$

and hence it suffices to calculate the norm of

$$\gamma_{\alpha,p} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} x+t & t \\ 1 & 1 \end{pmatrix}.$$

As argued in the proof of Lemma C.3, there are unipotent elements $u_1, u_2 \in \mathrm{GL}_2(\mathbb{Z}_p)$ such that

$$u_1 \begin{pmatrix} x+t & t \\ 1 & 1 \end{pmatrix} u_2 = \begin{pmatrix} x+t & 0 \\ 0 & \frac{x}{x+t} \end{pmatrix}.$$

Therefore, we obtain that

$$\gamma_{\alpha,p} \in M_p \begin{pmatrix} |t|_p & 0 \\ 0 & |x|_p/|t|_p \end{pmatrix} M_p = M_p \begin{pmatrix} |t|_p^2/|x|_p & 0 \\ 0 & 1 \end{pmatrix} M_p.$$

Hence, $\|\gamma_\alpha\|_p \le \max\{|x|_p, |t|_p^2/|x|_p\}$. Now, note that $|x|_p = |q|_p^{-|\alpha|}$ and $|t|_p \le |q|_p^{-|\alpha|}$. It follows that

$$\prod_p \|\gamma_\alpha\|_p \le \prod_p |q|_p^{-|\alpha|} = q^{|\alpha|}.$$

Since $\rho_{\min} = \rho_{\max} = q^{-1}$ and $d = 1$ in this case, it follows by (4.11) that we may take $A = 6$ and $L = 3$.

$\square$

4.2. Spectral gap for the averaging operator.

We wish to estimate the operator norm of $\alpha \cdot \mathcal{P}_\lambda$ on the subspace of $\mathrm{L}_{00}^2(X_S)$ consisting of functions which are invariant by a compact open subgroup of $W_{\mathrm{f}} < \mathbf{G}_{S_{\mathrm{f}}}$.

The main difficulty is that the subgroup generated by $\{\gamma_\omega : \omega \in \Lambda^*\}$ is not discrete or free in general (although the subsemigroup is). This causes difficulty in controlling the separation of the γ_ω's. To explain the idea, let us focus on the case $\alpha = \emptyset$. Observe that for each word ω,

$$\tilde{\gamma}_\omega := u(-b_\omega, \mathbf{0})\gamma_\omega \in \Gamma_S. \tag{4.12}$$

In particular, the subgroup generated by $\{\tilde{\gamma}_\omega : \omega \in \Lambda^*\}$ is discrete since it is contained in the lattice. Moreover, all the elements of the form $u(-b_\omega, \mathbf{0})$ belong to a compact neighborhood of identity (recall the first coordinate corresponds to the Archimedean place). This allows us to relate the spectral properties of $\mathcal{P}_\lambda$ to an operator which is supported on the lattice Γ_S.

The following is one of the key results of this article. The reader is referred to Proposition A.3 for sharper bounds for missing digit Cantor sets.

Proposition 4.3. *Assume that $\mathcal{F}$ has no exact overlaps; cf. (2.9). Let $W_{\mathrm{f}} < \mathbf{G}_{\mathrm{f}}$ be a compact open subgroup and $\ell = \dim(M_\infty)$. For every $\alpha \in \Lambda^*$ and every $k \in \mathbb{N}$, there exists a finite set $\Delta_k \subset \Gamma_S$, which is determined by α and the IFS, such that the following holds. Suppose $p > 1$ is given and let θ be the Hölder conjugate of p. Let $q = 2\theta/(\theta+1)$. Then, for every smooth W_{f}-invariant $\varphi \in \mathrm{L}_{00}^2(X_S)$,*

$$\left\|(\alpha \cdot \mathcal{P}_\lambda)^k(\varphi)\right\|_{L^2}^2 \ll_{\mathcal{F}, W_{\mathrm{f}}} \mathcal{S}_{2,\ell}(\varphi)^2 \left(\sum_{i \in \Lambda} \lambda_i^q\right)^{2k/q} \left(\sum_{g \in \Delta_k} \xi_{\mathbf{G}}^{p\epsilon(d)}(g)\right)^{\frac{1}{p}},$$

where $\epsilon(d)$ is defined in (3.11). In particular, if $\xi_{\mathbf{G}}^{\epsilon(d)} \in \ell^p(\Gamma_S)$, then

$$\left\|(\alpha \cdot \mathcal{P}_\lambda)^k(\varphi)\right\|_{L^2}^2 \ll_{\mathcal{F}, W_{\mathrm{f}}} \mathcal{S}_{2,\ell}(\varphi)^2 \left(\sum_{i \in \Lambda} \lambda_i^q\right)^{2k/q} \left\|\xi_{\mathbf{G}}^{\epsilon(d)}\right\|_{\ell^p(\Gamma_S)}.$$

Proof. In order to simplify notation, let

$$\eta = \xi_{\mathbf{G}}^{\epsilon(d)}, \qquad \mathcal{S} = \mathcal{S}_{2,\ell}.$$

For $g \in \mathbf{G}_S$, we write $g\varphi$ to denote $\varphi \circ g$. Fix $\alpha \in \Lambda^*$. Given $\omega \in \Lambda^*$, let γ_ω^α be as in (4.6). A direct computation shows that the Archimedean component of γ_ω^α is $k_\omega a(\rho_\omega)$ while its non-Archimedean component is given by

$$u\big(-f_\alpha f_\omega f_\alpha^{-1}(\mathbf{0})\big) \cdot k_\alpha k_\omega k_\alpha^{-1} \cdot a(\rho_\omega). \tag{4.13}$$

Let $\tau_\omega^\alpha = u\big(-f_\alpha f_\omega f_\alpha^{-1}(\mathbf{0})\big)k_\alpha k_\omega k_\alpha^{-1} k_\omega^{-1}$, so that $\tilde\gamma_\omega^\alpha := (\tau_\omega^\alpha, \mathrm{Id})\gamma_\omega^\alpha \in \Gamma_S$. It follows by (4.4) that

$$\tilde\gamma_\omega^\alpha = \tilde\gamma_{\omega_1}^\alpha \cdots \tilde\gamma_{\omega_k}^\alpha \tag{4.14}$$

for all $\alpha, \omega = (\omega_1, \dots, \omega_k) \in \Lambda^*$.

By calculating the translation vector of the similarities $f_\alpha f_\omega f_\alpha^{-1}$, one sees that $u(-f_\alpha f_\omega f_\alpha^{-1}(\mathbf{0}))$ is uniformly bounded in $\mathbf{G}_\infty$, independently of α and ω. Moreover, at the Archimedean place the elements $k_\bullet$ are all contained in a compact subgroup. Hence, it follows that $\{(\tau_\sigma^\beta, \mathrm{Id}) : \beta, \sigma \in \Lambda^*\}$ is contained in a bounded set $\mathcal{O} \subseteq \mathbf{G}_S$, which depends only on the IFS $\mathcal{F}$. By Lemma 3.3, there exists $C_\mathcal{O} \geq 1$ such that for any smooth function ψ and for any $g \in \mathcal{O} \cup \mathcal{O}^{-1}$,

$$\mathcal{S}(g\psi) \leq C_\mathcal{O} \mathcal{S}(\psi). \tag{4.15}$$

Let $\overline\tau_\omega^\alpha = (\tau_\omega^\alpha, \mathrm{Id})^{-1}$. It follows from Corollary 3.7 that

$$\big\|(\alpha \cdot \mathcal{P}_\lambda)^k(\varphi)\big\|^2 = \sum_{u,\omega \in \Lambda^k} \lambda_u \lambda_\omega \langle \gamma_u^\alpha \varphi, \gamma_\omega^\alpha \varphi \rangle = \sum_{u,\omega \in \Lambda^k} \lambda_u \lambda_\omega \langle \tilde\gamma_u^\alpha(\overline\tau_u^\alpha \varphi), \tilde\gamma_\omega^\alpha(\overline\tau_\omega^\alpha \varphi) \rangle$$

$$\ll_{W_{\mathrm{f}}} \sum_{u,\omega \in \Lambda^k} \lambda_u \lambda_\omega \eta\big(\tilde\gamma_\omega^\alpha(\tilde\gamma_u^\alpha)^{-1}\big) \mathcal{S}\left(\overline\tau_u^\alpha \varphi\right) \mathcal{S}\left(\overline\tau_\omega^\alpha \varphi\right).$$

By (4.15), we get

$$\big\|(\alpha \cdot \mathcal{P}_\lambda)^k(\varphi)\big\|^2 \ll_{\mathcal{F},W_{\mathrm{f}}} \mathcal{S}(\varphi)^2 \sum_{u,\omega \in \Lambda^k} \lambda_u \lambda_\omega \eta\big(\tilde\gamma_\omega^\alpha(\tilde\gamma_u^\alpha)^{-1}\big). \tag{4.16}$$

Denote by ν_α the measure supported on $\{\tilde\gamma_i^\alpha : i \in \Lambda\}$ such that $\nu_\alpha(\tilde\gamma_i^\alpha) = \lambda_i$. In particular, ν_α is supported on Γ_S. Moreover, in view of (4.14), we have for every $g \in \mathbf{G}_S$ and $k \in \mathbb{N}$ that

$$\nu_\alpha^{*k}(g) = \sum_{\omega \in \Lambda^k : \tilde\gamma_\omega^\alpha = g} \lambda_\omega, \tag{4.17}$$

where ν_α^{*k} denotes the kth convolution power of ν_α. Denote by $\check\nu_\alpha$ the adjoint of ν_α, defined as the push-forward of ν_α under the map $g \mapsto g^{-1}$. For $k \in \mathbb{N}$, let $\Delta_k \subset \Gamma_S$ be the (finite) support of the measure $\nu_\alpha^{*k} * (\check\nu_\alpha)^{*k}$. With this notation, we can rewrite the upper bound in (4.16) as

$$\sum_{u,\omega \in \Lambda^k} \lambda_u \lambda_\omega \eta(\tilde\gamma_\omega^\alpha(\tilde\gamma_u^\alpha)^{-1}) = \int_{\mathbf{G}_S} \eta(g) \, \mathrm{d}\big(\nu_\alpha^{*k} * (\check\nu_\alpha)^{*k}\big)(g).$$

Recall that θ denotes the Hölder conjugate of p. By Hölder's inequality, we obtain:

$$\int_{\mathbf{G}_S} \eta(g) \, \mathrm{d}\big(\nu_\alpha^{*k} * (\check\nu_\alpha)^{*k}\big)(g) = \sum_{g \in \Gamma_S} \big(\nu_\alpha^{*k} * (\check\nu_\alpha)^{*k}\big)(g)\eta(g)$$

$$\leqslant \left(\sum_{g \in \Gamma_S} \big(\big(\nu_\alpha^{*k} * (\check\nu_\alpha)^{*k}\big)(g)\big)^\theta \right)^{\frac{1}{\theta}} \left(\sum_{g \in \Delta_k} \eta^p(g) \right)^{\frac{1}{p}}.$$

By Young's inequality, applied with $q = 2\theta/(1+\theta)$,

$$\big\|\nu_\alpha^{*k} * (\check\nu_\alpha)^{*k}\big\|_{\ell^\theta(\Gamma_S)} \leqslant \big\|\nu_\alpha^{*k}\big\|_{\ell^q(\Gamma_S)} \big\|(\check\nu_\alpha)^{*k}\big\|_{\ell^q(\Gamma_S)} = \big\|\nu_\alpha^{*k}\big\|_{\ell^q(\Gamma_S)}^2.$$

Since $\mathcal{F}$ has no exact overlaps, the sub-semigroup generated by $\{\tilde\gamma_i^\alpha : i \in \Lambda\}$ is free. Indeed, this can be seen directly from the decomposition in (4.13) of the elements $\tilde\gamma_\omega^\alpha$. In particular, for all $u, \omega \in \Lambda^k$,

$$\tilde\gamma_u^\alpha = \tilde\gamma_\omega^\alpha \iff u = \omega.$$

Combined with (4.17), it follows that

$$\big\|\nu_\alpha^{*k}\big\|_{\ell^q(\Gamma_S)}^2 = \left(\sum_{\omega \in \Lambda^k} \lambda_\omega^q \right)^{2/q} = \left(\sum_{i \in \Lambda} \lambda_i^q \right)^{2k/q},$$

where the last equality follows by Lemma 2.4, applied with $\tau_i = \lambda_i^{q-1}$. This completes the proof.   □

4.3. **Summability of Matrix Coefficients.** We show that the matrix coefficients of Γ_S acting on $L^2_{00}(\mathbf{G}_S/\Gamma_S)$ belong to $\ell^p(\Gamma_S)$ for an explicit choice of p, Proposition 4.4. This verifies the hypothesis of the last assertion of Proposition 4.3 for this value of p, thus completing the proof of the bound on the norm of the operators $\mathcal{P}_\lambda$.

Proposition 4.4. *Let*

$$v(d) = \left(\left\lfloor \frac{d}{2} \right\rfloor + 1 \right) \left\lceil \frac{d}{2} \right\rceil, \qquad p = 2v(d).$$

Then, $\xi_{\mathbf{G}} \in \ell^{p+\varepsilon}(\Gamma_S)$ for all $\varepsilon > 0$.

As a first step, we bound the functions ξ_v in terms of the operator norm of the adjoint action. We recall that for any place v of $\mathbb{Q}$ the group $\mathbf{G}(\mathbb{Q}_v)$ admits a so-called KAK-decomposition; cf. Appendix C. More precisely, for any $g \in \mathbf{G}(\mathbb{Q}_v)$ there are $k_1, k_2 \in M_v$ and $a \in A_v^+$ such that

$$g = k_1 a k_2. \tag{4.18}$$

Proposition 4.5. *Let $D = \lfloor \frac{d+1}{2} \rfloor$. For all sufficiently small $\varepsilon > 0$ there exists a constant $\delta(\varepsilon) > 0$ such that for all $g \in \mathbf{G}(\mathbb{Q}_v)$,*

$$\delta_1 \|g\|_v^{-\frac{1}{2}D} \leq \xi_v(g) \leq \delta_2(\varepsilon) \|g\|_v^{-\frac{1}{2}+\varepsilon}. \tag{4.19}$$

Remark 4.6. Note that $D = 1$ if $d = 1$ or $d = 2$.

Proof. We start with the case where $v = p$ is a finite place. Given $g \in \mathbf{G}(\mathbb{Q}_p)$, let $a(g) \in A_p^+$ denote the Cartan element defined by (4.18) normalized so that the bottom right entry equals 1. Using [Oh02, Thm. 5.9] and [Oh98, Prop. 2.3], cf. [Oh02, §6.1], we find $\delta_1, \delta_2(\varepsilon) > 0$ such that

$$\delta_1 \left(\prod_{i=1}^{\lfloor \frac{d+1}{2} \rfloor} \frac{|a(g)_i|_p}{|a(g)_{d+1-(i-1)}|_p} \right)^{-\frac{1}{2}} \leq \xi_p(g) \leq \delta_2(\varepsilon) \left(\prod_{i=1}^{\lfloor \frac{d+1}{2} \rfloor} \frac{|a(g)_i|_p}{|a(g)_{d+1-(i-1)}|_p} \right)^{-\frac{1}{2}+\varepsilon}. \tag{4.20}$$

It remains to bound the product appearing in (4.20) in terms of $\|\mathrm{Ad}_g\|_p$. Using (4.18) assume without loss of generality that

$$g = \mathrm{diag}(p^{-n_1}, \ldots, p^{-n_d}, 1)$$

for integers $n_1 \geq \cdots \geq n_d \geq 0$. Then,

$$\prod_{i=1}^{\lfloor \frac{d+1}{2} \rfloor} \frac{|a(g)_i|_p}{|a(g)_{d+1-(i-1)}|_p} = p^{\eta(g)},$$

where

$$\eta(g) = n_1 - n_{d+1} + n_2 - n_d + \cdots + n_{\lfloor \frac{d+1}{2} \rfloor} - n_{d+2-\lfloor \frac{d+1}{2} \rfloor}$$

and in particular

$$n_1 \leq \eta(g) \leq D n_1. \tag{4.21}$$

Therefore, since $\|g\|_p = \|\mathrm{Ad}_g\|_p = p^{n_1}$, we obtain the claim for finite places of $\mathbb{Q}$.

If v is the infinite place the argument is very similar. We again have

$$\delta_1 \left(\prod_{i=1}^{\lfloor \frac{d+1}{2} \rfloor} \frac{a(g)_i}{a(g)_{d+1-(i-1)}} \right)^{-\frac{1}{2}} \leq \xi_\infty(g) \leq \delta_2(\varepsilon) \left(\prod_{i=1}^{\lfloor \frac{d+1}{2} \rfloor} \frac{a(g)_i}{a(g)_{d+1-(i-1)}} \right)^{-\frac{1}{2}+\varepsilon} \tag{4.22}$$

and

$$a(g)_1 \leq \prod_{i=1}^{\lfloor \frac{d+1}{2} \rfloor} \frac{a(g)_i}{a(g)_{d+1-(i-1)}} \leq a(g)_1^D. \tag{4.23}$$

We recall that $\|\mathrm{Ad}_g\|_\infty = \|\mathrm{Ad}_{a(g)}\|_\infty$ and $\mathrm{Ad}_{a(g)}$ is diagonalizable with eigenvalues

$$\sigma(\mathrm{Ad}_{a(g)}) = \left\{ \frac{a(g)_i}{a(g)_j} : 1 \leq i, j \leq d + 1 \right\}.$$

By definition, we get

$$\|\mathrm{Ad}_g\|_\infty^2 = \sum_{1 \leq i,j \leq d+1} \frac{a(g)_i^2}{a(g)_j^2}$$

and, in particular,

$$a(g)_1 \leq \|\mathrm{Ad}_g\|_\infty \leq (1 + \dim \mathbf{G})^{\frac{1}{2}} a(g)_1. \tag{4.24}$$

Combining this with (4.22), (4.23) and $\|g\|_\infty = \|\mathrm{Ad}_g\|_\infty$, the claim follows. $\qquad\square$

We record for later reference that (4.24) implies that, given $t \in (0, \infty)$ and $a(t)$ as in Section 4.1, we have

$$\|\mathrm{Ad}_{a(t)}\|_\infty = \|a(t)\|_\infty \asymp_d \max\{t, t^{-1}\}. \tag{4.25}$$

4.3.1. *Volume growth for norm balls in* $\mathbf{G}(\mathbb{Q}_p)$. In preparation of the proof of Proposition 4.4 we derive bounds on the volume of norm balls in $\mathbf{G}(\mathbb{Q}_p)$ for p a finite place of $\mathbb{Q}$. We fix a choice of a Haar measure m_p on $\mathbf{G}(\mathbb{Q}_p)$ such that $m_p(\mathbf{G}(\mathbb{Z}_p)) = 1$.

Lemma 4.7. *Let* $n \in \mathbb{N} \cup \{0\}$ *and set*

$$v(d) = \left(\left\lfloor \frac{d}{2} \right\rfloor + 1 \right) \left\lceil \frac{d}{2} \right\rceil.$$

Then for all $\varepsilon > 0$ *we have*

$$p^{v(d)n} \ll_d \mathrm{Vol}(\{g \in \mathbf{G}(\mathbb{Q}_p) : \|g\|_p = p^n\}) \ll_{d,\varepsilon} p^{(v(d)+\varepsilon)n}.$$

If $d = 1$, *the latter bound remains valid for* $\varepsilon = 0$.

Proof. Note that $\|g\|_p = 1$ if and only if $g \in \mathbf{G}(\mathbb{Z}_p)$. Hence we can assume that $n \in \mathbb{N}$. Let

$$I_n^d = \{(n_1, \ldots, n_d) \in (\mathbb{N} \cup \{0\})^d : n = n_1 \geq \cdots \geq n_d\}.$$

Given $\mathbf{n} \in I_n^d$, we let $a_{\mathbf{n}} = \mathrm{diag}(p^{-n_1}, \ldots, p^{-n_d}, 1)$. Then

$$\{g \in \mathbf{G}(\mathbb{Q}_p) : \|g\|_p = p^n\} = \bigsqcup_{\mathbf{n} \in I_n^d} K a_{\mathbf{n}} K. \tag{4.26}$$

It therefore remains to determine the cardinality of the set I_n^d and the Haar measure of sets of the form $K a_{\mathbf{n}} K$. It is known, cf. [Sl79, Lem. 4.1.1], that

$$\mathrm{Vol}(K a_{\mathbf{n}} K) \asymp \delta_{\mathbf{B}}(a_{\mathbf{n}}), \tag{4.27}$$

where $\delta_{\mathbf{B}}$ is the modular character on the image $\mathbf{B} \leq \mathbf{G}$ of the upper triangular subgroup in GL_{d+1}. Let $\mathbf{n} \in I_n^d$. One calculates (cf. Appendix D)

$$\delta_{\mathbf{B}}(a_{\mathbf{n}}) = \exp\left(\sum_{i=1}^d n_i(d + 2 - 2i) \log p \right). \tag{4.28}$$

As $n_i \geq 0$, the right-hand side attains its maximum at $\mathbf{n}^*$ given by

$$n_i^* = \begin{cases} n & \text{if } 2i \leq d + 2, \\ 0 & \text{else.} \end{cases}$$

One calculates

$$\sum_{i=1}^{d} n_i^*(d+2-2i) = n\left(\left\lfloor\frac{d}{2}\right\rfloor+1\right)\left\lceil\frac{d}{2}\right\rceil.$$

In particular, combining (4.26), (4.27) and (4.28), we obtain

$$p^{v(d)n} \ll_d \mathrm{Vol}(\{g \in \mathbf{G}(\mathbb{Q}_p) : \|g\|_p = p^n\}) \ll_d |I_n^d|p^{v(d)n}.$$

We next determine the cardinality of I_n^d. We first note that for $d = 1$ we clearly have $|I_n^d| = 1$ and therefore the last part of the lemma follows immediately. For general $d \in \mathbb{N}$ we note that I_n^d is precisely the set of ordered $d-1$-tuples of non-negative integers at most equal to n or, put differently, I_n^d identifies with the collection of multisets of cardinality $d-1$ with elements in $\{0,\ldots,n\}$. Therefore, we find

$$|I_n^d| = \binom{n+d-1}{d-1} \asymp_d n^{d-1}.$$

It follows that

$$\mathrm{Vol}(\{g \in \mathbf{G}(\mathbb{Q}_p) : \|g\|_p = p^n\}) \ll_d n^{d-1}p^{v(d)n} \ll_{d,\varepsilon} p^{(v(d)+\varepsilon)n}$$

and the lemma is proven. □

Corollary 4.8. *Let* $\eta_S : \mathbf{G}_S \to (0,\infty)$ *denote the function*

$$\eta_S(g) = \prod_{v \in S} \|g_v\|_v^{-1} \quad (g \in \mathbf{G}_S).$$

Then, $\eta_S \in \mathrm{L}^{v(d)+\varepsilon}(\mathbf{G}_S)$.

Proof. This follows relatively easily from the description of the Haar measure in terms of the KAK-decomposition for the infinite place, cf. [Kn01, Prop. 5.28], and from the bound in Lemma 4.7 for finite places; cf. Appendix E. □

4.4. **Proof of Proposition 4.4.** Note that for all $v \in S$ and for all $g_v \in \mathbf{G}(\mathbb{Q}_v)$ we have $\|g_v^{-1}\|_v = \|g_v\|_v$. Hence submultiplicativity implies that for all $\gamma_v \in \mathbf{G}(\mathbb{Q}_v)$ we have

$$\|g_v\|_v^{-1}\|\gamma_v\|_v \le \|g_v\gamma_v\|_v.$$

Using the folding-unfolding technique and Corollary 4.8 we find that for $q > v(d)$ we have

$$\infty > \int_{\mathbf{G}_S} \eta_S(g)^q \mathrm{d}g = \int_{X_S} \sum_{\gamma \in \Gamma_S} \eta_S(g\gamma)^q \mathrm{d}g\Gamma_S$$

and hence for almost all $g \in \mathbf{G}_S$ we have

$$\sum_{\gamma \in \Gamma_S} \eta_S(g\gamma)^q < \infty$$

by Fubini's theorem. By submultiplicativity, we have $\eta_S(g\gamma)^q \ge \eta_S(g)^q\eta_S(\gamma)^q$ and therefore

$$\sum_{\gamma \in \Gamma_S} \eta_S(\gamma)^q < \infty.$$

Using Proposition 4.5, it follows that

$$\sum_{\gamma \in \Gamma_S} \xi_{\mathbf{G}}(\gamma)^{\frac{2q}{1-2\varepsilon}} \ll_{\varepsilon,d} \sum_{\gamma \in \Gamma} \eta_S(\gamma)^q < \infty$$

and in particular $\xi_{\mathbf{G}} \in \ell^{2v(d)+\varepsilon}(\Gamma_S)$.

Remark 4.9. Note that the above argument works for any unimodular subgroup of $\mathbf{G}_S$ in place of the integer lattice Γ_S.

5. Expanding Horospheres and Congruence Covers

The goal of this section is to show that $a(t)$-translates of absolutely continuous measures on the horospherical group of $a(t)$ become equidistributed, in a suitable sense, towards the Haar measure on quotients of $\mathbf{G}_\infty$ by principal congruence subgroups of $\mathbf{G}(\mathbb{Z})$, with an emphasis on obtaining a uniform error rate and implied constants, independently of the congruence level. The main result of this section is Proposition 5.1.

Recall the notation introduced in Section 3.1. We fix a right-invariant Riemannian metric on $\mathbf{G}_\infty$. This metric induces a right invariant metric on the connected component $\mathbf{G}_\infty^+$ of $\mathbf{G}_\infty$. For any lattice $\Gamma \leq \mathbf{G}_\infty$, this induces a Riemannian metric on $\mathbf{G}_\infty/\Gamma$ such that the canonical projection is a local isometry. Given $x \in \mathbf{G}_\infty/\Gamma$, we denote by $\mathrm{inj}_\Gamma(x)$ the supremum over all radii R such that, for all $g \in \mathbf{G}_\infty^+$ contained in the ball of radius R at the origin, the map $g \mapsto gx$ is injective.

For a Lipschitz function φ, we write $\mathrm{Lip}(\varphi)$ for its Lipschitz constant and we denote the space of Lipschitz functions on $\mathbf{G}_\infty/\Gamma$ by $\mathrm{Lip}(\mathbf{G}_\infty/\Gamma) = \{\varphi \in \mathrm{C}(\mathbf{G}_\infty/\Gamma) : \mathrm{Lip}(\varphi) < \infty\}$.

Note that $X_\infty(N)$ is in general not connected for $N > 1$ and hence some care is needed in formulating equidistribution statements. The connected components of $X_\infty(N)$ correspond to the distinct orbits of $\mathbf{G}_\infty^+$. Each such component supports a unique $\mathbf{G}_\infty^+$-invariant Haar probability measure. Moreover, a function is **orthogonal to $\mathbf{G}_\infty^+$-invariant functions** in $\mathrm{L}^2(X_\infty(N))$ if and only if it has integral 0 on each connected component of $X_\infty(N)$.

The following is the main result of this section.

Proposition 5.1. *There exist $\kappa > 0, \ell \in \mathbb{N}$ and $C \geq 1$ such that the following holds. Let $\psi \in \mathrm{C}_c^\ell(\mathbb{R}^d)$ be a non-negative function of integral 1. Then, for every $N \in \mathbb{N}$, $\varphi \in \mathrm{B}_{2,\ell}^\infty(X_\infty(N)) \cap \mathrm{Lip}(X_\infty(N))$, and for every $t \geq 1$, the following holds for all $x \in X_\infty(N)$:*

$$\int \varphi\big(a(t)u(\mathbf{x})x\big)\psi(\mathbf{x})\,\mathrm{d}\mathbf{x}$$
$$= \int \varphi\,\mathrm{d}m_{\mathbf{G}_\infty^+ \cdot x} + O_\psi\Big(\sqrt{[\Gamma(1):\Gamma(N)]}\big(\mathcal{S}_{2,\ell}(\varphi) + \mathrm{Lip}(\varphi)\big)\max\{\mathrm{inj}_{\Gamma(N)}(x)^{-C},1\}\cdot t^{-\kappa}\Big),$$

where $m_{\mathbf{G}_\infty^+ \cdot x}$ is the unique $\mathbf{G}_\infty^+$-invariant probability measure on $\mathbf{G}_\infty^+ \cdot x$ and $\mathrm{inj}_{\Gamma(N)}(x)$ denotes the injectivity radius at x.

Without any attempt to optimize the exponents, we show Proposition 5.1 holds for any integer $\ell > d(d+1)/4$ and that κ can be chosen as follows:

$$\kappa = \frac{\kappa' - \varepsilon}{2 + 2d + 6\ell + d^2}, \tag{5.1}$$

for any $\varepsilon > 0$ (the implicit constant depends on ε and ℓ), where

$$\kappa' = \begin{cases} 25/64 & \text{if } d = 1, \\ 1/2 & \text{otherwise.} \end{cases} \tag{5.2}$$

The value of κ' comes from known bounds towards Selberg's eigenvalue conjecture [KS03]; cf. Proposition 5.3 below. It is possible to obtain much better values for κ (possibly at the cost of worse values of ℓ) via more analytic techniques similar to those in [Sa81, Bu90, FF03, St04, St13, Ed17].

Remark 5.2. The main point of Proposition 5.1 is the explicit dependence of the implied constant on N. This statement is well-known but we include a proof as we could not locate it in the literature. We note also that the implied constant depends on $\|\psi\|_{\mathrm{C}^\ell}$ and the radius of the smallest ball around the origin containing its support.

5.1. **Uniform Spectral Gap.** We start with a standard result which is a crucial ingredient to the proof of Proposition 5.1.

Proposition 5.3 (Uniform spectral gap)**.** *Let $\varepsilon > 0$. For all $N \in \mathbb{N}$, for all $t \geq 1$, and for all $\varphi, \psi \in \mathrm{B}_{2,\ell}^{\infty}(X_{\infty}(N))$ which are orthogonal to the $\mathbf{G}_{\infty}^{+}$-invariant functions, we have*

$$|\langle a(t)\varphi, \psi \rangle_2| \leq C\mathcal{S}_{2,\ell}(\varphi)\mathcal{S}_{2,\ell}(\psi)t^{-\kappa'+\varepsilon}.$$

where κ' is as in (5.2)*, ℓ is any integer larger than half the dimension of the maximal compact subgroup $M_{\infty} \leq \mathbf{G}_{\infty}$ and $C \geq 1$ depending only on d, ε and ℓ.*

Sketch of the Proof. This result is well-known and we only emphasize the fact that C is independent of N. We recall that $\mathrm{L}_{00}^2(X_{\infty}(N))$ denotes the orthogonal complement to the subspace of $\mathbf{G}_{\infty}^{+}$-invariant vectors in $\mathrm{L}^2(X_{\infty}(N))$ and that $\mathbf{G}_{\infty}^{+}$ is the image of $G = \mathrm{SL}_{d+1}(\mathbb{R})$ in $\mathbf{G}_{\infty}$. As $\mathbf{G}_{\infty}^{+}$ has index 2 inside $\mathbf{G}_{\infty}$, $X_{\infty}(N)$ consists of at most two connected components, each of which is isomorphic to G/Δ, where Δ is a congruence lattice in $\mathrm{SL}_{d+1}(\mathbb{Z})$. Moreover, when $d \geq 2$, Vogan's classification of the unitary dual of GL_{d+1} implies that matrix coefficients of a dense subset of vectors of any non-trivial, irreducible, unitary G-representation belong to $\mathrm{L}^{2d+\varepsilon}$ for every $\varepsilon > 0$; cf. [Oh98, Corollary C]. This in particular applies to the G-representations $\mathrm{L}_{00}^2(X_{\infty}(N))$. It then follows by [CHH88, Corollary on pg. 108] that for any two M_{∞}-finite vectors $v, w \in \mathrm{L}_{00}^2(X_{\infty}(N))$, we have

$$|\langle a(t)v, w \rangle| \leq (\dim\langle M_{\infty}v \rangle \dim\langle M_{\infty}w \rangle)^{\frac{1}{2}}\|v\|\|w\|\Xi\big(a(t)\big)^{\frac{1}{d}}, \tag{5.3}$$

where Ξ denotes the Harish-Chandra spherical function on G. Using [Wa, Thm. 4.5.3] and denoting by $\mathfrak{u}$ the Lie algebra of the subgroup $U = \{u(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^d\}$, we know that there is some $\delta > 0$ depending on d such that for $\|\cdot\|_{\infty}$ as in 3.2, we have

$$\Xi\big(a(t)\big) \ll_d e^{-\frac{1}{2}\operatorname{tr}\operatorname{ad}\log a(t)|_{\mathfrak{u}}}(1 + \log\|a(t)\|_{\infty})^{\delta} = t^{-\frac{d}{2}}(1 + \log t)^{\delta}.$$

In particular, we obtain that

$$|\langle a(t)v, w \rangle| \ll_{\varepsilon,d} (\dim\langle M_{\infty}v \rangle \dim\langle M_{\infty}w \rangle)^{\frac{1}{2}}\|v\|\|w\|t^{-\frac{1}{2}+\varepsilon}.$$

The statement for smooth vectors follows by the argument in [EMV09, Section 6.2.1] (with C depending only on ℓ).

In the case $d = 1$, it is shown in [KS03, Proposition 2] that the smallest non-zero eigenvalue of the Laplacian on $\mathbb{H}^2/\Delta$ is $\geq 975/4096$ for any congruence lattice Δ in $\mathrm{SL}_2(\mathbb{R})$. Using [Ra87, Theorem 2] and the formula for the Haar measure, this implies that smooth matrix coefficients of G belong to $\mathrm{L}^{64/25+\varepsilon}$ for all $\varepsilon > 0$. Hence, (5.3) follows in this case by [Sh00, Theorem 2.1]. The statement for smooth vectors follows upon combining (5.3) with [EMV09] as above. $\qquad\square$

5.2. **Proof of Proposition 5.1.** Without loss of generality, we will assume that φ is identically 0 on all connected components of $X_{\infty}(N)$, except the one containing x. By further replacing φ with $\varphi - \big(\int \varphi \, \mathrm{d}m_{\mathbf{G}_{\infty}^{+}\cdot x}\big)\chi_{\mathbf{G}_{\infty}^{+}\cdot x}$, we may assume it is orthogonal to $\mathbf{G}_{\infty}^{+}$-invariant functions in $\mathrm{L}^2(X_{\infty}(N))$.

We use the standard thickening technique to deduce Proposition 5.1 from Proposition 5.3. More precisely, using Proposition 5.3, one can deduce the following Proposition 5.4 which was originally obtained in [KM12].

We abuse notation and denote by $\mathcal{S}_{2,\ell}$ the $(2,\ell)$-Sobolev norm on $\mathbb{R}^d$.

Proposition 5.4 (Theorem 2.3, [KM12])**.** *There exists a constant $r_0 > 0$, depending only on $\mathbf{G}_{\infty}$, such that the following holds. Let $\psi \in \mathrm{C}_c^{\infty}(\mathbb{R}^d)$, $0 < r < r_0/2$, and $x \in X_{\infty}(N)$ for some $N \geq 1$. Suppose that ψ is supported in the ball of radius r around the origin in $\mathbb{R}^d$ and that the injectivity radius at x is at least $2r$. Let ℓ and κ' be as in Proposition 5.3. Then, for every*

$\varphi \in \mathrm{B}_{2,\ell}^{\infty}(X_{\infty}(N)) \cap \mathrm{Lip}(X_{\infty}(N))$ *such that φ is orthogonal to $\mathbf{G}_{\infty}^{+}$-invariant functions and for all $t \geq 1$ and $\varepsilon > 0$,*

$$\left| \int \varphi(a(t)u(\mathbf{x})x)\psi(\mathbf{x})\, d\mathbf{x} \right|$$
$$\ll_{\varepsilon} V_N\big(\mathcal{S}_{2,\ell}(\varphi) + \mathrm{Lip}(\varphi)\big)\left(r\int_{\mathbb{R}^d} |\psi| + r^{-(2\ell+k/2)}\mathcal{S}_{2,\ell}(\psi)t^{-\kappa'+\varepsilon} \right),$$

where $V_N = \sqrt{[\Gamma(1):\Gamma(N)]}$ and $k = \dim \mathbf{G}_{\infty} - d = d^2 + d$.

Proof. This result was obtained in [KM12, Theorem 2.3] in the case $N = 1$. We give a sketch of the required modifications. For general N, one replaces Theorem 2.1 in *loc. cit.* with Proposition 5.3 above.

The factor V_N arises as follows. Let $\Gamma^+(N) = \mathbf{G}_{\infty}^{+} \cap \Gamma(N)$ and $I_N^{+} = [\Gamma^+(1) : \Gamma^+(N)]$. We assume that the Haar measure on $\mathbf{G}_{\infty}^{+}$ is normalized so that it projects to a probability measure on $\mathbf{G}_{\infty}^{+}/\Gamma^+(1)$. Each connected component of $X_{\infty}(N)$ is isomorphic to $\mathbf{G}_{\infty}^{+}/\Gamma^+(N)$. Hence, in order to locally decompose the Haar measure on $\mathbf{G}_{\infty}^{+}$ into the product of the Lebesgue measure on $\mathbb{R}^d$ and the non-expanding subgroup in a compatible manner, cf. [KM12, Eq. (2.3)], the measure on the non-expanding subgroup needs to be scaled by $1/I_N^{+}$. This scaling implies that the Sobolev norm of the bump function on the non-expanding subgroup is scaled by the square root of I_N^{+}. Using that $[\mathbf{G}_{\infty} : \mathbf{G}_{\infty}^{+}] = 2$, one shows

$$I_N^{+} \ll [\Gamma(1) : \Gamma(N)].$$

We leave the details to the reader. Finally, one uses that $X_{\infty}(N)$ has at most two connected components to bound the Sobolev norms of the restriction of φ to $\mathbf{G}_{\infty}^{+} \cdot x$, as they occur in the proof of [KM12, Thm. 2.3], by $\mathcal{S}_{2,\ell}(\varphi)$. $\qquad\square$

Let $\kappa > 0$ be a parameter to be chosen later. Let ψ be as in the statement of Proposition 5.1. Let $r_0 > 0$ be the number provided by Proposition 5.4. Let $R > 0$ denote the infimum over all $\tilde{R} > 0$ such that ψ is supported in the ball $B(\mathbf{0}, \tilde{R})$ of radius $\tilde{R}$ around the origin. Define r by

$$r := \frac{1}{2}\min\left\{ 1, r_0, \inf_{\mathbf{y}\in B(\mathbf{0},4R)} \mathrm{inj}_{\Gamma(N)}(u(\mathbf{y})x) \right\}.$$

Note that for all $\mathbf{y} \in B(\mathbf{0}, 4R)$, $\mathrm{inj}_{\Gamma(N)}(x) \ll_R \mathrm{inj}_{\Gamma(N)}(u(\mathbf{y})x)$, where the implicit constant is independent of N. It follows that

$$r^{-1} \ll_R \max\left\{ 1, \mathrm{inj}_{\Gamma(N)}(x)^{-1} \right\}. \tag{5.4}$$

Let $V_N = \sqrt{[\Gamma(1):\Gamma(N)]}$. If $2R \leq t^{-\kappa}r$, using $t \geq 1$ we apply Proposition 5.4 with $t^{-\kappa}r$ in place of r to get

$$\int \varphi(a(t)u(\mathbf{x})x)\psi(\mathbf{x})\mathrm{d}\mathbf{x} \ll V_N(\mathcal{S}_{2,\ell}(\varphi) + \|\varphi\|_{\mathrm{Lip}})\left(t^{-\kappa}r\int \psi + (t^{-\kappa}r)^{-(2\ell+\frac{k}{2})}\mathcal{S}_{2,\ell}(\psi)t^{-\kappa'+\varepsilon} \right)$$
$$\leq V_N(\mathcal{S}_{2,\ell}(\varphi) + \|\varphi\|_{\mathrm{Lip}})\left(\max\{1, \mathcal{S}_{2,\ell}(\psi)\}\right)\left(t^{-\kappa}r + (t^{-\kappa}r)^{-(2\ell+\frac{k}{2})}t^{-\kappa'+\varepsilon} \right), \tag{5.5}$$

where we used our assumption that the integral of ψ is 1. By equating the terms involving t in the error above, we choose κ to satisfy

$$\kappa \leq \frac{\kappa' - \varepsilon}{1 + 2\ell + \frac{k}{2}}. \tag{5.6}$$

Since $r < 1$ and, in view of (5.4) and letting $c = 2\ell + k/2$, we obtain

$$\int \varphi(a(t)u(\mathbf{x})x)\psi(\mathbf{x})\mathrm{d}\mathbf{x} \ll V_N(\mathcal{S}_{2,\ell}(\varphi) + \|\varphi\|_{\mathrm{Lip}})\left(\max\{1, \mathcal{S}_{2,\ell}(\psi)\}\right)\max\left\{1, \mathrm{inj}_{\Gamma(N)}(x)^{-c}\right\}t^{-\kappa}.$$
(5.7)

If $2R > t^{-\kappa}r$, we choose a smooth partition of unity $\{\phi_i\}$ of $B(\mathbf{0}, 2R)$ with the following properties:

(1) Each ϕ_i is supported in a ball of radius $t^{-\kappa}r$ and satisfies $0 \le \phi_i \le 1$.
(2) For every i, $\mathcal{S}_{2,\ell}(\phi_i) \ll t^{\kappa(\ell-d/2)}r^{-(\ell-d/2)}$.
(3) The cardinality of the set of indices $j \ne i$ such that the supports of ϕ_i and ϕ_j intersect non-trivially is bounded above by a constant C_d, depending only on d.
(4) $\#\{\phi_i\} = O((R/t^{-\kappa}r)^d)$.

Denote by $\mathbf{y}_i \in B(\mathbf{0}, 4R)$ the center of a $t^{-\kappa}r$-ball containing the support of ϕ_i and by ψ_i the function $\phi_i\psi$. Let $x_i = u(\mathbf{y}_i)x$. It follows from the properties of the norm $\mathcal{S}_{2,\ell}(\cdot)$ that (cf. [KM12, Lemma 2.2])

$$\mathcal{S}_{2,\ell}(\psi_i) \ll \mathcal{S}_{2,\ell}(\phi_i)\|\psi\|_{\mathrm{C}^\ell} \ll t^{\kappa(\ell-d/2)}r^{-(\ell-d/2)}\|\psi\|_{\mathrm{C}^\ell}.$$

Fix i and apply Proposition 5.4 with $t^{-\kappa}r$, $\psi_i(\cdot + \mathbf{y}_i)$ and x_i in place of r, ψ and x respectively to get

$$\int \varphi(a(t)u(\mathbf{x})x)\psi_i(\mathbf{x})\,d\mathbf{x} = \int \varphi(a(t)u(\mathbf{x} - \mathbf{y_i})x_i)\psi_i(\mathbf{x})\,d\mathbf{x} = \int \varphi(a(t)u(\mathbf{x})x_i)\psi_i(\mathbf{x} + \mathbf{y}_i)\,d\mathbf{x}$$

$$= O\left(V_N(\mathcal{S}_{2,\ell}(\varphi) + \mathrm{Lip}(\varphi))\left(t^{-\kappa}r\int_{\mathbb{R}^d}\psi_i + r^{-(2\ell-k/2)}\mathcal{S}_{2,\ell}(\psi_i)_\ell t^{\kappa(2\ell-k/2)-\kappa'+\varepsilon}\right)\right)$$

$$= O\left(V_N(\mathcal{S}_{2,\ell}(\varphi) + \mathrm{Lip}(\varphi))\left(t^{-\kappa}r\int_{\mathbb{R}^d}\psi_i + r^{-(3\ell+k/2-d/2)}\|\psi\|_{\mathrm{C}^\ell}t^{\kappa(3\ell+k/2-d/2)-\kappa'+\varepsilon}\right)\right),$$

where k and κ are as in the proposition. Here we used the non-negativity of ψ_i. Using the fact that ϕ_i is a partition of unity and that $\int_{\mathbb{R}^d}\psi = 1$, we thus obtain

$$\int \varphi(a(t)u(\mathbf{x})x)\psi(\mathbf{x})\,d\mathbf{x} = \sum_i \int \varphi(a(t)u(\mathbf{x})x)\psi_i(\mathbf{x})\,d\mathbf{x}$$
(5.8)

$$= \sum_i O\left(V_N(\mathcal{S}_{2,\ell}(\varphi) + \mathrm{Lip}(\varphi))\left(t^{-\kappa}r\int_{\mathbb{R}^d}\psi_i + r^{-(3\ell+k/2-d/2)}\|\psi\|_{\mathrm{C}^\ell}t^{\kappa(3\ell+k/2-d/2)-\kappa'+\varepsilon}\right)\right)$$
(5.9)

$$= O\left(V_N(\mathcal{S}_{2,\ell}(\varphi) + \mathrm{Lip}(\varphi))\left(t^{-\kappa}r + R^d r^{-(3\ell+k/2+d/2)}\|\psi\|_{\mathrm{C}^\ell}t^{\kappa(3\ell+k/2+d/2)-\kappa'+\varepsilon}\right)\right).$$
(5.10)

Equating the powers of t in the two terms above, we obtain

$$\kappa = \frac{\kappa' - \varepsilon}{1 + d/2 + 3\ell + k/2},$$

which also satisfies (5.6). Using (5.4) it follows that

$$\int \varphi(a(t)u(\mathbf{x})x)\psi(\mathbf{x})\,d\mathbf{x} = O_R\left(V_N(\mathcal{S}_{2,\ell}(\varphi) + \mathrm{Lip}(\varphi))\max\{1, \|\psi\|_{\mathrm{C}^\ell}\}t^{-\kappa}\max\left\{\mathrm{inj}_{\Gamma(N)}(x)^{-c}, 1\right\}\right),$$

where $c = 3\ell + k/2 + d/2$. As $\min\{\mathrm{inj}_{\Gamma(N)}, 1\} \le 1$, letting $C = 3\ell + k/2 + d/2$, the claim follows.

6. Effective Equidistribution of Fractal Measures

The goal of this section is to prove that translates of certain self-similar measures become effectively equidistributed on the space of unimodular lattices, Theorem 6.1. This result is one of the main contributions of this article and constitutes the main ingredient in our proof of the analogue of Khintchine's theorem for fractal measures.

We fix a tuple $(\mathcal{F}, \lambda)$ as in Section 2.2 and we assume that $\mathcal{F}$ is rational (cf. (1.3)) and satisfies the open set condition. We denote by $\mu = \mu_{(\mathcal{F}, \lambda)}$ the associated self-similar probability measure. Throughout this section, we use r to denote the average contraction ratio

$$r := \sum_{i \in \Lambda} \lambda_i \rho_i. \tag{6.1}$$

Recall the definition of the spaces $\mathrm{B}^\infty_{\infty, \ell}$ in (3.7) and the definitions of O_α, ρ_α and b_α in (2.2) for $\alpha \in \Lambda^*$. We will also use the notation introduced in Sections 3.1 and 4.1 pertaining to the algebraic group $\mathbf{G} = \mathrm{PGL}_{d+1}$.

Let $\kappa > 0$, and $\ell \in \mathbb{N}$ be constants satisfying Proposition 5.1. Recall that ℓ can be chosen to be any integer with $\ell > d(d+1)/4$ (cf. discussion following Proposition 5.1). In particular, we choose

$$\ell = \dim M_\infty = \frac{d(d+1)}{2} \tag{6.2}$$

so that the orders of the Sobolev norms in Proposition 4.3 and 5.1 match. We assume without loss of generality that

$$\kappa \leq d(d+1). \tag{6.3}$$

Let p be the constant provided by Proposition 4.4. For $\varepsilon > 0$, denote by θ_ε the Hölder conjugate of $(p + \varepsilon)/\epsilon(d)$; cf. (3.11), i.e. θ_ε is the constant satisfying

$$\frac{1}{\theta_\varepsilon} + \frac{\epsilon(d)}{p + \varepsilon} = 1.$$

Let $q_\varepsilon = 2\theta_\varepsilon/(\theta_\varepsilon + 1) > 1$. Note that $p/\epsilon(d) > 1$ and in particular $1 < q_\varepsilon < \infty$. Let A and L be the constants satisfying Lemma 4.2.

Define positive constants $\sigma, o_\varepsilon, \upsilon$ by the following equations:

$$r^{-\sigma} = \left(\sum_{i \in \Lambda} \lambda_i^2 \rho_i^{-d}\right)^{1/2}, \quad r^{o_\varepsilon} = \left(\sum_{i \in \Lambda} \lambda_i^{q_\varepsilon}\right)^{1/q_\varepsilon}, \quad r^{-\upsilon} = \rho_{\min}^{-L/4} \sum_{i \in \Lambda} \lambda_i \rho_i^{-\ell}. \tag{6.4}$$

The positivity of υ follows from the open set condition; cf. Remark 6.5. Given $\alpha \in \Lambda^*$, we use the following notation:

$$h_\alpha := k_\alpha^{-1} a(1/\rho_\alpha) u(b_\alpha), \qquad x_\alpha := h_\alpha \Gamma(1).$$

Theorem 6.1. *Let μ be as above and suppose that*

$$\frac{2\sigma(o_\varepsilon + \upsilon)}{o_\varepsilon + \sigma} < \kappa, \tag{6.5}$$

for some $\varepsilon > 0$. Then, there exist $\delta > 0$ and $A_* \geq 1$ such that for every word $\alpha \in \Lambda^*$, $\varphi \in \mathrm{B}^\infty_{\infty, \ell}(X_\infty(1))$, and $t > 1$, the following holds:

$$\int \varphi\left(a(t)u(\mathbf{x})x_\alpha\right) d\mu(\mathbf{x}) = \int_{X_\infty(1)} \varphi + O\left(\rho_\alpha^{-A_*} \mathcal{S}_{\infty, \ell}(\varphi) \cdot t^{-\delta}\right),$$

where ℓ is as in (6.2). The implicit constant depends on $\mathcal{F}, \mu$ and ε.

Remark 6.2. We refer the reader to Appendix A, where we show that the above result holds under the hypothesis $\dim_H(\mathcal{K}) > 0.839$, in place of (6.5) in the special case of missing digit Cantor sets $\mathcal{K}$ in a prime base. In particular, in that case, we remove the dependence of the cutoffs on the

number of derivatives in the Sobolev norm in the parameter υ above. That part of the arguments in Appendix A is in fact valid for any IFS with equal contraction ratios.

Remark 6.3. Recall that for all $\alpha, \omega \in \Lambda^*$, we have $\lambda_{\alpha\omega} = \lambda_\alpha \lambda_\omega$ and $\rho_{\alpha\omega} = \rho_\alpha \rho_\omega$. This implies that for any $k \in \mathbb{N}$ the IFS $\mathcal{F}_k = \{f_\alpha : \alpha \in \Lambda^k\}$, given by considering the k-iterates of the IFS $\mathcal{F}$ with the probability vector λ^{*k} on Λ^k, has average contraction ratio r^k; cf. Lemma 2.4. The same lemma implies that the constants $\sigma, o_\varepsilon, \upsilon$ in (6.4) for $(\mathcal{F}_k, \lambda^{*k})$ are the same as the constants determined by $(\mathcal{F}, \lambda)$, i.e., Hypothesis (6.5) is invariant under iteration of the IFS.

Theorem 6.1 yields the following:

Corollary 6.4. *Let* $G = \mathrm{SL}_{d+1}(\mathbb{R})$, $\Gamma = \mathrm{SL}_{d+1}(\mathbb{Z})$. *For* $\alpha \in \Lambda^*$, *let* $k_\alpha \in \mathrm{SL}_{d+1}(\mathbb{R})$ *be defined as in* (4.1) *and let* $t_\alpha = \frac{d}{d+1} \log \rho_\alpha$. *Define*

$$x_\alpha^{\mathrm{u}} = k_\alpha^{-1} g_{-t_\alpha} u(b_\alpha) \Gamma \in G/\Gamma. \tag{6.6}$$

Suppose that (6.5) *holds and let* A_* *for some* $\varepsilon > 0$ *and* δ *be as in Theorem* 6.1. *Let* $\kappa_* = \frac{d+1}{d}\delta$. *Then,*

$$\int \varphi\left(g_t u(\mathbf{x}) x_\alpha^{\mathrm{u}}\right) \, d\mu(\mathbf{x}) = \int \varphi \, dm_{G/\Gamma} + O_\varepsilon\left(\rho_\alpha^{-A_*} \mathcal{S}_{\infty,\ell}(\varphi) \cdot e^{-\kappa_* t}\right), \tag{6.7}$$

for all $\varphi \in \mathrm{B}_{\infty,\ell}^\infty(G/\Gamma)$, $t > 0$, *and* $\alpha \in \Lambda^*$.

Proof. Recalling the notation in (1.6), we have that x_α^{u} is the image of $k_\alpha^{-1} a(1/\rho_\alpha) u(b_\alpha) \Gamma(1) \in X_\infty(1)$ under the identification $X_\infty(1) \cong G/\Gamma$ from Lemma 3.1. Then, the corollary follows by Theorem 6.1. The explicit value for κ_* is obtained by noting that the canonical projection maps $g_t \in G$ to $a(e^{(d+1)t/d})$ in $\mathbf{G}_\infty$. $\qquad \square$

6.1. Proof of Theorem B using Theorem 6.1.

Let κ, ℓ, L and p be as above. Assume that (1.4) holds with

$$\epsilon_0 = \min\left\{1, \frac{\kappa\epsilon(d)}{d\epsilon(d) + (4\ell + L)p}\right\}. \tag{6.8}$$

We show that this implies that (6.5) holds. The proof is purely formal, and relies on the fact that $(\rho_i^s)_{i \in \Lambda}$ form a probability vector. This follows from the open set condition; cf. [Mo46].

Let $\varepsilon > 0$ be sufficiently small so that

$$\left(\frac{d\log\rho_{\min}}{\log\lambda_{\max}} - 1\right)\frac{\log\lambda_{\min}}{s\log\rho_{\max}} \leqslant \delta_0 := \min\left\{1, \frac{\kappa\epsilon(d)}{d\epsilon(d) + (4\ell + L)(p + \varepsilon)}\right\}. \tag{6.9}$$

Such ε exists since the left-hand side is assumed to be strictly less than ϵ_0.

First, we find bounds on σ, o_ε and υ. Let $p_\varepsilon = (p + \varepsilon)/\epsilon(d)$. We note that $q_\varepsilon/(q_\varepsilon - 1) = 2p_\varepsilon$. Hence, since λ_i form a probability vector, we get $\lambda_{\min}^{1/2p_\varepsilon} \leqslant r^{o_\varepsilon} \leqslant \lambda_{\max}^{1/2p_\varepsilon}$. Similarly, we have $r^{-\upsilon} \leqslant \rho_{\min}^{-\ell - L/4}$ and $r^{-\sigma} \leq (\lambda_{\max}\rho_{\min}^{-d})^{1/2}$. It follows that

$$\frac{\log\lambda_{\max}}{2p_\varepsilon \log r} \leqslant o_\varepsilon \leqslant \frac{\log\lambda_{\min}}{2p_\varepsilon \log r}, \qquad \upsilon \leqslant \frac{(\ell + L/4)\log\rho_{\min}}{\log r}, \qquad \sigma \leqslant \frac{\log\lambda_{\max} - d\log\rho_{\min}}{-2\log r}.$$

Put together, and noting that $\log r \leq \log\rho_{\max} < 0$, we obtain

$$\frac{\sigma(o_\varepsilon + \upsilon)}{o_\varepsilon + \sigma} < \frac{\sigma(o_\varepsilon + \upsilon)}{o_\varepsilon} \leqslant \left(\frac{d\log\rho_{\min}}{\log\lambda_{\max}} - 1\right)\left(\frac{\log\lambda_{\min} + 2p_\varepsilon(\ell + L/4)\log\rho_{\min}}{2\log r}\right)$$

$$\leqslant \left(\frac{d\log\rho_{\min}}{\log\lambda_{\max}} - 1\right)\left(\frac{\log\lambda_{\min} + 2p_\varepsilon(\ell + L/4)\log\rho_{\min}}{2\log\rho_{\max}}\right). \tag{6.10}$$

Since $(\rho_i^s)_{i \in \Lambda}$ is a probability vector, we have $\lambda_{\min} \leq \rho_{\max}^s$. Hence, $\log \lambda_{\min}/s \log \rho_{\max} \geq 1$. By (6.9), we obtain

$$\frac{d \log \rho_{\min}}{\log \lambda_{\max}} - 1 \leqslant \delta_0,$$

which implies that $\rho_{\min}^d \geq \lambda_{\max}^{1+\delta_0} \geq \lambda_{\min}^{1+\delta_0}$. This shows that

$$\frac{\log \rho_{\min}}{\log \rho_{\max}} \leq \frac{1 + \delta_0}{d} \frac{\log \lambda_{\min}}{\log \rho_{\max}}.$$

Combined with (6.9) and (6.10), and using that $\delta_0 \leq 1$ and $s \leq d$, this shows that

$$
\begin{aligned}
\frac{\sigma(o_\varepsilon + \upsilon)}{o_\varepsilon + \sigma} &< \left(\frac{d \log \rho_{\min}}{\log \lambda_{\max}} - 1 \right) \left(\frac{1}{2} + \frac{p_\varepsilon(\ell + L/4)(1 + \delta_0)}{d} \right) \frac{\log \lambda_{\min}}{\log \rho_{\max}} \\
&\leqslant s\delta_0 \left(\frac{1}{2} + \frac{p_\varepsilon(\ell + L/4)(1 + \delta_0)}{d} \right) \\
&\leqslant \delta_0 \left(\frac{d}{2} + 2p_\varepsilon(\ell + L/4) \right) \leqslant \frac{\kappa}{2},
\end{aligned}
$$

where the last inequality follows by definition of δ_0. This proves that $(1.4) \implies (6.5)$.

6.2. **Set up.** The remainder of the section is dedicated to the proof of Theorem 6.1.

Let U be an open set satisfying the open set condition for $\mathcal{F}$. We fix a smooth non-negative compactly supported function $\tilde{\psi} \leq 1$ on $\mathbb{R}^d$ such that $\tilde{\psi}$ is not identically zero and its support is contained in U. Let $d\mathbf{x}$ denote the Lebesgue measure on $\mathbb{R}^d$ and define a probability measure ν by

$$d\nu(\mathbf{x}) = \psi(\mathbf{x})d\mathbf{x}, \qquad \text{where} \qquad \psi = \frac{\tilde{\psi}}{\int \tilde{\psi} d\mathbf{x}}. \tag{6.11}$$

For convenience, we also set

$$C_\nu = \frac{1}{\int \tilde{\psi} \, d\mathbf{x}}.$$

Let P_λ be the operator introduced in Definition 2.1. Note that $f_{\omega *}\nu$ is supported on $f_\omega(\mathrm{supp}(\psi))$ for all $\omega \in \Lambda^*$ and therefore there exists R depending on $\mathcal{F}$ and ψ such that $P_\lambda^m(\nu)$ is supported in the ball $B(\mathbf{0}, R) \subset \mathbb{R}^d$ of radius R around the origin; cf. [Hu81, Thm. 1]. We fix a smooth non-negative function ψ_2 such that $\mathrm{supp}(\psi_2) \subset B(\mathbf{0}, 2R)$ and $\psi_2|_{B(\mathbf{0},R)} = 1$.

The implicit constants in our error terms will depend on ν through the C^ℓ-norm of its density and the size of its support and the choice of ψ_2. These in turn can be made to depend only on the set U (and hence only on $\mathcal{F}$) by choosing ψ suitably.

By Theorem 2.3 and Lemma 2.2, there exists a constant $C \geq 1$, depending only on μ, ν and r such that for all $m \in \mathbb{N}$ and all bounded Lipschitz functions φ on $\mathbb{R}^d$, we have

$$|\mu(\varphi) - P_\lambda^m(\nu)(\varphi)| \leqslant Cr^m \mathrm{Lip}(\varphi). \tag{6.12}$$

We also record the following identity which follows from the definition of P_λ:

$$P_\lambda^{m+n}(\nu)(\varphi) = P_\lambda^m(\nu)(P_\lambda^n(\varphi)). \tag{6.13}$$

6.3. **Invariance by the S-arithmetic random walk and Cauchy-Schwarz.** Denote by $S := S(\mathcal{F})$ the minimal set of primes, along with ∞, such that

$$a(\rho_i), k_i, u(b_i) \in \Gamma_S = \mathbf{G}\left(\mathbb{Z}[S^{-1}]\right), \quad i \in \Lambda.$$

We denote by $S_{\mathrm{f}} \subset S$ the set of finite places. We view Γ_S as being diagonally embedded as a lattice in $\mathbf{G}_S$. We denote by x_0 and x_S the identity cosets in $X_\infty(1)$ and X_S respectively.

Before we turn to the proof of Theorem 6.1, we recall that by (3.4) the functions in $\mathrm{B}_{\infty,\ell}^\infty(X_\infty(1))$ are in one-to-one correspondence with the K_{f}-invariant functions in $\mathrm{B}_{\infty,\ell}^\infty(X_S)$, where K_{f} denotes

the maximal compact-open subgroup of $\mathbf{G}_{\mathrm{f}}$; cf. Section 3.1. From now on, we identify φ with the corresponding K_{f}-invariant function in $\mathrm{B}^{\infty}_{\infty,\ell}(X_S)$.

For the proof of Theorem 6.1, we fix a word $\alpha \in \Lambda^*$, $t > 1$ and $\varphi \in \mathrm{B}^{\infty}_{\infty,\ell}(X_{\infty}(1))$, where ℓ is as in Proposition 5.1. We assume without loss of generality that φ is real-valued and

$$\int_{X_{\infty}(1)} \varphi = 0. \tag{6.14}$$

For simplicity, we also use h_{α} and x_{α} to denote $(h_{\alpha}, \mathrm{Id}) \in \mathbf{G}_S$ and $(h_{\alpha}, \mathrm{Id})\Gamma_S$ respectively, where Id denotes the identity element in $\mathbf{G}$. Recall the maps γ_{ω} defined in (4.2) for $\omega \in \Lambda^*$ and the operators $\mathcal{P}_{\lambda}$. Note that γ_{ω} and $a(t,1)$ commute. Therefore one obtains from (4.3) the following equality, which is key to our proof:

$$\gamma_{\omega}a(t,1)u(\mathbf{x},\mathbf{0})\gamma_{\omega}^{-1}u(b_{\omega},\mathbf{0}) = a(t,1)u(f_{\omega}(\mathbf{x}),\mathbf{0}).$$

Since $\gamma_{\omega}^{-1}u(b_{\omega},\mathbf{0}) \in \Gamma_S$, this implies the following key identity:

$$\gamma_{\omega}a(t,1)u(\mathbf{x},\mathbf{0})x_S = a(t,1)u(f_{\omega}(\mathbf{x}),\mathbf{0})x_S. \tag{6.15}$$

Given a word ω, we denote by $\alpha\omega$ the word obtained by concatenating ω to the end of α. We claim that

$$a(t,1)u\big(f_{\omega}(\mathbf{x}),\mathbf{0}\big)x_{\alpha} = (k_{\alpha}^{-1},\mathrm{Id})a(t/\rho_{\alpha},1)u\big(f_{\alpha\omega}(\mathbf{x}),\mathbf{0}\big)x_S. \tag{6.16}$$

Indeed, since $f_{\alpha}(\mathbf{y}) = \rho_{\alpha}O_{\alpha}\mathbf{y} + b_{\alpha}$, we see that

$$u(\mathbf{y})h_{\alpha} = k_{\alpha}^{-1}a(1/\rho_{\alpha})u(f_{\alpha}(\mathbf{y})).$$

Applying this identity with $\mathbf{y} = f_{\omega}(\mathbf{x})$, we obtain

$$u(f_{\omega}(\mathbf{x}),\mathbf{0})x_{\alpha} = (k_{\alpha}^{-1},\mathrm{Id})a(1/\rho_{\alpha},1)u(f_{\alpha\omega}(\mathbf{x}),\mathbf{0})x_S \tag{6.17}$$

and hence (6.16) follows from the fact that $a(t,1)$ and (k_{α},Id) commute.

Hence, we obtain the following generalization of (6.15):

$$a(t,1)u\big(f_{\omega}(\mathbf{x}),\mathbf{0}\big)x_{\alpha} \overset{(6.16)}{=} (k_{\alpha}^{-1},\mathrm{Id})a(t/\rho_{\alpha},1)u\big(f_{\alpha\omega}(\mathbf{x}),\mathbf{0}\big)x_S$$

$$\overset{(6.15)}{=} (k_{\alpha}^{-1},\mathrm{Id})\gamma_{\alpha\omega}a(t/\rho_{\alpha},1)u\big(\mathbf{x},\mathbf{0}\big)x_S$$

$$\overset{(6.15)}{=} (k_{\alpha}^{-1},\mathrm{Id})\gamma_{\alpha\omega}\gamma_{\alpha}^{-1}a(t/\rho_{\alpha},1)u\big(f_{\alpha}(\mathbf{x}),\mathbf{0}\big)x_S$$

$$\overset{(6.17)}{=} (k_{\alpha}^{-1},\mathrm{Id})\gamma_{\alpha\omega}\gamma_{\alpha}^{-1}(k_{\alpha},\mathrm{Id})a(t,1)u(\mathbf{x},\mathbf{0})x_{\alpha}.$$

As $\gamma_{\alpha\omega} = \gamma_{\alpha}\gamma_{\omega}$, we therefore get

$$\big((k_{\alpha}^{-1},\mathrm{Id})\gamma_{\alpha}\cdot\gamma_{\omega}\cdot\gamma_{\alpha}^{-1}(k_{\alpha},\mathrm{Id})\big)\cdot a(t,1)u(\mathbf{x},\mathbf{0})x_{\alpha} = a(t,1)u(f_{\omega}(\mathbf{x}),\mathbf{0})x_{\alpha}. \tag{6.18}$$

Recall the operators $\alpha\cdot\mathcal{P}_{\lambda}$ in (4.7). Define $\Phi : \mathbb{R}^d \to \mathbb{C}$ by $\Phi(\mathbf{x}) = \varphi(a(t,1)u(\mathbf{x},\mathbf{0})x_{\alpha})$. Using the above identification and (6.18), for all $n \in \mathbb{N}$ we have

$$P_{\lambda}^n(\Phi)(\mathbf{x}) = \sum_{\omega\in\Lambda^n} \lambda_{\omega}(\Phi \circ f_{\omega})(\mathbf{x}) = \sum_{\omega\in\Lambda^n} \lambda_{\omega}\varphi(a(t,1)u(f_{\omega}(\mathbf{x}),\mathbf{0})x_{\alpha})$$

$$= \sum_{\omega\in\Lambda^n} \lambda_{\omega}\varphi\big((k_{\alpha}^{-1},\mathrm{Id})\gamma_{\alpha}\gamma_{\omega}\gamma_{\alpha}^{-1}(k_{\alpha},\mathrm{Id})a(t,1)u(\mathbf{x},\mathbf{0})x_{\alpha}\big)$$

$$= (\alpha\cdot\mathcal{P}_{\lambda})^n(\varphi)(a(t,1)u(\mathbf{x},\mathbf{0})x_{\alpha}).$$

In particular, for every probability measure ν on $\mathbb{R}^d$, $m, n \in \mathbb{N}$, and $t > 0$, one has

$$\int \varphi\,(a(t)u(\mathbf{x})x_{\alpha})\ \mathrm{d}P_{\lambda}^{m+n}(\nu)(\mathbf{x}) = \int (\alpha\cdot\mathcal{P}_{\lambda})^n(\varphi)\,(a(t,1)u(\mathbf{x},\mathbf{0})x_{\alpha})\ \mathrm{d}P_{\lambda}^m(\nu)(\mathbf{x}). \tag{6.19}$$

We fix some $m, n \in \mathbb{N}$ to be chosen towards the end of the proof.

Note that we may regard $\varphi \circ a(t)$ as a function on $\mathbb{R}^d$ by restriction to the closed orbit $U(\mathbb{R})x_\alpha$. Moreover, since conjugation by $a(t)$ expands $U(\mathbb{R})$ by a factor of t, we have

$$\mathrm{Lip}\left(\varphi \circ a(t)|_{U(\mathbb{R})x_\alpha}\right) \leqslant t \cdot \mathrm{Lip}\left(\varphi|_{U(\mathbb{R})a(t)x_\alpha}\right) \ll t \cdot \mathcal{S}_{\infty,\ell}(\varphi). \tag{6.20}$$

Combined with the estimate in (6.12), (6.20) implies

$$\int \varphi\left(a(t)u(\mathbf{x})x_\alpha\right)\,\mathrm{d}\mu = \int \varphi\left(a(t)u(\mathbf{x})x_\alpha\right)\,\mathrm{d}P_\lambda^{m+n}(\nu) + O\left(r^{m+n}t \cdot \mathcal{S}_{\infty,\ell}(\varphi)\right).$$

It then follows from Equation (6.19) that

$$\int \varphi\left(a(t)u(\mathbf{x})x_\alpha\right)\,\mathrm{d}\mu = \int \Theta(\mathbf{x})\,\mathrm{d}P_\lambda^m(\nu)(\mathbf{x}) + O\left(r^{m+n}t \cdot \mathcal{S}_{\infty,\ell}(\varphi)\right), \tag{6.21}$$

where for simplicity, we write

$$\Theta(\mathbf{x}) := (\alpha \cdot \mathcal{P}_\lambda)^n(\varphi)\left(a(t,1)u(\mathbf{x},\mathbf{0})x_\alpha\right).$$

Using the definition of P_λ^m in (2.5) and the measure $\mathrm{d}\nu(\mathbf{x}) = \psi(\mathbf{x})\mathrm{d}\mathbf{x}$, we obtain

$$\int \Theta(\mathbf{x})\,\mathrm{d}P_\lambda^m(\nu)(\mathbf{x}) = \int \sum_{\omega \in \Lambda^m} \lambda_\omega \Theta(f_\omega(\mathbf{x}))\psi(\mathbf{x})\,\mathrm{d}\mathbf{x}.$$

For each $m \in \mathbb{N}$ define

$$B_m := \mathrm{supp}\left(P_\lambda^m(\nu)\right).$$

Applying a change of variable $\mathbf{x} \mapsto f_\omega^{-1}(\mathbf{x})$ and noting that $\mathrm{d}f_\omega^{-1}(\mathbf{x}) = \rho_\omega^{-d}\mathrm{d}\mathbf{x}$ by invariance of the Lebesgue measure under rotations and translations, we obtain

$$\int \Theta(\mathbf{x})\,\mathrm{d}P_\lambda^m(\nu)(\mathbf{x}) = \int \Theta(\mathbf{x})\mathbf{1}_{B_m}(\mathbf{x})\,\mathrm{d}P_\lambda^m(\nu)(\mathbf{x})$$

$$= \int \sum_{\omega \in \Lambda^m} \lambda_\omega \Theta(f_\omega(\mathbf{x}))\mathbf{1}_{B_m}(f_\omega(\mathbf{x}))\psi(\mathbf{x})\,\mathrm{d}\mathbf{x}$$

$$= \int \Theta(\mathbf{x})\mathbf{1}_{B_m}(\mathbf{x})\sum_{\omega \in \Lambda^m} \lambda_\omega \rho_\omega^{-d}\psi(f_\omega^{-1}(\mathbf{x}))\,\mathrm{d}\mathbf{x}.$$

Hence, Cauchy-Schwarz gives

$$\left(\int \Theta(\mathbf{x})\,\mathrm{d}P_\lambda^m(\mathbf{x})\right)^2 \leqslant \underbrace{\left(\int_{B_m} \Theta(\mathbf{x})^2\,\mathrm{d}\mathbf{x}\right)}_{\text{Horospherical term}}\underbrace{\left(\int \Big(\sum_{\omega \in \Lambda^m} \lambda_\omega \rho_\omega^{-d}(\psi \circ f_\omega^{-1})(\mathbf{y})\Big)^2\,\mathrm{d}\mathbf{y}\right)}_{\text{Mass term}}. \tag{6.22}$$

6.4. **The Open Set Condition and the Mass Term.** The open set condition and the assumption that $\mathrm{supp}(\psi) \subseteq U$ imply that whenever $\alpha \neq \omega \in \Lambda^m$, $\psi \circ f_\omega^{-1}$ and $\psi \circ f_\alpha^{-1}$ have disjoint supports. Thus, expanding the squared sum in the mass term yields

$$\left(\sum_{\omega \in \Lambda^m} \lambda_\omega \rho_\omega^{-d}\psi(f_\omega^{-1}(\mathbf{x}))\right)^2 = \sum_{\omega \in \Lambda^m} \lambda_\omega^2 \rho_\omega^{-2d}\psi(f_\omega^{-1}(\mathbf{x}))^2.$$

Applying the change of variable $\mathbf{x} \mapsto f_\omega(\mathbf{x})$, the mass term becomes

$$\int \left(\sum_{\omega \in \Lambda^m} \lambda_\omega \rho_\omega^{-d}\psi(f_\omega^{-1}(\mathbf{x}))\right)^2\,\mathrm{d}\mathbf{x} = \sum_{\omega \in \Lambda^m} \lambda_\omega^2 \rho_\omega^{-d}\int \psi(\mathbf{x})^2\,\mathrm{d}\mathbf{x}. \tag{6.23}$$

Recall that $\psi = C_\nu\tilde{\psi}$, where $\tilde{\psi} \leq 1$. Hence, we have $\psi^2 \leq C_\nu^2\tilde{\psi}$. It follows that

$$\int \psi^2\,\mathrm{d}\mathbf{x} \leq C_\nu. \tag{6.24}$$

Applying Lemma 2.4 with $\tau_i = \lambda_i \rho_i^{-d}$ yields

$$\sum_{\omega \in \Lambda^m} \lambda_\omega^2 \rho_\omega^{-d} = \left(\sum_{i \in \Lambda} \lambda_i^2 \rho_i^{-d} \right)^m.$$

Hence, we obtain the following bound on the mass term

$$\int \left(\sum_{\omega \in \Lambda^m} \lambda_\omega \rho_\omega^{-d} \psi(f_\omega^{-1}(\mathbf{x})) \right)^2 \, d\mathbf{x} \leqslant C_\nu \left(\sum_{i \in \Lambda} \lambda_i^2 \rho_i^{-d} \right)^m = C_\nu r^{-2\sigma m}. \tag{6.25}$$

Remark 6.5. It is worth noting that the sum on the right side of (6.25) is always ≥ 1. Indeed, if s is the Hausdorff dimension of $\mathcal{K}$, then the open set condition implies that $\sum_{i \in \Lambda} \rho_i^s = 1$; cf. [Mo46]. Moreover, since $s \leq d$ and $\rho_i \leq 1$, we have $\rho_i^d \leq \rho_i^s$ for each $i \in \Lambda$. Note further that, since $(\lambda_i)_{i \in \Lambda}$ is a probability vector, the sum on the right hand side of (6.25) can be interpreted as an expected value with respect to the probability measure determined by $(\lambda_i)_{i \in \Lambda}$. Jensen's inequality thus implies

$$\sum_{i \in \Lambda} \lambda_i^2 \rho_i^{-d} \geqslant \frac{1}{\sum_{i \in \Lambda} \lambda_i \lambda_i^{-1} \rho_i^d} \geqslant \frac{1}{\sum_{i \in \Lambda} \rho_i^s} = 1.$$

Moreover, the inequality is strict unless $s = d$ and $\lambda_i = \rho_i^d$ for each $i \in \Lambda$.

6.5. **Uniform Spectral Gaps and the Horospherical Term.** Recall that B_k denotes the support of the measure $P_\lambda^k(\nu)$ for $k \in \mathbb{N}$. To estimate the horospherical term in (6.22), recall that for all $k \in \mathbb{N}$,

$$\mathbf{1}_{B_k}(\mathbf{x}) \leqslant \psi_2(\mathbf{x}).$$

Let $d\nu_2(\mathbf{x}) := \psi_2(\mathbf{x}) d\mathbf{x}$. By positivity of Θ^2 it follows that

$$\int_{B_m} \Theta(\mathbf{x})^2 \, d\mathbf{x} \leqslant \int \Theta(\mathbf{x})^2 \, d\nu_2(\mathbf{x}). \tag{6.26}$$

By Lemma 4.2, since φ is K_f-invariant, the function $(\alpha \cdot \mathcal{P}_\lambda)^n(\varphi)$ is invariant under a compact open subgroup $K_f[N]$ of K_f.

We wish to apply Proposition 5.1. By the double coset decomposition (3.5), $(\alpha \cdot \mathcal{P}_\lambda)^n(\varphi)$ can be regarded as a function on a finite, disjoint union of copies of $X_\infty(N)$, where $\Gamma(N)$ denotes the congruence lattice of level N. We note that if $g = a(t)u(\mathbf{x})h_\alpha$, then

$$K_f[N]g\Gamma_S = K_f[N](g, e_f)\Gamma_S = g.K_f[N](e_\infty, e_f)\Gamma_S$$

is identified with the point $g\Gamma(N) \in X_\infty(N)$ in the copy of $X_\infty(N)$ containing the identity double coset in $K_f[N] \backslash \mathbf{G}_S / \Gamma_S$. Moreover, since $g \in \mathbf{G}_\infty^+$, $g\Gamma(N)$ is contained in the connected component containing the identity coset inside this copy of $X_\infty(N)$. In particular, the integral on the right side of (6.26) can be seen as an integral over this component in this copy of $X_\infty(N)$. Let $\Gamma^+(N) = \mathbf{G}_\infty^+ \cap \Gamma(N)$ and $V_N = \sqrt{[\Gamma(1) : \Gamma(N)]}$. Letting D denote the total mass of ν_2, Proposition 5.1 implies

$$\frac{1}{D} \int (\alpha \cdot \mathcal{P}_\lambda)^n(\varphi)^2 \big(a(t)u(\mathbf{x})h_\alpha \Gamma(N) \big) \, d\nu_2(\mathbf{x}) = \int (\alpha \cdot \mathcal{P}_\lambda)^n(\varphi)^2 \, dm_{\mathbf{G}_\infty^+ / \Gamma^+(N)}$$

$$+ O\left(V_N \mathcal{S}_{\infty, \ell} \big((\alpha \cdot \mathcal{P}_\lambda)^n(\varphi)^2 \big) t^{-\kappa} \max\left\{ 1, \mathrm{inj}_{\Gamma(N)}(h_\alpha \Gamma(N))^{-C} \right\} \right), \tag{6.27}$$

where $m_{\mathbf{G}_\infty^+ / \Gamma^+(N)}$ is the $\mathbf{G}_\infty^+$-invariant Haar probability measure. Here, we used the fact that $\mathrm{Lip}(\cdot) + \mathcal{S}_{2,\ell}(\cdot) \ll \mathcal{S}_{\infty, \ell}(\cdot)$.

First, we bound the error terms in (6.27). Since $\Gamma(N)$ is a subgroup of $\Gamma(1)$, the injectivity radius at $h_\alpha \Gamma(N)$ in $X_\infty(N)$ is bounded from below by the injectivity radius at $x_\alpha = h_\alpha \Gamma(1)$ in $X_\infty(1)$. It

follows that $\mathrm{inj}_{\Gamma(N)}(h_\alpha \Gamma(N))^{-C} \leq \mathrm{inj}_{\Gamma(1)}(x_\alpha)^{-C}$. Moreover, since $k_\alpha^{-1} u(b_\alpha)$ is uniformly bounded in $\mathbf{G}_\infty$, we can find $C_0 \geq 1$ such that $\mathrm{inj}_{\Gamma(1)}(x_\alpha)^{-C} \ll \rho_\alpha^{-C_0}$. Noting that $\mathrm{inj}_{\Gamma(1)}^{-C} \gg 1$, we get

$$\max\left\{1, \mathrm{inj}_{\Gamma(N)}(h_\alpha \Gamma(N))^{-C}\right\} \ll \mathrm{inj}_{\Gamma(1)}(x_\alpha)^{-C} \ll \rho_\alpha^{-C_0}. \tag{6.28}$$

In order to bound $\mathcal{S}_{\infty,\ell}((\alpha \cdot \mathcal{P}_\lambda)^n(\varphi)^2)$, note that the Archimedean component of

$$(k_\alpha^{-1}, \mathrm{Id})\gamma_\alpha \gamma_\omega \gamma_\alpha^{-1}(k_\alpha, \mathrm{Id})$$

is equal to $k_\omega a(\rho_\omega)$; cf. (4.2) for the definition of γ_ω. Moreover, by M_∞-invariance of our chosen norm on the Lie algebra of $\mathbf{G}_\infty$, we have $\|\mathrm{Ad}(k_\omega a(\rho_\omega))\|_\infty = \rho_\omega^{-1}$. Hence, by Lemma 3.3, we have

$$\mathcal{S}_{\infty,\ell}((\alpha \cdot \mathcal{P}_\lambda)^n(\varphi)^2) \ll \mathcal{S}_{\infty,\ell}((\alpha \cdot \mathcal{P}_\lambda)^n(\varphi))^2 \leqslant \left(\sum_{\omega \in \Lambda^n} \lambda_\omega \|\mathrm{Ad}(k_\omega a(\rho_\omega))\|_\infty^\ell \mathcal{S}_{\infty,\ell}(\varphi)\right)^2$$

$$= \mathcal{S}_{\infty,\ell}(\varphi)^2 \left(\sum_{\omega \in \Lambda^n} \lambda_\omega \rho_\omega^{-\ell}\right)^2. \tag{6.29}$$

By Lemma 4.2, we have

$$V_N \ll \rho_\alpha^{-A/2} \rho_{\min}^{-nL/2}. \tag{6.30}$$

The next step is to bound the main term in (6.27) by applying the estimate on the spectral gap of $\alpha \cdot \mathcal{P}_\lambda$, Proposition 4.3. To this end, we wish to lift the integral in the main term from $\mathbf{G}_\infty^+/\Gamma^+(N)$ to X_S. We do so via a second application of the double coset decomposition (3.5). Let

$$C(S,d) = [\mathbf{G}_S : \mathbf{G}_S^+].$$

Then, $C(S,d)$ is an upper bound on the number of $\mathbf{G}_\infty^+$ orbits on the double quotient $K_{\mathrm{f}}[N]\backslash X_S$ by Proposition B.13. The Haar measure on $K_{\mathrm{f}}[N]\backslash X_S$ is the convex combination (with equal weights) of the $\mathbf{G}_\infty^+$-invariant probability measures on each of these orbits. Hence, it follows by positivity that

$$\int (\alpha \cdot \mathcal{P}_\lambda)^n(\varphi)^2 \, dm_{\mathbf{G}_\infty^+/\Gamma^+(N)} \leqslant C(S,d) \int_{K_{\mathrm{f}}[N]\backslash X_S} (\alpha \cdot \mathcal{P}_\lambda)^n(\varphi)^2 = C(S,d) \int_{X_S} (\alpha \cdot \mathcal{P}_\lambda)^n(\varphi)^2. \tag{6.31}$$

In the second equality, we used the invariance of $(\alpha \cdot \mathcal{P}_\lambda)^n(\varphi)^2$ by $K_{\mathrm{f}}[N]$.

Recall the constant p provided by Proposition 4.4 so that $\|\xi_{\mathbf{G}}\|_{\ell^{p+\varepsilon}(\Gamma_S)} < \infty$, for all $\varepsilon > 0$. Fix some $\varepsilon > 0$ so that our hypothesis (6.5) holds. Let θ_ε denote the Hölder conjugate of $(p+\varepsilon)/\epsilon(d)$, where $\epsilon(d)$ is given by (3.11).

Note that the $\mathbf{G}_\infty$-invariant probability measure on $X_\infty(1)$ agrees with the push-forward of the $\mathbf{G}_S$-invariant probability measure under the canonical projection with respect to K_{f}. It thus follows from Lemma 3.5 that $\varphi \in \mathrm{L}_{00}^2(X_S)^{K_{\mathrm{f}}}$. Hence, Proposition 4.3 implies that

$$\int_{X_S} (\alpha \cdot \mathcal{P}_\lambda)^n(\varphi)^2 \ll \mathcal{S}_{\infty,\ell}(\varphi)^2 \left(\sum_{i \in \Lambda} \lambda_i^{q_\varepsilon}\right)^{2n/q_\varepsilon} \|\xi_{\mathbf{G}}^{\epsilon(d)}\|_{\ell^{(p+\varepsilon)/\epsilon(d)}}$$

$$\ll_{\varepsilon,S,d} \mathcal{S}_{\infty,\ell}(\varphi)^2 \left(\sum_{i \in \Lambda} \lambda_i^{q_\varepsilon}\right)^{2n/q_\varepsilon}, \tag{6.32}$$

where $q_\varepsilon = 2\theta_\varepsilon/(\theta_\varepsilon + 1) > 1$.

Combining all the above estimates, we obtain the following bound on the horospherical term:

$$\text{Horospherical term} \ll_{\varepsilon,\nu,S,d} \mathcal{S}_{\infty,\ell}(\varphi)^2 \rho_\alpha^{-A/2-C_0}\left(\left(\sum_{i \in \Lambda} \lambda_i^{q_\varepsilon}\right)^{2n/q_\varepsilon} + t^{-\kappa}\rho_{\min}^{-nL/2}\left(\sum_{\omega \in \Lambda^n} \lambda_\omega \rho_\omega^{-\ell}\right)^2\right).$$

By Lemma 2.4, applied with $\tau_i = \rho_i^{-\ell}$, we have $\sum_{\omega \in \Lambda^n} \lambda_\omega \rho_\omega^{-\ell} = (\sum_{i \in \Lambda} \lambda_i \rho_i^{-\ell})^n$. Hence, recalling the constants σ, o_ε, and υ defined in (6.4), we see that the last factor in the above estimate becomes $r^{2no_\epsilon} + t^{-\kappa} r^{-2n\upsilon}$. For convenience, let $o = o_\varepsilon$ and define τ by the equation

$$t = r^{-\tau}.$$

Let $A_* = (C_0 + A/2)/2$. By combining the above estimate on the horospherical term with (6.22) and (6.25), the estimate in (6.21) becomes

$$\int \varphi(a(t)u(\mathbf{x})x_\alpha) \, d\mu \ll_{\nu,\varepsilon,S,d} \mathcal{S}_{\infty,\ell}(\varphi) \rho_\alpha^{-A_*} \left(r^{m+n-\tau} + r^{-\sigma m} \sqrt{r^{2no} + r^{\kappa\tau - 2n\upsilon}} \right). \tag{6.33}$$

To balance these rates, we choose n to be the largest natural number so that $2no \leq \kappa\tau - 2n\upsilon$. We then choose m to be the largest integer satisfying $m + n - \tau \leq -\sigma m + on$. Hence, n and m are given by

$$n = \left\lfloor \frac{\kappa\tau}{2(o+\upsilon)} \right\rfloor, \qquad m = \left\lfloor \frac{1}{1+\sigma} \left(\tau + (o-1) \left\lfloor \frac{\kappa\tau}{2(o+\upsilon)} \right\rfloor \right) \right\rfloor. \tag{6.34}$$

Note that $n \to \infty$ as $t \to \infty$. In Lemma 6.6 below, we show that $m \to \infty$ as $t \to \infty$ under our hypotheses. In particular, $m \geq 0$ for all large enough t.

With these choices, the right-hand side of (6.33) tends to 0 as $t \to \infty$ whenever

$$\delta := \frac{\kappa}{2(o+\upsilon)} + \frac{1}{1+\sigma} \left(1 + \frac{(o-1)\kappa}{2(o+\upsilon)} \right) - 1 > 0. \tag{6.35}$$

Rearranging, we obtain that $\delta > 0$ if and only if

$$\frac{2\sigma(o+\upsilon)}{o+\sigma} < \kappa,$$

which is satisfied by our hypothesis. Thus, the constant δ satisfies the conclusion of the theorem.

Lemma 6.6. *Our choice of m in (6.34) implies that $m \to \infty$ as $t \to \infty$.*

Proof. It suffices to show that $1 + \frac{(o-1)\kappa}{2(o+\upsilon)} > 0$. Suppose not and note that this implies that $2o + o\kappa \leq \kappa - 2\upsilon$. In particular, this implies that $\kappa > 2\upsilon$, since $(\kappa + 2)o > 0$. On the other hand, by Jensen's inequality, we have

$$r^{-\upsilon} = \rho_{\min}^{-L/4} \sum_{i \in \Lambda} \lambda_i \rho_i^{-\ell} \geqslant \rho_{\min}^{-L/4} \left(\sum_{i \in \Lambda} \lambda_i \rho_i \right)^{-\ell} = \rho_{\min}^{-L/4} r^{-\ell} \geqslant r^{-\ell}.$$

Since $\ell = d(d+1)/2$, cf. (6.2), it follows that $\upsilon \geq \frac{d(d+1)}{2}$. We thus get that $\kappa > d(d+1)$, contrary to our assumption that $\kappa \leq d(d+1)$ in (6.3). Thus, we conclude that $1 + \frac{\kappa(o-1)}{2(o+\upsilon)} > 0$ as desired. $\square$

7. Equidistribution of Random Walks

The goal of this section is to prove Theorem D. The argument is similar to the proof of Theorem 6.1 and hence we omit some of the details. The main difference is that we appeal to the equidistribution of translates of rational points on pieces of horocycles in place of the equidistribution of absolutely continuous measures.

The following is the precise substitute for Proposition 5.1 needed for the proof.

Proposition 7.1. *There exist $\sigma > 0, \ell \in \mathbb{N}$ and $c \geq 1$ such that the following holds. For every $p, m, N \in \mathbb{N}$, and $\varphi \in \mathrm{B}^\infty_{\infty,\ell}(X_\infty(N))$, the following holds for all $x \in X_\infty(N)$:*

$$p^{-m} \sum_{0 \le k < p^m} \varphi\left(a(p^m)u(kp^{-m})x\right)$$

$$= \int \varphi \, \mathrm{dm}_{\mathbf{G}_\infty^+ \cdot x} + O\left(\sqrt{[\Gamma(1):\Gamma(N)]}\mathcal{S}_{\infty,\ell}(\varphi)\max\{1, \mathrm{inj}_{\Gamma(N)}(x)^{-c}\} \cdot p^{-\sigma m}\right).$$

Proof. This statement is obtained in [ELS20, Prop. 5.3] for x being the identity coset. We outline the needed adjustments to the proof and omit the details. The generalization to points with non-periodic horocycle orbits follows the same reasoning needed when considering pieces of periodic orbits instead of the full orbit in *loc. cit.* The explicit dependence on N follows by using Proposition 5.1 in place of [ELS20, Eq. (16)].   □

Let $\mathcal{F}$ be a missing digit IFS, cf. Definition A.1, with digit set $\Lambda \subseteq \Delta := \{0, \ldots, p-1\}$ for some $p \in \mathbb{N}$. We will assume that p is an odd prime to obtain the best convergence rate available through our methods. The same argument works for general $p \in \mathbb{N}$ at the cost of worsening the equidistribution rate via the use of Proposition 4.3 instead of Proposition A.3.

Denote by s the Hausdorff dimension of the attractor $\mathcal{K}$ of $\mathcal{F}$. Let $0 < \varrho_0 < 1$ be a small parameter whose value is determined at the end of the proof and assume that $s > 1 - \varrho_0$.

Let λ be the uniform probability vector on Λ. Then, $\lambda_i = |\Lambda|^{-1} = p^{-s}$ for all $i \in \Lambda$ [Mo46]. Let S be the set of places for $\mathbb{Q}$ consisting of ∞ along with the prime divisors of p. Fix $\varphi \in \mathrm{B}_{\infty,\ell}^\infty(X_\infty(1))$ with integral 0.

Fix some $\alpha \in \Lambda^*$ and set $h_\alpha = a(\rho_\alpha^{-1})u(b_\alpha)$. As before, we will abuse notation and denote by x_α both the coset of h_α in $X_\infty(N)$ and of $(h_\alpha, e_{\mathrm{f}})$ in X_S. Recall the measure ν defined in (1.8). We shall show that for some $\kappa_0 > 0$, we have

$$\int \varphi \, \mathrm{d}(\nu^{*n} * \delta_{x_\alpha}) = O_{\varphi,\alpha,p}(p^{-\kappa_0 n}). \tag{7.1}$$

The image of ν in $X_\infty(1)$, also denoted ν, under the isomorphism $X_\infty(1) \cong \mathrm{SL}_2(\mathbb{R})/\mathrm{SL}_2(\mathbb{Z})$, cf. Lemma 3.1, satisfies

$$\int \varphi \, \mathrm{d}(\nu^{*n} * \delta_{x_\alpha}) = \sum_{\omega \in \Lambda^n} \lambda_\omega \varphi(a(t_n)u(b_\omega)x_\alpha),$$

where $t_n = p^n$. Recall that $b_\omega = f_\omega(\mathbf{0})$. Then, denoting by $\delta_{\mathbf{0}}$ the Dirac mass at $\mathbf{0}$, we observe that

$$\sum_{\omega \in \Lambda^n} \lambda_\omega \delta_{b_\omega} = P_\lambda^n(\delta_{\mathbf{0}}).$$

Let $0 < \varepsilon < 1/2$ be a parameter to be chosen later and define

$$k = \lfloor \varepsilon n \rfloor, \qquad m = n - k. \tag{7.2}$$

Applying the key identity (6.19) with $\nu = \delta_{\mathbf{0}}$, we obtain

$$\sum_{\omega \in \Lambda^n} \lambda_\omega \varphi(a(t_n)u(b_\omega)x_\alpha) = \int \varphi(a(t_n)u(\mathbf{x})x_\alpha) \, \mathrm{d}P_\lambda^n(\delta_{\mathbf{0}})(\mathbf{x})$$

$$= \int (\alpha \cdot \mathcal{P}_\lambda)^k(\varphi)(a(t_n)u(\mathbf{x})x_\alpha) \, \mathrm{d}P_\lambda^m(\delta_{\mathbf{0}})(\mathbf{x}). \tag{7.3}$$

Denote by $\chi_\mathcal{K}$ the indicator function of $\mathcal{K}$. For $\omega \in \Delta^m$, we denote by b_ω those rationals in $[0, 1)$ with denominator p^m and a numerator whose digit expansion mod p is given by ω. In particular, we have $\{b_\omega : \omega \in \Delta^m\} = \{k/p^m : 0 \le k < p^m\}$ and $\mathcal{K} \cap \{b_\omega : \omega \in \Delta^m\} = \{b_\omega : \omega \in \Lambda^m\}$.

Hence, the last integral in (7.3) can be rewritten as follows:

$$\int (\alpha \cdot \mathcal{P}_\lambda)^k(\varphi)(a(t_n)u(\mathbf{x})x_\alpha) \, \mathrm{d}P_\lambda^m(\delta_{\mathbf{0}}) = p^{-sm} \sum_{\omega \in \Delta^m} (\alpha \cdot \mathcal{P}_\lambda)^k(\varphi)(a(t_n)u(b_\omega)x_\alpha)\chi_\mathcal{K}(b_\omega).$$

By Cauchy-Schwarz, we get

$$\left(\sum_{\omega \in \Delta^m} (\alpha \cdot \mathcal{P}_\lambda)^k(\varphi)(a(t_n)u(b_\omega)x_\alpha)\chi_\mathcal{K}(b_\omega) \right)^2$$

$$\leqslant p^{sm} \sum_{\omega \in \Delta^m} (\alpha \cdot \mathcal{P}_\lambda)^k(\varphi)^2(a(t_n)u(b_\omega)x_\alpha)$$

$$= p^{sm} \sum_{\omega \in \Delta^m} \left((\alpha \cdot \mathcal{P}_\lambda)^k(\varphi)^2 \circ a(t_k)\right)(a(t_m)u(b_\omega)x_\alpha).$$

Arguing as in the proof of Theorem 6.1 using Lemma 4.2, we can regard the last sum above as taking place in $X_\infty(N)$, for a suitable N. Let $\Phi = (\alpha \cdot \mathcal{P}_\lambda)^k(\varphi)^2 \circ a(t_k)$. Then the last sum is obtained by summing the values of Φ over the rational points with denominator p^m on the horocycle orbit through the basepoint x_α after expansion by $a(p^m)$.

By effective equidistribution of those points on $X_\infty(N)$, cf. Proposition 7.1, and arguing as for (6.27) and (6.31), we obtain

$$p^{-m} \sum_{\omega \in \Delta^m} (\alpha \cdot \mathcal{P}_\lambda)^k(\varphi)^2 \circ a(t_k)(a(t_m)u(b_\omega)x_\alpha)$$

$$\ll_p \int_{X_S} (\alpha \cdot \mathcal{P}_\lambda)^k(\varphi)^2 + O\left(V_N \mathcal{S}_{\infty,\ell}\left((\alpha \cdot \mathcal{P}_\lambda)^k(\varphi)^2 \circ a(t_k)\right) t_m^{-\sigma} \max\left\{1, \mathrm{inj}_{\Gamma(N)}(x_\alpha)^{-c}\right\}\right),$$

where $V_N = \sqrt{\Gamma(1):\Gamma(N)}$, and we used the invariance of the Haar measure on X_S under $a(t_k)$. The dependence on p in the implied constant is through the index $[\mathbf{G}_S : \mathbf{G}_S^+]$.

Using Lemma 3.3, we can estimate the error term as follows, cf. (6.29):

$$\mathcal{S}_{\infty,\ell}\left((\alpha \cdot \mathcal{P}_\lambda)^k(\varphi)^2 \circ a(t_k)\right) \ll t_k^\ell \cdot \mathcal{S}_{\infty,\ell}((\alpha \cdot \mathcal{P}_\lambda)^k(\varphi)^2) \ll t_k^{3\ell} \cdot \mathcal{S}_{\infty,\ell}(\varphi)^2.$$

Let $\theta = \min\{s, 25/32\}$. By taking $\varrho_0 < 1/2$, we have $s > 1/2$. Hence, applying the spectral gap estimate for the operators $\alpha \cdot \mathcal{P}_\lambda$, Proposition A.3, we obtain

$$\int_{X_S} (\alpha \cdot \mathcal{P}_\lambda)^k(\varphi)^2 \ll_\varepsilon p^{-(\theta-\varepsilon)k} \mathcal{S}_{\infty,\ell}(\varphi)^2 \leqslant p^{-(1/2-\varepsilon)k} \mathcal{S}_{\infty,\ell}(\varphi)^2.$$

By Lemma 4.2 and arguing as in (6.30), we have $V_N \ll p^{3|\alpha|+3k/2}$, where $|\alpha|$ is the integer satisfying $\alpha \in \Lambda^{|\alpha|}$. Combining all the estimates, along with (7.2), and the facts $\mathrm{inj}_{\Gamma(1)} \leqslant \mathrm{inj}_{\Gamma(N)}$ and $\mathrm{inj}_{\Gamma(1)}^{-1} \gg 1$, we obtain

$$\int \varphi \, \mathrm{d}(\nu^{*n} * \delta_{x_\alpha}) \ll_{p,\varepsilon} p^{3|\alpha|/2} p^{\frac{(1-s)m}{2}} \sqrt{p^{-(1/2-\varepsilon)k} + p^{(3\ell+3/2)k-\sigma m}} \, \mathcal{S}_{\infty,\ell}(\varphi) \, \mathrm{inj}_{\Gamma(1)}(x_\alpha)^{-c/2}$$

$$\ll_p p^{3|\alpha|/2} p^{\frac{(1-s)(1-\varepsilon)n}{2}} \sqrt{p^{-(1/2-\varepsilon)\varepsilon n+1} + p^{((3\ell+3/2+\sigma)\varepsilon-\sigma)n}} \, \mathcal{S}_{\infty,\ell}(\varphi) \, \mathrm{inj}_{\Gamma(1)}(x_\alpha)^{-c/2}.$$

Setting $-(1/2-\varepsilon)\varepsilon = (3\ell+3/2+\sigma)\varepsilon-\sigma$, yields a quadratic equation in ε with one positive root given as follows: letting $o = 3\ell+2+\sigma$, then

$$\varepsilon = \frac{o - \sqrt{o^2 - 4\sigma}}{2}.$$

Note that we may assume that σ is small enough so that $\varepsilon < 1/2$. Hence, the above estimate becomes

$$\int \varphi \, \mathrm{d}(\nu^{*n} * \delta_{x_\alpha}) \ll_{p,\varepsilon} p^{3|\alpha|/2} p^{\frac{((1-s)(1-\varepsilon)-(1/2-\varepsilon)\varepsilon)n}{2}} \, \mathcal{S}_{\infty,\ell}(\varphi) \, \mathrm{inj}_{\Gamma(1)}(x_\alpha)^{-c/2}.$$

Then, noting that α is fixed, the estimate above tends to 0 when

$$1 - s < \varrho_0 := \frac{(1/2 - \varepsilon)\varepsilon}{1 - \varepsilon}.$$

This concludes the proof.

Remark 7.2. The above proof extends readily to IFS in higher dimensions which generate products of copies of missing digit sets. Moreover, with some additional effort, one can handle non-uniform probability vectors. The method is limited however to such special types of IFS, compared to the ones addressed by Theorem 6.1. First, by Theorem 2.3, the measures $P_\lambda^n(\delta_{\mathbf{0}})$ converge exponentially fast, with speed p^{-n}, towards the Hausdorff measure on $\mathcal{K}$. However, the Lipschitz constant of the functions $\varphi(a(t_n)u(\cdot)x_\alpha)$ is roughly p^n, which prevents us from deducing Theorem D from Theorem 6.1. This is also the reason we appeal to the equidistribution of rational points, Proposition 7.1, instead. For a general IFS, there is no natural analog of Proposition 7.1 for a "completed set" of the translation vectors $\{b_\omega : \omega \in \Lambda^n\}$, i.e., an analog of the full set of rational points of denominator p^n.

8. REDUCTION TO DYNAMICS

In this section, which largely follows [KM99], we set up some notation and background which allows us to connect ψ-approximability to cusp excursions, i.e., to homogeneous dynamics. The connection between Khintchine's Theorem on $\mathbb{R}$ and the geodesic flow on the modular surface was first observed in [Su82], where it was attributed to David Kazhdan. Throughout the remainder of the article, we let

$$G = \mathrm{SL}_{d+1}(\mathbb{R}), \qquad \Gamma = \mathrm{SL}_{d+1}(\mathbb{Z}).$$

Lemma 8.1 (Lemma 8.3, [KM99]). *Suppose $\psi : [1, \infty) \to \mathbb{R}_+$ is a continuous non-increasing function. Then, there exists a unique continuous function $r = r_\psi : [t_0, \infty) \to \mathbb{R}$, where $t_0 = -\frac{d}{d+1} \log \psi(1)$, such that*

(i) *The function*
$$\lambda(t) = t - r(t) \tag{8.1}$$
is strictly increasing and tends to ∞ as $t \to \infty$. Moreover, $\lambda(t_0) = 0$.

(ii) *The function*
$$L(t) = t + dr(t) \tag{8.2}$$
is non-decreasing.

(iii) *The functions $\lambda(t)$ and $L(t)$ are related by the following identity:*
$$\psi^d\left(e^{\lambda(t)}\right) = e^{-L(t)}, \qquad t \geq t_0. \tag{8.3}$$

(iv) *The function r is weakly monotone in the following sense: for every $t_2 \geq t_1 \geq t_0$, we have*
$$r(t_2) - r(t_1) \geqslant \frac{-1}{d}(t_2 - t_1). \tag{8.4}$$

Moreover, the function λ satisfies the following growth property for all $t_2 \geq t_1 \geq t_0$:
$$\lambda(t_2) - \lambda(t_1) \leqslant \frac{d+1}{d}(t_2 - t_1). \tag{8.5}$$

Proof. Items (i)-(iii) follow by [KM99, Lemma 8.3] with $x_0 = 1$ in the notation in *loc. cit.* The assertion $\lambda(t_0) = \log x_0 = 0$ follows from their proof. The last item now follows immediately from properties (8.1)-(8.3) via elementary manipulation. Let $t_2 \geq t_1 \geq t_0$. As ψ is non-increasing and strictly positive, (8.2) and (8.3) yield

$$1 \geq e^{-(L(t_2) - L(t_1))} = e^{-(t_2 + dr(t_2) - t_1 - dr(t_1))}$$

and thus (8.4) follows from monotonicity of the exponential function. For (8.5) one calculates

$$\lambda(t_2) - \lambda(t_1) = (t_2 - t_1) - \big(r(t_2) - r(t_1)\big) \leq (t_2 - t_1) + \frac{1}{d}(t_2 - t_1)$$

$$= \frac{d+1}{d}(t_2 - t_1).$$

$\square$

We record a corollary of the above lemma which we use frequently throughout our arguments.

Corollary 8.2. *Let ψ, t_0 and λ be as in Lemma 8.1. For each $n \in \mathbb{N}$, let t_n be such that $e^{\lambda(t_n)} = 2^n$. Then,*

$$t_n \geqslant t_0 + n\frac{d\log 2}{d+1}.$$

Proof. Lemma 8.1(iv) implies that $t_{n+1} - t_n \geqslant d\log 2/(d+1)$. The corollary follows by induction. $\square$

Recall that the map $g\Gamma \mapsto g\mathbb{Z}^{d+1}$ provides an identification of $X = G/\Gamma$ with the space of unimodular lattices in $\mathbb{R}^{d+1}$. A subgroup L of a lattice Δ in $\mathbb{R}^{d+1}$ is primitive if $L = \Delta \cap \mathbb{R} \cdot L$, where $\mathbb{R} \cdot L$ is the $\mathbb{R}$-span of L. We also recall the norms defined in Section 2.2. We define a function $d_1 : X \to [1, \infty)$ as follows:

$$d_1(g\Gamma) = \max\left\{\|v\|^{-1} : v \in g\mathbb{Z}^{d+1} - \{\mathbf{0}\}\right\}. \tag{8.6}$$

As of Mahler's compactness criterion the function d_1 is proper. For $\varepsilon > 0$, define

$$\mathscr{C}(\varepsilon) := \{x \in X : d_1(x) > 1/\varepsilon\}. \tag{8.7}$$

Then, the sets $X \setminus \mathscr{C}(\varepsilon)$ form a compact exhaustion of X.

Denote by $m_{G/\Gamma}$ the normalized G-invariant Haar probability measure on X. The next ingredient is an estimate on the measure of the sets $\mathscr{C}(\varepsilon)$ for the purpose of applying Borel-Cantelli arguments.

Proposition 8.3 (Proposition 7.1, [KM99])**.** *There exist constants $C_d, C_d' \geqslant 1$, depending only on d and the choice of norm on $\mathbb{R}^{d+1}$, so that*

$$C_d\varepsilon^{d+1} - C_d'\varepsilon^{2(d+1)} \leqslant m_{G/\Gamma}\left(\mathscr{C}(\varepsilon)\right) \leqslant C_d\varepsilon^{d+1},$$

for all $0 < \varepsilon < 1$. In fact, we may take $C_d = \mathfrak{c}_{d+1}/2\zeta(d+1)$, where $\mathfrak{c}_{d+1}$ is the volume of the unit ball in $\mathbb{R}^{d+1}$ in our fixed norm.

The following proposition, due to Kleinbock and Margulis, allows us to approximate the characteristic functions of the sets $\mathscr{C}(\varepsilon)$ by smooth functions. In the following statement, we identify $SO_d(\mathbb{R})$ with a subgroup of $SL_{d+1}(\mathbb{R})$ via the map

$$O \mapsto \begin{pmatrix} O & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix}. \tag{8.8}$$

Proposition 8.4. *For every $\eta > 0$ and $\ell \in \mathbb{N}$, there exists a constant $S_\eta \geq 1$ such that for every $\varepsilon > 0$, there are non-negative functions $\varphi_\varepsilon, \varphi_\varepsilon^+ \in \mathrm{B}_{\infty,\ell}^\infty(G/\Gamma)$ such that*

(1) If the norm defining d_1 (cf. (8.6)) is invariant under $SO_d(\mathbb{R})$, then so are φ_ε and φ_ε^+.
(2) $\varphi_\varepsilon \leqslant \chi_{\mathscr{C}(\varepsilon)} \leqslant \varphi_\varepsilon^+$.
(3) $\max\left\{\mathcal{S}_{\infty,\ell}(\varphi_\varepsilon), \mathcal{S}_{\infty,\ell}(\varphi_\varepsilon^+)\right\} \leq S_\eta$.
(4) $m_{G/\Gamma}(\varphi_\varepsilon^+) \leqslant (1+\eta)C_d\varepsilon^{d+1}$ and

$$\frac{1}{(1+\eta)}\big(C_d\varepsilon^{d+1} - C_d'\varepsilon^{2(d+1)}\big) \leqslant m_{G/\Gamma}(\varphi_\varepsilon),$$

where C_d and C_d' are the constants in Proposition 8.3.

Proof. The statement is standard, so we only sketch the proof. Let U_η denote the set of $g \in G$ whose operator norm induced from the norm on $\mathbb{R}^{d+1}$ is at most $a := (1 + \eta)^{1/(d+1)}$. In particular, for $x \in \mathscr{C}(\delta)$ and $g \in U_\eta$, we have $gx \in \mathscr{C}(a\delta)$. The construction proceeds by choosing a C^∞-bump function ψ_η supported in U_η, which is right $\mathrm{SO}_d(\mathbb{R})$-invariant and has integral 1. This is possible because the norm is $\mathrm{SO}_d(\mathbb{R})$-invariant. The desired functions φ_ε and φ_ε^+ are then given by convolving respectively the indicator functions of $\mathscr{C}(\varepsilon/a)$ and $\mathscr{C}(a\varepsilon)$ with ψ_η. The invariance of φ_ε and φ_ε^+ follows by right invariance of ψ_η. Note that $m_{G/\Gamma}(\varphi_\varepsilon^+) = m_{G/\Gamma}(\mathscr{C}(a\varepsilon))$ and similarly for φ_ε. Therefore, the measure estimates follow by Proposition 8.3. The Sobolev norms of the resulting functions can be bounded in terms of those of ψ_η as follows. Let α be a multi-index and denote by D^α a differential operator of order $\ell = |\alpha|$ on G, defined using α in terms of a basis of the Lie algebra. Then, using standard properties of the convolution, we have $D^\alpha(\psi_\eta * \chi_{\mathscr{C}(\varepsilon)}) = D^\alpha(\psi_\eta) * \chi_{\mathscr{C}(\varepsilon)}$. Hence, for $x \in G/\Gamma$,

$$\left| \int_G D^\alpha(\psi_\eta)(g^{-1}) \chi_\varepsilon(gx) \mathrm{d}g \right| \leq \|\chi_{\mathscr{C}(\varepsilon)}\|_\infty \|D^\alpha(\psi_\eta)\|_1 = \|D^\alpha(\psi_\eta)\|_1.$$

This shows that $\mathcal{S}_{\infty,\ell}(\varphi_\varepsilon) \ll \mathcal{S}_{1,\ell}(\psi_\eta)$. $\qquad\square$

9. THE CONVERGENCE THEOREM

The goal of this section is to obtain an analogue of the convergence part of Khintchine's theorem for measures whose translates become effectively equidistributed, Theorem 9.1. We note that we do not require that the measure in question is self-similar. We use the notation introduced in Section 8. We also use the subgroups g_t and $u(\mathbf{x})$ defined in (1.6).

Theorem 9.1 (A Convergence Theorem). *Let $\psi : \mathbb{N} \to \mathbb{R}_+$ be a non-increasing function. Suppose μ is a Borel probability measure on $\mathbb{R}^d$ such that μ satisfies the conclusion of Corollary 6.4 for $\alpha = \emptyset$ (i.e. for $x_\alpha^u = \Gamma \in G/\Gamma$). Then,*

$$\sum_{q \in \mathbb{N}} \psi^d(q) < \infty \implies \mu(W(\psi)) = 0.$$

Proof. For a lattice Δ in $\mathbb{R}^{d+1}$, we denote by $P(\Delta)$ the set of primitive vectors in Δ. In particular, $P(\mathbb{Z}^{d+1})$ consists of $v = (v_1, \ldots, v_{d+1}) \in \mathbb{Z}^{d+1}$ such that the greatest common divisor of $v_1, \ldots, v_{d+1}$ is 1. Let $\|\cdot\|$ denote the sup-norm on $\mathbb{R}^d$ and define the following sets:

$$A_n(\psi) := \left\{ \mathbf{x} \in \mathbb{R}^d : \exists (\mathbf{p}, q) \in P(\mathbb{Z}^{d+1}) \text{ s.t. } 0 < q < 2^{n+1} \text{ and } \|q\mathbf{x} - \mathbf{p}\| < \psi(2^n) \right\}. \tag{9.1}$$

We use A_n to denote $A_n(\psi)$ for simplicity. Then, we note that if $\mathbf{x} \in \mathbb{R}^d$ satisfies (1.1) for some $q \in [2^n, 2^{n+1})$, then monotonicity of ψ implies that $\mathbf{x} \in A_n$. Hence, we have that

$$W(\psi) \subseteq \limsup_{n \to \infty} A_n.$$

By the Borel-Cantelli Lemma, it suffices to show $\sum_{n \geq 1} \mu(A_n) < \infty$.

We shall view ψ as a continuous function on $[1, \infty)$ by linearly interpolating its values at $\mathbb{N}$. Let $r(t)$ and $\lambda(t)$ denote the functions provided by Lemma 8.1. For each sufficiently large $n \in \mathbb{N}$, we let t_n be such that $e^{\lambda(t_n)} = 2^n$. Note that (8.3) yields

$$e^{-(d+1)r(t_n)} = 2^n \psi^d(2^n). \tag{9.2}$$

For each $n \in \mathbb{N}$, define $V_n \subset \mathbb{R}^{d+1}$ by[2]

$$V_n = \left\{ w = (w_1, \ldots, w_{d+1}) \in \mathbb{R}^{d+1} : \|(w_1, \ldots, w_d)\| < e^{-r(t_n)}, \quad |w_{d+1}| < 2e^{-r(t_n)} \right\}.$$

---

[2]The extra factor of 2 in the bound on w_{d+1} ensures (9.4).

Denote by $\tilde{v}_n$ the indicator function of V_n. Denote by v_n the Siegel transform of $\tilde{v}_n$. More precisely, v_n is the function on G/Γ defined by

$$v_n(g\mathbb{Z}^{d+1}) = \sum_{w \in P(g\mathbb{Z}^{d+1})} \tilde{v}_n(w). \tag{9.3}$$

It follows from the definitions that for sufficiently large $n \in \mathbb{N}$, we have

$$A_n = \{\mathbf{x} : v_n(g_{t_n}u(\mathbf{x})\Gamma) \geq 1\}. \tag{9.4}$$

Next, we estimate the measure of A_n. Denote by $2V_n$ the box obtained by scaling the side lengths of V_n by 2. We let v_n^+ denote the Siegel transform of the indicator function of $2V_n$.

As the natural representation of G on $\mathbb{R}^{d+1}$ is continuous and using precompactness of V_n, there is a neighbourhood $\Theta \subseteq G$ of the identity such that $\Theta V_n \subseteq 2V_n$. We take θ to be a non-negative smooth function supported in the interior of Θ and having integral 1 with respect to the Haar measure on G, where the latter is normalized so that the induced measure on G/Γ is a probability measure.

Denote by $\tilde{\chi}_n$ and $\tilde{\chi}_n^+$ the indicator functions of the set of $x \in X$ such that $v_n(x) \geq 1$ and $v_n^+(x) \geq 1$ respectively. We let $\tilde{\varphi}_n^+ = \theta * \tilde{\chi}_n^+$ denote the convolution of θ with $\tilde{\chi}_n^+$. It follows that $\tilde{\chi}_n \leq \tilde{\varphi}_n^+$. Hence, using Corollary 6.4, we obtain

$$\mu(A_n) = \int \tilde{\chi}_n(g_{t_n}u(\mathbf{x})\Gamma)\, d\mu(\mathbf{x}) \leqslant \int \tilde{\varphi}_n^+(g_{t_n}u(\mathbf{x})\Gamma)\, d\mu$$

$$\leqslant \int \tilde{\varphi}_n^+\, dm_{G/\Gamma} + O(\mathcal{S}(\tilde{\varphi}_n^+)e^{-\kappa_* t_n}).$$

Note that $\mathcal{S}(\tilde{\varphi}_n^+) \ll \mathcal{S}(\theta)$ and the implied constant is independent of n; cf. Lemma 3.3. Moreover, using Fubini's theorem, the G-invariance of the Haar measure on G/Γ, and the fact that θ has integral 1, we get

$$\int \tilde{\varphi}_n^+\, dm_{G/\Gamma} = \int \tilde{\chi}_n\, dm_{G/\Gamma}.$$

Hence, since $\tilde{\chi}_n^+ \leq \frac{1}{2}v_n^+$ by symmetry of $2V_n$, Siegel's summation formula [Si45, Equation 25] implies that

$$\int \tilde{\varphi}_n^+\, dm_{G/\Gamma} \leqslant \frac{1}{2} \int v_n^+\, dm_{G/\Gamma} = \frac{1}{2\zeta(d+1)}\mathrm{Vol}(2V_n) = \frac{2^{d+2}\mathfrak{c}_d}{\zeta(d+1)}e^{-(d+1)r(t_n)},$$

where $\mathfrak{c}_d$ is the volume of the unit ball in $\mathbb{R}^d$.

Let $n \in \mathbb{N}$. By Corollary 8.2, we get that $t_n \geqslant t_0 + \frac{nd\log 2}{d+1}$. Combined with (9.2), we obtain

$$\sum_{n \geq 1} \mu(A_n) \ll \sum_{n \geq 1} 2^n \psi^d(2^n) + e^{-\sigma n},$$

for some $\sigma > 0$ and where the implied constant depends only on d, η, and ψ. Summability and monotonicity of ψ then shows that the measures of the sets A_n are summable concluding the proof. $\qquad\square$

10. Effective Double Equidistribution

In our application to Diophantine approximation, we need to apply a certain converse of the classical Borel-Cantelli (Prop. 11.1 below). As a replacement for the assumption on the independence of the events, we need a decay of correlation estimate, which we deduce from our equidistribution statement. The idea behind this deduction follows similar lines to [KSW17, Theorem 1.2], where a similar deduction is carried out for translates of absolutely continuous measures.

For the remainder of this section, we fix a tuple $(\mathcal{F}, \lambda)$ with self-similar measure μ and we let $\|\cdot\|$ denote the associated norms on $\mathbb{R}^d$ and $\mathbb{R}^{d+1}$; cf. Section 2.2. We will use the notation from Corollary 6.4 and we denote by $m_{G/\Gamma}$ the G-invariant Haar probability measure on G/Γ.

The following is the main result of this section.

Proposition 10.1. *Suppose that μ satisfies the conclusion of Corollary 6.4. Assume further that μ has null overlaps (cf. (2.8)). Then, there exist constants $\delta, \varepsilon_* > 0$ and $C_* \geq 1$ such that the following holds. For all non-negative functions $\varphi, \psi \in \mathrm{B}^\infty_{\infty,\ell}(G/\Gamma)$ which are invariant under $\{k_i : i \in \Lambda\}$, and for all $t \geq s > 0$, satisfying*

$$t \geqslant C_* s \qquad or \qquad s \leqslant t \leqslant (1 + \varepsilon_*)s,$$

we have

$$\int \varphi\left(g_t u(\mathbf{x})\Gamma\right) \psi\left(g_s u(\mathbf{x})\Gamma\right) \, d\mu \leqslant \int \varphi \, dm_{G/\Gamma} \int \psi\left(g_s u(\mathbf{x})\Gamma\right) \, d\mu + O\left(\mathcal{S}_{\infty,\ell}(\varphi)\mathcal{S}_{\infty,\ell}(\psi)e^{-\delta|t-s|}\right).$$

10.1. **Proof of Proposition 10.1 for long range correlations.** This subsection is dedicated to the proof under the assumption $t \geq C_* s$ for a suitable $C_* \geq 1$. The other case is handled in the next subsection and its proof is much simpler. We remark however that both cases require the effective equidistribution hypothesis.

To handle the case where the contraction ratios of the IFS are not all the same, we need the notion of complete prefix sets. We say $\alpha \in \Lambda^k$ is a **prefix** of $\omega = (\omega_i)_i \in \Lambda^{\mathbb{N}}$, if $\omega = (\alpha, T^k \omega)$, where $T : \Lambda^{\mathbb{N}} \to \Lambda^{\mathbb{N}}$ is the shift-map given by $(T\omega)_i = \omega_{i+1}$ for $i \in \mathbb{N}$. We say α is a prefix of a finite word ω if ω is strictly longer than α and ω is obtained from α by concatenating a finite word to the end of α. Following [KLW04], we make the following definition.

Definition 10.2. We say a finite set $P \subset \Lambda^*$ is a **complete prefix set** if for every $\omega \in \Lambda^{\mathbb{N}}$, there is a unique word $\alpha \in P$ which occurs as a prefix for ω.

Given $0 < \varepsilon < 1$, one can find a complete prefix set $P(\varepsilon)$ such that every word $\alpha \in P(\varepsilon)$ satisfies

$$\varepsilon \rho_{\min} \leq \rho_\alpha < \varepsilon. \tag{10.1}$$

For example, $P(\varepsilon)$ can be chosen as follows:

$$P(\varepsilon) = \{\alpha \in \Lambda^* : \alpha \text{ satisfies } (10.1) \text{ and no prefix of } \alpha \text{ satisfies } (10.1)\}. \tag{10.2}$$

One then checks that the sets $P(\varepsilon)$ chosen as above are complete prefix sets. We use those sets through the following lemma.

Lemma 10.3. *Assume μ has null overlaps and let P be a complete prefix set. Then, for every continuous function f on $\mathbb{R}^d$,*

$$\int f d\mu = \sum_{\alpha \in P} \int_{\mathcal{K}_\alpha} f d\mu.$$

Proof. Since μ has null overlaps, the collection $\{\mathcal{K}_\alpha : \alpha \in P\}$ forms a measurable partition of the support of μ. The lemma follows readily. $\qquad\square$

Finally, we need the following version of the mean value theorem.

Lemma 10.4. *Let $\psi \in \mathrm{C}^1(G/\Gamma)$. Then, for all $x \in G/\Gamma$ and $\mathbf{v} \in \mathbb{R}^d$,*

$$|\psi(u(\mathbf{v})x) - \psi(x)| \ll \mathcal{S}_{\infty,1}(\psi)\|\mathbf{v}\|.$$

Proof. Let $X_{\mathbf{v}} \in \mathrm{Lie}(G)$ be such that $u(\mathbf{v}) = \exp(X_{\mathbf{v}})$. Then, $\|X_{\mathbf{v}}\| \ll \|\mathbf{v}\|$. Viewing $X_{\mathbf{v}}$ as a differential operator, we have $X_{\mathbf{v}}\psi(x) = \lim_{t\to 0}(\psi(u(t\mathbf{v})x) - \psi(x))/t$. It follows that

$$|\psi(u(\mathbf{v})x) - \psi(x)| = \left|\int_0^1 X_{\mathbf{v}}\psi(u(t\mathbf{v})x) \, dt\right| \leq \|X_{\mathbf{v}}\psi\|_\infty \ll \mathcal{S}_{\infty,1}(\psi)\|\mathbf{v}\|,$$

where the last inequality follows by Lemma 3.3. $\qquad\square$

Let κ_*, ℓ and A_* be the constants provided by Corollary 6.4. Let $\kappa_\sharp = d\kappa_*/(d+1)$. Define

$$w = \frac{\kappa_\sharp t + s}{1 + \kappa_\sharp + A_*}.$$

To simplify notation, for $g \in G$, we use $\varphi(g)$ to denote $\varphi(g\Gamma)$. We further let

$$\xi_t := g_{dt/(d+1)} = \begin{pmatrix} e^{t/(d+1)}\mathrm{Id}_d & \mathbf{0} \\ \mathbf{0} & e^{-dt/(d+1)} \end{pmatrix}, \qquad \mathcal{S} := \mathcal{S}_{\infty,\ell}.$$

Let $\rho_{\min} = \min\{\rho_i : i \in \Lambda\}$. Let $P := P(e^{-w})$ be the complete prefix set defined in (10.2) with $\varepsilon = e^{-w}$. In what follows, we make repeated use of the fact that the norm $\mathcal{S}$ dominates the supremum norms and the Lipschitz constants of the functions φ and ψ.

Recall that $b_\alpha = f_\alpha(\mathbf{0})$. Note further that $\|f_\alpha(\mathbf{x}) - f_\alpha(\mathbf{y})\| = \rho_\alpha \|\mathbf{x} - \mathbf{y}\|$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ and $\alpha \in \Lambda^*$. Let $K = \sup_{\mathbf{x} \in \mathcal{K}} \|\mathbf{x}\|$. By Lemma 10.4, for each $\alpha \in P$ and for every $\mathbf{x} \in \mathcal{K}_\alpha$,

$$|\psi(\xi_s u(\mathbf{x})) - \psi(\xi_s u(b_\alpha))| \ll \mathcal{S}(\psi) \|e^s(\mathbf{x} - b_\alpha)\| = \mathcal{S}(\psi)e^s \rho_\alpha \|f_\alpha^{-1}(\mathbf{x}) - \mathbf{0}\|$$
$$\leqslant K\mathcal{S}(\psi)e^s \rho_\alpha \ll \mathcal{S}(\psi)e^{s-w}, \tag{10.3}$$

where we used that $f_\alpha^{-1}(\mathbf{x}) \in \mathcal{K}$ since $\mathbf{x} \in \mathcal{K}_\alpha$. It follows that

$$\left| \int_{\mathcal{K}_\alpha} \varphi(\xi_t u(\mathbf{x})) \psi(\xi_s u(\mathbf{x})) \, d\mu(\mathbf{x}) - \int_{\mathcal{K}_\alpha} \varphi(\xi_t u(\mathbf{x})) \psi(\xi_s u(b_\alpha)) \, d\mu(\mathbf{x}) \right| \ll \mathcal{S}(\varphi)\mathcal{S}(\psi)e^{s-w}\mu(\mathcal{K}_\alpha). \tag{10.4}$$

Next, we note that the definition of f_α and O_α in (2.2) implies that $f_\alpha^{-1}(\mathbf{x}) = \rho_\alpha^{-1}O_\alpha^{-1}(\mathbf{x} - b_\alpha)$. Moreover, for every $\mathbf{y} \in \mathbb{R}^d$, we have $k_\alpha^{-1}u(\mathbf{y})k_\alpha = u(O_\alpha^{-1}\mathbf{y})$. It follows that

$$k_\alpha^{-1}\xi_t u(\mathbf{x}) = \xi_{t+\log\rho_\alpha}\xi_{-\log\rho_\alpha}k_\alpha^{-1}u(\mathbf{x} - b_\alpha)u(b_\alpha) = \xi_{t+\log\rho_\alpha}u(f_\alpha^{-1}(\mathbf{x}))\underbrace{k_\alpha^{-1}\xi_{-\log\rho_\alpha}u(b_\alpha)}_{h_\alpha}.$$

We let $h_\alpha = k_\alpha^{-1}\xi_{-\log\rho_\alpha}u(b_\alpha)$. Lemma 2.6 and the invariance of φ by k_α imply

$$\int_{\mathcal{K}_\alpha} \varphi(\xi_t u(\mathbf{x}))d\mu = \int_{\mathcal{K}_\alpha} \varphi(\xi_{t+\log\rho_\alpha}u(f_\alpha^{-1}(\mathbf{x}))h_\alpha)d\mu = \mu(\mathcal{K}_\alpha)\int \varphi(\xi_{t+\log\rho_\alpha}u(\mathbf{x})h_\alpha)d\mu. \tag{10.5}$$

Recall we are assuming μ satisfies the conclusion of Corollary 6.4. Hence, we obtain

$$\int \varphi(\xi_{t+\log\rho_\alpha}u(\mathbf{x})h_\alpha)d\mu = \int \varphi \, dm_{G/\Gamma} + O\left(\rho_\alpha^{-A_*}\mathcal{S}(\varphi)e^{-\kappa_\sharp(t+\log\rho_\alpha)}\right)$$
$$= \int \varphi \, dm_{G/\Gamma} + O\left(\mathcal{S}(\varphi)e^{A_* w - \kappa_\sharp(t-w)}\right). \tag{10.6}$$

To combine the above estimates, we note that Lemma 10.3 implies

$$\int \varphi(\xi_t u(\mathbf{x})) \psi(\xi_s u(\mathbf{x})) \, d\mu = \sum_{\alpha \in P} \int_{\mathcal{K}_\alpha} \varphi(\xi_t u(\mathbf{x})) \psi(\xi_s u(\mathbf{x})) \, d\mu$$

Hence, using that $\psi \geq 0$, we obtain

$$\int \varphi\left(\xi_t u(\mathbf{x})\right) \psi\left(\xi_s u(\mathbf{x})\right) \, d\mu = \sum_{\alpha \in P} \int_{\mathcal{K}_\alpha} \varphi\left(\xi_t u(\mathbf{x})\right) \psi\left(\xi_s u(\mathbf{x})\right) \, d\mu$$

$$\overset{(10.4)}{\leqslant} \sum_{\alpha \in P} \psi\left(\xi_s u(b_\alpha)\right) \int_{\mathcal{K}_\alpha} \varphi\left(\xi_t u(\mathbf{x})\right) d\mu + O\left(\mathcal{S}(\varphi)\mathcal{S}(\psi)\left(e^{s-w}\right)\right)$$

$$\overset{(10.6)}{\leqslant} \left(\int \varphi \, dm_{G/\Gamma} + O\left(\mathcal{S}(\varphi)e^{A_* w - \kappa_\sharp(t-w)}\right)\right) \sum_{\alpha \in P} \psi\left(\xi_s u(b_\alpha)\right) \mu(\mathcal{K}_\alpha) + O\left(\mathcal{S}(\varphi)\mathcal{S}(\psi)\left(e^{s-w}\right)\right)$$

$$\leqslant \int \varphi \, dm_{G/\Gamma} \sum_{\alpha \in P} \psi\left(\xi_s u(b_\alpha)\right) \mu(\mathcal{K}_\alpha) + O\left(\mathcal{S}(\varphi)\mathcal{S}(\psi)\left(e^{s-w} + e^{A_* w - \kappa_\sharp(t-w)}\right)\right).$$

Define δ and C_* by

$$\delta = \frac{\kappa_\sharp}{2(1 + \kappa_\sharp + A_*)}, \qquad C_* = \frac{2A_* + \kappa_\sharp}{\kappa_\sharp}. \tag{10.7}$$

Suppose $t \geq C_* s$. Then, our choices of δ and w imply that

$$s - w \leq -\delta(t - s).$$

It follows that

$$\int \varphi\left(\xi_t u(\mathbf{x})\right) \psi\left(\xi_s u(\mathbf{x})\right) \, d\mu \leqslant \int \varphi \, dm_{G/\Gamma} \sum_{\alpha \in P} \psi\left(\xi_s u(b_\alpha)\right) \mu(\mathcal{K}_\alpha) + O\left(\mathcal{S}(\varphi)\mathcal{S}(\psi)e^{-\delta(t-s)}\right).$$

To conclude the proof, we note that (10.3) implies that

$$\psi(\xi_s u(b_\alpha)) \leqslant \frac{1}{\mu(\mathcal{K}_\alpha)} \int_{\mathcal{K}_\alpha} \psi(\xi_s u(\mathbf{x})) \, d\mu(\mathbf{x}) + O\left(\mathcal{S}(\psi)e^{s-w}\right),$$

for all $\alpha \in P$. Combined with the fact that $s - w \leq -\delta(t-s)$ and $\varphi \geq 0$, this implies that

$$\int \varphi \, dm_{G/\Gamma} \sum_{\alpha \in P} \psi\left(\xi_s u(b_\alpha)\right) \mu(\mathcal{K}_\alpha) \leqslant \int \varphi \, dm_{G/\Gamma} \int \psi\left(\xi_s u(\mathbf{x})\right) \, d\mu + O\left(\mathcal{S}(\varphi)\mathcal{S}(\psi)e^{-\delta(t-s)}\right).$$

This concludes the proof.

10.2. Decay of intermediate range correlations.
We retain the notation of the previous subsection and let

$$\gamma = \frac{(d+1)\kappa'}{2d}, \qquad \ell_1 = \ell(d+1)/d,$$

where κ' is given in (5.2).

Define a function Φ on G/Γ by

$$\Phi(x) = \psi(x)\left(\varphi(g_{t-s}x) - \int \varphi \, dm_{G/\Gamma}\right).$$

Then, we have

$$\int \varphi\left(g_t u(\mathbf{x})\right) \psi\left(g_s u(\mathbf{x})\right) \, d\mu(\mathbf{x}) = \int \varphi \, dm_{G/\Gamma} \int \psi(g_s u(\mathbf{x})) \, d\mu(\mathbf{x}) + \int \Phi(g_s u(\mathbf{x})\Gamma) \, d\mu(\mathbf{x}).$$

Since μ satisfies Corollary 6.4, we get

$$\int \Phi(g_s u(\mathbf{x})\Gamma) \, d\mu(\mathbf{x}) = \int \Phi \, dm_{G/\Gamma} + O\left(\mathcal{S}(\Phi)e^{-\kappa_* s}\right).$$

By properties of the Sobolev norm, Lemma 3.3, we get

$$\mathcal{S}(\Phi) \ll \mathcal{S}(\psi)\mathcal{S}(\varphi)e^{\ell_1(t-s)}.$$

Recall that G/Γ is isomorphic to $X_\infty(1)$ by Lemma 3.1. Hence, we may apply bounds on matrix coefficients provided by Proposition 5.3 to get

$$\int \Phi \, dm_{G/\Gamma} = \int \psi \cdot \left(\varphi \circ g_{t-s} - \int \varphi \, dm_{G/\Gamma} \right) dm_{G/\Gamma}$$

$$= \int \psi \cdot (\varphi \circ g_{t-s}) \, dm_{G/\Gamma} - \int \varphi \, dm_{G/\Gamma} \int \psi \, dm_{G/\Gamma} = O\left(\mathcal{S}(\varphi)\mathcal{S}(\psi)e^{-\gamma(t-s)} \right),$$

where we applied the proposition with $\varepsilon = \kappa'/2$. Letting ε_* be given by

$$\varepsilon_* = \frac{\kappa_\sharp}{\ell_1 + \gamma},$$

we obtain a decay rate of $e^{-\gamma(t-s)}$ whenever $s \le t \le (1 + \varepsilon_*)s$ as desired.

11. A Converse to Borel-Cantelli's Lemma

In this section, we obtain a generalization of the Borel-Cantelli Lemma; Proposition 11.1. This result allows us to overcome the lack of strong independence estimates for all pairs of times t and s in Proposition 10.1.

Proposition 11.1. *Suppose E_n is a sequence of measurable sets in a probability space (X, μ). Assume that there are constants $D \ge 1$, $0 < \sigma < 1$ and $0 < a \le 1/\sigma$ such that*

(1) $\mu(E_n) > 0$ for all $n \gg 1$ and $\sum_{n \in \mathbb{N}} \mu(E_n) = \infty$.
(2) There exist constants $C_*, C_\# \ge 1$ and $\varepsilon_* > 0$, such that for all $m, n \in \mathbb{N}$, with $m \gg 1$ and satisfying

$$n \ge C_* m \qquad or \qquad m \le n \le (1 + \varepsilon_*)m,$$

we have

$$\mu(E_m \cap E_n) \le C_\# \mu(E_m)\mu(E_n) + D\left(e^{-\sigma m}\mu(E_n) + e^{-\sigma(n-m)} \right).$$

(3) For all $m, n \in \mathbb{N}$ with $1 \ll m \le n$,

$$\mu(E_m \cap E_n) \le D\mu(E_m) \max\left\{ \mu(E_n)^\sigma, 2^{-\sigma(n-m)} \right\}.$$

(4) For all $m, n \in \mathbb{N}$ with $1 \ll m \le n \le m + \lceil -a\log\mu(E_m) \rceil$,

$$\mu(E_n) \le D\mu(E_m)^\sigma.$$

Then, $\mu(\limsup E_n) \ge 1/C_\#$.

Let $C_* \ge 1$ and $\varepsilon_* > 0$ be the constants in the statement. The idea is to choose a subset $\mathcal{J} \subset \mathbb{N}$ so that its elements are separated in such a way that we can apply our hypotheses on the decay of correlations while retaining the divergence of the sum of the measures. The main point in the construction below is that a C_*-adic interval (i.e. one of the form $[C_*^k, C_*^{k+1}]$) consists of $O(\log C_*/\log(1 + \varepsilon_*))$ many $(1 + \varepsilon_*)$-adic sub-intervals. The pigeonhole principle then allows us to choose only one such $(1 + \varepsilon_*)$-adic sub-interval from within each C_*-adic interval. This ensures that we only encounter long and intermediate range correlations so that we may apply Hypothesis 2. For the short range correlations (encountered only within our chosen $(1 + \varepsilon_*)$-adic intervals), we will apply Hypothesis 3. We now carry out the details.

Step 1 (Choosing a sub-collection): By enlarging C_*, we may assume without loss of generality that

$$C_* = (1 + \varepsilon_*)^{\ell_*},$$

for some $\ell_* \in \mathbb{N}$. For each integer $k \ge 0$, define S_k by

$$S_k = \max\left\{ \sum_{C_*^k(1+\varepsilon_*)^\ell \le n < \lfloor C_*^k(1+\varepsilon_*)^{\ell+1} \rfloor} \mu(E_n) : \ell \in \mathbb{N}, 0 \le \ell < \ell_* \right\}.$$

We define $q_k = C_*^k(1+\varepsilon_*)^\ell$, for some integer $0 \leq \ell < \ell_*$, to be the starting point of an interval that realizes the maximum S_k. More precisely, let q_k be chosen so that

$$S_k = \sum_{q_k \leqslant n < \lfloor(1+\varepsilon_*)q_k\rfloor} \mu(E_n), \qquad q_k = C_*^k(1+\varepsilon_*)^\ell, \qquad 0 \leqslant \ell < \ell_*.$$

We define $\mathcal{J}_k$ by

$$\mathcal{J}_k = \{n \in \mathbb{N} : q_k \leqslant n < \lfloor(1+\varepsilon_*)q_k\rfloor\}.$$

We note that, since the measures of E_n are not summable, we have

$$\infty = \sum_{n\in\mathbb{N}} \mu(E_n) = \sum_{k\geq 1}\sum_{0\leq\ell<\ell_*}\sum_{C_*^k(1+\varepsilon_*)^\ell \leqslant n < \lfloor C_*^k(1+\varepsilon_*)^{\ell+1}\rfloor} \mu(E_n)$$

$$\leqslant \ell_* \sum_{k\geq 1} S_k = \ell_* \sum_{n\in\mathcal{J}_k, k\geq 1} \mu(E_n).$$

In particular, the measures of the sets indexed by $\cup_k \mathcal{J}_k$ are not summable. Moreover, note that

$$\sum_{n\in\mathcal{J}_k, k\geq 1} \mu(E_n) = \sum_{n\in\mathcal{J}_{2k}, k\geq 1} \mu(E_n) + \sum_{n\in\mathcal{J}_{2k-1}, k\geq 1} \mu(E_n).$$

Hence, at least one of the above two sums is infinite. We assume that

$$\sum_{n\in\mathcal{J}} \mu(E_n) = \infty, \qquad \mathcal{J} := \bigcup_{k\geq 1} \mathcal{J}_{2k}. \tag{11.1}$$

The proof in the case where the sum over the sets with odd index diverges is identical. Observe that if $n, q \in \mathcal{J}_{2k}$ and $m \in \mathcal{J}_{2j}$ for some $j < k$, so that $q < n$, then

$$n \geqslant C_* m, \qquad q < n < (1+\varepsilon_*)q. \tag{11.2}$$

Step 2 (Reduction to independence estimates): We recall an inequality due to Chung-Erdős [CE52]: for all $M \leq N$ and positive measure sets F_r in a probability space (Ω, μ),

$$\mu\left(\bigcup_{r=M}^{N} F_r\right) \geqslant \left(\sum_{r=M}^{N} \mu(F_r)\right)^2 \bigg/ \sum_{M\leq r,s\leq N} \mu(F_r \cap F_s). \tag{11.3}$$

In what follows, we use the notation $\sum^\star$ to indicate that the sum is restricted to members of the set $\mathcal{J}$. We claim that for any fixed sufficiently large M

$$\sum_{M\leq m,n\leq N}^{\star} \mu(E_m \cap E_n) \leqslant (C_\# + o(1))\left(\sum_{M\leq n\leq N}^{\star} \mu(E_n)\right)^2, \qquad \text{as } N \to \infty. \tag{11.4}$$

To see that this claim implies the Proposition, note that this estimate combined with (11.3) implies that $\mu(\bigcup_{n\geq M, n\in\mathcal{J}} E_n) \geqslant 1/C_\#$, for all large enough $M \in \mathbb{N}$. Since $\limsup_{n\in\mathcal{J}} E_n$ is a decreasing intersection of sets of this form in a probability space, we obtain the desired result.

Also note that the conclusion of Proposition 11.1 follows trivially if $\mu(E_n) = 1$ for infinitely many n. Hence, we may assume for the remainder of the proof that $\mu(E_n) < 1$ for all n sufficiently large.

Recall the constants $\sigma, a > 0$ in the statement of the proposition. Fix natural numbers $M < N$ with M large enough so that all the hypotheses hold for $n \geq m \geq M$ and so that $\mu(E_m) < 1$ for all $m \geq M$. Let $f(m) = \lceil -a \log \mu(E_m)\rceil \geq 1$. Recall that by assumption, we have $\sigma < 1$ and $a \leq 1/\sigma$. Let $K = \lceil 1/a\sigma \rceil \in \mathbb{N}$. We will need the following elementary fact

$$\sum_{n\geq m} 2^{-\sigma(n-m)} = \frac{1}{1 - 2^{-\sigma}}, \qquad \forall m \in \mathbb{N}. \tag{11.5}$$

Step 3 (Short range independence): Applying Hypothesis 4 iteratively and using that $\sigma < 1$, we find that, for all $m \geq M$ and $k \geq 0$,

$$\mu(E_{m+kf(m)}) \leqslant D^k \mu(E_m)^{\sigma^k}.$$

Hence, for all $m \leq n < m + Kf(m)$, if $0 \leq k \leq K - 1$ satisfies $kf(m) \leq n - m < (k+1)f(m)$, then

$$\mu(E_n) \leqslant D\mu(E_{m+kf(m)})^\sigma \leqslant D^{k\sigma+1}\mu(E_m)^{\sigma^{k+1}} \leqslant D^{(K-1)\sigma+1}\mu(E_m)^{\sigma^K}.$$

Combined with Hypothesis 3 and using that $\sigma < 1$, we obtain

$$\sum_{\substack{M \leq m \leq N, \\ m \leq n < m+Kf(m)}}^{\star} \mu(E_m \cap E_n) \leqslant D \sum_{M \leq m \leq N}^{\star} \mu(E_m) \sum_{m \leq n < m+Kf(m)}^{\star} \left(\mu(E_n)^\sigma + 2^{-\sigma(n-m)}\right)$$

$$\leqslant D^{K+1} \sum_{M \leq m \leq N}^{\star} \mu(E_m)^{1+\sigma^{K+1}} Kf(m) + \frac{D}{1-2^{-\sigma}} \sum_{M \leq m \leq N}^{\star} \mu(E_m).$$

Note that $f(m) \leq 1 - a\log\mu(E_m)$. Let $\varepsilon_0 = \sigma^{K+1}$ and $C_0 \geq 1$ so that $\log x \leq C_0 x^{\varepsilon_0}$ for all $x \geq 1$. Then,

$$\sum_{\substack{M \leq m \leq N, \\ m \leq n < m+Kf(m)}}^{\star} \mu(E_m \cap E_n)$$

$$\leqslant D^{K+1} Ka \sum_{M \leq m \leq N}^{\star} \mu(E_m)^{1+\varepsilon_0} \left(\log\frac{1}{\mu(E_m)} + \frac{1}{a}\right) + \frac{D}{1-2^{-\sigma}} \sum_{M \leq m \leq N}^{\star} \mu(E_m)$$

$$\ll_{D,a,\sigma} \sum_{M \leq m \leq N}^{\star} \mu(E_m). \tag{11.6}$$

Step 4 (Long range independence): To estimate the sum over pairs of events which are separated by more than $Kf(m)$, we use Hypothesis 2. We first bound the contributions of the error terms. Note that

$$\sum_{m=M}^{N}{}^{\star} \sum_{n=m+Kf(m)}^{N}{}^{\star} e^{-\sigma(n-m)} \leqslant \sum_{m=M}^{N}{}^{\star} \sum_{k=Kf(m)}^{\infty}{}^{\star} e^{-\sigma k} \ll \sum_{m=M}^{N}{}^{\star} e^{\log\mu(E_m)} = \sum_{m=M}^{N}{}^{\star} \mu(E_m), \tag{11.7}$$

where the implied constant depends only on σ. For the other error term in Hypothesis 2, we have

$$\sum_{m=M}^{N}{}^{\star} \sum_{n=m+Kf(m)}^{N}{}^{\star} e^{-\sigma m}\mu(E_n) \leqslant \frac{1}{1-e^{-\sigma}} \sum_{M \leq m \leq N}^{\star} \mu(E_m). \tag{11.8}$$

Recall that our choice of $\mathcal{J}$ implies that the pairs $n, m \in \mathcal{J}$ with $m \leq n \leq N$ satisfy the separation inequalities (11.2). In particular, we may apply Hypothesis 2, combined with (11.7) and (11.8), to obtain

$$\sum_{m=M}^{N}{}^{\star} \sum_{n=m+Kf(m)}^{N}{}^{\star} \mu(E_m \cap E_n) \leqslant C_{\#} \sum_{M \leq m,n \leq N}^{\star} \mu(E_m)\mu(E_n) + C_1 \sum_{M \leq m \leq N}^{\star} \mu(E_m), \tag{11.9}$$

for some constant $C_1 \geq 1$, depending only on D and σ. Finally, by (11.1), we have that the sum of the measures of E_m diverges when restricting to $m \in \mathcal{J}$. Therefore, for every fixed M,

$$\sum_{M \leq m \leq N}^{\star} \mu(E_m) = o\left(\left(\sum_{M \leq m \leq N}^{\star} \mu(E_m)\right)^2\right), \qquad \text{as } N \to \infty. \tag{11.10}$$

Hence, (11.4) follows from combining (11.6) and (11.9) with (11.10), thus concluding the proof.

12. The Divergence Theorem

The goal of this section is to obtain an analogue of the divergence part of Khintchine's Theorem for self-similar measures whose translates become effectively equidistributed, Theorem 12.1. Together with Theorem 9.1, this provides a complete analogue of Khintchine's theorem for this class of measures and completes the proof of Theorem A.

Throughout this section, we fix $(\mathcal{F}, \lambda)$ and the associated norms on $\mathbb{R}^d$ and $\mathbb{R}^{d+1}$, denoted $\|\cdot\|$, as in Section 2.2. We let $\mu = \mu_{(\mathcal{F}, \lambda)}$ denote the associated self-similar probability measure.

Theorem 12.1 (A Divergence Theorem). *Let $\psi : \mathbb{N} \to \mathbb{R}_+$ be a non-increasing function and let μ be as above. Assume that $\mathcal{F}$ is irreducible and satisfies the open set condition. Assume further that μ satisfies the conclusion of Corollary 6.4 for functions φ which are invariant under $\{k_i : i \in \Lambda\}$. Then,*

$$\sum_{q \in \mathbb{N}} \psi^d(q) = \infty \implies \mu\left(W(\psi)\right) = 1.$$

Remark 12.2. It can be shown that if the dimension of the attractor of $\mathcal{F}$ is $> d-1$, then $\mathcal{F}$ is automatically irreducible. In particular, irreducibility holds for any IFS satisfying Hypothesis 1.4 in the introduction in view of our choice of $\epsilon_0 \in (0,1)$ in (6.8).

Restricting to functions which are invariant under $\{k_i : i \in \Lambda\}$ in the statement of Theorem 12.1 allows us to deduce it as a direct consequence of the following apriori weaker statement.

Proposition 12.3 (Main Proposition). *Let $\psi : \mathbb{N} \to \mathbb{R}_+$ be a non-increasing function. There exists a constant $\delta_{\psi,d} > 0$ such that for any measure μ satisfying the hypotheses of Theorem 12.1, we have*

$$\sum_{q \in \mathbb{N}} \psi^d(q) = \infty \implies \mu\left(W(\psi)\right) \geq \delta_{\psi,d}.$$

12.1. **Proof of Theorem 12.1 assuming Proposition 12.3.** Let $\delta_\psi > 0$ be the constant provided by Proposition 12.3 and set

$$\mu_\theta := (f_\theta)_* \mu.$$

By Lemma 2.7, to show that $\mu(W(\psi)) = 1$, it suffices to prove that $\mu_\theta(W(\psi)) \geq \delta_\psi$. This lower bound in turn follows by Proposition 12.3 upon verifying that the measures $\mu_\theta := (f_\theta)_* \mu$ satisfy its hypotheses for all $\theta \in \Lambda^*$.

Fix $\theta \in \Lambda^*$. Then, μ_θ is fully supported on $\mathcal{K}_\theta$. Moreover, μ_θ is self-similar with respect to the IFS $\mathcal{F}_\theta = \left\{f_i^\theta := f_\theta f_i f_\theta^{-1} : i \in \Lambda\right\}$ and the same probability vector (λ_i). One also checks that μ_θ has null overlaps. We claim that μ_θ satisfies the conclusion of Corollary 6.4 for functions which are invariant under $\{k_i : i \in \Lambda\}$

Indeed, let $\alpha \in \Lambda^n$ for some n and let $\varphi \in \mathrm{B}_{\infty,\ell}^\infty(G/\Gamma)$ be a function that is invariant under $\{k_i : i \in \Lambda\}$. The similarity f_α^θ takes the form $\rho_\alpha^\theta O_\alpha^\theta + b_\alpha^\theta$, where

$$\rho_\alpha^\theta = \rho_\alpha, \quad O_\alpha^\theta = O_\theta O_\alpha O_\theta^{-1}, \quad b_\alpha^\theta = f_\theta f_\alpha f_\theta^{-1}(\mathbf{0}).$$

We let $k_\alpha^\theta = k_\theta k_\alpha k_\theta^{-1}$. If $\theta = \emptyset$, we will usually omit the superscript.

We let $t_\alpha = \frac{d}{d+1} \log \rho_\alpha$ and we set $h_\alpha^\theta = (k_\alpha^\theta)^{-1} g_{-t_\alpha} u(b_\alpha^\theta)$. By (6.6), we need to verify the conclusion of Corollary 6.4 holds for $\mu = \mu_\theta$ and for basepoints of the form $h_\alpha^\theta \Gamma$. We recall the following key identity, related to (4.3):

$$u(f_\alpha^\theta \mathbf{x})(h_\alpha^\theta)^{-1} = k_\alpha^\theta g_{t_\alpha} u(\mathbf{x}). \tag{12.1}$$

This implies that $h_\alpha^\theta = h_\theta^{-1} h_\alpha h_\theta$. Indeed, the key identity implies that for all $\mathbf{x} \in \mathbb{R}^d$

$$u(f_\alpha^\theta \mathbf{x}) h_\theta^{-1} h_\alpha^{-1} h_\theta = k_\alpha^\theta g_{t_\alpha} u(\mathbf{x}) = u(f_\alpha^\theta \mathbf{x})(h_\alpha^\theta)^{-1}.$$

In particular, the basepoints in Corollary 6.4 that we need to examine for the IFS $\mathcal{F}_\theta$ take the form

$$x_\alpha^{\theta,\mathrm{u}} = (k_\alpha^\theta)^{-1} g_{-t_\alpha} u(b_\alpha^\theta) \Gamma = h_\theta^{-1} h_\alpha h_\theta \Gamma.$$

By another application of (12.1), we have, for all $t \in \mathbb{R}$,

$$g_t u(f_\theta \mathbf{x}) h_\alpha^\theta = k_\theta g_{t+t_\theta} u(\mathbf{x}) h_\alpha h_\theta. \tag{12.2}$$

Denote by $\theta\alpha$ the word obtained by concatenating α to the end of θ. It follows from the definitions that $h_\alpha h_\theta = h_{\theta\alpha}$. Hence, since μ satisfies Corollary 6.4 by hypothesis and since φ is k_θ-invariant, we obtain

$$\int \varphi\big(g_t u(\mathbf{x}) x_\alpha^{\theta,\mathrm{u}}\big) \, d\mu_\theta(\mathbf{x}) = \int \varphi\big(g_t u(f_\theta \mathbf{x}) x_\alpha^{\theta,\mathrm{u}}\big) \, d\mu = \int \varphi(g_{t+t_\theta} u(\mathbf{x}) h_{\theta\alpha} \Gamma) \, d\mu$$

$$= \int \varphi \, dm_{G/\Gamma} + O(\rho_{\theta\alpha}^{-A_*} \mathcal{S}_{\infty,\ell}(\varphi) e^{-\kappa_*(t+t_\theta)}).$$

Finally, we note that $\rho_{\theta\alpha} = \rho_\theta \rho_\alpha$. This shows that μ_θ satisfies the conclusion of Corollary 6.4 for functions invariant by $\{k_i : i \in \Lambda\}$ and concludes the proof of Theorem 12.1.

Remark 12.4. In the case where the IFS is *rational*, one of the referees suggested an alternative argument for upgrading from $\mu(W(\psi)) \geq \delta_\psi$ to $\mu(W(\psi)) = 1$, which we now sketch. Let $\bar\psi : \mathbb{N} \to \mathbb{R}_+$ be a non-increasing function satisfying

(1) $\sum_{q \geq 1} \bar\psi^d(q) \geq 1$.
(2) For all $T > 0$, we have $\lim_{q \to \infty} \bar\psi(q)/\psi(Tq) = 0$.

In particular, our arguments show that $\mu(W(\bar\psi)) \geq \delta_\psi$. Note further that rationality of the IFS implies that if x is in $W(\bar\psi)$, then $h(x)$ is in $W(\psi)$ for any h in the set of maps generated by the IFS. Ergodicity of the associated operator P_λ along with positivity of the measures of $W(\psi)$ and $W(\bar\psi)$ imply that $W(\psi)$ has full measure. We thank the referee for this suggestion.

12.2. **Preliminary reductions.** The remainder of the section is dedicated to the proof of Proposition 12.3. We retain the notation of Section 9 pertaining to the homogeneous space G/Γ.

Recall that $P(\mathbb{Z}^{d+1})$ is the set of primitive vectors in $\mathbb{Z}^{d+1}$. Let a non-increasing approximation function ψ_0 be given so that $\sum_q \psi_0^d(q) = \infty$. Let $C_0 \geq 1$ be a constant satisfying

$$C_0^{-1} \|\cdot\| \leq \|\cdot\|_0 \leq C_0 \|\cdot\|,$$

where $\|\cdot\|_0$ is the sup-norm on $\mathbb{R}^d$.

Lemma 12.5. *In the proof of Proposition 12.3, we may assume that*

$$\psi_0^d(2^n) \leqslant 1/2^n \qquad (n \in \mathbb{N}). \tag{12.3}$$

Proof. Suppose that $\psi_0^d(2^n) \geq 2^{-n}$ for some $n \in \mathbb{N}$. Then, Dirichlet's Theorem implies that for *every* $\mathbf{x} \in \mathbb{R}^d$, there exists $q \in [1, 2^n]$ and $\mathbf{p} \in \mathbb{Z}^d$, so that

$$\|q\mathbf{x} - \mathbf{p}\|_0 \leq 2^{-n/d} \leq \psi_0(2^n).$$

In particular, if $\psi_0^d(2^n) \geq 2^{-n}$ for infinitely many $n \in \mathbb{N}$, then $W(\psi_0) = \mathbb{R}^d$, and the conclusion of Proposition 12.3 follows. Hence, we may assume (12.3) fails only for finitely many n. Since changing finitely many values of ψ_0 does not change the set $W(\psi_0)$, we may assume (12.3) holds for all n. $\qquad\square$

Let $\psi = \psi_0/C_0$ and note that $\sum_q \psi^d(q) = \infty$. For $n \in \mathbb{N}$, define

$$A_n^*(\psi) := \left\{ \mathbf{x} \in \mathbb{R}^d : \exists (\mathbf{p}, q) \in P(\mathbb{Z}^{d+1}) \text{ s.t. } 2^{n-1} \leq |q| < 2^n \text{ and } \|q\mathbf{x} - \mathbf{p}\| < \psi(2^n) \right\}. \tag{12.4}$$

We then define $W^*(\psi)$ to be

$$W^*(\psi) = \limsup_{n \to \infty} A_n^*(\psi).$$

By monotonicity of ψ, we have $W^*(\psi) \subseteq W(\psi_0)$,[3] and hence it suffices to show that $W^*(\psi)$ has full measure. By Lemma 12.5 and using $C_0 \geq 1$, we can assume without loss of generality that

$$\psi^d(2^n) \leqslant 1/2^n \qquad (n \in \mathbb{N}). \tag{12.5}$$

The remainder of this section is dedicated to verifying the hypotheses of Proposition 11.1 for the sets $A_n^*(\psi)$, which we denote A_n^* for simplicity.

It will be convenient for our arguments to also have a lower estimate on ψ; cf. Lemma 12.15. This is done in the following lemma.

Lemma 12.6. *In the proof of Proposition 12.3, we may assume that*

$$\psi^d(q) \geqslant \frac{1}{q \log^{1.1} q}. \tag{12.6}$$

Proof. Let $\psi_2^d(q) = 1/(q \log^{1.1} q)$, $\psi_3 = \max\{\psi, \psi_2\}$. Suppose that $\mu(W(\psi_3)) = 1$. By summability of ψ_2^d and the Convergence Theorem (Theorem 9.1), we have $\mu(W(\psi_2)) = 0$. Using $W(\psi_3) = W(\psi) \cup W(\psi_2)$, it follows that $\mu(W(\psi)) = 1$.
$\square$

Hence, throughout the remainder of the proof, we assume that (12.6) holds.

Similarly to Section 9, we view ψ as a continuous function on $[1, \infty)$ by linearly interpolating its values at $\mathbb{N}$. Denote by $r(t)$ the function obtained from ψ by Lemma 8.1. Let $\lambda(t)$ be the strictly increasing function provided by the same lemma. Define an increasing sequence of times t_n by

$$e^{\lambda(t_n)} = 2^n. \tag{12.7}$$

Let $\ell \in \mathbb{N}$ and $\kappa_* > 0$ be the constants provided by Corollary 6.4. In order to simplify notation, we let $\mathcal{S} = \mathcal{S}_{\infty, \ell}$.

Define the function d_1 in (8.6) using our fixed norm on $\mathbb{R}^{d+1}$ and set

$$\eta = \left(\frac{4}{3}\right)^{1/(d+2)} - 1. \tag{12.8}$$

Applying Proposition 8.4 with this η, we obtain, for each $\varepsilon > 0$, functions φ_ε and φ_ε^+ with uniformly bounded Sobolev norms. For each $n \in \mathbb{N}$, we let

$$\varphi_n := \varphi_{e^{-r(t_n)}}, \qquad \varphi_n^+ := \varphi_{e^{-r(t_n)}}^+, \qquad \chi_n := \chi_{\mathscr{C}(e^{-r(t_n)})}, \tag{12.9}$$

where $\chi_{\mathscr{C}(e^{-r(t_n)})}$ denotes the indicator function of $\mathscr{C}(e^{-r(t_n)})$.

Viewing $\mathrm{SO}_d(\mathbb{R})$ as a subgroup of G via the embedding in (8.8), we see it leaves the norm on $\mathbb{R}^{d+1}$ invariant; cf. Section 2.2. In particular, the functions φ_n and φ_n^+ are invariant by $\mathrm{SO}_d(\mathbb{R})$ by Proposition 8.4. Hence, we may apply our hypothesis that μ satisfies Corollary 6.4 for functions invariant by $\{k_i : i \in \Lambda\} \subset \mathrm{SO}_d(\mathbb{R})$ to φ_n and φ_n^+.

12.3. **Divergence of the sum of measures.** Our first lemma allows us to verify the first hypothesis of Proposition 11.1.

Lemma 12.7. *There exists $C_{\mathcal{F}} \geq 1$ such that for all $n \in \mathbb{N}$,*

$$C_d e^{-(d+1)r(t_n)}/6 - C_d' e^{-2(d+1)r(t_n)} - C_{\mathcal{F}} e^{-\kappa_* t_n} \leqslant \mu(A_n^*) \leqslant 2 C_d e^{-(d+1)r(t_n)} + C_{\mathcal{F}} e^{-\kappa_* t_n},$$

where C_d and C_d' are the constants in Proposition 8.3.

---

[3]Recall that $W(\psi_0)$ was defined in (1.1) using the sup-norm on $\mathbb{R}^d$.

Proof. Fix $n \in \mathbb{N}$. We begin by proving the lower bound. For all $n \in \mathbb{N}$, we define $U_n \subset \mathbb{R}^{d+1}$ by (cf. Section 2.2)

$$U_n = \left\{ w = (w_1, \ldots, w_{d+1}) \in \mathbb{R}^{d+1} : \|(w_1, \ldots, w_d)\| < e^{-r(t_n)}, |w_{d+1}| < e^{-r(t_n)} \right\}.$$

Similarly define

$$V_n = \left\{ w = (w_1, \ldots, w_{d+1}) \in U_n : |w_{d+1}| < 2^{-1} e^{-r(t_n)} \right\}.$$

Denote by $\tilde{u}_n$ and $\tilde{v}_n$ the indicator functions of U_n and V_n and by u_n and v_n the Siegel transforms of $\tilde{u}_n$ and $\tilde{v}_n$ respectively; cf. (9.3).

Consider the following sets:

$$I_n := \left\{ \mathbf{x} \in \mathbb{R}^d : \chi_n(g_{t_n} u(\mathbf{x})\Gamma) = 1 \right\} = \left\{ \mathbf{x} \in \mathbb{R}^d : u_n(g_{t_n} u(\mathbf{x})\Gamma) \geq 1 \right\}.$$

The second equality follows from the fact that a lattice admits a non-trivial short vector if and only if it admits a short primitive vector. We also define sets E_n by

$$E_n := \left\{ \mathbf{x} \in \mathbb{R}^d : v_n(g_{t_n} u(\mathbf{x})\Gamma) \geq 1 \right\}.$$

Note that $I_n \setminus E_n \subseteq A_n^*$ and therefore

$$\mu(A_n^*) \geqslant \mu(I_n) - \mu(E_n). \tag{12.10}$$

We first bound the measure of I_n from below. By definition we have $\varphi_n \leq \chi_n$ and hence $\mu(I_n) \geqslant \int \varphi_n(g_{t_n} u(\mathbf{x})\Gamma) \, d\mu(\mathbf{x})$. Let $C_\mathcal{F} \geq 1$ be chosen such that it bounds the implicit constant in Corollary 6.4 from above and such that $\mathcal{S}(\varphi_n) \leq C_\mathcal{F}$ for all $n \in \mathbb{N}$. Since μ satisfies Corollary 6.4 by our hypothesis,

$$\mu(I_n) \geqslant \int \varphi_n(g_{t_n} u(\mathbf{x})\Gamma) \, d\mu(\mathbf{x}) \geqslant \int \varphi_n \, dm_{G/\Gamma} - C_\mathcal{F} e^{-\kappa_* t_n}. \tag{12.11}$$

Since φ_n was chosen using Proposition 8.4, we obtain

$$\mu(I_n) \geqslant \frac{C_d}{1+\eta} e^{-(d+1)r(t_n)} - \frac{C_d'}{1+\eta} e^{-2(d+1)r(t_n)} - C_\mathcal{F} e^{-\kappa_* t_n}, \tag{12.12}$$

where $C_d = \mathfrak{c}_{d+1}/2\zeta(d+1)$ and C_d' are the constants provided by Proposition 8.4. Here, $\mathfrak{c}_{d+1}$ is the volume of the unit ball in $\mathbb{R}^{d+1}$ with respect to our norm.

Next, we bound the measure of E_n from above. The idea is similar to the proof of Theorem 9.1. Using continuity of the natural representation of G on $\mathbb{R}^d$ and the fact that V_n is open with compact closure, we can choose a neighbourhood $\Theta_\eta \subseteq G$ of the identity such that $\Theta_\eta V_n \subseteq (1+\eta)V_n$. In what follows, θ_η is a non-negative smooth function on G with support contained in Θ_η and of integral 1 with respect to the Haar measure on G, which we normalize so that the induced measure on G/Γ is a probability measure.

We let v_n^+ denote the Siegel transform of the indicator function on $(1 + \eta)V_n$, and we let $\tilde{\chi}_n$ and $\tilde{\chi}_n^+$ denote the indicator functions on the set of $x \in X$ such that $v_n(x) \geq 1$ and $v_n^+(x) \geq 1$ respectively. Let $\tilde{\varphi}_n^+ = \theta_\eta * \tilde{\chi}_n^+$. The argument above then implies that $\tilde{\chi}_n \leq \tilde{\varphi}_n^+$.

Using Corollary 6.4 it follows that

$$\mu(E_n) = \int \tilde{\chi}_n(g_{t_n} u(\mathbf{x})\Gamma) d\mu(\mathbf{x}) \leqslant \int \tilde{\varphi}_n^+(g_{t_n} u(\mathbf{x})\Gamma) d\mu(\mathbf{x})$$

$$\leqslant \int \tilde{\varphi}_n^+ \, dm_{G/\Gamma} + O(\mathcal{S}(\tilde{\varphi}_n^+) e^{-\kappa_* t_n}).$$

By Lemma 3.3(4), we have that $\mathcal{S}(\tilde{\varphi}_n^+)$ depends only on $\mathcal{S}(\theta_\eta)$. In particular, by enlarging $C_\mathcal{F}$ if necessary, we may assume that $\mathcal{S}(\tilde{\varphi}_n^+) \leq C_\mathcal{F}$ for all n. Note that by symmetry of norm balls we

have $\tilde{\chi}_n^+ \le \frac{1}{2} v_n^+$. Using Fubini's and Siegel's theorems, cf. [Si45, Equation 25], we find

$$\int \tilde{\varphi}_n^+ \, dm_{G/\Gamma} \le \frac{1}{2} \int v_n^+ \, dm_{G/\Gamma} = \frac{1}{2\zeta(d+1)} \mathrm{Vol}((1+\eta)V_n) = \frac{(1+\eta)^{d+1} \mathfrak{c}_{d+1}}{4\zeta(d+1)} e^{-(d+1)r(t_n)},$$

where $\mathfrak{c}_{d+1}$ is the volume of the unit ball in $\mathbb{R}^{d+1}$. Hence,

$$\mu(E_n) \le \frac{C_d(1+\eta)^{d+1}}{2} e^{-(d+1)r(t_n)} + C_{\mathcal{F}} e^{-\kappa_* t_n}, \tag{12.13}$$

where C_d is the same constant as in (12.12). Using the choice of η, the lower bound in the lemma now follows by combining (12.10), (12.12), and (12.13).

The upper bound follows upon observing that $\mu(A_n^*) \le \int \varphi_n^+(g_{t_n} u(\mathbf{x})\Gamma) d\mu$, where φ_n^+ was chosen in (12.9). Then one applies Corollary 6.4, Proposition 8.3, and the properties of φ_n^+ as in Proposition 8.4 to conclude. $\square$

We note that the lower bound in Lemma 12.7 may not be positive for all n, which causes issues for the independence arguments. To this end we would like to restrict ourselves to a set of indices which avoids this problem. In order to do that, we do however need that $r(t_n) \to \infty$ as $n \to \infty$.

Lemma 12.8. *We can assume without loss of generality that*

$$\liminf_{n\to\infty} r(t_n) = \infty. \tag{12.14}$$

Proof. Suppose that $\liminf_{n\to\infty} r(t_n) = R$ for some $0 \le R < \infty$. Let I_n as in the proof of Lemma 12.7. One checks that[4]

$$I_n = \left\{ \mathbf{x} \in \mathbb{R}^d : \exists (\mathbf{p}, q) \in P(\mathbb{Z}^{d+1}) \text{ s.t. } 0 < |q| < 2^n \text{ and } \|q\mathbf{x} - \mathbf{p}\| < \psi(2^n) \right\}.$$

Thus, $W(\psi) \supseteq \limsup I_n$. It is then elementary to check that

$$\mu(W(\psi)) \ge \limsup_n \mu(I_n).$$

Let $\varphi_{e^{-R}}$ be the smooth function obtained by applying Proposition 8.4 with η as in (12.8). In particular, recalling the definition of φ_n in (12.9), the functions φ_n converge pointwise to $\varphi_{e^{-R}}$ along a subsequence. Recall by Corollary 8.2 that $t_n \to \infty$. Thus the bound (12.11) and the dominated convergence theorem yield

$$\limsup_n \mu(I_n) \ge \int \varphi_{e^{-R}} \, dm_{G/\Gamma} > 0,$$

and the conclusion of Proposition 12.3 follows. $\square$

We will from now on assume (12.14). Define $\mathcal{G}_0 \subseteq \mathbb{N}$ by

$$\mathcal{G}_0 = \left\{ n \in \mathbb{N} : C_d e^{-(d+1)r(t_n)}/12 \ge C_d' e^{-2(d+1)r(t_n)} + C_{\mathcal{F}} e^{-\kappa_* t_n} + C_{\mathcal{F}} e^{-\delta n} \right\}, \tag{12.15}$$

where $\delta = \kappa_* d \log 2/(d+1)$ and C_d and C_d' are the constants in Lemma 12.7. Using Lemma 12.7, one obtains that for all $n \in \mathcal{G}_0$

$$C_d e^{-(d+1)r(t_n)}/12 \le \mu(A_n^*) \le 4C_d e^{-(d+1)r(t_n)}. \tag{12.16}$$

Lemma 12.9. *The set $\mathbb{N} \setminus \mathcal{G}_0$ is finite.*

Proof. As of Corollary 8.2, the last two terms in the defining inequality in (12.15) are $O_{\mathcal{F},\psi,d}(e^{-\delta n})$. On the other hand, by definition of t_n, by (12.6), and by Lemma 8.1 we have

$$e^{-(d+1)r(t_n)} = 2^n \psi^d(2^n) \tag{12.17}$$

---

[4]The sets I_n differ from A_n^* in removing the lower bound restriction on the denominators q.

and therefore $e^{-(d+1)r(t_n)} \gg n^{-1.1}$ for all $n \in \mathbb{N}$. Thus we obtain

$$C_d e^{-(d+1)r(t_n)}/24 \geq C_{\mathcal{F}}(e^{-\kappa_* t_n} + e^{-\delta n})$$

for sufficiently large n. The claim now follows from combining this with Lemma 12.8. □

Corollary 12.10. $\sum_{n \in \mathcal{G}_0} \mu(A_n^*) = \infty$.

Proof. Since by hypothesis ψ is monotone and satisfies $\sum_{q \geq 1} \psi^d(q) = \infty$, it follows from (12.17) that

$$\sum_{n \in \mathbb{N}} e^{-(d+1)r(t_n)} = \infty.$$

The claim thus follows from Lemma 12.9 and (12.16). □

12.4. **Weak Quasi-independence.** The goal of this subsection is to verify the Hypothesis 3 of Proposition 11.1 regarding the weak quasi-independence hypothesis of the sets A_n^* in the short range.

Proposition 12.11. *There exist constants $C \geq 1$ and $\gamma > 0$ such that for all $m, n \in \mathcal{G}_0$ with $m \leq n$,*

$$\mu(A_m^* \cap A_n^*) \leqslant C \rho_{\min}^{-\gamma} \mu(A_m^*) \max \left\{ \mu(A_n^*)^{\gamma/d}, 2^{-\gamma(n-m)} \right\},$$

where $\rho_{\min}$ is the smallest contraction ratio of the IFS $\mathcal{F}$.

Remark 12.12. The constant $\rho_{\min}$ is not an intrinsic parameter to μ and hence is distinguished from the constant C in the statement.

We remark that the proof of Proposition 12.11 relies on a doubling estimate for the measures of the sets A_n^*; cf. Proposition 12.14. This step in turn relies on our effective equidistribution hypothesis. Additionally, a key ingredient in the proof is the following simplex lemma, whose idea is attributed to Davenport [Da64]. This observation has found numerous applications in problems in Diophantine approximation.

Proposition 12.13. *Suppose $B \subset \mathbb{R}^d$ is a ball of radius $r > 0$ in our fixed norm and let $N \geq 1$. Denote by $\mathcal{Q}(N)$ the set of all the rational points $\mathbf{p}/q$ satisfying $0 < |q| < N$ and $\mathbf{p} \in \mathbb{Z}^d$. Assume that $\mathrm{Vol}(B) < \frac{1}{d! N^{d+1}}$. Then, there exists an affine hyperplane $\mathcal{L}$ such that $B \cap \mathcal{Q}(N) \subset \mathcal{L}$.*

Proof. The lemma is well-known and we include a proof for completeness. If $B \cap \mathcal{Q}(N)$ contains d or fewer points, then the lemma follows in this case. Otherwise, suppose that $\mathbf{p}_i/q_i \in B \cap \mathcal{Q}(N)$, $1 \leq i \leq d+1$, are distinct points which are not contained in any proper affine hyperplane. Denote by Δ the d-dimensional simplex with vertices given by the points $\mathbf{p}_i/q_i$. Then, Δ is contained in B and hence

$$|\Delta| \leqslant \mathrm{Vol}(B), \tag{12.18}$$

where $|\Delta|$ denotes the volume of Δ.

For each i, write $\mathbf{p}_i = (p_{i,1}, \ldots, p_{i,d})$. The product of $d!$ and the volume of Δ equals the absolute value of the determinant of the matrix whose i^{th} row is $(1, p_{i,1}/q_i, \ldots, p_{i,d}/q_i)$. It follows that

$$d!|\Delta| \geqslant \frac{1}{|q_1 \cdots q_{d+1}|} \geqslant \frac{1}{N^{d+1}}. \tag{12.19}$$

We obtain a contradiction upon combining (12.18) and (12.19) with our hypothesis that $\mathrm{Vol}(B) < \frac{1}{d! N^{d+1}}$. □

The next ingredient is the following doubling property of the measures of the sets A_k^*.

Proposition 12.14. *For every $A \geq 1$ there is $A' \geq 1$ such that*

$$\mu(A_m^*(A\psi)) \leqslant A' \mu(A_m^*(\psi)), \tag{12.20}$$

whenever $m \in \mathcal{G}_0$.

Proof. Let $\tilde{\psi} = A\psi$ and let $\tilde{r}$ and $\tilde{\lambda}$ be the corresponding functions provided by Lemma 8.1. Let $\tilde{t}_m$ be chosen so that $e^{\tilde{\lambda}(\tilde{t}_m)} = 2^m$. By Lemma 12.7, we have

$$\mu(A_m^*(\tilde{\psi})) \leqslant 2C_d e^{-(d+1)\tilde{r}(\tilde{t}_m)} + C_{\mathcal{F}} e^{-\kappa_* \tilde{t}_m}.$$

It then follows from the relation $e^{-(d+1)r(t)} = e^{\lambda(t)}\psi^d(e^{\lambda(t)})$ (which also holds for $\tilde{r}$ and $\tilde{\lambda}$ in place of r and λ respectively) that

$$e^{-(d+1)\tilde{r}(\tilde{t}_m)} = 2^m A\psi^d(2^m) = Ae^{-(d+1)r(t_m)} \leqslant A\frac{12}{C_d}\mu(A_m^*(\psi)),$$

where we used (12.16) and the fact that $m \in \mathcal{G}_0$ for the last inequality. Moreover, again using Lemma 8.1(iv) and an induction argument, there is a constant $\tilde{\tau}_0$ depending solely on A and ψ such that $\tilde{t}_m \geqslant \tilde{\tau}_0 + md\log 2/(d+1)$. Since $m \in \mathcal{G}_0$, it follows that

$$C_{\mathcal{F}} e^{-\kappa_* \tilde{t}_m} \leqslant \frac{e^{-\kappa_* \tilde{\tau}_0} C_d}{12} e^{-(d+1)r(t_m)} \leqslant e^{-\kappa_* \tilde{\tau}_0}\mu(A_m^*(\psi)).$$

We, thus, obtain

$$\mu(A_m^*(\tilde{\psi})) \leqslant \left[24A + e^{-\kappa_* \tilde{\tau}_0}\right]\mu(A_m^*(\psi)). \tag{12.21}$$

$\square$

Proof of Proposition 12.11. For $y \in \mathbb{R}^d$ and $r \geq 0$, we write $B(y,r)$ for the ball around y of radius r in our fixed norm $\|\cdot\|$ on $\mathbb{R}^d$. For all $k \in \mathbb{N}$, we can write A_k^* as a union of boxes as follows:

$$A_k^* = \bigcup_{2^{k-1} \leq |q| < 2^k} \bigcup_{\mathbf{p} \in \mathbb{Z}^d : (\mathbf{p},q) \in P(\mathbb{Z}^{d+1})} B\left(\frac{\mathbf{p}}{q}, \frac{\psi(2^k)}{|q|}\right), \tag{12.22}$$

where for $\mathbf{p} = (p_1, \ldots, p_d)$, we use $(\mathbf{p}, q)$ to denote the vector $(p_1, \ldots, p_d, q)$. We denote by $\mathcal{Z}(k)$ the set of centers $\mathbf{p}/q$ of the boxes in the union in (12.22). Denote by Λ the finite index set of the maps in the IFS $\mathcal{F}$. We let K denote the diameter of the fractal $\mathcal{K}$ (in the metric induced by $\|\cdot\|$) and denote by K_* the maximum of 1 and K.

Denote by $\mathfrak{c}_d$ the volume of the unit ball in $\mathbb{R}^d$ in the norm $\|\cdot\|$. Fix $m \leq n$, with $m, n \in \mathcal{G}_0$, and define

$$\delta = \min\left\{\frac{\psi(2^m)}{2^m}, 2^{-\frac{d+1}{d}n}\right\}. \tag{12.23}$$

The monotonicity of ψ implies that $\psi(2^m)/2^m \geq \psi(2^n)/2^n$. Moreover, in view of (12.5), we have that $2^{-\frac{d+1}{d}n} \geq \psi(2^n)/2^n$ and thus we have

$$\frac{\psi(2^n)}{2^n} \leqslant \delta. \tag{12.24}$$

Let $P(\delta)$ be the complete prefix set defined in (10.2), with $\varepsilon = \delta$. By Lemma 10.3, we have

$$\mu(A_m^* \cap A_n^*) = \sum_{\alpha \in P(\delta)} \mu(\mathcal{K}_\alpha \cap A_m^* \cap A_n^*) \leqslant \sum_{\substack{\alpha \in P(\delta), \\ \mathcal{K}_\alpha \cap A_m^* \neq \emptyset}} \mu(\mathcal{K}_\alpha \cap A_n^*). \tag{12.25}$$

Fix $\alpha \in P(\delta)$. Note that the diameter of $\mathcal{K}_\alpha$ is $\leq K_* \rho_\alpha$. Hence, using (12.24), for any ball B_α of radius $5K_* \delta$ and center in $\mathcal{K}_\alpha$, we have

$$\left\{\mathbf{x} \in \mathbb{R}^d : d(\mathbf{x}, \mathcal{K}_\alpha) \leqslant \frac{2\psi(2^n)}{2^{n-1}}\right\} \subseteq B_\alpha,$$

where d denotes the distance in the metric induced by $\|\cdot\|$. Let $B(\alpha, n) = \mathcal{Z}(n) \cap B_\alpha$. We note that if $\mathcal{K}_\alpha$ meets a box $B(\mathbf{p}/q, \psi(2^n)/|q|)$ for some $\mathbf{p}/q \in \mathcal{Z}(n)$, then $\mathbf{p}/q \in B(\alpha, n)$. Recall that there is $M_d \in \mathbb{N}$ depending only on d such that B_α can be covered using at most M_d balls of volume strictly

smaller than $2^{-(d+1)n}/d!$. It follows from Proposition 12.13 that there is a collection $\mathfrak{L}_\alpha = \{\mathcal{L}_i\}$ consisting of at most M_d hyperplanes so that

$$B(\alpha, n) \subset \bigcup_{\mathcal{L}_i \in \mathfrak{L}_\alpha} \mathcal{L}_i.$$

In particular, letting $\varepsilon = \psi(2^n)/2^{n-1}$, this shows that for all $\mathbf{p}/q \in \mathcal{Z}(n)$, we have

$$\mathcal{K}_\alpha \cap B\left(\frac{\mathbf{p}}{q}, \frac{\psi(2^n)}{|q|}\right) \subseteq \bigcup_{\mathcal{L}_i \in \mathfrak{L}_\alpha} \mathcal{K}_\alpha \cap \mathcal{L}_i^{(\varepsilon)}, \tag{12.26}$$

where $\mathcal{L}_i^{(\varepsilon)}$ is the open ε-neighborhood of $\mathcal{L}_i$.

Since $\mathcal{F}$ is irreducible and satisfies the open set condition, Proposition 2.8 shows that μ is (C, γ)-absolutely decaying for some $C \geq 1$ and $\gamma > 0$. Combined with (12.25) and (12.26), this yields

$$\mu(A_m^* \cap A_n^*) \leqslant \sum_{\substack{\alpha \in P(\delta), \\ \mathcal{K}_\alpha \cap A_m^* \neq \emptyset}} \sum_{\mathcal{L}_i \in \mathfrak{L}_\alpha} \mu(\mathcal{K}_\alpha \cap \mathcal{L}_i^{(\varepsilon)}) \leqslant C \sum_{\substack{\alpha \in P(\delta), \\ \mathcal{K}_\alpha \cap A_m^* \neq \emptyset}} |\mathfrak{L}_\alpha| \left(\frac{\varepsilon}{\rho_\alpha}\right)^\gamma \mu(\mathcal{K}_\alpha)$$

$$\leqslant \frac{C_1}{\rho_{\min}^\gamma} \left(\frac{\psi(2^n)/2^n}{\delta}\right)^\gamma \sum_{\substack{\alpha \in P(\delta), \\ \mathcal{K}_\alpha \cap A_m^* \neq \emptyset}} \mu(\mathcal{K}_\alpha), \tag{12.27}$$

where we set $C_1 = 2^\gamma C M_d$.

The next ingredient is to observe that if $\alpha \in P(\delta)$ and $\mathcal{K}_\alpha$ intersects A_m^* non-trivially, then we have that $\mathcal{K}_\alpha$ is contained in $A_m^*((K_* + 1)\psi)$ by choice of δ, where $A_m^*((K_* + 1)\psi)$ is as in (12.4) with the function $(K_* + 1)\psi$ in place of ψ. Hence, by Proposition 12.14, there exists $A' \geq 1$, depending only on ψ and K_*, such that

$$\sum_{\substack{\alpha \in P(\delta), \\ \mathcal{K}_\alpha \cap A_m^* \neq \emptyset}} \mu(\mathcal{K}_\alpha) \leqslant A' \mu(A_m^*). \tag{12.28}$$

We can now conclude the proof of Proposition 12.11. First we note that as ψ is by assumption non-increasing and as $m \leq n$, we have

$$\frac{\psi(2^n)/2^n}{\psi(2^m)/2^m} \leqslant 2^{-(n-m)}. \tag{12.29}$$

This estimate takes care of the case $\delta = \psi(2^m)/2^m$. For the second case, we recall that $e^{\lambda(t_n)} = 2^n$. Hence (8.1)-(8.3) imply that

$$\frac{\psi(2^n)/2^n}{2^{-(d+1)n/d}} = 2^{n/d}\psi(2^n) = e^{-\frac{d+1}{d}r(t_n)}.$$

Hence, it follows by (12.16), since $n \in \mathcal{G}_0$, that

$$\left(\frac{\psi(2^n)/2^n}{2^{-(d+1)n/d}}\right)^\gamma \leqslant \left(\frac{12}{C_d}\mu(A_n^*)\right)^{\gamma/d}. \tag{12.30}$$

The lemma follows upon combining (12.27), (12.28), (12.29), and (12.30). $\qquad \square$

12.5. Quasi-independence and Weak Monotonicity.
The goal of this subsection is to verify Hypotheses 2 and 4 of Proposition 11.1.

Lemma 12.15. *There exist constants $C_*' \geq 1$ and $\sigma, \varepsilon_*' > 0$ such that the following holds. For all sufficiently large $m, n \in \mathcal{G}_0$ satisfying*

$$n \geqslant C_*' m, \qquad or \qquad m \leqslant n \leqslant (1 + \varepsilon_*')m, \tag{12.31}$$

we have

$$\mu(A_m^* \cap A_n^*) \leqslant 576\mu(A_m^*)\mu(A_n^*) + C_{\mathcal{F}}' \left(e^{-\sigma m}\mu(A_n^*) + e^{-\sigma(n-m)}\right)$$

for some constant $C_{\mathcal{F}}' \geq 1$.

Proof. We start with an outline of the argument. Similarly to the convergence case, we approximate $\mu(A_k^*)$ by the average with respect to μ of a smooth function on G/Γ over pushed periodic horocycle at time t_k and then apply Proposition 10.1 in order to derive the desired inequality. In order to do this, we first need to find C_*' and ε_*' so that the separation of m and n implies the separation for t_m and t_n as required by Proposition 10.1. The rest of the proof is then concerned with deriving the upper bound in Lemma 12.15 using the assumed equidistribution of the fractal measure on expanding horospheres. The main terms coming from the equidistribution give rise to the constant 576 and – via Proposition 11.1 – thus to the uniform lower bound in Proposition 12.3.

Let $m \leq n$ be in $\mathcal{G}_0$. Using our hypotheses on μ, we choose constants C_* and ε_* as provided by the conclusion of Proposition 10.1.

First, we choose parameters $C_*' \geq 1$ and $\varepsilon_*' > 0$ so that the separation assumption (12.31) implies the corresponding inequalities for t_n and t_m.

By Lemma 8.1, we have

$$t_n - L(t_n) = -dr(t_n) = -d(t_n - \lambda(t_n)).$$

Since $\lambda(t_n) = n\log 2$ and $L(t_n) = -\log\psi^d(2^n)$, we get

$$t_n = \frac{d\log 2}{d+1}n - \frac{1}{d+1}\log\psi^d(2^n).$$

It follows that

$$t_n - t_m = \frac{d\log 2}{d+1}(n-m) + \frac{1}{d+1}\log\frac{\psi^d(2^m)}{\psi^d(2^n)}. \tag{12.32}$$

Recall that $\psi^d(2^m) \leq 1/2^m$ and $\psi^d(q) \geq 1/q\log^{1.1}q$ by (12.5) and (12.6) respectively. Moreover, by Corollary 8.2, we have that

$$t_m \geqslant t_0 + md\log 2/(d+1) \qquad (m \in \mathbb{N}). \tag{12.33}$$

Let $\varepsilon_*' = d\varepsilon_*/2(d+1)$ and suppose that $n - m \leq \varepsilon_*'m$. Then,

$$t_n - t_m \leqslant (n-m)\log 2 + \frac{1.1}{d+1}\log n$$

$$\leqslant m\varepsilon_*'\log 2 + \frac{1.1}{d+1}\log((1+\varepsilon_*')m)$$

$$\leqslant \varepsilon_*(t_m - t_0)/2 + \frac{1.1}{d+1}\log((1+\varepsilon_*')m).$$

In view of (12.33), we have that

$$\frac{1.1}{d+1}\log((1+\varepsilon_*')m) - \varepsilon_*t_0/2 \leqslant \varepsilon_*t_m/2,$$

for all m large enough. Hence, it follows that $t_n - t_m \leq \varepsilon_*t_m$ for m sufficiently large.

Now, suppose $n \geq C_*'m$ with C_*' still to be determined. Using that $t_n \geq t_0 + nd\log 2/(d+1)$, it follows that for large enough m,

$$t_n \geqslant \frac{dC_*'\log 2}{2(d+1)}m.$$

Arguing as above using (12.32) to estimate $t_m - t_0$, we find that $t_m \leqslant m$ whenever m is large enough. Choosing C_*' to be large enough, depending on C_*, we see that $t_n \geqslant C_*t_m$ when $m \gg 1$.

We now proceed to applying Proposition 10.1. Let $k \in \mathbb{N}$. Using the notation in (12.9), we note that if θ_k is the indicator function of A_k^*, then $\theta_k(\mathbf{x}) \leqslant \chi_k(g_{t_k}u(\mathbf{x})\Gamma)$ for all $\mathbf{x} \in \mathbb{R}^d$. Moreover, since $\chi_k \leq \varphi_k^+$ for all $k \in \mathbb{N}$, we obtain

$$\mu(A_m^* \cap A_n^*) \leqslant \int \varphi_m^+(g_{t_m}u(\mathbf{x})\Gamma)\varphi_n^+(g_{t_n}u(\mathbf{x})\Gamma) \, d\mu(\mathbf{x}). \tag{12.34}$$

Recall that Proposition 8.4(3) implies $\mathcal{S}(\varphi_k^+) \ll 1$, uniformly over $k \in \mathbb{N}$. Hence, by Proposition 10.1, there exist constants $\delta > 0$ and $\tilde{C}_{\mathcal{F}} \geq 1$ such that

$$\int \varphi_m^+(g_{t_m}u(\mathbf{x})\Gamma)\varphi_n^+(g_{t_n}u(\mathbf{x})\Gamma) \, d\mu(\mathbf{x})$$
$$\leqslant \int \varphi_m^+(g_{t_m}u(\mathbf{x})\Gamma) \, d\mu \int \varphi_n^+ \, dm_{G/\Gamma} + \tilde{C}_{\mathcal{F}}e^{-\delta|t_n - t_m|}. \tag{12.35}$$

Next, we find an upper bound for the right side of (12.35) as in the conclusion of the lemma. By (12.32) and monotonicity of ψ, we have that

$$t_n - t_m \geqslant \frac{d \log 2}{d+1}(n - m).$$

Hence, for $\sigma \leq \frac{d}{d+1}\delta \log 2$, we obtain

$$e^{-\delta|t_n - t_m|} \leqslant e^{-\sigma|n-m|}. \tag{12.36}$$

Moreover, by definition of φ_n^+ in (12.9) and Proposition 8.4(4), we have that

$$\int \varphi_n^+ \, dm_{G/\Gamma} \leqslant (1+\eta)C_d e^{-(d+1)r(t_n)} \leqslant 2C_d e^{-(d+1)r(t_n)}.$$

Hence, since $n \in \mathcal{G}_0$, we may apply (12.16) to get that

$$\int \varphi_n^+ \, dm_{G/\Gamma} \leqslant 24\mu(A_n^*). \tag{12.37}$$

To bound the term $\int \varphi_m^+(g_{t_m}u(\mathbf{x})\Gamma) \, d\mu$, we use the effective equidistribution hypothesis on μ. After possibly enlarging the constant $\tilde{C}_{\mathcal{F}}$ in order to subsume the implicit constant in Corollary 6.4, we get

$$\int \varphi_m^+(g_{t_m}u(\mathbf{x})\Gamma) \, d\mu \leqslant \int \varphi_m^+ \, dm_{G/\Gamma} + \tilde{C}_{\mathcal{F}}e^{-\kappa_* t_m}.$$

Arguing as above, since $m \in \mathcal{G}_0$, we see that

$$\int \varphi_m^+(g_{t_m}u(\mathbf{x})\Gamma) \, d\mu \leqslant 24\mu(A_m^*) + \tilde{C}_{\mathcal{F}}e^{-\kappa_* t_m}, \tag{12.38}$$

Finally, using (12.33) once again, there is a constant $A \geq 1$, depending only on ψ and d, such that

$$e^{-\kappa_* t_m} \leq Ae^{-\sigma m} \tag{12.39}$$

for any $\sigma \leq \kappa_* d \log 2/(d+1)$.

Let $\sigma = \min\{\delta, \kappa_*\} d \log 2/(d+1)$ and $C_{\mathcal{F}}' = 24A\tilde{C}_{\mathcal{F}}$. Combining the estimates (12.34)–(12.39), we obtain

$$\mu(A_m^* \cap A_n^*) \leqslant \left(24\mu(A_m^*) + A\tilde{C}_{\mathcal{F}}e^{-\sigma m}\right)24\mu(A_n^*) + \tilde{C}_{\mathcal{F}}e^{-\sigma(n-m)}$$
$$\leqslant 576\mu(A_m^*)\mu(A_n^*) + C_{\mathcal{F}}'\left(e^{-\sigma m} + e^{-\sigma(n-m)}\right),$$

where we used the trivial bound $\mu(A_n^*) \leq 1$. The lemma follows. $\qquad\square$

Since the function $r(t)$ may fail to be monotone, the measures of the sets A_n^* may also fail to decrease monotonically to 0. The next lemma allows us to control this failure of monotonicity in short intervals of natural numbers.

Lemma 12.16. *Let $\gamma \geq 0$ be arbitrary. For every $0 < \varepsilon < 0.1$, there exists a constant $C_{d,\gamma} \geq 1$, so that for all $m \leq n \in \mathcal{G}_0$,*

$$n - m \leqslant \gamma + \varepsilon r(t_m) \implies \mu(A_n^*) \leqslant C_{d,\gamma} \mu(A_m^*)^\sigma,$$

where $\sigma = 1 - 0.1/(d+1)$.

Proof. Fix $\varepsilon \in (0, 0.1)$ and assume that $n - m \leqslant \gamma + \varepsilon r(t_m)$. We would like to bound $t_n - t_m$ from above. Arguing similarly to the proof of Lemma 12.15, using (12.32), we find

$$t_n - t_m \leqslant \frac{(\gamma + \varepsilon r(t_m))d \log 2}{d+1} + \frac{1}{d+1} \log \frac{\psi^d(2^m)}{\psi^d(2^n)}.$$

Since $2^k \psi^d(2^k) = e^{-(d+1)r(t_k)}$ for all $k \in \mathbb{N}$ and $\varepsilon \leq 0.1$, we obtain

$$t_n - t_m \leqslant \gamma + 0.1 r(t_m) + r(t_n) - r(t_m).$$

Recall that $r(t_n) - r(t_m) \geq -(t_n - t_m)/d$ by (8.4). Hence, we obtain

$$r(t_n) - r(t_m) \geqslant -(\gamma + 0.1 r(t_m))/(d+1). \tag{12.40}$$

Moreover, since $m, n \in \mathcal{G}_0$, (12.16) implies $C_d e^{-(d+1)r(t_m)} \leq 8\mu(A_m^*)$ and $\mu(A_n^*) \leq 4C_d e^{-(d+1)r(t_n)}$. Combined with (12.40), we obtain

$$\mu(A_n^*) \leqslant C_{d,\gamma} \mu(A_m^*)^\sigma,$$

where $C_{d,\gamma} = 32 C_d^{1-\sigma} e^\gamma$ and $\sigma = 1 - 0.1/(d+1)$. This completes the proof. $\qquad\square$

Corollary 12.17. *The collection $\{A_n^* : n \in \mathbb{N}\}$ satisfies Hypothesis 4 of Proposition 11.1.*

Proof. Let $m \in \mathbb{N}$. It suffices to verify the hypothesis holds when m is large enough. As $\mathcal{G}_0$ is cofinite, we can assume without loss of generality that $m \in \mathcal{G}_0$. Hence by (12.16) we have

$$-\frac{1}{d+1} \log \mu(A_m^*) \leqslant r(t_m) - \frac{1}{d+1} \log \left(\frac{C_d}{12} \right).$$

Given $\varepsilon \in (0, 1)$, let $a_\varepsilon = \frac{\varepsilon}{d+1}$ and

$$\gamma = \max \left\{ 0, -\frac{1}{d+1} \log \left(\frac{C_d}{12} \right) \right\}.$$

Then, for all sufficiently large $m \in \mathcal{G}_0$ and for all $n \geq m$ we have

$$n - m \leq \lceil -a_\varepsilon \log \mu(A_m^*) \rceil \implies n - m \leq \gamma + 1 + \varepsilon r(t_m).$$

Hence, the Corollary follows from Lemma 12.16. $\qquad\square$

12.6. **Proof of Proposition 12.3.** The results of this section verify Hypotheses 1-4 of Proposition 11.1 for the sequence of events A_n^*. In particular, Lemma 12.15 shows we may take $C_\# = 576$ in the notation of Proposition 11.1. Hence, we get that $\mu(W(\psi)) \geqslant 1/576$ as claimed.

APPENDIX A. SPECTRAL GAP FOR MISSING DIGIT CANTOR SETS

The goal of this section is to prove Theorem C providing a stronger version of our Khintchine and equidistribution theorems in the special case of missing digit Cantor sets. This is done by weakening the hypothesis (6.5) in Theorem 6.1. A key input is a sharper estimate on the spectral gap of the operators $\mathcal{P}_\lambda$, Proposition A.3. Additionally, we take advantage of the equal contraction ratios to show that, in fact, the Sobolev norm (not just the L^2-norm) of a suitable variant of the operators $\mathcal{P}_\lambda^n$ decays in n. Finally, we require a sharper form of Proposition 5.1 due to Strömbergsson as well as bounds towards Selberg's eigenvalue conjecture by Kim-Sarnak.

First, we recall the definition of a missing digit Cantor set.

Definition A.1. A set $\mathcal{K} \subset [0,1]$ is a **missing digit Cantor set** if there exists a prime number $p \geq 3$ and $\emptyset \neq \Lambda \subseteq \{0, \ldots, p-1\}$ such that $\mathcal{K}$ consists of those $x \in [0,1]$ whose digits in their base p expansion all belong to Λ. A **missing digit IFS** (with attractor $\mathcal{K}$) is defined as follows:

$$\mathcal{F} = \left\{ f_i(x) = \frac{x+i}{p} : i \in \Lambda \right\}. \tag{A.1}$$

Throughout the remainder of this section, we fix a missing digit Cantor set $\mathcal{K}$ in base p and digit set Λ along with its associated missing digit IFS $\mathcal{F}$.

In particular, in our notation, $\rho = \rho_i = 1/p$ and $b_i = i/p$. One checks that this IFS satisfies the open set condition. In particular, we have

$$s := \dim_H(\mathcal{K}) = \log|\Lambda| / \log p.$$

By [Mo46], the s-dimensional Hausdorff measure of $\mathcal{K}$ is positive and finite. We denote by μ the restriction of this measure to $\mathcal{K}$, normalized to be a probability measure. By [Hu81], μ is the self-similar measure associated to the probability vector $\lambda_i = \rho^s, i \in \Lambda$.

The following is the precise form of Theorem C.

Theorem A.2. *The conclusions of Theorem A and Theorem 6.1 hold for μ as above whenever*

$$s > 0.839. \tag{A.2}$$

Since we showed that Theorem 6.1 implies Theorem A, we only need to verify that in this special case the former holds under the condition (A.2).

A wasteful step in the proof of Theorem 6.1 is (6.29). To improve this estimate, we introduce slightly different operators than $\alpha \cdot \mathcal{P}_\lambda$ which take advantage of the equal contraction ratios. For $\omega \in \Lambda^n$, we define

$$\tau_\omega = u(\mathbf{0}, -b_\omega)a(1, \rho_\omega) = u(\mathbf{0}, -b_\omega)a(1, p^{-n}). \tag{A.3}$$

Note that τ_ω has trivial Archimedean component. For $\alpha \in \Lambda^*$, let $\alpha \cdot \mathcal{Q}_\lambda$ denote the averaging operator defined analogously to $\alpha \cdot \mathcal{P}_\lambda$ in (4.7) with τ_ω in place of γ_ω. Note that $a(\rho_\omega, 1)\tau_\omega = \gamma_\omega$; cf. (4.2). In particular, for any function φ on X_S and every $n \in \mathbb{N}$, we have

$$(\alpha \cdot \mathcal{P}_\lambda)^n(\varphi)(x) = (\alpha \cdot \mathcal{Q}_\lambda)^n(\varphi)\big(a(p^{-n}, 1)x\big). \tag{A.4}$$

A.1. **Sharper version of Proposition 4.3.**

The following result provides a sharper rate of decay of the operator norm of $\mathcal{P}_\lambda^n$. It holds without restrictions on the dimension of the Cantor set.

Proposition A.3. *Let $\varepsilon > 0$ and $\delta_\varepsilon = \frac{25}{32} - 2\varepsilon$. For all $n \in \mathbb{N}$, $\alpha \in \Lambda^*$, and for every smooth K_f-invariant function $\varphi \in L_{00}^2(X_S)$, we have*

$$\big\| (\alpha \cdot \mathcal{P}_\lambda)^n(\varphi) \big\|_{L^2}^2 \ll_{\varepsilon,p,s} \mathcal{S}_{2,1}(\varphi)^2 (p^{-(s-\varepsilon)n} + p^{-\delta_\varepsilon(n+|\alpha|)}).$$

The same estimate holds for $\mathcal{Q}_\lambda$ in place of $\mathcal{P}_\lambda$.

Proof. Note that in view of bounds towards the Generalized Ramanujan Conjectures (GRC) for SL_2 in [KS03, Proposition 2], for K_{f}-invariant functions, one can take the bound in Corollary 3.7 to be $\xi_{\mathbf{G}}^{25/32}$ instead of $\xi_{\mathbf{G}}^{1/2}$ (GRC predicts the exponent should be 1); cf. [Ve10, Lem. 9.1]. In what follows, we let $m := |\alpha|$. Given $\omega \in \Lambda^n$, in analogy to the proof of Proposition 4.3, we denote $\gamma_\omega^\alpha = \gamma_\alpha \gamma_\omega \gamma_\alpha^{-1}$ and similarly $\tau_\omega^\alpha = \gamma_\alpha \tau_\omega \gamma_\alpha^{-1}$. Expanding $(\alpha \cdot \mathcal{P}_\lambda)^n$ according to (4.7), it follows from Corollary 3.7 and Proposition 4.5 that

$$
\left\| (\alpha \cdot \mathcal{P}_\lambda)^n (\varphi) \right\|^2 = \sum_{\eta, \omega \in \Lambda^n} \lambda_\eta \lambda_\omega \langle \gamma_\eta^\alpha \varphi, \gamma_\omega^\alpha \varphi \rangle \ll \mathcal{S}_{2,1}(\varphi)^2 \sum_{\eta, \omega \in \Lambda^n} \lambda_\eta \lambda_\omega \xi_{\mathbf{G}}^{25/32} \left(\gamma_\omega^\alpha (\gamma_\eta^\alpha)^{-1} \right)
$$

$$
= \mathcal{S}_{2,1}(\varphi)^2 \sum_{\eta, \omega \in \Lambda^n} \lambda_\eta \lambda_\omega \xi_{\mathbf{G}}^{25/32} \left(u(p^{-m}(b_\eta - b_\omega)) \right)
$$

$$
\ll_\varepsilon \mathcal{S}_{2,1}(\varphi)^2 \sum_{\eta, \omega \in \Lambda^n} \lambda_\eta \lambda_\omega \left\| u(p^{-m}(b_\eta - b_\omega)) \right\|_p^{-\frac{25}{64} + \varepsilon},
$$

where $\left\| u(p^{-m}(b_\eta - b_\omega)) \right\|_p$ denotes the norm of the adjoint action of $u(p^{-m}(b_\eta - b_\omega))$ on the Lie algebra of $\mathbf{G}(\mathbb{Q}_p)$. Note further that the above estimate holds for $\mathcal{Q}_\lambda$ since $\gamma_\omega^\alpha (\gamma_\eta^\alpha)^{-1} = \tau_\omega^\alpha (\tau_\eta^\alpha)^{-1}$. Hence, it suffices to bound the above average.

To calculate the adjoint norm, we find a polar decomposition of $u(p^{-m}(b_\eta - b_\omega))$. Note that for all $\mathbf{x} \in \mathbb{Q}_p$ with $|\mathbf{x}|_p > 1$, we have

$$
\begin{pmatrix} 1 & 0 \\ \frac{-1}{\mathbf{x}+1} & 1 \end{pmatrix} u(\mathbf{x}) \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & \frac{-\mathbf{x}}{\mathbf{x}+1} \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 + \mathbf{x} & 0 \\ 0 & \frac{1}{\mathbf{x}+1} \end{pmatrix}.
$$

Then, since $|1/(\mathbf{x}+1)|_p < 1$ and $|\mathbf{x}/(\mathbf{x}+1)|_p = 1$, we obtain

$$
u(\mathbf{x}) = m_1 a\left((1+\mathbf{x})^2 \right) m_2
$$

for some $m_1, m_2 \in \mathbf{G}(\mathbb{Z}_p)$ (recall that $\mathbf{G} = \mathrm{PGL}_2$). Using (3.3), we get

$$
\| u(\mathbf{x}) \|_p = \left\| a\left((\mathbf{x}+1)^2 \right) \right\|_p = |\mathbf{x}|_p^2.
$$

It follows that

$$
\left\| (\alpha \cdot \mathcal{P}_\lambda)^n (\varphi) \right\|^2 \ll_\varepsilon \mathcal{S}_{2,1}(\varphi)^2 \left(\sum_{\eta = \omega \in \Lambda^n} \lambda_\eta \lambda_\omega + p^{-(\frac{25}{32} - 2\varepsilon)m} \sum_{\eta \neq \omega \in \Lambda^n} \lambda_\eta \lambda_\omega |b_\eta - b_\omega|_p^{-\frac{25}{32} + 2\varepsilon} \right).
$$

Fix some $\varepsilon > 0$ and let $\delta = 25/32 - 2\varepsilon$. For each $\eta \neq \omega \in \Lambda^n$, define

$$
d(\eta, \omega) = \max \{ 1 \le i \le n : \eta_i \neq \omega_i \}.
$$

Recall that $b_j = j/p$ for all $j \in \Lambda$. Let $\eta \neq \omega \in \Lambda^n$ and let $d = d(\eta, \omega)$. A simple calculation then shows that

$$
b_\eta - b_\omega = \sum_{i=1}^n (b_{\eta_i} - b_{\omega_i}) p^{-i+1} = \sum_{i=1}^d (\eta_i - \omega_i) p^{-i} = p^{-d} \sum_{i=1}^d (\eta_i - \omega_i) p^{d-i}.
$$

By definition, we have $\eta_d \neq \omega_d$. This implies that the integer $\sum_{i=1}^d (\eta_i - \omega_i) p^{d-i}$ is coprime to p, i.e., a unit in $\mathbb{Z}_p = \{ \mathbf{x} \in \mathbb{Q}_p : |\mathbf{x}|_p \le 1 \}$. Thus, it follows that

$$
\sum_{\eta \neq \omega \in \Lambda^n} \lambda_\eta \lambda_\omega |b_\eta - b_\omega|_p^{-\delta} = \sum_{\eta \neq \omega \in \Lambda^n} \lambda_\eta \lambda_\omega p^{-\delta d(\eta, \omega)} = \sum_{\eta \in \Lambda^n} \lambda_\eta \sum_{j=1}^n p^{-\delta j} \sum_{\substack{\omega \in \Lambda^n \\ d(\eta, \omega) = j}} \lambda_\omega.
$$

We now specialize to the case where λ is the uniform probability vector with weight $1/|\Lambda|$. Then, for each $\eta \in \Lambda^n$, we have

$$
\lambda_\eta = |\Lambda|^{-n} = p^{-sn},
$$

where $s = \log |\Lambda| / \log p$. Hence, we obtain

$$\sum_{\eta \neq \omega \in \Lambda^n} \lambda_\eta \lambda_\omega \, |b_\eta - b_\omega|_p^{-\delta} = p^{-2sn} \sum_{j=1}^n p^{-\delta j} \sum_{\eta \in \Lambda^n} \underbrace{|\{\omega \in \Lambda^n : d(\eta, \omega) = j\}|}_{= |\Lambda|^{j-1}(|\Lambda| - 1)}$$

$$\leqslant p^{-2sn} \sum_{j=1}^n p^{-\delta j} \sum_{\eta \in \Lambda^n} |\Lambda|^j = p^{-2sn} \sum_{j=1}^n p^{-\delta j} |\Lambda|^{n+j}.$$

If $s \neq \delta$, then, using that $|\Lambda| = p^s$, we obtain

$$\sum_{\eta \neq \omega \in \Lambda^n} \lambda_\eta \lambda_\omega \, |b_\eta - b_\omega|_p^{-\delta} \leqslant p^{-sn+s-\delta} \sum_{j=0}^{n-1} p^{(s-\delta)j} = p^{-sn+s-\delta} \frac{p^{(s-\delta)n} - 1}{p^{(s-\delta)} - 1} \leqslant \frac{p^{s-\delta}}{|p^{s-\delta} - 1|} \big(p^{-\delta n} + p^{-sn} \big).$$

Otherwise, if $s = \delta$, we get a bound of the form np^{-sn}. Finally, we note that

$$\sum_{\eta = \omega \in \Lambda^n} \lambda_\eta \lambda_\omega = p^{-sn}.$$

$\square$

In the proof of Theorem A.2, we will need an estimate on the decay of the L^4-norm of the operators $\alpha \cdot \mathcal{Q}_\lambda$. We deduce this estimate in the following corollary.

Corollary A.4. *For all $q \geq 2$, $n \in \mathbb{N}$, $\alpha \in \Lambda^*$, and for every bounded smooth K_{f}-invariant function $\varphi \in \mathrm{L}_{00}^2(X_S)$, we have*

$$\big\| (\alpha \cdot \mathcal{Q}_\lambda)^n(\varphi) \big\|_{\mathrm{L}^q}^q \ll_{\varepsilon,p,s,q} \mathcal{S}_{2,1}(\varphi)^2 \, \|\varphi\|_\infty^{q-2} \cdot p^{-2o_\varepsilon n},$$

for every $\varepsilon > 0$, where $2o_\varepsilon = \min \{25/32, s\} - \varepsilon$.

Proof. The case when $q = 2$ is exactly Proposition A.3. Hence, we may assume $q > 2$. Let μ_S denote the $\mathbf{G}_S$-invariant probability measure on X_S. Using Fubini's Theorem one checks that

$$\big\| (\alpha \cdot \mathcal{Q}_\lambda)^n(\varphi) \big\|_{\mathrm{L}^q}^q = \int_0^\infty \mu_S(x : |(\alpha \cdot \mathcal{Q}_\lambda)^n(\varphi)(x)|^q > t) \, dt$$

$$= \int_0^{\|\varphi\|_\infty^q} \mu_S(x : |(\alpha \cdot \mathcal{Q}_\lambda)^n(\varphi)(x)|^q > t) \, dt,$$

where we used that $\|\alpha \cdot \mathcal{Q}_\lambda^n(\varphi)\|_\infty \leq \|\varphi\|_\infty$. Hence, by Proposition A.3 and Chebychev's inequality, we have for all $t > 0$,

$$\mu_S(x : |(\alpha \cdot \mathcal{Q}_\lambda)^n(\varphi)(x)|^q > t) \ll_{\varepsilon,p,s} \mathcal{S}_{2,1}(\varphi)^2 p^{-2o_\varepsilon n} t^{-2/q}.$$

Hence, since $q > 2$, we obtain

$$\big\| (\alpha \cdot \mathcal{Q}_\lambda)^n(\varphi) \big\|_{\mathrm{L}^q}^q \ll_{\varepsilon,p,s} \mathcal{S}_{2,1}(\varphi)^2 p^{-2o_\varepsilon n} \|\varphi\|_\infty^{q-2} \frac{q}{q-2}.$$

$\square$

A.2. **Sharper version of Proposition 5.1.** The following result provides a sharper value of κ constituting the rate of equidistribution of horospherical measures on congruence covers.

Proposition A.5 (Prop. 3.1, [St13]). *Let $\Delta \leq \Gamma(1)$ be a congruence lattice and $X_\Delta = \mathbf{G}_\infty / \Delta$. Then, for every $\varphi \in \mathrm{B}_{2,3}^\infty(X_\Delta)$, $x \in X_\Delta$ and $t \geq 1$,*

$$\int_0^1 \varphi(a(t)u(\mathbf{x})x) \, d\mathbf{x} = \int \varphi \, dm_{\mathbf{G}_\infty^+ \cdot x} + O\big(V_\Delta \cdot \mathcal{S}_{2,3}(\varphi) \cdot t^{-\kappa} \cdot \mathscr{Y}_\Delta^{1/2}(x)\big),$$

where $V_\Delta = \sqrt{[\Gamma(1) : \Delta]}$, $\mathscr{Y}_\Delta$ is a positive proper function on X_Δ and if $\lambda_1 \in (0, 1/4)$ is a uniform lower bound on the non-zero eigenvalues of the Laplacian on X_Δ for all Δ, then

$$\kappa = \frac{1 - \sqrt{1 - 4\lambda_1}}{2}.$$

The implied constant is independent of Δ.

Proof. The statement in [St13, Prop. 3.1] is stated in a slightly different form, we outline the needed modifications. First, the results in *loc. cit.* are stated for quotients of $\mathrm{SL}_2(\mathbb{R})$. Recall that $\mathbf{G}_\infty^+$ is the image of $\mathrm{SL}_2(\mathbb{R})$ inside $\mathbf{G}_\infty$ and is a normal subgroup of index 2. In particular, for each Δ, X_Δ consists of at most two connected components, each of which is isomorphic to $\mathrm{SL}_2(\mathbb{R})/\Delta'$, where $\Delta' \leq \mathrm{SL}_2(\mathbb{Z})$ is a congruence lattice. We define $\mathscr{Y}_\Delta$ to be $\mathscr{Y}_{\Delta'}$ (in the notation of [St13, Eq. (11)]) on each of the connected components of X_Δ.

The measure on X_Δ defining the L^2-Sobolev norms $\|\cdot\|_{W_k}$ in *loc. cit.* has total mass $\asymp V_\Delta^2$. In particular, this norm is equivalent to $V_\Delta \cdot \mathcal{S}_{2,k}$ for all $k \in \mathbb{N}$; cf. discussion following [St13, Eq. (9)]. Note further that the statement is made for long horocycle orbits starting from a point p. The above statement is obtained from this result with $T = t$ and with $pa(T^{-1})$ in place of p in the notation in *loc. cit* using standard conjugation relations of $a(t)$ and $u(\mathbf{x})$.

Next, we note that the implied constant in [St13, Prop. 3.1] can be made independent of Δ. The dependence on the lattice comes from [St13, Lem. 2.1]. Note that the bounds in [St13, Lem. 2.2, 2.3] are not needed for our weaker error term $t^{-\kappa}\mathscr{Y}_\Delta^{1/2}(x)$.

The dependence in [St13, Lem. 2.1] arises from a choice of an injectivity radius to allow for a thick-thin decomposition of $\mathbb{H}^2/\Delta'$ in order to apply the Sobolev embedding theorem (cf. the choice of ϵ in the proof of [St13, Lem. 2.1] given in [FF03, Lem. 5.3]). As X_Δ are all covers of $X_\infty(1) \cong \mathrm{SL}_2(\mathbb{R})/\mathrm{SL}_2(\mathbb{Z})$, a choice of an injectivity radius in $X_\infty(1)$ works for all of X_Δ.

Hence, the error term can be obtained by applying [St13, Lem. 2.1] to Burger's integral formula in [St13, Eq. (23)] combined with the estimates on the height function in [St13, pg 303] and the estimates on the intertwining operators given in [St13, Eq. (22)] (or [Bu90, pg. 791] with $\alpha = \kappa$ in the notation of [Bu90]) as is done in [St13]. One uses [St13, Lem. 2.2] to ensure pointwise convergence of the last integral in [St13, Eq. (23)] to a bounded continuous function as is done towards the end of the proof so that the above bounds apply.

Finally, we note that the order 3 Sobolev norm in the statement (as opposed to $\mathcal{S}_{2,4}$ in *loc. cit.*) arises from only applying the bounds of [St13, Lem. 2.1] in the proof of [St13, Prop. 3.1].
□

A.3. **Proof of Theorem A.2.** We outline the needed modifications of the proof of Theorem 6.1 in this setting. We retain the notation in that proof, in particular the constants a, b_ε and c in the statement of Theorem 6.1. We begin by noting that the average contraction ratio r is p^{-1} in the case at hand.

Since missing digit Cantor sets satisfy the open set condition with the open set $(0, 1)$, the proof of Theorem 6.1 shows that we can take the absolutely continuous measure ν to be the Lebesgue measure on the unit interval. In this case, the Mass Term in (6.22) takes the form

$$\text{Mass Term} = p^{2\sigma m}, \qquad 2\sigma = 1 - s.$$

By Lemma 4.2 and using (A.4), $(\alpha \cdot \mathcal{Q}_\lambda)^n(\varphi)$ can be regarded as a function on $\mathbf{G}_\infty/\Delta$, for some congruence lattice Δ. Hence, we may apply Proposition A.5 in place of Proposition 5.1 to obtain

the following replacement of (6.27):

$$\int_0^1 (\alpha \cdot \mathcal{P}_\lambda)^n (\varphi)^2 \big(a(t)u(\mathbf{x})h_\alpha\Delta\big) \, \mathrm{d}\mathbf{x} = \int_0^1 (\alpha \cdot \mathcal{Q}_\lambda)^n (\varphi)^2 \big(a(tp^{-n})u(\mathbf{x})h_\alpha\Delta\big) \, \mathrm{d}\mathbf{x}$$

$$= \int (\alpha \cdot \mathcal{Q}_\lambda)^n (\varphi)^2 \, \mathrm{d}m_{\mathbf{G}_\infty^+/\Delta^+}$$

$$+ O\left(V_\Delta \mathcal{S}_{2,\ell}\big((\alpha \cdot \mathcal{Q}_\lambda)^n (\varphi)^2\big) p^{\kappa n} t^{-\kappa} \mathscr{Y}_\Delta^{1/2}(h_\alpha\Delta)\right), \quad \text{(A.5)}$$

where $\ell = 3$, $\Delta^+ = \mathbf{G}_\infty^+ \cap \Delta$. Here, we use the fact that $h_\alpha \in \mathbf{G}_\infty^+$ so that $\mathbf{G}_\infty^+ \cdot h_\alpha\Delta \cong \mathbf{G}_\infty^+/\Delta^+$.

Note further that, by [St13, Eq. (11)-(13)], $\mathscr{Y}_\Delta(x) \leq \mathscr{Y}_{\Gamma(1)}(x) \ll \|\mathrm{Ad}_g\|$, where $g \in \mathbf{G}_\infty$ is any representative of x and $\|\mathrm{Ad}_g\|$ denotes the norm of its adjoint action. In particular, the estimate $\mathscr{Y}_\Delta^{1/2}(x_\alpha) \ll \rho_\alpha^{-C_0}$ holds for a suitable $C_0 \geq 1$ in place of the estimate (6.28).

The key point in introducing the operators $\mathcal{Q}_\lambda$ is as follows. Since multiplication by elements of $\mathbf{G}_\mathrm{f}$ commutes with differential operators on $\mathbf{G}_\infty$, one checks using Lemma 3.3 that

$$\mathcal{S}_{2,\ell}\big((\alpha \cdot \mathcal{Q}_\lambda)^n (\varphi)^2\big) \ll \mathcal{S}_{4,\ell}\big((\alpha \cdot \mathcal{Q}_\lambda)^n (\varphi)\big)^2.$$

Moreover, note that $\mathcal{D}\varphi$ has mean 0 for any differential operator $\mathcal{D}$. This can be checked by induction on the degree of the operator using the dominated convergence theorem, invariance of the Haar measure, and the limit definition of Lie derivatives; cf. proof of Lemma 10.4. In particular, Lemma 3.5 implies that the lift of $\mathcal{D}\varphi$ to X_S belongs to $\mathrm{L}_{00}^2(X_S)$.

Hence, by Corollary A.4, applied with $q = 4$, for any $\mathcal{D}$ of degree $\leq \ell$, we have

$$\|(\alpha \cdot \mathcal{Q}_\lambda)^n (\mathcal{D}\varphi)\|_{\mathrm{L}^4}^4 \ll_{\varepsilon,p,s} \mathcal{S}_{\infty,1}(\mathcal{D}\varphi)^4 p^{-2o_\varepsilon n} \leqslant \mathcal{S}_{\infty,\ell+1}(\varphi)^4 p^{-2o_\varepsilon n},$$

where

$$2o_\varepsilon := \min\{25/32, s\} - \varepsilon.$$

It follows that

$$\mathcal{S}_{2,\ell}\big((\alpha \cdot \mathcal{Q}_\lambda)^n (\varphi)^2\big) \ll \mathcal{S}_{4,\ell}\big((\alpha \cdot \mathcal{Q}_\lambda)^n (\varphi)\big)^2 \ll_{\varepsilon,p,s} \mathcal{S}_{\infty,\ell+1}(\varphi)^2 p^{-o_\varepsilon n}.$$

Additionally, by Lemma 4.2, the congruence lattice Δ can be chosen so that $V_\Delta \ll p^{3|\alpha|+3n/2}$. Finally, estimating the main term in (A.5) using Proposition A.3, we obtain the following sharper bound on the horospherical term:

$$\text{Horospherical term} \ll_{\varepsilon,p} \mathcal{S}_{\infty,\ell+1}(\varphi)^2 p^{(3+C_0)|\alpha|} \left(p^{-2o_\varepsilon n} + p^{2n\upsilon - \kappa\tau}\right),$$

where

$$2\upsilon = 3/2 + \kappa - o_\varepsilon, \qquad p^\tau = t.$$

By known bounds towards Selberg's eigenvalue conjecture due to [KS03, Proposition 2], we can take $\lambda_1 \geq 975/4096$ in Proposition A.5. In particular, we may take $\kappa = 25/64$.

By combining the above estimates and balancing the rates as is done in the proof of Theorem 6.1 (cf. discussion following (6.33)), we see that the conclusion of that theorem holds in our setting if

$$2\sigma(o_\varepsilon + \upsilon) < \kappa(\sigma + o_\varepsilon) \iff (1-s)(o_\varepsilon + 3/2) < 2\kappa o_\varepsilon,$$

for some $\varepsilon > 0$ and with our choices of $\sigma, o_\varepsilon, \upsilon$ and κ as above. This condition is in turn satisfied under our hypothesis (A.2) as can be shown by a direct calculation.

A.4. A version of Lebesgue density.

In this subsection, we verify the version of Lebesgue density theorem for Bernoulli measures on symbolic spaces used in the proof of Lemma 2.7.

Let Λ be a finite set and λ be a probability vector on Λ. For $\alpha \in \Sigma := \Lambda^\mathbb{N}$ and $k \in \mathbb{N}$, denote by $\Sigma(\alpha, k)$ the cylinder set given by the prefix of α of length k and denote this prefix by $\alpha|_k$. We endow Λ with the discrete topology and Σ with the associated product topology.

Lemma A.6. *Suppose $B \subseteq \Sigma$ is a Borel set. Then, for $\lambda^{\mathbb{N}}$-almost every $x \in B$,*

$$\lim_{k \to \infty} \frac{\lambda^{\mathbb{N}}(B \cap \Sigma(x,k))}{\lambda^{\mathbb{N}}(\Sigma(x,k))} = 1.$$

Proof. We deduce this result from the corresponding well-known Lebesgue density theorem for Radon measures on the real line. Let $p = 2|\Lambda|$ and consider the auxiliary IFS given by

$$\mathcal{F} = \{f_i(x) = (x+i)/p : 0 \le i \le p-1, i \text{ is even}\}.$$

Let $\mathcal{K}$ be its attractor and note that the images of $\mathcal{K}$ under distinct maps in $\mathcal{F}$ are disjoint. Let $\pi : \Sigma \to \mathbb{R}$ be the coding map defined by $\pi(\alpha) = \lim_{k\to\infty} f_{\alpha|_k}(0)$ and $\mu = \pi_* \lambda^{\mathbb{N}}$ be the self-similar measure. Then, π is a homeomorphism onto its image $\mathcal{K}$; cf. [Hu81, Thm. 3.1.(3) and Thm. 4.4.(4)]. Hence, it suffices to show that

$$\lim_{k \to \infty} \frac{\mu(\pi(B) \cap \pi(\Sigma(\alpha,k)))}{\mu(\pi(\Sigma(\alpha,k)))} = 1$$

for $\lambda^{\mathbb{N}}$-almost every $\alpha \in B$. Let $\alpha \in B$. To relate the images of cylinder sets under π to intervals in $\mathbb{R}$, one first checks that $\pi(\Sigma(\alpha,k))$ is contained in the image of $[0,1]$ under $f_{\alpha|_k}$. Hence, by definition of $\mathcal{F}$, given any $\beta \in \Sigma$ such that $\Sigma(\beta,k) \ne \Sigma(\alpha,k)$, the distance between $\pi(\Sigma(\beta,k))$ and $\pi(\Sigma(\alpha,k))$ is at least p^{-k}. It follows that

$$\pi(\Sigma(\alpha,k)) = B(\pi(\alpha), p^{-k}) \cap \mathcal{K},$$

where $B(\pi(\alpha), p^{-k})$ denotes the open interval around $\pi(\alpha)$ of radius p^{-k}. It follows by Lebesgue's density theorem for Radon measures on $\mathbb{R}$ that

$$\lim_{k \to \infty} \frac{\mu(\pi(B) \cap \pi(\Sigma(\alpha,k)))}{\mu(\pi(\Sigma(\alpha,k)))} = \lim_{k \to \infty} \frac{\mu(\pi(B) \cap B(\pi(\alpha), p^{-k}))}{\mu(B(\pi(\alpha), p^{-k}))} = 1,$$

for $\lambda^{\mathbb{N}}$-almost every $\alpha \in B$. Note that we are allowed to use open balls in this application of Lebesgue density since μ is non-atomic. Indeed, it suffices to note that $\lambda^{\mathbb{N}}(\Sigma(\beta,k)) \le \lambda_{\max}^k \to 0$ for any $\beta \in \Sigma$, where $\lambda_{\max}$ denotes the largest component of λ. $\qquad\square$

Appendix B. Congruence quotients

The goal of this appendix is to give proofs of several facts presented in Section 3.1 and used in the proof of Theorem 6.1. In Corollary B.11, we establish the correspondence between compact-open subgroups of $\mathbf{G}(\mathbb{A}_f)$ and principal congruence subgroups of $\mathbf{G}(\mathbb{Z})$ which underlies the double coset decomposition (3.4). In Proposition B.13, we prove the uniform bound on the number of connected components of $K_{S_f}[N]\backslash \mathbf{G}_S/\Gamma_S$ used in (6.31). At the end of this section, we will define general congruence subgroups; this extension is immediate but we include it for completess.

B.1. **Integral structures.** We begin by making an explicit choice of the integral structure on $\mathbf{G}$ used to define congruence groups. Given a ring R, we let

$$V_R = \text{Mat}_{d+1}(\mathbb{Z}) \otimes_{\mathbb{Z}} R.$$

Then V_R is an R-algebra which is a free R-module of rank $(d+1)^2$. The algebra V_R allows us to realize $\mathbf{G}(k)$ as a linear group whenever k is a field. More explicitly, we fix a faithful k-representation of $\mathbf{G}$ by choosing the standard basis $\mathcal{E}_{d+1}$ of V_k which gives rise to an isomorphism $\text{GL}(V_k) \cong \text{GL}_{(d+1)^2}(k)$ and we define

$$\mathbf{G} = \{g \in \text{GL}_{(d+1)^2} : \forall u, v \in \mathcal{E}_{d+1} \quad g(uv) = (gu)(gv)\}.$$

In what follows, we let $\Phi : \text{GL}_{d+1} \to \mathbf{G}$ denote the k-representation given by

$$\Phi(x)(v) = xvx^{-1} \quad (x \in \text{GL}_{d+1}, v \in \text{Mat}_{d+1}). \tag{B.1}$$

By the Skolem-Noether theorem we have $\Phi(\text{GL}_{d+1}(k)) = \mathbf{G}(k)$ for any field k and in particular $\mathbf{G}(k) \cong \text{GL}_{d+1}(k)/k^{\times}$, where we identify $k^{\times}$ with the scalar diagonal matrices in $\text{GL}_{d+1}(k)$.

We record the following consequence of the above discussion which is used to apply the results of [GMO08].

Lemma B.1. *The group* $\mathbf{G}$ *is a connected group over* k.

Proof. As $\mathbf{G}$ is an affine k-group, we only have to prove that it is connected. Recall that GL_{d+1} is an irreducible affine $\mathbb{Q}$-group. To this end, note that GL_{d+1} is the principal open set defined by the polynomial $\det(x_{ij})$, i.e.,

$$\mathrm{GL}_{d+1}(k) = \left\{ (x_{ij}) \in k^{(d+1)^2} : \det(x_{ij}) \neq 0 \right\}.$$

This is a Zariski-open subset of affine space. As affine space is irreducible, every open subset of affine space is irreducible and hence $\mathrm{GL}_{d+1}(k)$ is irreducible. In particular, it follows that GL_{d+1} is a connected group over k; cf. [Bo91, Prop. I.1.2]

By the Skolem-Noether theorem, $\mathbf{G}$ is therefore the image of a connected group under the morphism (B.1) and as morphisms map Zariski-connected sets to Zariski-connected sets, the claim follows. $\qquad\square$

In what follows, we let $D = (d+1)^2$. We identify $\mathbf{G}(k)$ with its image in $\mathrm{GL}_D(k)$ given by the basis $\mathcal{E}_{d+1}$.

Definition B.2. Let k be a field and let $R \hookrightarrow k$ a subring. Then

$$\mathbf{G}(R) = \mathbf{G}(k) \cap \mathrm{GL}_D(R).$$

We denote by $\mathcal{V}_{\mathrm{f}} \subseteq \mathbb{N}$ the set of finite rational primes and we let $\mathcal{V} = \mathcal{V}_{\mathrm{f}} \cup \{\infty\}$. The following definition of adelic points and integral adelic points of a $\mathbb{Q}$-group is formulated for a general algebraic $\mathbb{Q}$-subgroup $\mathbf{H}$ of GL_D. It encompasses in particular the cases $\mathbf{H} = \mathbf{G}$ and $\mathbf{H} = \mathrm{GL}_D$.

Definition B.3. Let $S \subseteq \mathcal{V}$, $S_{\mathrm{f}} = S \setminus \{\infty\}$, and $\mathbf{H} \leq \mathrm{GL}_D$ be a $\mathbb{Q}$-subgroup. We set

$$\mathbf{H}(\mathbb{Z}_{S_{\mathrm{f}}}) = \prod_{p \in S_{\mathrm{f}}} \mathbf{H}(\mathbb{Z}_p),$$

$$\mathbf{H}(\mathbb{Q}_{S_{\mathrm{f}}}) = \left\{ (g_p)_{p \in S_{\mathrm{f}}} \in \prod_{p \in S_{\mathrm{f}}} \mathbf{H}(\mathbb{Q}_p) : g_p \in \mathbf{H}(\mathbb{Z}_p) \text{ for all but finitely many } p \in S_{\mathrm{f}} \right\}.$$

If $\infty \in S$, then $\mathbf{H}(\mathbb{Q}_S) = \mathbf{H}(\mathbb{R}) \times \mathbf{H}(\mathbb{Q}_{S_{\mathrm{f}}})$. If $S_{\mathrm{f}} = \mathcal{V}_{\mathrm{f}}$, we set $\mathbf{H}(\widehat{\mathbb{Z}}) = \mathbf{H}(\mathbb{Z}_{S_{\mathrm{f}}})$, $\mathbf{H}(\mathbb{A}_{\mathrm{f}}) = \mathbf{H}(\mathbb{Q}_{S_{\mathrm{f}}})$, $\mathbf{H}(\mathbb{A}) = \mathbf{H}(\mathbb{R}) \times \mathbf{H}(\mathbb{A}_{\mathrm{f}})$, and $\mathbf{H}(\mathbb{R} \times \widehat{\mathbb{Z}}) = \mathbf{H}(\mathbb{R}) \times \mathbf{H}(\widehat{\mathbb{Z}})$.

B.2. **The fundamental compact-open subgroups.** Given $p \in \mathcal{V}_{\mathrm{f}}$ and $v \in \mathbb{N}$ we define a map

$$\pi_{p,v} : \mathrm{Mat}_D(\mathbb{Z}_p) \to \mathrm{Mat}_D(\mathbb{Z}/p^v\mathbb{Z})$$

by coordinate-wise reduction mod p^v. This map clearly defines a ring homomorphism and for any $x \in \mathrm{Mat}_D(\mathbb{Z}_p)$ we have

$$(\det \circ \pi_{p,v})(x) \equiv \det(x) \bmod p^v.$$

As $\mathbb{Z}_p^\times$ and $(\mathbb{Z}/p^v\mathbb{Z})^\times$ consist precisely of the elements whose projections mod p^v do not vanish, this induces a group homomorphism $\pi_{p,v} : \mathrm{GL}_D(\mathbb{Z}_p) \to \mathrm{GL}_D(\mathbb{Z}/p^v\mathbb{Z})$. For the sake of completeness, we argue that it is surjective. To this end one notes that $\mathbb{Z}/p^v\mathbb{Z}$ is a semi-local ring, so that $\mathrm{SL}_D(\mathbb{Z}/p^v\mathbb{Z})$ is generated by elementary matrices; cf. [HO89, Thm. 4.3.9]. Therefore $\pi_{p,v}$ restricts to an epimorphism from $\mathrm{SL}_D(\mathbb{Z}_p)$ to $\mathrm{SL}_D(\mathbb{Z}/p^v\mathbb{Z})$. Now one uses that

$$\mathrm{GL}_D(\mathbb{Z}/p^v\mathbb{Z}) \cong (\mathbb{Z}/p^v\mathbb{Z})^\times \ltimes \mathrm{SL}_D(\mathbb{Z}/p^v\mathbb{Z}),$$

where $(\mathbb{Z}/p^v\mathbb{Z})^\times$ identifies with the set of matrices of the form

$$\left\{ \begin{pmatrix} a & 0 \\ 0 & \mathrm{Id}_{D-1} \end{pmatrix} : a \in (\mathbb{Z}/p^v\mathbb{Z})^\times \right\}.$$

In what follows, we will denote $L_p[p^v] = \ker \pi_{p,v}$ and $L_p[1] = \mathrm{GL}_D(\mathbb{Z}_p)$. We set

$$K_p[p^v] = L_p[p^v] \cap \mathbf{G}(\mathbb{Z}_p).$$

Lemma B.4. *The family $\{L_p[p^v]\}$ is a basis of open neighbourhoods of the identity in $\mathrm{GL}_D(\mathbb{Z}_p)$. In particular, the family $\{K_p[p^v] : v \in \mathbb{N}\}$ is a basis of open neighbourhoods of the identity in $\mathbf{G}(\mathbb{Z}_p)$.*

Proof. The group $L_p[p^v] \leq \mathrm{GL}_D(\mathbb{Z}_p)$ is closed and has finite index, therefore it is open. The topology on $\mathrm{GL}_D(\mathbb{Z}_p)$ is induced by the metric $\|\cdot\|_p$ on $\mathrm{Mat}_D(\mathbb{Z}_p)$ given by

$$\|x\|_p = \max\{|x_{i,j}|_p : 1 \leq i,j \leq D\} \quad (x \in \mathrm{Mat}_D(\mathbb{Z}_p)).$$

Let $v \in \mathbb{N}$ and $x \in \mathrm{Mat}_D(\mathbb{Z}_p)$, then

$$\|\mathrm{Id}_D - x\|_p \leq p^{-v} \iff x \in \mathrm{Id}_D + p^v \mathrm{Mat}_D(\mathbb{Z}_p).$$

In particular, the collection

$$\{\mathrm{Id}_D + p^v \mathrm{Mat}_D(\mathbb{Z}_p) : v \in \mathbb{N}\}$$

is a basis of open neighbourhoods of the identity in $\mathrm{GL}_D(\mathbb{Z}_p)$. $\qquad\qquad\square$

Given $S_{\mathrm{f}} \subseteq \mathcal{V}_{\mathrm{f}}$, we denote by $\mathcal{I}_{S_{\mathrm{f}}}$ the set of natural numbers whose prime factorization involves only primes contained in S_{f}. Given $N \in \mathcal{I}_{S_{\mathrm{f}}}$, we define $(v_p(N))_{p \in \mathcal{I}_{S_{\mathrm{f}}}}$ by $N = \prod_{p \in S_{\mathrm{f}}} p^{v_p(N)}$. We set

$$L_{S_{\mathrm{f}}}[N] = \prod_{p \in S_{\mathrm{f}}} L_p[p^{v_p(N)}]$$

and $K_{S_{\mathrm{f}}}[N] = L_{S_{\mathrm{f}}}[N] \cap \mathbf{G}(\mathbb{Z}_{S_{\mathrm{f}}})$. If $S_{\mathrm{f}} = \mathcal{V}_{\mathrm{f}}$, we write $L_{\mathrm{f}}[N]$ and $K_{\mathrm{f}}[N]$ for $L_{S_{\mathrm{f}}}[N]$ and $K_{S_{\mathrm{f}}}[N]$ respectively.

Corollary B.5. *The family $\{L_{S_{\mathrm{f}}}[N] : N \in \mathcal{I}_{S_{\mathrm{f}}}\}$ forms a basis of compact open neighbourhoods of the identity in $\mathrm{GL}_D(\mathbb{Z}_{S_{\mathrm{f}}})$. In particular, the family $\{K_{S_{\mathrm{f}}}[N] : N \in \mathcal{I}_{S_{\mathrm{f}}}\}$ forms a basis of compact open neighbourhoods of the identity in $\mathbf{G}(\mathbb{Z}_{S_{\mathrm{f}}})$.*

Proof. The groups $L_{S_{\mathrm{f}}}[N]$ are open by definition of the product topology. Compactness follows from Tychonov's theorem. In order to prove that they form a neighbourhood basis, let $V \subseteq \mathrm{GL}_D(\mathbb{Z}_{S_{\mathrm{f}}})$ be an open neighbourhood of the identity. Then there is a finite set $T_{\mathrm{f}} \subseteq S_{\mathrm{f}}$ and for all $p \in T_{\mathrm{f}}$ an open neighbourhood V_p of the identity in $\mathrm{GL}_D(\mathbb{Z}_p)$ such that

$$\prod_{p \in T_{\mathrm{f}}} V_p \times \prod_{p \in S_{\mathrm{f}} \setminus T_{\mathrm{f}}} \mathrm{GL}_D(\mathbb{Z}_p) \subseteq V.$$

Given $p \in T_{\mathrm{f}}$, let $v_p \in \mathbb{N}$ be such that $L_p[p^{v_p}] \subseteq V_p$ and define $N = \prod_{p \in T_{\mathrm{f}}} p^{v_p}$. Then, $L_{S_{\mathrm{f}}}[N] \subseteq V$ by definition. $\qquad\qquad\square$

B.3. **Principal congruence subgroups.** Similar to what was done in Section B.2, we can define for any $N \in \mathbb{N}$, with $N \geq 2$, the group homomorphism $\varpi_N : \mathrm{GL}_D(\mathbb{Z}) \to \mathrm{GL}_D(\mathbb{Z}/N\mathbb{Z})$ given by projection mod N. Note that ϖ_N is *not* surjective. We let $\Lambda(N) = \ker \varpi_N$ and $\Gamma(N) = \Lambda(N) \cap \mathbf{G}(\mathbb{Z})$. We also define $\Lambda(1) = \mathrm{GL}_D(\mathbb{Z})$ and $\Gamma(1) = \mathbf{G}(\mathbb{Z})$.

Definition B.6. Let $N \in \mathbb{N}$. The family $\{\Gamma(N) : N \in \mathbb{N}\}$ is called the family of principal congruence subgroups. A subgroup $\Delta \leq \Gamma(1)$ is a congruence subgroup if it contains a principal congruence subgroup.

Let $S_{\mathrm{f}} \subseteq \mathcal{V}_{\mathrm{f}}$. In what follows, we will view

$$\mathbb{Z}[S_{\mathrm{f}}^{-1}] = \mathbb{Z}[\tfrac{1}{p} : p \in S_{\mathrm{f}}]$$

as a subring of $\mathbb{Q}_{S_{\mathrm{f}}}$ by embedding it diagonally. Similarly, $\mathrm{GL}_D(\mathbb{Z}[S_{\mathrm{f}}^{-1}])$ and $\mathbf{G}(\mathbb{Z}[S_{\mathrm{f}}^{-1}])$ become subgroups of $\mathrm{GL}_D(\mathbb{Q}_{S_{\mathrm{f}}})$ and $\mathbf{G}(\mathbb{Q}_{S_{\mathrm{f}}})$.

We are now ready to prove the first main result of this section.

Proposition B.7. *Let $S_f \subseteq \mathcal{V}_f$ and $N \in \mathcal{I}_{S_f}$. Then, $\Gamma(N) = K_{S_f}[N] \cap \mathbf{G}(\mathbb{Z}[S_f^{-1}])$. In particular, for all N, we have*

$$\Gamma(N) = K_f[N] \cap \mathbf{G}(\mathbb{Q}).$$

Proof. It suffices to show that

$$\Lambda(N) = L_{S_f}[N] \cap \mathrm{GL}_D(\mathbb{Z}[S_f^{-1}]).$$

If $g \in \Lambda(N)$, i.e., $g \in \mathrm{GL}_D(\mathbb{Z})$ and $g \equiv \mathrm{Id}_D \bmod N$, then clearly for all $p|N$ we have $g \equiv \mathrm{Id}_D \bmod p^{v_p(N)}$. As $\det g \in \{\pm 1\} \subseteq \mathbb{Z}_p^\times$ for all $p \in S_f$, we have $g \in \mathrm{GL}_D(\mathbb{Z}_p)$ for all $p \in S_f$. Combining these two facts, we obtain that $\Lambda(N) \subseteq L_{S_f}[N] \cap \mathrm{GL}_D(\mathbb{Z}[S_f^{-1}])$.

Before we turn to the opposite inclusion, we note that

$$\mathrm{GL}_D(\mathbb{Z}_{S_f}) \cap \mathrm{GL}_D(\mathbb{Z}[S_f^{-1}]) = \mathrm{GL}_D(\mathbb{Z}).$$

This, in particular, implies the result in the special case $N = 1$. The inclusion $\mathrm{GL}_D(\mathbb{Z}) \subseteq \mathrm{GL}_D(\mathbb{Z}_{S_f}) \cap \mathrm{GL}_D(\mathbb{Z}[S_f^{-1}])$ is clear. For the opposite inclusion, one first notes that $\mathbb{Z}[S_f^{-1}] \cap \mathbb{Z}_{S_f} = \mathbb{Z}$ and hence $\mathrm{GL}_D(\mathbb{Z}_{S_f}) \cap \mathrm{GL}_D(\mathbb{Z}[S_f^{-1}]) \subseteq \mathrm{Mat}_D(\mathbb{Z})$. Let $g \in \mathrm{GL}_D(\mathbb{Z}_{S_f}) \cap \mathrm{GL}_D(\mathbb{Z}[S_f^{-1}])$. Then, $\det g \in \mathbb{Z}$. On the other hand $\det g \in \mathbb{Z}_{S_f}^\times$, i.e., we have $\det g \in \mathbb{Z}_p^\times$ for all $p \in S_f$. This means that $\det g$ is coprime to p for all $p \in S_f$. But, since $\det g \in \mathbb{Z}[S_f^{-1}]^\times$, we get $\det g \in \{\pm 1\}$. It follows that $\mathrm{GL}_D(\mathbb{Z}_{S_f}) \cap \mathrm{GL}_D(\mathbb{Z}[S_f^{-1}]) \subseteq \mathrm{GL}_D(\mathbb{Z})$.

Let now $g \in L_{S_f}[N] \cap \mathrm{GL}_D(\mathbb{Z}[S_f^{-1}])$. In particular

$$g \in \mathrm{GL}_D(\mathbb{Z}_{S_f}) \cap \mathrm{GL}_D(\mathbb{Z}[S_f^{-1}]) = \mathrm{GL}_D(\mathbb{Z}).$$

Therefore reduction mod N is just the standard reduction. By assumption we have for all $p \in S_f$ that $g \equiv \mathrm{Id}_D \bmod p^{v_p(N)}$ and in particular $g \in \Lambda(N)$. $\qquad\square$

B.4. Finiteness of class number and principal congruence subgroups.

Similarly to what we did earlier, we will now regard $\mathbb{Q}$ as a subfield of the ring $\mathbb{A} = \mathbb{R} \times \mathbb{A}_f$ by diagonal embedding. Notice that this embedding differs from the composition of embeddings $\mathbb{Q} \hookrightarrow \mathbb{A}_f \hookrightarrow \mathbb{A}$. Similarly, we can view $\mathbf{G}(\mathbb{Q})$ as a subgroup of $\mathbf{G}(\mathbb{A})$. It was proven by Borel and Harish-Chandra that $\mathbf{G}(\mathbb{Q})$ is a lattice in $\mathbf{G}(\mathbb{A})$.

Proposition B.8. *Let*

$$X_{\mathbb{A},1} = \mathbf{G}(\widehat{\mathbb{Z}}) \backslash \mathbf{G}(\mathbb{A}) / \mathbf{G}(\mathbb{Q}).$$

Then, $\mathbf{G}_\infty$ acts transitively on $X_{\mathbb{A},1}$, i.e., $\mathbf{G}_\infty \backslash X_{\mathbb{A},1}$ is a singleton.

Proof. We first claim that $\mathbf{G}(\mathbb{A}) = \Phi(\mathrm{GL}_{d+1}(\mathbb{A}))$. To this end, let $g \in \mathbf{G}(\mathbb{A})$ and using Skolem-Noether choose $x_p \in \mathrm{GL}_{d+1}(\mathbb{Q}_p)$, $p \in \mathcal{V}_f$, such that $g_p = \Phi(x_p)$. By definition we have that $g_p \in \mathbf{G}(\mathbb{Z}_p)$ for all but finitely many $p \in \mathcal{V}_f$. Recall that $g_p \in \mathbf{G}(\mathbb{Z}_p)$ implies that we can assume $x_p \in \mathrm{GL}_{d+1}(\mathbb{Z}_p)$; cf. the proof of Lemma 3.1. It follows in particular that $g = \Phi(x)$ for some $x \in \mathrm{GL}_{d+1}(\mathbb{A})$.

Let $g \in \mathbf{G}(\mathbb{A})$ arbitrary and choose $x \in \mathrm{GL}_{d+1}(\mathbb{A})$ such that $g = \Phi(x)$. By [PR94, Prop. 8.1], we know that GL_{d+1} has class number one, i.e., $x = k\gamma$, where $k \in \mathrm{GL}_{d+1}(\mathbb{R} \times \widehat{\mathbb{Z}})$ and $\gamma \in \mathrm{GL}_{d+1}(\mathbb{Q})$. In particular, we have $g = \Phi(k)\Phi(\gamma)$. Note that $\Phi(\gamma) \in \mathrm{GL}_D(\mathbb{Q})$ by rationality of the representation Φ. Note that

$$k_p \mathrm{Mat}_{d+1}(\mathbb{Z}_p) k_p^{-1} = \mathrm{Mat}_{d+1}(\mathbb{Z}_p)$$

and hence $\Phi(k_p) \in \mathrm{GL}_D(\mathbb{Z}_p)$ for all $p \in \mathcal{V}_f$, i.e., $\Phi(k) \in \mathrm{GL}_D(\mathbb{R} \times \widehat{\mathbb{Z}})$. Therefore

$$g \in \mathbf{G}(\mathbb{R} \times \widehat{\mathbb{Z}})\mathbf{G}(\mathbb{Q}).$$

$\qquad\square$

Proposition B.9. *Let $N \in \mathbb{N}$. Then, the double quotient*

$$X_{\mathbb{A},N} = {}_{K_{\mathrm{f}}[N]}\backslash {}^{\mathbf{G}(\mathbb{A})}/_{\mathbf{G}(\mathbb{Q})}$$

is a finite union of $\mathbf{G}_\infty$-orbits. Let $x \in X_{\mathbb{A},N}$, then

$$\mathbf{G}_\infty.x \cong {}^{\mathbf{G}_\infty}/_{\Gamma(N)}$$

as $\mathbf{G}_\infty$-spaces, i.e., $X_{\mathbb{A},N}$ is a disjoint union of finitely many copies of $\mathbf{G}_\infty/\Gamma(N)$.

Proof. Recall that $K_{\mathrm{f}}[1] = \mathbf{G}(\widehat{\mathbb{Z}})$ by definition. By Proposition B.8, we know that $\mathbf{G}_\infty\backslash X_{\mathbb{A},1}$ is a singleton. As $K_{\mathrm{f}}[N] \leq K_{\mathrm{f}}[1]$ is a finite index subgroup, the finiteness of $\mathbf{G}_\infty\backslash X_{\mathbb{A},N}$ follows immediately.

For the second part, using Proposition B.8, let $\mathcal{R}_{D,N} \subseteq \mathbf{G}(\widehat{\mathbb{Z}})$ be a set of representatives for the double quotient $(\mathbf{G}_\infty \times K_{\mathrm{f}}[N])\backslash \mathbf{G}(\mathbb{A})/\mathbf{G}(\mathbb{Q})$. After possibly multiplying x by an element in $\mathbf{G}_\infty$, we can assume that $x = K_{\mathrm{f}}[N]\eta\mathbf{G}(\mathbb{Q})$ for some $\eta \in \mathcal{R}_{D,N}$. We will show that in this case

$$\mathrm{Stab}_{\mathbf{G}_\infty}(x) = \Gamma(N).$$

Let $g_\infty \in \mathbf{G}_\infty$, then

$$g_\infty \cdot K_{\mathrm{f}}[N]\eta\mathbf{G}(\mathbb{Q}) = K_{\mathrm{f}}[N]\eta\mathbf{G}(\mathbb{Q}) \iff \exists \gamma \in \mathbf{G}(\mathbb{Q}) \exists k \in K_{\mathrm{f}}[N],\ (g_\infty, \eta) = (\gamma, k\eta\gamma)$$
$$\iff g_\infty \in \mathbf{G}(\mathbb{Q}) \cap \eta^{-1}K_{\mathrm{f}}[N]\eta.$$

Now note that $K_{\mathrm{f}}[N]$ is the kernel of the group homomorphism

$$\Psi : \mathbf{G}(\widehat{\mathbb{Z}}) \to \prod_{p|N} \mathbf{G}(\mathbb{Z}/p^{v_p(N)}\mathbb{Z}), \qquad \Psi((k_p)_{p\in\mathcal{V}_{\mathrm{f}}}) = (k_p \bmod p^{v_p(N)}\mathbb{Z})_{p|N}.$$

In particular, $K_{\mathrm{f}}[N]$ is a normal subgroup and thus $\eta^{-1}K_{\mathrm{f}}[N]\eta = K_{\mathrm{f}}[N]$. Hence, if $x \in X_{\mathbb{A},N}$ is arbitrary, letting $\eta \in \mathcal{R}_{D,N}$ be such that $x = K_{\mathrm{f}}[N]\eta\mathbf{G}(\mathbb{Q})$, Proposition B.7 implies that

$$g_\infty \in \mathrm{Stab}_{\mathbf{G}_\infty}(x) \iff g_\infty \in \mathbf{G}(\mathbb{Q}) \cap \eta^{-1}K_{\mathrm{f}}[N]\eta = \Gamma(N).$$

$\square$

B.5. **Correspondence in the S-arithmetic setup.** We deduce analogous results to those obtained in the previous section for quotients of $\mathbf{G}_S$. In particular, the decomposition in (3.5) follows by Corollary B.11. First, we need the following.

Corollary B.10. *Let $S_{\mathrm{f}} \subseteq \mathcal{V}_{\mathrm{f}}$ finite and let $S = \{\infty\} \cup S_{\mathrm{f}}$. Then,*

$$\mathbf{G}(\mathbb{R} \times \mathbb{Z}_{S_{\mathrm{f}}})\Gamma_S = \mathbf{G}(\mathbb{Q}_S).$$

Proof. Let

$$M_{S_{\mathrm{f}}} = \prod_{p\in\mathcal{V}_{\mathrm{f}}\backslash S_{\mathrm{f}}} \mathbf{G}(\mathbb{Z}_p).$$

We first claim that

$${}_{M_{S_{\mathrm{f}}}}\backslash {}^{\mathbf{G}(\mathbb{A})}/_{\mathbf{G}(\mathbb{Q})} \cong {}^{\mathbf{G}(\mathbb{Q}_S)}/_{\Gamma_S}$$

as $\mathbf{G}(\mathbb{Q}_S)$-spaces. To this end, we note that $\mathbf{G}(\mathbb{Q}_S)$ acts transitively on the left hand side by Proposition B.8. Denote by x_0 the identity coset in the left-hand side double quotient. In particular, it remains to show that

$$\mathrm{Stab}_{\mathbf{G}(\mathbb{Q}_S)}(x_0) = \Gamma_S.$$

Let e denote the identity in $\prod_{p\notin S} \mathbf{G}(\mathbb{Z}_p)$ and $g \in \mathrm{Stab}_{\mathbf{G}(\mathbb{Q}_S)}(x_0)$. Arguing as in the proof of Proposition B.7, we get that $g \in M_{S_{\mathrm{f}}} \cap \mathbf{G}(\mathbb{Q}) = \Gamma_S$ as desired.

$\square$

The following is the analogue of Proposition B.9 and implies (3.4).

Corollary B.11. *Let $N \in \mathbb{N}$ and $S_{\mathrm{f}} \subseteq \mathcal{V}_{\mathrm{f}}$ such that $v_p(N) \neq 0 \implies p \in S_{\mathrm{f}}$. Let $\mathcal{R}_{D,N}$ be as in the proof of Proposition B.9. Then, the projection $\mathcal{R}_{D,N,S_{\mathrm{f}}}$ of $\mathcal{R}_{D,N}$ to $K_{S_{\mathrm{f}}}[1]$ is a set of representatives of the $\mathbf{G}_\infty$-orbits in*

$$X_{\mathbb{Q}S,N} = {}_{K_{S_{\mathrm{f}}}[N]}\backslash \mathbf{G}_S/_{\Gamma_S}.$$

Moreover, the map $\mathcal{R}_{D,N} \to \mathcal{R}_{D,N,S_{\mathrm{f}}}$ is a bijection and $X_{\mathbb{Q}S,N}$ is a disjoint union of $|\mathcal{R}_{D,N}|$-many copies of $\mathbf{G}_\infty/\Gamma(N)$.

Proof. Note that $K_{\mathrm{f}}[N] = K_{S_{\mathrm{f}}}[N] \times M_{S_{\mathrm{f}}}$. Therefore, as in the proof of Corollary B.10, we obtain

$$X_{\mathbb{Q}S,N} \cong {}_{K_{S_{\mathrm{f}}}[N]}\backslash (M_{S_{\mathrm{f}}}\backslash \mathbf{G}(\mathbb{A})/\mathbf{G}(\mathbb{Q})) \cong {}_{K_{\mathrm{f}}[N]}\backslash \mathbf{G}(\mathbb{A})/_{\mathbf{G}(\mathbb{Q})}.$$

Looking at these isomorphisms more explicitly, it is easy to check that $\mathcal{R}_{D,N,S_{\mathrm{f}}}$ is a set of representatives which is in one-to-one correspondence with $\mathcal{R}_{D,N}$. We leave the rest of the proofs to the reader. $\qquad\square$

B.6. **Counting connected components.** In this section, we aim at finding a uniform bound on the number of connected components of

$$X_{\mathbb{Q}S,N} = {}_{K_{S_{\mathrm{f}}}[N]}\backslash \mathbf{G}_S/_{\Gamma_S},$$

independently of N, where we assume that $S_{\mathrm{f}} \subseteq \mathcal{V}_{\mathrm{f}}$ is finite, $S = S_{\mathrm{f}} \cup \{\infty\}$, and $N \in \mathcal{I}_{S_{\mathrm{f}}}$. The main result is Proposition B.13.

We first need a lemma about $\mathbf{G}_S^+$-orbits. Given $p \in \mathcal{V}_{\mathrm{f}}$ and $n \in \mathbb{N}$, let $S_p[p^n] \leq \mathrm{SL}_{d+1}(\mathbb{Z}_p)$ be the kernel of the homomorphism $\mathrm{SL}_{d+1}(\mathbb{Z}_p) \to \mathrm{SL}_{d+1}(\mathbb{Z}/p^n\mathbb{Z})$ given by reduction mod p^n. For $n = 0$, we let $S_p[p^n] = \mathrm{SL}_{d+1}(\mathbb{Z}_p)$. Recall the representation $\Phi : \mathrm{GL}_{d+1} \to \mathbf{G}$ defined in (B.1). A calculation shows that $\Phi(S_p[p^n]) \subseteq K_p[p^n]$. In what follows, we let $K_p[p^n]^+ = \Phi(S_p[p^n])$ and, given $N \in \mathcal{I}_{S_{\mathrm{f}}}$, we define

$$K_{S_{\mathrm{f}}}[N]^+ = \prod_{p \in S_{\mathrm{f}}} K_p[p^{v_p(N)}]^+ \subseteq K_{S_{\mathrm{f}}}[N].$$

In fact, $K_{S_{\mathrm{f}}}[N]^+$ is normal in $\mathbf{G}(\mathbb{Z}_{S_{\mathrm{f}}})$. To this end, one checks that $S_p[p^n] \lhd \mathrm{GL}_{d+1}(\mathbb{Z}_p)$ is a normal subgroup and then uses that

$$\mathbf{G}(\mathbb{Z}_p) = \Phi(\mathrm{GL}_{d+1}(\mathbb{Z}_p))$$

as already argued in the proof of Lemma 3.1.

Similarly, we note that $\Phi(\mathrm{SL}_{d+1}(\mathbb{Z}[S_{\mathrm{f}}^{-1}])) \subseteq \Gamma_S$, and we will denote

$$\Gamma_S^+ = \Phi(\mathrm{SL}_{d+1}(\mathbb{Z}[S_{\mathrm{f}}^{-1}])).$$

As SL_{d+1} has the strong approximation property, we find

$$\mathbf{G}_S^+ = (\mathbf{G}_\infty^+ \times K_{S_{\mathrm{f}}}[N]^+)\Gamma_S^+ \tag{B.2}$$

for all $N \in \mathcal{I}_{S_{\mathrm{f}}}$; cf. [PR94, Thm. 7.12].

Lemma B.12. *Let $S_{\mathrm{f}} \subseteq \mathcal{V}_{\mathrm{f}}$, $S = S_{\mathrm{f}} \cup \{\infty\}$, $N \in \mathcal{I}_{S_{\mathrm{f}}}$, and $x \in \mathbf{G}_S/\Gamma_S$. Then,*

$$\mathbf{G}_S^+.x = (\mathbf{G}_\infty^+ \times K_{S_{\mathrm{f}}}[N]^+).x.$$

Proof. This follows immediately from Corollary B.10 and (B.2). To this end let $g \in \mathbf{G}_\infty \times K_{S_{\mathrm{f}}}[N]$ and $\eta \in \mathbf{G}(\mathbb{Z}_{S_{\mathrm{f}}})$ such that $x = g\eta\Gamma_S$. As $K_{S_{\mathrm{f}}}[N]^+$ is normal in $\mathbf{G}(\mathbb{Z}_{S_{\mathrm{f}}})$, it follows that

$$\mathbf{G}_S^+.x = g\eta\mathbf{G}_S^+\Gamma_S = g\eta(\mathbf{G}_\infty^+ \times K_{S_{\mathrm{f}}}[N]^+)\Gamma_S = (\mathbf{G}_\infty^+ \times K_{S_{\mathrm{f}}}[N]^+).x. \tag{B.3}$$

$\qquad\square$

In what follows, we let

$$\mathbf{G}_{\mathrm{char},S} = {}_{\mathbf{G}_S^+}\backslash \mathbf{G}_S/_{\Gamma_S}$$

and we note that $\mathbf{G}_{\mathrm{char}}$ is a finite set whose cardinality is bounded by the index $[\mathbf{G}_S : \mathbf{G}_S^+]$.

Proposition B.13. *We have*

$$\mathbf{G}_\infty^+\backslash X_{\mathbb{Q}_S,N} \cong K_{S_{\mathrm{f}}}[N]\mathbf{G}_S^+\backslash\mathbf{G}_S/\Gamma_S \cong K_{S_{\mathrm{f}}}[N]\backslash\mathbf{G}_{\mathrm{char},S}.$$

In particular, given a finite set $S \subseteq \mathcal{V}$ containing ∞, the number of connected components of $X_{\mathbb{Q}_S,N}$ is at most $[\mathbf{G}_S : \mathbf{G}_S^+]$.

Proof. For the first bijection, we recall from the proof of Lemma B.12 that $K_{S_{\mathrm{f}}}[N]^+ \le K_{S_{\mathrm{f}}}[N]$ and therefore (B.3) yields that for all $g \in \mathbf{G}_S$

$$(\mathbf{G}_\infty^+ \times K_{S_{\mathrm{f}}}[N])g\Gamma_S = K_{S_{\mathrm{f}}}[N](\mathbf{G}_\infty^+ \times K_{S_{\mathrm{f}}}[N]^+)g\Gamma_S$$
$$= K_{S_{\mathrm{f}}}[N]\mathbf{G}_S^+ g\Gamma_S.$$

The second bijection follows by definition of $\mathbf{G}_{\mathrm{char},S}$. The last part of the statement follows by definition of $\mathbf{G}_{\mathrm{char},S}$. $\square$

The following corollary follows immediately.

Corollary B.14. *Let $W \subseteq K_{S_{\mathrm{f}}}$ be an open subgroup. Then the quotient $W\backslash\mathbf{G}_S/\Gamma_S$ is a union of at most $[\mathbf{G}_S : \mathbf{G}_S^+]$-many $\mathbf{G}_\infty^+$-orbits.*

Appendix C. KAK-decomposition and norms

In this section we will introduce a function on $\mathbf{G}(\mathbb{Q}_v)$ which measures the size of an element. These functions are used in Section 4.3 to define norm-balls on $\mathbf{G}(\mathbb{Q}_S)$. The main input is the KAK-decomposition, which is well-known for $v = \infty$. We only discuss the case where v is a finite place of $\mathbb{Q}$. In what follows, p is a natural prime.

Lemma C.1. *Let $D \in \mathbb{N}$ and denote by $\|\cdot\|_p : \mathrm{Mat}_D(\mathbb{Q}_p) \to [0, \infty)$ the norm defined for $x = (x_{ij})$ by*

$$\|x\|_p = \max\{|x_{ij}|_p : 1 \le i, j \le D\}.$$

Then, $\|\cdot\|_p$ is an operator norm. Moreover, for all $k \in \mathrm{GL}_D(\mathbb{Z}_p)$ we have $\|k\|_p = 1$.

Proof. The fact that $\|\cdot\|_p$ is an operator norm follows from the ultrametric property of the p-adic absolute value. We leave the details to the reader. For the second part we note that for any $x \in \mathrm{GL}_D(\mathbb{Z}_p)$ we have $|\det x|_p = 1$ and $\|x\|_p \le 1$. On the other hand we have $|\det x|_p \le \|x\|_p^D$ by the ultrametric property. Hence $\|k\|_p = 1$ for all $k \in \mathrm{GL}_D(\mathbb{Z}_p)$. $\square$

In the following discussion, Lemma C.1 is used in the form of the following corollary.

Corollary C.2. *Let $D \in \mathbb{N}$, $x \in \mathrm{Mat}_D(\mathbb{Q}_p)$ and $k_1, k_2 \in \mathrm{GL}_D(\mathbb{Z}_p)$. Then*

$$\|k_1 x k_2\|_p = \|x\|_p.$$

Proof. Since operator norms are submultiplicative, we have

$$\|x\|_p \le \|k_1^{-1}\|_p\|k_1 x k_2\|_p\|k_2^{-1}\|_p = \|k_1 x k_2\|_p \le \|k_1\|_p\|x\|_p\|k_2\|_p = \|x\|_p.$$

$\square$

Lemma C.3. *Let $x \in \mathrm{GL}_2(\mathbb{Q}_p)$. Then there are $m_1, m_2 \in \mathrm{GL}_2(\mathbb{Z}_p)$ and $n, k \in \mathbb{Z}$ such that*

$$x = m_1 \begin{pmatrix} p^n & 0 \\ 0 & p^k \end{pmatrix} m_2.$$

Proof. Let $x = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Assume that $\|x\|_p = |a|_p$ and write $a = p^m u$, $b = p^n v$ and $c = p^\ell w$ with $u, v, w \in \mathbb{Z}_p^\times$. Then

$$\begin{pmatrix} 1 & 0 \\ -p^{\ell-m}\frac{w}{u} & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & -p^{n-m}\frac{v}{u} \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & d' \end{pmatrix}.$$

As we assumed that a was maximal, it follows that the two unipotent matrices lie in $\mathrm{GL}_2(\mathbb{Z}_p)$. Denote them by m_1, m_2 respectively. The resulting diagonal matrix is of the form

$$\begin{pmatrix} a & 0 \\ 0 & d' \end{pmatrix} = \begin{pmatrix} u & 0 \\ 0 & z \end{pmatrix} \begin{pmatrix} p^m & 0 \\ 0 & p^k \end{pmatrix}$$

with $u, z \in \mathbb{Z}_p^\times$, and thus we have shown that

$$x = m_1^{-1} \begin{pmatrix} u & 0 \\ 0 & z \end{pmatrix} \begin{pmatrix} p^n & 0 \\ 0 & p^k \end{pmatrix} m_2^{-1}.$$

As $m_1 \mathrm{diag}(u, z) \in \mathrm{GL}_2(\mathbb{Z}_p)$, we obtain the claim under the assumption that $\|x\|_p = |a|_p$. If $\|x\|_p \neq |a|_p$, then we distinguish two cases. If $\|x\|_p = |d|_p$, then we conjugate x by the matrix $m = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, so that the maximal entry comes to lie in the top left corner. If $\|x\|_p = |b|_p$ or $\|x\|_p = |c|_p$, then it is a property of non-archimedian absolute values that for either $x \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ or $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} x$ the top left entry will be maximal. These operations all follow from multiplying x with matrices in $\mathrm{GL}_2(\mathbb{Z}_p)$ and thus the claim is proven. $\qquad\square$

In what follows, given $x \in \mathrm{GL}_{d+1}(\mathbb{Q}_p)$, we denote by $[x] \in \mathbf{G}(\mathbb{Q}_p)$ its image under Φ as introduced in the proof of Proposition B.8.

Proposition C.4. *Let p be a finite rational prime and assume that $g \in \mathbf{G}(\mathbb{Q}_p)$. Then there exist uniquely determined nonnegative integers $n_1 \geq \cdots \geq n_d$ as well as elements $k_1, k_2 \in \mathbf{G}(\mathbb{Z}_p)$ such that*

$$g = k_1 [\mathrm{diag}(p^{-n_1}, \ldots, p^{-n_d}, 1)] k_2.$$

Proof. As $[\mathrm{GL}_{d+1}(\mathbb{Z}_p)] = \mathbf{G}(\mathbb{Z}_p)$, cf. the proof of Proposition B.8, it suffices to prove the existence of a decomposition of any element $x \in \mathrm{GL}_{d+1}(\mathbb{Q}_p)$ into a product of the form $k_1 D k_2$, where D is a diagonal matrix whose non-zero entries are powers of p for a set of exponents uniquely determined (with multiplicity) by x. If this is the case, we can use elements in $\mathrm{GL}_{d+1}(\mathbb{Z}_p)$ to arrange the diagonal entries in decreasing order with respect to the p-adic absolute value. Furthermore, there will be an element ω in the center of $\mathrm{GL}_{d+1}(\mathbb{Q}_p)$ such that $\omega^{-1} D$ is of the form required by the proposition and we note that $[\omega^{-1} D] = [D]$.

So let $x \in \mathrm{GL}_{d+1}(\mathbb{Q}_p)$ be arbitrary and in view of Lemma C.3 assume that $d \geq 2$. For the existence of a decomposition of x, we apply elements in $\mathrm{GL}_{d+1}(\mathbb{Z}_p)$ so that $\|x\|_p = |x_{11}|_p$. Then one can use the copies of $\mathrm{SL}_2(\mathbb{Z}_p)$ in $\mathrm{GL}_{d+1}(\mathbb{Q}_p)$ associated with spans of pairs of the standard basis to reduce the matrix x to a matrix of the form

$$x' = \begin{pmatrix} x_{11} & 0 \\ 0 & y \end{pmatrix},$$

where $y \in \mathrm{GL}_d(\mathbb{Q}_p)$ and $\|y\|_p \leq |x_{11}|_p$. Using $\mathbb{Q}_p = \bigsqcup_{n \in \mathbb{Z}} p^n \mathbb{Z}_p^\times$ and multiplication by a diagonal matrix with diagonal entries contained in $\mathbb{Z}_p^\times$, we can assume without loss of generality that $x_{11} = \|g\|_p^{-1}$ in the expression for x' obtained above. Now we proceed by induction on d. $\qquad\square$

Let us give a more intrinsic interpretation of Proposition C.4. In what follows, we consider

$$\mathfrak{pgl}_{d+1}(\mathbb{Q}_p) = \{v \in \mathrm{Mat}_{d+1}(\mathbb{Q}_p) : \mathrm{tr}(v) = 0\}.$$

Then $\mathrm{Ad} : \mathbf{G}(\mathbb{Q}_p) \to \mathrm{GL}(\mathfrak{pgl}_{d+1}(\mathbb{Q}_p))$ given by

$$\mathrm{Ad}_g(v) = g v g^{-1} \qquad (v \in \mathfrak{pgl}_{d+1}(\mathbb{Q}_p))$$

is a well-defined, faithful representation. We let $\|\cdot\|_p$ be the norm on $\mathrm{GL}(\mathfrak{pgl}_{d+1}(\mathbb{Q}_p))$ induced by the operator norm $\|\cdot\|_p$ on $\mathrm{GL}(\mathrm{Mat}_{d+1}(\mathbb{Q}_p))$ via restriction of the isomorphism to $\mathfrak{pgl}_{d+1}(\mathbb{Q}_p)$. To this end we note that any isomorphism of $\mathfrak{pgl}_{d+1}(\mathbb{Q}_p)$ extends trivially to the center of $\mathrm{Mat}_{d+1}(\mathbb{Q}_p)$.

Lemma C.5. *Let $k_1, k_2 \in \mathbf{G}(\mathbb{Z}_p)$ and let $n_1 \geq \cdots \geq n_d$ nonnegative integers. Set*

$$g = k_1[\mathrm{diag}(p^{-n_1}, \ldots, p^{-n_d}, 1)]k_2.$$

Then $\|\mathrm{Ad}_g\|_p = \|\mathrm{Ad}_{g^{-1}}\|_p = p^{n_1}$.

Proof. By Proposition C.4 and Corollary C.2, we can assume without loss of generality that $g = [\mathrm{diag}(p^{-n_1}, \ldots, p^{-n_d}, 1)]$. We let $n_{d+1} = 0$. An elementary calculation shows that the operator Ad_g acting on $\mathrm{Mat}_{d+1}(\mathbb{Q}_p)$ is diagonalizable with eigenvalues $p^{n_i - n_j}$ for : $1 \leq i, j \leq d + 1$. Therefore Ad_g has maximal eigenvalue p^{n_1} and $\|\mathrm{Ad}_g\|_p = p^{n_1}$. As $\sigma(\mathrm{Ad}_g)$ is symmetric under multiplicative inversion, we also have $\|\mathrm{Ad}_{g^{-1}}\|_p = p^{n_1}$. $\square$

Appendix D. Modular character on the Borel subgroup

The goal of this section is to prove (4.28). Recall first that the p-adic value satisfies that for all $x \in \mathbb{Q}_p$ and for all $f \in \mathrm{C}_c(\mathbb{Q}_p)$

$$|x|_p \int_{\mathbb{Q}_p} f(xt)\mathrm{d}m_{\mathbb{Q}_p}(t) = \int_{\mathbb{Q}_p} f(t)\mathrm{d}m_{\mathbb{Q}_p}(t),$$

where $m_{\mathbb{Q}_p}$ denotes any choice of a Haar measure on $\mathbb{Q}_p$. This implies that up to normalization the Haar measure on $\mathbb{Q}_p^\times$ is given by

$$\int_{\mathbb{Q}_p^\times} f(y)\mathrm{d}m_{\mathbb{Q}_p^\times}(y) = \int_{\mathbb{Q}_p^\times} \frac{f(y)}{|y|_p}\mathrm{d}m_{\mathbb{Q}_p}(y) \quad (f \in \mathrm{C}_c(\mathbb{Q}_p^\times)).$$

We let $B_p = \mathbf{B}(\mathbb{Q}_p)$ and note that $B_p = A_p U_p$, where A_p denotes the image of the diagonal subgroup of $\mathrm{GL}_{d+1}(\mathbb{Q}_p)$ and U_p is the (injective) image of the subgroup of upper triangular unipotent matrices. Note that U_p is homeomorphic to $\mathbb{Q}_p^{d'}$ with $d' = \frac{1}{2}d(d+1)$ and that the push forward m_{U_p} of the Haar measure on $\mathbb{Q}_p^{d'}$ to U_p defines a Haar measure on U_p. Therefore one finds that for a matrix $b = au \in B_p$ with $u \in U_p$ and $a = \mathrm{diag}(a_1, \ldots, a_d, 1)$ we have

$$\mathrm{d}m_{B_p}^{\mathrm{left}}(b) \propto \mathrm{d}m_{U_p}(u) \prod_{i=1}^d \frac{\mathrm{d}m_{\mathbb{Q}_p}(a_i)}{|a_i|_p^{d+2-i}}, \qquad \mathrm{d}m_{B_p}^{\mathrm{right}}(b) \propto \mathrm{d}m_{U_p}(u) \prod_{i=1}^d \frac{\mathrm{d}m_{\mathbb{Q}_p}(a_i)}{|a_i|^i}.$$

As the modular function satisfies $\mathrm{d}m_{B_p}^{\mathrm{right}}(b) \propto \delta_{\mathbf{B}}(b)\mathrm{d}m_{B_p}^{\mathrm{left}}(b)$, we find

$$\delta_{\mathbf{B}}(b) = \prod_{i=1}^d |a_i|_p^{d+2-2i}.$$

Equation (4.28) now follows by plugging in $a_{\mathbf{n}}$ for b.

Appendix E. Integrability of Matrix Coefficients

In this section, we show that $\eta_S \in \mathrm{L}^{v(d)+\varepsilon}(\mathbf{G}_S)$. Given $v \in S$, let $\eta_v : \mathbf{G}(\mathbb{Q}_v) \to (0, \infty)$ denote

$$\eta_v(g_v) = \|g_v\|_v^{-1} \quad (g_v \in \mathbf{G}(\mathbb{Q}_v)).$$

As $\eta_{\mathbf{G}}$ is the product of the various η_v, $v \in S$, and as the Haar measure on $\mathbf{G}_S$ is the product measure, it suffices to show that $\eta_v \in \mathrm{L}^{v(d)+\varepsilon}(\mathbf{G}_v)$ for all places v of $\mathbb{Q}$.

E.1. **Integrability in the Archimedean place.** We recall that the Haar measure on $\mathbf{G}_\infty$ is explicitly given by the formula

$$\int_{\mathbf{G}_\infty} f(g)\mathrm{d}g = \int_K \int_{A_\infty^+} \int_K f(k_1 a k_2) \prod_{\alpha \in \Sigma_\infty^+} (\sinh \alpha(\log a))^{m_\alpha} \mathrm{d}k_1 \mathrm{d}a \mathrm{d}k_2 \quad (f \in C_c(\mathbf{G}_\infty)),$$

where m_α denotes the multiplicity of the positive root $\alpha \in \Sigma_\infty^+$ and where the Haar measure on A is the push-forward of the Lebesgue measure on the Cartan subalgebra under the exponential map; cf. [Kn01, Prop. 5.28]. Note that in our situation $m_\alpha = 1$ for all $\alpha \in \Sigma_\infty^+$. By definition of the hyperbolic sine we have

$$\prod_{\alpha \in \Sigma_\infty^+} \sinh \alpha(\log a) \leq e^{\beta(\log a)} \quad (a \in A_\infty^+),$$

where $\beta = \sum_{\alpha \in \Sigma_\infty^+} \alpha$. We recall that for any $\alpha \in \Sigma_\infty^+$ there are $1 \leq i < j \leq d+1$ such that for all $a = \mathrm{diag}(a_1, \ldots, a_{d+1}) \in A_\infty^+$, we have $\alpha(\log a) = \log a_i - \log a_j$. Hence, we get

$$\beta(\log a) = \sum_{i=1}^{d+1} (d + 2 - 2i) \log a_i.$$

As argued in the proof of Lemma 4.7 and recalling that we parametrize A_∞^+ such that $a_{d+1} = 1$, we thus find that $\beta(\log a) \leq v(d) \log a_1$. Using the lower bound in (4.24), it follows that $\eta_\infty(k_1 a k_2) \leq a_1^{-1}$ and thus

$$\int_{\mathbf{G}_\infty} \eta_\infty(g)^q \mathrm{d}g = \int_K \int_{A_\infty^+} \int_K \eta_\infty(k_1 a k_2)^q \prod_{\alpha \in \Sigma_\infty^+} \sinh \alpha(\log a) \mathrm{d}k_1 \mathrm{d}a \mathrm{d}k_2 \ll_\vartheta \int_{A_\infty^+} a_1^{-q+v(d)} \mathrm{d}a.$$

By definition of A_∞^+ and recalling that the Haar measure on the connected component of the diagonal subgroup of $\mathbf{G}_\infty$ is the push-forward of the Lebesgue measure on the Lie algebra

$$\mathfrak{a} \cong \{(t_1, \ldots, t_d, t_{d+1}) : t_1 + \cdots + t_{d+1} = 0\}$$

under the exponential map, we have for any $s > 0$ that

$$\int_{A_\infty^+} a_1^{-s} \mathrm{d}a = \int_0^\infty \int_{t_d}^\infty \cdots \int_{t_2}^\infty e^{-st_1} \mathrm{d}t_1 \cdots \mathrm{d}t_d = \frac{1}{s^d}.$$

Hence, whenever $q > v(d)$, then

$$\int_{\mathbf{G}_\infty} \eta_\infty(g)^q \mathrm{d}g < \infty.$$

E.2. **Integrability in the finite places.** In what follows, we note that

$$\mathbf{G}(\mathbb{Q}_p) = \bigsqcup_{n \in \mathbb{N}_0} S_n, \qquad S_n = \{g \in \mathbf{G}(\mathbb{Q}_p) : \|g\|_p = p^n\}.$$

Moreover, Lemma 4.7 yields that for all $n \in \mathbb{N}_0$ we have

$$\int_{\mathbf{G}(\mathbb{Q}_p)} \eta_p(g)^q \mathrm{d}g = \sum_{n \in \mathbb{N}_0} \int_{S_n} \eta_p(g)^q \mathrm{d}g \ll_d \sum_{n \in \mathbb{N}_0} p^{-nq} \mathrm{Vol}(S_n) \ll_{\varepsilon,d} \sum_{n \in \mathbb{N}_0} p^{n(-q+v(d)+\varepsilon)}.$$

It follows that $\eta_p \in L^{v(d)+\varepsilon}(\mathbf{G}(\mathbb{Q}_p))$ whenever $q > v(d)$.

E.3. **Integrability of matrix coefficients.** This argument was mentioned in the proof of Proposition 5.3 and it follows readily from Section E.1. Recall from Proposition 4.5 that

$$\xi_\infty(g)^{2/(1-2\varepsilon)} \ll_{\varepsilon,d} \eta_\infty(g) \quad (g \in \mathbf{G}_\infty).$$

Hence for all $0 < \varepsilon < \frac{1}{2}$ and for $q > \frac{2v(d)}{1-2\varepsilon}$ we find

$$\int_{\mathbf{G}_\infty} \xi_\infty(g)^q \mathrm{d}g \ll_{\varepsilon,d} \int_{\mathbf{G}_\infty} \eta_\infty(g)^{(\frac{1}{2}-\varepsilon)q} \mathrm{d}g < \infty$$

as $(\frac{1}{2} - \varepsilon)q > v(d)$ and $\eta_\infty \in \mathrm{L}^{v(d)+\varepsilon}(\mathbf{G}_\infty)$.

Remark E.1. The same argument works for finite places.

REFERENCES

[AB19] D. Allen and B. Bárány. Diophantine Approximation on Fractals: Hausdorff measures of shrinking targets on self-conformal sets. Preprint (`https://arXiv.org/abs/1911.03410`), 2019.

[ACY20] D. Allen, S. Chow, and H. Yu. Dyadic Approximation in the Middle-Third Cantor Set. Preprint (`https://arxiv.org/abs/2005.09300`), 2020.

[BQ12] Y. Benoist and J.-F. Quint. Introduction to random walks on homogeneous spaces. *Jpn. J. Math.* 7 (2012), no. 2, 135–166.

[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.

[BQ13] Y. Benoist and J.-F. Quint. Stationary measures and invariant subsets of homogeneous spaces (II). *J. Amer. Math. Soc.* 26 (2013), no 3, 659–734.

[BQ13] Y. Benoist and J.-F. Quint. Stationary measures and invariant subsets of homogeneous spaces (III). *Ann. of Math.* (2) 178 (2013), no 3, 1017–1059.

[Be12] V. Beresnevich. Rational points near manifolds and metric Diophantine approximation. *Ann. of Math.* (2) 175 (2012), no. 1, 187–235.

[BDV07] V. Beresnevich, D. Dickinson, and S. Velani. Diophantine approximation on planar curves and the distribution of rational points (with an Appendix by R. C. Vaughan). *Ann. of Math.* (2) 166 (2007), no. 2, 367–426.

[BVVZ17] V. Beresnevich, R. C. Vaughan, S. Velani, and Zorin. Diophantine approximation on manifolds and the distribution of rational points: contributions to the convergence theory. *Int. Math. Res. Not.* (2017), no. 10, 2885–2908.

[Bo91] A. Borel. Linear algebraic groups. Second edition. Graduate Texts in Mathematics, 126. *Springer-Verlag, New York*, 1991.

[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.

[Bu90] M. Burger. Horocycle flow on geometrically finite surfaces. *Duke Math. J.* 61 (1990), no. 3, 779–803.

[Ca50] J. W. S. Cassels. Some metrical theorems in Diophantine approximation. I. *Proc. Cambridge Philos. Soc.* 46 (1950), 209–218.

[CE52] K. L. Chung and P. Erdős. On the application of the Borel-Cantelli lemma. *Trans. Amer. Math. Soc.* 72 (1952), 179–186.

[CY19] S. Chow and L. Yang. An effective Ratner equidistribution theorem for multiplicative Diophantine approximation on planar lines. *preprint* (2019).

[CHH88] M. Cowling, U. Haagerup, and R. Howe Almost L^2 matrix coefficients. *J. Reine Angew. Math.* 387 (1988), 97–110.

[Da64] H. Davenport. A note on Diophantine approximation. II. *Mathematika* 11 (1964), 50–58.

[Ed17] S. C. Edwards. On the rate of equidistribution of expanding translates of horospheres in $\Gamma\backslash G$. Preprint (`https://arxiv.org/abs/1701.04977`), 2017.

[EFS11] M. Einsiedler, L. Fishman, and U. Shapira. Diophantine approximations on fractals. *Geom. Funct. Anal.* 21 (2011), no. 1, 14–35.

[ELS20] M. Einsiedler, M. Luethi, and N. Shah. Primitive rational points on expanding horocycles in products of the modular surface with the torus. *Ergod. Th. & Dynam. Sys.,* First View (2020), 1–45.

[EMV09] M. Einsiedler, G. A. Margulis, and A. Venkatesh. Effective equidistribution for closed orbits of semisimple groups on homogeneous spaces. *Invent. Math.* 177 (2009), no. 1, 137–212.

[FF03] L. Flaminio and G. Forni. Invariant distributions and time averages for horocycle flows. *Duke Math. J.* 119 (2003), no. 3, 465–526.

[Ga62] P. Gallagher. Metric simultaneous Diophantine approximation. *J. London Math. Soc.* 37 (1962), 387–390.

[GMO08] A. Gorodnik, F. Maucourant, and H. Oh/ Manin's and Peyre's conjectures on rational points and adelic mixing. *Ann. Sci. Éc. Norm. Supér.* (4) 41 (2008), no. 3, 383–435.

[Gr95] S. Graf. On Bandt's tangential distribution for self-similar measures. *Monatsh. Math.* 120 (1995), no. 3–4, 223–246.

[HO89] A. J. Hahn and O. T. O'Meara. The classical groups and K-theory. With a foreword by J. Dieudonné. Grundlehren der Mathematischen Wissenschaften, 291. *Springer-Verlag, Berlin*, 1989.

[Ha06] J. Han. The general linear group over a ring. *Bull. Korean Math. Soc.* 43 (2006), no. 3,

[Hu20] JJ. Huang. The density of rational points near hypersurfaces. *Duke Math. J.* 169 (2020), no. 11, 2045–2077.

[Hu81] J. E. Hutchinson. Fractals and self-similarity. *Indiana Univ. Math. J.* 30 (1981), no. 5, 713–747.

[Kh26] A. Khintchine. Zur metrischen Theorie der Diophantischen Approximationen. *Math. Z.* 24 (1926), no. 1, 706–714.

[KS03] H. Kim and P. Sarnak. Refined estimates towards the Ramanujan and Selberg conjectures. Appendix to: H. Kim. Functoriality for the exterior square of GL$_4$. *J. Amer. Math. Soc.* 16 (2003), no. 1, 139–183.

[KLW04] D. Kleinbock, E. Lindenstrauss, and B. Weiss. On fractal measures and Diophantine approximation. *Selecta Math. (N.S.)* 10 (2004), no. 4, 479–523.

[KM98] D. Kleinbock and G. A. Margulis. Flows on homogeneous spaces and Diophantine approximation on manifolds. *Ann. of Math.* (2) 148 (1998), no. 1, 339–360.

[KM99] D. Kleinbock and G. A. Margulis. Logarithm laws for flows on homogeneous spaces. *Invent. Math.* 138 (1999), no. 3, 451–494.

[KM12] D. Kleinbock and G. A. Margulis. On effective equidistribution of expanding translates of certain orbits in the space of lattices. *Number theory, analysis and geometry*, 385–396, Springer, New York, 2012.

[KSW17] D. Kleinbock, R. Shi, and B. Weiss. Pointwise equidistribution with an error rate and with respect to unbounded functions. *Math. Ann.* 367 (2017), no. 1–2, 857–879.

[Kn01] A. W. Knapp. Representation theory of semisimple groups. An overview based on examples. Reprint of the 1986 original. Princeton Landmarks in Mathematics. *Princeton University Press, Princeton, NJ*, 2001.

[Li06] E. Lindenstrauss. Invariant measures and arithmetic quantum unique ergodicity. *Ann. of Math.* (2) 163 (2006), no. 1, 165–219.

[LSV07] J. Levesley, C. Salp, and S. L. Velani. On a problem of K. Mahler: Diophantine approximation and Cantor sets. *Math. Ann.* 338 (2007), no. 1, 97–118.

[Ma84] K. Mahler. Some suggestions for further research. *Bull. Austral. Math. Soc.* 29 (1984), no. 1, 101–108.

[Mo46] P. A. P. Moran. Additive functions of intervals and Hausdorff measure. *Proc. Cambridge Philos. Soc.* 42 (1946), 15–23.

[Oh98] H. Oh. Tempered subgroups and representations with minimal decay of matrix coefficients. *Bull. Soc. Math. France* 126 (1998), no. 3, 355–380.

[Oh02] H. Oh. Uniform pointwise bounds for matrix coefficients of unitary representations and applications to Kazhdan constants. *Duke Math. J.* 113 (2002), no. 1, 133–192.

[PR94] V. Platonov and A. Rapinchuk. Algebraic groups and number theory. Translated from the 1991 Russian original by Rachel Rowen. Pure and Applied Mathematics, 139. *Academic Press, Inc., Boston, MA*, 1994.

[Ra87] M. Ratner. The rate of mixing for geodesic and horocycle flows. *Ergodic Theory Dynam. Systems* 7 (1987), no. 2, 267–288.

[PV05] A. Pollington and S. L. Velani. Metric Diophantine approximation and "absolutely friendly" measures. *Selecta Math. (N.S.)* 11 (2005), no. 2, 297–307.

[Sa81] P. Sarnak. Asymptotic behavior of periodic orbits of the horocycle flow and Eisenstein series. *Comm. Pure Appl. Math.* 34 (1981), no. 6, 719–739.

[Sh00] Y. Shalom. Rigidity, unitary representations of semisimple groups, and fundamental groups of manifolds with rank one transformation group. *Ann. of Math. (2)* 152 (2000), no. 1, 113–182.

[Si45] C. L. Siegel. A mean value theorem in geometry of numbers. *Ann. of Math.* (2) 46 (1945), 340–347.

[Sl79] A. J. Silberger. Introduction to harmonic analysis on reductive p-adic groups. Based on lectures by Harish-Chandra at the Institute for Advanced Study, 1971–1973. Mathematical Notes, 23. *Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo*, 1979.

[St04] A. Strömbergsson. On the uniform equidistribution of long closed horocycles. *Duke Math. J.* 123 (2004), no. 3, 507–547.

[St13] A. Strömbergsson. On the deviation of ergodic averages for horocycle flows. *J. Mod. Dyn.* 7 (2013), no. 2, 291–328.

[SW19] D. Simmons and B. Weiss. Random walks on homogeneous spaces and Diophantine approximation on fractals. *Invent. Math.* 216 (2019), no. 2, 337–394.

[Su82] D. Sullivan. Disjoint spheres, approximation by imaginary quadratic numbers, and the logarithm law for geodesics. *Acta Math.* 149 (1982), no. 3-4, 215–237.

[VV06] R. C. Vaughan and S. Velani. Diophantine approximation on planar curves: the convergence theory. *Invent. Math.* 166 (2006), no. 1, 103–124.

[Ve10] A. Venkatesh. Sparse equidistribution problems, period bounds and subconvexity. *Ann. of Math. (2)* 172 (2010), no. 2, 989–1094.

[Wa] N. R. Wallach. Real Reductive Groups I. Pure and Applied Mathematics, 132. *Academic Press, Inc., Boston, MA*, 1988.

[We01] B. Weiss. Almost no points on a Cantor set are very well approximable. *R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci.* 457 (2001), no. 2001, 949–952.

[We02] B. Weiss. Dynamics on parameter spaces: submanifold and fractal subset questions. Rigidity in dynamics and geometry (Cambridge, 2000), 425–440.

[Yu20] H. Yu. Rational points near self-similar sets. Preprint. (2020).

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF UTAH, SALT LAKE CITY, UT
E-mail address: khalil@math.utah.edu

DEPARTMENT OF MATHEMATICS, TEL AVIV UNIVERSITY
E-mail address: manuelluthi@mail.tau.ac.il