Annals of Mathematics **181** (2015), 243–301
http://dx.doi.org/10.4007/annals.2015.181.1.4

Random walks in Euclidean space

By Péter Pál Varjú

Abstract

Fix a probability measure on the space of isometries of Euclidean space $\mathbf{R}^d$. Let $Y_0 = 0, Y_1, Y_2, \ldots \in \mathbf{R}^d$ be a sequence of random points such that Y_{l+1} is the image of Y_l under a random isometry of the previously fixed probability law, which is independent of Y_l. We prove a Local Limit Theorem for Y_l under necessary nondegeneracy conditions. Moreover, under more restrictive but still general conditions we give a quantitative estimate which describes the behavior of the law of Y_l on scales $e^{-cl^{1/4}} < r < l^{1/2}$.

1. Introduction

Let $X_1, X_2, \ldots$ be independent identically distributed random isometries of Euclidean space $\mathbf{R}^d$. Let $x_0 \in \mathbf{R}^d$ be any point, and consider the sequence of points

$$Y_0 = x_0, \ldots, Y_l = X_l(X_{l-1}(\ldots(x_0))), \ldots .$$

We call this sequence the *random walk* started from the point x_0, and Y_l is its lth step.

The purpose of this paper is to understand the distribution of Y_l. This problem can be traced back to Arnold and Krylov [2] who studied the mixing of the random walk on the sphere where the steps are rotations. They asked if their results extend to isometries of Euclidean or hyperbolic space.

Existing results in the literature can be divided into two classes. Some papers describe the behavior of the measure on scale $O(1)$, and others do it on scale $O(\sqrt{l})$. We begin by discussing the first category.

Każdan [16] and Guivarc′h [15] proved a Ratio Limit Theorem for $d = 2$. This result describes the local behavior of the distribution of Y_l. It states that the conditional distribution of Y_l on a fixed compact set is asymptotically uniform, i.e., Lebesgue. More precisely, for any two smooth compactly supported

I acknowledge the support of the European Research Council (Advanced Research Grant 267259) and the Simons Foundation.

functions f and g, we have

$$\lim_{l\to\infty}\frac{\mathbf{E}[f(Y_l)]}{\mathbf{E}[g(Y_l)]}=\frac{\int f(x)dx}{\int g(x)dx}$$

provided that the denominator on the right do not vanish. The law of X_1 of course needs to satisfy some natural nondegeneracy conditions for which we refer the reader to the original articles. The proofs rely on the fact that SO(2) is commutative.

In the papers [3], [17], [20] the Local Limit Theorem is generalized to higher dimension, however the arguments require some restrictive assumption on the law of X_1, e.g., absolute continuity, which implies that the group generated by the support of X_1 contains translations. In the absence of translations new ideas are required to obtain the Local Limit Theorem in full generality, which is the main goal of our paper.

Very recently, Conze and Guivarc′h [9] proved a Ratio Limit Theorem under a certain assumption on the associated random walk on SO(d). This assumption may hold in full generality, but so far it has been verified only under special circumstances. (We elaborate on this assumption in Section 4 after Theorem A.) Their approach also does not rely on translations but differs from the methods of this paper.

Tutubalin [23] proved a Central Limit Theorem for dimension $d = 2$ and $d = 3$, which was later generalized to higher dimension by Gorostiza [12] and Roynette [22]. The Central Limit Theorem describes the behavior of the distribution of Y_l on scale $O(\sqrt{l})$. More precisely, it claims that $Y_l/\sqrt{l}$ converges weakly to a Gaussian distribution if Y_1 has finite second moments. The Central Limit Theorem was revisited by many authors; see, e.g., [14], [21], [18] and [1]. In these works the Central Limit Theorem was even generalized to cases when the distribution Y_1 has infinite second moment and the limit distribution is not Gaussian.

To formulate our results we need to make some nondegeneracy condition on the law of X_i. We say that the law of X_i is *degenerate* if there is a proper closed subset A of $\mathbf{R}^d$ and an isometry $\gamma \in \mathrm{Isom}(\mathbf{R}^d)$ such that Y_l is almost surely contained in $\gamma^l(A)$. We say that the law of X_i is *nondegenerate* if it is not degenerate. Before we state the main result of the paper, we state two simpler results that can be deduced from our method. The following version of the Central Limit Theorem follows from our work.

THEOREM 1 (Central Limit Theorem). *Let X_i, Y_i and x_0 be as above. Suppose that Y_1 has finite second moments and the law of X_i is nondegenerate. Then there is a vector $v_0 \in \mathbf{R}^d$ such that the distribution of $(Y_l - lv_0)/\sqrt{l}$ weakly converges to a Gaussian distribution.*

The limit distribution of course depends on the distributions of X_i. We do not describe this dependence explicitly, but the mean and covariance matrix could be computed from the proof. Here we only mention that the covariance matrix is invariant under the rotation parts of the elements in the support of X_i.

This result is not new; it is covered by some of the references mentioned earlier. We revisit Tutubalin's argument in the greater generality considered in this paper. (Tutubalin assumes that $\mathrm{supp}\,(X_i)$ generates a dense subgroup of $\mathrm{Isom}(\mathbf{R}^d)$. Moreover, he discusses the cases $d = 2, 3$ only, although he does not seem to use this restriction in an essential way.) Our main purpose is to obtain quantitative bounds that will be necessary for proving the error estimates in Theorem 3 below.

THEOREM 2 (Local Limit Theorem). *Let X_i, Y_i and x_0 be as above. Suppose that Y_0 has finite moments of order d^2+3d+1 and X_i are nondegenerate. Let f be any continuous and compactly supported function. Then there is $v_0 \in \mathbf{R}^d$ and $c > 0$ depending only on the distribution of X_i such that*

$$\lim_{l\to\infty} l^{d/2}\mathbf{E}[f(Y_l - lv_0)] = c\int f(x)dx.$$

We remark that v_0 is the same as in the previous theorem and c can be computed from the covariance matrix of the limit distribution. Moreover, it turns out from the proof that v_0 is almost surely fixed by the rotation part of X_1, hence it is 0, if the rotation part of the support of X_1 is sufficiently rich.

When $v_0 = 0$, the Local Limit Theorem can be interpreted as follows. The probability that Y_l belongs to a fixed compact set with smooth boundary is asymptotic to $cl^{-d/2}$ times the Lebesgue measure of the set.

In the Local Limit Theorem, we need the finiteness of high order moments for technical reasons. However, if the group generated by the support of X_i is dense in $\mathrm{Isom}(\mathbf{R}^d)$, then our arguments imply the Local Limit Theorem under the assumption of finite second moment only. In fact, this is true under much weaker assumptions on the group generated by $\mathrm{supp}\,(X_i)$ (see Theorem 3).

Now we formulate the main result of the paper, which gives a quantitative description of the distribution of Y_l on multiple scales. However, we need a more restrictive assumption, which we call (SSR). We postpone the definition to the next section, where we explain the notation used through the paper. For now, we only mention that (SSR) holds, for example, if $\mathrm{supp}\,(X_i)$ generates a dense subgroup of $\mathrm{Isom}(\mathbf{R}^d)$ and $d \geq 3$. In addition, we can improve the error terms under stronger conditions, i.e., symmetricity or (E). These will be defined in the next section as well.

THEOREM 3. *Let X_i, Y_i and x_0 be as above. Suppose that the law of X_i is nondegenerate, satisfies* (SSR) *and Y_l has finite moments of order α for some $\alpha > 2$. Furthermore, let f be a smooth function of compact support. Then* *there is a point $y_0 \in \mathbf{R}^d$, a quadratic form $\Delta(x,x)$ and constants C_Δ and $c>0$ that depend only on the law of X_i such that*

$$\text{(1)}\quad \mathbf{E}[f(Y_l)] = C_\Delta l^{-d/2}\int f(x)e^{-\Delta(x-y_0,x-y_0)/l}dx \\ + O(l^{-\frac{d+\min\{1,\alpha-2\}}{2}} + |x_0|^2 l^{-\frac{d+2}{2}})\|f\|_1 + O(e^{-cl^{1/4}})\|f\|_{W^{2,(d+1)/2}}.$$

In addition, if μ is symmetric or satisfies (E), *we have*

$$\mathbf{E}[f(Y_l)] = C_\Delta l^{-d/2}\int f(x)e^{-\Delta(x-y_0,x-y_0)/l}dx \\ + O(l^{-\frac{d+\min\{2,\alpha-2\}}{2}} + |x_0|^2 l^{-\frac{d+2}{2}})\|f\|_1 + O(e^{-cl^{1/4}})\|f\|_{W^{2,(d+1)/2}}.$$

The implied constants depend only on the law of X_i.

A few remarks are in order about the conclusion of this theorem. The norm $\|\cdot\|_1$ is the L^1-norm, and $\|\cdot\|_{W^{2,(d+1)/2}}$ is an L^2 Sobolev norm defined by

$$\|f\|^2_{W^{2,(d+1)/2}} = \int |\widehat{f}(\xi)|^2(1+|\xi|)^{d+1}d\xi.$$

(The exponent $d+1$ could be replaced by any number greater than d and the theorem would still hold.)

The first term on the right-hand sides is the main term — it is the integral of f with respect to a Gaussian measure centered at y_0, and with covariance matrix Δ/l; C_Δ is simply a normalizing factor. It will follow from the proof that Δ is invariant under the rotation parts of elements of the support of μ.

The other two terms are error terms. The first is responsible for the large scale, and the second is for the small scale behavior of the random walk. To illustrate this, fix a smooth compactly supported function F, and consider the family $f_l(x) = r_l^{-d}F(x/r_l)$ associated to a sequence of scales r_l. (That is, the diameter of the support of f_l is proportional to r_l.) It is easily seen that as long as $r_l < \sqrt{l}$, the order of magnitude of the main term is $l^{-d/2}$, the first error term is $l^{-\frac{d+\min\{1,\alpha-2\}}{2}}$ while the second error term is $e^{-cl^{1/4}}r_l^{-d-1/2}$. This shows that the theorem gives a good approximation in the scale range $\sqrt{l} \ge r_l \ge e^{-c'l^{1/4}}$.

The factor $O(e^{-cl^{1/4}})$ is probably not optimal. In fact, our proofs lead to better estimates in some cases. This is discussed in detail in Section 4 after Theorem A, after the necessary background is explained.

Acknowledgments. I am greatly indebted to Elon Lindenstrauss for telling me about the problem and his continued interest in my project. I am also grateful to Jean Bourgain and Emmanuel Breuillard for helpful conversations about various aspects of the problem. I thank Noam Berger, Alex Lubotzky and Nikolay Nikolov for suggesting references [8], [4] and [13] respectively.

I thank the referee for his or her very careful reading of the paper and for suggestions that greatly improved the presentation.

I am grateful for the hospitality of the Mathematical Sciences Research Institute, Berkeley, CA.

2. Notation and outline

We identify the isometry group of the d-dimensional Euclidean space with the semi-direct product $\mathrm{Isom}(\mathbf{R}^d) = \mathbf{R}^d \rtimes \mathrm{O}(d)$. For $\gamma = (v, \theta) \in \mathbf{R}^d \rtimes \mathrm{O}(d)$ and a point $x \in \mathbf{R}^d$, we write

$$\gamma(x) = v + \theta x,$$

and we define the product of two isometries by

$$(v_1, \theta_1)(v_2, \theta_2) = (v_1 + \theta_1 v_2, \theta_1\theta_2).$$

If γ is an isometry, we write $v(\gamma)$ for the translation component and $\theta(\gamma)$ for the rotation component of γ in the above semi-direct decomposition.

Let μ be a probability measure on $\mathrm{Isom}(\mathbf{R}^d)$. Define the *convolution* $\mu * \mu$ in the usual way by

$$\int_{\mathrm{Isom}(\mathbf{R}^d)} f(\gamma) d\mu * \mu(\gamma) = \int_{\mathrm{Isom}(\mathbf{R}^d)} \int_{\mathrm{Isom}(\mathbf{R}^d)} f(\gamma_1\gamma_2) d\mu(\gamma_1) d\mu(\gamma_2)$$

for $f \in C(\mathrm{Isom}(\mathbf{R}^d))$, and write

$$\mu^{*(l)} = \underbrace{\mu * \cdots * \mu}_{l-\text{fold}}$$

for the l-fold convolution. With this notation, $\mu^{*(l)}$ is the distribution of the product of l independent random element of $\mathrm{Isom}(\mathbf{R}^d)$ of law μ. We define the measure $\widetilde{\mu}$ by the formula

$$\int_{\mathrm{Isom}(\mathbf{R}^d)} f(\gamma) d\widetilde{\mu}(\gamma) = \int_{\mathrm{Isom}(\mathbf{R}^d)} f(\gamma^{-1}) d\mu(\gamma)$$

for $f \in C(\mathrm{Isom}(\mathbf{R}^d))$ and say that μ is *symmetric* if $\widetilde{\mu} = \mu$. The measure μ also acts on measures on $\mathbf{R}^d$ in the following way: If ν is a measure on $\mathbf{R}^d$, we can define another measure $\mu.\nu$ on $\mathbf{R}^d$ by

$$\int_{\mathbf{R}^d} f(x) d\mu.\nu(x) = \int_{\mathbf{R}^d} \int_{\mathrm{Isom}(\mathbf{R}^d)} f(\gamma(x)) d\mu(\gamma) d\nu(x)$$

for $f \in C(\mathbf{R}^d)$.

We write δ_{x_0} for the Dirac delta measure concentrated at the point x_0. With this notation, the law of Y_l, the lth step of the random walk, is $\mu^{*(l)}.\delta_{x_0}$.

Write $\theta(\mu)$ for the projection of μ on $\mathrm{O}(d)$; i.e., for $f \in C(\mathrm{O}(d))$,

$$\int_{\mathrm{O}(d)} f(\sigma)d\theta(\mu)(\sigma) = \int_{\mathrm{Isom}(\mathbf{R}^d)} f(\theta(\gamma))d\mu(\gamma).$$

Denote by $G \subset \mathrm{Isom}(\mathbf{R}^d)$ the closure of the group generated by $\mathrm{supp}\,(\widetilde{\mu}*\mu)$. Fix any element $\gamma_0 \in \mathrm{supp}\,\mu$. Then it is clear that $\mathrm{supp}\,\mu \subset \gamma_0 G$.

We can replace μ by $\mu' = \mu^{*(k)}$ for some fixed integer $k > 1$ without loss of generality, since Y_{lk+j}, the $lk + j$th step of the original random walk, is the lth step of the modified random walk started from the random point Y_j. If we do so, then we replace G by the group G' as defined as the closure of the group generated by $\mathrm{supp}\,(\widetilde{\mu^{*(k)}} * \mu^{*(k)})$. It can be seen easily that G' is the closure of the group generated by

$$G \cup \gamma_0^{-1} G \gamma_0 \cup \cdots \cup \gamma_0^{-k+1} G \gamma_0^{k-1}.$$

In Lemma 5 we will see that if we choose k sufficiently large, then $\theta(G')$ is normalized by $\theta(\gamma_0^k)$. To keep this section compact, we postpone the statement and proof of Lemma 5, as well as Lemmata 4 and 6, which we will mention in the next pages.

Denote by $K \subset \mathrm{O}(d)$ the closure of the group generated $\mathrm{supp}\,\theta(\widetilde{\mu} * \mu)$. By the previous paragraph, we can (and will throughout the paper) assume without loss of generality that K is normalized by $\theta_0 := \theta(\gamma_0)$. Denote by K° the connected component of K. Denote by μ_K the Haar measure on the group K.

Now we list the various conditions that we will stipulate on μ in various parts of the paper. Some of these were already mentioned in Theorem 3.

(C) ("Centered") The barycenter of the image of the origin in $\mathbf{R}^d$ under μ is the origin; i.e.,

$$\int \gamma(0)d\mu(\gamma) = 0.$$

(E) ("Even") The action of K on $\mathbf{R}^d$ is "even", i.e., for every $v \in \mathbf{R}^d$, there is $\theta_v \in K$ such that $\theta_v v = -v$.

(SSR) ("Semi-simple rotations") K° is semi-simple, and there is no nonzero point in $\mathbf{R}^d$ that is fixed by K°.

We also recall the conditions we already defined for convenient reference. We say that μ is *nondegenerate* if there is no proper closed subset $A \subset \mathbf{R}^d$ and an isometry $\gamma \in \mathrm{Isom}(\mathbf{R}^d)$ such that $\mu^{*(l)}.\delta_{x_0}$ is almost surely contained in $\gamma^l(A)$.

It will be useful for us in many places in the paper to symmetrize μ by replacing it with $\widetilde{\mu} * \mu$. Unfortunately, the measure we obtain this way might be degenerate. Consider the following example in $\mathbf{R}^2$: Let γ_1 and γ_2 be two rotations about two different centers through the same angle that is not a rational multiple of π. We leave it to the reader to verify that γ_1 and γ_2

generate a dense subgroup in the orientation preserving isometries and hence the measure $\mu = (\delta_{\gamma_1} + \delta_{\gamma_2})/2$ is nondegenerate. However, for any k, $\widetilde{\mu^{*(k)}} * \mu^{*(k)}$ is supported on pure translations. Moreover, we can choose γ_1 and γ_2 to have matrices with rational entries, and then the translations in the support of $\widetilde{\mu^{*(k)}} * \mu^{*(k)}$ will all be rational. Hence they preserve the lattice $(1/q)\mathbf{Z}^2$, where q is the common denominator. This shows that $\widetilde{\mu^{*(k)}} * \mu^{*(k)}$ is degenerate.

For the above reason, we introduce a different notion that is easily seen to descend to $\widetilde{\mu} * \mu$. We say that μ is *almost nondegenerate* if for every point $x \in \mathbf{R}^d$, the set $\{\gamma(x) : \gamma \in \operatorname{supp} \mu\}$ does not lie in a proper affine subspace. As we will see in Lemma 6, if μ is nondegenerate, then $\mu^{*(k)}$ is almost nondegenerate for some integer $k \geq 1$, but it may happen that μ itself is *not* almost nondegenerate. The implication in the other direction is often true, as well. In particular, almost nondegeneracy is sufficient for most of the paper, except for Section 8.2.

We say that μ have *finite moments of order* $\alpha > 0$ if

$$\int |v(\gamma)|^\alpha d\mu(\gamma) < \infty.$$

A few remarks are in order regarding the role of these conditions. *Nondegeneracy* is clearly necessary for the Local Limit Theorem. However, we cannot impose it always for reasons discussed above. On the other hand, almost nondegeneracy is required throughout the paper. Condition (SSR) is needed to control the behavior of $\mu^{*(l)}$ on very small scales (up to $e^{-cl^{1/4}}$). Under this assumption we can utilize some powerful results about random walks on semisimple compact Lie groups. We assume (SSR) throughout Section 4 and some other parts of the paper. *Symmetry* or (E) allow us to improve the error terms in Theorem 3. They will be assumed in certain parts of Section 5 to show that the cubic terms in certain Taylor expansions cancel with each other. We assume throughout the paper that μ has *finite moments of order* 2. In Section 5 we assume finite moments of order α for $2 \leq \alpha \leq 4$, and the quality of our error terms depend on α. To be able to conclude the Local Limit Theorem without using (SSR) we assume the finiteness of higher order moments in Section 8.2. Finally, (C) is an assumption that does not restrict generality, as we will see in Lemma 4. Therefore we assume it throughout the paper to simplify our arguments.

Now we introduce some further notation and indicate the general strategy of the proof of Theorem 3. Recall that the distribution of the random walk started at the point x_0 after l-steps is the measure $\mu^{*(l)}.\delta_{x_0}$. As a consequence of the definitions, we see that

$$\mu^{*(l+1)}.\delta_{x_0} = \mu.(\mu^{*(l)}.\delta_{x_0}).$$

Hence our main goal is to understand the operation $\nu \mapsto \mu.\nu$.

This is achieved by studying the Fourier transform, which is given by the formula

$$\widehat{\nu}(\xi) = \int e(\langle \xi, x\rangle) d\nu(x),$$

where $e(x) := e^{-2\pi i x}$. For the Fourier transform of $\mu.\nu$, we get

$$\begin{aligned} (\mu.\nu)^{\wedge}(\xi) &= \int e(\langle \xi, \gamma(x)\rangle) d\mu(\gamma) d\nu(x) \\ &= \int e(\langle \xi, v(\gamma) + \theta(\gamma)(x)\rangle) d\mu(\gamma) d\nu(x) \\ &= \int e(\langle \xi, v(\gamma)\rangle) \widehat{\nu}(\theta(\gamma)^{-1}\xi) d\mu(\gamma). \end{aligned} \tag{2}$$

This formula shows that the action of μ on the Fourier transform of ν can be disintegrated with respect to spheres centered at the origin. For every $r \geq 0$, we define a unitary representation of the group $\mathrm{Isom}(\mathbf{R}^d)$ on the space $L^2(S^{d-1})$. Let

$$\rho_r(\gamma)\varphi(\xi) = e(r\langle \xi, v(\gamma)\rangle)\varphi(\theta(\gamma)^{-1}\xi) \tag{3}$$

for $\gamma \in \mathrm{Isom}(\mathbf{R}^d)$, $\varphi \in L^2(S^{d-1})$ and $\xi \in S^{d-1}$. We also define the operator

$$S_r(\varphi) = \int \rho_r(\gamma)(\varphi) d\mu(\gamma). \tag{4}$$

For a function $\varphi \in C(\mathbf{R}^d)$ and $r \geq 0$, we denote by $\mathrm{Res}_r\varphi$ its restriction to the sphere of radius r. That is, $\mathrm{Res}_r : C(\mathbf{R}^d) \to C(S^{d-1})$ is an operator defined by $[\mathrm{Res}_r\varphi](\xi) = \varphi(r\xi)$ for $|\xi| = 1$. With this notation, we can write (2) as

$$\mathrm{Res}_r(\widehat{\mu.\nu})(\xi) = S_r(\mathrm{Res}_r\widehat{\nu})(\xi).$$

Operators similar to S_r were introduced by Každan [16] and Guivarc′h [15]. Guivarc′h proved in the $d = 2$ case, when K is Abelian, that $\|S_r\| < 1 - cr^2$ for $r < 1$ and $\|S_r\| < 1 - c_r$ for $r \geq 1$, where $c > 0$ is a constant depending only on μ, while c_r also depend on r. These estimates are sufficient for proving a Ratio Limit Theorem and, as Breuillard [7] pointed out, combined with the Central Limit Theorem, it is sufficient even for a Local Limit Theorem. We are unable to prove such strong estimates, but we will prove in Section 4 a weaker version: Proposition 7, which is still sufficient for our application. In brief, we prove the estimate with constants c and c_r that (mildly) depend on the oscillations of φ. The proof is based on mixing properties of random walks on semi-simple compact Lie groups (see Theorem A below).

Using the estimates given in Section 4, we can show that the Fourier transform of the random walk after l steps "lives in" the ball of radius $l^{-1/2}\log l$. These estimates alone are sufficient for the Ratio Limit Theorem but not for the Central or Local Limit Theorems. The frequency range $r > l^{-1/2}\log l$ is responsible for the second error term in Theorem 3.

We need a more precise understanding of the Fourier transform of $\mu^{*(l)}.\delta_{x_0}$ in the range $r \le l^{-1/2} \log l$. This frequency range contributes the main term and the first error term in Theorem 3. In Section 5, we give Tutubalin's [23] argument for the Central Limit Theorem in the more general setting that we consider and obtain error estimates. In brief, this argument is based on decomposing $L^2(S^{d-1})$ as the orthogonal sum of several subspaces and using the Taylor expansion of the function $e(x)$ showing that these subspaces are almost invariant for S_r. We show that the contribution of only one of these subspaces is significant and that on this subspace rotations act trivially. Hence the problem is reduced to the easy case of sums of independent random variables.

There is some interdependency between the arguments of Sections 4 and 5. We explain this to demonstrate that our proof is not circular. The arguments of Section 4 depend on the Central Limit Theorem, which in turn depends on Section 5. However, Proposition 15 is sufficient for the Central Limit Theorem; its refinement, Proposition 24, is not needed. Among the results of Section 4, only Proposition 24 depends on the arguments of Section 4.

We will encounter L^2 spaces on various submanifolds of $\mathbf{R}^d$. We always consider them with respect to the "natural" measure; i.e., which is invariant under isometries. When the manifold is compact we normalize the measure to be probability.

Throughout the paper the letters c, C and various subscripted versions refer to constants and parameters. The same symbol occurring in different places need not have the same value unless the contrary is explicitly stated. For convenience, we use lower case for constants that are best thought of to be small and upper case for those that are best thought of to be large. In addition, we occasionally use Landau's O and o notation.

The organization of the rest of the paper that we have not explained yet is as follows. In Section 6 we combine the estimates of Section 5 and 4 to conclude Theorem 3. In Section 7 we derive Theorem 1 as a corollary of the results in Section 5. Finally, in Section 8 we prove Theorem 2. When K is Abelian and its action on $\mathbf{R}^d$ has a trivial component, some additional difficulties arise that prevent us from using the method of Guivarc′h [15]. To address these issues, we use Taylor expansions in Section 8.2 motivated by Tutubalin's paper. For this argument, we need to assume the finiteness of high order moments.

3. Justifying the simplifying assumptions

We prove three technical lemmata in this section that we referred to in the previous section. Their common feature is that they allow us to make certain simplifying assumptions on the law μ generating the random walk without loss of generality.

First we prove that there is a suitable choice of origin for the coordinate system so that assumption (C) is satisfied. Then we prove that our assumption that K is normalized by θ_0 is justified if we replace μ by a convolution power of itself. Finally we prove that the same replacement allows us to assume that μ is almost nondegenerate.

LEMMA 4. *Assume that there is no point in $\mathbf{R}^d$ except for the origin that is fixed by all elements of K. Then there is a unique point $x \in \mathbf{R}^d$ such that*

$$\int \gamma(x) d\mu(\gamma) = x.$$

The conclusion implies that if we change our coordinate system, and set x to be the origin, then (C) is satisfied.

Proof. Consider the map $\mathbf{R}^d \to \mathbf{R}^d$:

$$T(x) = \int \gamma(x) d\mu(\gamma) - \int \gamma(0) d\mu(\gamma) = \int \theta(\gamma) x d\mu(\gamma).$$

It is clear that T is a linear transformation.

We show that $x - T(x)$ has trivial kernel. Suppose that $x = T(x)$ for some $x \in \mathbf{R}^d$. Since $|\theta(\gamma)x| = |x|$ for all γ, and $T(x)$ is the average of these points, we must have $\theta(\gamma)x = x$ for μ-almost all γ. By our assumption, $x = 0$, and hence the kernel of $x - T(x)$ is indeed trivial.

Therefore there is a unique point x such that $x - T(x) = \int \gamma(0) d\mu(\gamma)$, and this is exactly what we wanted to prove. $\square$

LEMMA 5. *Let $K < \mathrm{O}(d)$ be a compact group, and let $\theta_0 \in \mathrm{O}(d)$. There is a positive integer l such that θ_0 normalizes the group generated by*

$$K \cup \theta_0^{-1} K \theta_0 \cup \cdots \cup \theta_0^{-(l-1)} K \theta_0^{l-1}.$$

Proof. It is a well-known fact that if K is a compact Lie group, then there is a chain of normal subgroups $K^\circ \triangleleft H \triangleleft K$ such that H/K° is commutative and $[K : H] < C_d$ for a constant C_d depending on d. For a proof in the context of algebraic groups that carries over to compact groups without any changes, see [4, Th. J].

Write K_l for the closure of the group generated by

$$K \cup \theta_0^{-1} K \theta_0 \cup \cdots \cup \theta_0^{-(l-1)} K \theta_0^{l-1},$$

and write $K_l^\circ \triangleleft K_l$ for its connected component.

Let $l_0 \geq 1$ be an integer such that $K_l^\circ = K_{l_0}^\circ$ for $l \geq l_0$. (The sequence K_l° stabilizes since $\dim K_l^\circ$ may grow at most finitely many times.) Denote by L the closure of the union of the groups K_l. Then $K_{l_0}^\circ \triangleleft L$ since $K_{l_0}^\circ$ is normal in all K_l for $l \geq l_0$.

We show that $L^\circ/K_{l_0}^\circ$ is commutative. For any $l \geq l_0$ and $g, h \in K_l$, we have

$$[g^{C_d!}, h^{C_d!}] \in K_{l_0}^\circ;$$

hence this property descends to L. Since all elements in a connected compact Lie group are $C_d!$ powers, we have $[L^\circ, L^\circ] < K_{l_0}^\circ$; thus $L^\circ/K_{l_0}^\circ$ is indeed commutative. Note that L and L° are both normalized by θ_0, which is of crucial importance for what follows.

Write $H_l = K_l \cap L^\circ$. Then clearly $K_l^\circ \lhd H_l \lhd K_l$, $[K_l : H_l] \leq [L : L^\circ]$ and H_l/K_l° is commutative for $l \geq l_0$. Let $l_1 \geq l_0$ be such that $[K_{l_1} : H_{l_1}] = [L : L^\circ]$, and let $g_1, \dots, g_m$ be a system of representatives for H_{l_1} cosets in K_{l_1}.

We show that $\exp(H_l/K_{l_0}^\circ)$ is constant for $l \geq l_1 + 1$. The *exponent* $\exp(G)$ of a group G is the smallest integer n such that $g^n = 1$ for all $g \in G$. Since the elements of H_l approximate those of L°, this would imply that $\exp(L^\circ/K_{l_0}^\circ) < \infty$. Then $L^\circ = K_{l_0}^\circ$, as both of them are connected Lie groups. Thus $H_l = K_{l_0}^\circ$ for all $l \geq l_0$, and $K_l = \{g_1, \dots, g_m\} K_{l_0}^\circ$ for $l \geq l_1$. That is, the sequence K_l stabilizes, which was to be proved.

Let $l \geq l_1 + 1$. Then all elements of K_{l+1} are of the form

$$g = \prod_\alpha \gamma_\alpha^{-1}(g_{i_\alpha} h_{i_\alpha})\gamma_\alpha,$$

where $h_{i_\alpha} \in H_l$ and $\gamma_\alpha \in \{1, \theta_0\}$. For each α, we can write

$$\gamma_\alpha^{-1}(g_{i_\alpha} h_{i_\alpha})\gamma_\alpha = g_{j_\alpha} h_{j_\alpha} \gamma_\alpha^{-1} h_{i_\alpha} \gamma_\alpha,$$

where g_{j_α} is the appropriate coset representative and

$$h_{j_\alpha} = g_{j_\alpha}^{-1} \gamma_\alpha^{-1} g_{i_\alpha} \gamma_\alpha \in H_{l_1+1} < H_l.$$

We bring all g_{j_α} to the left-hand side of the product and get that each element of H_{l+1} is of the form

$$h = \prod_\beta \gamma_\beta^{-1} h_\beta \gamma_\beta,$$

where $h_\beta \in H_l$ and $\gamma_\beta \in \{1, \theta_0\} \cdot K_{l_1}$. Thus all $\gamma_\beta^{-1} h_\beta \gamma_\beta$ are in L°; in particular, they commute modulo $K_{l_0}^\circ$. In addition, the degree of each $h_\beta \cdot K_{l_0}^\circ \in H_l/K_{l_0}^\circ$ divides $\exp(H_l/K_{l_0}^\circ)$; hence so is the degree of $h \cdot K_{l_0}^\circ \in H_{l+1}/K_{l_0}^\circ$. This implies that $\exp(H_{l+1}/K_{l_0}^\circ) = \exp(H_l/K_{l_0}^\circ)$, which was to be proved. □

LEMMA 6. *Suppose that μ is nondegenerate. Then there is a positive integer l such that $\mu^{*(l)}$ is almost nondegenerate.*

Proof. By the nondegeneracy assumption, it follows that for each point $x \in \mathbf{R}^d$, there is $l(x)$ such that the set $\{\gamma(x) : \gamma \in \operatorname{supp} \mu^{*(l(x))}\}$ is not contained in a proper affine subspace. Indeed, assume to the contrary that this fails, and l_0 and W are such that $\{\gamma(x) : \gamma \in \operatorname{supp} \mu^{*(l_0)}\}$ spans W and W is of largest possible dimension. Then $\gamma(x)$ is $d\mu^{*(l)}(\gamma)$-almost surely contained in

$\gamma_0^{l-l_0}(W)$, where $\gamma_0 \in \operatorname{supp}\mu$ is arbitrary. This contradicts the nondegeneracy of μ.

It is left to show that $l(x)$ is bounded on $\mathbf{R}^d$. It is easy to see that $\{x : l(x) \le L\}$ is a Zariski open set for every $L \in \mathbf{Z}$. As $L \to \infty$ this is an ascending chain that eventually covers $\mathbf{R}^d$. Therefore the claim follows from the Noetherian property of Zariski open sets. □

4. **Estimates for high frequencies**

The goal of this section is to estimate the norm of the operator S_r defined in Section 2. We are not able to show that $\|S_r\| < 1$, but we can give an estimate for $\|S_r\varphi\|_2$ in terms of the following Lipschitz type norm of φ:

$$\|\varphi\|_{\mathrm{Lip}(K)} := \|\varphi\|_\infty + \sup_{\xi \in S^{d-1}, \theta \in K\backslash\{1\}} \frac{|\varphi(\xi) - \varphi(\theta(\xi))|}{\mathrm{dist}(1,\theta)},$$

where $\mathrm{dist}(\cdot,\cdot)$ is a distance function on K that is induced by the invariant Riemannian metric on K. Note that there is a constant C depending on the geometry of the embedding of K inside $\mathrm{O}(d)$ such that $|\xi - \theta(\xi)| \le C\mathrm{dist}(1,\theta)$ for every ξ and θ. Thus $\|\varphi\|_{\mathrm{Lip}(K)} \le C\|\varphi\|_{\mathrm{Lip}}$ for any function, where $\|\cdot\|_{\mathrm{Lip}}$ is the ordinary Lipschitz norm on the sphere.

Proposition 7. *Suppose that μ is almost nondegenerate, has finite moments of order* 2 *and satisfies* (SSR). *Then there is a constant $c > 0$ depending only on μ such that the following hold. Let $\varphi \in L^2(S^{d-1})$ with $\|\varphi\|_2 = 1$. Then*

$$\|S_r\varphi\|_2 \le 1 - c\min\left\{r^2, \frac{1}{\log^3((r+1)\|\varphi\|_{\mathrm{Lip}(K)} + 2)}\right\}. \tag{5}$$

This estimate allows us to control the Fourier transform of the random walk in the frequency range $e^{cl^{1/4}} > r > l^{-1/2}\log l$.

Mixing properties of random walks on semi-simple compact Lie groups is a crucial ingredient of our proof. We state the result that we use in the next theorem. The proof is given in the paper [24, Cor. 7]. A quantitatively weaker version, but essentially sufficient for our purpose, could be deduced from the Solovay-Kitaev algorithm, at least in the case $K = \mathrm{SU}(d)$. The Solovay-Kitaev algorithm was first described in an e-mail discussion list by Solovay in 1995. Kitaev independently discovered it and published it in 1997 [19]. For a recent exposition, see [10]. See also the paper of Dolgopyat [11, Ths. A.2 and A.3], which provides similar estimates.

Theorem A. *Let K be a compact Lie group with semi-simple connected component. Let μ be a symmetric probability measure on K such that* $\operatorname{supp}\mu$ *generates a dense subgroup in K. Then there is a constant $c > 0$ depending* *only on μ such that the following hold. Let $\varphi \in L^2(K)$ be a function such that $\|\varphi\|_2 = 1$ and $\int \varphi dm_K = 0$. Then*

$$\left\| \int \varphi(\theta^{-1}\sigma) d\mu(\theta) \right\|_2 < 1 - \frac{c}{\log^2(\|\varphi\|_{\mathrm{Lip}} + 2)}. \tag{6}$$

Recall that m_K denotes the Haar measure on K.

As we mentioned in the introduction, Conze and Guivarc′h [9] proved the Ratio Limit Theorem under a certain assumption. This assumption is that $K = \mathrm{SO}(d)$, and $\theta(\mu)$ satisfies (6) with $1 - c$ on the right independently of φ. We add that Bourgain and Gamburd [6], [5] proved (in the $K = \mathrm{SU}(d)$ case) that if μ satisfies some additional conditions (e.g., the support of μ consists of matrices with algebraic entries), then the stronger version of (6) needed by Conze and Guivarc′h holds.

If one improves the estimate in Theorem A, then our argument presented below provides better estimates in Proposition 7 and Theorem 3. In particular, if one can replace the right-hand side of (6) with $1 - c\log^{-A}(\|\varphi\|_{\mathrm{Lip}} + 2)$, then one can write $1 - c\min\{r^2, \log^{-A-1}((1+r)\|\varphi\|_{\mathrm{Lip}(K)} + 2)\}$ on the right-hand side of (5) and $O(e^{-cl^{1/(A+2)}})\|f\|_W$ instead of the second error term in (1). In fact, Theorem A is proved with better bounds for most Lie groups — except for those that project onto SO(3). For details, we refer to [24]. Moreover, for certain generators (e.g., when they are given with algebraic entries), the estimates are available even with $A = 0$, as we discussed above.

The rest of this section is devoted to the proof of Proposition 7. A simple observation shows that it is enough to prove it for symmetric measures. Indeed, we have

$$\|S_r\varphi\|_2^2 = \langle S_r\varphi, S_r\varphi\rangle = \langle \varphi, S_r^* S_r \varphi\rangle \le \|S_r^* S_r \varphi\|_2, \tag{7}$$

and $S_r^* S_r$ is the operator analogous to S_r corresponding to the symmetric measure $\widetilde{\mu} * \mu$.

We check that the assumptions of Proposition 7 hold for $\widetilde{\mu} * \mu$ if they hold for μ. Since $1 \in \mathrm{supp}\,(\theta(\widetilde{\mu} * \mu))$, $\mathrm{supp}\,(\theta(\widetilde{\mu} * \mu)) \subset \mathrm{supp}\,(\theta((\widetilde{\mu} * \mu)^{*(2)}))$; hence $\mathrm{supp}\,(\theta((\widetilde{\mu} * \mu)^{*(2)}))$ generates a dense subgroup of K, so (SSR) holds for $\widetilde{\mu} * \mu$. We have

$$\begin{aligned} \int |v(\gamma_1 \cdot \gamma_2)|^2 d\widetilde{\mu}(\gamma_1)\mu(\gamma_2) &\le \int |v(\gamma_1) + v(\gamma_2)|^2 d\widetilde{\mu}(\gamma_1)\mu(\gamma_2) \\ &\le 2\int |v(\gamma_1)|^2 + |v(\gamma_2)|^2 d\widetilde{\mu}(\gamma_1)\mu(\gamma_2) \le \infty, \end{aligned}$$

so $\widetilde{\mu} * \mu$ also has finite second moments. Let $\gamma_1 \in \mathrm{supp}\,(\widetilde{\mu})$ be arbitrary. Then for any point $x \in \mathbf{R}^d$, the set

$$\{\gamma(x) : \gamma \in \mathrm{supp}\,(\widetilde{\mu} * \mu)\} \supset \{\gamma_1(\gamma(x)) : \gamma \in \mathrm{supp}\,(\mu)\}$$

cannot be contained in a proper affine subspace. Thus $\widetilde{\mu} * \mu$ is also almost nondegenerate. For the rest of the section, we write μ for $\widetilde{\mu} * \mu$ and S_r for $S_r^* S_r$. In addition, this argument shows that we can assume that S_r is nonnegative.

By Lemma 4, we can change the origin in such a way that (C) holds for μ. Denote by u the new origin in the old coordinate system. Then the isometry (v, θ) becomes $(v - u + \theta u, \theta)$ in the new coordinates. Hence the operator S_r will be replaced by the operator

$$\begin{aligned} S'_r \varphi(\xi) &= \int e(r\langle v(\gamma) - u + \theta(\gamma) u, \xi\rangle)\varphi(\theta(\gamma)^{-1}\xi) d\mu(\gamma) \\ &= e(r\langle -u, \xi\rangle) S_r(e(r\langle u, \xi\rangle)\varphi(\xi)). \end{aligned}$$

By setting $\varphi'(\xi) = e(r\langle u, \xi\rangle)\varphi(\xi)$, we see that $\|S_r\varphi\|_2 = \|S'_r\varphi'\|_2$. Note that

$$\|e(r\langle u, \xi\rangle)\varphi(\xi)\|_{\mathrm{Lip}(K)} \le C((r+1)\|\varphi\|_\infty + \|\varphi\|_{\mathrm{Lip}(K)}) \le C(r+1)\|\varphi\|_{\mathrm{Lip}(K)},$$

where C is a constant depending only on u. This shows that if Proposition 7 holds for S'_r and φ', then it also holds for S_r and φ.

From now on, until the end of the section, we assume that μ is symmetric, almost nondegenerate, has finite second moments and satisfies (C) and (SSR). Moreover, we assume that S_r is selfadjoint and nonnegative. By the above discussion, these assumptions are justified.

Until the end of the section, we fix $r > 0$ and a function $\varphi \in \mathrm{Lip}(S^{d-1})$ and prove Proposition 7 for these. The strategy of the proof is the following. We fix two integers

$$l_1 = [C_1(r^{-2} + \log^3(\|\varphi\|_{\mathrm{Lip}(K)} + 2))], \quad l_2 = [C_2(r^{-2} + \log^3(\|\varphi\|_{\mathrm{Lip}(K)} + 2))],$$

where C_1, C_2 are suitably chosen large constants depending on μ but not on φ or r. We will show that the set of isometries that almost fix φ in the ρ_r representation is of $\mu^{*(l)}$ measure at most $9/10$ for $l = l_1$ or $l = l_2$. This implies Proposition 7 by a standard argument. More precisely, we prove the following lemma.

Lemma 8. *Define the set*

$$B(\varepsilon) := \{\gamma \in \mathrm{Isom}(\mathbf{R}^d) : \|\rho_r(\gamma)\varphi - \varphi\|_2 < \varepsilon\}.$$

If ε is sufficiently small depending on μ, and C_1 is sufficiently large depending on ε and μ, and C_2 is sufficiently large depending on C_1, ε and μ, then

$$\mu^{*(l_i)}(B(\varepsilon)) < 9/10 \tag{8}$$

holds for $i = 1$ or $i = 2$.

We show how to deduce Proposition 7 from Lemma 8.

Proof of Proposition 7. Let l_i be the one for which (8) holds, and assume that l_i is even for simplicity. Then we can write

$$\begin{aligned}
\|S_r^{l_i/2}(\varphi)\|_2^2 &= \langle S_r^{l_i/2}(\varphi), S_r^{l_i/2}(\varphi)\rangle = \langle S_r^{l_i}(\varphi),\varphi\rangle \\
&= \int \langle \rho_r(\gamma)\varphi,\varphi\rangle d\mu^{*(l_i)}(\gamma) \le \frac{1}{2}\int \langle \rho_r(\gamma)\varphi + \rho_r(\gamma^{-1})\varphi,\varphi\rangle d\mu^{*(l_i)}(\gamma) \\
&\le \frac{1}{2}\int_{\mathrm{Isom}(\mathbf{R}^d)\backslash B(\varepsilon)} \langle \rho_r(\gamma)\varphi + \rho_r(\gamma^{-1})\varphi,\varphi\rangle d\mu^{*(l_i)}(\gamma) + \mu^{*(l_i)}(B(\varepsilon)) \\
&\le (1-\varepsilon^2/2)/10 + 9/10.
\end{aligned}$$

To deduce the last inequality, we used the identity

$$\langle \rho_r(\gamma)\varphi + \rho_r(\gamma^{-1})\varphi,\varphi\rangle = 2 - \|\rho_r(\gamma)\varphi - \varphi\|_2^2.$$

We concluded that $\|S_r^{l_i/2}(\varphi)\|_2 \le e^{-c}$ for some $c > 0$ depending on μ. By selfadjointness of S_r we can deduce that $\|S_r(\varphi)\|_2 \le e^{-2c/l_i}$. We compare this with the definition of l_i, which finishes the proof. □

The rest of the section is devoted to the proof of Lemma 8. We begin with a simple lemma that shows that the length of the translation part of γ is proportional to $\sqrt{l}$ with $\mu^{*(l)}$ probability at least 9/10.

LEMMA 9.

$$\int |v(\gamma)|^2 d\mu^{*(l)}(\gamma) = l\cdot \int |v(\gamma)|^2 d\mu(\gamma).$$

Proof. For $l = 1$, the statement is obvious; for l larger the proof is by induction:

$$\begin{aligned}
\int |v(\gamma)|^2 d\mu^{*(l+1)}(\gamma) &= \iint |v(\gamma_1\gamma_2)|^2 d\mu(\gamma_2)d\mu^{*(l)}(\gamma_1) \\
&= \iint |v(\gamma_1) + \theta(\gamma_1)v(\gamma_2)|^2 d\mu(\gamma_2)d\mu^{*(l)}(\gamma_1) \\
&= \iint |\theta(\gamma_1^{-1})v(\gamma_1) + v(\gamma_2)|^2 d\mu(\gamma_2)d\mu^{*(l)}(\gamma_1) \\
&= \iint \left(|v(\gamma_1)|^2 + |v(\gamma_2)|^2 + 2\langle\theta(\gamma_1^{-1})v(\gamma_1), v(\gamma_2)\rangle\right) d\mu(\gamma_2)d\mu^{*(l)}(\gamma_1).
\end{aligned}$$

Integrating out γ_2, the third term in the last line vanishes by (C). This proves the lemma by induction. □

For the rest of the section, we assume to the contrary that (8) fails, and we proceed by various lemmata that show under the indirect hypothesis that $B(\varepsilon')$ contains larger and larger families of isometries if ε' becomes larger and larger (but still small). We will reach contradiction when we show that we can find translations of length comparable to r^{-1} in many directions. The following simple property of the set $B(\varepsilon)$ will be used repeatedly.

LEMMA 10. *For any numbers $\varepsilon_1, \varepsilon_2$, we have $B(\varepsilon_1)B(\varepsilon_2) \subset B(\varepsilon_1 + \varepsilon_2)$.*

Proof. Let $\gamma_1 \in B(\varepsilon_1)$ and $\gamma_2 \in B(\varepsilon_2)$ be arbitrary. Then by the triangle inequality, we have

$$\|\rho_r(\gamma_1\gamma_2)\varphi - \varphi\|_2 \leq \|\rho_r(\gamma_1\gamma_2)\varphi - \rho_r(\gamma_1)\varphi\|_2 + \|\rho_r(\gamma_1)\varphi - \varphi\|_2. \tag{9}$$

Since ρ_r is a unitary representation, the first term on the right is equal to $\|\rho_r(\gamma_2)\varphi - \varphi\|_2 \leq \varepsilon_2$. Thus (9) $\leq \varepsilon_1 + \varepsilon_2$, which proves the lemma. □

In the first lemma, we conclude that $B(4\varepsilon)$ contains isometries with an arbitrary prescribed rotation part and translation part proportional to $\sqrt{l}$.

LEMMA 11. *Suppose that (8) fails for $i = 1$ and some $\varepsilon > 0$. Suppose further that C_1 is sufficiently large depending on μ and ε. Then there exist a constant C that depends only on μ such that the following holds. There is a set $X \subset B(4\varepsilon)$ such that*

$$\theta(X) = K \qquad \text{and} \qquad |v(\gamma)| < C\sqrt{l_1} \quad \text{for} \quad \gamma \in X.$$

Proof. We deduce the lemma from Theorem A. Let $\mathcal{B}$ be the $\varepsilon\|\varphi\|_{\mathrm{Lip}(K)}^{-1}$ neighborhood of the identity in K°. It follows from the definitions that

$$\|\rho_r(\theta)\varphi - \varphi\|_\infty \leq \varepsilon$$

for every $\theta \in \mathcal{B}$. Thus we have $\mathcal{B} \subset B(\varepsilon)$.

Take an approximate identity ψ on K, which has the following properties:

$$\mathrm{supp}\,(\psi) \subset \mathcal{B}, \quad \int \psi dm_K = 1 \quad \text{and} \quad \|\psi\|_{\mathrm{Lip}} \leq C\|\varphi\|_{\mathrm{Lip}(K)}^{1+\dim K}.$$

Note that these imply that $\|\psi\|_2 \leq C\|\varphi\|_{\mathrm{Lip}(K)}^{(\dim K)/2}$. These constants again depend only on K and ε. Now we apply Theorem A successively l_1 times starting with the function $(1-\psi)/\|1-\psi\|_2$ and get

$$\left\|1 - \int \psi(\theta^{-1}\sigma)d\theta(\mu)^{*(l_1)}(\theta)\right\|_2 \leq \frac{1}{10}$$

provided C_1 is sufficiently large depending only on μ and ε. Recall that $l_1 > C_1 \log^3(\|\varphi\|_{\mathrm{Lip}(K)} + 2)$. We note that taking the average of translates of ψ may only decrease the Lipschitz norm.

Now let $Y \subset B(\varepsilon)$ be such that $\mu^{*(l_1)}(Y) > 8/10$, and

$$|v(\gamma)| < C\sqrt{l_1}/2 \quad \text{for} \quad \gamma \in Y.$$

For a sufficiently large C depending on (the second moment of) μ, this is possible due to the assumption that (8) fails and Lemma 9. Denote by ν the measure we obtain from $\theta(\mu^{*(l_1)})$ if we restrict it to the set Y, and normalize it to get a probability measure. Then we have

$$\left\| \int \psi(\theta^{-1}\sigma) d\nu(\theta) \right\|_2 \leq (1 + \frac{1}{10}) \cdot \frac{10}{8} < \sqrt{2}.$$

Thus

$$m_K(\theta(Y)\mathcal{B}) \geq m_K \left(\operatorname{supp} \left(\int \psi(\theta^{-1}\sigma) d\nu(\theta) \right) \right) > \frac{1}{2},$$

which proves the lemma with the choice $X = Y\mathcal{B}Y\mathcal{B}$. □

LEMMA 12. *There are constants $c, C > 0$ that depend only on μ such that we have*

$$\mu^{*(l)}(\gamma : |\langle v(\gamma), u_0 \rangle| > c\sqrt{l} \text{ and } |v(\gamma)| < C\sqrt{l}) > 1/2$$

for any sufficiently large (depending only on μ) integer l and any $u_0 \in S^{d-1}$.

Proof. The lemma is an easy consequence of the Central Limit Theorem, i.e., Theorem 1. □

We could, off course, replace 1/2 in the lemma with any number less than 1. Now we can show that under our standing assumption that (8) fails, $B(5\varepsilon)$ contains a nontrivial translation.

LEMMA 13. *Suppose that (8) fails with some $1/2 > \varepsilon > 0$ for both $i = 1$ and $i = 2$. Suppose further that C_1 is sufficiently large so that Lemma 11 holds and C_2 is sufficiently large depending on μ and C_1. Then there are constants $c, C > 0$ depending only on μ such that for any $u_0 \in S^{d-1}$, there is an element $\gamma_1 \in B(5\varepsilon)$ with the following properties:*

$$|v(\gamma_1)| < C\sqrt{l_2}, \quad \langle v(\gamma_1), u_0 \rangle > c\sqrt{l_2} \quad \text{and} \quad \theta(\gamma_1) = 1.$$

Proof. Denote by c_0 and C_0 the constants from Lemma 12. Using that lemma with $l = l_2$ and the failure of (8) for $i = 2$, we find an element $\gamma_2 \in B(\varepsilon)$ such that

$$|\langle v(\gamma_2), u_0 \rangle| > c_0\sqrt{l_2} \quad \text{and} \quad |v(\gamma_2)| < C_0\sqrt{l_2}.$$

On the other hand, applying to Lemma 11, we can find $\gamma_3 \in B(4\varepsilon)$ that satisfies

$$\theta(\gamma_3) = \theta(\gamma_2)^{-1} \quad \text{and} \quad |v(\gamma_3)| < C_0'\sqrt{l_1},$$

where C_0' is the constant C from that lemma.

We demand that $l_2/l_1 = C_2/C_1$ is so large that

$$c_0\sqrt{l_2} > 2C_0'\sqrt{l_1}.$$

Then it is an easy calculation to verify that $\gamma_1 = \gamma_2\gamma_3$ has the claimed properties. □

Recall the definition of l_2 — in particular, that it implies $\sqrt{l_2} > \sqrt{C_2} r^{-1}$. The next lemma shows that we can find a translation γ_1' with properties similar to those of γ_1 in the previous lemma, but which is shorter.

LEMMA 14. *Under the same hypothesis as in Lemma 13, there is a constant $c > 0$ that depends only on μ, and there is an element $\gamma_1' \in B(26\varepsilon)$ with the following properties*:

$$|v(\gamma_1')| < r^{-1}/2, \quad \langle v(\gamma_1'), u_0 \rangle > cr^{-1} \quad \text{and} \quad \theta(\gamma_1') = 1.$$

Proof. Let $\gamma_1 \in B(5\varepsilon)$ be an isometry with the properties stated in Lemma 13, and write $v = v(\gamma_1)$. For simplicity, we assume that $\langle v, u_0 \rangle > 0$; the other case is similar. By the assumption (SSR), we have

$$\int \langle \theta v, u_0 \rangle dm_{K^\circ}(\theta) = 0.$$

Thus, there is $\theta_1 \in K^\circ$, such that $\langle \theta_1 v, u_0 \rangle \leq 0$.

There is a curve $\Theta : [0,1] \to K^\circ$ such that $\Theta(0) = 1$ and $\Theta(1) = \theta_1$, and the length of the curve $[0,1] \to \Theta(t)v$ is less than $C|v|$, where C depends only on the embedding of K° to $\mathrm{O}(d)$ and hence on μ. Then there is a sequence of rotations

$$\sigma_0 = 1, \sigma_1, \sigma_2, \ldots, \sigma_N = \theta_1 \in K^\circ$$

with $N \leq 2Cr|v| + 1$ such that for any $1 \leq i \leq N$,

$$|\sigma_i v - \sigma_{i-1} v| < r^{-1}/2.$$

By the triangle inequality, there is an index $1 \leq i \leq N$ such that

$$\langle \sigma_{i-1} v - \sigma_i v, u_0 \rangle \geq \langle v, u_0 \rangle / N \geq cr^{-1},$$

with a suitably small constant $c > 0$. (c depends on C and the constants appearing in the previous lemma.)

Now let $g_i \in B(4\varepsilon)$ be such that $\theta(g_i) = \sigma_i$; such elements can be found by virtue of Lemma 11. The proof is finished by an easy verification of the stated properties for the element

$$\gamma_1' := g_{i-1} \gamma_1 g_{i-1}^{-1} g_i \gamma_1^{-1} g_i^{-1}. \qquad \square$$

Proof of Lemma 8. We assume to the contrary that (8) fails for both $i = 1$ and $i = 2$. Let $c > 0$ be the constant from Lemma 14. Clearly, there is a point $u_0 \in S^{d-1}$ such that

$$\int_{|\xi - u_0| < c/2} |\varphi|^2 d\xi > c', \tag{10}$$

with a constant $c' > 0$ that depends only on c and d.

By Lemma 14, there is an element $\gamma_1' \in B(26\varepsilon)$ such that $\theta(\gamma_1') = 1$ and $v' := v(\gamma_1')$ satisfies $|v'| < r^{-1}/2$, and $|\langle v', u_0\rangle| > cr^{-1}$. This leads to the inequality

$$\|\varphi - \rho_r(\gamma_1')\varphi\|_2 = \int_{S^{d-1}} |(1 - e(r\langle \xi, v'\rangle)\varphi(\xi)|^2 d\xi < (26\varepsilon)^2.$$

If $|\xi - u_0| < c/2$, then $c/2 < |r\langle \xi, v'\rangle| < 1/2$. This and (10) give

$$|1 - e(c/2)|^2 c' < (26\varepsilon)^2,$$

which is a contradiction if we choose ε to be sufficiently small. Since c and c' depend only on μ, it follows that ε depends only on μ. We chose C_1 depending on μ and ε in Lemma 11 and C_2 depending on C_1 and μ in Lemma 13. Thus all these parameters depend only on μ. □

5. Estimates for low frequencies

We recall some of our notation: μ is a fixed probability measure on $\mathrm{Isom}(\mathbf{R}^d)$, K is the closure of the rotation group generated by $\operatorname{supp}\theta(\widetilde{\mu} * \mu)$, and we assume that $\operatorname{supp}\theta(\mu) \subset \theta_0 K$, where $\theta_0 \in \mathrm{O}(d)$ is a rotation that normalizes K.

Fix a point $x_0 \in \mathbf{R}^d$, the starting point of the random walk, and fix a real number $r \geq 0$. Define

$$\psi_0(\xi) = e(r\langle x_0, \xi\rangle) = \mathrm{Res}_r(\widehat{\delta_{x_0}})(\xi)$$

for $\xi \in S^{d-1}$, which is the Fourier transform of the measure δ_{x_0} restricted to the sphere of radius r. Our objective in this section is to estimate $S_r^l \psi_0 = \mathrm{Res}_r(\widehat{\nu}_l)$. The estimate will be useful in the range $r < l^{-1/2}\log l$, i.e., when the frequency is sufficiently small.

The next proposition shows that $S_r^l\psi_0$ is approximated by the Fourier transform of a Gaussian distribution with covariance matrix Δ that depends on μ.

Proposition 15. *Assume that μ is almost nondegenerate, has finite moments of order α for some $\alpha \geq 2$ and satisfies* (C). *There is a constant C and a symmetric positive definite quadratic form $\Delta(\xi, \xi)$ on $\mathbf{R}^d$ invariant under the action of K and θ_0 such that the following holds*:

$$\|S_r^l\psi_0 - e^{-r^2 l\Delta}\|_2 < C(r^{\min\{1,\alpha-2\}} + |x_0|^2 r^2). \tag{11}$$

Moreover, if μ is symmetric or satisfies (E), *then we have the better bound*:

$$\|S_r^l\psi_0 - e^{-r^2 l\Delta}\|_2 < C(r^{\min\{2,\alpha-2\}} + |x_0|^2 r^2). \tag{12}$$

C and Δ depend only on μ. When $\alpha = 2$, we can replace the right-hand sides of (11) *and* (12) *by $o(1) + C(|x_0|^2 r^2)$ as $r \to 0$.*

The role of the last sentence is simply that we can conclude the Central Limit Theorem even in the case $\alpha = 2$.

The rest of this section is devoted to the proof of this proposition. We will give a slight improvement in Section 5.4 for the range $r > l^{-1/2}$. However, this improvement requires the assumption (SSR), and it is based on the results of Section 4.

Throughout this section we make the following assumptions. We assume that r is small, i.e. $r < c\min\{1, |x_0|^{-1}\}$, where c is a suitable small constant. For r larger, the statement of the proposition is vacuous. We assume that μ is almost nondegenerate, has finite moments of order $\alpha \geq 2$ and satisfies (C). In addition, at certain parts we assume that μ is symmetric or satisfies (E), but we always mention these explicitly.

The argument is based on Tutubalin's paper [23]. The most significant difference is that we consider the following, more general, decomposition of the space $\mathcal{H} := L^2(S^{d-1})$. This is due to the fact that we do not assume $K = \mathrm{SO}(d)$.

Let $\mathcal{H}_0$ be the subspace of functions $\varphi \in \mathcal{H}$ that are fixed by the action of K; i.e., $\varphi(\theta\xi) = \varphi(\xi)$ for every $\theta \in K$. For later reference we note that if $\varphi \in \mathcal{H}$, then the orthogonal projection of φ to $\mathcal{H}_0$ is obtained by the formula

$$\int \varphi(\theta\xi) dm_K(\theta). \tag{13}$$

Denote by $\mathcal{P}_k \subset \mathcal{H}$ the space of functions that are restrictions of degree k polynomials to S^{d-1}. We define the spaces $\mathcal{H}_k$, $k \geq 1$ recursively. Once $\mathcal{H}_k$ is defined, let $\mathcal{H}_{k+1}$ be the orthogonal complement of $\mathcal{H}_k$ in the space

$$\mathrm{span}\,\{\psi\varphi : \psi \in \mathcal{P}_{k+1}, \varphi \in \mathcal{H}_0\},$$

where $\mathrm{span}\,\{\cdot\}$ denotes the smallest closed subspace that contains the functions inside the brackets. Since $\mathcal{P}_k \subset \mathcal{H}_0 \oplus \cdots \oplus \mathcal{H}_k$, we indeed have

$$\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_1 \oplus \cdots .$$

Let $\mathcal{H}_\infty = \mathcal{H}_4 \oplus \mathcal{H}_5 \oplus \cdots$. Finally, let $P_i : \mathcal{H} \to \mathcal{H}_i$ be the orthogonal projection operator for each $i \in \{0, 1, \ldots, \infty\}$.

In the special case $K = \mathrm{SO}(d)$, $\mathcal{H}_k$ is the familiar space of spherical harmonics of degree k that was considered in Tutubalin's paper.

Taking $r = 0$, it is easy to see that the above subspaces are invariant for S_0. Below, we will show that they are "almost invariant" for small r; more precisely, we will bound the norm of $P_i S_r P_j$ by a polynomial of r for $i \neq j$. Additionally, we will see that the norm of $P_i S_r P_i$ for all $1 \leq i < \infty$ is strictly less than 1. However, the dependence on i would require a more careful analysis. Fortunately, we do not need to do this here, since as we will see, the contribution of the spaces $\mathcal{H}_i$, $i \geq 4$ is negligible compared to other error terms.

(This is why we introduced the notation for the space $\mathcal{H}_\infty$.) These estimates, which are simply based on Taylor expansion, will be given in Section 5.1.

To simplify notation, we write $\psi_l = (P_0 S_r P_0)^l \psi_0$ for $l \geq 1$. We will use the almost invariance of S_r mentioned in the previous paragraph to show that ψ_l is a good approximation to $S_r^l \psi_0$. This is done in two steps. We set $P = P_0 + P_1 + P_2 + P_3$, and we consider another sequence, defined by $\psi_l' = (P S_r P)^l \psi_0$ for $l \geq 0$. The next two lemmata, which will be proved in Section 5.2, claim that ψ_l' approximates $S_r^l \psi_0$ and ψ_l approximates ψ_l'.

Lemma 16. *There is a constant $C > 0$ such that the following holds:*

$$\|\psi_l' - S_r^l \psi_0\|_2 \leq C(r^{\min\{\alpha-2,2\}} + (|x_0|r)^4).$$

When $\alpha = 2$, we can replace the right-hand side by $o(1) + C(|x_0|r)^4$ as $r \to 0$.

Lemma 17. *There are constants $C, c > 0$ depending only on μ such that the following holds for $l \geq C \log(r^{-1}|x_0| + 2)$:*

$$\|\psi_l - \psi_l'\|_2 \leq C e^{-cr^2 l} r.$$

If μ is symmetric, then

$$\|\psi_l - \psi_l'\|_2 \leq C e^{-cr^2 l} r^2.$$

If μ satisfies (E) (*but not necessarily symmetric*), *then*

$$\|\psi_l - \psi_l'\|_2 \leq C e^{-cr^2 l} (r^{\min\{\alpha-1,2\}} + |x_0| r^2).$$

In light of these lemmata, it remains to understand the operator $P_0 S_r P_0$. This is essentially a multiplication operator as the next formula shows. For $\varphi \in \mathcal{H}_0$,

$$P_0 S_r P_0 \varphi(\xi) = \iint e(r\langle \sigma\xi, v(\gamma)\rangle)\varphi(\theta(\gamma)^{-1}\sigma\xi)\, d\mu(\gamma)\, dm_K(\sigma) \tag{14}$$
$$= F(\xi)\varphi(\theta_0^{-1}\xi),$$

where

$$F(\xi) = \iint e(r\langle \sigma\xi, v(\gamma)\rangle)\, d\mu(\gamma)\, dm_K(\sigma). \tag{15}$$

Recall that $\operatorname{supp} \theta(\mu) \subset \theta_0 K$ and θ_0 normalizes K.

Based on this formula and the Taylor expansion of the function F, we will prove the following lemma in Section 5.3.

Lemma 18. *There are constants $C, c > 0$ and a quadratic form Δ on $\mathbf{R}^d$ depending only on μ such that*

$$\|\psi_l - e^{-lr^2\Delta}\|_2 < C e^{-clr^2} (r^{\min\{1,\alpha-2\}} + |x_0|^2 r^2).$$

Δ is invariant under K and θ_0. When $\alpha = 2$, we can replace the right-hand side by $o(1) + C(|x_0|^2 r^2)$ as $r \to 0$.

If μ is symmetric or satisfies (E), *then we have the better estimate*

$$\|\psi_l - e^{-lr^2\Delta}\|_2 < Ce^{-clr^2}(r^{\min\{2,\alpha-2\}} + |x_0|^2 r^2).$$

Proposition 15 immediately follows from Lemmata 16–18.

5.1. *Taylor expansion and approximate invariance.* We give some estimates for the norm of the operators $P_i S_r P_j$ in this section. These will be deduced from the following lemma, which is based on the Taylor series expansion of the function $e(r\langle \xi, v(\gamma)\rangle)$, which is the multiplier in the representation ρ_r.

LEMMA 19. *There is an absolute constant $C > 0$ such that for any $\varphi \in \mathcal{H}$ with $\|\varphi\|_2 = 1$ and $\gamma \in \mathrm{Isom}(\mathbf{R}^d)$ with $\theta(\gamma) \in \theta_0 K$, we have*

$$\|P_i\rho_r(\gamma)P_i\varphi - \rho_0(\gamma)P_i\varphi\|_2 < Cr|v(\gamma)|, \tag{16}$$

$$\|P_j\rho_r(\gamma)P_i\varphi\|_2 < \min\{1, C(r|v(\gamma)|)^{|i-j|}\}. \tag{17}$$

Proof. For the proof, we can assume that $\varphi \in \mathcal{H}_i$; that is, $P_i\varphi = \varphi$. By Taylor's theorem,

$$\begin{aligned} \rho_r(\gamma)\varphi(\xi) &= e(r\langle \xi, v(\gamma)\rangle)\varphi(\theta(\gamma)^{-1}\xi) \\ &= \left[\sum_{m=0}^{M-1} C_m r^m \langle \xi, v(\gamma)\rangle^m + O(r^M|v(\gamma)|^M)\right]\varphi(\theta(\gamma)^{-1}\xi), \end{aligned} \tag{18}$$

where $C_0 = 1$, and C_m and the implied constant are absolute.

To deduce (16), take $M = 1$ in (18), apply P_i to both sides and subtract

$$\rho_0(\gamma)\varphi(\xi) = \varphi(\theta(\gamma)^{-1}\xi) = P_i(\varphi(\theta(\gamma)^{-1}\xi)).$$

To deduce (17) when $j > i$, take $M = j - i$. Write

$$q(\xi) = \sum_{m=0}^{j-i-1} C_m r^m \langle \xi, v(\gamma)\rangle^m \in \mathcal{P}_{j-i-1}.$$

Since $\varphi \in \mathcal{H}_i$, we have

$$\varphi = p_1\psi_1 + \cdots + p_k\psi_k,$$

with some $p_1, \ldots, p_k \in \mathcal{P}_i$ and $\psi_1, \ldots, \psi_k \in \mathcal{H}_0$. Then

$$\begin{aligned} &\sum_{m=0}^{j-i-1} C_m r^m \langle \xi, v(\gamma)\rangle^m \varphi(\theta(\gamma)^{-1}\xi) \\ &\qquad = q(\xi)[p_1(\theta(\gamma)^{-1}\xi)\psi_1(\theta_0^{-1}\xi) + \cdots + p_k(\theta(\gamma)^{-1}\xi)\psi_k(\theta_0^{-1}\xi)], \end{aligned}$$

where $q(\xi)p_n(\theta(\gamma)^{-1}\xi) \in \mathcal{P}_{j-1}$ for any $1 \le n \le k$. Thus after applying P_j, these terms vanish in (18). Then we get

$$P_j\rho_r(\gamma)\varphi = O(r^{j-i}|v(\gamma)|^{j-i})\varphi(\theta(\gamma)^{-1}\xi),$$

which proves (17) when $j > i$.

Now let $j < i$, and let $\psi \in \mathcal{H}_j$ with $\|\psi\|_2 = 1$ be such that

$$\|P_j \rho_r(\gamma)\varphi\|_2 = \langle \rho_r(\gamma)\varphi, \psi\rangle = \langle \varphi, \rho_r(\gamma^{-1})\psi\rangle \le \|P_i \rho_r(\gamma^{-1})\psi\|_2.$$

Then the claim follows from (17) applied for ψ and γ^{-1} and the role of i and j reversed. $\square$

LEMMA 20. *There are constants $c < 1$ and C depending only on μ such that the following hold*:

$$\|P_i S_r P_i\| \le c \quad \text{for} \quad r < c \quad \text{and} \quad 1 \le i \le 3,$$
$$\|P_i S_r P_j\| \le C r^{\min\{|i-j|,\alpha\}}.$$

When $|i-j| > \alpha$, we can replace the second estimate by $o(r^{\min\{|i-j|,\alpha\}})$.

Proof. To prove the first inequality, we integrate (16) with respect to $d\mu(\gamma)$:

$$\text{(19)} \qquad \|P_i S_r P_i \varphi - S_0 P_i \varphi\|_2 = \left\| \int P_i \rho_r(\gamma) P_i \varphi - \rho_0(\gamma) P_i \varphi d\mu(\gamma) \right\|_2 < Cr \int |v(\gamma)| d\mu(\gamma).$$

This inequality shows that it is enough to estimate the norm of S_0 on $\mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \mathcal{H}_3$. Denote by $\mathcal{P}'$ the orthogonal complement of $\mathcal{H}_0 \cap \mathcal{P}_3$ in $\mathcal{P}_3$. For each $\varphi \in \mathcal{P}'$, $\|S_0^* S_0 \varphi\|_2 < \|\varphi\|_2$, because otherwise φ would be invariant under K; i.e., we would have $\varphi \in \mathcal{H}_0$. Since $\mathcal{P}'$ is finite dimensional, there is a constant $c < 1$ such that $\|S_0^* S_0 \varphi\|_2 < c\|\varphi\|_2$ for $\varphi \in \mathcal{P}'$.

Let $\varphi_1, \dots, \varphi_k$ be an orthonormal basis of $\mathcal{P}'$ consisting of eigenfunctions of $S_0^* S_0$. Observe that the spaces $\varphi_i \cdot \mathcal{H}_0$ are eigenspaces of $S_0^* S_0$ with the same eigenvalues as φ_i. Note that any $\varphi \in \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \mathcal{H}_3$ is of the form $p_1\psi_1 + \dots + p_k\psi_k$ with $p_i \in \mathcal{P}'$ and $\psi_i \in \mathcal{H}_0$. Hence the eigenspaces $\varphi_i \cdot \mathcal{H}_0$ span $\mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \mathcal{H}_3$. Then we are able to conclude that

$$\|P_i S_0 P_i\|^2 \le \|P_i S_0^* S_0 P_i\| < c$$

for $i = 1, 2, 3$. This combined with (19) proves the first claim provided r is sufficiently small.

For the second claim, we integrate (17):

$$\|P_j S_r P_i\| < \int \min\{1, C(r|v(\gamma)|)^{|i-j|}\} d\mu(\gamma).$$

If $|i-j| \le \alpha$, we can simply write

$$\|P_j S_r P_i\| < C r^{|i-j|} \int |v(\gamma)|^{|i-j|} d\mu(\gamma),$$

and the claim follows from the moment condition on μ. If $|i-j| > \alpha$, we write

$$\|P_j S_r P_i\| < Cr^\alpha \int |v(\gamma)|^\alpha \min\{(r|v(\gamma)|)^{-\alpha}, (r|v(\gamma)|)^{|i-j|-\alpha}\} d\mu(\gamma).$$

Observe that

$$\min\{(r|v(\gamma)|)^{-\alpha}, (r|v(\gamma)|)^{|i-j|-\alpha}\} \leq 1$$

and that it tends to 0 for all γ as $r \to 0$. Now the claim follows by the dominated convergence theorem. $\square$

The bound on $P_1 S_r P_0$ in Lemma 20 is not optimal. Indeed, it is easy to see that the linear terms in the Taylor expansions cancel thanks to condition (C).

Lemma 21. *There is a constant C such that*

$$\|P_1 S_r P_0\| \leq Cr^2.$$

If μ also satisfies (E), *then we get the better bound*

$$\|P_1 S_r P_0\| \leq Cr^{\min\{3,\alpha\}}.$$

Proof. Take $\varphi \in \mathcal{H}_0$, and as in the proof of Lemma 19, write the Taylor expansion

$$\rho_r(\gamma)\varphi(\xi) = \left[1 + C_1 r\langle \xi, v(\gamma)\rangle + C_2 r^2 \langle \xi, v(\gamma)\rangle^2 + O(r^3|v(\gamma)|^3)\right] \varphi(\theta_0^{-1}\xi).$$

Similarly to the proof of Lemma 20, we integrate this inequality. To get the first claim, we only need to note that

$$\int \langle \xi, v(\gamma)\rangle d\mu(\gamma) = \left\langle \xi, \int v(\gamma) d\mu(\gamma) \right\rangle = 0$$

because of assumption (C).

Assumption (E) implies that $\mathcal{H}_0$ consists of even functions, and hence $\mathcal{H}_1$ contains only odd ones. Since

$$\int \langle \xi, v(\gamma)\rangle^2 d\mu(\gamma) \cdot \varphi(\theta_0^{-1}\xi)$$

is an even function of ξ, it is in the kernel of P_1. This establishes the second claim. $\square$

We also need a norm estimate for $P_0 S_r P_0$. As we remarked in (14), this operator is essentially a multiplication operator by the function F defined in (15). Hence what we need to understand is the behavior of F near the origin.

Lemma 22. *There is a constant C such that*

$$|F(\xi) - (1 - r^2\Delta(\xi,\xi))| < Cr^{\min\{3,\alpha\}},$$

where $\Delta(\xi,\xi)$ is a positive definite quadratic form depending on μ. If $\alpha < 3$, then the above bound can be improved to $o(r^\alpha)$. Furthermore, if μ is symmetric or satisfies (E), *then we have the improved bound*

$$|F(\xi) - (1 - r^2\Delta(\xi,\xi))| < Cr^{\min\{4,\alpha\}}.$$

As an immediate corollary, we get that $\|P_0 S_r P_0\| < 1 - cr^2$ for some constant $c > 0$.

Proof. We expand (15), the definition of F, in Taylor series the same way as we did in the previous lemmata:

$$F(\xi) = \iint 1 + C_1 r\langle\sigma\xi, v(\gamma)\rangle - C_2 r^2\langle\sigma\xi, v(\gamma)\rangle^2 + C_3 r^3\langle\sigma\xi, v(\gamma)\rangle^3 dm_K(\sigma)d\mu(\gamma) + O(r^{\min\{\alpha,4\}}), \tag{20}$$

where C_1, C_2, C_3 are absolute constants, $C_2 > 0$, and the implied constant is depends only on μ.

First we note that as in the proof of Lemma 21, (C) implies that

$$\int \langle\sigma\xi, v(\gamma)\rangle d\mu(\gamma) = 0$$

for all σ and ξ. Hence the linear term vanishes in the Taylor expansion (20).

Second, the quadratic term in (20),

$$\Delta(\xi,\xi) := \iint C_2 r^2\langle\sigma\xi, v(\gamma)\rangle^2 dm_K(\sigma)d\mu(\gamma),$$

is clearly a K invariant positive semi-definite quadratic form. We only need to show that it is strictly positive definite. Denote by V the maximal subspace of $\mathbf{R}^d$ on which $\Delta(\xi,\xi)$ vanishes. By the definition of Δ, all $v(\gamma)$ is orthogonal to V, which would contradict almost nondegeneracy if $V \neq \{0\}$.

If $\alpha \geq 3$ and (E) is satisfied, then for all ξ, there is $\sigma_\xi \in K$ such that $\sigma_\xi\xi = -\xi$. Then

$$2\int \langle\sigma\xi, v(\gamma)\rangle^3 dm_K(\sigma) = \int \langle\sigma\xi, v(\gamma)\rangle^3 + \langle\sigma\sigma_\xi\xi, v(\gamma)\rangle^3 dm_K(\sigma) = 0;$$

hence the cubic term in (20) vanishes.

Finally, if μ is symmetric and $\alpha \geq 3$, then we also have

$$2\iint \langle\sigma\xi, v(\gamma)\rangle^3 dm_K(\sigma)d\mu(\gamma) = \iint \langle\xi, \sigma v(\gamma)\rangle^3 + \langle\xi, \sigma\theta(\gamma)v(\gamma^{-1})\rangle^3 dm_K(\sigma)d\mu(\gamma) = 0.$$

The first equality follows since μ is symmetric and m_K is invariant under multiplication by $\theta(\gamma)$ from the left. The second one follows from $\theta(\gamma)v(\gamma^{-1}) = -v(\gamma)$. The claim now follows from (20). $\square$

5.2. *Approximating S_r by $P_0 S_r P_0$.* The purpose of this section is to prove Lemmata 16 and 17. For both of them, we need the next lemma, which provides some estimates for the projections of ψ'_l to the spaces $\mathcal{H}_i$.

LEMMA 23. *There are constants c and C such that the following hold:*

$$\|P_1\psi'_l\|_2 \le C(r^2 + e^{-cl}|x_0|r), \tag{21}$$

$$\|P_i\psi'_l\|_2 \le C(r^{\min\{i,\alpha\}} + e^{-cl}|x_0|^i r^i) \tag{22}$$

for $i \in \{2,3\}$ and $l \ge 0$.

For $l \ge C\log(r^{-1}|x_0|+2)$, we have

$$\|P_0\psi'_l\|_2 \le Ce^{-cr^2 l}, \tag{23}$$

$$\|P_1\psi'_l\|_2 \le Cr^2 e^{-cr^2 l}, \tag{24}$$

$$\|P_i\psi'_l\|_2 \le Cr^{\min\{i,\alpha\}} e^{-cr^2 l} \tag{25}$$

for $i \in \{2,3\}$. Moreover, if μ satisfies (E), *then we can replace r^2 by $r^{\min\{3,\alpha\}}$ in* (21) *and* (24).

The following proof is very technical, although the idea behind it is very simple. The argument is based on induction on l, the norm estimates of the previous section and triangle inequality.

We first give a brief sketch, which explains why the induction works. For simplicity, take $x_0 = 0$ and suppose that the lemma holds for some

$$\log(r^{-1}|x_0|+2) < l < r^{-2}.$$

(For simplicity, we consider only this range in this informal discussion.) We can write

$$P_i\psi'_{l+1} = \sum_{j=0}^{3} P_i S_r P_j \psi'_l.$$

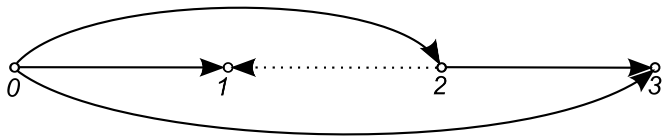
We use the induction hypothesis and the lemmata of the previous section to bound the terms.

What we need for the argument to work are essentially the inequalities

$$(1 - \|P_i S_r P_i\|)X_i > \sum_{j\ne i} \|P_i S_r P_j\| \|P_j\psi'_l\|_2, \tag{26}$$

where X_i is the bound for $\|P_i\psi'_l\|_2$ claimed in (23)–(25). Notice that if we plug in the estimates that we obtained in the previous section, then all terms on the right-hand side are of the same or smaller order of magnitude than the left-hand side for any i and j. For example, take $i = 2$: Then $(1-\|P_2 S_r P_i\|)X_i \ge Cr^2$ by Lemma 20. For the terms on the right, we have $\|P_2 S_r P_j\|\|P_j\psi'_l\|_2 \le Cr^{|2-j|}X_j$ by Lemma 20 again. We see that $\|P_2 S_r P_0\|\|P_0\psi'_l\|_2 \le Cr^2$ and all the other terms are of lower order.

In the following diagram we draw a directed edge from vertex j to i if the corresponding term on the right of (26) is of the same order as the left-hand side:



(The dotted edge is present when (E) is assumed.) Since it does not have a directed cycle, we can set the constants in (23)–(25) following the edges of the diagram such that the induction will work.

Proof of Lemma 23. We do not give a separate proof for the last statement, but it will be clear that using the improved estimates in Lemma 21 available under (E), we get the better error terms. Let β_0 and B_0 be constants whose existence is guaranteed by Lemmata 20–22 such that

$$\|P_0 S_r P_0\| \le e^{-\beta_0 r^2}, \tag{27}$$

$$\|P_i S_r P_i\| \le e^{-\beta_0} \quad \text{for } i \in \{1,2,3\}, \tag{28}$$

$$\|P_i S_r P_j\| \le B_0 r^{\min\{\alpha, |i-j|\}}, \tag{29}$$

$$\|P_1 S_r P_0\| \le B_0 r^2. \tag{30}$$

The proof is by induction, and we begin with (21) and (22). We suppose that $|x_0| \ge 1$. In the opposite case the argument is identical; we only need to replace $|x_0|^2$ and $|x_0|^3$ everywhere by $|x_0|$. We show that

$$\|P_1 \psi_l'\|_2 \le C_1 (r^2 + e^{-\beta_0 l/2} |x_0| r),$$
$$\|P_i \psi_l'\|_2 \le C_i (r^{\min\{i,\alpha\}} + e^{-\beta_0 l/2} |x_0|^i r^i)$$

for $i \in \{2,3\}$ and $l \ge 0$, where $C_i > 1$ are suitable constants to be specified later. For $l = 0$, the claim is verified easily by the Taylor expansion of ψ_0.

Suppose that the claim holds for l; we prove it for $l+1$. To estimate

$$\|P_i \psi_{l+1}'\|_2 \le \sum_{j=0}^{3} \|P_i S_r P_j\| \cdot \|P_j \psi_l'\|, \tag{31}$$

we use the induction hypothesis for $\|P_j \psi_l'\|_2$ and the norm estimates (27)–(30). For $i = 1$, we write

$$\begin{aligned}
\|P_1 \psi_{l+1}'\|_2 &\le B_0 r^2 + e^{-\beta_0} C_1 (r^2 + e^{-\beta_0 l/2} |x_0| r) \\
&\quad + B_0 r C_2 (r^2 + e^{-\beta_0 l/2} |x_0|^2 r^2) + B_0 r^2 C_3 (r^{\min\{3,\alpha\}} + e^{-\beta_0 l/2} |x_0|^3 r^3) \\
&\le \Big([B_0 (1 + r C_2 + r^2 C_3) + e^{-\beta_0} C_1] e^{\beta_0/2} \Big) \cdot (r^2 + e^{-\beta_0 (l+1)/2} |x_0| r).
\end{aligned}$$

To obtain the last line, we use inequalities of type

$$(r^2 + e^{-\beta_0 l/2}|x_0|^j r^j) \le e^{\beta_0/2}(r^2 + e^{-\beta_0(l+1)/2}|x_0|^j r^j)$$

and also $r|x_0| < 1$ that we can suppose without loss of generality, as we mentioned at the beginning of Section 5. For $i = 2$,

$$\begin{aligned}\|P_2\psi'_{l+1}\|_2 \le{}& B_0 r^2 + B_0 r C_1 (r^2 + e^{-\beta_0 l/2}|x_0| r),\\ &+ e^{-\beta_0} C_2 (r^2 + e^{-\beta_0 l/2}|x_0|^2 r^2) + B_0 r C_3 (r^{\min\{3,\alpha\}} + e^{-\beta_0 l/2}|x_0|^3 r^3),\\ &\le \Big([B_0(1 + C_1 + rC_3) + e^{-\beta_0} C_2] e^{\beta_0/2}\Big)\cdot(r^2 + e^{-\beta_0(l+1)/2}|x_0|^2 r^2).\end{aligned}$$

We derive the last line the same way as before, but we also use the inequality $|x_0| \ge 1$. For $i = 3$,

$$\begin{aligned}&\|P_3\psi'_{l+1}\|_2 \le B_0 r^{\min\{3,\alpha\}} + B_0 r^2 C_1 (r^2 + e^{-\beta_0 l/2}|x_0| r)\\ &+ B_0 r C_2 (r^2 + e^{-\beta_0 l/2}|x_0|^2 r^2) + e^{-\beta_0/2} C_3 (r^{\min\{3,\alpha\}} + e^{-\beta_0 l/2}|x_0|^3 r^3)\\ &\le \Big([B_0(1 + C_1 + C_2) + e^{-\beta_0} C_3] e^{\beta_0/2}\Big)\cdot(r^{\min\{3,\alpha\}} + e^{-\beta_0(l+1)/2}|x_0|^3 r^3).\end{aligned}$$

Now the claim is satisfied if we take

$$\begin{aligned}C_1 &= 2e^{\beta_0/2} B_0/(1 - e^{-\beta_0/2}),\\ C_2 &= 2e^{\beta_0/2} B_0 (1 + C_1)/(1 - e^{-\beta_0/2}),\\ C_3 &= e^{\beta_0/2} B_0 (1 + C_1 + C_2)/(1 - e^{-\beta_0/2})\end{aligned}$$

and r is so small that $rC_2 + r^2 C_3 < 1$ and $rC_3 < 1$.

The proof of (23)–(25) is very similar. We begin by choosing l_0 such that $e^{-\beta_0 l_0/2}|x_0| r < r^2$, but $l_0 < 2\beta_0^{-1}\log(|x_0|/r) + 1$. We show by induction that for $l \ge l_0$ the following holds:

$$\begin{aligned}\|P_0\psi'_l\|_2 &\le C'_0 e^{-\beta_0 r^2 l/2},\\ \|P_1\psi'_l\|_2 &\le C'_1 r^2 e^{-\beta_0 r^2 l/2},\\ \|P_i\psi'_l\|_2 &\le C'_i r^{\min\{i,\alpha\}} e^{-\beta_0 r^2 l/2}\end{aligned}$$

for $i \in \{2, 3\}$ and $l \ge 0$, where $C'_0 = e^{\beta_0 r^2 l_0/2}$ and for $i > 0$, $C'_i \ge C_i e^{\beta_0 r^2 l_0/2}$ are suitable constants to be specified later. We note that

$$\beta_0 l_0/2 \le \log(|x_0|/r) + 1 \le \log(r^{-2}) + 1 < r^{-2}$$

since $|x_0| r < 1$. Hence $e^{\beta_0 r^2 l_0/2} < e$, so the above bounds on the constants C'_i are independent of x_0 and r.

From the first part of the proof it follows that the claim holds for $l = l_0$. Now we suppose that it holds for a particular $l \ge l_0$ and prove it for $l + 1$. As above, we use (31) along with the induction hypothesis and (27)–(30). We get

$$\begin{aligned}\|P_0\psi'_{l+1}\|_2 &\le e^{-\beta_0 r^2}C'_0 e^{-\beta_0 r^2 l/2} + B_0 r C'_1 r^2 e^{-\beta_0 r^2 l/2}\\ &\quad + B_0 r^2 C'_2 r^2 e^{-\beta_0 r^2 l/2} + B_0 r^{\min\{3,\alpha\}} C'_3 r^{\min\{3,\alpha\}} e^{-\beta_0 r^2 l/2}\\ &\le \left([C'_0 e^{-\beta_0 r^2} + B_0(r^3 C'_1 + r^4 C'_2 + r^{\min\{6,2\alpha\}} C'_3)]e^{\beta_0 r^2/2}\right)\cdot e^{-\beta_0 r^2(l+1)/2},\end{aligned}$$

$$\begin{aligned}\|P_1\psi'_{l+1}\|_2 &\le B_0 r^2 C'_0 e^{-\beta_0 r^2 l/2} + e^{-\beta_0} C'_1 r^2 e^{-\beta_0 r^2 l/2}\\ &\quad + B_0 r C'_2 r^2 e^{-\beta_0 r^2 l/2} + B_0 r^2 C'_3 r^{\min\{3,\alpha\}} e^{-\beta_0 r^2 l/2}\\ &\le \left([e^{-\beta_0} C'_1 + B_0(C'_0 + r C'_2 + r^{\min\{3,\alpha\}} C'_3)]e^{\beta_0 r^2/2}\right)\cdot r^2 e^{-\beta_0 r^2(l+1)/2},\end{aligned}$$

$$\begin{aligned}\|P_2\psi'_{l+1}\|_2 &\le B_0 r^2 C'_0 e^{-\beta_0 r^2 l/2} + B_0 r C'_1 r^2 e^{-\beta_0 r^2 l/2}\\ &\quad + e^{-\beta_0} C'_2 r^2 e^{-\beta_0 r^2 l/2} + B_0 r C'_3 r^{\min\{3,\alpha\}} e^{-\beta_0 r^2 l/2}\\ &\le \left([e^{-\beta_0} C'_2 + B_0(C'_0 + r C'_1 + r^{\min\{2,\alpha-1\}} C'_3)]e^{\beta_0 r^2/2}\right)\cdot r^2 e^{-\beta_0 r^2(l+1)/2},\end{aligned}$$

$$\begin{aligned}\|P_3\psi'_{l+1}\|_2 &\le B_0 r^{\min\{3,\alpha\}} C'_0 e^{-\beta_0 r^2 l/2} + B_0 r^2 C'_1 r^2 e^{-\beta_0 r^2 l/2}\\ &\quad + B_0 r C'_2 r^2 e^{-\beta_0 r^2 l/2} + e^{-\beta_0} C'_3 r^{\min\{3,\alpha\}} e^{-\beta_0 r^2 l/2}\\ &\le \left([e^{-\beta_0} C'_3 + B_0(C'_0 + r C'_1 + C'_2)]e^{\beta_0 r^2/2}\right)\cdot r^{\min\{3,\alpha\}} e^{-\beta_0 r^2(l+1)/2}.\end{aligned}$$

Now choose the constants in such a way that

$$C'_1 \ge \frac{2C'_0 B_0 e^{\beta_0 r^2/2}}{(1-e^{-\beta_0+\beta_0 r^2/2})},\quad C'_2 \ge \frac{2C'_0 B_0 e^{\beta_0 r^2/2}}{(1-e^{-\beta_0+\beta_0 r^2/2})},$$
$$C'_3 \ge \frac{2B_0(C'_0 + C'_2)e^{\beta_0 r^2/2}}{(1-e^{-\beta_0+\beta_0 r^2/2})},$$

and observe that the claim holds for $l+1$ if r is sufficiently small. □

Proof of Lemma 16. By the triangle inequality, we have

$$\begin{aligned}\|\psi'_l - S_r^l\psi_0\|_2 &\le \sum_{k=0}^{l-1}\|S_r^{l-k-1}(S_r - PS_rP)(PS_rP)^k\psi'_0\|_2\\ &\le \sum_{k=0}^{l-1}\|(S_r - PS_rP)\psi'_k\|_2.\end{aligned}$$

To estimate the terms, we write

$$\begin{aligned}\|(S_r - PS_rP)\psi'_k\|_2 &= \left\|S_r P_\infty\psi'_k + \sum_{j=0}^{3}(S_rP_j - PS_rP_j)P_j\psi'_k\right\|_2\\ &\le \|P_\infty\psi'_k\|_2 + \sum_{j=0}^{3}\|P_\infty S_r P_j\|\cdot\|P_j\psi'_k\|_2.\end{aligned}\tag{32}$$

Recall that P_∞ is the projection to the complement of $\mathcal{H}_0 \oplus \cdots \oplus \mathcal{H}_3$. Note that $P_\infty \psi'_k = 0$ for $k \geq 1$ since $\psi'_k \in \mathcal{H}_0 \oplus \cdots \oplus \mathcal{H}_3$.

We use Lemmata 23 and 20. For $1 \leq k \leq C \log(r^{-1}|x_0| + 2)$, we have

$$\|(S_r - PS_rP)\psi'_k\|_2 < C(r^{\min\{4,\alpha\}} + |x_0|^3 r^4 e^{-ck}), \tag{33}$$

while for $k = 0$, we have to add $\|P_\infty \psi'_k\|_2 \leq C|x_0|^4 r^4$ to the above estimate. For $k \geq C \log(r^{-1}|x_0| + 2)$, we have

$$\|(S_r - PS_rP)\psi'_k\|_2 < Cr^{\min\{4,\alpha\}} e^{-cr^2 k}. \tag{34}$$

Summing for k, we get the statement of the lemma.

When $\alpha = 2$, the constants in Lemma 20 are arbitrarily small as $r \to 0$. If we plug these in (32), we see that the constants in (33) and (34) are also arbitrarily small. □

Proof of Lemma 17. By the triangle inequality,

$$\begin{aligned}
\|\psi_l - \psi'_l\|_2 &= \|(P_0 S_r P_0)^l \psi_0 - (PS_rP)^l \psi_0\|_2 \\
&\leq \sum_{k=0}^{l-1} \|(P_0 S_r P_0)^{l-k-1}(P_0 S_r P_0 - PS_rP)(PS_rP)^k \psi_0\|_2 \\
&\leq \sum_{k=0}^{l-2} \|P_0 S_r P_0\|^{l-k-1} \cdot \|(P_0 S_r P_0 - P_0 S_r P)\psi'_k\|_2 \\
&\quad + \|(P_0 S_r P_0 - PS_rP)\psi'_{l-1}\|_2.
\end{aligned}$$

As in the previous proof, we write

$$\|(P_0 S_r P_0 - P_0 S_r P)\psi'_k\|_2 \leq \sum_{j=1}^{3} \|P_0 S_r P_j\| \cdot \|P_j \psi'_k\|_2.$$

Again, we use Lemmata 23, 20, 21 and the estimate $\|P_0 S_r P_0\| \leq 1 - cr^2$, which follows from Lemma 22. For $k \leq \log(r^{-1}|x_0| + 2)$, we can write

$$\|P_0 S_r P_0\|^{l-k-1} \cdot \|(P_0 S_r P_0 - P_0 S_r P)\psi'_k\|_2 \leq Ce^{-cr^2 l}(|x_0| r^2 e^{-ck} + r^3),$$

while for $k \geq \log(r^{-1}|x_0| + 2)$, we get

$$\|P_0 S_r P_0\|^{l-k-1} \cdot \|(P_0 S_r P_0 - P_0 S_r P)\psi'_k\|_2 \leq Cr^3 e^{-cr^2(l-1)}$$

and

$$\begin{aligned}
\|(P_0 S_r P_0 - PS_rP)\psi'_{l-1}\|_2 &\leq \|(P_0 S_r P_0 - P_0 S_r P)\psi'_{l-1}\|_2 \\
&\quad + \|(P_0 S_r P - PS_rP)\psi'_{l-1}\|_2 \leq Cr^3 e^{-cr^2(l-1)} + Cr^2 e^{-cr^2(l-1)}.
\end{aligned}$$

Summing up for k, we get the first statement of the lemma.

If μ is symmetric, then S_r is selfadjoint; hence

$$\|P_0 S_r P_1\| = \|P_1 S_r P_0\| \leq Cr^2.$$

Using this instead of Lemma 20, we get

$$\|P_0 S_r P_0\|^{l-k-1} \cdot \|(P_0 S_r P_0 - P_0 S_r P)\psi'_k\|_2 \leq C r^4 e^{-cr^2(l-1)},$$

and the better estimate follows after summation.

If μ satisfies (E) instead, then the better estimate in Lemma 23 gives

$$\|P_0 S_r P_0\|^{l-k-1} \cdot \|(P_0 S_r P_0 - P_0 S_r P)\psi'_k\|_2 \leq C r^{\min\{4,\alpha+1\}} e^{-cr^2(l-1)}. \qquad \square$$

5.3. *Proof of Lemma* 18. Again, we only show the first inequality in the lemma; the second follows along the same lines by applying the improved estimate in Lemma 22.

Using the Taylor series expansion of ψ_0 together with (C) and finite second moments, we see that

$$\psi_0(\xi) = 1 + O(|x_0|^2 r^2), \tag{35}$$

where the implied constant only depends on μ. Furthermore, we have by (14) that

$$\psi_l(\xi) = \left[\prod_{j=0}^{l-1} F(\theta_0^{-j}\xi)\right] \psi_0(\theta_0^{-l}\xi), \tag{36}$$

where F is given by (15).

Let Δ be the quadratic form appearing in Lemma 22. Define Δ_0 to be its symmetrization by the group generated by θ_0; i.e.,

$$\Delta_0(\xi,\xi) := \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \Delta(\theta_0^{-i}\xi, \theta_0^{-i}\xi).$$

In light of (35) and (36) it is enough to show that

$$\left|\prod_{j=0}^{l-1} F(\theta_0^{-j}\xi) - e^{-lr^2\Delta_0(\xi,\xi)}\right| < C e^{-clr^2} r^{\min\{1,\alpha-2\}} \tag{37}$$

for all ξ.

The rest of the proof is devoted to this inequality. By Lemma 22, we have

$$\sum_{j=0}^{l-1} \log F(\theta_0^{-j}\xi) = -r^2 \sum_{j=0}^{l-1} \Delta(\theta_0^{-j}\xi, \theta_0^{-j}\xi) + O(l r^{\min\{3,\alpha\}}). \tag{38}$$

Denote by W the space of quadratic forms on $\mathbf{R}^d$, and denote by $\Theta_0 \in \mathrm{End}(W)$ the linear transformation induced by θ_0. It is easily seen that Θ_0 is diagonalizable and all its eigenvalues are on the unit circle of $\mathbf{C}$. Denote by W_0 the 1-eigenspace of Θ_0. Hence Δ_0 is the projection of Δ to W_0. Then on the orthogonal complement of W_0, we have

$$\sum_{j=0}^{l-1} \Theta_0^{-j} = \frac{\Theta_0^{-l} - 1}{\Theta_0^{-1} - 1}.$$

Thus it follows that

$$\sum_{j=0}^{l-1} \Delta(\theta_0^{-j}\xi, \theta_0^{-j}\xi) = \sum_{j=0}^{l-1} \Theta_0^{-j}\Delta(\xi,\xi) = l\Delta_0(\xi,\xi) + O(1).$$

The implied constant depends on the distance of the nontrivial eigenvalues of Θ_0 to 1.

If we combine our inequalities, we get

$$\sum_{j=0}^{l-1} \log F(\theta_0^{-j}\xi) = -lr^2\Delta_0(\xi,\xi) + O(r^2 + lr^{\min\{3,\alpha\}}).$$

This immediately implies that there is a constant $c > 0$ such that

$$\prod_{j=0}^{l-1} F(\theta_0^{-j}\xi) < e^{-clr^2}.$$

If we use the inequality $|e^A - e^B| < (A - B)\max\{e^A, e^B\}$, then we get

$$\left|\prod_{j=0}^{l-1} F(\theta_0^{-j}\xi) - e^{-lr^2\Delta_0(\xi,\xi)}\right| < Ce^{-clr^2}(r^2 + lr^{\min\{3,\alpha\}}). \tag{39}$$

To obtain (37) and hence the lemma, we only need to note that

$$e^{-clr^2} l \leq \frac{2}{c} \cdot e^{-clr^2/2} r^{-2}.$$

If $\alpha = 2$, the term $O(lr^{\min\{3,\alpha\}})$ in (38) can be improved to $o(lr^2)$ by Lemma 22. Hence the right-hand side of (39) can be improved to $o(1)$ as claimed.

5.4. *Some improvements using* (SSR). The purpose of this section is to give the following slight improvement of the bounds in Proposition 15. The proof depends on the results of Section 4, so it is important to note that Proposition 15 itself is enough for the arguments of Section 4. In fact, we can even get Theorem 3 without the results of this section at the modest expense of multiplying the first error term by $\log l$.

PROPOSITION 24. *Assume that μ is almost nondegenerate, has finite moments of order $\alpha \geq 2$ and satisfies* (C) *and* (SSR). *Then there are constants C, c and a symmetric positive definite quadratic form $\Delta(\xi,\xi)$ on $\mathbf{R}^d$ invariant under the action of K and θ_0 such that the following holds*:

$$\|S_r^l\psi_0 - e^{-r^2 l\Delta}\|_2 < C(r^{\min\{1,\alpha-2\}} + |x_0|^2r^2)\cdot(e^{-clr^2} + r^{10d}) \tag{40}$$

for $r < l^{-1/3}$. The constants C, c and the form Δ depend only on μ. Moreover, if μ is symmetric or satisfies (E), *then we have the better bound*:

$$\|S_r^l\psi_0 - e^{-r^2 l\Delta}\|_2 < C(r^{\min\{2,\alpha-2\}} + |x_0|^2r^2)\cdot(e^{-clr^2} + r^{10d}).$$

The quantities $l^{-1/3}$ and r^{10d} appearing in the proposition could be replaced by other powers of l and r. Notice that the estimate differs only by the factor $(e^{-clr^2}+r^{10d})$ compared to Proposition 15. We indicate how to modify the argument in the previous sections to obtain this improvement. We only need to sharpen Lemma 16 by the same factor.

First we note that Proposition 24 gives an improvement only if $l > r^{-2}$, and we recall that $l < r^{-3}$ is assumed in the proposition. Hence, we only consider the range $r^{-2} < l < r^{-3}$.

Similarly to the proof of Lemma 16, we write

$$\|S_r^l\psi_0 - (PS_rP)^l\psi_0\|_2 \le \sum_{j=0}^{l-1} \|S_r^{l-j-1}(S_r - PS_rP)\psi_j'\|_2.$$

To obtain the improvement of Lemma 16, we need to show the following improved estimates for the terms in the above sum:

$$\|S_r^{l-j-1}(S_r - PS_rP)\psi_j'\|_2 \le C(r^{\min\{4,\alpha\}} + |x_0|^2 r^2 e^{-cj}) \cdot (e^{-clr^2} + r^{20d}). \tag{41}$$

We have already seen that

$$\|(S_r - PS_rP)\psi_j'\|_2 < C(r^{\min\{\alpha,4\}} + |x_0|^2 r^2 e^{-cj})e^{-cr^2 j}.$$

We utilize Proposition 7 to estimate the norm of this function when S_r is applied.

If we have

$$\|S_r^m(S_r - PS_rP)\psi_j'\|_2 \le C(r^{\min\{4,\alpha\}} + |x_0|^2 r^2 e^{-cj}) \cdot r^{20d}$$

for any $m < l-j-1$, then (41) immediately follows. In the opposite case, we have $\|S_r^m(S_r - PS_rP)\psi_j'\|_2 \ge r^{4+20d}$. We will show below that in this case we have

$$\left\| \frac{S_r^m(S_r - PS_rP)\psi_j'}{\|S_r^m(S_r - PS_rP)\psi_j'\|_2} \right\|_{\mathrm{Lip}(K)} \le Cr^{-20d-4}m. \tag{42}$$

Since $r^2 < \log^{-3}((r+1)Cr^{-20d-4}m + 2)$, we have

$$\|S_r^{m+1}(S_r - PS_rP)\psi_j'\|_2 < e^{-cr^2}\|S_r^m(S_r - PS_rP)\psi_j'\|_2$$

by Proposition 7. Repeated application of this inequality gives the claim (41).

We turn to the proof of (42), which finishes the proof of Proposition 24. We introduce some notation. Let $\varphi \in C(S^{d-1})$ and $\xi \in S^{d-1}$, and write

$$\varphi^\xi(\theta) := \varphi(\theta\xi) \in C(K).$$

We note that

$$\sup_{\xi \in S^{d-1}} \|\varphi^\xi\|_{\mathrm{Lip}} \le \|\varphi\|_{\mathrm{Lip}(K)} \le \|\varphi\|_\infty + \sup_{\xi\in S^{d-1}} \|\varphi^\xi\|_{\mathrm{Lip}}. \tag{43}$$

Write

$$\mathcal{P}_3^\xi := \{\varphi^\xi : \varphi \in \mathcal{P}_3\}.$$

This is the space of polynomials of degree at most 3 restricted to the K-orbit of ξ pulled back to the group K. Denote by P^ξ the orthogonal projection to the space $\mathcal{P}_3^\xi$ in $L^2(K)$. We note that a function $\varphi \in C(S^{d-1})$ is in $\mathcal{H}_0 \oplus \cdots \oplus \mathcal{H}_3$ if and only if we have $\varphi^\xi \in \mathcal{P}_3^\xi$ for all $\xi \in S^{d-1}$. Moreover, $(P\varphi)^\xi = P^\xi \varphi^\xi$ for every $\varphi \in C(S^{d-1})$.

We first estimate $\|\psi_j'\|_{\mathrm{Lip}(K)}$. Note that any function in $\mathcal{P}_3^\xi$ is a polynomial of degree at most C_d in the entries of the matrices representing the elements of $K < \mathrm{O}(d)$. Here C_d is a number depending only on d. Then the space spanned by all P^ξ for $\xi \in S^{d-1}$ is finite dimensional; hence we can write $\|\varphi\|_{\mathrm{Lip}} \le C\|\varphi\|_2$ for any function $\varphi \in \mathcal{P}_3^\xi$ with a number C independent of ξ. This follows from the fact that any two norms on a finite-dimensional space are equivalent.

We show that $\|(\psi_j')^\xi\|_2 \le 1$ for every $\xi \in S^{d-1}$. By the previous paragraph, this implies that $\|(\psi_j')^\xi\|_{\mathrm{Lip}} \le C$ and also $\|\psi_j'\|_{\mathrm{Lip}(K)} \le C$ by (43). Clearly $(\psi_0')^\xi = \psi_0^\xi$ is of L^2 norm 1. The claim will follow by induction if we show that $\|(P\varphi)^\xi\|_2 \le 1$ and $\|(S_r\varphi)^\xi\|_2 \le 1$ for all functions $\varphi \in C(S^{d-1})$ that satisfy $\|\varphi^\xi\|_2 \le 1$ for all $\xi \in S^{d-1}$. The first inequality holds since P^ξ is a projection. For the second inequality, we observe

$$(S_r\varphi)^\xi(\sigma) = \int e(r\langle\sigma\xi, v(\gamma)\rangle)\varphi^{\theta_0^{-1}\xi}(\theta(\gamma)^{-1}\sigma\theta_0)d\mu(\gamma). \tag{44}$$

Recall that $\operatorname{supp}\theta(\mu) \subset \theta_0 K$, and θ_0 normalizes K so $\theta(\gamma)^{-1}\sigma\theta_0 \in K$ for $\gamma \in \operatorname{supp}\mu$. The right-hand side of (44) is simply the average of functions of norm 1, so the inequality we need holds.

The proof of (42) will be completed by the following lemma.

Lemma 25. *We have*

$$\|S_r\varphi\|_\infty \le \|\varphi\|_\infty \quad \text{and} \quad \|S_r\varphi\|_{\mathrm{Lip}(K)} \le Cr\|\varphi\|_\infty + \|\varphi\|_{\mathrm{Lip}(K)}$$

for any function $\varphi \in C(S^{d-1})$, where C is a number depending only on μ.

Proof. The first claim is trivial. For the second one, we write

$$\begin{aligned}
&|\rho_r(\gamma)\varphi(\sigma\xi) - \rho_r(\gamma)\varphi(\xi)| \\
&\qquad = |e(r\langle v(\gamma), \sigma\xi\rangle)\varphi(\theta(\gamma^{-1})\sigma\xi) - e(r\langle v(\gamma), \xi\rangle)\varphi(\theta(\gamma^{-1})\xi)| \\
&\qquad \le |e(r\langle v(\gamma), \sigma\xi\rangle) - e(r\langle v(\gamma), \xi\rangle)||\varphi(\theta(\gamma^{-1})\xi)| \\
&\qquad\quad + |\varphi(\theta(\gamma^{-1})\sigma\xi) - \varphi(\theta(\gamma^{-1})\xi)|.
\end{aligned}$$

The first term is estimated by $Cr|v(\gamma)| \cdot |\xi - \sigma\xi| \cdot \|\varphi\|_\infty$. For the second term we observe that $\theta(\gamma^{-1})\sigma\xi = [\theta(\gamma^{-1})\sigma\theta(\gamma)]\theta(\gamma^{-1})\xi$; hence

$$|\varphi(\theta(\gamma^{-1})\sigma\xi) - \varphi(\theta(\gamma^{-1})\xi)| \le \|\varphi\|_{\mathrm{Lip}(K)} \cdot \mathrm{dist}(1, \sigma).$$

We integrate γ and use the moment condition

$$|S_r\varphi(\sigma\xi) - S_r\varphi(\xi)| \le (Cr\|\varphi\|_\infty + \|\varphi\|_{\mathrm{Lip}(K)}) \cdot \mathrm{dist}(1,\sigma).$$

This proves the lemma. □

6. The main theorem

We turn to the proof of the main result of this paper, Theorem 3. As usual, we denote by μ the common law of X_i. By assumption, μ has finite moments of order $\alpha > 2$ and satisfies (SSR). Without loss of generality, we can replace μ by $\mu^{*(l_0)}$ for some fixed integer l_0; hence by Lemmata 6 and 5, we can assume that μ is almost nondegenerate and that K is normalized by θ_0. Furthermore, we assume that μ also satisfies (C) and prove the estimates with $y_0 = 0$. Lemma 4 in the previous section shows that we can reduce the general case of the theorem to this one by changing the coordinate system.

Denote by $\nu_l = \mu^{*(l)}.\delta_{x_0}$ the distribution of the random walk after l steps starting from the point x_0. To evaluate the left-hand sides of the formulae in the statement in Theorem 3, we use Plancherel's formula:

$$\mathbf{E}[f(Y_l)] = \int f(x)d\nu_l = \int \widehat{f}(\xi)\widehat{\nu_l}(\xi)d\xi.$$

We break the latter integral into two regions. First we consider $|\xi| < l^{-1/3}$ and use Proposition 24 in this region.

Recall from Section 2 that $\mathrm{Res}_r : C(\mathbf{R}^d) \to C(S^{d-1})$ is the restriction to the sphere of radius r and that

$$\int_{|\xi|=r} |\varphi(\xi)|^2 d\xi = r^{d-1}\mathrm{Vol}(S^{d-1})\|\mathrm{Res}_r\varphi\|_2^2.$$

(The factor $\mathrm{Vol}(S^{d-1})$ is due to our normalization convention for the L^2 norm on S^{d-1}.) Recall also that

$$\psi_0(\xi) = \psi_{0,r}(\xi) = e(r\langle x_0,\xi\rangle) = \mathrm{Res}_r(\widehat{\delta_{x_0}})(\xi)$$

and

$$\mathrm{Res}_r\widehat{\nu_l} = S_r^l\psi_{0,r}.$$

For $r \le l^{-1/3}$, we write

$$\int_{|\xi|=r} \widehat{f}(\xi)\widehat{\nu_l}(\xi)d\xi = r^{d-1}\int_{S^{d-1}} [\mathrm{Res}_r\widehat{f}](\xi)\cdot[S_r^l\psi_{0,r}](\xi)d\xi$$

$$= \int_{|\xi|=r} \widehat{f}(\xi)e^{-l\Delta(\xi,\xi)}d\xi + r^{d-1}\int_{S^{d-1}} [\mathrm{Res}_r\widehat{f}](\xi)\Psi(\xi)d\xi, \tag{45}$$

where $\Delta(\xi,\xi)$ is the quadratic form that appears in Proposition 24 and $\Psi = S_r^l\psi_0 - e^{-r^2l\Delta}$. By Proposition 24,

$$\|\Psi\|_2 \le C(r^{\min\{1,\alpha-2\}} + |x_0|^2r^2)(e^{-clr^2} + r^{10d}).$$

Using $|\widehat{f}(\xi)| \le \|f\|_1$ and the Cauchy-Schwartz inequality, we can bound the second term in (45) by

$$Cr^{d-1}(r^{\min\{1,\alpha-2\}} + |x_0|^2 r^2)(e^{-clr^2} + r^{10d})\|f\|_1.$$

Integrating for $0 \le r \le l^{-1/3}$, we can write

$$\begin{aligned} &\int_{|\xi| \le l^{-1/3}} \widehat{f}(\xi)\widehat{\nu}_l(\xi)d\xi \\ &\qquad = \int_{\xi \in \mathbf{R}^d} \widehat{f}(\xi)e^{-l\Delta(\xi,\xi)}d\xi + O(l^{-\frac{d+\min\{1,\alpha-2\}}{2}} + |x_0|^2 l^{-\frac{d+2}{2}})\|f\|_1. \end{aligned} \tag{46}$$

It is well known that $e^{-l\Delta(\xi,\xi)}$ is the Fourier transform of a Gaussian measure; i.e., there is a quadratic form Δ' and a constant $C_{\Delta'}$ such that

$$\int \widehat{f}(\xi)e^{-l\Delta(\xi,\xi)}d\xi = C_{\Delta'}l^{-d/2}\int f(x)e^{-\Delta'(x,x)/l}dx.$$

We recognize the first term on the right of (46) as the main term in Theorem 3, while the second one is the first error term. It is also clear that if μ is symmetric or satisfies (E) and we use the improved bounds of Proposition 24, then we get the improved error term claimed in the theorem.

It is left to show that

$$\left|\int_{|\xi| \ge l^{-1/3}} \widehat{f}(\xi)\widehat{\nu}_l(\xi)d\xi\right| \le Ce^{-cl^{1/4}}\|\varphi\|_W^{2,(d+1)/2}, \tag{47}$$

and the proof will be finished.

To this end, we prove a lemma using Proposition 7.

Lemma 26. *Let l be an integer, and suppose that $e^{l^{1/4}} > r > l^{-1/3}$ and $|x_0| < e^{l^{1/4}}$. As above, write $\psi_{0,r}(\xi) = e(r\langle x_0, \xi\rangle)$ for $|\xi| = 1$. There is a constant c depending only on μ such that*

$$\|S_r^l\psi_{0,r}\|_2 \le e^{-cl^{1/4}}.$$

Proof. Choose $1 > c > 0$ to be sufficiently small, to be specified below. Assume to the contrary that the statement is false for some r, l and x_0. Then for each $j \le l$, we have

$$\|S_r^j\psi_{0,r}\|_2 > e^{-l^{1/4}};$$

otherwise we get a contradiction from $\|S_r\| \le 1$. (Recall that $c < 1$.)

To use Proposition 7, we need to estimate the Lipschitz norm of the function $S_r^j\psi_{0,r}$. Using $\|\psi_{0,r}\|_{\mathrm{Lip}} \le Cr|x_0|$ and Lemma 25 repeatedly, we get

$$\|S_r^j\psi_{0,r}\|_{\mathrm{Lip}(K)} \le Cr(j + |x_0|) \le Ce^{2l^{1/4}}$$

for all $j \le l$. Now define $\varphi_j = S_r^j\psi_{0,r}/\|S_r^j\psi_{0,r}\|_2$.

Then we have

$$(r+1)\|\varphi_j\|_{\mathrm{Lip}} < Ce^{4l^{1/4}}.$$

We can apply Proposition 7 for φ_j and get

$$\|S_r\varphi_j\|_2 \le e^{-c'l^{-3/4}},$$

where c' is a number depending only on the constant c from Proposition 7. Note that $r^2 \ge l^{-2/3} \ge l^{-3/4}$.

If we multiply these inequalities together for $1 \le j \le l$, we get

$$\|S_r^l\psi_{0,r}\|_2 \le e^{-c'l^{1/4}},$$

a contradiction if we choose c to be less than c'. □

Similarly as above, we use the Cauchy-Schwartz inequality:

$$\int_{|\xi|=r} \widehat{f}(\xi)\widehat{\nu}_l(\xi)d\xi = r^{d-1}\int_{S^{d-1}} \mathrm{Res}_r\widehat{f}(\xi)\cdot S_r^l\psi_{0,r}(\xi)d\xi$$
$$\le r^{(d-1)/2}\left(\int_{|\xi|=r}|\widehat{f}(\xi)|^2d\xi\right)^{1/2}\cdot \mathrm{Vol}(S^{d-1})^{1/2}\|S_r^l\psi_{0,r}\|_2.$$

We integrate for $r > l^{-1/3}$ and then use the Cauchy-Schwartz inequality again:

$$\int_{|\xi|\ge l^{-1/3}} \widehat{f}(\xi)\widehat{\nu}(\xi)d\xi \le C\int_{l^{-1/3}}^{\infty}\left(r^{d+1}\int_{|\xi|=r}|\widehat{f}(\xi)|^2d\xi\right)^{1/2}\cdot\left(r^{-1}\|S_r^l\psi_{0,r}\|_2\right)dr$$
$$\le C\left(\int_{|\xi|\ge l^{-1/3}}|\widehat{f}(\xi)|^2|\xi|^{d+1}d\xi\right)^{1/2}\cdot\left(\int_{l^{-1/3}}^{\infty}\|S_r^l\psi_{0,r}\|_2^2r^{-2}dr\right)^{1/2}.$$

The first integral on the right-hand side is bounded by $\|f\|_{W^{2,(d+1)/2}}$. By Lemma 26, we have

$$\|S_r^l\psi_{0,r}\|_2^2r^{-1/2} < e^{-cl^{1/4}}$$

for all r. Indeed, one can use the lemma for $r \le e^{-l^{1/4}}$, and simply $\|S_r^l\psi_{0,r}\|_2 \le 1$ for larger r. This implies (47), and Theorem 3 is proved.

7. The Central Limit Theorem

The purpose of this section is to prove Theorem 1. Recall the notation from Sections 1 and 2. We will deduce the theorem from Proposition 15 very similarly to the methods of the previous section.

Notice that the limiting distribution of $Y_l/l^{1/2}$ does not depend on the starting point x_0. Indeed, let Y_l' be the random walk obtained from the same X_l, but from a different point x_0'. Since isometries preserve distance, we have

$$|Y_l/\sqrt{l} - Y_l'/\sqrt{l}| = |x_0 - x_0'|/l^{1/2} \to 0.$$

For the rest of this section, take $x_0 = 0$.

Denote by $W \subset \mathbf{R}^d$ the linear subspace of vectors fixed by K. By Lemma 4, we can choose the origin in such a way that

$$v_0 := \mathbf{E}(Y_1) = \int \gamma(0)d\mu(\gamma) \in W.$$

Define the random isometries X_l' by $X_l'(x) = X_l(x) - v_0$. Since $v_0 \in W$, we have

$$Y_l' := X_l'(X_{l-1}'(\dots(0))) = Y_l - lv_0. \tag{48}$$

Denote by μ' the law of the random isometry X_1'. Then μ' satisfies (C). In what follows, we assume that $\mu = \mu'$, and we prove the theorem with $v_0 = 0$. In light of (48), this implies the theorem without the assumption $\mu = \mu'$ as well.

Let $\nu_l = \mu^{*(l)}.\delta_0$ be the law of Y_l. Denote by $\psi_0 \in L^2(S^{d-1})$ the constant function $\psi_0(\xi) \equiv 1$. Similarly to Section 6, we can write

$$\mathrm{Res}_r\widehat{\nu}_l(\xi) = S_r^l\psi_0(\xi).$$

Fix an arbitrary constant $R > 0$. Let Δ be the quadratic form from Proposition 15. Let λ be the Gaussian measure with Fourier transform

$$\widehat{\lambda}(\xi) = e^{-\Delta(\xi,\xi)}.$$

Proposition 15 implies

$$\|\mathrm{Res}_{r/\sqrt{l}}\widehat{\nu}_l - \mathrm{Res}_r\widehat{\lambda}\|_2 \to 0$$

as $l \to \infty$, uniformly for $r < R$. Let f be a function such that $\widehat{f}(\xi) = 0$ for $|\xi| > R$. Then by Plancherel's formula and the Cauchy-Schwartz inequality,

$$\begin{aligned}\left|\int f(x/\sqrt{l})d\nu_l(x) - \int f(x)d\lambda(x)\right| &= \left|\int \widehat{f}(\xi)(\widehat{\nu}_l(\xi/\sqrt{l}) - \widehat{\lambda}(\xi))d\xi\right| \\ &\le \|f\|_1 \cdot \int_{|\xi|<R} |\widehat{\nu}_l(\xi/\sqrt{l}) - \widehat{\lambda}(\xi)|d\xi \\ &\le \|f\|_1 \cdot C_{R,d}\int_0^R \|\mathrm{Res}_{r/\sqrt{l}}\widehat{\nu}_l - \mathrm{Res}_r\widehat{\lambda}\|_2 dr \to 0,\end{aligned}$$

where $C_{R,d}$ is a constant depending on R and d.

Since we can approximate any continuous function by those that have compactly supported Fourier transform, the proof is complete.

8. The Local Limit Theorem

We finish the paper with the proof of Theorem 2. The proof is again based on Plancherel's formula and estimates on $\widehat{\nu}_l$, the Fourier transform of the law of the random walk. For the frequency range $|\xi| \le l^{-1/2+\varepsilon}$, we again use Proposition 15. However, we do not assume that μ satisfies (SSR), so we need to find a suitable replacement for Proposition 7 in estimating $\widehat{\nu}_l(\xi)$ in the frequency range $l^{-1/2+\varepsilon} \le |\xi| \le R$, where R is an arbitrary fixed number. Our bounds will depend on R in an uncontrolled fashion, so we will be able to conclude the Local Limit Theorem only on scales $O(1)$ in contrast to Theorem 3.

We introduce some notation. Recall that G is the closure of the group generated by $\operatorname{supp} \widetilde{\mu} * \mu$ and $\operatorname{supp} \mu$ is contained in the coset $\gamma_0 G$. We denote by K the closure of $\theta(G)$ and by K° its connected component. By Lemma 5, we can assume that K is normalized by $\theta(\gamma_0)$.

It is easy to see that we can decompose $\mathbf{R}^d$ as an orthogonal sum of subspaces $V_{ss} \oplus V_a \oplus V_o$ such that the action of K° is semi-simple on V_{ss}, Abelian on V_a and trivial on V_o. Since K° is invariant under conjugation by elements of K and $\theta(\gamma_0)$, it follows that these subspaces are invariant under K and $\theta(\gamma_0)$, as well. We denote by S^i the unit sphere in V_i, where i is either ss, a or o. We denote by π_i the orthogonal projection $\mathbf{R}^d \to V_i$. We write $\theta_i(\gamma)$ for the restriction of $\theta(\gamma)$ to V_i for $\gamma \in G$, and we also write $v_i(\gamma) = \pi_i(v(\gamma))$. In addition, by abuse of notation, we write $\pi_i(\gamma)$ for the isometry $x \mapsto v_i(\gamma) + \theta_i(\gamma)x$ on V_i. In addition, we will denote by $\pi_i(\mu)$ the probability measure on $\operatorname{Isom}(V_i)$ that is the pushforward of μ under π_i.

We give an estimate on $\widehat{\nu}_l(\xi)$ in the region $|\pi_{ss}(\xi)| + |\pi_a(\xi)| \ge l^{-1/2} \log^{1/2} l$ in Section 8.1. The methods will be similar to Section 4. We define unitary representations ρ_{r_{ss}, r_a, r_o} of G and consider operators similar to S_r. We show that if a function is almost fixed by such an operator, then it must be almost fixed by $\rho_r(\gamma)$ for pure translations γ pointing in any direction in $V_{ss} \oplus V_a$. The results of Section 4 can be reused to find translations in V_{ss}, and the method of Guivarc'h [15] can be used to produce translations in V_a. The essence of the latter method is taking commutators of isometries with commuting rotation part.

We estimate $\widehat{\nu}_l(\xi)$ in the region $|\pi_{ss}(\xi)| + |\pi_a(\xi)| \le l^{-1/2} \log^{1/2} l$ in Section 8.2. We need to use different methods since it may happen that all pure translations in G are orthogonal to V_o. For example, consider the group generated by a one-parameter family of screw rotations and all translations perpendicular to their axes. If G is this group, then μ will be nondegenerate. So instead of finding translations, we will approximate $\widehat{\nu}_l(\xi)$ by polynomials of a suitably large but fixed degree in the $\pi_{ss}(\xi)$ and $\pi_a(\xi)$ variables using Taylor expansion. This allows us to work with operators on finite-dimensional spaces and use continuity arguments.

We combine the above mentioned estimates to conclude Theorem 2 in Section 8.3.

8.1. *Estimates using translations.* The purpose of this section is to prove the following estimate on the L^2 average of $\widehat{\nu}_l$ on the direct product of spheres in V_{ss}, V_a and V_o.

Proposition 27. *With the assumptions of Theorem 2 and the notation explained at the beginning of Section 8 the following holds. There are constants $C, c > 0$ depending only on μ and R such that*

$$r_{ss}^{1-\dim V_{ss}} r_a^{1-\dim V_a} r_o^{1-\dim V_o} \int_{|\pi_{ss}(\xi)|=r_{ss}, |\pi_a(\xi)|=r_a, |\pi_o(\xi)|=r_o} |\widehat{\nu}_l(\xi)|^2 d\xi \leq C e^{-c\min\{(r_{ss}+r_a)^2 l, l^{1/4}\}}$$

holds for all $0 < r_{ss}, r_a, r_o < R$.

We fix three nonnegative real parameters r_{ss}, r_a and r_o. Analogously to ρ_r, we introduce the unitary representation ρ_{r_{ss},r_a,r_o} of the group generated by G and γ_0 on the space $L^2(S^{ss} \times S^a \times S^o)$ via the following formula:

$$\begin{aligned}\rho_{r_{ss},r_a,r_o}(\gamma)\varphi(\xi_{ss}, \xi_a, \xi_o) = e(&r_{ss}\langle \xi_{ss}, v_{ss}(\gamma)\rangle + r_a\langle \xi_a, v_a(\gamma)\rangle + r_o\langle \xi_o, v_o(\gamma)\rangle) \\ &\cdot \varphi(\theta_{ss}(\gamma)^{-1}\xi_{ss}, \theta_a(\gamma)^{-1}\xi_a, \theta_o(\gamma)^{-1}\xi_o).\end{aligned}$$

Here $\varphi \in L^2(S^{ss} \times S^a \times S^o)$, $\xi_{ss} \in S^{ss}$, $\xi_a \in S^a$ and $\xi_o \in S^o$. This representation corresponds to the action of γ on the Fourier transform of a measure restricted to the product of the spheres of radii r_{ss}, r_a, r_o (resp.) in V_{ss}, V_a, V_o (resp.). For a measure η on the group generated by G and γ_0, we define the operators

$$\rho_{r_{ss},r_a,r_o}(\eta) = \int \rho_{r_{ss},r_a,r_o}(\gamma) d\eta(\gamma) \tag{49}$$

acting on $L^2(S^{ss} \times S^a \times S^o)$. These are analogues of S_r, and similarly as in Proposition 7, we prove the following estimate for them.

Proposition 28. *Suppose that μ is almost nondegenerate and has finite moments of order 2. Then for every $R > 0$, there is a constant $c > 0$ depending only on μ and R such that the following hold.*

Let $R \geq r_{ss}, r_a, r_o \geq 0$. Let $\varphi \in \mathrm{Lip}(S^{ss} \times S^a \times S^o)$ with $\|\varphi\|_2 = 1$. Then

$$\|\rho_{r_{ss},r_a,r_o}(\mu)\varphi\|_2 \leq 1 - c\min\{r_{ss}^2 + r_a^2, \log^{-3}(\|\varphi\|_{\mathrm{Lip}} + 2)\}.$$

Proposition 27 can be deduced from Proposition 28 exactly the same way as the proof of Lemma 26. The rest of the section is devoted to the proof of Proposition 28. We fix r_{ss}, r_a, r_o and write ρ instead of ρ_{r_{ss},r_a,r_o}, saving a considerable amount of ink. The hypothesis of Proposition 28 on μ is assumed throughout the section.

Our first goal is to replace μ with a symmetric measure μ_1 such that $\operatorname{supp}\theta(\mu_1) \subset K^\circ$.

Lemma 29. *We can write $(\widetilde{\mu} * \mu)^{*(L)} = p\mu_1 + q\mu_2$, with $1 \geq p > 0$, where μ_1 and μ_2 are probability measures on $\mathrm{Isom}(\mathbf{R}^d)$ and $L \geq 1$ is an integer depending on μ. Furthermore, μ_1 is almost nondegenerate, symmetric, has* *finite moments of order* 2*, and the closure of the group generated by* $\operatorname{supp}\theta(\mu_1)$ *is* K°*.*

Proof. We fix an integer L and write

$$G^\circ = \{\gamma \in G : \theta(\gamma) \in K^\circ\},$$

let $p = (\widetilde{\mu} * \mu)^{*(L)}(G^\circ)$ and let μ_1 be $1/p$ times the restriction of $(\widetilde{\mu} * \mu)^{*(L)}$ to G°. The only nontrivial property to check is that almost nondegeneracy holds if L is sufficiently large. It is enough to check for an arbitrary point x the condition that the points $\gamma(x)$ for $\gamma \in \operatorname{supp}\mu_1$ do not lie in a proper affine subspace if L is sufficiently large possibly depending on x. Then the claim follows from the same Noetherian property argument as in Lemma 6.

Denote by o the order of K/K°. Using the Central Limit Theorem for the measure $\widetilde{\mu} * \mu$, we can find an integer L_0 and a finite set

$$A \subset \{\gamma(x) : \gamma \in \operatorname{supp}(\widetilde{\mu} * \mu)^{*(L_0)}\}$$

that approximates an $(o+1) \times \cdots \times (o+1)$ grid. The approximation can be arbitrarily good if L_0 is sufficiently large. All that we need is that a proper affine subspace intersects A in at most $|A|/(o+1)$ points.

Then by the pigeon hole principle, there is $\theta_1 \in K$ such that

$$B := \{\gamma(x) : \gamma \in \operatorname{supp}(\widetilde{\mu} * \mu)^{*(L_0)}, \theta(\gamma) \in \theta_1 K^\circ\}$$

is not contained in a proper affine subspace. Now the claim follows for $L = 2L_0$. Indeed, take any $\gamma_1 \in \operatorname{supp}(\widetilde{\mu} * \mu)^{*(L_0)}$ with $\theta(\gamma_1) \in \theta_1 K^\circ$ and observe that $\gamma_1^{-1}(B)$ is in the set of images of x under elements of $\operatorname{supp}(\mu_1)$. □

For the rest of the proof we work with μ_1, and assume that it satisfies the properties claimed in Lemma 29. Moreover, we assume that μ_1 has property (C), which is justified by Lemma 4 after changing the origin. Then we also need to multiply the function φ appearing in Proposition 28 with a character, possibly increasing its Lipschitz norm by a factor depending on μ and R. (Compare with the discussion on page 256 in Section 4.) We set out to prove an inequality analogous to the one claimed in Proposition 28 for the operator $\rho(\mu_1)$.

We fix $\varphi \in C^1(S^{ss} \times S^a \times S^o)$. As in Section 5, the heart of the proof is the study of the set

$$B(\varepsilon) := \{\gamma \in \operatorname{Isom}(\mathbf{R}^d) : \|\rho(\gamma)\varphi - \varphi\|_2 < \varepsilon\}.$$

The next two lemmata are obtained by a simple variation on the arguments in Section 5.

Lemma 30. *Let* $\varepsilon > 0$*, and let* $l_1 = C_1(r_{ss}^{-2} + \log^3(\|\varphi\|_{\mathrm{Lip}} + 2))$*, where* C_1 *is a suitably large constant depending on* μ *and* ε*. Suppose that* $\mu_1^{*(l_1)}(B(\varepsilon)) >$

9/10. *Then there is a set* $X \subset B(64\varepsilon)$ *such that*

$$\theta_{ss}(X) = \pi_{ss}(K^\circ) \quad \text{and} \quad \pi_a(X) = \{1\} = \pi_o(X).$$

Proof. Following the proof of Lemma 11, it is easy to find a subset $X_0 \subset B(4\varepsilon)$ such that $\pi_{ss}(\theta(X_0)) = \pi_{ss}(K^\circ)$. Consider $X_1 = [X_0, X_0]$ and $X = [X_1, X_1]$. Clearly $\pi_o(X_1) = \{1\}$, and $\pi_a(X_1)$ consists of translations. Therefore $\pi_a(X) = \{1\} = \pi_o(X)$, and $X \subset B(64\varepsilon)$ follows from the triangle inequality. Since every element is a commutator in a connected semi-simple compact Lie group (see [13]), we have $\pi_{ss}(X) = \pi_{ss}(K^\circ)$, which finishes the proof. □

Lemma 31. *Let* $\varepsilon > 0$ *be arbitrary and* l_1, *and let* C_1 *be as in the previous lemma. Let* $l_2 = C_2(r_{ss}^{-2} + \log^3(\|\varphi\|_{\mathrm{Lip}} + 2))$, *where* C_2 *is a suitably large constant depending on* μ_1 *and* C_1. *Suppose that* $\mu_1^{*(l_i)}(B(\varepsilon)) > 9/10$ *for* $i = 1, 2$. *Then there is a constant* $c > 0$ *depending on* μ_1 *such that the following hold. For any unit vector* $u_0 \in S^{ss}$, *there is an element* $\gamma_1' \in B(386\varepsilon)$ *such that*

$$\theta_{ss}(\gamma_1') = 1, \quad |v_{ss}(\gamma_1')| < r_{ss}^{-1}/2,$$
$$\langle v_{ss}(\gamma_1'), u_0\rangle > c r_{ss}^{-1}, \quad \pi_a(\gamma_1') = 1 = \pi_0(\gamma_1').$$

Proof. Consider the projection to V_{ss}, and repeat the argument in Lemmata 12–14, except instead of the set X constructed in Lemma 11, use the one constructed in Lemma 30. We need to use six elements of X, and since now they are in $B(64\varepsilon)$ instead of $B(4\varepsilon)$, the resulting element γ_1' will be in $B(386\varepsilon)$. Recall from the proof of Lemma 14 that γ_1' is of the form $g_1\gamma_1 g_1^{-1} g_2 \gamma_1^{-1} g_2^{-1}$, where $g_1, g_2 \in X$. Since $\pi_a(g_1) = \pi_a(g_2) = 1$, we have $\pi_a(\gamma_1') = \pi_a(\gamma_1)\pi_a(\gamma_1^{-1}) = 1$, and a similar calculation applies to the projection to V_o. This finishes the proof. □

In the above lemma we constructed a translation in V_{ss}. The next goal will be to construct a translation in V_a. This is done in the next two lemmata by adapting the method of Guivarc′h [15]. Denote by G_1 the closure of the group generated by supp(μ_1).

Lemma 32. $\pi_a([G_1, G_1])$ *is the additive group of the vector space* V_a.

Here $[G_1, G_1]$ denotes the derived subgroup of G_1, not just the set of commutators.

Proof. Clearly $H := \pi_a([G_1, G_1])$ is a subgroup of the additive group of V_a, and it is invariant under the action of $\pi_a(K^\circ)$. Since $\pi_a(K^\circ)$ is connected, every connected component of H is invariant under $\pi_a(K^\circ)$. Every such component is an affine subspace of V_a. The point of such an affine subspace that is closest to the origin is a fixed point of $\pi_a(K^\circ)$. By the definition of V_a, the only fixed point is the origin. Therefore it follows that H is a linear subspace of V_a invariant under the action of $\pi_a(K^\circ)$.

Assume to the contrary that H is a proper subspace of V_a. Let W be a two-dimensional subspace of V_a that is invariant under $\pi_a(K)$ and orthogonal to H. By projecting the translation part to W, G_1 naturally embeds to $\mathrm{Isom}(W)$; denote by G_W the image. Then G_W is commutative (since it has a trivial commutator) but has a nontrivial rotation part (since $\pi_a(K^\circ)$ acts on W nontrivially); hence it consists of rotations around the same point $x \in W$. This means that μ_1-almost every image of x is orthogonal to W, a contradiction to almost nondegeneracy. □

LEMMA 33. *Let ε, l_1 and C_1 be as in Lemma 30. Suppose that $\mu_1^{*(l_1)}(B(\varepsilon)) > 9/10$. Then for every $u_0 \in V_a$, there are $c > 0$, $v \in V_a$ with $|v - u_0| < |u_0|/10$ and an integer L such that the following holds. Let M be an arbitrary positive integer, and assume that $\mu_1^{*(2)}(B(\varepsilon/M)) > 1 - c$. Then there is $\gamma_1' \in B(L\varepsilon)$ such that*

$$v_a(\gamma_1') = Mv, \quad \theta_a(\gamma_1') = 1, \quad \pi_{ss}(\gamma_1') = 1 = \pi_o(\gamma_1').$$

The vector v may depend on φ, but c and L depend only on μ and u_0.

This lemma allows us to find pure translations in $B(L\varepsilon)$ approximating an arbitrary direction in V_a. This (or Lemma 31) leads to contradiction if we set the parameters in such a manner that $M|u_0| \approx 1/r_a$ and $L\varepsilon$ is sufficiently small, so one of the assumptions of Lemma 33 must fail. We can derive the claim of Proposition 28 from either $\mu_1^{*(l_1)}(B(\varepsilon)) < 9/10$ or $\mu_1^{*(2)}(B(\varepsilon/M)) < 1 - c$. In the second case, for example, we can get $\|\rho(\mu_1)\varphi\|_2 \le 1 - c\varepsilon^2/M^2$. This will imply the claim if we set $M \approx \max\{1, r_a^{-1}\}$.

Observe that the numbers c and L depend on u_0 in an uncontrolled way. Hence it is important to note that we will apply the lemma with choosing u_0 from a fixed finite collection depending on the parameter R.

Proof. There is $\gamma_2 \in G_1$ such that $\theta_a(\gamma_2)$ does not have any fixed vectors in V_a except for 0. This is an open condition, so we can assume that $\gamma_2 \in \operatorname{supp} \mu_1^{*(m)}$ for some integer m depending on μ. Thus $\gamma_2 = g_1 \cdots g_m$ for some $g_i \in \operatorname{supp} \mu_1$.

There is a vector $u_1 \in V_a$ such that $u_0 = u_1 - \theta_a(\gamma_2)u_1$. (Since $\theta_a(\gamma_2)$ has no fixed vectors, $1 - \theta_a(\gamma_2)$ has trivial kernel.) We can find a small ball U_i around each g_i such that $|u_1 - \theta_a(g_1' \cdots g_m')u_1 - u_0| < |u_0|/20$ for any choice of $g_i' \in U_i$. We set the constant c in the lemma so that $\mu_1^{*(2)}(U_i) > c$ for all i. This allows us to find an element

$$\gamma_2' = g_1' \cdots g_m' \in B(m\varepsilon/M)$$

such that $|u_1 - \theta_a(\gamma_2')u_1 - u_0| < |u_0|/20$.

For reasons that will be clear at the end of the proof, we now search for an element $\gamma_3' \in B(L\varepsilon/M)$ such that $v_a(\gamma_3')$ approximates u_1 instead of u_0, which is the objective in the lemma. By Lemma 32, we have elements $\gamma_4, \gamma_5 \in G_1$ such that $u_1 = \pi_a([\gamma_4, \gamma_5])$. Using an argument very similar to the one above, we can approximate γ_4 and γ_5 by elements in $B(m\varepsilon/M)$ (by taking m larger and c smaller perhaps). Then we can find a vector v_1 and an element $\gamma_3' \in B(4m\varepsilon/M)$ such that $v_a(\gamma_3') = v_1$, $|v_1 - u_1| < |u_0|/40$ and $\theta_a(\gamma_3') = 1$.

Now we make use of the set X constructed in Lemma 30 to cancel the rotation parts of γ_2' and γ_3' in the V_{ss} component. Let $h_2, h_3 \in X$ be such that $\theta_{ss}(h_2) = \theta_{ss}(\gamma_2')^{-1}$ and $\theta_{ss}(h_3) = \theta_{ss}(\gamma_3')^{-M}$. Then $h_2 \cdot \gamma_2'$ and $h_3 \cdot (\gamma_3')^M$ act on $V_{ss} \oplus V_o$ by translation; hence

$$\gamma_1' := [h_3 \cdot (\gamma_3')^M, h_2\gamma_2']$$

acts trivially on V_{ss} and V_o. On the other hand, an easy calculation shows that

$$\pi_a(\gamma_1') = (M(v_1 - \theta_a(\gamma_2')v_1), 1)$$

and

$$|v_1 - \theta_a(\gamma_2')v_1 - u_0| \le |u_1 - \theta_a(\gamma_2')u_1 - u_0| + 2|u_0|/40 \le |u_0|/10$$

and $\gamma_1' \in B((10m + 256)\varepsilon)$, which was to be proved. □

Proof of Proposition 28. Without any significant changes to the argument in the proof of Proposition 7, we can deduce from Lemma 31 the estimate

$$\|\rho(\mu_1)\varphi\|_2 \le 1 - c' \min\{r_{ss}^2, \log^{-3}(\|\varphi\|_{\mathrm{Lip}} + 2)\}. \tag{50}$$

We suppress the details but carry out a similar argument that proves

$$\|\rho(\mu_1)\varphi\|_2 \le 1 - c' \min\{r_a^2, \log^{-3}(\|\varphi\|_{\mathrm{Lip}} + 2)\}. \tag{51}$$

There is a unit vector $u \in S^a$ such that

$$\int_{\xi_{ss} \in S^{ss}, \xi_o \in S^o, |\xi_a - u| < 1/10} |\varphi(r_{ss}\xi_{ss}, r_a\xi_a, r_o\xi_o)|^2 d\xi_{ss} d\xi_a d\xi_o > \varepsilon_0^2 \tag{52}$$

for a constant ε_0 that depends only on the dimension of V_a. Moreover, we can choose u from a fixed finite sufficiently dense subset of S^a. If $r_a < 1$, let $M = \lfloor 10 r_a^{-1} \rfloor$, and let $M = 1$ otherwise. If $r_a < 1$, then let $u_0 = u/20$, otherwise let $u_0 = 5u/\lceil 10 r_a \rceil$. Let C_1, c and L be the constants from Lemma 33 with this choice of u_0. (Note that the possible values for u_0 are in a finite set that depends only on the dimension of V_a and R.) Set $\varepsilon = \varepsilon_0/L$.

Assume to the contrary that $\mu_1^{*(l_1)}(B(\varepsilon)) > 9/10$ and $\mu_1^{*(2)}(B(\varepsilon/M)) > 1 - c$. Then we can apply Lemma 33. Let $v \in V_a$ and γ_1' be as in the lemma.

Then

$$\varepsilon_0^2 = L^2\varepsilon^2 \geq \|\rho(\gamma_1')\varphi - \varphi\|_2^2$$
$$\geq \int_{\xi_{ss}\in S^{ss},\xi_o\in S^o,|\xi_a-u|<1/10} |1-e(\langle Mv, r_a\xi_a\rangle)|^2|\varphi(r_{ss}\xi_{ss}, r_a\xi_a, r_o\xi_o)|^2 d\xi_{ss}d\xi_a d\xi_o.$$

With the above definitions, Mu_0 and Mv approximate $r_a^{-1}u/2$. Hence $e(\langle Mv, \xi_a\rangle)$ is close to -1; in particular, $|1 - e(\langle Mv, \xi_a\rangle)| > 1$ in the domain of integration. This is in contradiction to (52). (This somewhat vague discussion can be made precise by a straightforward calculation.)

Now there are two possibilities. Either $\mu_1^{*(l)}(B(\varepsilon)) \leq 9/10$, which implies (51) as we have seen in the proof of Proposition 7, or else $\mu_1^{*(2)}(B(\varepsilon/M)) \leq 1-c$. If $\gamma_1^{-1}\cdot\gamma_2 \notin B(\varepsilon/M)$, then

$$\mathrm{Re}(\langle\rho(\gamma_2)\varphi, \rho(\gamma_1)\varphi\rangle) = \mathrm{Re}(\langle\rho(\gamma_1^{-1}\cdot\gamma_2)\varphi, \varphi\rangle) \leq 1 - \varepsilon^2/2M^2.$$

Hence (51) follows:

$$\|\rho(\mu_1)\varphi\|_2^2 = \int \mathrm{Re}(\langle\rho(\gamma_2)\varphi, \rho(\gamma_1)\varphi\rangle)d\mu_1(\gamma_1)d\mu_1(\gamma_2) \leq 1 - c\varepsilon^2/2M^2.$$

Note that $1/M \geq \min\{r_a, 1\}$.

If p and q are as in Lemma 29 and L' is the number L from that lemma, then we can conclude from (50) and (51) that

$$\|(\rho(\mu)^*\rho(\mu))^{L'}\varphi\|_2 = \|p\rho(\mu_1)\varphi + q\rho(\mu_2)\varphi\|_2$$
$$\leq 1 - pc'\min\{r_{ss}^2 + r_a^2, \log^{-3}(\|\varphi\|_{\mathrm{Lip}} + 2)\},$$

which in turn implies the proposition. □

8.2. *Estimates using continuity arguments.* We continue to use the notation, V_{ss}, V_a, V_o, etc. introduced in the beginning of Section 8. Our goal is to prove the following estimate, which complements the results of the previous section.

PROPOSITION 34. *Assume that μ is nondegenerate and has finite moments of order α for some $\alpha \geq 2$, and let $R > 0$ be a number. Then there is a number C depending on μ, x_0, R and α such that the following holds. Let $0 \leq r_{ss}, r_a \leq R$ be numbers, l a positive integer and $0 \leq s, \delta \leq 1$ numbers such that $l > C\log(s^{-1})\delta^{-2}$, $s > r_{ss} + r_a$ and $C^{-1} \geq \delta \geq C(r_{ss} + r_a)$. Then*

$$r_{ss}^{1-\dim V_{ss}} r_a^{1-\dim V_a} \int_{\pi_{ss}(\xi)=r_{ss},\pi_a(\xi)=r_a,\delta\leq|\pi_o(\xi)|\leq R} |\widehat{\nu}_l(\xi)|d\xi$$
$$< C\log(s^{-1})^{1/2}s\delta^{\dim V_o} + C\log(s^{-1})^{\alpha/2}s^\alpha\delta^{-\alpha-2}.$$

We indicate the approximate values of the parameters that we will set in the next section. We take $s \approx (\log^{1/2} l)l^{-1/2}$ and $\delta \approx l^{-\beta}$, where β is slightly smaller than 1/2.

To outline the idea of the proof, we temporarily assume that $\theta_o(G)$ is trivial and $\gamma_0 = 1$. (We will reduce the problem to this situation by defining a measure μ_1 similar to the one we had in the previous section.) We can restrict the action of G on Fourier space to sets of the form $\{\xi : |\pi_{ss}(\xi)| = r_{ss}, |\pi_a(\xi)| = r_a, \pi_o(\xi) = \xi_o\}$. This gives rise to a unitary representation

$$\rho_{r_{ss},r_a,\xi_o}(\gamma)\varphi(\xi_{ss}, \xi_a) = e(\langle r_{ss}\xi_{ss} + r_a\xi_a + \xi_o, v(\gamma)\rangle)\varphi(\theta_{ss}^{-1}(\gamma)\xi_{ss}, \theta_a^{-1}(\gamma)\xi_a)$$

of G for each $r_{ss}, r_a \geq 0$ and $\xi_o \in V_o$ acting on the space $L^2(S^{ss} \times S^a)$.

We will study the operators $\rho_{r_{ss},r_a,\xi_o}(\mu)$ defined analogously to (49). We consider the finite-dimensional subspace $\mathcal{P}_{\alpha-1} \subset L^2(S^{ss} \times S^a)$, which we define as the restriction of polynomials of degree at most $\alpha - 1$ to $S^{ss} \times S^a$. This space is invariant for $\rho_{0,0,\xi_o}(\mu)$, and we will show that there are only finitely many "bad" points in the ball $\{\xi_o : |\xi_o| \leq R\}$ such that $\|\rho_{0,0,\xi_o}(\mu)|_{\mathcal{P}_\alpha}\| = 1$. We will also understand the behavior of the function $\|\rho_{0,0,\xi_o}(\mu)|_{\mathcal{P}_\alpha}\|$ in small neighborhoods of those "bad" points. We then combine this with a continuity argument (essentially using that the above norm function is continuous and attains its extrema on compact sets) to obtain bounds for $\|\rho_{0,0,\xi_o}(\mu)|_{\mathcal{P}_\alpha}\|$ for ξ_o not too close to the "bad" points.

We also show that $\rho_{r_{ss},r_a,\xi_o}(\mu)$ is a small perturbation of $\rho_{0,0,\xi_o}(\mu)$ and the norm bounds are valid for the former operator as well. Then we show that we can approximate $\widehat{\nu}_l$ by polynomials of degree $\alpha - 1$ in the $\pi_{ss}(\xi) + \pi_a(\xi)$ coordinates, and using the norm bounds of $\rho_{r_{ss},r_a,\xi_o}(\mu)$ iteratively, we get the desired bound on $\widehat{\nu}_l$.

We need to give a separate argument in the neighborhood of "bad" points. The bounds in this case will be substantially weaker. We will show that the only "bad" point for $\alpha = 0$ is the origin, and we can do a similar argument as above.

We note that there are examples when "bad" points do occur. We recommend to the reader to analyze the instructive example mentioned earlier: when G is generated by a one-parameter family of skew rotations and all translations perpendicular to the axes.

Some of the above ideas are related to the arguments of Section 5 and hence motivated by Tutubalin's paper [23].

We give a lemma similar to Lemma 29, which introduces the measure μ_1 mentioned above.

Lemma 35. *Let μ be as in Proposition 34. Then we can write $\widetilde{\mu}^{*(L)} * \mu^{*(L)} = p\mu_1 + q\mu_2$, with $1 \geq p > 0$, where μ_1 and μ_2 are probability measures on* $\mathrm{Isom}(\mathbf{R}^d)$ *and $L \geq 1$ is an integer depending on μ, R and x_0. In addition,* *the set*

$$\{v_o(\gamma) : \gamma \in \operatorname{supp} \mu_1, |v_o(\gamma)| < 1/(2R)\} \tag{53}$$

is not contained in a proper affine subspace of V_o. *Furthermore,* μ_1 *is symmetric, has finite moments of order* α, $1 \in \operatorname{supp} \mu_1$ *and the closure of the group generated by* $\operatorname{supp} \theta(\mu_1)$ *is* K°.

Proof. The proof is very similar to that of Lemma 29. The main difference is that we use nondegeneracy instead of the Central Limit Theorem. We fix a sufficiently large integer L. We write

$$G^\circ = \{\gamma \in G : \theta(\gamma) \in K^\circ\}.$$

Let $p = \widetilde{\mu}^{*(L)} * \mu^{*(L)}(G^\circ)$, and let μ_1 be $1/p$ times the restriction of $\widetilde{\mu}^{*(L)} * \mu^{*(L)}$ to G° The only nontrivial property to check is that (53) is not contained in a proper affine subspace if L is sufficiently large.

Denote by o the order of K/K°. Fix an arbitrary $\gamma_0 \in \operatorname{supp} \mu$, and let x_0 be the starting point of the random walk. We show that if L is sufficiently large, then we can find a set

$$A \subset \{\pi_o(\gamma(x_0)) : \gamma \in \operatorname{supp} \mu^{*(L)}\}$$

that approximates an $(o+1) \times \cdots \times (o+1)$ grid contained in a $1/(4R)$ neighborhood of $\pi_o(\gamma_0^L(x_0))$. The approximation can be arbitrarily good, and what we need below is that no proper affine subspace contains more than $|A|/(o+1)$ points of A.

To this end, we consider an $(o+1) \times \cdots \times (o+1)$ grid A' contained in the $1/(4R)$ neighborhood of $\pi_o(x_0)$, and for each point $x \in A'$, let D_x be the complement of a small open neighborhood of x. By nondegeneracy, we know that if L is sufficiently large, then for any $x \in A'$, we have

$$\{\pi_o(\gamma(x_0)) : \gamma \in \operatorname{supp} \mu^{*(L)}\} \not\subset \pi_o(\gamma_0^L)(D_x).$$

This implies the claim on the existence of the set A.

By the pigeon hole principle, there is $\theta_1 \in K$ such that

$$B := \{\pi_o(\gamma(x_0)) : \gamma \in \operatorname{supp} \mu^{*(L)}, \theta(\gamma) \in \theta(\gamma_0)\theta_1 K^\circ\} \cap A$$

is not contained in a proper affine subspace. We choose an arbitrary element $\gamma_1 \in \operatorname{supp} \mu^{*(L)}$ with $\theta(\gamma_1) \in \theta(\gamma_0)\theta_1 K^\circ$ and $\pi_o(\gamma_1(x_0)) \in B$. We observe that

$$\pi_o(\gamma_1^{-1})(B) \subset \{\pi_o(\gamma(x_0)) : \gamma \in \operatorname{supp} \mu_1\}.$$

We also note that the set $\pi_o(\gamma_1^{-1})(B)$ contains x_0 by construction and its diameter is at most $1/(2R)$. Since $\pi_o(\gamma)$ is a translation for all $\gamma \in \operatorname{supp} \mu_1$, we have $v_o(\gamma) = \pi_o(\gamma(x_0)) - x_0$ for those γ. Therefore $(\pi_o(\gamma_1^{-1})(B) - x_0) \subset$ (53), which is not contained in a proper affine subspace. This finishes the proof. □

We start the program described above by giving a few lemmata on the properties of the operators $\rho_{r_{ss},r_a,\xi_o}(\mu_1)$. Throughout the section, we assume that μ_1 satisfies the properties stated in Lemma 35 and, in addition, that it satisfies (C). By changing the origin, $\widehat{\nu_l}$ gets multiplied by a character, and this does not change the statement of Proposition 34. Hence the assumption (C) is justified by Lemma 4. We first show that $\rho_{r_{ss},r_a,\xi_o}(\mu_1)$ is a small perturbation of $\rho_{0,0,\xi_o}(\mu_1)$.

LEMMA 36. *There is a constant C depending only on μ_1 such that*

$$\|\rho_{r_{ss},r_a,\xi_o}(\mu_1)-\rho_{0,0,\xi_o}(\mu_1)\|<C(r_{ss}+r_a).$$

Proof. Let $\varphi\in L^2(S^{ss}\times S^a)$. Then by Taylor's theorem,

$$\begin{aligned}\rho_{r_{ss},r_a,\xi_o}(\mu_1)\varphi(\xi_{ss},\xi_a)=\int&(1+O(r_{ss}|v_{ss}(\gamma)|+r_a|v_a(\gamma)|))e(\langle\xi_o,v_o(\gamma)\rangle)\\&\cdot\varphi(\theta_{ss}(\gamma)^{-1}\xi_{ss},\theta_a(\gamma)^{-1}\xi_a)d\mu_1(\gamma).\end{aligned}$$

Then

$$\|\rho_{r_{ss},r_a,\xi_o}(\mu_1)\varphi-\rho_{0,0,\xi_o}(\mu_1)\varphi\|_2\le C\int(r_{ss}|v_{ss}(\gamma)|+r_a|v_a(\gamma)|)\|\varphi\|_2d\mu_1(\gamma),$$

which proves the claim. □

The next lemma is about the behavior of $\rho_{r_{ss},r_a,\xi_o'}(\mu_1)$ in a neighborhood of a "bad" point.

LEMMA 37. *There are constants c and C that depend only on μ_1 and R such that the following holds. Suppose that $\varphi\in L^2(S^{ss}\times S^a)$ and $|\xi_o|\le R$ are such that $\rho_{0,0,\xi_o}(\mu_1)\varphi=\varphi$. Then*

$$\|\rho_{r_{ss},r_a,\xi_o'}(\mu_1)\varphi\|_2<1-c|\xi_o-\xi_o'|^2+C(r_{ss}^2+r_a^2)$$

for every $r_{ss},r_a\ge 0$ and $\xi_o'\in V_o$ with $|\xi_o'|<R$.

Proof. Since $\rho_{0,0,\xi_o}(\mu_1)$ is an average of unitary operators, we must have $\rho_{0,0,\xi_o}(\gamma)\varphi=\varphi$ for all $\gamma\in\operatorname{supp}(\mu_1)$. Then

$$\begin{aligned}\rho_{r_{ss},r_a,\xi_o'}(\mu_1)\varphi&=\int\rho_{r_{ss},r_a,\xi_o'}(\gamma)\rho_{0,0,\xi_o}(\gamma^{-1})\varphi d\mu_1(\gamma)\\&=\varphi\cdot\int e(r_{ss}\langle\xi_{ss},v_{ss}(\gamma)\rangle+r_a\langle\xi_a,v_a(\gamma)\rangle+\langle\xi_o'-\xi_o,v_o(\gamma)\rangle)d\mu_1(\gamma)\\&=\varphi\cdot[\int(1-2\pi i r_{ss}\langle\xi_{ss},v_{ss}(\gamma)\rangle-2\pi i r_a\langle\xi_a,v_a(\gamma)\rangle)\\&\qquad\cdot e(\langle\xi_o'-\xi_o,v_o(\gamma)\rangle)d\mu_1(\gamma)+O(r_a^2+r_{ss}^2)].\end{aligned}\tag{54}$$

For every positive c_0, we can find C' such that the following estimate holds for the linear term in (54):

(55)

$$\begin{aligned}&\left|\int(2\pi i r_{ss}\langle\xi_{ss}, v_{ss}(\gamma)\rangle + 2\pi i r_a\langle\xi_a, v_a(\gamma)\rangle)e(\langle\xi_o' - \xi_o, v_o(\gamma)\rangle)d\mu_1(\gamma)\right|\\ &\le\left|\int 2\pi i r_{ss}\langle\xi_{ss}, v_{ss}(\gamma)\rangle + 2\pi i r_a\langle\xi_a, v_a(\gamma)\rangle d\mu_1(\gamma)\right| + C(r_{ss}+r_a)|\xi_o' - \xi_o|\\ &\le C'(r_{ss}+r_a)^2 + c_0|\xi_o' - \xi_o|^2.\end{aligned}$$

For the second inequality, we used (C) to show that the first term vanishes, and we used the inequality between the geometric and arithmetic means to estimate the second term.

Consider the function

$$\Phi(\xi) = \int e(\langle\xi, v_o(\gamma)\rangle)d\mu_1(\gamma)$$

on V_o. Note that Φ depends only on μ_1. Combining (54) and (55) we get

$$\|\rho_{r_{ss},r_a,\xi_o'}(\mu_1)\varphi\|_2 \le \Phi(\xi_o' - \xi_o)\cdot\varphi + C'(r_{ss}^2 + r_a^2) + c_0|\xi_o' - \xi_o|^2.$$

If $\Phi(\xi) = 1$ for some $\xi \neq 0$, then $\langle\xi, v_o(\gamma)\rangle$ is an integer for all $\gamma \in \operatorname{supp}(\mu_1)$. This is impossible because it contradicts the property that (53) is not contained in a proper affine subspace. Using Taylor series expansion, (C) and the moment condition, we can show that $\Phi(\xi) \le 1 - c_1|\xi|^2$ for some $c_1 > 0$ and $|\xi| < R$. These estimates prove the lemma if we set $c_0 < c_1$. □

Let $X_\alpha \subset V_o$ be the set of those ξ_o for which there is $\varphi \in \mathcal{P}_{\alpha-1}$ such that $\rho_{0,0,\xi_o}(\mu_1)\varphi = \varphi$. This is the set whose elements we called "bad" points above. We note that $\|\rho_{0,0,\xi_o}(\mu_1)\varphi\|_2 = \|\varphi\|_2$ implies $\rho_{0,0,\xi_o}(\mu_1)\varphi = \varphi$ since $1 \in \operatorname{supp}\mu_1$. If $\rho_{0,0,\xi_o}(\mu_1)\varphi = \varphi$ and $\rho_{0,0,\xi_o'}(\mu_1)\varphi' = \varphi'$ for $\xi_o \neq \xi_o'$, then φ and φ' are both eigenfunctions of $\rho_{0,0,\xi_o}(\mu_1)$ with different eigenvalues; hence they are orthogonal. Indeed, $\rho_{0,0,\xi_o}(\mu_1)\varphi' = \varphi'\cdot\int e(\langle\xi_o - \xi_o', v_o(\gamma)\rangle)d\mu_1(\gamma)$, as the previous proof shows. Since $\mathcal{P}_{\alpha-1}$ is finite dimensional, X_α is finite.

We now combine Lemma 37 with a continuity argument to obtain norm estimates for ρ_{r_{ss},r_a,ξ_o} on $\mathcal{P}_{\alpha-1}$.

Lemma 38. *There are numbers $c, C > 0$ depending only on μ_1, R and α such that the following holds. Let r_{ss}, r_a be numbers and $\xi_o \in V_o$ such that $|\xi_o| \le R$ and $\operatorname{dist}(\xi_o, X_\alpha) > C(r_{ss}+r_a)$. Then*

$$\|\rho_{r_{ss},r_a,\xi_o}(\mu_1)\varphi\|_2 \le (1 - c\operatorname{dist}(\xi_o, X_\alpha)^2)\|\varphi\|_2$$

for any $\varphi \in \mathcal{P}_{\alpha-1}$.

Proof. We assume that $\|\varphi\|_2 = 1$. For each point $\xi_o' \in X_\alpha$, we choose a compact set $D_{\xi_o'} \subset V_o$ such that their union cover the R-ball and ξ_o' is the only element of X_α in $D_{\xi'_o}$. Denote by $D_{\xi''_o}$ one of the regions that contain ξ_o. Write W for the 1-eigenspace of $\rho_{0,0,\xi''_o}(\mu_1)$ in $\mathcal{P}_{\alpha-1}$, and write U for the orthogonal complement. Write π_W and π_U for the orthogonal projections respectively. Set $a = \|\pi_W\varphi\|_2$ and $b = \|\pi_U\varphi\|_2$.

Since W and U are invariant under $\rho_{0,0,\xi_o}(\mu_1)$, we have

$$\pi_U\rho_{0,0,\xi_o}(\mu_1)\pi_W\varphi = 0 = \pi_W\rho_{0,0,\xi_o}(\mu_1)\pi_U\varphi.$$

The function $\|\rho_{0,0,\xi_o}|_U\|$ is continuous in $D_{\xi''_o}$. Denote by $1-c_1$ its maximum. Observe that $c_1 > 0$ and it depends only on R, α, μ_1 and ξ''_o, and that there are a finite number of possibilities for ξ''_o, so c_1 can be bounded below by a positive number depending only on R, α, μ_1. Then

$$\|\rho_{0,0,\xi_o}(\mu_1)\pi_U\varphi\|_2 < (1-c_1)b.$$

Combining the above inequalities with Lemma 36, we get

$$\|\pi_U\rho_{r_{ss},r_a,\xi_o}(\mu_1)\pi_W\varphi\|_2 \le C(r_{ss}+r_a)a, \tag{56}$$

$$\|\pi_W\rho_{r_{ss},r_a,\xi_o}(\mu_1)\pi_U\varphi\|_2 \le C(r_{ss}+r_a)b, \tag{57}$$

$$\|\pi_U\rho_{r_{ss},r_a,\xi_o}(\mu_1)\pi_U\varphi\|_2 < (1-c_1/2)b \tag{58}$$

if r_{ss} and r_a are sufficiently small (depending on c_1).

We get

$$\|\pi_W\rho_{r_{ss},r_a,\xi_o}(\mu_1)\pi_W\varphi\|_2 < (1-c\,\mathrm{dist}(\xi_o,X_\alpha)^2)a \tag{59}$$

from Lemma 37.

Combining estimates (56– 59) we can write

$$\begin{aligned}
\|\rho_{r_{ss},r_a,\xi_o}(\mu_1)\varphi\|_2^2 \le{}& [(1-c\,\mathrm{dist}(\xi_o,X_\alpha)^2)a + C(r_{ss}+r_a)b]^2\\
&+[(1-c_1/2)b + C(r_{ss}+r_a)a]^2\\
\le{}& (1-c\,\mathrm{dist}(\xi_o,X_\alpha)^2)a^2 + (1-c_1/2)b^2\\
&+4C(r_{ss}+r_a)ab + C^2(r_{ss}+r_a)^2\\
\le{}& \left(1-\frac{c\,\mathrm{dist}(\xi_o,X_\alpha)^2}{2}\right)a^2 + \left(1-c_1/2+\frac{C_2(r_{ss}+r_a)^2}{\mathrm{dist}(\xi_o,X_\alpha)^2}\right)b^2\\
&+C^2(r_{ss}+r_a)^2.
\end{aligned}$$

We used the inequality between the geometric and the arithmetic means in the last line. We can assume $10C_2(r_{ss}+r_a)^2 < c_1\,\mathrm{dist}(\xi_o,X_\alpha)^2$, and the lemma follows. □

The following lemma allows us to approximate $\widehat{\nu}_l$ by polynomials in the ξ_{ss} and ξ_a variables using Taylor expansion.

LEMMA 39. *Let μ be a probability measure on* $\mathrm{Isom}(\mathbf{R}^d)$ *with finite moments of order α, and suppose that there are no points but the origin that is* *fixed by all $\theta(\gamma)$ for $\gamma \in \operatorname{supp}\mu$. Then there is a constant C depending on α and μ such that*

$$\int |v(\gamma)|^\alpha d\mu^{*(l)}(\gamma) \leq Cl^{\alpha/2}.$$

Proof. Changing the origin changes the αth order moments by an additive constant at most, so for the purposes of this proof, we can assume that μ satisfies (C) due to Lemma 4. Let $X_1, \ldots, X_l$ be independent random isometries with law μ. Consider the sequence of random vectors

$$Y_l = v(X_1 \cdots X_l) = v(X_1) + \theta(X_1)v(X_2) + \cdots + \theta(X_1)\cdots\theta(X_{l-1})v(X_l).$$

By (C), these form a martingale, and its conditional moments of order α are uniformly bounded. Thus the lemma follows from Burkholder's inequality; see [8, Th. 3.2]. □

Note that if we apply the above lemma for the measures $\pi_{ss}(\mu)$ and $\pi_a(\mu)$, then we get

$$\int |v_{ss}(\gamma)|^\alpha + |v_a(\gamma)|^\alpha d\mu^{*(l)}(\gamma) \leq Cl^{\alpha/2}. \tag{60}$$

We denote by Y_α the largest subset of X_α invariant under $\theta_o(\operatorname{supp}\mu)$, and we write $Z_\alpha = X_\alpha \backslash Y_\alpha$. Let δ be a number that satisfies the inequalities $C^{-1} \geq \delta \geq C(r_{ss} + r_a)$ as in Proposition 34, where C is a number that may depend on μ, x_0, R and α. We write

$$\begin{aligned}
D_1 &= \{\xi : |\pi_{ss}(\xi)| = r_{ss}, |\pi_a(\xi)| = r_a, \operatorname{dist}(\pi_o(\xi), Y_\alpha\backslash\{0\}) \leq \delta\},\\
D_2 &= \{\xi : |\pi_{ss}(\xi)| = r_{ss}, |\pi_a(\xi)| = r_a, \operatorname{dist}(\pi_o(\xi), Y_\alpha) \geq \delta, |\pi_o(\xi)| \leq R\},\\
D_3 &= \{\xi : |\pi_{ss}(\xi)| = r_{ss}, |\pi_a(\xi)| = r_a, \operatorname{dist}(\pi_o(\xi), X_\alpha) \geq \delta, |\pi_o(\xi)| \leq R\}.
\end{aligned}$$

Observe that $D_1 \cup D_2$ is the domain of integration in Proposition 34. We also note that D_1 and D_2 are invariant under $\theta(\operatorname{supp}\mu)$, while D_3 is invariant under $\theta(\operatorname{supp}\mu_1)$. These features will be important in what follows. We denote by $\|\cdot\|_{L^2(D_i)}$ the L^2 norm with respect to the natural volume measure on these manifolds normalized to have total mass 1. That is, we have $\|1\|_{L^2(D_i)} = 1$ by our convention.

We now estimate the Fourier transform of $\mu_1.\nu_l$ on D_3 using Lemma 38 and approximating ν_l by polynomials in $\pi_{ss}(\xi) \oplus \pi_a(\xi)$.

Lemma 40. *There are numbers c, C depending only on μ_1, x_0, R and α such that for any integer l, we either have*

$$\left\| \int e(\langle v(\gamma), \xi\rangle)\widehat{\nu_l}(\theta(\gamma)^{-1}\xi)d\mu_1(\gamma)\right\|_{L^2(D_3)} \leq (1 - c\delta^2)\|\widehat{\nu_l}\|_{L^2(D_3)}$$

or

$$\|\widehat{\nu_l}\|_{L^2(D_3)} \leq C\delta^{-2}(r_{ss} + r_a)^\alpha l^{\alpha/2}.$$

Proof. Note that

$$\widehat{\nu}(\xi) = \int e(\langle v(\gamma) + \theta(\gamma)x_0, \xi\rangle) d\mu^{*(l)}(\gamma), \tag{61}$$

where x_0 is the starting point of the random walk. We fix a point $\xi_o \in V_o$ such that $\operatorname{dist}(\xi_o, X_\alpha) > \delta$. We take the Taylor expansion of (61) around ξ_o. Using Lemma 39 and its corollary (60), we find a polynomial $\varphi_{\xi_o} \in \mathcal{P}_{\alpha-1}$ such that

$$|\widehat{\nu}_l(\xi) - \varphi_{\xi_o}(\pi_{ss}(\xi)/r_{ss}, \pi_a(\xi)/r_a)| \leq C(r_{ss} + r_a)^\alpha l^{\alpha/2} \tag{62}$$

for all ξ satisfying $\pi_o(\xi) = \xi_o$, $\pi_{ss}(\xi) = r_{ss}$ and $\pi_a(\xi) = r_a$.

By Lemma 38, we have

$$\|\rho_{r_{ss}, r_a, \xi_o}(\mu_1)\varphi_{\xi_o}\|_2 \leq (1 - c\delta^2)\|\varphi_{\xi_o}\|_2. \tag{63}$$

We integrate (62) and (63) for ξ_o:

$$\begin{aligned}
&\left\| \int e(\langle v(\gamma), \xi\rangle) \widehat{\nu}_l(\theta(\gamma)^{-1}\xi) d\mu_1(\gamma) \right\|_{L^2(D_3)} \\
&\quad \leq \left\| \int e(\langle v(\gamma), \xi\rangle) \varphi_{\pi_o(\xi)}(\theta_{ss}(\gamma)^{-1}\pi_{ss}(\xi)/r_{ss}, \theta_a(\gamma)^{-1}\pi_a(\xi)/r_a) d\mu_1(\gamma) \right\|_{L^2(D_3)} \\
&\qquad + C(r_{ss} + r_a)^\alpha l^{\alpha/2} \leq (1 - c\delta^2)\|\widehat{\nu}_l\|_{L^2(D_3)} + C(r_{ss} + r_a)^\alpha l^{\alpha/2}.
\end{aligned}$$

This finishes the proof. □

We give a similar estimate on D_1. The argument is essentially the same, but for $\alpha = 1$.

Lemma 41. *There are numbers c, C depending only on μ_1, x_0, R and α such that for any integer l, we either have*

$$\left\| \int e(\langle v(\gamma), \xi\rangle) \widehat{\nu}_l(\theta(\gamma)^{-1}\xi) d\mu_1(\gamma) \right\|_{L^2(D_1)} \leq (1 - c)\|\widehat{\nu}_l\|_{L^2(D_1)}$$

or

$$\|\widehat{\nu}_l\|_{L^2(D_1)} \leq C(r_{ss} + r_a) l^{1/2}.$$

Proof. First we prove that $X_1 = \{0\}$. Indeed, let $\xi_o \in X_1$. Then for every $\gamma \in \operatorname{supp} \mu_1$, we have $\varphi = \rho_{0,0,\xi_o}(\gamma)\varphi$ for a constant function φ. Hence $e(\langle v_o(\gamma), \xi_o\rangle) = 1$ for all such γ. Since the set (53) is not contained in a proper affine subspace, this implies that $\xi_o = 0$, proving the claim.

We now fix a point $\xi_o \in V_o$ such that $\operatorname{dist}(\xi_o, X_\alpha) \leq \delta$. Similarly to (62), we can find a constant function φ_{ξ_o} such that $|\widehat{\nu}_l(\xi) - \varphi_{\xi_o}| \leq C(r_{ss} + r_a) l^{1/2}$ for all ξ satisfying $\pi_o(\xi) = \xi_o$, $\pi_{ss}(\xi) = r_{ss}$ and $\pi_a(\xi) = r_a$.

Note that there is a number $c_1 > 0$ depending only on μ_1, R and α such that $\operatorname{dist}(\xi_o, X_1) = |\xi_o| > c_1$. Then by Lemma 38, we have

$$\|\rho_{r_{ss}, r_a, \xi_o}(\mu_1)\varphi_{\xi_o}\|_2 \leq (1 - c)\|\varphi_{\xi_o}\|_2.$$

As in the proof of Lemma 40, we can deduce

$$\left\|\int e(\langle v(\gamma),\xi\rangle)\widehat{\nu}_l(\theta(\gamma)^{-1}\xi)d\mu_1(\gamma)\right\|_{L^2(D_1)} \le (1-c)\|\widehat{\nu}_l\|_{L^2(D_1)} + C(r_{ss}+r_a)l^{1/2},$$

which proves the claim. $\square$

We turn Lemma 40 into an estimate on $\widehat{\nu}_{l+L}$, where L is the number from Lemma 35. We use a trick similar to (7). Notice that the estimate is useful only if a large proportion of the L^2 mass of $\widehat{\nu}_l$ is on D_3.

LEMMA 42. *There are numbers c, C depending only on μ, x_0, R and α such that for any integer l, we either have*

$$\|\widehat{\nu}_{l+L}\|_{L^2(D_2)} \le \|\widehat{\nu}_l\|_{L^2(D_2)} - c\delta^2\|\widehat{\nu}_l\|_{L^2(D_3)}$$

or

$$\|\widehat{\nu}_l\|_{L^2(D_3)} \le C\delta^{-2}(r_{ss}+r_a)^\alpha l^{\alpha/2}.$$

Proof. We suppose that the second alternative of the conclusion does not hold. Then it follows from Lemma 40 that

$$\left\|\int e(\langle v(\gamma),\xi\rangle)\widehat{\nu}_l(\theta(\gamma)^{-1}\xi)d\mu_1(\gamma)\right\|_{L^2(D_2)} \le \|\widehat{\nu}_l\|_{L^2(D_2)} - c\delta^2\|\widehat{\nu}_l\|_{L^2(D_3)}.$$

We use here that both D_2 and D_3 are invariant under $\theta(\operatorname{supp}\mu_1)$ and their volumes are bounded by a constant multiple of each other.

Recall that $\widetilde{\mu}^{*(L)} * \mu^{*(L)} = p\mu_1 + q\mu_2$; hence

$$\begin{aligned}(64)\qquad &\left\|\int e(\langle v(\gamma),\xi\rangle)\widehat{\nu}_l(\theta(\gamma)^{-1}\xi)d\widetilde{\mu}^{*(L)} * \mu^{*(L)}(\gamma)\right\|_{L^2(D_2)} \\ &\qquad \le \|\widehat{\nu}_l\|_{L^2(D_2)} - cp\delta^2\|\widehat{\nu}_l\|_{L^2(D_3)}.\end{aligned}$$

We can write

$$\begin{aligned}\|\nu_{l+L}\|^2_{L^2(D_2)} &= \fint_{D_2}\left|\int e(\langle v(\gamma),\xi\rangle)\widehat{\nu}_l(\theta(\gamma)^{-1}\xi)d\mu^{*(L)}(\gamma)\right|^2 d\xi \\ &= \fint_{D_2}\iint e(\langle v(\gamma_1),\xi\rangle)\widehat{\nu}_l(\theta(\gamma_1)^{-1}\xi) \\ &\quad \cdot e(-\langle v(\gamma_2),\xi\rangle)\overline{\widehat{\nu}_l(\theta(\gamma_2)^{-1}\xi)}d\mu^{*(L)}(\gamma_2)d\mu^{*(L)}(\gamma_1)d\xi \\ &= \fint_{D_2}\iint e(\langle -\theta(\gamma_2)^{-1}v(\gamma_2) + \theta(\gamma_2)^{-1}v(\gamma_1),\xi\rangle) \\ &\quad \cdot \widehat{\nu}_l(\theta(\gamma_1)^{-1}\theta(\gamma_2)\xi)\overline{\widehat{\nu}_l(\xi)}d\mu^{*(L)}(\gamma_2)d\mu^{*(L)}(\gamma_1)d\xi \\ &= \fint_{D_2}\int e(\langle v(\gamma),\xi\rangle)\widehat{\nu}_l(\theta(\gamma)^{-1}\xi)d\widetilde{\mu}^{*(L)} * \mu^{*(L)}(\gamma)\overline{\widehat{\nu}_l(\xi)}d\xi \\ &\le \|\widehat{\nu}_l\|_{L^2(D_2)} \cdot \left\|\int e(\langle v(\gamma),\xi\rangle)\widehat{\nu}_l(\theta(\gamma)^{-1}\xi)d\widetilde{\mu}^{*(L)} * \mu^{*(L)}(\gamma)\right\|_{L^2(D_2)}\end{aligned}$$

$$\leq \|\widehat{\nu}_l\|_{L^2(D_2)}(\|\widehat{\nu}_l\|_{L^2(D_2)} - cp\delta^2\|\widehat{\nu}_l\|_{L^2(D_3)})$$
$$\leq \left(\|\widehat{\nu}_l\|_{L^2(D_2)} - \frac{cp\delta^2}{2}\|\widehat{\nu}_l\|_{L^2(D_3)}\right)^2,$$

which proves the lemma. We used the symbol $⨍$ to denote integration with respect to the normalized volume measure of total mass 1. □

The same proof using Lemma 41 instead of Lemma 40 gives the following.

LEMMA 43. *There are numbers c, C depending only on μ, x_0, R and α such that for any integer l, we either have*

$$\|\widehat{\nu}_{l+L}\|_{L^2(D_1)} \leq (1-c)\|\widehat{\nu}_l\|_{L^2(D_1)}$$

or

$$\|\widehat{\nu}_l\|_{L^2(D_1)} \leq C(r_{ss}+r_a)l^{1/2}.$$

Now we use rotations $\theta_o(\gamma)$ for $\gamma \in \operatorname{supp}\mu$ to move the L^2 mass away from Z_α. This allows us to prove that there is a number $0 \leq k \leq |Z_\alpha|$ such that the L^2 mass of $\widehat{\nu}_{l+k}$ on D_2 is not concentrated near the points in Z_α, and we can upgrade Lemma 42 into the following lemma.

LEMMA 44. *There are numbers c, C depending only on μ, x_0, R and α such that for any integer l, we either have*

$$\|\widehat{\nu}_{l+|Z_\alpha|+L}\|_{L^2(D_2)} \leq (1-c\delta^2)\|\widehat{\nu}_l\|_{L^2(D_2)}$$

or

$$\|\widehat{\nu}_l\|_{L^2(D_2)} \leq C\delta^{-2}(r_{ss}+r_a)^\alpha l^{\alpha/2}.$$

Proof. It is clear that $\|\widehat{\nu}_l\|_{L^2(D_2)}$ decreases as l grows, so it will be sufficient to prove the inequality for $\widehat{\nu}_{l+k+L}$ for some $0 \leq k \leq |Z_\alpha|$.

We show that there is some $c > 0$ depending only on R, μ and α such that there is $0 \leq k \leq |Z_\alpha|$ such that

$$\|\widehat{\nu}_{l+k}\|_{L^2(D_3)} \geq c\|\widehat{\nu}_l\|_{L^2(D_2)} \quad \text{or} \quad \|\widehat{\nu}_{l+k}\|_{L^2(D_2)} \leq (1-c)\|\widehat{\nu}_l\|_{L^2(D_2)}. \tag{65}$$

This combined with Lemma 42 finishes the proof.

Suppose that the first inequality in (65) does not hold for $k = 0$, say with $c = 1/2$. Then there is $\xi_o \in Z_\alpha$ such that

$$\|1_{\{\xi:|\pi_o(\xi)-\xi_o|\leq\delta\}}\widehat{\nu}_l\|^2_{L^2(D_2)} \geq \frac{\|\widehat{\nu}_l\|^2_{L^2(D_2)}}{2|Z_\alpha|}.$$

Here $1_{\{\xi:|\pi_o(\xi)-\xi_o|\leq\delta\}}$ denotes the indicator function of the set $\{\xi : |\pi_o(\xi)-\xi_o| \leq \delta\}$. It is clear that there is some $c_0 > 0$ depending only on μ, R and α such that there is $k \leq |Z_\alpha|$ and $\gamma_0 \in \operatorname{supp}(\mu^{*(k)})$ with $\operatorname{dist}(\theta_o(\gamma_0)^{-1}\xi_o, X_\alpha) > c_0$. We can assume that $\delta < c_0/10$. Hence there is a neighborhood of γ_0 in $\operatorname{Isom}(\mathbf{R}^d)$, which we denote by U, such that $\mathrm{dist}(\theta_o(\gamma)^{-1}\xi_o, X_\alpha) > 2\delta$ for each $\gamma \in U$. Thus for each $\gamma \in U$, we have

$$\|1_{\{\xi:\mathrm{dist}(\pi_o(\xi),X_\alpha)\le\delta\}}(\xi)\widehat{\nu}_l(\theta(\gamma)^{-1}\xi)\|^2_{L^2(D_2)} \le (1-1/(2|Z_\alpha|))\|\widehat{\nu}_l\|^2_{L^2(D_2)}.$$

Recall that

$$\widehat{\nu}_{l+k}(\xi) = \int e(\langle\xi, v(\gamma)\rangle\widehat{\nu}_l(\theta(\gamma)^{-1}\xi)d\mu^{*(k)}(\gamma).$$

Since $\mu^{*(k)}(U) \ge c_1 > 0$ for some number c_1 depending only on μ, R and α, we have

$$\|1_{\{\xi:\mathrm{dist}(\pi_o(\xi),X_\alpha)\le\delta\}}\widehat{\nu}_{l+k}\|^2_{L^2(D_2)} \le (1-c_1/(2|Z_\alpha|))\|\widehat{\nu}_l\|^2_{L^2(D_2)}.$$

This implies (65) if the number c there is sufficiently small. □

Proof of Proposition 34. We apply Lemma 43 iteratively for $l = 0, L, 2L, \ldots, \lceil A\log s^{-1}\rceil L$. If the first alternative of the lemma holds always, then we have $\|\widehat{\nu}_{\lceil A\log s^{-1}\rceil L}\|_{L^2(D_1)} \le s^{cA}$. In the opposite case, we get

$$\begin{aligned}\|\widehat{\nu}_{\lceil A\log s^{-1}\rceil L}\|_{L^1(D_1)} &\le \|\widehat{\nu}_{\lceil A\log s^{-1}\rceil L}\|_{L^2(D_1)}\\ &\le C(r_{ss}+r_a)(\lceil A\log s^{-1}\rceil L)^{1/2} \le CA^{1/2}\log^{1/2}(s^{-1})s.\end{aligned}$$

If we choose A sufficiently large depending on R, μ and α, the last expression will be larger than $s^{A/c}$.

We also apply Lemma 44 iteratively for $l = 0, L+|Z_\alpha|, 2(L+|Z_\alpha|), \ldots, \lceil A\delta^{-2}\log s^{-1}\rceil(L+|Z_\alpha|)$. If the first alternative of the lemma holds always, then we have $\|\widehat{\nu}_{\lceil A\delta^{-2}\log s^{-1}\rceil(L+|Z_\alpha|)}\|_{L^2(D_2)} \le s^{cA}$. In the opposite case, we get

$$\begin{aligned}\|\widehat{\nu}_{\lceil A\delta^{-2}\log s^{-1}\rceil(L+|Z_\alpha|)}\|_{L^1(D_2)} &\le \|\widehat{\nu}_{\lceil A\delta^{-2}\log s^{-1}\rceil(L+|Z_\alpha|)}\|_{L^2(D_2)}\\ &\le C\delta^{-2}(r_{ss}+r_a)^\alpha(\lceil A\delta^{-2}\log s^{-1}\rceil(L+|Z_\alpha|))^{\alpha/2}\\ &\le CA^{\alpha/2}\delta^{-\alpha-2}\log^{\alpha/2}(s^{-1})s^\alpha.\end{aligned}$$

If we choose A sufficiently large depending on R, μ and α, the last expression will be larger than $s^{A/c}$. Summing the above two estimates and taking into account that $\mathrm{Vol}(D_1) \le Cr_{ss}^{\dim V_{ss}-1}r_a^{\dim V_a-1}\delta^{\dim V_o}$ and $\mathrm{Vol}(D_2) \le Cr_{ss}^{\dim V_{ss}-1}r_a^{\dim V_a-1}$, we get the claim. □

8.3. *Proof of the Local Limit Theorem.* Recall from the statement of the theorem that $X_1, X_2\ldots$ are independent identically distributed random isometries. By the assumptions of the theorem, the common law of X_i is nondegenerate and has finite moments of order $\alpha > d^2+3d$.

By Lemma 4 we can choose the origin in such a way that $v := \mathbf{E}[X_1(x_0)-x_0]$ is fixed by K. Now let $\gamma_v \in \mathrm{Isom}(\mathbf{R}^d)$ be translation by $-v$. Consider the random isometries $X_i\cdot\gamma_v$, and denote by μ their common law. Then μ also satisfies (C) besides nondegeneracy and the above moment condition, and clearly it is enough to prove the theorem for these modified random isometries.

We can approximate any compactly supported continuous function in L^∞ norm by functions that have smooth (say C^∞) and compactly supported Fourier transform. Therefore we consider an arbitrary function f such that $\widehat{f}$ is smooth and compactly supported, and we prove the conclusion of Theorem 2 for it. Then this will prove the theorem by approximation. Let $R > 0$ be a number such that the support of $\widehat{f}$ is contained in the ball of radius R around the origin.

We again write $\nu_l = \mu^{*(l)}.\delta_{x_0}$ and use Plancherel's formula

$$\int f(x) d\nu_l(x) = \int \widehat{f}(\xi)\widehat{\nu}_l(\xi) d\xi.$$

Let Δ be the quadratic form from Proposition 15. It is easily seen that

$$\lim_{l\to\infty} l^{d/2} \int \widehat{f}(\xi) e^{-l\Delta(\xi,\xi)} d\xi = c\widehat{f}(0),$$

where c is a constant depending on Δ. Since $\widehat{f}(0) = \int f(x)dx$, it is enough to show that

$$\lim_{l\to\infty} l^{d/2} \int \widehat{f}(\xi)(\widehat{\nu}(\xi) - e^{-l\Delta(\xi,\xi)}) d\xi = 0.$$

The rest of the proof is devoted to estimating the above integral. We break it up into several regions. Let $\delta = l^{-\beta}$ with $\beta > d/(2d+2)$ and also

$$\beta(\alpha+2) - \frac{\alpha}{2} < -\frac{d}{2},$$

which is possible since $\alpha > d^2 + 3d$. (This will also be the δ that we set in Proposition 34.) The first region is defined as $\Omega_1 := \{\xi : |\xi| \le \delta\}$. Proposition 15 implies that

$$r^{-d+1} \int_{|\xi|=r} |\widehat{\nu}_l(\xi) - e^{-l\Delta(\xi,\xi)}|^2 d\xi \le Cr^2.$$

By the Cauchy-Schwartz inequality, we have

$$r^{-d+1} \int_{|\xi|=r} |\widehat{\nu}_l(\xi) - e^{-l\Delta(\xi,\xi)}| d\xi \le Cr.$$

After integrating for $0 \le r \le \delta = l^{-\beta}$ and using $|\widehat{f}(\xi)| \le \|f\|_1$, we get

$$\left| \int_{\Omega_1} \widehat{f}(\xi)(\widehat{\nu}(\xi) - e^{-l\Delta(\xi,\xi)}) d\xi \right| \le C\|f\|_1 l^{-\beta(d+1)}.$$

Since $\beta > d/(2d+2)$, the right side is $o(l^{-d/2})$.

Recall the notation from the beginning of Section 8, where we decomposed $\mathbf{R}^d$ as an orthogonal sum $V_{ss} \oplus V_a \oplus V_o$. To simplify the notation, we write $\xi = (\xi_{ss}, \xi_a, \xi_o)$, where ξ_i is the component of ξ in the corresponding subspace V_i.

The second region we consider is

$$\Omega_2 := \{\xi = (\xi_{ss}, \xi_a, \xi_o) : |\xi_{ss}| + |\xi_a| > C_0 l^{-1/2} \log^{1/2} l, |\xi| < R\},$$

where C_0 is a suitable constant depending on μ. (This region has an overlap with the first one.) We integrate the bound in Proposition 27 for $C_0 l^{-1/2}\log^{1/2} l < r_{ss}+r_a < R$ and $0\le r_o\le R$ and obtain

$$\int_{\Omega_2} |\widehat{\nu}_l(\xi)|^2 d\xi \le C\mathrm{Vol}(\Omega_2) e^{-c(C_0 l^{-1/2}\log^{1/2} l)^2 l} \le C\mathrm{Vol}(\Omega_2) l^{-(d+1)}$$

if we take C_0 sufficiently large. The number $d+1$ in the exponent is arbitrary. Using the Cauchy-Schwartz inequality as above and $|\widehat{f}(\xi)|\le \|f\|_1$, we get

$$\left|\int_{\Omega_2} \widehat{f}(\xi)(\widehat{\nu}(\xi)-e^{-l\Delta(\xi,\xi)})d\xi\right| \le C\mathrm{Vol}(\Omega_2)\|f\|_1 l^{-(d+1)/2}.$$

Note that $e^{-l\Delta(\xi,\xi)}$ is negligible in the region of integration. The right-hand side is again $o(l^{-d/2})$.

The third region we consider is given by the inequalities

$$\Omega_3 = \{\xi : |\xi_{ss}|+|\xi_a| < C_0 l^{-1/2}\log^{1/2} l,\ \delta < |\xi_o| < R\}.$$

Note that $C_0 l^{-1/2}\log^{1/2} l$ is much smaller than δ, so if $\xi\notin\Omega_1$, that is $|\xi|>2\delta$, then $\xi\in\Omega_2\cup\Omega_3$. We use Proposition 34 with $s = C_0 l^{-1/2}\log^{1/2} l$ and integrate the bound for $0 < r_{ss}+r_a < s$ and $\delta\le r_o\le R$ and get

$$\int_{\Omega_3} |\widehat{\nu}_l(\xi)| d\xi \le C s^{\dim V_{ss}+\dim V_a}(\log^{1/2}(s^{-1}) s\delta^{\dim V_o} + \log^{\alpha/2}(s^{-1}) s^{\alpha}\delta^{-\alpha-2}). \tag{66}$$

For the first term on the right, we write

$$\begin{aligned} &C s^{\dim V_{ss}+\dim V_a}\log^{1/2}(s^{-1}) s\delta^{\dim V_o} \\ &\qquad = C l^{d/2} l^{(1/2-\beta)\dim V_o} l^{-1/2}\log^{1+(\dim V_{ss}+\dim V_a)} l. \end{aligned}$$

Since $\beta > d/(2d+2)$, we have $\dim V_o(1/2-\beta) < 1/2$, so the right-hand side is $o(l^{-d/2})$. For the second term in (66), we write

$$\begin{aligned} &C s^{\dim V_{ss}+\dim V_a}\log^{\alpha/2}(s^{-1}) s^{\alpha}\delta^{-\alpha-2} \\ &\qquad \le C l^{-(\dim V_{ss}+\dim V_a)/2} l^{\beta(\alpha+2)-\alpha/2}\log^{\alpha+(\dim V_{ss}+\dim V_a)/2} l. \end{aligned}$$

Since $\beta(\alpha+2)-\alpha/2 < -d/2$, the right-hand side is again $o(l^{-d/2})$.

Using $|\widehat{f}(\xi)|\le\|f\|_1$ again, we get

$$\left|\int_{\Omega_3} \widehat{f}(\xi)(\widehat{\nu}(\xi)-e^{-l\Delta(\xi,\xi)})d\xi\right| \le o(l^{-d/2})\|f\|_1.$$

Combining the estimates for the three regions above, we get the theorem.

References

[1] M. Ådahl, I. Melbourne, and M. Nicol, Random iteration of Euclidean isometries, *Nonlinearity* **16** (2003), 977–987. MR 1975792. Zbl 1037.37006. http://dx.doi.org/10.1088/0951-7715/16/3/311.

[2] V. I. Arnol′d and A. L. Krylov, Uniform distribution of points on a sphere and certain ergodic properties of solutions of linear ordinary differential equations in a complex domain, *Dokl. Akad. Nauk SSSR* **148** (1963), 9–12. MR 0150374. Zbl 0237.34008.

[3] P. Baldi, P. Bougerol, and P. Crépel, Théorème central limite local sur les extensions compactes de $\mathbf{R}^d$, *Ann. Inst. H. Poincaré Sect. B* **14** (1978), 99–111. MR 0501239. Zbl 0382.60013.

[4] H. Bass, Theorems of Jordan and Burnside for algebraic groups, *J. Algebra* **82** (1983), 245–254. MR 0701045. Zbl 0506.20016. http://dx.doi.org/10.1016/0021-8693(83)90182-5.

[5] J. Bourgain and A. Gamburd, A spectral gap theorem in SU(d), *J. Eur. Math. Soc.* (*JEMS*) **14** (2012), 1455–1511. MR 2966656. Zbl 1254.43010. http://dx.doi.org/10.4171/JEMS/337.

[6] J. Bourgain and A. Gamburd, On the spectral gap for finitely-generated subgroups of SU(2), *Invent. Math.* **171** (2008), 83–121. MR 2358056. Zbl 1135.22010. http://dx.doi.org/10.1007/s00222-007-0072-z.

[7] E. Breuillard, Random walks on Lie groups. Available at http://www.math.u-psud.fr/~breuilla/part0gb.pdf.

[8] D. L. Burkholder, Distribution function inequalities for martingales, *Ann. Probability* **1** (1973), 19–42. MR 0365692. Zbl 0301.60035. http://dx.doi.org/10.1214/aop/1176997023.

[9] J.-P. Conze and Y. Guivarc'h, Ergodicity of group actions and spectral gap, applications to random walks and Markov shifts, *Discrete Contin. Dyn. Syst.* **33** (2013), 4239–4269. MR 3038061. Zbl 06224557. http://dx.doi.org/10.3934/dcds.2013.33.4239.

[10] C. M. Dawson and M. A. Nielsen, The Solovay-Kitaev algorithm, *Quantum Inf. Comput.* **6** (2006), 81–95. MR 2212257. Zbl 1152.81706. Available at http://www.rintonpress.com/xqic6/qic-6-1/081-095.pdf.

[11] D. Dolgopyat, On mixing properties of compact group extensions of hyperbolic systems, *Israel J. Math.* **130** (2002), 157–205. MR 1919377. Zbl 1005.37005. http://dx.doi.org/10.1007/BF02764076.

[12] L. G. Gorostiza, The central limit theorem for random motions of d-dimensional Euclidean space, *Ann. Probability* **1** (1973), 603–612. MR 0353408. Zbl 0263.60010. http://dx.doi.org/10.1214/aop/1176996889.

[13] M. Gotô, A Theorem on compact semi-simple groups, *J. Math. Soc. Japan* **1** (1949), 270–272. MR 0033829. Zbl 0041.36208. http://dx.doi.org/10.2969/jmsj/00130270.

[14] A. Grintsyavichyus, The domain of normal attraction of a stable law for the group of motions of a Euclidean space, *Litovsk. Mat. Sb.* **25** (1985), 39–52. MR 0823642.

[15] Y. Guivarc'h, Equirépartition dans les espaces homogènes, in *Théorie Ergodique* (Actes Journées Ergodiques, Rennes, 1973/1974), *Lecture Notes in Math.* **532**, Springer-Verlag, New York, 1976, pp. 131–142. MR 0480860. Zbl 0368.28024. http://dx.doi.org/10.1007/BFb0080176.

[16] D. A. Každan, Uniform distribution on a plane, *Trudy Moskov. Mat. Obšč.* **14** (1965), 299–305. MR 0193187. Zbl 0225.26018.

[17] Yu. S. Khokhlov, A local limit theorem for the composition of random motions of Euclidean space, *Dokl. Akad. Nauk SSSR* **260** (1981), 295–299. MR 0630143. Zbl 0494.60026.

[18] Yu. S. Khokhlov, The domain of normal attraction of a semistable distribution on a semidirect product compact group and $\mathbf{R}^d$, *J. Math. Sci.* **76** (1995), 2147–2152. MR 1356652. http://dx.doi.org/10.1007/BF02363227.

[19] A. Y. Kitaev, Quantum computations: algorithms and error correction, *Uspekhi Mat. Nauk* **52** (1997), 53–112. MR 1611329. Zbl 0917.68063. http://dx.doi.org/10.1070/RM1997v052n06ABEH002155.

[20] V. M. Maximov, Local theorems for Euclidean motions. I, *Z. Wahrsch. Verw. Gebiete* **51** (1980), 27–38. MR 0566105. Zbl 0427.60010. http://dx.doi.org/10.1007/BF00533814.

[21] S. T. Rachev and J. E. Yukich, Rates of convergence of α-stable random motions, *J. Theoret. Probab.* **4** (1991), 333–352. MR 1100238. Zbl 0724.60019. http://dx.doi.org/10.1007/BF01258741.

[22] B. Roynette, Théorème central-limite pour le groupe des déplacements de $\mathbf{R}^d$, *Ann. Inst. H. Poincaré Sect. B* **10** (1974), 391–398 (1975). MR 0375422. Zbl 0324.60026. Available at http://www.numdam.org/item?id=AIHPB_1974__10_4_391_0.

[23] V. N. Tutubalin, The central limit theorem for random motions of Euclidean space, *Vestnik Moskov. Univ. Ser. I Mat. Meh.* **22** (1967), 100–108. MR 0242233.

[24] P. P. Varjú, Random walks in compact groups, *Doc. Math.* **18** (2013), 1137–1175. MR 3138842. Zbl 1278.60011. Available at http://www.math.uiuc.edu/documenta/vol-18/35.pdf.

(Received: September 12, 2012)
(Revised: November 4, 2013)

Centre for Mathematical Sciences, Cambridge, England and
Einstein Institute of Mathematics, The Hebrew University of Jerusalem, Jerusalem, Israel
E-mail: pv270@dpmms.cam.ac.uk