GAFA, Geom. funct. anal.
Vol. 11 (2001) 953 – 970
1016-443X/01/050953-18 $ 1.50+0.20/0

GAFA Geometric And Functional Analysis

DIOPHANTINE PROPERTIES OF ELEMENTS OF SO(3)

V. Kaloshin and I. Rodnianski

Abstract

A number $\alpha \in \mathbb{R}$ is diophantine if it is not well approximable by rationals, i.e. for some $C, \varepsilon > 0$ and any relatively prime $p, q \in \mathbb{Z}$ we have $|\alpha q - p| > Cq^{-1-\varepsilon}$. It is well-known and is easy to prove that almost every α in $\mathbb{R}$ is diophantine. In this paper we address a noncommutative version of the diophantine properties. Consider a pair $A, B \in SO(3)$ and for each $n \in \mathbb{Z}_+$ take all possible words in A, A^{-1}, B, and B^{-1} of length n, i.e. for a multiindex $\mathcal{I} = (i_1, j_1, \dots, i_m, j_m)$ define $|\mathcal{I}| = \sum_{k=1}^{m}(|i_k| + |j_k|) = n$ and $W_n(A, B) = \{W_{\mathcal{I}}(A, B) = A^{i_1} B^{j_1} \dots A^{i_m} B^{j_m}\}_{|\mathcal{I}|=n}$.

Gamburd–Jakobson–Sarnak [GJS] raised the problem: prove that for Haar almost every pair $A, B \in SO(3)$ the closest distance of words of length n to the identity, i.e. $s_{A,B}(n) = \min_{|\mathcal{I}|=n} \|W_{\mathcal{I}}(A, B) - E\|$, is bounded from below by an exponential function in n. This is the analog of the diophantine property for elements of $SO(3)$. In this paper we prove that $s_{A,B}(n)$ is bounded from below by an exponential function in n^2. We also exhibit obstructions to a "simple" proof of the exponential estimate in n.

1 Introduction

The classical result of metric number theory on Diophantine properties of numbers says the following: for any $\varepsilon > 0$ and a.e. $\alpha \in \mathbb{R}$ there is a constant $C = C(\alpha) > 0$ such that the map $n\alpha (\mathrm{mod}\ 1)$ satisfies the property $n\alpha (\mathrm{mod}\ 1) > C|n|^{-1-\varepsilon}$ for every integer n [Kh].

Diophantine properties of numbers arise in various problems in metric number theory [Kh], smooth dynamical systems, holomorphic dynamics [HK], KAM theory [He], [L], [Mo], and others.

Generalizations of the metric number theory led to the development of the theory of simultaneous Diophantine approximations and even Dio-

The first author is partially supported by the Sloan Dissertation Fellowship and the American Institute of Mathematics Five-year Fellowship.

phantine approximations on manifolds. In the latter case consider manifold $M \subset \mathbb{R}^n$ defined by n analytic functions $f_1, \dots, f_n : U \subset \mathbb{R}^d \to \mathbb{R}$, $M = \{\mathbf{f}(x) : x \in U\}$. Assume that functions $1, f_1, \dots, f_n$ are linearly independent over $\mathbb{R}$. One of the central questions of the theory is the following conjecture made by Sprindžuk in 1980 and recently proved by D. Kleinbock and G. Margulis [KlM]:

Any manifold $M \subset \mathbb{R}^n$ of the above type is extremal, i.e. for almost all $\mathbf{y} \in M$ and any $\epsilon > 0$ there exists a positive constant $D(\mathbf{y})$ such that for all $\mathbf{q} \in \mathbb{Z}^n$ and $p \in \mathbb{Z}$

$$|\mathbf{q} \cdot \mathbf{y} + p| \geq \frac{D(\mathbf{y})}{\|\mathbf{q}\|^{n(1+\epsilon)}} . \tag{1}$$

Here $\mathbf{q} \cdot \mathbf{y} = \sum_{i=1}^n q_i y_i$ and $\|\mathbf{q}\| = \max_{1 \leq i \leq n} |q_i|$.

In fact, Kleinbock-Margulis prove even a stronger statement that M is strongly extremal (approximation in the sense of (1) is replaced by the notion of *multiplicative approximation*). The proof is based on the correspondence between the approximation properties of vectors $\mathbf{y} \in \mathbb{R}^n$ and the behavior of certain orbits in the space of unimodular lattices in $\mathbb{R}^{n+1}$.

The analogue of the Diophantine property can be also formulated in the noncommutative setting. As far as we know very little is known in this case. However, some intuition has already been developed for the group $SU(2)(SO(3))$. We say that $g_1, \dots, g_k \in SU(2)$ are Diophantine if there exists a positive constant $D(g_1, \dots, g_k)$ such that for any $n \geq 1$ and any word W_n in $g_1, \dots, g_k$ of length n

$$\|W_n \pm E\| \geq D^{-n} .$$

Our interest to the problem of Diophantine approximations on the group $SO(3)$ stems mainly from the question formulated in the list of open problems in the paper of A. Gamburd, D. Jakobson, and P. Sarnak (Problem 4): *The Haar generic elements* $(g_1, g_2, \dots, g_k) \in SU(2)^k$ *in the sense of measure are Diophantine* [GJS]. The paper [GJS] provides an elementary solution of Ruziewicz problem asserting that the Haar measure on $\mathbb{S}^2$ is the unique finitely additive $SO(3)$ invariant measure defined on Lebesgue sets.

In what follows it is more convenient for us to pass to the group $SO(3)$ and restrict our attention to the case of two generators. Consider a subgroup F generated by two elements $A, B \in SO(3)$. The group $SO(3)$ would have a Diophantine property if for almost all rotations $A, B \in SO(3)$ in the sense of measure and all reduced words $W_n \in F$ of length n in A, B, A^{-1}, B^{-1},

$$\|W_n - E\| \geq D(A, B)^{-n} \tag{2}$$

for some positive constant $D(A,B)$. The presence of the words of the form $ABA^{-1}B^{-1}$ and like indicates that F has to be a free subgroup. It is a classical fact that the set of elements $A, B \in SO(3)$ which do not generate a free subgroup is a countable union of analytic sets of codimension one (see also Lemma 2 for an independent demonstration). To see this it suffices to establish the existence of just one free subgroup of rank two. The first explicit construction of such a subgroup was given by Hausdorff in 1914 in his work on Hausdorff–Banach–Tarski paradox. A free subgroup F of rank two in $SO(3)$ allows one to construct four disjoint subsets of the sphere $\mathbb{S}^2$ such that after rotating these subsets by the elements of F one obtains two copies of $\mathbb{S}^2$ minus a countable set. Modulo the issue of the countable set it follows that there is no finitely additive invariant measure defined on *all* sets of $\mathbb{S}^2$. It also follows that any finitely additive $SO(3)$ invariant measure defined on Lebesgue sets is absolutely continuous with respect to the Lebesgue measure. The book of Wagon [W] provides a good account of the subject.

The Ruziewicz problem is to show that any finitely additive, rotation invariant measure necessarily coincides with the Haar measure on $\mathbb{S}^2$. In the general setting, the problem is formulated for the finitely additive $SO(n+1)$ invariant measure on $\mathbb{S}^n$. It is interesting to note that in dimension 1 Banach provided a negative solution to the Ruziewicz problem. G. Margulis [M] and D. Sullivan [S] used Kazhdan property (T) to give the positive answer in dimensions $n \geq 4$. For the dimensions $n = 2, 3$ the affirmative solution had been given by V. Drinfeld [Dr].

The solution of Ruziewicz problem in dimensions $n \geq 2$ can be reduced to the problem of finding a free subgroup $F \in SO(n+1)$ with a *spectral gap* property [R]. Namely, consider the subspace $L^2_0(\mathbb{S}^n) = \{f \in L^2(\mathbb{S}^n) : \int_{\mathbb{S}^n} f\, d\mu = 0\}$. Then F is said to have a spectral gap property if there exists a positive constant c such that for any $f \in L^2_0(\mathbb{S}^n)$ there exists an element $g \in F$ such that $\|f \circ g - f\| \geq c\|f\|$. After passing from $SO(3)$ to its double cover $SU(2)$ the above can also be reformulated in terms of the spectra of the irreducible representations of $SU(2)$ restricted to the element $z = g_1 + g_1^{-1} + \cdots + g_k + g_k^{-1}$. Namely, let π_N denote the irreducible representation of $SU(2)$ realized as a linear action on the space of homogeneous polynomials in two variables of degree N. Define $\hat{z}(\pi_N) = \pi_N(g_1) + \pi_N(g_1^{-1}) + \cdots + \pi_N(g_k) + \pi_N(g_k^{-1})$ to be an $(N+1) \times (N+1)$ matrix. Then we say that a subgroup F generated by $g_1, \ldots, g_k$ has a gap if

$$\limsup_{N\to\infty} \|\hat{z}(\pi_N)\| < \|z\|\,.$$

A. Lubotzky, R. Phillips, and P. Sarnak constructed explicit examples of elements $g_1, \dots, g_k \in SU(2)$ with $k \geq 3$ generating a subgroup with a gap. For those generators $\|\hat{z}(\pi_N)\| \leq 2\sqrt{2k-1} < 2k$, [LuPS1].

Lubotzky–Phillips–Sarnak also show that the sequence of measures $\mu_N(z)$ associated with the eigenvalue distributions of $\hat{z}(\pi_N)$ has two accumulation points as $N \to \infty$. Namely, they prove that there exist two measures $\nu^{\text{even}}(z)$ and $\nu^{\text{odd}}(z)$ such that $\mu_{2N}(z) \to \nu^{\text{even}}(z)$ and $\mu_{2N+1}(z) \to \nu^{\text{odd}}(z)$. Moreover, the rate of the convergence depends on the Diophantine properties of the generators $g_1, \dots, g_k$ of F. In addition, they show that a free subgroup generated by the elements $g_1, \dots, g_k \in SU(2)$ represented by matrices with algebraic entries is Diophantine, i.e. there exists a constant $D(g_1, \dots, g_k)$ such that that the bound (2) holds for any positive integer n and any word W_n in $g_1, \dots, g_k$ of length n.

Let $\mathcal{H}_s(\mathbb{S}^2)$ be the s-th Sobolev space with the standard norm $\|\cdot\|_s$ and $C^\infty(\mathbb{S}^2)$ be the space of C^∞-functions with zero mean. D. Dolgopyat in [Do], provided that $g_1, \dots, g_k \in SU(2)$ or $\in SO(3)$ generate a dense subgroup in $SU(2)$ or $SO(3)$, obtained upper estimates for $\|f \circ z - f\|_s$ in terms of $\|f\|_{s+\alpha}$. Namely, for some $\alpha > 0$ and any s there is C_s such that $\|f \circ z - f\|_s \leq C_s \|f\|_{s+\alpha}$.

In this paper we take a first step in an attempt to understand the Diophantine properties of the group $SO(3)$. We establish that almost all pairs of elements $(A, B) \in SO(3)$ generate the subgroups satisfying a *weak* Diophantine condition when the conjectured exponent n in (2) is replaced by n^2. Although, the results below are stated for the rank two subgroups of $SO(3)$ they can be easily generalized to include the case of the group $SU(2)$ and a higher number of generators.

It follows from the pigeonhole principle and the compactness of $SO(3)$ that the exponential estimate (2) (not super-exponential and not polynomial) is the optimal one since the number of words of length n grows exponentially with n. It is an easy exercise to show that for a Baire generic (residual) set of pairs $A, B \in SO(3)$ Diophantine condition is not satisfied. Therefore, the problem about Diophantine properties of elements of $SO(3)$ is another example of a property which fails on a Baire generic set, but holds on a set of full measure. Numerous examples of this phenomena appear in dynamical systems and topology (see [O], [HuSY], and [K]).

As we mentioned above, in this paper we obtain the first result on Diophantine properties of elements of $SO(3)$. Consider $SO(3)$ with the Haar measure μ on it. We show that for a.e. pair $(A,\ B) \in SO(3) \times SO(3)$

there is a constant $D > 0$ such that for any n and any word $W_n(A, B)$ of length n in A and B we have

$$\|W_n(A, B) \pm E\| \geq D^{-n^2}. \tag{3}$$

Let us describe the approach we use to prove the result and discuss the difficulties arising in the attempt to get the exponential lower bound as in (2). Let $A, B \in SO(3)$ be two distinct elements, $n \in \mathbb{Z}_+$ be an integer, and $W_n(A, B)$ be a word of length n in A and B. Denote by α and β the angles of rotations of A and B respectively, and by γ the angle between the axes of A and B. Without loss of generality we can assume that the axis of rotation of A, denote v_A, is the OX-axis in the ambient space $\mathbb{R}^3$ and the axis of rotation of B, denote v_B, belongs to the (x, y)-plane forming angle γ with v_A in the clockwise direction. Notice that now any word $W_n(A, B)$ is uniquely defined by the choice of coordinate system and a triple

$$(\alpha, \beta, \gamma) \in \mathbb{T}_\alpha \times \mathbb{T}_\beta \times \mathbb{T}_\gamma = \mathbb{T}^3 . \tag{4}$$

Denote $W_n(A, B) = W_n(\alpha, \beta, \gamma)$. Now consider the 3-dimensional torus $\mathbb{T}^3$ as a parameter space with the Lebesgue measure Leb_3 on it. It is clear that a set of the full product Haar measure $\mu \times \mu$ on $SO(3) \times SO(3)$ corresponds to a set of the full Lebesgue measure Leb_3 on $\mathbb{T}^3$.

The proof presented below is based on the standard Borel–Cantelli arguments. The rough sketch is as follows. Fix a word $W_n(\alpha, \beta, \gamma)$ of length n in A and B. The goal is to estimate the measure of the set of parameters $(\alpha, \beta, \gamma) \in \mathbb{T}^3$ for which $W_n(\alpha, \beta, \gamma)$ is at most D^{-n^2} away from E. Let $m_n(D)$ be the upper bound for the measure of the union of these sets over all words of length n. By Borel–Cantelli if $\sum_n m_n(D) < \infty$, then for a.e. $(\alpha, \beta, \gamma) \in \mathbb{T}^3$ property (3) holds for all except finitely many words. Increasing D, we can satisfy these finite number of conditions and complete the proof.

It turns out that the distance of $W_n(A, B)$ to E can be represented as a trigonometric polynomial $P_n(\alpha, \beta, \gamma)$ of degree n in α, β, and γ with integer coefficients. Fix $\beta = \beta^*$ and $\gamma = \gamma^*$ and consider the measure of α's for which $P_n(\alpha, \beta^*, \gamma^*)$ is D^{-n^2}-small. If a nontrivial polynomial $P_n(\alpha, \beta^*, \gamma^*)$ with integer coefficients has a zero of order n in α then the measure $|\{\alpha : |P_n(\alpha, \beta^*, \gamma^*)| < D^{-n^2}\}|$ can be as big as D^{-n}. Suppose we can prove that D^{-n} is the upper bound. Since, there are at most 4^n words $W_n(A, B)$ of length n we obtain that the total "bad" measure of words of length n is at most $(4/D)^n$ and is exponentially small for $D > 4$.

One can think that the polynomial $P_n(\alpha, \beta^*, \gamma^*)$ with a zero of high order in α corresponds to the fact that the word $W_n(A, B)$ "sticks" in a

neighborhood of E and leaves this neighborhood slowly as the parameters α, β, γ vary. This shows that a possible presence of high order degeneracies for the polynomial representing the distance from a word $W_n(A, B)$ to E creates difficulties for the estimates of the measure of a set where $W_n(A, B)$ is close to E. In particular, possible high degeneracies stand in the way of proving the desired optimal result (2).

In the last section we present a collection of words $W_n(A, B)$ of length n for which polynomial $P_n(\alpha, \beta, \gamma)$ does have a zero of order $\sqrt{n}$. This shows that it is indeed possible for a word $W_n(A, B)$ to "stick" in a neighborhood of E. This degenerate collection is constructed using the commutants $[A, B] = ABA^{-1}B^{-1}$. The degeneracies of high orders for the trigonometric polynomials $P_n(\alpha, \beta, \gamma)$ do occur. This seems to be the main obstacle to obtain the "true" exponential Diophantine property (2) for subgroups of $SO(3)$.

2 Statement of the Result

Let $A, B \in SO(3)$ be two distinct elements and $m \in \mathbb{Z}_+$. Denote $\mathcal{I}_m = (s_1, r_1, \dots, s_m, r_m)$ a set of integers such that the s_1 and r_m might be zero and the other $2m-2$ are nonzero, $|\mathcal{I}_m| = \sum_p (|s_p| + |r_p|)$, and $W_{\mathcal{I}_m}(A, B) = A^{s_1} B^{r_1} \dots A^{s_m} B^{r_m}$. Thus $W_{\mathcal{I}_m}(A, B)$ corresponds to the word defined by the multiindex $\mathcal{I}_m$.

Theorem 1. *For any element $C \in SO(3)$ and $\mu \times \mu$-a.e. pair $(A, B) \in SO(3) \times SO(3)$ there is a constant $D = D(A, B) > 0$ such that*

$$\min_{m<n} \min_{\{\mathcal{I}_m : |\mathcal{I}_m| = n\}} \left\| W_{\mathcal{I}_m}(A, B) - C \right\| \geq D^{-n^2} \quad \text{for all} \quad n \in \mathbb{Z}_+ . \tag{5}$$

In other words, for any element C and a μ-generic choice of a pair A and B one can find $D = D(A, B) > 0$ with the property that all possible words in A, A^{-1}, B, B^{-1} of length n cannot approximate C better than D^{-n^2}.

Remark 1. *The most interesting case when C is the identity.*

Recall that the pair the pair of rotations $(A, B) \in SO(3) \times SO(3)$ is uniquely defined by a triple $(\alpha, \beta, \gamma) \in \mathbb{T}^3$. The parameters α, β are the angles of rotations of A, B, and γ is the angle between their axes of rotations. Reformulate Theorem 1 in a different form.

Theorem 2. *For any element $C \in SO(3)$ and Lebesgue a.e. $(\alpha, \beta, \gamma) \in \mathbb{T}^3$ there is a constant $D = D(\alpha, \beta, \gamma) > 0$ such that*

$$\min_{m<n} \min_{\{\mathcal{I}_m : |\mathcal{I}_m| = n\}} \left\| W_{\mathcal{I}_m}(\alpha, \beta, \gamma) - C \right\| \geq D^{-n^2} \quad \text{for all} \quad n \in \mathbb{Z}_+ . \tag{6}$$

Fix a word $W_{\mathcal{I}_m}(\alpha,\beta,\gamma)$. The idea of the proof is to show that outside of some small measure set in $\mathbb{T}^3$ the size of the derivative

$$\left\|W_{\mathcal{I}_m}(\alpha,\beta,\gamma)'_\alpha\right\|^2 = D^\alpha_{\mathcal{I}_m}(\alpha,\beta,\gamma) \tag{7}$$

is not too small. When the derivative with respect to α is not too small the word $W_{\mathcal{I}_m}(\alpha,\beta,\gamma)$ varies sufficiently fast with α and passes the "dangerous" D^{-n^2}-neighborhood of the rotation C sufficiently quickly. This implies smallness of the "prohibited" set in the parameter space (α,β,γ).

Fix $n \in \mathbb{Z}_+$ and denote the set of multiindices $\mathcal{R}_n = \cup_{m<n}\{\mathcal{I}_m{:}|\mathcal{I}_m|{=}n\}$. Define

$$\begin{aligned}&\Phi_{\mathcal{I}_m}(D,C) = \left\{(\alpha,\beta,\gamma)\in\mathbb{T}^3 : \|W_{\mathcal{I}_m}(\alpha,\beta,\gamma) - C\| \le D^{-n^2}\right\}\\ &\Phi_n(D,C) = \cup_{\mathcal{I}_m\in\mathcal{R}_n}\Phi_{\mathcal{I}_m}(D,C)\,.\end{aligned} \tag{8}$$

If for some $D^* > 0$ we prove that

$$\sum_{n=1}^{\infty} Leb_3\left\{\Phi_n(D^*,C)\right\} < \infty\,, \tag{9}$$

then for Leb_3-a.e. $(\alpha,\beta,\gamma)\in\mathbb{T}^3$ (resp. $\mu\times\mu$-a.e. $(A,B)\in SO(3)\times SO(3)$) there is $D = D(\alpha,\beta,\gamma) \ge D^*$ (resp. $D = D(A,B)$) such that (6) is satisfied.

To estimate the measure of $\Phi_n(D,C)$ it suffices to fix a word of length n, multiindexed say by $\mathcal{I}_m$, and estimate the measure of the set $\Phi_{\mathcal{I}_m}(D,C)$. Define the set of parameters, where the derivative with respect to α is small

$$\Phi^\alpha_{\mathcal{I}_m} = \left\{(\alpha,\beta,\gamma)\in\mathbb{T}^3 : D_{\mathcal{I}_m}(\alpha,\beta,\gamma) \le D^{-2n^2/3}\right\}. \tag{10}$$

Denote $H(\mathbb{R})$ the ring of quaternions $q = x_0 + ix_1 + jx_2 + kx_3$, $x_p\in\mathbb{R}$. Let $\bar q = x_0 - (ix_1 + jx_2 + kx_3)$ and $N(q) = q\bar q$. Denote $SH(\mathbb{R}) = \{q\in H(\mathbb{R}) : N(q) = 1\}$ the set of unit quaternions. It is well known that there is a representation of $SO(3)$ as $SH(\mathbb{R})$ in the following form:

$$q = \cos\alpha + \sin\alpha(iv_1 + jv_2 + kv_3)\,, \tag{11}$$

where α is the angle of rotation and (v_1,v_2,v_3) is the unit vector in $\mathbb{R}^3$ corresponding to the axis of rotation of an element from $SO(3)$.

LEMMA 1. *With the above notation*

$$\left\|W_{\mathcal{I}_m}(\alpha,\beta,\gamma)''_{\alpha\alpha}\right\|^2 \le |\mathcal{I}_m|^4. \tag{12}$$

Proof. This follows from the quaternion representation (11). Indeed, our choice of the ambient coordinate system gives

$$\begin{aligned}W_{\mathcal{I}_m}(\alpha,\beta,\gamma) = (\cos s_1\alpha + i\sin s_1\alpha)\big(\cos r_1\beta + \sin r_1\beta(i\cos\gamma + j\sin\gamma)\big)\\ \dots(\cos s_m\alpha + i\sin s_m\alpha)\big(\cos r_m\beta + \sin r_m\beta(i\cos\gamma + j\sin\gamma)\big)\,.\end{aligned} \tag{13}$$

Differentiating this expression twice with respect to α yields

$$\left\|W_{\mathcal{I}_m}(\alpha,\beta,\gamma)''_{\alpha\alpha}\right\|^2 \le \left(\sum_{p=1}^{m}|s_p|\right)^4 \le |\mathcal{I}_m|^4. \tag{14}$$

LEMMA 2. *For a nontrivial multiindex $\mathcal{I}_m$ the map $W_{\mathcal{I}_m} : SO(3)\times SO(3) \to SO(3)$ is analytic and nonconstant. This, in particular, implies that a pair of random elements of $SO(3)$ forms a free group.*

REMARK 2. *The conclusion of Lemma 2 is a well-known fact. In particular, the statement that almost all subgroups in $SO(3)$ are free can be reduced to simply showing that there exists a free subgroup in $SO(3)$. The latter is a classical question which was solved positively first by F. Hausdorff in 1914 [Ha]. Note that H. Auerbach in [Au] showed that for compact simply connected Lie group G Haar almost every pair A, B generates a free group whose closure in the whole G. This implies Lemma 2. We present here a very explicit (constructive) independent proof of Lemma 2.*

Proof. Consider the representation (13). To show that a trigonometric function is nontrivial with respect to, say α, it is sufficient to establish that the highest frequency in α has a nonzero functional coefficient. We shall compute this functional coefficient, namely, the coefficient in front the monomial $\exp\left(i\,\mathrm{sign}(s_m)\sum_{p=1}^{m}|s_p|\alpha\right)$. Notice that in the case of positive s the following permutation

$$e^{is\alpha}(\cos r\beta + i\sin r\beta\cos\gamma) = (\cos r\beta + i\sin r\beta\cos\gamma)e^{is\alpha} \tag{15}$$

does not change the sign of s. In the case of negative s we can change the sign of s to positive using the following permutation

$$e^{is\alpha}j\sin r\beta\sin\gamma = j\sin r\beta\sin\gamma e^{-is\alpha}\,. \tag{16}$$

Now we describe the procedure of permuting terms with α to the right and *particular terms with β and γ* to the left so that after such permutations the only term which has α is on the right end of the word and equals $\exp\left(i\,\mathrm{sign}(s_m)\sum_{p=1}^{m}|s_p|\alpha\right)$.

The first step of permutation: Consider the signs of s_1 and s_2. If they are different, then we change the sign of the s_1-term by choosing the permutation (16), otherwise, we choose (15). In both cases $s = s_1$ and $r = r_1$. After the permutation, the first term with α from the left is $\exp\left(i\,\mathrm{sign}(s_2)\sum_{p=1}^{2}|s_p|\alpha\right)$.

The second step of permutation: Consider the signs of s_2 and s_3. Use the recipe of the first step. The permutation gives the third term

$\exp\big(i\,\mathrm{sign}(s_3)\sum_{p=1}^{3}|s_p|\alpha\big)$ and so on. Therefore, the only term which has $\exp\big(i\,\mathrm{sign}(s_m)\sum_{p=1}^{m}|s_p|\alpha\big)$ equals

$$\prod_{\{p:s_ps_{p-1}>0\}}(\cos r_p\beta + i\sin r_p\beta\cos\gamma)$$
$$\times\prod_{\{p:s_ps_{p-1}<0\}} j\sin r_p\beta\sin\gamma\quad\exp\left(i\,\mathrm{sign}(s_m)\sum_{p=1}^{m}|s_p|\alpha\right).$$

In general, the multiplication for quaternions is noncommutative, but terms under each product sign $\prod$ commute with each other so that the expression above makes unique sense. This proves that the polynomial $W_{\mathcal{I}_m}(\alpha,\beta,\gamma)$ is not identically constant. This completes the proof. □

LEMMA 3. *Let $|\mathcal{I}_m| = n$ be a multiindex. Then*

$$Leb_3\big\{\Phi_{\mathcal{I}_m}(D,C)\big\} \le Leb_3\{\Phi^{\alpha}_{\mathcal{I}_m}(D)\} + 4D^{-n^2/3}n^4. \tag{17}$$

Proof. In the complement to the set $\Phi^{\alpha}_{\mathcal{I}_m}(D)$ we have the estimates

$$\big\|W_{\mathcal{I}_m}(\alpha,\beta,\gamma)'_{\alpha}\big\|^2 \ge D^{-2n^2/3}\quad\text{and}\quad\big\|W_{\mathcal{I}_m}(\alpha,\beta,\gamma)''_{\alpha\alpha}\big\|^2 \le n^4. \tag{18}$$

Recall that $(\alpha,\beta,\gamma)\in\mathbb{T}_{\alpha}\times\mathbb{T}_{\beta}\times\mathbb{T}_{\gamma}$ as in (4). For each pair $(\beta,\gamma)\in\mathbb{T}_{\beta}\times\mathbb{T}_{\gamma}$ split the circle $\mathbb{T}_{\alpha}$ into $2D^{n^2/3}n^4$ intervals of equal length. Choose one interval and denote it by I. If there is a point in $(\alpha^*,\beta,\gamma)\in I$ which belongs to the complement of the set $\Phi^{\alpha}_{\mathcal{I}_m}(D)$, then by the Taylor formula along with (18) we have for each point in I

$$\big\|W_{\mathcal{I}_m}(\alpha,\beta,\gamma)'_{\alpha}\big\|^2 \ge \tfrac{D^{-n^2/3}}{2}\,. \tag{19}$$

Therefore, the Taylor formula implies that measure of $\alpha\in I\setminus\Phi^{\alpha}_{\mathcal{I}_m}(D)$ such that

$$\big\|W_{\mathcal{I}_m}(\alpha,\beta,\gamma) - C\big\| \le D^{-n^2} \tag{20}$$

is at most $2D^{-2n^2/3}$. Collecting all segments I and applying the Fubini theorem we complete the proof. □

Denote

$$\Phi^{\alpha}_{n}(D) = \cup_{\mathcal{I}_m\in\mathcal{R}_n}\Phi^{\alpha}_{\mathcal{I}_m}(D)\,. \tag{21}$$

Lemma 3 allows one to reduce the proof of (9) to the proof of

$$\sum_{n=1}^{\infty} Leb_3\{\Phi^{\alpha}_{n}(D^*)\} < \infty\,. \tag{22}$$

We prove the convergence next. First, we reformulate the problem.

LEMMA 4. *Any word $W_{\mathcal{I}_m}$ of length n ($|\mathcal{I}_m| = n$) corresponds to a polynomial $P_{\mathcal{I}_m}(x_\alpha, y_\alpha, x_\beta, y_\beta, x_\gamma, y_\gamma)$ of degree $2n + 2m$ with integer coefficients such that*

$$\left\|W_{\mathcal{I}_m}(\alpha,\beta,\gamma)'_\alpha\right\|^2 = P_{\mathcal{I}_m}(\cos\alpha, \sin\alpha, \cos\beta, \sin\beta, \cos\gamma, \sin\gamma)\,. \tag{23}$$

Proof. Consider the quaternion representation (13), differentiate it with respect to α, and take the sum of squares of its components. Then express $\cos s_p\alpha$ and $\sin s_p\alpha$ (resp. $\cos r_p\beta$ and $\sin r_p\beta$) as polynomials in $\cos\alpha$ and $\sin\alpha$ (resp. $\cos\beta$ and $\sin\beta$). Since all of the above operations are conducted with integer-coefficient trigonometric expressions, the result is a polynomial $P_{\mathcal{I}_m}(\cos\alpha, \sin\alpha, \cos\beta, \sin\beta, \cos\gamma, \sin\gamma)$ with integer coefficients. □

The main idea of our argument is that *a polynomial with integer coefficients cannot be small on a set of large measure*. In our notation for $|\mathcal{I}_m| = n$,

$$\Phi^\alpha_{\mathcal{I}_m}(D) = \big\{(\alpha,\beta,\gamma) \in \mathbb{T}^3 : \\ P_{\mathcal{I}_m}(\cos\alpha, \sin\alpha, \cos\beta, \sin\beta, \cos\gamma, \sin\gamma) \le D^{-2n^2/3}\big\}\,.$$

PROPOSITION 1. *There exists a positive constant D^* such that such that for all multiindexes $\mathcal{I}_m$ of a sufficiently large length $n = |\mathcal{I}_m|$, the measure of the set $\Phi^\alpha_{\mathcal{I}_m}(D^*)$ obeys the estimate*

$$Leb_3\big\{\Phi^\alpha_{\mathcal{I}_m}(D^*)\big\} \le 5^{-n}\,. \tag{24}$$

Since the number of words of length n is at most 4^n, Proposition 1 immediately implies (22) and thus (9). As we have mentioned before, estimate (9) leads to the conclusion of Theorems 2 and 1.

The proof of Proposition 1 requires two additional technical lemmas.

3 Estimates for the Polynomials in One Variable and the Elimination of Variables

The first result is the estimate on the size of the set where a polynomial in one variable is small.

LEMMA 5 ([DM], [KlM]). *Let $F(x)$ be a polynomial of degree $\le n$. Denote $\|F\|_B := \max_{x\in B}|F(x)|$. Then for any open interval B*

$$Leb_1\big\{x \in B : |F(x)| \le \epsilon\big\} \le 2n(n+1)^{1/n}\left(\tfrac{\epsilon}{\|F\|_B}\right)^{1/n}|\{B\}|\,.$$

Lemma 5 is proved in the paper of S. Dani and G. Margulis [DM]. For more general statements in this direction see also Kleinbock–Margulis [KlM].

There are several technical difficulties that complicate matters in our setup. We need to show that a certain polynomial in several variables does not spend too much time in the neighborhood of zero. In addition, the polynomial in question is trigonometric which means that some of the variables are dependent. To resolve the latter we apply the procedure known as the elimination of variables [Mu] along with quantitative estimates described in Lemma 6 below.

LEMMA 6. *Let $P(x,y,u,v) = \sum_{k=0}^{r} p_k(x,u,v)\, y^k$. Assume that the coefficients $p_k(x,u,v)$ are polynomials of (x,u,v) of degree $\le s$ with respect to each variable:* $\deg_{x,u,v} p_k \le s$.

Assume also that the following estimates are true for some constant $H \ge 1$:

$$\max_{x,u,v\in[-1,1]} \bigl|p_k(x,u,v)\bigr| \le H\,, \quad \forall k = 0,\dots,r\,. \tag{25}$$

Then there exists a polynomial $R_\epsilon(x,u,v)$ (which is, in fact, the so-called resultant of the polynomials $P(x,y,u,v) - \epsilon$ and $y^2 + x^2 - 1$) in variables (x,u,v) such that

1) *If for some y the polynomials $P(x,y,u,v) - \epsilon = y^2 + x^2 - 1 = 0$, then $R_\epsilon(x,u,v) = 0$.*
2) *The polynomial R_ε can be decomposed into 3 polynomials independent of ε : $R_\epsilon = R + \epsilon Q_1 + \epsilon^2 Q_2$, and (here the polynomial $R(x,u,v)$ is the resultant of $P(x,y,u,v)$ and $y^2 + x^2 - 1$)*

$$\begin{aligned} &\max_{x,u,v\in[-1,1]} \bigl|R(x,u,v)\bigr| \le 2^r (rH)^2\,, \\ &\max_{x,u,v\in[-1,1]} \bigl|Q_i(x,u,v)\bigr| \le 2^r (rH)^{2-i}\,, \quad i = 1,2\,. \end{aligned} \tag{26}$$

In addition, if we define the following polynomial of (u,v):

$$P_1(u,v) := \bigl(4(r+s)\bigr)! \int_{-1}^{1} \bigl(R(x,u,v)\bigr)^2 dx\,. \tag{27}$$

Then

3) $\deg_{u,v} P_1 \le 4(r+s)$.

Writing the polynomial $P_1(u,v)$ in the form $P_1(u,v) = \sum_{k=0}^{4(r+s)} p_{1k}(u) v^k$, we also have the estimates

$$\max_{u\in[-1,1]} |p_{1k}(u)| \le \frac{\bigl((4(r+s)-k)!\bigr)^3}{k!} 4^{r+1} (rH)^4\,, \quad \forall k = 0,\dots,4(r+s)\,. \tag{28}$$

Proof. The polynomial $R_\epsilon(x,u,v)$ is constructed as the resultant of the the polynomials $P(x,y,u,v)-\epsilon$ and y^2+x^2-1. We define $R_\epsilon(x,u,v)=\det \mathbb{A}$, where the $(r+2)\times(r+2)$ matrix $\mathbb{A}$ has the form

$$\mathbb{A}:=\begin{pmatrix} 1 & 0 & x^2-1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & 1 & 0 & x^2-1 \\ p_r(\cdot) & \dots & \dots & \dots & p_0(\cdot)-\epsilon & 0 \\ 0 & p_r(\cdot) & \dots & \dots & \dots & p_0(\cdot)-\epsilon \end{pmatrix} \tag{29}$$

Any solution y of the system $P(x,y,u,v)-\epsilon=y^2+x^2-1=0$ produces a nontrivial kernel containing the vector $(y^{r+1},y^r,\dots,1)$ of the matrix $\mathbb{A}$. Therefore, if for fixed (x,u,v) such a y exists, the resultant $R_\epsilon(x,u,v)$ vanishes.

In general, the resultant $R[P_1,P_2](x)$ of two polynomials $P_1(x,y)$ and $P_2(x,y)$ is a useful algebraic tool allowing to reduce the question of the existence of common y roots of $P_1(x,y)$ and $P_2(x,y)$ for a fixed x_0 to the vanishing of the resultant $R[P_1,P_2](x_0)=0$.

The estimate (28) is the only nontrivial remaining statement of this lemma. Its proof is based on the application of the Markov inequality:

$$\max_{x\in[-1,1]}|F'(x)|\le n^2\max_{x\in[-1,1]}|F(x)|$$

which holds for any polynomial F of degree n. It easily follows from the first inequality in (26) and (27) that

$$\max_{u,v\in[-1,1]}\big|P_1(u,v)\big|\le\big(4(r+s)\big)!\,4^{r+1}(rH)^4\,.$$

The coefficients $p_{1k}(u)$ can be found from the identity

$$p_{1k}(u)=\tfrac{1}{k!}\tfrac{d^k}{dv^k}P_1(u,0)$$

Using Markov's inequality for the polynomial $P_1(u,v)$ of degree $4(r+s)$ k times we conclude that

$$\big|p_{1k}(u)\big|\le\frac{((4(r+s)-k)!)^3}{k!}4^{r+1}(rH)^4\,,\quad\forall k=0,\dots,4(r+s)\,.\qquad\square$$

4 Proof of Proposition 1

We need to estimate the size of the set $\Phi^\alpha_{\mathcal{I}_m}(D)$ of the parameters $(\alpha,\beta,\gamma)\in T^3$ such that the polynomial $|P_{\mathcal{I}_m}(\cos\alpha,\sin\alpha,\cos\beta,\sin\beta,\cos\gamma,\sin\gamma)|\le$

$D^{-2n^2/3}$. The above set has essentially the same measure as the set

$$\mathbb{K} := \big\{(x_\alpha, x_\beta, x_\gamma) \in [-1,1]^3 : P_{\mathcal{I}_m}(x_\alpha, y_\alpha, x_\beta, y_\beta, x_\gamma, y_\gamma) - \epsilon = 0\,,$$
$$y_\alpha^2 + x_\alpha^2 - 1 = 0\,, \quad y_\beta^2 + x_\beta^2 - 1 = 0\,, \quad y_\gamma^2 + x_\gamma^2 - 1 = 0\,,$$
$$\text{for some } \ \epsilon \le D^{-n^2/3}\big\}\,.$$

We will apply the process of the elimination of variables and Lemma 5 three times in a row. First, let us list the properties of the polynomial $P_{\mathcal{I}_m}(x_\alpha, y_\alpha, x_\beta, y_\beta, x_\gamma, y_\gamma)$.

1) $\deg_{x_\alpha, y_\alpha, x_\beta, y_\beta, x_\gamma, y_\gamma} P \le 2n$.
2) $P_{\mathcal{I}_m}(x_\alpha, y_\alpha, x_\beta, y_\beta, x_\gamma, y_\gamma) = \sum_{k=0}^{n} p_k(x_\alpha, x_\beta, y_\beta, x_\gamma, y_\gamma) y_\alpha^k$,
3) $\max_{(x_\alpha, x_\beta, y_\beta, x_\gamma, y_\gamma)} |p_k(x_\alpha, x_\beta, y_\beta, x_\gamma, y_\gamma)| \le H := (2^n\, n)^2$, $\forall k = 0,...,n$.

Apply Lemma 6 to the polynomials $\sum_{k=0}^{n} p_k(x_\alpha, x_\beta, y_\beta, x_\gamma, y_\gamma) y_\alpha^k$ and $y_\alpha^2 + x_\alpha^2 - 1$ with $s = r = 2n$ and $H = (2^n n)^2$. From the properties of the resultant $R_\epsilon(x_\alpha, x_\beta, y_\beta, x_\gamma, y_\gamma)$ defined in Lemma 6 it follows that

$$\mathbb{K} \subset \big\{(x_\alpha, x_\beta, x_\gamma) \in [-1,1]^3 : R_\epsilon(x_\alpha, x_\beta, y_\beta, x_\gamma, y_\gamma) = 0\,,$$
$$y_\beta^2 + x_\beta^2 - 1 = 0\,,\ y_\gamma^2 + x_\gamma^2 - 1 = 0\,,\ \text{for some } \epsilon \le D^{-n^2/3}\big\}\,. \quad (30)$$

Using estimates (26) we conclude that

$$\mathbb{K} \subset \big\{(x_\alpha, x_\beta, x_\gamma) \in [-1,1]^3 : R(x_\alpha, x_\beta, y_\beta, x_\gamma, y_\gamma) \le \delta\,,$$
$$y_\alpha^2 + x_\alpha^2 - 1 = 0\,,\ y_\gamma^2 + x_\gamma^2 - 1 = 0\big\} \quad (31)$$
$$\delta := D^{-n^2/3}\big(2^{2n}(2n\,H) + 2^{2n}(2n\,H)^2 D^{-n^2/3}\big)\,.$$

Observe that for $n \ge 20$ and with a choice of $D \ge 5$, the new small parameter $\delta \le D^{-n^2/6}$.

Fix $(x_\beta, y_\beta, x_\gamma, y_\gamma)$ satisfying $y_\beta^2 + x_\beta^2 - 1 = 0$, $y_\gamma^2 + x_\gamma^2 - 1 = 0$, and apply Lemma 5 to the polynomial $R(x_\alpha, x_\beta, y_\beta, x_\gamma, y_\gamma)$ with respect to x_α. For fixed x_β, x_γ let

$$\mathbb{K}_{x_\beta, x_\gamma} := \big\{x_\alpha \in [-1,1] : (x_\alpha, x_\beta, x_\gamma) \in \mathbb{K}\big\}\,.$$

Lemma 5 then implies that

$$Leb_1\{\mathbb{K}_{x_\beta, x_\gamma}\} \le 16n(8n+1)^{1/8n} \left(\frac{\delta}{\|R(\cdot, x_\beta, y_\beta, x_\gamma, y_\gamma)\|}\right)^{1/8n}. \quad (32)$$

Note that Leb_1 and Leb_2 denote one and two-dimensional Lebesgue measures correspondingly.

Define

$$\mathbb{K}^1 := \big\{(x_\beta, x_\gamma) \in [-1,1]^2 : \|R(\cdot, x_\beta, y_\beta, x_\gamma, y_\gamma)\| \le \delta^{1/2}, \\ y_\beta^2 + x_\beta^2 - 1 = 0\,,\ y_\gamma^2 + x_\gamma^2 - 1 = 0\big\}\,.$$

The Fubini Theorem implies that

$$Leb_3\{\mathbb{K}\} \le 2Leb_2\{\mathbb{K}^1\} + Leb_3\Big\{\bigcup_{(x_\beta,x_\gamma)\notin\mathbb{K}^1} \mathbb{K}_{x_\beta,x_\gamma}\Big\}\,. \tag{33}$$

Observe also that by the Fubini Theorem and (32) the set $\bigcup_{(x_\beta,x_\gamma)\notin\mathbb{K}^1} \mathbb{K}_{x_\beta,x_\gamma}$ obeys the following estimate on its size:

$$Leb_3\Big\{\bigcup_{(x_\beta,x_\gamma)\notin\mathbb{K}^1} \mathbb{K}_{x_\beta,x_\gamma}\Big\} \le 64n(8n+1)^{1/8n}\delta^{1/16n}. \tag{34}$$

To estimate the size of the set $\mathbb{K}^1$ we employ the conclusions of the second part of Lemma 6. Define $P_{1,\mathcal{I}_m}(x_\beta, y_\beta, x_\gamma, y_\gamma)$ from the resultant $R(x_\alpha, x_\beta, y_\beta, x_\gamma, y_\gamma)$ as in (27):

$$P_{1,\mathcal{I}_m}(x_\beta, y_\beta, x_\gamma, y_\gamma) := (16n)! \int_{-1}^{1} \big|R(x_\alpha, x_\beta, y_\beta, x_\gamma, y_\gamma)\big|^2 dx_\alpha\,.$$

The constant in front of the integral is introduced so that the resulting polynomial is still a polynomial with integer coefficients. Clearly,

$$\mathbb{K}^1 \subset \big\{(x_\beta, x_\gamma) \in [-1,1]^2 : |P_{1,\mathcal{I}_m}(x_\beta, y_\beta, x_\gamma, y_\gamma)| \le 2(16n)!\delta\,, \\ y_\beta^2 + x_\beta^2 - 1 = 0\,,\ y_\gamma^2 + x_\gamma^2 - 1 = 0\big\}\,. \tag{35}$$

Combining estimates (33), (34), and (35) we conclude that for $n \ge 20$ with a choice of $D \ge 10^{100}$ we have

$$Leb_3\{\Phi^\alpha_{\mathcal{I}_m}(D)\} \le 2\,Leb_2\big\{(x_\beta, x_\gamma) \in [-1,1]^2 :\, |P_{1,\mathcal{I}_m}(x_\beta, y_\beta, x_\gamma, y_\gamma)| \\ \le D^{-n^2/10}\,,\ y_\beta^2 + x_\beta^2 - 1 = 0\,,\ y_\gamma^2 + x_\gamma^2 - 1 = 0\big\} + 10^{-n}.$$

The problem is now reduced to a similar two-dimensional question. We are in position to apply another round of Lemma 6 and Lemma 5. Reiterate the arguments above for the polynomial $P_{1,\mathcal{I}_m}(x_\beta, y_\beta, x_\gamma, y_\gamma)$ with the following properties:

1) $\deg_{x_\beta,y_\beta,x_\gamma,y_\gamma} P_{1,\mathcal{I}_m} \le 16n$.
2) $P_{1,\mathcal{I}_m}(x_\beta, y_\beta, x_\gamma, y_\gamma) = \sum_{k=0}^{16n} p_{1k}(x_\beta, x_\gamma, y_\gamma) y_\beta^k$,
3) $\max_{x_\beta,x_\gamma,y_\gamma\in[-1,1]^3} |p_{1k}(x_\beta, x_\gamma, y_\gamma)| \le H_1 := ((16n)!)^3 4^{2n+1}(2nH)^4$.

Define the resultant $R_1(x_\beta,x_\gamma,y_\gamma)$ of the polynomials $P_{1,\mathcal{I}_m}(x_\beta,y_\beta,x_\gamma,y_\gamma)$ and $y_\beta^2+x_\beta^2-1$. We obtain

$$\mathbb{K}_{x_\gamma} := \big\{x_\beta \in [-1,1] : (x_\beta, x_\gamma) \in \mathbb{K}^1\big\},$$

$$Leb_2\big\{(x_\beta,x_\gamma)\in[-1,1]^2 : |P_{1,\mathcal{I}_m}(x_\beta,y_\beta,x_\gamma,y_\gamma)| \le D^{-n^2/10}\,,$$
$$y_\beta^2+x_\beta^2-1=0\,,\ y_\gamma^2+x_\gamma^2-1=0\big\} \le 2\,Leb_1\{\mathbb{K}^2\} + Leb_2\Big\{\bigcup_{x_\gamma\notin\mathbb{K}^2}\mathbb{K}_{x_\gamma}\Big\},$$

$$\mathbb{K}^2 := \big\{x_\gamma\in[-1,1] : \|R_1(\cdot,x_\gamma,y_\gamma)\| \le \delta_1^{1/2}\,,\ y_\gamma^2+x_\gamma^2-1=0\big\},$$

$$\delta_1 := D^{-n^2/10}\big(2^{32n}16n\,H_1 + 2^{32n}(16n\,H_1)^2D^{-n^2/10}\big),$$

$$Leb_2\Big\{\bigcup_{x_\gamma\notin\mathbb{K}^2}\mathbb{K}_{x_\gamma}\Big\} \le 4(16n)\big(4(8n)+1\big)^{1/64n}\delta_1^{1/64n}.$$

Observe that for all $n\ge 20$ and $D\ge 10^{100}$, the small parameter $\delta_1 \le D^{-n^2/20}$. Hence, choosing constant $D\ge 10^{10^3}$ we obtain

$$Leb_3\{\Phi^\alpha_{\mathcal{I}_m}(D)\} \le 2\,Leb_1\big\{x_\gamma\in[-1,1] : |P_{2,\mathcal{I}_m}(x_\gamma,y_\gamma)| \le D^{-n^2/30}\,,$$
$$y_\gamma^2+x_\gamma^2-1=0\big\} + 10^{-n} + 10^{-n}\,, \quad (36)$$

where the polynomial $P_{2,\mathcal{I}_m}(x_\gamma,y_\gamma)$ is formed from the resultant $R_1(x_\beta,x_\gamma,y_\gamma)$ as in (27). Finally, we can eliminate y_γ with the help of Lemma 6 applied to the polynomials $P_{2,\mathcal{I}_m}(x_\gamma,y_\gamma)$ and $y_\gamma^2+x_\gamma^2-1$. As a consequence, we obtain the resultant $R_2(x_\gamma)$ of the above polynomials. Then applying Lemma 5 we can find a positive constant $\delta_2 \le D^{-n^2/50}$ such that

$$Leb_3\{\Phi^\alpha_{\mathcal{I}_m}(D)\} \le \Big(\tfrac{\delta_2}{\|R_2(\cdot)\|}\Big)^{1/512n} + 10^{-n} + 10^{-n}\,. \quad (37)$$

The resultant $R_2(x_\gamma)$ is a polynomial with integer coefficients of degree at most $512n$. Therefore, $(1024n)!\int_{-1}^1 |R_2(x_\gamma)|^2dx_\gamma$ is a non-negative integer. If it is positive, the desired estimate immediately follows from (37) with a choice of constant $D^*=10^{10^4}$. To show that $R_2(x_\gamma)$ is not identically 0 we argue as follows.

The polynomial $R_2(x_\gamma)$ was obtained via combination of the elimination of variables (forming the resultant) and the integration as in (27). Certainly, integration cannot produce the identically zero polynomial from a nonzero one. Therefore, we need to justify the "non-degeneracy" of the elimination. The basic property of the resultant $R[P_1,P_2](x)$ of two polynomials $P_1(x,y)$ and $P_2(x,y)$ is that $R[P_1,P_2](x_0)$ equals 0 if and only if for some $y\in\mathbb{C}$ we have $P_1(x_0,y)=P_2(x_0,y)=0$ ([Mu, p. 34]). In our case one of polynomials,

say P_2, is x^2+y^2-1. If $R[P_1,P_2](x)\equiv 0$, then $x=\cos\alpha, y=\sin\alpha$, and $P_1(\cos\alpha,\sin\alpha)$ vanishes on the open set $\alpha\in U\subset\mathbb{R}$, which implies that P_1 is identically zero. This is in contradiction with the non-degeneracy of $W_{\mathcal{I}_m}(\alpha,\beta,\gamma)$ (see Lemma 2).

5 Degenerate Words

In this section for $n=4^m\in Z_+$ we construct a collection of words $W_n(\alpha,\beta,\gamma)$ of length n such that the polynomial $P_n(\alpha,\beta,\gamma)$ describing the distance of $W_n(\alpha,\beta,\gamma)$ to E, as defined above, has a zero of order $\sqrt{n}$ with respect to α at any point of the form $(0,\beta,\gamma)$.

Recall that $W_n(\alpha,\beta,\gamma)=W_n(A,B)$ is a word in A and B, defined by the angle of rotation α of A, the angle of rotation β of B and the angle γ between the axis of rotations of A and B (see the introduction). Denote by $[A,B]=ABA^{-1}B^{-1}$ the commutant formed by A and B. The idea of the construction is the following remark: For a sufficiently small α the angle of rotation of the commutant $[A,B]$ is of order at most α^2. This can be seen, for example, from the quaternion representation (11).

Consider two rotations $A,B\in SO(3)$. Define the maps

$$\phi:\begin{pmatrix}A\\B\end{pmatrix}\mapsto\begin{pmatrix}[A,B]\\[B,A^{-1}]\end{pmatrix},\qquad \psi:\begin{pmatrix}A\\B\end{pmatrix}\mapsto\begin{pmatrix}[A^{-1},B^{-1}]\\[B^{-1},A]\end{pmatrix},\tag{38}$$

which map a pair of rotations into a pair of commutant words. We take all possible compositions (iterations) $\phi\circ\phi\circ\psi\ldots$ of the maps ϕ and ψ of total length k, thus creating 2^k words of length at most 4^k. Let us loosely denote by A_k,B_k the words obtained by the total of k iterations of ϕ and ψ.

Observe that A_1 and B_1 are rotations by the angle of the order at most α^2 provided that α is sufficiently small, A_2 and B_2 are rotations by the angle of the order at most α^4, and A_k and B_k are rotations by the angle of the order at most α^{2^k}. Therefore, with $n=4^k$, we have constructed a collection of words with the degree of degeneration $\sqrt{n}$.

We can also add that it seems possible to choose a composition of the maps ϕ,ψ in such a way that the corresponding word $W_n(A,B)$ has a zero of order $n^{(\rho+1)/2}=2^{2(\rho+1)k}$ (with the constant $\rho=(\sqrt{5}-1)/2$ the golden mean) at the point $(\alpha,\beta,\gamma)=(0,\beta,\gamma)$, thus constructing a word with even a higher degree of degeneration.

All degenerations described here occur in a neighborhood of zero. It is an interesting question whether there are zeroes of high order far away from the identity element in $SO(3)$.

Acknowledgments We would like to thank Peter Sarnak for drawing our attention to the problem, stimulating discussions, and encouragement. We also thank Dmitri Jakobson and the referee for valuable suggestions.

References

[Au] H. Auerbach, Sur grooupes lineares bornes (III), Studia Math. 5 (1934), 43–49.

[DM] S. Dani, G. Margulis, Limit distributions of orbits of unipotent flows and values of quadratic forms, I.M. Gelfand Seminar, Adv. Soviet Math. 16, Part 1, Amer. Math. Soc., Providence, RI (1993), 91–137.

[Do] D. Dolgopyat, On mixing properties of compact group extensions of hyperbolic systems, Israel J. Math., to appear.

[Dr] V. Drinfeld, Finitely additive measures on $\mathbb{S}^2$ and $\mathbb{S}^3$, invariant with respect to rotations, Func. Anal. and its Appl. 18 (1984), 245–246.

[HK] B. Hasselblatt, A. Katok, Introduction to the Modern Theory of Dynamical Systems, Encyclopedia of Math and its App. 54, 802pp., Cambridge University Press, Cambridge, 1995. PAGENOS?

[Ha] F. Hausdorff, Grunzuge der Mengenlehre, Leipzig, 1914.

[He] M. Herman, Sur les courbes invariantes par les difféomorphismes de l'anneau, vol. I & II, Asterisque 103–104 (1983); 144 (1986).

[HuSY] B. Hunt, T. Sauer, J. Yorke, Prevalence: a translation-invariant "almost every" on infinite-dimensional spaces, Bull. Amer. Math. Soc. 27:2 (1992), 217–238; 28:2 (1993), 306–307.

[GJS] A. Gamburd D. Jakobson, P., Sarnak, Spectra of elements in the group ring of $SU(2)$, J. Eur. Math. Soc. (JEMS) 1:1 (1999), 51–85.

[K] V. Kaloshin, Some prevalent properties of smooth dynamical systems, Proc. of Steklov Math. Inst. 213 (1997), 123–151.

[Kh] A. Khintchine, Continued Fractions (translated by P. Wynn), P. Noordhoff Ltd., Groningen, 1963.

[KlM] D. Kleinbock, G. Margulis, Flows on homogeneous spaces and Diophantine approximation on manifolds, Ann. of Math. 148:1 (1998), 339–360.

[L] V. Lazutkin, KAM Theory and Semiclassical Approximations to Eigenfunctions, Ergeb. Math. Grenzgeb. (3) 24, Springer-Verlag, Berlin, 1993.

[Lu] A. Lubotzky, Discrete Groups, Expanding Graphs and Invariant Measures, Progress in Mathematics 125, Birkhäuser, Basel, 1994.

[LuPS1] A. Lubotzky, R. Phillips, P. Sarnak, Hecke operators and distributing points on the sphere, I, Comm. Pure Appl. Math. 39, suppl. (1986), 149–186.

[LuPS2] A. Lubotzky, R. Phillips, P. Sarnak, Hecke operators and distributing points on the sphere, II, Comm. Pure Appl. Math. 40:4 (1987), 401–420.

[M] G. Margulis, Some remarks on invariant means, Monatschefte fur Mathematik 90 (1980), 233–235.

[Mo] J. Moser, Stable and Random Motion in Dynamical Systems, Princeton University Press, Princeton, 1973.

[Mu] D. Mumford, Algebraic Geometry I, Complex Projective Varieties, Springer-Verlag, New York, 1976.

[O] J. Oxtoby, Measure and Category, Grad. Texts in Math. 2, Springer-Verlag, New York-Berlin, 1980.

[R] J. Rosenblatt, Uniqueness of invariant means for measure preserving transformations, Trans. AMS 265 (1981), 623–636.

[S] D. Sullivan, For $n > 3$ there is only one finitely additive rotationally invariant measure on the n-sphere on all Lebesgue measurable sets, Bull. AMS 1 (1981), 121–123.

[W] S. Wagon, The Banach–Tarski Paradox, Cambridge University Press, Cambridge, 1993.

V. Kaloshin, Department of Mathematics, Princeton University, Princeton, NJ 08544, USA kaloshin@math.princeton.edu

I. Rodnianski, Department of Mathematics, Princeton University, Princeton, NJ 08544, USA irod@math.princeton.edu

Submitted: June 2000
Revised version: January 2001