DYNAMICS AND LITTLEWOOD'S CONJECTURE

SHIN KIM

Abstract. We consider the action of the group, A, of positive diagonal matrices on the space $X = SL(3,\mathbb{R})/SL(3,\mathbb{Z})$. Manfred Einsiedler, Anatole Katok, and Elon Lindenstrauss showed that if μ is an A-invariant and ergodic probability measure on X and there exists a one-parameter subgroup of A acting on X with positive entropy, then μ must be an algebraic measure. This result is used to conclude that the set of counterexamples to Littlewood's conjucture has Hausdorff dimension zero if the set is nonempty. We will discuss some parts of their work.

Contents

1. Littlewood's conjecture

For any real number x, we let $\|x\| = \min\{x - p : p \in \mathbb{Z}\}$ be the distance between x and the integer closest to x. The following statement about simultaneous approximation by rationals is known as Littlewood's conjecture.

Conjecture 1.1. *For any real numbers α and β,*

$$\liminf_{q\to\infty} q\|q\alpha\|\|q\beta\| = 0.$$

To provide some context for this conjecture, we review some basic results in Diophantine approximation. Let α be any real number. We denote the greatest integer less than or equal to α as $\lfloor\alpha\rfloor$. The fractional part of α is defined by $\{\alpha\} = \alpha - \lfloor\alpha\rfloor$. The theorem below is known as Dirichlet's theorem.

Theorem 1.2. *Let Q be any integer greater than 1. Then there exist integers p and q such that $1 \le q < Q$ and $\mid \alpha q - p \mid\le \frac{1}{Q}$.*

Date: May 30, 2018.

Proof. The theorem holds whenever α is an integer or whenever α is a rational number whose denominator is less than Q when α is expressed in the lowest terms. So, we assume that α is either an irrational number, or a rational number whose lowest possible denominator is greater than or equal to Q. Then, the $Q+1$ numbers

$$0, 1, \{\alpha\}, \{2\alpha\}, \ldots, \{(Q-1)\alpha\}$$

are distinct numbers in the interval $[0,1]$. We divide the interval $[0,1]$ into Q intervals given by

$$I_i = \begin{cases} [\frac{i-1}{Q}, \frac{i}{Q}) & \text{if } i < Q \\ [\frac{Q-1}{Q}, 1] & \text{if } i = Q. \end{cases}$$

By the pigeonhole principle, one of the Q subintervals contains at least two of the $Q+1$ numbers. It follows that, there exist integers r_1, r_2, s_1 ands_2 such that $0 \leq r_1, r_2 < Q$, $r_1 > r_2$, and

$$\mid (r_1\alpha - s_1) - (r_2\alpha - s_2) \mid \leq \frac{1}{Q}.$$

Setting $q = r_1 - r_2$ and $p = s_1 - s_2$, we see that $1 \leq q < Q$ and $\mid q\alpha - p \mid \leq \frac{1}{Q}$. □

Corollary 1.3. *Suppose that α is irrational. Then, there exist infinitely many pairs p and q of relatively prime integers with $\mid \alpha - \frac{p}{q} \mid < \frac{1}{q^2}$*

In particular, Corollary 1.3 implies that

$$\liminf_{q \to \infty} q\|q\alpha\| \leq 1$$

for any irrational number α. However, it is not true that there are infinitely many distinct rationals $\frac{p}{q}$ such that $\mid \alpha - \frac{p}{q} \mid < \frac{c}{q^2}$ for an arbitrary fixed constant $c > 0$.

Lemma 1.4. *Let α be a real quadratic irrational that is a root of a non-zero polynomial, $P(T) = aT^2 + bT + c$, with integer coefficients. Let D be the discriminant. Then, for any $A > \sqrt{D}$, the inequality*

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{Aq^2}$$

has only finitely many solutions.

Proof. Suppose that $\frac{p}{q}$ is a rational number such that $\mid \alpha - \frac{p}{q} \mid < \frac{1}{Aq^2}$. Then,

$$\begin{aligned} \left|P\left(\frac{p}{q}\right)\right| &= \left|a\left(\frac{p}{q}\right)^2 + b\left(\frac{p}{q}\right) + c\right| \\ &= \frac{|ap^2 + bpq + cq^2|}{q^2} \\ &\geq \frac{1}{q^2}, \end{aligned}$$

because $P(\frac{p}{q}) \neq 0$ and $ap^2+bpq+cq^2$ is an integer. Write $P(T) = a(T-\alpha)(T-\alpha')$. Then the discrfiminant of $P(T)$ is $D = a^2(\alpha-\alpha')^2$. We observe that

$$\begin{aligned}
\Big|P\Big(\frac{p}{q}\Big)\Big| &= \Big|a\Big(\frac{p}{q}-\alpha\Big)\Big(\frac{p}{q}-\alpha'\Big)\Big| \\
&\leq \frac{1}{Aq^2}\Big|a\Big(\frac{p}{q}-\alpha'\Big)\Big| \\
&= \frac{1}{Aq^2}\Big|a\Big(\alpha'-\alpha+\alpha-\frac{p}{q}\Big)\Big| \\
&\leq \frac{1}{Aq^2}|a(\alpha'-\alpha)| + \frac{1}{Aq^2}\Big|a\Big(\alpha-\frac{p}{q}\Big)\Big| \\
&\leq \frac{\sqrt{D}}{Aq^2} + \frac{|a|}{A^2q^4}
\end{aligned}$$

Note that $(\frac{\sqrt{D}}{Aq^2} + \frac{|a|}{A^2q^4})/(\frac{1}{q^2}) = \frac{\sqrt{D}}{A} + \frac{|a|}{A^2q^2}$ converges to $\frac{\sqrt{D}}{A} < 1$ as q approaches ∞. Thus, there exists an integer N, which depends only on D, A, and a, so that q has to be less than N. $\square$

Let $x = \frac{1}{2}(\sqrt{5}-1)$. Then, x is a root of the polynomial $Q(T) = T^2+T-1$, and the disciriminant of $Q(T)$ is 5. So, the lemma above implies that if $A > \sqrt{5}$, the set

$$\{\frac{p}{q} \in \mathbb{Q} : \Big|x - \frac{p}{q}\Big| < \frac{1}{Aq^2}\}$$

is finite. This proves the following theorem, due to Hurwitz.

Theorem 1.5. *If $c < \frac{1}{\sqrt{5}}$, then there exists an irrational number α such that there are at most finitely many distinct rationals $\frac{p}{q}$ that satisfy*

$$\Big|\alpha - \frac{p}{q}\Big| < \frac{c}{q^2}.$$

Since $x = \frac{1}{2}(\sqrt{5}-1)$ is an irrational number, our observation above implies that there exists some constant $r > 0$, which depends on x, such that

$$\Big|x - \frac{p}{q}\Big| > \frac{r}{q^2},$$

for all $\frac{p}{q} \in \mathbb{Q}$. Equivalently,

$$\|qx\| > \frac{r}{q},$$

for all $q \geq 1$. We say that a number such as x is *badly approximable* by rationals and we define

$$\mathbf{Bad} = \{x \in \mathbb{R} : \text{ there exists } r = r(x) > 0 \text{ with } \|qx\| > \frac{r}{q} \text{ for all } q \geq 1\}$$

to be the set of all badly approximable numbers.

If $x \in \mathbf{Bad}$, then $\liminf_{q\to\infty} q\|qx\| > 0$. On the other hand, if x is an irrational number and $x \notin \mathbf{Bad}$, then we can find a sequence of distinct naturals (q_n) such that $q_n\|q_n x\|$ converges to 0. Thus, the set of counterexamples to the one dimensional analogue of Littlewood's conjecture is precisely $\mathbf{Bad}$. Then, we can interpret Littlewood's conjecture as saying that any two real numbers can be simultaneously approximated "well" by two rational numbers with the same denominator.

In this paper, we will discuss how results from dynamics imply that the set of counterexamples to Littlewood's conjecture is "small". To begin, we introduce some notions from dynamics that we will use.

2. Measure theoretic entropy and topological entropy

Let (Y, μ) be a probability space, and let $T : Y \to Y$ be a measurable map. We say that μ is *T-invariant* if $\mu(T^{-1}(E)) = \mu(E)$ for any measurable set $E \subseteq Y$. A T-invariant measure μ is *ergodic* if any measurable function $f : Y \to \mathbb{R}$ such that $f(T(y)) = f(y)$ for μ-a.e. $y \in Y$ is constant almost everywhere.

Let μ be a T-invariant probability measure. A *finite partition* of Y is a finite collection,

$$\mathcal{P} = \{C_1, \dots, C_n\},$$

of essentially disjoint measurable sets such that $\bigcup_{i=1}^n C_i$ is a set of full measure in Y. Similarly, a *countable partition* of Y is a countable collection of essentially disjoint measurable sets whose union is a set of full measure in Y. For a partition $\mathcal{P}$ of Y, we define the *entropy of $\mathcal{P}$* by

$$H_\mu(\mathcal{P}) = -\sum_{C \in \mathcal{P}} \mu(C) \log(\mu(C)).$$

If $\mathcal{P}$ and $\mathcal{Q}$ are partitions of Y, then the *common refinement* of $\mathcal{P}$ and $\mathcal{Q}$ is defined by

$$\mathcal{P} \vee \mathcal{Q} = \{C \cap D : C \in \mathcal{P} \text{ and } D \in \mathcal{Q}\}.$$

Let $\mathcal{P}$ be a parition of Y with finite entropy. For each $n \in \mathbb{N}$,

$$T^{-n}\mathcal{P} = \{T^{-n}(C) : C \in \mathcal{P}\}$$

is a partition of Y. The *measure theoretic entropy of T relative to $\mathcal{P}$* is the limit

$$h_\mu(T, \mathcal{P}) = \lim_{N \to \infty} \frac{1}{N} H_\mu\left(\bigvee_{n=0}^{N-1} T^{-n}\mathcal{P}\right).$$

The two lemmas below show that this limit exists.

Lemma 2.1. $H_\mu(\mathcal{P} \vee \mathcal{Q}) \le H_\mu(\mathcal{P}) + H_\mu(\mathcal{Q})$.

Proof. See Theorem 4.3 of [14]. □

Lemma 2.2. *Let (a_n) be a sequence of non-negative real numbers such that*

$$0 \le a_{n+m} \le a_n + a_m$$

for all $n, m \ge 0$. Then,

$$\lim_{n \to \infty} \frac{a_n}{n} = \inf_{n \ge 0} \frac{a_n}{n}.$$

Proof. See Theorem 4.9 of [14]. □

The *measure theoretic entropy* is defined by

$$h_\mu(T) = \sup\{h_\mu(T, \mathcal{P}) : \mathcal{P} \text{ is a partition with } H_\mu(\mathcal{P}) < \infty\}.$$

Now, suppose that (Y, μ) is also a compact metric space with a metric d and that $T : Y \to Y$ is a continuous map. For each $n \in \mathbb{N}$, we define $d_n : Y \times Y \to [0, \infty)$ by

$$d_n(x, y) = \max_{0 \le k < n} d(T^k(x), T^k(y)).$$

Then, d_n is a metric on Y and $d_{n+1} \geq d_n$ for each $n \in \mathbb{N}$. Furthermore, d_n generates the same topology as d. To see this, let B be a basis element for the topology generated by d_n. Then, $B = \{x \in Y : d_n(x, y) < r\}$ for some $y \in Y$ and $r > 0$. Let $x' \in B$. Because T^k is continuous at x' for each $1 \leq k \leq n$, we can choose $r' > 0$ so that if $d(x, x') < r'$ then $d_n(x, x') < r - d_n(x', y)$. This implies that the open ball centered at x' with d-radius equal to r' is contained in B.

Fix $\epsilon > 0$. A subset $A \subseteq Y$ is *(n, ϵ)-separated* if $d_n(x, y) \geq \epsilon$ whenever x and y are any two distinct points in A. Any (n, ϵ)-separated set is finite because it is a discrete set in (Y, d_n) and (Y, d_n) is compact by our observation above. Define $sep(n, \epsilon, T)$ to be the maximum cardinality of an (n, ϵ)-separated set.

Define $cov(n, \epsilon, T)$ to be the minimum cardinality of a covering of Y by sets of d_n-diameter less than ϵ. Since we can cover Y by finitely many open balls with d_n-radius less than $\frac{\epsilon}{2}$, $cov(n, \epsilon, T)$ is a finite quantity.

Lemma 2.3. *For any $n \in \mathbb{N}$ and $\epsilon > 0$,*

$$cov(n, 2\epsilon, T) \leq sep(n, \epsilon, T) \leq cov(n, \epsilon, T).$$

Proof. Let $\{U_1, \ldots, U_k\}$ be a cover of Y by sets of d_n-diameter less than ϵ with $k = cov(n, \epsilon, T)$. Suppose that $A \subseteq Y$ is a finite subset with $|A| > k$. Then, there exists i with $1 \leq i \leq k$ and $a, a' \in A$ such that $a, a' \in U_i$. So, A cannot be (n, ϵ)-separated. Thus, $sep(n, \epsilon, T) \leq cov(n, \epsilon, T)$.

Now, let A be an (n, ϵ)-separated set with $|A| = sep(n, \epsilon, T)$. Then the collection of open balls $\{B(a, \epsilon)\}_{a \in A}$ is a covering of Y. Suppose that the collection does not cover Y. Then, there exists $z \in Y \setminus A$ such that $d_n(z, a) \geq \epsilon$ for all $a \in A$. This contradicts our assumption because $A \cup \{z\}$ is an (n, ϵ)-separated set that is strictly larger than A. Therefore, $cov(n, 2\epsilon, T) \leq sep(n, \epsilon, T)$. □

For each $\epsilon > 0$, we define

$$h_\epsilon(T) = \limsup_{n \to \infty} \frac{1}{n} \log(cov(n, \epsilon, T)).$$

Since $cov(n, \epsilon, T) \leq cov(n, \epsilon', T)$ whenever $\epsilon > \epsilon'$, $h_\epsilon(T)$ increases monotonically as ϵ approaches zero. We define the *topological entropy of* T to be the limit

$$h_{top}(T) = \lim_{\epsilon \to 0^+} h_\epsilon(T).$$

The lemma above implies that

$$h_{top}(T) = \lim_{\epsilon \to 0^+} \limsup_{n \to \infty} \frac{1}{n} \log(sep(n, \epsilon, T)).$$

As an aside, the topological entropy is indeed a "topological invariant" in the following sense.

Proposition 2.4. *$h_{top}(T)$ does not depend on the choice of a particular metric generating the topology of X.*

Proof. See Proposition 2.5.3 of [1]. □

There are a few important facts about entropy that we will need later.

Proposition 2.5. *Let (Y, d) be a compact metric space and let $T : Y \to Y$ be a continuous map. Then,*

$$h_{top}(T) = \sup\{h_\mu(T) : \mu \text{ is a } T\text{-invariant probability measure on } Y\}.$$

Proof. See Theorem 8.6 of [14]. □

Proposition 2.6. *Let Y be a compact metric space and let $T : Y \to Y$ be a continuous map. If μ and ν are T-invariant probability measure on Y and $p \in [0, 1]$, then*

$$h_{p\mu+(1-p)\nu}(T) = ph_\mu(T) + (1-p)h_\nu(T).$$

Proof. See Theorem 8.1 of [14]. □

Proposition 2.7. *Let Y be a compact metric space and let $T : Y \to Y$ be a continuous map. Let μ be a T-invariant probability measure on Y. Then, there exists a probability space (Ξ, ν) and a measurable map from Ξ to the space of T-invariant and ergodic probability meaures on Y given by $\xi \mapsto \mu_\xi$ such that*

$$\mu = \int_\Xi \mu_\xi d\nu(\xi).$$

Moreover,

$$h_\mu(T) = \int_\Xi h_{\mu_\xi}(T) d\nu(\xi).$$

Proof. For the first statement, see Theorem 6.10 and the following remark of [14]. For the second statement, see Theorem 8.4 of [14]. □

Remark 2.8. Let (Z, d) be a compact metric space and let $S : Z \to Z$ be a homeomorphism. We say that S is *expansive* if there exists $\delta > 0$ such that whenever x and y are distinct points in Z there exists $n \in \mathbb{Z}$ with $d(S^n(x), S^n(y)) > \delta$. Let $M(Z, S)$ denote the space of S-invariant probability measures on Z. We view $M(Z, S)$ as a topological subspace of the space of Radon measures on Z equipped with the weak-$*$ topology. Define $\Phi : M(Z, S) \to \mathbb{R}$ by $\Phi(\rho) = h_\rho(S)$. Theorem 8.2 of [14] states that if S is expansive then Φ is upper semi-continuous. This fact and Proposition 2.6 are used to prove Theorem 8.4 of [14]. A similar argument is used later on in this paper to construct a measure with positive entropy.

The notion of ergodicity also makes sense in a different setting. Let Z be a locally compact space, let ν be a probability measure on Z, and let H be a locally compact group. Suppose that H acts continuously on Z. We say that ν is H*-invariant* if $h_*\nu = \nu$ for all $h \in H$ where $h_*\nu$ is the measure given by

$$h_*\nu(E) = \nu(h^{-1}E) \text{ for any Borel set } E \subseteq Z.$$

An H-invariant measure ν is *ergodic* if any measurable function $f : Z \to \mathbb{R}$ on Z such that $f(hz) = f(z)$ for all $h \in H$ and ν-a.e. $z \in Z$ is constant almost everywhere.

In the case when $H = \{a_s\}_{s\in\mathbb{R}}$ is a one-parameter group, one can show that the quantity $\frac{1}{s}h_\nu(a_s)$ is independent of $s \neq 0$. So, we may define the measure theoretic entropy of the flow $\{a_s\}$to be

$$h_\nu(a_\circ) = h_\nu(a_1).$$

A similar results holds for the topological entropy of a flow.

3. Hausdorff dimension and upper box dimension

Let (Y, d) be a compact metric space, and let $F \subseteq Y$ be a subset. Fix $s \geq 0$. If $\{U_i\}_{i=1}^{\infty}$ is countable collection of sets and $\delta > 0$ is a positive number, then we say that $\{U_i\}$ is a *δ-cover* of F whenever $F \subseteq \bigcup_{i=1}^{\infty} U_i$ and $diam(U_i) < \delta$ for each $i \in \mathbb{N}$. For any $\delta > 0$, we define

$$\mathcal{H}^s_\delta(F) = \inf \big\{ \sum_{i=1}^{\infty} diam(U_i)^s : \{U_i\} \text{ is a } \delta\text{-cover of } F \big\}.$$

If $0 < \delta' < \delta$, then any δ'-cover of F is also a δ-cover of F. This implies that $\mathcal{H}^s_\delta \leq \mathcal{H}^s_{\delta'}$. Then, we can define the *s-dimensional Hausdorff measure* of F by

$$\mathcal{H}^s(F) = \lim_{\delta \to 0^+} \mathcal{H}^s_\delta(F).$$

Let $\{U_i\}$ be a δ-cover of F. If $t > s$, then

$$diam(U_i)^t = diam(U_i)^s diam(U_i)^{t-s} \leq diam(U_i)^s \delta^{t-s}$$

for each $i \in \mathbb{N}$. It follows that

$$\mathcal{H}^t_\delta(F) \leq \delta^{t-s} \mathcal{H}^s_\delta(F).$$

Similarly, if $t' < s$, then

$$\mathcal{H}^{t'}_\delta(F) \geq \frac{1}{\delta^{s-t'}} \mathcal{H}^s_\delta(F).$$

So, if $0 < \mathcal{H}^s(F) < \infty$, then $\mathcal{H}^t(F) = 0$ whenever $t > s$ and $\mathcal{H}^{t'}(F) = \infty$ whenever $t' < s$. Therefore,

$$\inf\{s : \mathcal{H}^s(F) = 0\} = \sup\{s : \mathcal{H}^s F = \infty\}$$

and we define the *Hausdorff dimension* of F, or $dim_H(F)$, to be this common value.

Proposition 3.1. *The Hausdorff dimension has the following properties.*

(1) *If E and F are subsets of Y and $E \subseteq F$, then $dim_H(E) \leq dim_H(F)$.*
(2) *If $F_1, F_2, \dots$ is a sequence of subsets of Y, then $dim_H(\bigcup_{i=1}^{\infty} F_i) = \sup_{i \in \mathbb{N}} dim_H(F_i)$.*

Proof. See page 29 of [6]. □

Let $\epsilon > 0$ be a postive number. A set $E \subseteq F$ is *ϵ-separated* if $d(x, y) \geq \epsilon$ for any $x, y \in E$ with $x \neq y$. Note that any ϵ-separated subset of F is finite because Y is compact. Define $sep(\epsilon, F)$ to be the maximum cardinality of ϵ-separated subsets of F. A collection $\{U_i\}_{i \in I}$ is an *ϵ-cover* of F if $diam(U_i) < \epsilon$ for all $i \in I$ and $\bigcup_{i \in I} U_i = F$. Define $cov(\epsilon, F)$ to be the minimum cardinality of ϵ-covers of F. Note that $cov(\epsilon, F)$ is finite because Y is compact.

Define the *upper box dimension* of F by

$$dim_{ub}(F) = \limsup_{\epsilon \to 0^+} \frac{\log(sep(\epsilon, B))}{-\log(\epsilon)}.$$

Proposition 3.2. *The upper box dimension has the following properties.*

(1) *If E and F are subsets of Y and $E \subseteq F$, then $dim_{ub}(E) \leq dim_{ub}(F)$.*
(2) *If E and F are subsets of Y, then $dim_{ub}(E \cup F) = \max\{dim_{ub}(E), dim_{ub}(F)\}$.*

The upper box dimension tends to be easier to compute than the Hausdorff dimension.

Lemma 3.3. *Suppose that (ϵ_k) is a sequence in $(0,1)$ such that $\epsilon_{k+1} \geq c\epsilon_k$ for some $0 < c < 1$ and (ϵ_k) decreases to zero. Then,*

$$dim_{ub}(F) = \limsup_{k\to\infty} \frac{\log(sep(\epsilon_k, F))}{-\log(\epsilon_k)}.$$

Proof. For any k large enough so that $c > \epsilon_{k+1}$ and $\epsilon \in [\epsilon_{k+1}, \epsilon_k]$,

$$\begin{aligned}\frac{\log(sep(\epsilon, F))}{-\log(\epsilon)} &\leq \frac{\log sep(\epsilon_{k+1}, F)}{-\log(\epsilon_k)} \\ &= \frac{\log(sep(\epsilon_{k+1}, F))}{-\log(\epsilon_{k+1}) + \log(\frac{\epsilon_{k+1}}{\epsilon_k})} \\ &\leq \frac{\log(sep(\epsilon_{k+1}, F))}{-\log(\epsilon_{k+1}) + \log(c)}\end{aligned}$$

Then,

$$\limsup_{\epsilon\to 0} \frac{\log(sep(\epsilon, F))}{-\log(\epsilon)} \leq \sup_{k\geq n} \frac{\log(sep(\epsilon_k, F))}{-\log(\epsilon_k) + \log(c)} \text{ for all } n \in \mathbb{N}.$$

Therefore,

$$\limsup_{\epsilon\to 0} \frac{\log(sep(\epsilon, F))}{-\log(\epsilon)} \leq \limsup_{k\to\infty} \frac{\log(sep(\epsilon_k, F))}{-\log(\epsilon_k) + \log(c)}.$$

We choose a subsequence $(\epsilon_{k_j})_{j\in\mathbb{N}}$ such that

$$\lim_{j\to\infty} \frac{\log(sep(\epsilon_{k_j}, F))}{-\log(\epsilon_{k_j}) + \log(c)} = \limsup_{k\to\infty} \frac{\log(sep(\epsilon_k, F))}{-\log(\epsilon_k) + \log(c)}.$$

Then,

$$\begin{aligned}\lim_{j\to\infty} \frac{\log(sep(\epsilon_{k_j}, F))}{-\log(\epsilon_{k_j})} &= \lim_{j\to\infty} \frac{\log(sep(\epsilon_{k_j}, F))}{-\log(\epsilon_{k_j}) + \log(c)} \frac{-\log(\epsilon_{k_j}) + \log(c)}{-\log(\epsilon_{k_j})} \\ &= \lim_{j\to\infty} \frac{\log(sep(\epsilon_{k_j}, F))}{-\log(\epsilon_{k_j}) + \log(c)}.\end{aligned}$$

Therefore,

$$\begin{aligned}\limsup_{k\to\infty} \frac{\log(sep(\epsilon_k, F))}{-\log(\epsilon_k) + \log(c)} &= \lim_{j\to\infty} \frac{\log(sep(\epsilon_{k_j}, F))}{-\log(\epsilon_{k_j})} \\ &\leq \limsup_{\epsilon\to 0} \frac{\log(sep(\epsilon, F))}{-\log(\epsilon)}.\end{aligned}$$

The result follows because

$$\limsup_{k\to\infty} \frac{\log(sep(\epsilon_k, F))}{-\log(\epsilon_k) + \log(c)} = \limsup_{k\to\infty} \frac{\log(sep(\epsilon_k, F))}{-\log(\epsilon_k)}$$

□

Furthermore, the upper box dimension is related to the Hausdorff dimension by the following facts.

Lemma 3.4. *For any $\epsilon > 0$,*

$$cov(2\epsilon, F) \leq sep(\epsilon, F) \leq cov(\epsilon, F).$$

Proof. The lemma follows by arguments analogous to the ones given in the proof of lemma 2.3. □

Proposition 3.5. *For any subet $F \subseteq Y$,*

$$dim_H(F) \leq dim_{ub}(F).$$

In particular, if the upper box dimension of a set F is zero, then the Hausdorff dimension of F is zero.

Proof. By lemma 3.2,

$$\frac{\log(cov(2\epsilon, F))}{-\log(\epsilon)} \leq \frac{\log(sep(\epsilon, F))}{-\log(\epsilon)} \leq \frac{\log(cov(\epsilon, F))}{-\log(\epsilon)}.$$

Additionally,

$$\frac{\log(cov(2\epsilon, F))}{-\log(2\epsilon)} = \frac{\log(cov(2\epsilon, F))}{-\log(\epsilon) - \log(2)}.$$

Therefore,

$$dim_{ub}(B) = \limsup_{\epsilon \to 0^+} \frac{\log(cov(\epsilon, F))}{-\log(\epsilon)}.$$

Since F can be covered by $cov(\delta, F)$ sets of diameter less than δ,

$$\mathcal{H}^s_\delta(F) \leq cov(\delta, F)\delta^s,$$

for any nonnegative number s. So, if $1 < \mathcal{H}^s(F)$ for some $s \geq 0$, then

$$\log(cov(\delta, F)) + s\log(\delta) > 0$$

for all δ sufficiently small. Thus,

$$s \leq \limsup_{\delta \to 0^+} \frac{\log(cov(\delta), F)}{-\log(\delta)}.$$

□

In general, the Hausdorff dimension of a subset does not equal the upper box dimension of the subset. For example, consider the case when $Y = [0, 1]$ and $F = \{0\} \cup \{\frac{1}{n} : n \in \mathbb{N}\}$. Proposition 3.3 and the fact that single element subsets have zero Hausdorff dimension imply that countable sets have zero Hausdorff dimension. So, $dim_H(F) = 0$. However, we will show that $dim_{ub}(F) = \frac{1}{2}$.

Let $\delta \in (0, \frac{1}{2})$ be given. Let m be the integer such that $\frac{1}{m(m+1)} < \delta \leq \frac{1}{(m-1)m}$. Then, any subset $U \subseteq Y$ with $diam(U) < \delta$ can contain at most one of the points in $\{1, \frac{1}{2}, \ldots, \frac{1}{m}\}$. So, $cov(\delta, F) \geq m$. Additionally, $[0, \frac{1}{m}]$ can be covered by $m + 1$ intervals of length δ. So, $cov(\delta, F) \leq 2m$. Consequently,

$$\frac{\log(m)}{\log(m(m+1))} \leq \frac{\log(cov(\delta, F))}{-\log(\delta)} \leq \frac{\log(2m)}{\log((m-1)m)}.$$

For each $k \in \mathbb{N}$ with $k \geq 2$, set $\epsilon_k = \frac{1}{2}\frac{1}{(k-1)k} + \frac{1}{2}\frac{1}{k(k+1)}$. Then, $\frac{1}{k(k+1)} < \epsilon_k \leq \frac{1}{(k-1)k}$ for each k. Additionally, $\epsilon_{k+1} \geq \frac{k}{k+2}\epsilon_k$ for each k so that $\epsilon_{k+1} \geq \frac{1}{2}\epsilon_k$ for all

$k \geq 2$. Therefore,

$$\begin{aligned}
dim_{ub}(F) &= \limsup_{k\to\infty} \frac{\log(cov(\epsilon_k, F))}{\log(\epsilon_k)} \text{ by lemma 3.4} \\
&\leq \limsup_{k\to\infty} \frac{\log(2k)}{\log((k-1)k)} \\
&= \limsup_{k\to\infty} \frac{\log(2)+\log(k)}{2\log(k)} \frac{2\log(k)}{\log(k-1)+\log(k)} \\
&= \left(\lim_{k\to\infty} \frac{\log(2)+\log(k)}{2\log(k)}\right)\left(\lim_{k\to\infty} \frac{2\log(k)}{\log(k-1)+\log(k)}\right) \\
&= \frac{1}{2}
\end{aligned}$$

Similarly,

$$\begin{aligned}
dim_{ub}(F) &= \limsup_{k\to\infty} \frac{\log(cov(\epsilon_k, F))}{\log(\epsilon_k)} \\
&\geq \liminf_{k\to\infty} \frac{\log(k)}{\log(k(k+1))} \\
&= \frac{1}{2}.
\end{aligned}$$

4. Partial result towards Littlewood's conjecture

Set $G = SL(3,\mathbb{R})$, $\Gamma = SL(3,\mathbb{Z})$, and $X = G/\Gamma$. Let d_X be a metric on X induced by a right-invariant metric, d_G, on G. Let $A \leqslant G$ be the subgroup given by

$$A = \{a(s,t) : s,t \in \mathbb{R}\},$$

where

$$a(s,t) = \begin{pmatrix} e^{-s-t} & 0 & 0 \\ 0 & e^s & 0 \\ 0 & 0 & e^t \end{pmatrix},$$

for all real numbers s and t. We say that an A-invariant measure μ on X is *algebraic* if there exists $g \in G$ and a closed subgroup $L \leqslant G$ such that the following properties hold:

- $Lg\Gamma$ admits a unique L-invariant probability measure,
- μ is supported on $Lg\Gamma$, and
- the restriction of μ to $Lg\Gamma$ is the unique L-invariant probability measure on $Lg\Gamma$.

Manfred Einsiedler, Anatole Katok, and Elon Lindenstrauss used the following theorem to prove a partial result towards the Littlewood conjecture.

Theorem 4.1. *Let μ be an A-invariant and ergodic probability measure on X. Suppose that there exists a one-parameter subgroup $\{a_t\}_{t\in\mathbb{R}}$ of A such that $h_\mu(a_\circ) > 0$. Then, μ is an A-invariant algebraic measure which is not supported on a periodic A-orbit. In fact, μ is the G-invariant probability measure on X.*

If μ is the G-invariant measure on X, then the support of μ is the entire space X. The Godement compactness criterion implies that $X = G/\Gamma$ is compact if and only if the only unipotent element of Γ is the identity element. Since $\begin{pmatrix}1 & 1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}$ is a nontrivial unipotent element of Γ, X is not compact. Therefore, μ is not compactly supported.

Corollary 4.2. *Suppose that μ satisfies the hypotheses of Theorem 4.1. Then μ is not compactly supported.*

We recall that a lattice in $\mathbb{R}^n$ is a set of all integral combinations of n linearly independent vectors $v_1, \dots v_n$. We denote the space of all lattices with covolume one in $\mathbb{R}^n$ by $\mathcal{L}_n$. A set of n vectors determine a lattice with covolume one if and only if the n-by-n matrix whose columns are the n vectors has determinant ± 1. Moreover, given two n-by-n matrices A and B in $SL(n, \mathbb{R})$, the columns of A and the columns of B determine the same lattice if and only if AB^{-1} is an element of $SL(n, \mathbb{Z})$. Thus, $SL(n, \mathbb{R})/SL(n, \mathbb{Z})$ is identified with $\mathcal{L}_n$ by the map given by $gSL(n, \mathbb{Z}) \mapsto g(\mathbb{Z}^n)$.

This identification induces a topology on $\mathcal{L}_n$. We say that a subset $E \subset \mathcal{L}_n$ is bounded if the closure of E is compact with respect to this topology. For each lattice $\Lambda \in \mathcal{L}_n$, we define

$$\delta_{\mathbb{R}^n}(\Lambda) = \min_{v \in \Lambda \setminus \{0\}} \|v\|_\infty.$$

Proposition 4.3 (Mahler compactness criterion)**.** *$E \subseteq \mathcal{L}_n$ is bounded if and only if there exists $\epsilon > 0$ such that $\delta_{\mathbb{R}^n}(\Lambda) > \epsilon$ for all $\Lambda \in E$*

Proof. See page 53 of [8]. □

For any pair of real numbers α and β, we define $x_{\alpha,\beta}$ by

$$x_{\alpha,\beta} = \begin{pmatrix}1 & 0 & 0\\ \alpha & 1 & 0\\ \beta & 0 & 1\end{pmatrix}\Gamma.$$

Let A^+ be the semigroup in G given by

$$A^+ = \{a(s,t) : s, t > 0\}.$$

The following proposition states that the pairs of real numbers that satisfy the equation in Littlewood's conjecture are precisely those that correspond to the unbounded orbits under the action of coordinate dilations in the space $\mathcal{L}_3$.

Proposition 4.4. *The pair of real numbers (α, β) satisfies*

$$\liminf_{n\to\infty} n\|n\alpha\|\|n\beta\| = 0$$

if and only if $A^+ x_{\alpha,\beta}$ is unbounded.

Proof. If either α or β is a rational number, then we can find an integer vector v such that $x_{\alpha,\beta}(v)$ has at least one zero in its entries. Then, we can use Proposition 4.3 to check that $A^+ x_{\alpha,\beta}$ is unbounded. So, we may assume that both α and β are irrational numbers.

Suppose that $A^+ x_{\alpha,\beta}$ is unbounded. Let $\epsilon \in (0, \frac{1}{2})$ be given. By Proposition 4.3, there exists positive numbers $s, t > 0$ such that

$$\delta_{\mathbb{R}^3}((a(s,t)x_{\alpha,\beta})(\mathbb{Z}^3)) < \epsilon^{\frac{1}{3}}.$$

So, there exists $(n, k, l) \in \mathbb{Z}^3 \setminus \{0\}$ such that

$$\begin{aligned}
\epsilon^{\frac{1}{3}} &> \|(a(s,t)x_{\alpha,\beta})(n,k,l)\|_\infty \\
&= \| \begin{pmatrix} e^{-s-t} & 0 & 0 \\ e^s\alpha & e^s & 0 \\ e^t\beta & 0 & e^t \end{pmatrix} \begin{pmatrix} n \\ k \\ l \end{pmatrix} \|_\infty \\
&= \|(ne^{-s-t}, ne^s\alpha + ke^s, ne^t\beta + le^t)\|_\infty
\end{aligned}$$

Then, n must be nonzero; otherwise, either $k \neq 0$ so that $|ne^s\alpha + ke^s| = |ke^s| > 1$, or $l \neq 0$ so that $|ne^t\beta + le^t| = |le^t| > 1$. Additionally, we can assume that n is positive because we can replace (n, k, l) with $(-n, -k, -l)$ without changing the inequality above. Thus,

$$\begin{aligned}
n\|n\alpha\|\|n\beta\| &= ne^{-s-t}(e^s\|n\alpha\|)(e^t\|n\beta\|) \\
&\leq ne^{-s-t}(e^s|n\alpha + k|)(e^t|n\beta + l|) \\
&< \epsilon
\end{aligned}$$

Since α and β are irrational, $n\|n\alpha\|\|n\beta\| \neq 0$. Then, by repeating the argument above, we can find a positive integer n' so that

$$0 < n'\|n'\alpha\|\|n'\beta\| < n\|n\alpha\|\|n\beta\|.$$

In this manner, we can inductively choose a sequence of distinct positive integers (n_k) such that $n_k\|n_k\alpha\|\|n_k\beta\|$ converges to 0 as k approaches ∞. Thus, α and β satisfy the desired equation.

On the other hand, suppose that α and β satisfy

$$\liminf_{n\to\infty} n\|n\alpha\|\|n\beta\| = 0.$$

Let $\epsilon \in (0, \frac{1}{2})$ be given. Then, there exists $n \in \mathbb{N}$ such that

$$n\|n\alpha\|\|n\beta\| < \epsilon^6.$$

We can find integers k and l such that $\|n\alpha\| = |n\alpha - k|$, and $\|n\beta\| = |n\beta - l|$. So,

$$n|n\alpha - k||n\beta - l| < \epsilon^6.$$

If $\max[|n\alpha - k|, |n\beta - l|] < \epsilon$, then we can choose positive numbers s and t such that

$$e^s|n\alpha - k| = \epsilon, \text{ and } e^t|n\beta - k| = \epsilon.$$

This implies that

$$ne^{-s-t} < \frac{\epsilon^6}{e^s|n\alpha - k|e^t|n\beta - k|} < \epsilon.$$

Now, suppose that $\max[|n\alpha - k|, |n\beta - l|] \geq \epsilon$. Without loss of generality, we may assume that $|n\alpha - k| \geq \epsilon$. In this case, $n|n\beta - l| < \epsilon^5$. Let Q be the smallest integer larger than $\frac{1}{2\epsilon}$. By Theorem 1.2, there exists a positive integer $q < Q$ so that

$$\|qn\alpha\| < \frac{1}{Q} < 2\epsilon.$$

Note that $Q < \frac{1}{\epsilon}$ by our choice of ϵ. Then,

$$qn|qn\beta - ql| < \epsilon^3 < \epsilon.$$

So, we can choose positive numbers s and t such that

$$e^s\|qn\alpha\| = 2\epsilon, \text{ and } e^t|qn\beta - ql| = \epsilon.$$

Moreover,

$$qn\|qn\alpha\||qn\beta - ql| < qn|qn\alpha - qk||qn\beta - ql| < \epsilon^3.$$

Therefore,

$$\|(qne^{-s-t}, qne^s\alpha - k'e^s, qne^t\beta - qle^t)\|_\infty \leq 2\epsilon,$$

where k' is the integer closest to $qn\alpha$. □

For any $\sigma, \tau > 0$ and $t \in \mathbb{R}$, set $a_{\sigma,\tau}(t) = a(\sigma t, \tau t)$. For any $0 < l < 1$, define K_l to be the closure of

$$\{x_{\alpha,\beta} \in X : \delta_{\mathbb{R}^3}(a(s,t)x_{\alpha,\beta}(\mathbb{Z}^3)) \geq l \text{ for all } s, t > 0\}.$$

For any $0 < l < 1$, K_l is a compact metric space by Proposition 4.3, and $a(s,t)K_l \subseteq K_l$ for all $s, t > 0$. Note that if $\bigcup_{0<l<1} K_l$ is empty, then Littlewood's conjecture is true. For our purposes, we assume that $\bigcup_{0<l<1} K_l$ is nonempty.

Proposition 4.5. *Let $a \in A$ and $K \subseteq X$ be compact with $aK \subseteq K$. Then the map $\mu \mapsto h_\mu(a_{|K})$ is an upper semi-continuous map from the space of a-invariant probability measures on K with the weak-* topology to the nonnegative real numbers.*

Proof. See Corollary 9.3 of [4]. □

Proposition 4.6. *For any $\sigma, \tau > 0$ and $l > 0$, the topological entropy of $a_{\sigma,\tau}$ acting on the compact set K_l is 0.*

Proof. Let $\sigma, \tau > 0$ and $l > 0$ be given. Suppose, for contradiction, that the topological entropy of $a(\sigma, \tau)$ acting on K_l is positive. By Proposition 2.5, there exists an $a(\sigma,\tau)$-invariant probability measure ν' on K_l such that $h_{\nu'}(a(\sigma,\tau)) > 0$. Note that ν' extends to a Radon probability measure ν on X supported on K_l.

For each $n \in \mathbb{N}$, define the map $\psi_n : C(K_l) \to \mathbb{R}$ by

$$\psi_n(f) = \frac{1}{n^2}\int_0^n \int_0^n \int_{K_l} f(x)\, d(a(s,t)_*\nu')(x)\, ds\, dt$$

for all $f \in C(K_l)$. Note that

$$\begin{aligned}\psi_n(f) &= \frac{1}{n^2}\int_0^n \int_0^n \int_{a(-s,-t)K_l} f(a(s,t)x)\, d\nu'(x)\, ds\, dt \\ &= \frac{1}{n^2}\int_0^n \int_0^n \int_{K_l} f(a(s,t)x)\, d\nu'(x)\, ds\, dt,\end{aligned}$$

because $K_l \subseteq a(-s,-t)K_l$ for all $s, t > 0$. Then, ψ_n is a positive linear functional and $\|\psi_n\| = 1$. By the Riesz-Kakutani representation theorem, ψ_n corresponds to a unique probability measure ν'_n. By construction, ν'_n is $a(\sigma,\tau)$-invariant. Then, Proposition 2.6 and Proposition 4.5 imply that,

$$h_{\nu'_n}(a(\sigma,\tau)) = \frac{1}{n^2}\int_0^n \int_0^n h_{a(s,t)_*\nu'}(a(\sigma,\tau))\, ds\, dt.$$

For any $s, t > 0$, $h_{\nu'}(a(\sigma,\tau)) = h_{a(s,t)_*\nu'}(a(\sigma,\tau))$; in fact, $a(s,t)$ defines an isomorphism between (K_l, ν') and $(K_l, a(s,t)_*\nu')$ for all $s, t > 0$. Therefore,

$$h_{\nu'_n}(a(\sigma,\tau)) = h_{\nu'}(a(\sigma,\tau)).$$

The space of Radon measures on K_l is the dual space of $C(K_l)$, which is a separable space. Thus, the unit sphere in the space of Radon measures on K_l is metrizable in the weak-$*$ topology. Since $(\nu'_n)_{n\in\mathbb{N}}$ is a sequence contained in the unit sphere, the Banach-Alaogulu theorem implies that there exists a subsequence

$(\nu'_{n_j})_{j\in\mathbb{N}}$ that converges in the weak-$*$ topology. Let ν'_∞ be the weak-$*$ limit. Then, ν'_∞ is $a(\sigma,\tau)$-invariant because, for any $g \in C(K_l)$,

$$\begin{aligned}\int_{K_l} g(x)\, d(a(\sigma,\tau)_*\nu'_\infty)(x) &= \int_{a(-\sigma,-\tau)K_l} g(a(\sigma,\tau)x)\, \nu'_\infty(x)\\ &= \lim_{j\to\infty}\int_{a(-\sigma,-\tau)K_l} g(a(\sigma,\tau)x)\, d\nu'_{n_j}(x)\\ &= \lim_{j\to\infty}\int_{K_l} g(x)\, d(a(\sigma,\tau)_*\nu'_{n_j})(x)\\ &= \lim_{j\to\infty}\int_{K_l} g(x)\, d\nu'_{n_j}(x)\\ &= \int_{K_l} g(x)\, d\nu'_\infty(x).\end{aligned}$$

Furthermore, Propostion 4.5 implies that

$$h_{\nu'_\infty}(a(\sigma,\tau)) \geq \limsup_{j\to\infty} h_{\nu'_{n_j}}(a(\sigma,\tau)) = h_{\nu'}(a(\sigma,\tau)) > 0.$$

We denote the extension of ν'_∞ to a Radon probability measure on X supported on K_l by ν_∞. Let $f \in C_c(X)$ and $s_0, t_0 \in \mathbb{R}$. We observe that

$$\begin{aligned}&\int_X f(x)\, d(a(s_0,t_0)_*\nu_\infty)(x)\\ &= \int_X f(a(s_0,t_0)x)\, d\nu_\infty(x)\\ &= \lim_{j\to\infty}\frac{1}{n_j^2}\int_0^{n_j}\int_0^{n_j}\int_X f(a(s_0+s,t_0+t)x)\, d\nu(x)\, ds\, dt\\ &= \lim_{j\to\infty}\frac{1}{n_j^2}\int_{t_0}^{n_j+t_0}\int_{s_0}^{n_j+s_0}\int_X f(a(s,t)x)\, d\nu(x)\, ds\, dt\end{aligned}$$

Then,

$$\begin{aligned}&|\int_X f(x)\, d\nu_\infty(x) - \int_X f(x)\, d(a(s_0,t_0)_*\mu_\infty)(x)|\\ &\leq \lim_{j\to\infty}\frac{1}{n_j^2}\int\int_{R_j}\int_X |f(x)|\, d\nu(x)\, ds\, dt,\end{aligned}$$

where $R_j = ([0,n_j]\times[0,n_j])\triangle([t_0,n_j+t_0]\times[s_0,n_j+s_0])$. The Lebesgue measure of R_j is at most $2(|t_0|n_j + |s_0|n_j)$. Therefore,

$$\begin{aligned}&\lim_{j\to\infty}\frac{1}{n_j^2}\int\int_{R_j}\int_X |f(x)|\, d\nu(x)\, ds\, dt,\\ &\leq \lim_{j\to\infty}\frac{2(|t_0|n_j+|s_0|n_j)\|f\|_\infty}{n_j^2}\\ &= 0.\end{aligned}$$

As a result,

$$\int_X f(x)\, d\nu_\infty(x) = \int_X f(x)\, d(a(s_0,t_0)_*\nu_\infty)(x)$$

for all $f \in C_c(X)$ and $s_0, t_0 \in \mathbb{R}$. Then, for every $s_0, t_0 \in \mathbb{R}$,

$$\nu_\infty = a(s_0, t_0)_* \nu_\infty$$

because $a(s_0, t_0)_* \nu_\infty$ is a Radon measure. It follows that ν_∞ is A-invariant.

Let $\nu_\infty = \int_\Xi \nu_\xi \, dm(\xi)$ be the ergodic decomposion of ν_∞. By Proposition 2.7,

$$h_{\nu_\infty}(a(\sigma, \tau)) = \int_\Xi h_{\nu_\xi}(a(\sigma, \tau)) \, dm(\xi).$$

Additionally,

$$h_{\nu_\infty}(a(\sigma, \tau)) = h_{\nu'_\infty}(a(\sigma, \tau)).$$

Therefore, we can choose an A-invariant ergodic measure ν_ξ such that

$$h_{\nu_\xi}(a_{\sigma,\tau}) = h_{\nu_\xi}(a(\sigma, \tau)) > 0 \text{ and } \nu_\xi(X \setminus K_l) = 0.$$

This contradicts Corollary 4.2. $\square$

For any $r > 0$ and $x \in X$, let B_r^G denote the open ball in G of d_G-radius r at the identity and let $B_r(x) = B_r^G x$. We will assume the following lemma.

Lemma 4.7. *For every $r > 0$, there exists a constant $c_0 \geq 1$ such that*

$$c_0^{-1} \|g - h\| \leq d_G(g, h) \leq c_0 \|g - h\|$$

for all $g, h \in B_r^G$, where $\|A\| = \max_{i,j} |a_{ij}|$ for $A = (a_{ij}) \in M_3(\mathbb{R})$.

For any $0 < l < 1$, define C_l to be the closure of

$$\{x \in X : \delta_{\mathbb{R}^3}(a(s,t)x(\mathbb{Z}^3)) \geq l \text{ for all } s, t > 0\}.$$

It is a fact that, for every $x \in X$, there exists some $r = r(x) > 0$ small enough such that the map given by $g \mapsto gx$ is an isometry between B_r^G and $B_r(x)$. For any $0 < l < 1$, we can choose a unifrom r such that B_r^G is isomorphic to $B_r(x)$ for all $x \in C_l$, because C_l is compact by the Mahler compactness critierion.

Fix $a = a(1,1)$. We note that if $x \in X$, $g \in B_r^G$ and $y = gx$ then $ay = (aga^{-1})ax$. We define the *unstable subgroup for conjugation with a* by

$$U = \{g \in G : a^n g a^{-n} \to e \text{ as } n \to -\infty\},$$

where e is the identity. Because

$$a^{-1} g a = \begin{pmatrix} g_{11} & e^3 g_{12} & e^3 g_{13} \\ e^{-3} g_{21} & g_{22} & g_{23} \\ e^{-3} g_{31} & g_{32} & g_{33} \end{pmatrix}, \text{ for any } g = (g_{ij}) \in G,$$

U consists of matrices of the form

$$\begin{pmatrix} 1 & 0 & 0 \\ * & 1 & 0 \\ * & 0 & 1 \end{pmatrix}.$$

In particular, U is homeomorphic to $\mathbb{R}^2$. Moreover, if $x = e\Gamma \in X$ is the identity coset then

$$Ux \cap C_l = K_l, \text{ for all } 0 < l < 1.$$

The following lemma and proposition show that K_l is a countable union of sets with zero upper box dimension for all $0 < l < 1$.

Lemma 4.8. *Let $C \subseteq X$ be compact with $aC \subseteq C$, and let $r = r(C)$ be a positive number such that $g \mapsto gx$ is an isometry from B_r^G to $B_r(x)$ for all $x \in C$. Then, there exists $\lambda > 1$ and $c_2 > 0$ so that for any small enough $\epsilon > 0$, any $z \in C$, any $f \in B_r^U$, and any integer $N \geq 1$ with $d_X(fz, z) \geq \lambda^{-N}\epsilon$, there exists a nonnegative integer $n < N$ with $d_X(a^n fz, a^n z) \geq c_2\epsilon$.*

Proof. See lemma 8.4 of [4]. □

Proposition 4.9. *Let $C \subseteq X$ be compact with $aC \subseteq C$. Then, one of the following properties holds.*

(1) *The intersection $Ux \cap C$ of the unstable manifold Ux with C is a countable union of compact sets with upper box dimension zero for every $x \in X$.*
(2) *The action of a on C has positive topological entropy.*

Proof. Suppose that there exists some $\epsilon > 0$ such that $P_y = C \cap (B_\epsilon^U y)$ has upper box dimension 0 for every $y \in C$. Let $x \in X$ such that $Ux \cap C$ is nonempty. Suppose that $D \subseteq U$ is a compact subset of U. Then, $Dx \cap C$ is also a compact subset of C, so there exists a finite cover $\{U_i\}_{i=1}^m$ of $Dx \cap C$ such that $U_i = P_{y_i}$ for some $y_i \in C$ for every $1 \leq i \leq m$. Since upper box dimension is finitely additive, the upper box dimension of $Dx \cap C$ is 0. Then, $Ux \cap C$ is a countable union of compact sets of box dimension zero because U is homeomorphic to $\mathbb{R}^2$.

On the other hand, suppose that for any $\epsilon > 0$ there exists $y = y(\epsilon) \in C$ such that the box dimension of P_y is positive. Choose $r > 0$ and $\epsilon' > 0$ such that r and ϵ' satisfy the hypothesis of lemma 4.8 and $2\epsilon' \leq r$. Let y' be the point in C such that $dim_{box}(P_{y'})$ is nonzero. Fix $b \in (0, dim_{ub}(P_{y'}))$. For every $N > 0$, let $F_N \subseteq P_{y'}$ be a $\epsilon'\lambda^{-N}$-separated set of maximal cardinality. Since $b < dim_{ub}(P_{y'})$, lemma 3.3 implies that there exist infinitely many integers N with

$$b \leq \frac{\log(sep(\epsilon'\lambda^{-N}, P_{y'}))}{-\log(\epsilon'\lambda^{-N})}.$$

We observe that

$$\begin{aligned}
&-\log(\epsilon'\lambda^{-N})b \leq \log(sep(\epsilon'\lambda^{-N}, P_{y'})) \\
&\Rightarrow \log((\epsilon')^{-b}\lambda^{Nb}) \leq \log(|F_N|) \\
&\Rightarrow (\epsilon')^{-b}\lambda^{Nb} \leq |F_N|.
\end{aligned}$$

Let (N_k) be an increasing sequence of integers such that $(\epsilon')^{-b}\lambda^{N_k b} \leq |F_{N_k}|$ for each $k \in \mathbb{N}$. Fix $k \in \mathbb{N}$, and choose two distinct points p and q in F_{N_k}. Then, $d_X(p, q) \geq \epsilon'\lambda^{-N_k}$ and there exist $g, h \in B_{\epsilon'}^U$ such that $p = gy'$ and $q = hy'$. We note that $f = gh^{-1} \in B_r^U$. Lemma 4.8 implies that there exists a nonnegative integer n such that $n < N_k$ and $d_X(a^n fq, a^n q) \geq c_2\epsilon'$. Therefore, F_{N_k} is $(N_k, c_2\epsilon')$-separated.

Finally, we see that

$$\begin{aligned} h_{top}(a_{|C}) &= \lim_{\epsilon\to 0^+}\limsup_{n\to\infty}\frac{1}{n}\log(sep(n,\epsilon,a)) \\ &\geq \limsup_{n\to\infty}\frac{1}{n}\log(sep(n,c_2\epsilon',a)) \\ &\geq \limsup_{k\to\infty}\frac{1}{N_k}\log(sep(N_k,c_2\epsilon',a)) \\ &\geq \limsup_{k\to\infty}\frac{1}{N_k}\log(|F_{N_k}|) \\ &\geq b\log(\lambda)+\limsup_{k\to\infty}\frac{\log((\epsilon')^{-b})}{N_k} > 0. \end{aligned}$$

□

For each $0 < l < 1$, define

$$S_l = \{(\alpha,\beta)\in[0,1]\times[0,1] : x_{\alpha,\beta}\Gamma \in K_l\}.$$

Now, fix $0 < l < 1$ and choose $r > 0$ such that the map $g \mapsto gx$ is an isometry between B_r^G and $B_r(x)$ for all $x \in K_l$. By lemma 4.7, there exists a constant $c_0 \geq 1$ such that

$$c_0^{-1}\max[|\alpha-\alpha'|,|\beta-\beta'|] \leq d_G(x_{\alpha,\beta},x_{\alpha',\beta'}) \leq c_0\max[|\alpha-\alpha'|,|\beta-\beta'|].$$

for any $x_{\alpha,\beta}\in K_l$ and $x_{\alpha',\beta'}\in B_r(x_{\alpha,\beta})$. Additionally, each element of S_l corresponds to a unique element in K_l. Thus, if $\epsilon < r$ and $E \subseteq S_l$ is an ϵ-separated set, then

$$F = \{x_{\alpha,\beta}\Gamma : (\alpha,\beta)\in E\}$$

is an $\frac{\epsilon}{c_0}$-separated set in K_l. It follows that

$$dim_{ub}(S_l) \leq dim_{ub}(K_l) = 0.$$

By Proposition 3.5,

$$dim_H(S_l) = 0.$$

Corollary 4.10. *Let $S = \{(\alpha,\beta)\in[0,1]\times[0,1] : A^+x_{\alpha,\beta}$ is bounded$\}$. Then, the Hausdorff dimension of S is 0.*

Therefore, theorem 4.2 implies that the set of counterexamples to the Littlewood conjecture has Hausdorff dimension zero. In the remaining sections, we will discuss some of the ideas involved in the proof of theorem 4.2.

5. Measure rigidity

For notational convenience, define

$$\Sigma = \{\mathbf{t} = (t_1,t_2,t_3)\in\mathbb{R}^3 : t_1+t_2+t_3 = 0\}.$$

For each $\mathbf{t}\in\Sigma$ set $\alpha^{\mathbf{t}} = a(t_2,t_3)$.

Let $H \leqslant G$ be a subgroup. An element $g \in G$ *normalizes* H if $gHg^{-1} = H$, and a subgroup $L \leqslant G$ normalizes H if any element of L normalizes H. The *normalizer of* H is defined by

$$N(H) = \{g \in G : gHg^{-1} = H\}.$$

An element $g \in G$ *centralizes* H if $gh = hg$ for all $h \in H$, and the *centralizer of* H is defined by

$$C(H) = \{g \in G : gh = hg \text{ for all } h \in H\}.$$

Both $N(H)$ and $C(H)$ are subgroups of G containing H. We say that H is *unipotent* if $(h - Id)^n = 0$ for some $n \in \mathbb{N}$. If A normalizes H, then for every $x \in X$ and $a \in A$,

$$a(Hx) = H(ax).$$

As a result, the foliation of X into H-orbits is invariant under the action of A.

Let μ be an A-invariant probability measure on X. For any unipotent subgroup $U \leqslant G$ normalized by A, there exists a system $\{\mu_x^U\}_{x\in X}$ of Radon measures on U and an A-invariant subset $X' \subseteq X$ of full measure so that the following properties hold.

(1) The map $x \mapsto \mu_x^U$ is measurable.
(2) For every $\epsilon > 0$ and $x \in X'$, $\mu_x^U(B_\epsilon^U) > 0$.
(3) For every $x \in X'$ and $u \in U$ with $ux \in X'$, $\mu_x^U \propto u_*\mu_{ux}^U$, where $u_*\mu_{ux}^U$ is the push forward of μ_{ux}^U under right multiplication by u.
(4) We normalize the measures so that $\mu_x^U(B_1^U) = 1$ for every $x \in X'$. Then, for any $x \in X$ and $\mathbf{t} \in \Sigma$, $\mu_{\alpha^{\mathbf{t}}x}^U$ is proportional to the push forward of μ_x^U under conjugation by $\alpha^{\mathbf{t}}$.
(5) If $U \subseteq C(\alpha\mathbf{t})$, then $\mu_{\alpha^{\mathbf{t}}x}^U = \mu_x^U$.
(6) μ is U-invariant if and only if μ_x^U is the Haar measure on U for almost every $x \in X'$.
(7) μ_x^U is atomic if and only if μ_x^U is supported on the identity $e \in U$. In this case, we say that μ_x^U is trivial.

Now, let (i, j) be a pair of distinct integers with $1 \le i, j \le 3$. We let E_{ij} denote the matrix whose entries are 1 at the (i, j)-th entry and 0 everywhere else, and we define

$$U_{ij} = \{u_{ij}(s) : s \in \mathbb{R}\}, \text{ where } u_{ij}(s) = Id + sE_{ij}.$$

We observe that

$$\begin{pmatrix} e^{t_1} & 0 & 0 \\ 0 & e^{t_2} & 0 \\ 0 & 0 & e^{t_3} \end{pmatrix} u_{ij}(s) \begin{pmatrix} e^{-t_1} & 0 & 0 \\ 0 & e^{-t_2} & 0 \\ 0 & 0 & e^{-t_3} \end{pmatrix} = u_{ij}(e^{t_i - t_j}s).$$

In particular, A normalizes U_{ij} so that the orbits of U_{ij} form an A-invariant foliation. We denote the foliation of X into U_{ij}-orbits by F_{ij}. The leaves of this foliation are one-dimensional because U_{ij} is a one-parameter subgroup of G. Additionally, U_{ij} is a unipotent subgroup of G. In this case, we write $\{\mu_x^{U_{ij}}\}_{x\in X} = \{\mu_x^{ij}\}_{x\in X}$.

We will assume the following adaptation of the Ledrappier-Young formula ([7]).

Proposition 5.1. *For any pair of indices, (i, j), there are constants $s_{ij}(\mu) \in [0, 1]$ that satisfy the following properties.*

(1) $s_{ij}(\mu) = 0$ if and only if μ_x^{ij} is trivial for almost every $x \in X$.
(2) $s_{ij}(\mu) = 1$ if and only if μ_x^{ij} is Haar for almost every $x \in X$.
(3) For any $\boldsymbol{t} \in \Sigma$

$$h_\mu(\alpha^{\boldsymbol{t}}) = \sum_{i,j} s_{ij}(\mu)(t_i - t_j)^+,$$

where $(r)^+ = max(0, r)$ for each $r \in \mathbb{R}$.

Proof. See lemma 6.2 of [3] $\square$

The following theorem is an important ingredient in the proof of Theorem 4.1.

Theorem 5.2. *Let μ be an A-invariant and ergodic probability measure on X. For every pair (i,j) of distinct indices, there are two mutually exclusive possibilities; either μ_x^{ij} and μ_x^{ji} are trivial for almost every $x \in X$, or μ_x^{ij} and μ_x^{ji} are Haar for almost every $x \in X$.*

Suppose that there exists a pair of distinct indices (a,b) such that μ_x^{ab} is not trivial for all x in some subset of positive measure. Define

$$E = \{x \in X : \mu_x^{ab} \text{ is trivial }\}.$$

Then, E is an essentially A-invariant subset of X. Indeed, if $x \in E \cap X'$, then for any $\mathbf{t} \in \Sigma$ there exists a constant $C > 0$ that depends on x and $\mathbf{t}$ such that

$$\begin{aligned}
\mu_{\alpha^{\mathbf{t}}x}^{ab}(U_{ab} \setminus \{0\}) &= C\mu_x^{ab}(\alpha^{-\mathbf{t}}(U_{ab} \setminus \{0\})\alpha^{\mathbf{t}}) \\
&= C\mu_x^{ab}(U_{ab} \setminus \{0\}) \\
&= 0.
\end{aligned}$$

So, the ergodicity of μ implies that E is a null set. In other words, μ_x^{ab} is nonatomic almost everywhere.

Then, the proof of Theorem 5.2 can be broken up into two cases. Let $1 \leq a,b,c \leq 3$ be distinct indices and suppose that μ_x^{ab} is nonatomic almost everywhere. The high entropy case is when either μ_x^{ac} or μ_x^{cb} is nonatomic for almost every $x \in X$. The low entropy case is when both μ_x^{ac} and μ_x^{cb} are trivial for almost every $x \in X$. In both cases, we can conclude that μ_x^{ab} and μ_x^{ba} is Haar almost everywhere.

For now, we assume that Theorem 5.2 holds. Suppose that μ is an A-invariant and ergodic probability measure on X such that $h_\mu(a_\circ) > 0$ for some one parameter subgroup, $\{a_t\}_{t\in\mathbb{R}}$, of A. We write $a_1 = \alpha^{\mathbf{t}}$ for some $\mathbf{t} \in \Sigma$. Since $h_\mu(\alpha^{\mathbf{t}}) > 0$, there exists a distinct pair of indices (a,b) such that $t_a > t_b$ and $s_{ab}(\mu) > 0$ by Proposition 5.1. By Theorem 5.2, μ_x^{ab} and μ_x^{ba} are Haar almost everywhere. Then, μ is invariant under the actions of U_{ab} and U_{ba}. If μ_x^{ac} is Haar almost everywhere, then μ is also invariant under the actions of U_{bc} because U_{bc} is contained in the subgroup generated by U_{ba} and U_{ac}. By Theorem 5.2, μ is invariant under U_{ca} and U_{bc}. Then, μ is G-invariant because the collection of subgroups

$$\{U_{ij} : (i,j) \text{ is a pair of distinct indices}\}$$

generates G. An analogous argument holds for the case when μ_x^{bc} is Haar almost everyhwere.

On the other hand, suppose that μ_x^{ac} and μ_x^{bc} are trivial almost everywhere. Without loss of generality, we may assume that $a = 1$ and $b = 2$. Define H to be the subgroup of G consisting of elements of the form $\begin{pmatrix} * & * & 0 \\ * & * & 0 \\ 0 & 0 & 1 \end{pmatrix}$. It is a fact that U_{12} and U_{21} generate H. So, μ is H-invariant and H is generated by unipotent one-parameter subgroups. Additionally, H is normalized by A. It is also a fact that H is the maximal proper subgroup of G that satisfies these three conditions. Then, Theorem 6.1 of [4] implies that there exists a subgroup $L \leqslant G$ such that $H \leqslant L$, L is normalized by A, and almost every H-ergodic component of μ is the L-invariant measure on a closed L-orbit. With more work, we can show that H must equal L using the maximailty of H. Define

$$X_H = \{x \in X : Hx \text{ is closed and of finite volume }\}.$$

Then, the support of μ is contained in X_H.

Let $x \in X_H$ and write $x = g\Gamma$ where $g \in G$. It is a fact that we can choose the representative g to be a matrix with rational entries. Set $z = (1, 1, 0)$ and $m' = g^{-1}z$. Then, m' is a rational vector so there exists an integer k such that $m = km'$ is an integer vector and $gm = kz$. Define a one parameter subgroup $\{b_s\}_{s\in\mathbb{R}} \leqslant A$ by

$$b_s = \alpha^{s\mathbf{t}} \text{ where } \mathbf{t} = (-1, -1, 2).$$

Then,

$$b_s gm = b_s kz = (e^{-s}k, e^{-s}k, 0) \text{ for all } s \in \mathbb{R}.$$

By Mahler's compactness criterion, $b_s x \to \infty$ as $t \to \infty$; in other words, for any compact subset $K \subseteq X$, $b_s x \notin K$ for all s sufficiently large. In particular, if $K \subseteq X_H$ is any compact subset, then $x \in K$ cannot return to K infinitely often under the action of b_1. This contradicts Poincaré recurrence because we can approximate the measure of X_H by compact sets contained in X_H and X_H is a full measure set. Therefore, μ must be G-invariant.

In this way, Theorem 5.2 implies Theorem 4.1. In the remaining two sections, we will discuss the proof of Theorem 5.2.

6. The high entropy case

Let $\lambda = (i, j)$ be a pair of distinct indices. Let k be the other index and let ξ be a pair of distinct indices such that $\xi \neq (i, j)$ and $\xi \neq (j, i)$. If $\xi = (k, i)$ or $\xi = (k, j)$ choose $\mathbf{t} = (t_1, t_2, t_3) \in \Sigma$ such that $t_i = t_j = 1$. Otherwise, choose $\mathbf{t}$ such that $t_i = t_j = -1$. Then, $\alpha^{\mathbf{t}}$ acts isometrically on the leaves of $F_\lambda = F_{ij}$. and contracts the leaves of F_ξ.

For each $x \in X$, we define the measure ν_x^λ on $\mathbb{R}$ by

$$\nu_x^\lambda(A) = \mu_x^\lambda(\{exp(sE_\lambda) : s \in A\})$$

for any Borel set $A \subseteq \mathbb{R}$. It is a fact that ν_x^λ is a Radon measure on $\mathbb{R}$ for each $x \in X$.

Lemma 6.1. *For any $f \in C_c(\mathbb{R})$, the map $\Phi_f : X \to \mathbb{R}$ given by*

$$\Phi_f(x) = \int_{\mathbb{R}} f \, d\nu_x^\lambda$$

is measurable.

Proof. See the proof of proposition 5.1 in [3]. □

Proposition 6.2. *For some null set $N \subseteq X$ and for any two $x, y \in X \setminus N$ such that there exists $y' = exp(rE_\lambda)x \in F_\lambda(x)$ with $y' \in F_\xi(y)$,*

$$\nu_y^\lambda(A) = C\nu_x^\lambda(A + r)$$

for any Borel set $A \subseteq \mathbb{R}$ and some constant $C > 0$.

Proof. Let N_0 be the null set such that for any $x, y \in X \setminus N_0$,

$$\mu_x^\lambda \propto u_* \mu_y^\lambda \text{ whenever } y = ux \text{ for some } u \in U_\lambda, \text{ and}$$
$$\mu_x^\lambda = \mu_y^\lambda \text{ whenever } y = \alpha^{n\mathbf{t}} \text{ for some } n \in \mathbb{Z}.$$

It is a fact that $C_c(\mathbb{R})$ is a separable space. So we can choose a countable dense subset, $\{f_k\}_{k\in\mathbb{N}}$, of $C_c(\mathbb{R})$. For each $j \in \mathbb{N}$, we use Luzin's theorem to choose a compact set, K_j, such that the restriction $(\Phi_{f_k})_{|K_j}$ is continuous for all $k \in \mathbb{N}$ and $\mu(K_j) > 1 - \frac{1}{j}$. By replacing K_j with $\bigcup_{i=1}^{j} K_i$, we may assume that the sequence (K_j) is increasing.

For each $j \in \mathbb{N}$, define $g_j : X \to \mathbb{R}$ by

$$g_j(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \chi_{K_j}(\alpha^{k\mathbf{t}} x).$$

By the Pointwise Ergodic Theorem, $g_j(x) = \mu(K_j)$ for almost every $x \in X$. Define

$$L_j = \{x \in X : g_j(x) \leq \frac{1}{2}\}.$$

Note that L_j is a measurable set for each $j \in \mathbb{N}$ and $(L_j)_{j\in\mathbb{N}}$ is an increasing sequence. We observe that

$$\begin{aligned}
1 &= \lim_{j\to\infty} \mu(K_j) \\
&= \lim_{j\to\infty} \int_X g_j d\mu \\
&\leq \lim_{j\to\infty} \Big(\mu(X \setminus L_j) + \frac{1}{2}\mu(L_j)\Big) \\
&\leq 1 - \frac{1}{2} \lim_{j\to\infty} \mu(L_j).
\end{aligned}$$

Therefore, $\lim_{j\to\infty} \mu(L_j) = 0$ so that $N = N_0 \cup \big(\bigcap_{j=1}^{\infty} L_j\big)$ is a null set.

Now, suppose that $x, y \notin N$ and that there exists an element $y' = exp(rE_\lambda)$ in $F_\lambda(x) \cap F_\xi(y)$. By the construction of N, there exists $j_0 \in \mathbb{N}$ such that $x, y \notin L_{j_0}$. So, there exists $\epsilon > 0$ such that

$$\min[g_{j_0}(x), g_{j_0}(y)] \geq \frac{1}{2} + \epsilon.$$

Then, there exists $N \in \mathbb{N}$ such that for all $n \geq N$

$$\sum_{k=0}^{n-1} \chi_{K_{j_0}}(\alpha^{k\mathbf{t}} x) \geq \frac{n + n\epsilon}{2}, \text{ and}$$
$$\sum_{k=0}^{n-1} \chi_{K_{j_0}}(\alpha^{k\mathbf{t}} y) \geq \frac{n + n\epsilon}{2}.$$

This implies that

$$|\{k \in \{1, \dots, n\} : \alpha^{k\mathbf{t}} x, \alpha^{k\mathbf{t}} y \in K_j\}| \geq \lfloor n\epsilon \rfloor,$$

for all $n \geq N$. Therefore we can choose an increasing sequence of integers $(n_i)_{i\in\mathbb{N}}$ so that

$$\alpha^{n_i\mathbf{t}} x, \alpha^{n_i\mathbf{t}} y \in K_j \text{ for all } i \in \mathbb{N}.$$

Since K_j is compact, we can replace $(n_i)_{i\in\mathbb{N}}$ with a subsequence if necessary and assume that $(\alpha^{n_i\mathbf{t}}x)_{i\in\mathbb{N}}$ and $(\alpha^{n_i\mathbf{t}}y)_{i\in\mathbb{N}}$ converge to $\overline{x}$ and $\overline{y}$, respectively.

By our choice of N,

$$\nu^\lambda_{\alpha^{n_i\mathbf{t}}x} = \nu^\lambda_x \text{ for all } i \in \mathbb{N}.$$

By our choice of K_j,

$$\int_{\mathbb{R}} f_k\, d\nu^\lambda_{\overline{x}} = \lim_{i\to\infty} \int_{\mathbb{R}} f_k\, d\nu^\lambda_{\alpha^{n_i\mathbf{t}}x} = \int_{\mathbb{R}} f_k\, d\nu^\lambda_x \text{ for all } k \in \mathbb{N}.$$

Because $\{f_k\}_{k\in\mathbb{N}}$ is dense in $C_c(\mathbb{R})$,

$$\int_{\mathbb{R}} f\, d\nu^\lambda_{\overline{x}} = \int_{\mathbb{R}} f\, d\nu^\lambda_x \text{ for all } f \in C_c(\mathbb{R}).$$

Therefore $\nu^\lambda_{\overline{x}} = \nu^\lambda_x$. Similarly, $\nu^\lambda_{\overline{y}} = \nu^\lambda_y$. Since $\alpha^{\mathbf{t}}$ contracts $F_\xi(y)$,

$$d(\alpha^{n_i\mathbf{t}}y, \alpha^{n_i\mathbf{t}}y') \to 0 \text{ as } i \to \infty.$$

Additionally, $\alpha^{n_i\mathbf{t}}y' = exp(rE_\lambda)\alpha^{n_i\mathbf{t}}x$ for each $i \in \mathbb{N}$. It follows that

$$exp(rE_\lambda)\overline{x} = \lim_{i\to\infty} \alpha^{n_i\mathbf{t}}y' = \overline{y}.$$

Thus,

$$\nu^\lambda_y = \nu^\lambda_{\overline{y}} = \nu^\lambda_{exp(rE_\lambda)\overline{x}} \propto (T_{-r})_*\nu^\lambda_{\overline{x}} = (T_{-r})_*\nu^\lambda_x,$$

where $T_{-r} : \mathbb{R} \to \mathbb{R}$ is translation by $-r$. □

A notable special case of the proposition above is when $r = 0$ and $y' = x$.

Corollary 6.3. *There exists a null set $N \subseteq X$ so that for any two $x, y \in X \setminus N$ such that $x \in F_\xi(y)$,*

$$\nu^\lambda_x = \nu^\lambda_y.$$

Lemma 6.4. *Let $F_{\lambda_1}, \ldots, F_{\lambda_j}$ be different foliations. Let N be a null set. For each $1 \le i \le j$, define*

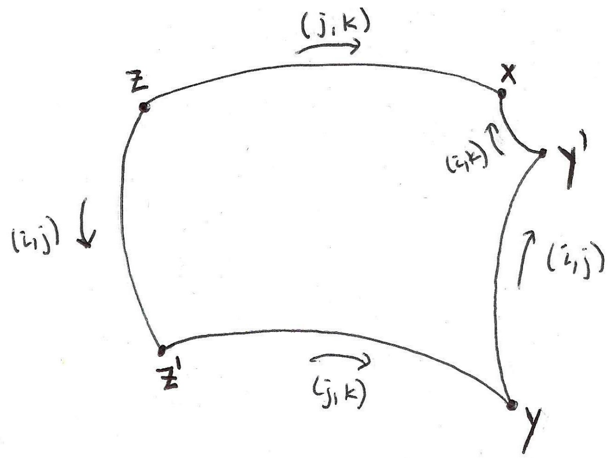
$$N(\lambda_i, x) = \{g \in U_{\lambda_i} : gx \in N\}.$$

Then, there exists a null set N' such that $N \subseteq N'$ and $\mu^{\lambda_i}_x(N(\lambda_i, x)) = 0$ for all $x \notin N'$ and $1 \le i \le j$.

Proof. See lemma 3.1 in [3]. □

Now, suppose that $1 \le i, j, k \le 3$ are three different indices such that μ^{ij}_x and μ^{jk}_x are nonatomic almost everywhere. Choose a null set $N_0 \subseteq X$ such that the the properties listed in Section 5 and the statement of Proposition 6.2 holds for all possible choices of λ, and μ^{ij}_x and μ^{jk}_x are nontrivial for all $x \in N_0$. By lemma 6.4, there exists a null set N such that $N_0 \subseteq N$ and $\mu^{ab}_x(N((a,b),x) = 0$ for all $x \notin N$ and $1 \le a, b \le 3$ with $a \ne b$, where

$$N((a,b),x) = \{g \in U_{ab} : gx \in N\}.$$



Let $z \notin N$ and $\epsilon > 0$. By the construction of N, μ_z^{ij} is nonatomic. Then, there exists $r \in (-\sqrt{\epsilon}, \sqrt{\epsilon}) \setminus \{0\}$ such that

$$z' = exp(rE_{ij})z \in F_{ij}(z) \setminus N,$$

because every neighorhood of the identity element in U_{ij} has positive measure. Additionally, Proposition 6.3 implies that $\nu_z^{jk} = \nu_{z'}^{jk}$. So, we can choose $s \in (-\sqrt{\epsilon}, \sqrt{\epsilon}) \setminus \{0\}$ such that

$$\begin{aligned} x &= exp(sE_{jk})z \in F_{jk}(z) \setminus N, \text{and} \\ y &= exp(sE_{jk})z' \in F_{jk}(z') \setminus N. \end{aligned}$$

We observe that

$$\begin{aligned} y &= exp(sE_{jk})exp(rE_{ij})z \\ &= exp(sE_{jk})exp(rE_{ij})exp(-sE_{jk})x \\ &= (Id + sE_{jk})(Id + rE_{ij})(Id - sE_{jk})x \\ &= (Id + rE_{ij})(Id - rsE_{ik})x, \end{aligned}$$

since

$$(Id - rE_{jk})(Id + sE_{jk})(Id + rE_{ij})(Id - sE_{jk}) = Id - rsE_{ik}.$$

Set

$$y' = (Id - rsE_{ik})x \in F_{ik}(x),$$

so that $y = exp(rE_{ij})y'$. By Proposition 6.2, there exists $D > 0$ such that

$$\nu_x^{ik}(A) = D\nu_y^{ik}(A - rs),$$

for any Borel set $A \subseteq \mathbb{R}$. Additionally, $\nu_z^{ik} = \nu_{z'}^{ik}, \nu_x^{ik} = \nu_z^{ik}$, and $\nu_y^{ik} = \nu_{z'}^{ik}$ by Corollary 6.3. Therefore, there exists $D > 0$ so that

$$\nu_z^{ik}(A + rs) = D\nu_z^{ik}(A),$$

for any Borel set $A \subseteq \mathbb{R}$. Note that $0 < |rs| < \epsilon$.

For each $t \in \mathbb{R}$, define $T_t : \mathbb{R} \to \mathbb{R}$ by

$$T_t(x) = x + t$$

and define the Borel measure $(T_t)_*\nu_z^{ik}$ by

$$(T_t)_*\nu_z^{ik}(B) = \nu_z^{ik}(T_{-t}(B)) = \nu_z^{ik}(B-t).$$

Define $G(\nu_z^{ik}) \subseteq \mathbb{R}$ by

$$G(\nu_z^{ik}) = \{t \in \mathbb{R} : (T_t)_*\nu_z^{ik} \propto \nu_z^{ik}\}.$$

Then, $G(\nu_z^{ik})$ is a subgroup of $\mathbb{R}$. We have shown above that 0 is an accumulation point of $G(\nu_z^{ik})$. This implies that $G(\nu_z^{ik})$ is a dense in $\mathbb{R}$.

We claim that $G(\nu_z^{ik})$ is a closed in $\mathbb{R}$. To see this, let (t_n) be a sequence in $G(\nu_z^{ik})$ that converges to $t_0 \in \mathbb{R}$. For each $t \in G(\nu_z^{ik})$ let C_t be the positive number such that $(T_t)_*\nu_z^{ik} = C_t\nu_z^{ik}$. Then, for any $f \in C_c(\mathbb{R})$,

$$\begin{aligned}\int_{\mathbb{R}} f(x+t_0)\, d\nu_z^{ik}(x) &= \lim_{n\to\infty}\int_{\mathbb{R}} f(x+t_n)\, d\nu_z^{ik}(x) \\ &\quad \text{by the Dominated Convergence Theorem} \\ &= \lim_{n\to\infty} C_{t_n}\int_{\mathbb{R}} f(x)\, d\nu_z^{ik}(x)\end{aligned}$$

Therefore, $t_0 \in G(\nu_z^{ik})$ and $C_{t_0} = \lim_{n\to\infty} C_{t_n}$. As a result, $G(\nu_z^{ik})$ is a closed subgroup of $\mathbb{R}$. It follows that $G(\nu_z^{ik}) = \mathbb{R}$.

We need to use the Lebesgue differentiation theorem, which is stated here.

Proposition 6.5. *Let ν_1 and ν_2 be two locally finite Borel measures on $\mathbb{R}$.*

(1) The limit

$$\rho(x) = \lim_{r\to 0}\frac{\nu_1((x-r,x+r))}{\nu_2((x-r,x+r))}$$

exists ν_2-a.e. and $\rho : \mathbb{R} \to [0,\infty)$ is a ν_2 measurable function.

(2) The set

$$S = \{x : \rho(x) = \infty\}$$

is measurable with respect to ν_1.

(3) If $(\nu_1)_{|S}$ is the restriction of ν_1 to S, then

$$\nu_1 = \rho\nu_2 + (\nu_1)_{|S}.$$

Let ρ and S be defined as in the statement of Proposition 6.5 with $\nu_1 = \nu_z^{ik}$ and $\nu_2 = m$, where m is the Lebesgue measure. Suppose for contradiction that ν_z^{ik} is not absolutely continuous with respect to m. Then, S has positive measure. Fix $x \in S$ and let $y \in \mathbb{R}$ be given. Set $t = y - x$. Because $t \in G(\nu_z^{ik})$,

$$(T_t)_*\nu_z^{ik} = C_t\nu_z^{ik}$$

Then,

$$\begin{aligned}\rho(y) &= \lim_{r\to 0}\frac{\nu_z^{ik}((y-r,y+r))}{m((y-r,y+r))} \\ &= \lim_{r\to 0}\frac{C_t^{-1}(T_t)_*\nu_z^{ik}((y-r,y+r))}{m((y-r,y+r))} \\ &= \lim_{r\to 0}\frac{C_t^{-1}\nu_z^{ik}((x-r,x+r))}{m((x-r,x+r))} \\ &= C_t^{-1}\rho(x).\end{aligned}$$

Therefore, S must equal $\mathbb{R}$. This contradicts Proposition 6.5. So, ν_z^{ik} is absolutely continuous with respect to m. By a similar argument, m is absolutely continuous with respect to ν_z^{ik}. Therefore, there exists a measurable function $h_z : \mathbb{R} \to (0, \infty)$ so that $d\nu_z^{ik} = h_z \, dm$.

Let $t \in \mathbb{R}$. We see that

$$d(T_t)_* \nu_z^{ik} = C_t \, d\nu_z^{ik} = C_t h_z \, dm.$$

On the other hand, for any $f \in C_c(\mathbb{R})$,

$$\begin{aligned} \int_{\mathbb{R}} f(x) \, d(T_t)_* \nu_z^{ik}(x) &= \int_{\mathbb{R}} f(x+t) \, d\nu_z^{ik}(x) \\ &= \int_{\mathbb{R}} f(x+t) h_z(x) \, dm(x) \\ &= \int_{\mathbb{R}} f(x) h_z(x-t) \, dm(x). \end{aligned}$$

It follows that

$$h_z(x-t) = C_t h_z(x)$$

for almost every $x \in \mathbb{R}$.

Note that the map $t \mapsto C_t$ satisfies $C_{s+t} = C_s C_t$ for all $s, t \in \mathbb{R}$. Additionally, we proved that this map is continous when we showed that $G(\nu_z^{ik})$ is closed. Then, the map $t \mapsto log(C_t)$ is a countinous additive function. It is a fact that all continuous additive functions are linear. As a result, there exists $\beta \in \mathbb{R}$ such that $\log(C_t) = \beta t$ or $C_t = e^{\beta t}$ for all $t \in \mathbb{R}$. This implies that

$$h_z(x-t) = e^{\beta t} h_z(x)$$

for any $t \in \mathbb{R}$ and for almost every $x \in \mathbb{R}$. Since h is positive almost everywhere, there exists some constant $D > 0$ so that

$$h_z(x) = De^{\beta x}$$

for all $x \in \mathbb{R}$. We have shown that, for any $z \notin N$, there exist constants $D > 0$ and $\beta \in \mathbb{R}$ depending on z so that

$$d\nu_z^{ik} = De^{\beta x} \, dm.$$

Let $\epsilon > 0$ be given. Suppose for contradiction that the set

$$E_\epsilon = \{z \in X \setminus N : \left| \log \left(\frac{h_z(x-1)}{h_z(x)} \right) \right| > \epsilon \text{ for almost every } x \in \mathbb{R}\}$$

has positive measure with respect to μ.

Define $F : X \setminus N \to \mathbb{R}$ by

$$F(z) = \limsup_{n \to \infty} \frac{\log \nu_z^{ik}([-n, n])}{2n}.$$

We claim that F is a measurable function. For each $n, j \in \mathbb{N}$, set

$$\mathcal{O}_j^n = \left[-n - \frac{1}{j}, n + \frac{1}{j} \right]$$

and let $g_j^n : \mathbb{R} \to \mathbb{R}$ be a continuous function such that $g_j^n = 1$ on $[-n, n]$ and $g_j^n = 0$ outside $\mathcal{O}_j^n$. Then, $g_j^n \to \chi_{[-n,n]}$ in $L^1(\mathbb{R}, \nu_z^{ik})$ for all $z \in X \setminus N$ because ν_z^{ik}

is a Radon measure. Hence,

$$F(z) = \limsup_{n\to\infty} \frac{\log(\limsup_{j\to\infty} \Phi_{g_j^n}(z))}{2n}$$

is measurable.

Let $z \in E_\epsilon$. Then, $d\nu_z^{ik} = De^{\beta x}\, dm$ for some $D > 0$ and $|\beta| > \epsilon$. We consider the case when $\beta > 0$. For any $r \in \mathbb{R}$,

$$\begin{aligned} \frac{\log(\nu_z^{ik}([-r,r]))}{2r} &= \frac{\log(\int_{-r}^{r} 1\, d\nu_z^{ik})}{2r} \\ &= \frac{\log(\int_{-r}^{r} De^{\beta x}\, dm(x))}{2r} \\ &= \frac{\log(D/\beta) + \log(e^{\beta r} - e^{-\beta r})}{2r}. \end{aligned}$$

Additionally, for any $r \in \mathbb{R}$,

$$\frac{\log(e^{\beta r} - 1)}{2r} \leq \frac{\log(e^{\beta r} - e^{-\beta r})}{2r} \leq \frac{\log(e^{\beta r})}{2r} = \frac{\beta}{2}$$

and if r is large enough so that $e^{\beta r} > 1$ then

$$\log(e^{\beta r} - 1) = \beta r + \frac{1}{\xi} \text{ for some } \xi \in (e^{\beta r} - 1, e^{\beta r})$$

by Taylor's theorem. Therefore,

$$F(z) = \limsup_{r\to\infty} \frac{\log \nu_z^{ik}([-r,r])}{2r} = \frac{\beta}{2}.$$

Similarly, when $\beta < 0$,

$$F(z) = \limsup_{r\to\infty} \frac{\log \nu_z^{ik}([-r,r])}{2r} = -\frac{\beta}{2}.$$

It follows that

$$F(z) = \limsup_{r\to\infty} \frac{\log \nu_z^{ik}([-r,r])}{2r} > \frac{\epsilon}{2},$$

for all $z \in E_\epsilon$.

Choose $\mathbf{t} = (t_1, t_2, t_3) \in \Sigma$ so that $t_i > t_k$. Let $z \in E_\epsilon$ be a typical point for Poincaré recurrence. Then, there exists an increasing sequence of natural numbers (n_k) such that $\alpha^{n_k \mathbf{t}} z \notin N$. On the other hand, for any $z \notin N$, $n \in \mathbb{N}$, and $\mathbf{t} \in \Sigma$,

$$\begin{aligned} F(\alpha^{n\mathbf{t}} z) &= \limsup_{r\to\infty} \frac{\log(\nu_{\alpha^t z}^{ik}([-r,r]))}{2r} \\ &= \limsup_{r\to\infty} \frac{\log(C) + \log(\nu_z^{ik}([-re^{n(t_k - t_i)}, re^{n(t_k - t_i)}]))}{2r} \\ &= e^{n(t_k - t_i)} \limsup_{r\to\infty} \frac{\log(C) + \log(\nu_z^{ik}([-re^{n(t_k - t_i)}, re^{n(t_k - t_i)}]))}{2re^{n(t_k - t_i)}} \\ &= e^{n(t_k - t_i)} F(z). \end{aligned}$$

Then, $F(\alpha^{n_k \mathbf{t}} z) \to 0$ as $k \to \infty$. So, $\alpha^{n_k \mathbf{t}} z \notin E_\epsilon$ for all k large enough. This is a contradiction. Therefore, E_ϵ must be a null set. Since $\epsilon > 0$ was given, we have proved the following result.

Proposition 6.6. *Let $1 \leq i, j, k \leq 3$ be three different indices. If μ_x^{ij} and μ_x^{jk} are nonatomic almost everywhere, then μ_x^{ik} is the Haar measure almost everywhere.*

Theorem 6.7. *Let a, b, and c be distinct indices. Suppose that μ_x^{ab} is nontrivial almost everywhere and that either μ_x^{ac} or μ_x^{cb} is nontrivial almost everywhere. Then both μ_x^{ab} and μ_x^{ba} are Haar measures almost everywhere and μ is invariant under the action of the group generated by U_{ab} and U_{ba}.*

Proof. By Proposition 6.6 and Proposition 5.1,

$$s_{ca}(\mu) > 0 \Rightarrow s_{cb}(\mu) = 1 \text{ and } s_{bc}(\mu) > 0 \Rightarrow s_{ac}(\mu) = 1.$$

Define $\mathbf{t} = (t_1, t_2, t_3)$ and $\mathbf{t}' = (t'_1, t'_2, t'_3)$ by $t_a = \frac{2}{3}$, $t_b = -\frac{1}{3}$, $t_c = -\frac{1}{3}$, $t'_a = -\frac{1}{3}$, $t'_b = \frac{2}{3}$, and $t'_c = -\frac{1}{3}$. By Proposition 5.1,

$$\begin{aligned}
h_\mu(\alpha^{\mathbf{t}}) &= s_{ab}(\mu) + s_{ac}(\mu), \\
h_\mu(\alpha^{-\mathbf{t}}) &= s_{ba}(\mu) + s_{ca}(\mu), \\
h_\mu(\alpha^{\mathbf{t}'}) &= s_{ba}(\mu) + s_{bc}(\mu), \text{ and} \\
h_\mu(\alpha^{-\mathbf{t}'}) &= s_{ab}(\mu) + s_{cb}(\mu).
\end{aligned}$$

By the properties of measure theoretic entropy,

$$h_\mu(\alpha^{\mathbf{t}}) = h_\mu(\alpha^{-\mathbf{t}}) \text{ and } h_\mu(\alpha^{\mathbf{t}'}) = h_\mu(\alpha^{-\mathbf{t}'}).$$

Therefore,

$$\begin{aligned}
s_{ab}(\mu) + s_{ac}(\mu) &= s_{ba}(\mu) + s_{ca}(\mu) \text{ and} \\
s_{ba}(\mu) + s_{bc}(\mu) &= s_{ab}(\mu) + s_{cb}(\mu).
\end{aligned}$$

Note that $s_{ab}(\mu) > 0$ by Proposition 5.1. Then,

$$s_{ac}(\mu) = 1 \Rightarrow s_{ca}(\mu) > 0 \Rightarrow s_{cb}(\mu) = 1 \Rightarrow s_{bc}(\mu) > 0 \Rightarrow s_{ac}(\mu) = 1.$$

In particular, $s_{ac}(\mu) \geq s_{ca}(\mu)$, so that

$$s_{ba}(\mu) \geq s_{ab}(\mu) > 0.$$

By Proposition 6.6 and Proposition 5.1,

$$s_{cb}(\mu) > 0 \Rightarrow s_{ca}(\mu) = 1 \text{ and } s_{ac}(\mu) > 0 \Rightarrow s_{bc}(\mu) = 1.$$

As a result,

$$s_{bc}(\mu) = 1 \Rightarrow s_{cb}(\mu) > 0 \Rightarrow s_{ca}(\mu) = 1 \Rightarrow s_{ac}(\mu) > 0 \Rightarrow s_{bc}(\mu) = 1.$$

It follows that

$$s_{ac}(\mu) = 1 \Leftrightarrow s_{ac}(\mu) > 0 \Leftrightarrow s_{cb}(\mu) > 0 \Leftrightarrow s_{cb}(\mu) = 1.$$

Thus, μ_x^{ac} and μ_x^{cb} are Haar almost everywhere by Proposition 5.1.

Now, we can reverse the roles of $s_{ab}(\mu)$ and $s_{ac}(\mu)$ and apply the arguments above to conclude that μ_x^{ab} is Haar almost everywhere. Note that the inequalities above also imply that μ_x^{ca} and μ_x^{bc} are Haar almost everywhere. This allows us to conclude that μ_x^{ba} is Haar almost everywhere. □

7. The low entropy case

We will breifly overview the theorems used in the low entropy case. Let a, b and c be distinct indices and suppose that μ_x^{ab} is nontrivial almost everywhere but μ_x^{ac} and μ_x^{cb} are trivial almost everywhere. We recall from the proof of theorem 6.7 that

$$\begin{aligned} s_{ab}(\mu) + s_{ac}(\mu) &= s_{ba}(\mu) + s_{ca}(\mu) \text{ and} \\ s_{ba}(\mu) + s_{bc}(\mu) &= s_{ab}(\mu) + s_{cb}(\mu). \end{aligned}$$

Then, $s_{ba}(\mu) \leq s_{ab}(\mu)$, since $s_{ac}(\mu) = 0$. Similarly, $s_{ab}(\mu) \leq s_{ba}(\mu)$, because $s_{bc}(\mu) = 0$. Then,

$$s_{ba}(\mu) = s_{ab}(\mu).$$

Therefore, it suffices to show that μ_x^{ab} is Haar almost everywhere.

We define

$$\begin{aligned} A'_{ab} &= \{\alpha^{\mathbf{t}} \in A : t_a = t_b\} \text{ and} \\ C_{ab} &= C(< U_{ab}, U_{ba} >) = C(U_{ab}) \cap C(U_{ba}). \end{aligned}$$

Let $K \subseteq X$ be a compact subset. We say that *the A'_{ab} returns to K are strong exceptional* if there exists $\delta > 0$ so that for all $x, x' \in K$ and $\alpha^{\mathbf{t}} \in A'_{ab}$ with $x' = \alpha^{\mathbf{t}} x \in B_\delta(x) \cap K$, every $g \in B_\delta^G$ with $x' = gx$ satisfies $g \in C_{ab}$.

Proposition 7.1. *The following two conditions are equivalent.*

(1) *Almost every ergodic component of μ with respect to A'_{ab} is supported on a single C_{ab}-orbit.*
(2) *For every $\epsilon > 0$, there exists a compact set K with measure $\mu(K) > 1 - \epsilon$ so that the A'_{ab}-returns to K are strong exceptional.*

Proof. See Proposition 4.3 of [4]. □

The main theorem in the low entropy case states that μ is U_{ab} invariant if the two equivalent conditions in Proposition 7.1 fails.

Theorem 7.2. *Suppose that μ_x^{ab} are nontrivial almost everywhere and that μ_x^{ij} are trivial almost everywhere for every pair of indices (i, j) such that $(i, j) \neq (a, b)$ and either $i = a$ or $j = b$. Then one of the following properties holds.*

(1) *Almost every ergodic component of μ with respect to A'_{ab} is supported on a single C_{ab}-orbit.*
(2) *μ is U_{ab}-invariant.*

Proof. See Section 4 of [4]. □

Then, we can conclude that μ is U_{ab} invariant because there is no element $\gamma \in \Gamma$ that satisfies the conclusion of the following theorem.

Theorem 7.3. *Suppose that ν is an A'_{ab} invariant probability measure on X and that supp$(\nu) \subseteq C_{ab}x$ for some $x \in X$. Then, there exists an element $\gamma \in \Gamma$ with the following properties.*

(1) *γ is diagonalizable over $\mathbb{R}$.*
(2) *± 1 are not eigenvalues of γ.*
(3) *γ has one eigenvalue with multiplicity two and another simple eigenvalue.*

Proof. See Theorem 5.1 of [4]. □

8. Further work

We have merely cited some interesting results in this paper, and it will be worthwhile to learn the proofs of these results. The proof of Proposition 5.1 uses the idea that the conditional measure for foliations into higher dimensional leaves is a product of the conditional measures on the one-dimensional leaves considered above ([3]). Deducing Theorem 4.1 from Theorem 5.2 uses results from the theory of algebraic groups ([4]). The techniques involved in the proof of Theorem 7.2 are related to the works of Marina Ratner on unipotent flows ([4]).

Moreover, there are other interesting topics in number theory related to the ideas discussed here; the type of dynamics considered in this paper can also be observed in the study of automorphic forms and ideal classes in number fields ([12]).

Acknowledgments. I would like to thank my advisor Aaron Brown for the instructive weekly discussions in his office and for his feedback on this paper.

References

[1] M. Brin and G. Stuck. Introduction to Dynamical Systems. Cambridge University Press. 2002.
[2] A. Brown et al. Entropy, Lyapunov Exponents, and Rigidity of Group Actions. arXiv. 2018.
[3] M. Einsiedler and A. Katok. Invariant Measures on G/Γ for Split Simple Lie Groups G. Wiley Periodicals. 2003.
[4] M. Einsiedler, A. Katok, and E. Lindenstrauss. Invariant Measures and the Set of Exceptions to Littlewood's Conjecture. Annals of Mathematics. 2006.
[5] M. Einsiedler and T. Ward. Ergodic Theory with a View towards Number Theory. Springer. 2011.
[6] K. Falconer. Fractal Geometry: Mathematical Foundations and Applications. John Wiley & Sons. 1990.
[7] L.-S. Young and F. Ledrappier. The Metric Entropy of Diffeomorphisms. II. Relations between Entropy, Exponents and Dimension. Annals of Mathematics. 1985.
[8] D. Morris. Introduction to Arithmetic Groups. arXiv. 2001.
[9] M. Queffélec. An Introduction to Littlewood's Conjecture. Sociéte Mathématique de France. 2009.
[10] H. Royden and P. Fitzpatrick. Real Analysis. Pearson Education. 2010.
[11] W. M. Schmidt. Diophantine Approximation. Spiringer-Verlag. 1980.
[12] A. Vankatesh. The Work of Einsiedler, Katok and Lindenstrauss on the Littlewood Conjecture.
[13] K. Verma. A Primer on the Functional Equation $f(x + y) = f(x) + f(y)$. Resonance. 2017.
[14] P. Walters. An Introduction to Ergodic Theory. Spiringer-Verlag. 1982.