ON LOGARITHM LAWS FOR DYNAMICAL SYSTEMS

D. Y. Kleinbock and G. A. Margulis

Research announcement, as of October 31, 1995

Abstract. We generalize and sharpen D. Sullivan's logarithm law for geodesics by specifying conditions on a measure preserving system (X, μ, g) and a family of subsets $\{A_n \mid n \in \mathbb{N}\}$ of X under which $\#\{1 \le n \le N \mid g^n x \in A_n\}$ is for μ-a.e. $x \in X$ asymptotically (as $N \to \infty$) equal to $\sum_{n=1}^{N} \mu(A_n)$. These conditions can be verified for various special cases of systems $(X, \mu, g, \{A_n\})$, including actions of nonquasiunipotent elements g of semisimple Lie groups G on homogeneous spaces $X = G/\Gamma$, with $\{A_n\}$ being certain subsets of X approaching infinity as $n \to \infty$.

Introduction

1. Let Γ be a discrete group of hyperbolic isometries of $\mathbb{H}^{d+1}$ such that $V = \mathbb{H}^{d+1}/\Gamma$ is not compact and has finite volume. For $(p, \xi) \in S(V)$ (the unit tangent bundle of V), let $\gamma_t(p, \xi)$ be the geodesic on V through p in the direction of ξ. Dennis Sullivan in [Su] proved the following

Theorem. *For any $p \in V$ and almost all directions ξ*

$$\limsup_{t\to\infty} \frac{\operatorname{dist}\big(\gamma_t(p,\xi), p\big)}{\log t} = 1/d\,. \tag{1}$$

His proof uses mixing of the geodesic flow on $S(V)$ together with certain intrinsic properties of hyperbolic manifolds.

In this note, we sketch an alternative proof of Sullivan's logarithm law. We are able to bypass the geometric ("disjoint circle") part of Sullivan's proof by making use of quantitative characterization (cf. [Mo, KSp, KM]) of mixing of the geodesic flow. This way, much stronger results than those of Sullivan can be obtained. Furthermore, our main results (Theorems 3.1 and 3.2) can be applied to a large class of dynamical systems. In particular, one can establish

2. Theorem. *Let g be an endomorphism of a d-dimensional torus $T \cong \mathbb{R}^d/\mathbb{Z}^d$ given by $g(x_1, \dots, x_d) = (m_1 x_1, \dots, m_d x_d) \mod 1$, where m_i are integers greater than 1. Then for almost all $x \in T$*

$$\limsup_{n\to\infty} \frac{\log\left(\frac{1}{\operatorname{dist}(g^n x, 0)}\right)}{\log n} = 1/d\,; \tag{2}$$

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

furthermore, there exist C_1, $C_2 > 0$ *such that for almost all* $x \in T$

$$\liminf_{N\to\infty} \frac{\#\{1 \le n \le N \mid \operatorname{dist}(g^n x, 0) \le n^{-1/d}\}}{\log N} \ge C_1$$

and

$$\limsup_{N\to\infty} \frac{\#\{1 \le n \le N \mid \operatorname{dist}(g^n x, 0) \le n^{-1/d}\}}{\log N} \le C_2\,.$$

3. Theorem. *Let* G *be a connected semisimple Lie group without compact factors,* Γ *an irreducible lattice in* G*,* $\{g_t\}$ *a one-parameter nonquasiunipotent subgroup of* G*,* μ *the normalized Haar measure on* G/Γ*. Assume that the action of* $\{g_t\}$ *on* G/Γ *has property* (EM)[1]*. Also let* $\Delta : G/\Gamma \to \mathbb{R}_+$ *be a function such that*

(i) *it is uniformly continuous,*

(ii) *its distribution function* $F_\Delta(z) \stackrel{\text{def}}{=} \mu\big(\{x \in G/\Gamma \mid \Delta(x) \le z\}\big)$ *is uniformly continuous, and*

(iii) *there exists* $\lim_{z\to\infty} \frac{1}{z} \log\big(\frac{1}{1-F_\Delta(z)}\big) = d$*.*

Then for almost all $x \in G/\Gamma$

$$\limsup_{t\to\infty} \frac{\Delta(g_t x)}{\log t} = 1/d\,. \tag{3}$$

Moreover, if (iii) *is strengthened to*

(iii′) *there exist* C_1*,* C_2*,* $d > 0$ *such that for any* $z > 0$

$$C_1 e^{-dz} \le 1 - F_\Delta(c) = \mu\big(\{x \in G/\Gamma \mid \Delta(x) \ge z\}\big) \le C_2 e^{-dz}\,,$$

then for almost all $x \in G/\Gamma$*,* $\liminf_{N\to\infty} \dfrac{\#\{1 \le n \le N \mid \Delta(g_n x) \ge \frac{1}{d}\log n\}}{\log N} \ge C_1$ *and* $\limsup_{N\to\infty} \dfrac{\#\{1 \le n \le N \mid \Delta(g_n x) \ge \frac{1}{d}\log n\}}{\log N} \le C_2$*.*

Note that Theorem 1 can be derived from Theorem 3 by taking $G = SO_{d+1,1}(\mathbb{R})$ and Δ coming from $\operatorname{dist}(p, p_0)$ for fixed $p_0 \in V \cong K\backslash G/\Gamma$, K a maximal compact subgroup of G. More generally, one gets

4. Corollary. *For any locally symmetric space* V *of noncompact type, there exists* $d = d(V)$ *such that for any* $p \in V$ *and almost all directions* ξ*,* (1) *is satisfied.*

The constant $d(V)$ can be explicitly calculated in any given special case.

Another class of applications of Theorem 3 is given by the function Δ on the space $SL_d(\mathbb{R})/SL_d(\mathbb{Z})$ of unimodular lattices in $\mathbb{R}^d$ defined by

$$\Delta(\Lambda) \stackrel{\text{def}}{=} \max_{\mathbf{x}\in\Lambda\smallsetminus\{0\}} \log\big(\tfrac{1}{\|\mathbf{x}\|}\big)\,,$$

where $\|\cdot\|$ is any norm on $\mathbb{R}^d$; this function can be shown to satisfy the assumptions (i), (ii) and (iii′) of Theorem 3. Then the correspondence [Da, K] between Diophantine approximation and flows on spaces of lattices provides a direct connection with

[1] This condition, which was introduced in [KM], implies that correlation coefficients of smooth functions on G/Γ decay exponentially. It holds, for instance, if the semisimple part of $\{g_t\}$ projects noncompactly onto each simple factor of G.

metrical theory of Diophantine approximation. In particular, a modification of Theorem 3 can be used to prove Khintchine's Divergence Theorem (see [Ca]), as well as some of its refinements and higher dimensional generalizations.

The paper is organized as follows. In §1 we show that upper estimates for the limits in all the above theorems follow from the standard Borel-Cantelli lemma. On the other hand, lower estimates essentially reduce to proving that for certain families of subsets A_n of a probability space (X, μ) with $\sum_{n=1}^\infty \mu(A_n) = \infty$, almost all points belong to infinitely many A_n. In §2 we state a partial (quasi-independent) converse to Borel-Cantelli lemma, and then sharpen it in §3, specifying a condition on the family $\{A_n\}$ under which $\#\{n \le N \mid x \in A_n\}$ is for μ-a.e. $x \in X$ asymptotically (as $N \to \infty$) equal to $\sum_{n=1}^N \mu(A_n)$. In §4 we show that the latter condition, in its turn, holds whenever the correlations $|\mu(A_m \cap A_n) - \mu(A_m)\mu(A_n)|$ decay with certain rate as $|m-n| \to \infty$. Finally, in §5 we indicate how one can use dynamical properties of the actions considered in Theorems 1, 2 and 3 to verify the "correlation decay" condition introduced in §4.

§1. General setting and upper estimates

1.1. Let (X, μ) be a probability space. We will consider families $\mathcal{H} = \{h_n \mid n \in \mathbb{N}\}$ of nonnegative integrable[2] functions on X. The function $w : \mathbb{N} \to \mathbb{R}_+$ sending n to $\int_X h_n \, d\mu$ will be called the *weight function* of the family $\mathcal{H}$. We will say that the family $\mathcal{H}$ is *bounded* by $L > 0$ if $w(n) \le L$ for any $n \in \mathbb{N}$.

A natural way to obtain bounded families of functions is to consider characteristic functions of measurable subsets A_n of X from a family $\mathcal{A} = \{A_n \mid n \in \mathbb{N}\}$. The weight function $w(n) = \mu(A_n)$ of $\{1_{A_n} \mid n \in \mathbb{N}\}$ will be called the *weight function* of $\mathcal{A}$.

We will say that a family of functions or sets is *summable* if its weight function is, i.e. if $\sum_{n=1}^\infty w(n) < \infty$, and *nonsummable* otherwise.

Main example. If φ is any nonnegative measurable function on X and $f : \mathbb{N} \to \mathbb{R}_+$ any function, one can consider the family of sets

$$\mathcal{A}_{\varphi,f} \stackrel{\text{def}}{=} \big\{\{x \in X \mid \varphi(x) \le f(n)\} \mid n \in \mathbb{N}\big\},$$

its weight function being $F_\varphi \circ f$, where F_φ is the distribution function of the random variable φ.

1.2. Another main example. Let g be a μ-preserving self-map of X. Then given any family $\mathcal{H} = \{h_n \mid n \in \mathbb{N}\}$ of functions on X or a family $\mathcal{A} = \{A_n \mid n \in \mathbb{N}\}$ of subsets of X, one can consider *twisted* families

$$\mathcal{H}^g \stackrel{\text{def}}{=} \{g^{-n} h_n \mid n \in \mathbb{N}\} \quad \text{and} \quad \mathcal{A}^g \stackrel{\text{def}}{=} \{g^{-n} A_n \mid n \in \mathbb{N}\}.$$

By g-invariance of μ, the weight function of the twisted family is equal to that of the original one.

[2]Throughout the sequel all the functions h_n will be assumed nonnegative and integrable.

1.3. Given a family $\mathcal{A} = \{A_n \mid n \in \mathbb{N}\}$ with weight function w and a μ-generic point $x \in X$, one may want to look at the asymptotics of $\#\{1 \le n \le N \mid x \in A_n\}$ in comparison with the sum of measures of the sets A_n, $1 \le n \le N$, as $N \to \infty$. This is for example the subject of the classical Borel-Cantelli Lemma. In order to state its useful quantitative version, define functions

$$\Psi_{\mathcal{A},N}(x) \stackrel{\text{def}}{=} \frac{\#\{1 \le n \le N \mid x \in A_n\}}{\sum_{n=1}^N \mu(A_n)} = \frac{\sum_{n=1}^N 1_{A_n}(x)}{\sum_{n=1}^N w(n)}.$$

More generally, for $N \in \mathbb{N}$ and a family $\mathcal{H} = \{h_n \mid n \in \mathbb{N}\}$ of functions on X with weight function w, define

$$\Psi_{\mathcal{H},N} \stackrel{\text{def}}{=} \frac{\sum_{n=1}^N h_n}{\sum_{n=1}^N w(n)}.$$

In both cases we will make the ratio equal to 1 if the denominator is zero.

The next lemma is a straightforward quantitative extension of a part of Borel-Cantelli Lemma that does not assume (quasi-)independence of functions h_n (cf. [KSt, part (a) of the theorem]).

Lemma. *Let (X,μ) be a probability space, $\mathcal{H} = \{h_n \mid n \in \mathbb{N}\}$ a family of functions on X. Then for any $\varepsilon \ge 0$*

$$\mu\left(\left\{x \in X \,\middle|\, \liminf_{N\to\infty} \Psi_{\mathcal{H},N}(x) \le 1+\varepsilon\right\}\right) > \frac{\varepsilon}{1+\varepsilon};$$

in particular, $\liminf_{N\to\infty} \Psi_{\mathcal{H},N}$ is finite almost everywhere and is not greater than 1 on a set of positive measure.

A (less trivial) counterpart that gives lower estimate for $\limsup_{N\to\infty} \Psi_{\mathcal{H},N}(x)$ will be considered in §2. Let us now state a special case of Lemma 1.3.

1.4. Corollary. *Let (X,μ) be a probability space, $\{h_n \mid n \in \mathbb{N}\}$ a summable family of functions on X with weight function w. Then for any $\varepsilon \ge 0$*

$$\mu\left(\left\{x \in X \,\middle|\, \sum_{n=1}^{\infty} h_n(x) \le (1+\varepsilon)\sum_{n=1}^{\infty} w(n)\right\}\right) > \frac{\varepsilon}{1+\varepsilon},$$

in particular, $\sum_{n=1}^\infty h_n(x) < \infty$ for μ-a.e. $x \in X$.

In the case of a summable family $\{A_n\}$ of subsets of X, one gets

$$\mu\left(\left\{x \in X \,\middle|\, \#\{n \in \mathbb{N} \mid x \in A_n\} \le (1+\varepsilon)\sum_{n=1}^{\infty} w(n)\right\}\right) > \frac{\varepsilon}{1+\varepsilon}$$

and

$$\#\{n \in \mathbb{N} \mid x \in A_n\} < \infty \quad \text{for } \mu\text{-a.e. } x \in X, \tag{1.1}$$

the last assertion being the conclusion of the easy part of the classical Borel-Cantelli Lemma.

Let us now show how this corollary gives upper estimates for the limits in each of the three theorems from the introduction.

1.5. Example. Take $X = S(V)$ as in Theorem 1, μ the Liouville measure on $S(V)$. Fix $p_0 \in V$ and $\varepsilon > 0$, and let $\varphi\big((p,\xi)\big) = e^{-\operatorname{dist}(p,p_0)}$ and

$$f(n) = n^{-(1+\varepsilon)/d}\,. \tag{1.2}$$

Then (cf. [Su, §9]) the distribution function F_φ of φ satisfies

$$C_1 z^d \le F_\varphi(z) \le C_2 z^d \quad \text{for some } C_1,\ C_2 > 0 \text{ and all } 0 < z < 1\,, \tag{1.3}$$

therefore the family $\mathcal{A}_{\varphi,f}$ is summable, and if one lets g be the time-one map of the geodesic flow on $S(V)$, an application of Corollary 1.4 to the twisted family $(\mathcal{A}_{\varphi,f})^g$ yields

$$\#\{n \in \mathbb{N} \mid \varphi(g^n x) \le n^{-(1+\varepsilon)/d}\} < \infty \text{ for } \mu\text{-a.e. } x \in X\,. \tag{1.4}$$

From (1.4) one easily deduces that for μ-a.e. $(p,\xi) \in S(V)$

$$\frac{\operatorname{dist}\big(\gamma_n(p,\xi), p_0\big)}{\log n} < (1+\varepsilon)/d \quad \text{for large enough } n\,,$$

and since ε is arbitrary, lim sup of the left hand side is not greater than $1/d$. To derive the upper estimate for the limit in Theorem 1 from the above statement, it suffices to observe that for any two points p_1, p_2 of V

- the functions $\operatorname{dist}(\cdot,p_1)$ and $\operatorname{dist}(\cdot,p_2)$ are asymptotically equivalent, and
- for any geodesic γ starting from p_1 there is a geodesic starting from p_2 which stays at a bounded distance from γ.

1.6. Example. Take $X = T$ and g as in Theorem 2, μ the Lebesgue measure on T. Let $\varphi(x) = \operatorname{dist}(x,0)$, and take f as in (1.2). The distribution function F of φ clearly satisfies (1.3), hence the family $\mathcal{A}_{\varphi,f}$ is summable, and Corollary 1.4 applied to $(\mathcal{A}_{\varphi,f})^g$ gives (1.4). Thus for almost all $x \in T$ and all $\varepsilon > 0$

$$\frac{\log\left(\frac{1}{\operatorname{dist}(g^n x,0)}\right)}{\log n} < (1+\varepsilon)/d \quad \text{for large enough } n\,,$$

and the inequality in Theorem 2 follows.

1.7. Example. Take G, Γ, μ and Δ as in Theorem 3, $X = G/\Gamma$, $g = g_1$, and let $\varphi(x) = e^{-\Delta(x)}$. From the assumption (iii) it follows that the distribution function F_φ "almost" satisfies (1.3); more precisely, for any $\delta > 0$ there exist C_1, $C_2 > 0$ such that $C_1 z^{d(1-\delta)} \le F_\varphi(z) \le C_2 z^{d(1+\delta)}$ for all $0 < z < 1$. The latter condition can be replaced by (1.3) under the assumption (iii$'$). In either case the family $\mathcal{A}_{\varphi,f}$, with f as in (1.2), is summable, and the rest of the argument goes as in the preceding example.

The goal of the subsequent exposition will be to state conditions on the map g and the families $\mathcal{H}$ (resp. $\mathcal{A}$) under which nonsummability of $\mathcal{H}$ (resp. $\mathcal{A}$) implies that $\sum_{n=1}^\infty h_n(g^n x)$ (resp. $\#\{n \in \mathbb{N} \mid g^n x \in A_n\}$) is infinite for a.e. $x \in X$. Or, better yet, as $N \to \infty$, one may want to describe the asymptotic behavior of $\Psi_{\mathcal{H}^g,N}(x)$ (resp. $\Psi_{\mathcal{A}^g,N}(x)$) for almost all $x \in X$. This will prove, among other things, lower estimates for the limits in (1), (2) and (3).

§2. Kochen-Stone families and lower estimates

2.1. One can easily find many examples of families $\mathcal{H}$ of functions on a probability space (X,μ) for which $\Psi_{\mathcal{H},N} \underset{\mu\text{-a.e.}}{\longrightarrow} 0$ as $N\to\infty$. It is also well known that a lower estimate for $\limsup_{N\to\infty}\Psi_{\mathcal{H},N}$ follows from certain conditions on second moments of the functions h_n. Specifically, we make the following definition, For a family $\mathcal{H}=\{h_n \mid n\in\mathbb{N}\}$ of functions on a probability space (X,μ) and $C\ge 1$, say that $\mathcal{H}$ is C-*Kochen-Stone*[3] (will be abbreviated as C-KS) if

$$\liminf_{N\to\infty}\int_X (\Psi_{\mathcal{H},N})^2\,d\mu \le C\,. \tag{C-KS}$$

This in particular implies $h_n\in L^2(X)$ for all $n\in\mathbb{N}$. We will say that a family of sets is C-KS if the family of its characteristic functions is.

The following lemma is a quantitative sharpening of parts (b) and (c) of the theorem in [KSt] (cf. also [Su, §1] for related ideas).

Lemma. *Let (X,μ) be a probability space, $C\ge 1$, $\mathcal{H}=\{h_n \mid n\in\mathbb{N}\}$ a C-KS family of functions on X. Then*

(a) *for any $\varepsilon\le 1$*

$$\mu\left(\left\{x\in X \,\middle|\, \limsup_{N\to\infty}\Psi_{\mathcal{H},N}(x) > 1-\varepsilon\right\}\right) \ge \frac{\varepsilon^2}{C-1+\varepsilon^2}\,, \tag{2.1}$$

moreover, $\limsup_{N\to\infty}\Psi_{\mathcal{H},N}$ *is not less than* 1 *on a set of positive measure;*

(b) *if $\mathcal{H}$ is nonsummable, then $\sum_{n=1}^\infty h_n=\infty$ on a set of measure at least $1/C$.*

2.2. Corollary. *Assume in addition that* $\limsup_{N\to\infty}\Psi_{\mathcal{H},N}$ *is constant on a set of measure* 1*. Then*

(a) $\limsup_{N\to\infty}\Psi_{\mathcal{H},N}(x)\ge 1$ *for μ-a.e. $x\in X$;*

(b) *if $\mathcal{H}$ is nonsummable, then*

$$\sum_{n=1}^{\infty} h_n(x)=\infty \text{ for } \mu\text{-a.e. } x\in X. \tag{2.2}$$

The corresponding statements for families of sets are immediate. Note that by Lemma 2.1, 1-KS nonsummable families satisfy (2.2) without any additional assumptions.

In [Su], using geometric argument D. Sullivan essentially verified the condition (C-KS) for certain families of subsets of $S(V)$. From that he easily derived the lower estimate for the limit in (1). Indeed, take X, μ, g and φ as in Example 1.5, and let $f(n)=n^{-1/d}$. Then the family $\mathcal{A}=\mathcal{A}_{\varphi,f}$ is nonsummable, and from ergodicity of g-action on X it follows that $\limsup_{N\to\infty}\Psi_{\mathcal{A}^g,N}$ must be constant almost everywhere. Thus if one believes that $\mathcal{A}^g$ is C-KS for some C, Corollary 2.2 applies and one has

$$\#\{n\in\mathbb{N}\mid \varphi(g^n x)\le n^{-1/d}\}=\infty \text{ for } \mu\text{-a.e. } x\in X\,.$$

[3]We would like to mention that this condition seems to have been introduced independently by several authors, see [Sp, p. 317] for a historical overview

Therefore for μ-a.e. $(p, v) \in S(V)$

$$\frac{\operatorname{dist}\big(\gamma_n(p,v), p_0\big)}{\log n} \geq 1/d \quad \text{for infinitely many } n\,,$$

which forces lim sup of the left hand side to be not less than $1/d$.

We will see later that the fact that the family $\mathcal{A}^g$ as above is Kochen-Stone follows from dynamics of g-action on $S(V)$. In fact, we will show that this family, as well as the corresponding families from Examples 1.6 and 1.7, satisfy stronger conditions, using which one can describe asymptotics of $\#\{1 \leq n \leq N \mid \varphi(g^n x) \leq n^{-1/d}\}$ as $N \to \infty$ for almost all x.

§3. Strongly Kochen-Stone families and asymptotics for $\Psi_{\mathcal{H},N}$

3.1. Let (X, μ) be a probability space, $\mathcal{H}$ a nonsummable family of functions on X with weight function w. For $C \geq 1$, we will say that $\mathcal{H}$ is *strongly C-Kochen-Stone* if lim inf in (C-KS) can be replaced by lim sup and the error can be estimated in terms of growth of $\sum_{n=1}^N w(n)$. More precisely, if for some q, $Q > 0$ one has

$$\int_X (\Psi_{\mathcal{H},N})^2\, d\mu \leq C + \frac{Q}{\big(\sum_{n=1}^N w(n)\big)^q}\,.$$

We will now state a quantitative sharpening of Lemma 2.1.

Theorem. *Let (X, μ) be a probability space, $\mathcal{H} = \{h_n \mid n \in \mathbb{N}\}$ a bounded non-summable strongly* 1-KS *family of functions on X. Then*

$$\Psi_{\mathcal{H},N} \underset{\mu\text{-a.e.}}{\longrightarrow} 1 \text{ as } N \to \infty\,. \tag{3.1}$$

This is a quasi-independent analogue of Theorem 6.6 from [Du]; our proof essentially follows the lines of Durrett's argument.

Proof. For $N \in \mathbb{N}$, we will suppress $\mathcal{H}$ in $\Psi_{\mathcal{H},N}$ and let $S_N(x) = \sum_{n=1}^N h_n(x)$ and $c_N = \sum_{n=1}^N w(n)$. Without loss of generality one can assume that $q \leq 2$. Let L be an upper bound for $w(n)$, $n \in \mathbb{N}$. Then one can define an increasing sequence $\{N_k\}$, $k \in \mathbb{N}$, by

$$N_k = \min\{N \mid c_N \geq Lk^{2/q}\}\,.$$

Since $\mathcal{H}$ is strongly 1-KS and $\int_X \Psi_N\, d\mu = 1$, one has

$$\operatorname{Var}(\Psi_N) \leq \frac{Q}{(c_N)^q}\,,$$

therefore by Chebyshev's inequality, for any $\varepsilon > 0$ and $k \in \mathbb{N}$

$$\mu\big(\{x \in X \mid |\Psi_{N_k}(x) - 1| > \varepsilon\}\big) \leq \frac{\operatorname{Var}(\Psi_{N_k})}{\varepsilon^2} \leq \frac{Q}{\varepsilon^2 (c_{N_k})^q} \leq \frac{Q}{\varepsilon^2 L^q k^2}\,.$$

Thus the family $\{x \in X \mid |\Psi_{N_k}(x) - 1| > \varepsilon\}$ is summable, so by (1.1), $\#\{k \in \mathbb{N} \mid |\Psi_{N_k}(x) - 1| > \varepsilon\}$ is finite for μ-a.e. $x \in X$. Since ε is arbitrarily small, it follows that $\Psi_{N_k} \underset{\mu\text{-a.e.}}{\longrightarrow} 1$ as $k \to \infty$.

Finally we take any other sequence $\{m_j\}$ with $m_j \to \infty$ as $j \to \infty$, and for any j, define $k(j) \stackrel{\text{def}}{=} \max\{k \mid N_k \le m_j\}$; clearly $k(j) \to \infty$ as $j \to \infty$. Then μ-almost everywhere one has

$$\Psi_{m_j} = \frac{S_{m_j}}{c_{m_j}} \le \frac{S_{N_{k(j)+1}}}{c_{N_{k(j)}}} = \frac{c_{N_{k(j)+1}}}{c_{N_{k(j)}}} \cdot \frac{S_{N_{k(j)+1}}}{c_{N_{k(j)+1}}} \le \frac{L\big(k(j)+1\big)^{2/q} + L}{Lk(j)^{2/q}} \Psi_{N_{k(j)+1}}$$
$$= \left(\left(1 + \frac{1}{k(j)}\right)^{2/q} + \frac{1}{k(j)^{2/q}}\right) \Psi_{N_{k(j)+1}} \to 1 \text{ as } j \to \infty\,. \qquad \text{Similarly,}$$
$$\Psi_{m_j} \ge \frac{S_{N_{k(j)}}}{c_{N_{k(j)+1}}} \ge \frac{k(j)^{2/q}}{\big(k(j)+1\big)^{2/q} + 1} \Psi_{N_{k(j)}} \to 1 \text{ as } j \to \infty\,,$$

which completes the proof of (3.1). □

3.2. For two families $\mathcal{H} = \{h_n \mid n \in \mathbb{N}\}$ and $\mathcal{H}' = \{h'_n \mid n \in \mathbb{N}\}$ with weight functions w and w' respectively and for $c \le 1$, say that $\mathcal{H}'$ is *c-majorated* by $\mathcal{H}$ (denoted by $\mathcal{H}' \underset{c}{\subset} \mathcal{H}$) if

$$h'_n \le h_n \quad \text{and} \quad c \cdot w(n) \le w'(n) \quad \text{for all } n \in \mathbb{N}\,.$$

As before, we will say that families of sets have the above property if the families of their characteristic functions do. From the above definition it follows that if $\mathcal{H}' \underset{c}{\subset} \mathcal{H}$, $\mathcal{H}$ is bounded (resp. summable) if and only if $\mathcal{H}'$ is too.

Theorem. *Let (X, μ) be a probability space, $\mathcal{H} = \{h_n \mid n \in \mathbb{N}\}$ a bounded non-summable family of functions on X.*

(a) *If for $c \le 1$ there exists a strongly* 1-KS *family $\mathcal{H}'$ c-majorated by $\mathcal{H}$, then $\liminf_{N\to\infty} \Psi_{\mathcal{H},N} \ge c$ μ-almost everywhere; in particular,* (2.2) *holds.*

(b) *If for $c \le 1$ there exists a strongly* 1-KS *family $\mathcal{H}'$ such that $\mathcal{H}$ is c-majorated by $\mathcal{H}'$, then $\limsup_{N\to\infty} \Psi_{\mathcal{H},N}(x) \le 1/c$ μ-almost everywhere.*

(c) *If for any $c < 1$ there exist two strongly* 1-KS *families $\mathcal{H}'$ and $\mathcal{H}''$ such that $\mathcal{H}' \underset{c}{\subset} \mathcal{H} \underset{c}{\subset} \mathcal{H}''$, then* (3.1) *holds.*

Proof. Assuming $\mathcal{H}' \underset{c}{\subset} \mathcal{H}$, one has μ-almost everywhere

$$\frac{1}{c}\Psi_{\mathcal{H},N} = \frac{1}{c}\frac{\sum_{n=1}^N h_n}{\sum_{n=1}^N w(n)} \ge \frac{\sum_{n=1}^N h_n}{\sum_{n=1}^N w'(n)} \ge \frac{\sum_{n=1}^N h'_n}{\sum_{n=1}^N w'(n)} = \Psi_{\mathcal{H}',N} \underset{\text{by Theorem 3.1}}{\longrightarrow} 1$$

as $N \to \infty$, hence (a). The proof of (b) is completely similar, and (c) is clearly a direct consequence of (a) and (b). □

§4. Correlation decay implies Kochen-Stone

4.1. Let (X, μ) be a probability space, $\mathcal{H}$ a family of functions on X with weight function w. For $M \ge 1$ and a function $\alpha : \mathbb{Z}_+ \to \mathbb{R}_+$, we will say that $\mathcal{H}$ *has property* (M-α-CD)[4] if for any $m, n \in \mathbb{N}$

$$(h_m, h_n) \le Mw(m)w(n) + \alpha(|m-n|)\sqrt{w(m)w(n)}\,.$$

Similarly, we will say that a family of sets *has property* (M-α-CD) if the family of their characteristic functions does.

[4] This is an abbreviation for "correlation decay". The reader should however bear in mind that this property, especially if $M > 1$, has little to do with actual decay of correlation coefficients.

Example. Take $X = [0,1]$, μ the Lebesgue measure, g the map $x \to 2x \mod 1$ and $l : \mathbb{N} \to \mathbb{Z}_+$ any function. Consider the family $\mathcal{A} = \{[0, 2^{-l(n)}] \mid n \in \mathbb{N}\}$. Then the twisted family $\mathcal{A}^g$ has property (1-α-CD) with $\alpha(k) = 2^{-k/2}$. Indeed, take $m \geq n \in \mathbb{N}$ and denote $i = l(m)$ and $j = l(n)$. Then $w(m) = 2^{-i}$ and $w(n) = 2^{-j}$, and one can easily see that for $i \leq j$, $\mu\left(g^{-m}[0, 2^{-i}] \cap g^{-n}[0, 2^{-j}]\right) = \mu\left(g^{-(m-n)}[0, 2^{-i}] \cap [0, 2^{-j}]\right)$ is equal to

$$\begin{cases} 2^{-j}, & 0 \leq m-n \leq j-i \\ 2^{-i} 2^{-\min(m-n,j)}, & m-n > j-i \end{cases} \leq 2^{-j} 2^{-i} + 2^{-(m-n)/2} \sqrt{2^{-i} 2^{-j}}\,,$$

while for $i > j$

$$\mu\left(g^{-(m-n)}[0, 2^{-i}] \cap [0, 2^{-j}]\right) = 2^{-i} 2^{-\min(m-n,j)} \leq 2^{-i} 2^{-j} + 2^{-(m-n)} \sqrt{2^{-i} 2^{-j}}\,.$$

4.2. Theorem. *Let (X, μ) be a probability space, $\mathcal{H}$ a family of functions on X with weight function w, $M \geq 1$ and $\alpha : \mathbb{Z}_+ \to \mathbb{R}_+$ with $\sum_{k=0}^{\infty} \alpha(k) < \infty$. Assume that $\mathcal{H}$ has property (M-α-CD). Then*

(a) *$\mathcal{H}$ is C-Kochen-Stone for $C = M + \dfrac{2\sum_{k=0}^{\infty} \alpha(k)}{\sum_{n=1}^{\infty} w(n)}$;*

(b) *if $\mathcal{H}$ is nonsummable, it is strongly M-Kochen-Stone.*

4.3. Corollary. *For a probability space (X, μ) and a bounded nonsummable family $\mathcal{H}$ of functions on X, assume that $\mathcal{H}$ has property* (1-α-CD) *for some $\alpha : \mathbb{Z}_+ \to \mathbb{R}_+$ with $\sum_{k=0}^{\infty} \alpha(k) < \infty$. Then* (3.1) *holds.*

Proof. Combine Theorems 4.2 and 3.1. ☐

4.4. Corollary. *For a random sequence x of zeroes and ones, let $Z_n(x)$ be the number of successive zeroes after the nth position, and let $l : \mathbb{N} \to \mathbb{Z}_+$ be any function such that $\sum_{n=1}^{\infty} 2^{-l(n)} = \infty$. Then*

$$\lim_{N\to\infty} \frac{\#\{n \in \{1, \ldots, N\} \mid Z_n(x) \geq l(n)\}}{\sum_{n=1}^{N} 2^{-l(n)}} = 1$$

for almost all sequences x.

Proof. Take $X = [0,1]$, $g : x \to 2x \mod 1$ and $\mathcal{A} = \left\{[0, 2^{-l(n))}] \mid n \in \mathbb{N}\right\}$. It was shown in Example 2.1 that $\mathcal{A}^g$ has property (1-α-CD) for $\alpha(k) = 2^{-k/2}$, so one can apply Corollary 4.3 to get the desired result. ☐

4.5. *Proof of Theorem 4.2.* With the notation introduced in the proof of Theorem 3.1, one has for any $N \in \mathbb{N}$

$$\int_X (S_N)^2 \, d\mu = \int_X \left(\sum_{n=1}^{N} h_n\right)^2 d\mu = \sum_{m,n=1}^{N} (h_m, h_n)$$

$$\text{(by (1-}\alpha\text{-CD))} \quad \leq \sum_{m,n=1}^{N} w(m)w(n) + \sum_{m,n=1}^{N} \alpha(|m-n|) \sqrt{w(m)w(n)}\,.$$

The first summand in the right hand side is clearly equal to $(c_N)^2$, while the second one is not greater than

$$2\sum_{k=0}^{N-1}\alpha(k)\sum_{m=1}^{N-k}\sqrt{w(m)w(m+k)} \underset{\text{(by Schwarz)}}{\leq} 2\sum_{k=0}^{N-1}\alpha(k)\left(\sum_{m=1}^{N-k}w(m)\sum_{m=1}^{N-k}w(m+k)\right)^{1/2},$$

which, in its turn, is not greater than $2c_N\sum_{k=0}^{\infty}\alpha(k)$. Therefore

$$\int_X(\Psi_N)^2\,d\mu = \frac{1}{(c_N)^2}\int_X(S_N)^2\,d\mu \leq 1+\frac{2\sum_{k=0}^{\infty}\alpha(k)}{c_N},$$

which proves both (a) and (b). $\square$

§5. Applications to dynamical systems and Diophantine approximation

5.1. Let X, μ, g and φ be as in any of the Examples 1.5, 1.6 or 1.7. In the sequel to this paper we will show that there exist c, $0<c<1$, and a summable function $\alpha(k)$ such that if $f:\mathbb{N}\to\mathbb{R}_+$ is any function and $\mathcal{A}=\mathcal{A}_{\varphi,f}$ is the corresponding family of sets, one can find two families $\mathcal{H}'$ and $\mathcal{H}''$ of functions on X such that $\mathcal{H}'\underset{c}{\subset}\mathcal{A}\underset{c}{\subset}\mathcal{H}''$, and both $(\mathcal{H}')^g$ and $(\mathcal{H}'')^g$ have property (1-α-CD). Furthermore, in Examples 1.5 and 1.7 one can choose c as close to 1 as one wishes. If $\sum_{n=1}^{\infty}\big(f(n)\big)^d=\infty$, this, in view of Theorems 4.2 and 3.2, will describe asymptotics of $\#\{1\leq n\leq N \mid \varphi(g^n x)\leq f(n)\}$. The choice $f(n)=n^{-1/d}$ gives lower estimates for the limits in (1), (2) and (3), with

$$\begin{cases} \#\{1\leq n\leq N \mid \operatorname{dist}(\gamma_n(p,v),p)\geq \frac{1}{d}\log n\} & \text{in Theorem 1}\\ \#\{1\leq n\leq N \mid \operatorname{dist}(g^n x,0)\leq n^{-1/d}\} & \text{in Theorem 2}\\ \#\{1\leq n\leq N \mid \Delta(g^n x)\geq \frac{1}{d}\log n\} & \text{in Theorem 3 assuming (iii}')\end{cases}$$

being almost surely bounded between $\mathrm{const}_1\cdot\log N$ and $\mathrm{const}_2\cdot\log N$.

In the setting of Theorem 2 (Example 1.6), this can be done via approximation of the sets from $\mathcal{A}$ by products of m_i-adic intervals and using the argument outlined in Example 4.1. On the other hand, in the setting of Theorem 3 / Example 1.7 (and hence Theorem 1 / Example 1.5), one can use the assumptions (i) and (ii) on Δ to approximate the sets from $\mathcal{A}$ by smooth functions, and then apply estimates from [KM, §2.4] on exponential decay of their correlation coefficients.

5.2. One can also use the special case $G=SL_d(\mathbb{R})$, $\Gamma=SL_d(\mathbb{R})$, $\Delta(\Lambda)=\max_{\mathbf{x}\in\Lambda\smallsetminus\{0\}}\log\left(\frac{1}{\|\mathbf{x}\|}\right)$ of the setting of Theorem 3 to get certain number-theoretic applications of Theorem 3.2, for example, prove

Khintchine's Diverence Theorem (see [Kh, C])**.** *Let $\psi:\mathbb{N}\to\mathbb{R}_+$ be a nonincreasing function such that $\sum_{n=0}^{\infty}\psi(n)^k=\infty$. Then the set of inequalities*

$$|\alpha_i n-m_i|<\psi(n)\,,\quad 1\leq i\leq k$$

has infinitely many solutions $n\in\mathbb{N}$ and $m_i\in\mathbb{Z}$ for a.e. $(\alpha_1,\dots,\alpha_k)\in\mathbb{R}^k$.

This can be done by reducing this theorem to a problem involving flows in the space of unimodular lattices in $\mathbb{R}^{k+1}$ via the correspondence first discovered by S.G. Dani ([Da], cf. also [K]). Note that the full strength of W. Schmidt's [Sc] generalization of this theorem is so far beyond our reach; however a substantial part of Schmidt's result can be derived by our methods. Other related facts (some of them new) from metrical theory of Diophantine approximation can also be obtained.

Acknowledgements

The first named author is grateful to Alfred P. Sloan Foundation for financial support, and to Yuval Peres for helpful discussions.

References

[C] J. W. S. Cassels, *An introduction to Diophantine approximation*, Cambridge Tracts in Math., vol. 45, Cambridge Univ. Press, Cambridge, 1957.

[Da] S. G. Dani, *Divergent trajectories of flows on homogeneous spaces and Diophantine approximation*, J. Reine Angew. Math. **359** (1985), 55–89.

[Du] R. Durrett, *Probability: theory and examples*, Wadsworth & Brooks/Cole, 1991.

[K] D. Kleinbock, *Flows on homogeneous spaces and Diophantine properties of matrices*, preprint (1995).

[Kh] A. Khintchine, *Zur metrischen Theorie der Diophantischen Approximationen*, Math. Z. **24** (1926), 706–714.

[KM] D. Y. Kleinbock and G. A. Margulis, *Bounded orbits of nonquasiunipotent flows on homogeneous spaces*, Adv. in Soviet Math., Amer. Math. Soc., Providence, RI (to appear).

[KSp] A. Katok and R. Spatzier, *First cohomology of Anosov actions of higher rank Abelian groups and applications to rigidity*, Inst. Hautes Études Sci. Publ. Math. **79** (1994), 131–156.

[KSt] S. Cochen and C. Stone, *A note on the Borel-Cantelli lemma*, Ill. J. Math. **8** (1964), 248-251.

[Mo] C. C. Moore, *Exponential decay of correlation coefficients for geodesic flows*, Group Representations, Ergodic Theory, Operator Algebras and Mathematical Physics, Math. Sci. Res. Inst. Publ., vol. 6, Springer-Verlag, Berlin and New York, 1987, pp. 163–181.

[Sc] W. M. Schmidt, *A metrical theorem in Diophantine approximation*, Canadian J. Math. **12** (1960), 619–631.

[Sp] F. Spitzer, *Principles of random walk*, Van Nostrand, Princeton, 1964.

[Su] D. Sullivan, *Disjoint spheres, approximation by imaginary quadratic numbers, and the logarithm law for geodesics*, Acta Math. **149** (1982), 215-237.

D. Y. Kleinbock, Department of Mathematics, Yale University, New Haven, CT 06520

E-mail address: `kleinboc@math.yale.edu`

G. A. Margulis, Department of Mathematics, Yale University, New Haven, CT 06520

E-mail address: `margulis@math.yale.edu`