Limit theorems for toral translations

Dmitry Dolgopyat and Bassam Fayad

Contents

The authors thank Mariusz Lemańczyk, Jens Marklof, Jean Paul Thouvenot, and Ilya Vinogradov for useful discussions. The research of DD was partially supported by the NSF. The research of BF was partially supported by ANR-11-BS01-0004.

1. Introduction

One of the surprising discoveries of dynamical systems theory is that many deterministic systems with non-zero Lyapunov exponents satisfy the same limit theorems as the sums of independent random variables. Much less is known for the zero exponent case where only a few examples have been analyzed. In this survey we consider the extreme case of toral translations where each map not only has zero exponents but is actually an isometry. These systems were studied extensively due to their relations to number theory, to the theory of integrable systems and to geometry. Surprisingly many natural questions are still open. We review known results as well as the methods to obtain them and present a list of open problems. Given a vast amount of work on this subject, it is impossible to provide a comprehensive treatment in this short survey. Therefore we treat the topics closest to our research interests in more detail while some other subjects are mentioned only briefly. Still we hope to provide the reader with the flavor of the subject and introduce some tools useful in the study of toral translations, most notably, various renormalization techniques.

Let $X = \mathbb{T}^d$, μ be the Haar measure on X and $T_\alpha(x) = x + \alpha$.

The most basic question in smooth ergodic theory is the behavior of ergodic sums. Given a map T and a zero mean observable $A(\cdot)$ let

$$A_N(x) = \sum_{n=0}^{N-1} A(T^n x) \tag{1}$$

If there is no ambiguity, we may write A_N for $A_N(x)$. Conversely we may use the notation $A_N(\alpha, x)$ to indicate that the underlying map is the translation of vector α. The uniform distribution of the orbit of x by T is characterized by the convergence to 0 of $A_N(x)/N$. In the case of toral translations T_α with irrational frequency vector α the uniform distribution holds for all points x. The study of the ergodic sums is then useful to quantify the rate of uniform distribution of the Kronecker sequence $n\alpha$ mod 1 as we will see in Section 3 where discrepancy functions are discussed. The question about the distribution of ergodic sums is analogous to the the Central Limit Theorem in probability theory. One can also consider analogues of other classical probabilistic results. In this survey we treat two such questions. In Section 4 we consider so called Poisson regime where (1) is replaced by $\sum_{n=0}^{N-1} \chi_{\mathcal{C}_N}(T_\alpha^n x)$ and the sets $\mathcal{C}_N$ are scaled in such a way that only finite number of terms are non-zero for typical x. Such sums appear in several questions in mathematical physics, including quantum chaos [**91**] and Boltzmann-Grad limit of several mechanical systems [**93**]. They also describe the resonances in the study of ergodic sums for toral translations as we will see in Section 7. In Section 8 we consider Borel-Cantelli type questions where one takes a sequence of shrinking sets and studies a number of times a typical orbit hits whose sets. These questions are intimately related to some classical problems in the theory of Diophantine approximations.

The ergodic sums above toral translations also appear in natural dynamical systems such as skew products, cylindrical cascades and special flows. Discrete time systems related to ergodic sums over translations are treated in Section 9 while flows are treated in Section 10. These systems give additional motivation to study the ergodic sums (1) for smooth functions having singularities of various types: power, fractional power, logarithmic... Ergodic sums for functions with singularities are discussed in Section 2. Finally in Section 11 we present the results related to action of several translations at the same time.

Notations.

We say that a vector $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{R}^d$ is irrational if $\{1, \alpha_1, \dots, \alpha_d\}$ are linearly independent over $\mathbb{Q}$.

For $x \in \mathbb{R}^d$, we use the notation $\{x\} := (x_1, \dots, x_d)$ mod (1). We denote by $\|x\|$ the closest signed distance of some $x \in \mathbb{R}$ to the integers.

Assuming that $d \in \mathbb{N}$ is fixed, for $\sigma > 0$ we denote by $\mathcal{D}(\sigma) \subset \mathbb{R}^d$ the set of Diophantine vectors with exponent σ, that is

$$\mathcal{D}(\sigma) = \{\alpha : \exists C \ \ \forall k \in \mathbb{Z}^d - 0, m \in \mathbb{Z} \quad |(k, \alpha) - m| \geq C|k|^{-d-\sigma}\} \tag{2}$$

Let us recall that $\mathcal{D}(\sigma)$ has a full measure if $\sigma > 0$, while $\mathcal{D}(0)$ is an uncountable set of zero measure and $\mathcal{D}(\sigma)$ is empty for $\sigma < 0$. The set $\mathcal{D}(0)$ is called the set of constant type vector or badly approximable vectors. An irrational vector α that is not Diophantine for any $\sigma > 0$ is called Liouville.

We denote by $\mathfrak{C}$ the standard Cauchy random variable with density $\frac{1}{\pi(1+x^2)}$. Normal random variable with zero mean and variance D^2 will be denoted by $\mathfrak{N}(D^2)$. Thus $\mathfrak{N}(D^2)$ has density $\frac{1}{2\pi D}e^{-x^2/2D^2}$. We will write simply $\mathfrak{N}$ for $\mathfrak{N}(1)$. Next, $\mathfrak{P}(X, \mu)$ will denote the Poisson process on X with measure μ (we refer the reader to Section 5 for the definition and basic properties of Poisson processes).

2. Ergodic sums of smooth functions with singularities

2.1. Smooth observables. For toral translations, the ergodic sums of smooth observables are well understood. Namely if A is sufficiently smooth with zero mean then for almost all α, A is a coboundary, that is, there exists $B(\alpha, x)$ such that

$$A(x) = B(x + \alpha, \alpha) - B(x, \alpha). \tag{3}$$

Namely if $A(x) = \sum_{k\neq 0} a_k e^{2\pi i(k,x)}$ then

$$B(\alpha, x) = \sum_{k\neq 0} b_k e^{2\pi i(k,x)} \text{ where } b_k = \frac{a_k}{e^{i2\pi(k,\alpha)} - 1}.$$

The above series converges in L^2 provided $\alpha \in \mathcal{D}(\sigma)$ and $A \in \mathcal{H}^\sigma = \{A : \sum_k |a_k|k|^{(\sigma+d)}|^2 < \infty\}$. Note that (3) implies that

$$A_N(x) = B(x + N\alpha, \alpha) - B(x, \alpha)$$

giving a complete description of the behavior of ergodic sums for almost all α. In particular we have

Corollary 1. *If α is uniformly distributed on $\mathbb{T}^d$ then $A_N(x)$ has a limiting distribution as $N \to \infty$, namely*

$$A_N \Rightarrow B(y, \alpha) - B(x, \alpha)$$

where (y, α) is uniformly distributed on $\mathbb{T}^d \times \mathbb{T}^d$.

Proof. We need to show that as $N \to \infty$ the random vector $(\alpha, N\alpha)$ converge to a vector with coordinates independent random variables uniformly distributed on $\mathbb{T}^d \times \mathbb{T}^d$. To this end it suffices to check that if $\phi(x, y)$ is a smooth function on $\mathbb{T}^d \times \mathbb{T}^d$ then

$$\lim_{N\to\infty} \int_{\mathbb{T}^d} \phi(\alpha, N\alpha) d\alpha = \int_{\mathbb{T}^d\times\mathbb{T}^d} \phi(\alpha, \beta) d\alpha d\beta$$

but this is easily established by considering the Fourier series of ϕ. $\square$

We will see in Section 10 how our understanding of ergodic sums for smooth functions can be used to derive ergodic properties of area preserving flows on $\mathbb{T}^2$ without fixed points.

On the other hand there are many open questions related to the case when the observable A is not smooth enough for (3) to hold. Below we mention several classes of interesting observables.

2.2. Observables with singularities. Special flows above circle rotations and under ceiling functions that are smooth except for some singularities naturally appear in the study of conservative flows on surfaces with fixed points.

Another motivation for studying ergodic sums for functions with singularities is the case of meromorphic functions, whose sums appear in questions related to both number theory [**48**] and ergodic theory [**106**].

2.2.1. *Observables with logarithmic singularities.* In the study of conservative flows on surfaces, non degenerate saddle singularities are responsible for logarithmic singularities of the ceiling function.

Ceiling functions with logarithmic singularities also appear in the study of multi-valued Hamiltonians on the two torus. In [**3**], Arnold investigated such flows and showed that the torus decomposes into cells that are filled up by periodic orbits and one open ergodic component. On this component, the flow can be represented as a special flow over an interval exchange map of the circle and under a ceiling function that is smooth except for some logarithmic singularities. The singularities can be asymmetric since the coefficient in front of the logarithm is twice as big on one side of the singularity as the one on the other side, due to the existence of homoclinic loops (see Figure 1).

More motivations for studying function with logarithmic singularities as well as some numerical results for rotation numbers of bounded type are presented in [**69**].

A natural question is to understand the fluctuations of the ergodic sums for these functions as the frequency α of the underlying rotation is random as well as the base point x. Since Fourier coefficients of the symmetric logarithm function have the asymptotics similar to that of the indicator function of an interval one may expect that the results about the latter that we will discuss in Section 3 can be extended to the former.

QUESTION 1. *Suppose that A is smooth away from a finite set of points $x_1, x_2 \dots x_k$ and near x_j, $A(x) = a_j^{\pm} \ln |x - x_j| + r_j(x)$ where $-$ sign is taken if $x < x_j$, $+$ sign is taken if $x > x_j$ and r_j are smooth*

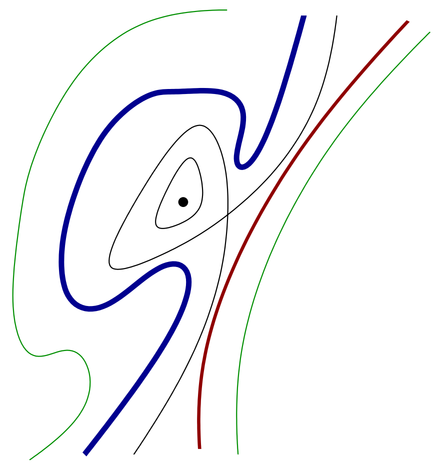

FIGURE 1. Multivalued Hamiltonian flow. Note that the orbits passing to the left of the saddle spend approximately twice longer time comparing to the orbits passing to the right of the saddle and starting at the same distance from the separatrix since they pass near the saddle twice.

functions. What can be said about the distribution of $A_N(\alpha, x)/\ln N$ *as* x *and* α *are random?*

2.2.2. *Observables with power like singularities.* When considering conservative flows on surfaces with degenerate saddles one is led to study the ergodic sums of observables with integrable power like singularities (more discussion of these flows will be given in Section 10). Special flows above irrational rotations of the circle under such ceiling functions are called Kocergin flows.

The study of ergodic sums for smooth ergodic flows with nondegenerate hyperbolic singular points on surfaces of genus $p \geq 2$ shows that these flows are in general not mixing (see Section 10). *A contrario* Kocergin showed that special flows above irrational rotations and under ceiling functions with integrable power like singularities are always mixing. This is due to the important deceleration next to the singularity that is responsible for a *shear* along orbits that separates the points sufficiently to produce mixing. In other words, the mixing is due to large oscillations of the ergodic sums. In this note we will be frequently interested in the distribution properties of these sums.

One may also consider the case of non-integrable power singularities since they naturally appear in problems of ergodic theory and number theory. The following result answers a question of [**48**].

THEOREM 2. *([**120**]) If A has one simple pole on $\mathbb{T}^1$ and (α, x) is uniformly distributed on $\mathbb{T}^2$ then $\frac{A_N}{N}$ has limiting distribution as $N \to \infty$.*

The function A in Theorem 2 has a symmetric singularity of the form $1/x$ that is the source of cancellations in the ergodic sums.

QUESTION 2. *What happens for an asymmetric singularity of the type $1/|x|$?*

QUESTION 3. *What happens in the quenched setting where α is fixed?*

We now present several generalizations of Theorem 2.

THEOREM 3. *Let $A = \tilde{A}(x) + \frac{c_- \chi_{x<x_0} + c^+ \chi_{x>x_0}}{|x-x_0|^a}$ where $\tilde{A}$ is smooth and $a > 1$.*

(a) If (α, x) is uniformly distributed on $\mathbb{T}^2$ then $\dfrac{A_N}{N^a}$ converges in distribution.

(b) For almost every x fixed, if α is uniformly distributed on $\mathbb{T}$ then $\dfrac{A_N(\alpha, x)}{N^a}$ converge to the same limit as in part (a).

THEOREM 4. [**87**] *If A has zero mean and is smooth except for a singularity at 0 of type $|x|^{-a}$, $a \in (0, 1)$ then A_N/N^a converges in distribution.*

The proof of Theorem 4 is inspired by the proof of Theorem 10 of Section 3 which will be presented in Section 7.3.

Marklof proved in [**92**] that if $\alpha \in \mathcal{D}(\sigma)$ with $\sigma < (1-a)/a$, then for A as in Theorem 4 $A_N(\alpha, \alpha)/N \to 0$.

QUESTION 4. *What happens for other angles α and other type of singularities, including the non integrable ones for which the ergodic theorem does not necessarily hold.*

Another natural generalization of Theorem 2 is to consider meromorphic functions. Let A be such a function with highest pole of order m. Thus A can be written as

$$A(x) = \sum_{j=1}^{r} \frac{c_j}{(x - x_j)^m} + \tilde{A}(x)$$

where the highest pole of $\tilde{A}$ has order at most $m - 1$.

Theorem 5. *(a) Let A be fixed and let α be distributed according to a smooth density on $\mathbb{T}$. Then for any $x \in \mathbb{T}$, $\dfrac{A_N(\alpha, x)}{N^m}$ has a limiting distribution as $N \to \infty$.*

(b) Let $\tilde{A}, c_1, \ldots c_r$ be fixed while $(\alpha, x, x_1 \ldots x_r)$ are distributed according to a smooth density on on $\mathbb{T}^{r+2}$ then $\dfrac{A_N(\alpha, x)}{N^m}$ has a limiting distribution as $N \to \infty$.

(c) If $(x_1, x_2 \ldots x_r)$ is a fixed irrational vector then for almost every $x \in \mathbb{T}$ the limit distribution in part (a) is the same as the limit distribution in part (b).

Proofs of Theorems 3 and 5 are sketched in Section 7.

It will be apparent from the proof of Theorem 5 that the limit distribution in part (a) is not the same **for all** $x_1, x_2 \ldots x_j$. For example if $x_j = jx_1$ leads to an exceptional distribution since a close approach to x_1 and x_2 by the orbit of x should be followed by a close approach to x_j for $j \geq 3$. We will see that phenomena appears in many limit theorems (see e.g Theorem 9, Theorem 25 and Question 52, Theorem 38 and Question 38, as well as [**93**]).

Question 5. *What can be said about more general meromorphic functions such as $\sin 2\pi x/(\sin 2\pi x + 3\cos 2\pi y)$ on $\mathbb{T}^d$ with $d > 1$?*

3. Ergodic sums of characteristic functions. Discrepancies

The case where $A = \chi_\Omega$ is a classical subject in number theory. Define the discrepancy function

$$D_N(\Omega, \alpha, x) = \sum_{n=0}^{N-1} \chi_\Omega(x + n\alpha) - N\frac{\mathrm{Vol}(\Omega)}{\mathrm{Vol}(\mathbb{T}^d)}.$$

Uniform distribution of the sequence $x + k\alpha$ on $\mathbb{T}^d$ is equivalent to the fact that, for regular sets Ω, $D_N(\Omega, \alpha, x)/N \to 0$ as $N \to \infty$. A step further in the description of the uniform distribution is the study of the rate of convergence to 0 of $D_N(\Omega, \alpha, x)/N$.

In $d = 1$ it is known that if $\alpha \in \mathbb{T} - \mathbb{Q}$ is fixed, the discrepancy $D_N(\Omega, \alpha, x)/N$ displays an oscillatory behavior according to the position of N with respect to the denominators of the best rational approximations of α. A great deal of work in Diophantine approximation has been done on estimating the discrepancy function in relation with the arithmetic properties of $\alpha \in \mathbb{T}$, and more generally for $\alpha \in \mathbb{T}^d$.

3.1. The maximal discrepancy. Let

$$\overline{D}_N(\alpha) = \sup_{\Omega \in \mathbb{B}} D_N(\Omega, \alpha, 0) \tag{4}$$

where the supremum is taken over all sets Ω in some natural class of sets $\mathbb{B}$, for example balls or boxes (product of intervals).

The case of (straight) boxes was extensively studied, and growth properties of the sequence $\overline{D}_N(\alpha)$ were obtained with a special emphasis on their relations with the Diophantine approximation properties of α. In particular, following earlier advances of [**75, 48, 99, 64, 117**] and others, [**8**] proves

THEOREM 6. *Let*

$$\overline{D}_N(\alpha) = \sup_{\Omega - box} D_N(\Omega, \alpha, 0)$$

Then for any positive increasing function ϕ we have

(5)

$$\sum_n \phi(n)^{-1} < \infty \iff \frac{\overline{D}_N(\alpha)}{(\ln N)^d \phi(\ln \ln N)} \begin{array}{l} \textit{is bounded for} \\ \textit{almost every } \alpha \in \mathbb{T}^d. \end{array}$$

In dimension $d = 1$, this result is the content of Khinchine theorems obtained in the early 1920's [**64**], and it follows easily from well-known results from the metrical theory of continued fractions (see for example the introduction of [**8**]). The higher dimensional case is significantly more difficult and the cited bound was only obtained in the 1990s.

The bound in (5) focuses on how bad can the discrepancy become along a subsequence of N, for a fixed α in a full measure set. In a sense, it deals with the worst case scenario and does not capture the oscillations of the discrepancy. On the other hand, the restriction on α is necessary, since given any $\varepsilon_n \to 0$ it is easy to see that for $\alpha \in \mathbb{T}$ *sufficiently Liouville*, the discrepancy (relative to intervals) can be as bad as $N_n \varepsilon_n$ along a suitable sequence N_n (large multiples of Liouville denominators).

For $d = 1$, it is not hard to see, using continued fractions, that for any $\alpha :$ $\limsup \frac{\overline{D}_N(\alpha)}{\ln N} > 0$, $\liminf \overline{D}_N(\alpha) \leq C$; and for $\alpha \in \mathcal{D}(0)$ $\limsup \frac{\overline{D}_N(\alpha)}{\ln N} < +\infty$. The study of higher dimensional counterparts to these results raises several interesting questions.

QUESTION 6. *Is it true that* $\limsup \frac{\overline{D}_N(\alpha)}{\ln^d N} > 0$ **for all** $\alpha \in \mathbb{T}^d$*?*

QUESTION 7. *Is it true that there exists α such that* $\limsup \frac{\overline{D}_N(\alpha)}{\ln^d N} < +\infty$*?*

QUESTION 8. *What can one say about* $\liminf \frac{\overline{D}_N(\alpha)}{a_N}$ *for a.e.* α, *where* a_N *is an adequately chosen normalization? for every* α*?*

QUESTION 9. *Same questions as Questions 6–8 when boxes are replaced by balls.*

QUESTION 10. *Same questions as Questions 6–8 for the isotropic discrepancy, when boxes are replaced by the class of all convex sets* [**79**].

3.2. Limit laws for the discrepancy as α is random. In this survey, we will mostly concentrate on the distribution of the discrepancy function as α is random. The above discussion naturally raises the following question.

QUESTION 11. *Let* α *be uniformly distributed on* $\mathbb{T}^d$. *Is it true that* $\frac{\overline{D}_N(\alpha)}{\ln^d N}$ *converges in distribution as* $N \to \infty$*?*

Why do we need to take α random? The answer is that for fixed α the discrepancy does not have a limit distribution.

For example for $d = 1$ the Denjoy-Koksma inequality says that

$$|A_{q_n} - q_n \int A(x)dx| \leq 2V$$

where q_n is the n-th partial convergent to α and V denotes the total variation of A. In particular $D_{q_n}(I, \alpha, x)$ can take at most 3 values.

In higher dimensions one can show that if Ω is either a box or any other convex set then for almost all α and almost all tori, when x is random the variable

$$\frac{D_N(\Omega, \mathbb{R}^d/L, \alpha, \cdot)}{a_N}$$

does not converge to a non-trivial limiting distribution for any choice of $a_N = a_N(\alpha, L)$ (see discussion in the introduction of [**29**]).

QUESTION 12. *Is this true* **for all** α, L*?*

QUESTION 13. *Study the distributions which can appear as weak limits of* $\dfrac{D_N(\Omega, \alpha, \cdot)}{a_N}$, *in particular their relation with number theoretic properties of* α.

Let us consider the case $d = 1$ (so the sets of interest are intervals and we will write I instead of Ω.) It is easy to see that all limit distributions are atomic for all I iff $\alpha \in \mathbb{Q}$.

QUESTION 14. *Is it true that all limit distributions are either atomic or Gaussian for almost all* I *iff* α *is of bounded type?*

Evidence for the affirmative answer is contained in the following results.

THEOREM 7. *([**55**]) If $\alpha \notin \mathbb{Q}$ and $I = [0, 1/2]$ then there is a sequence N_j such that* $\dfrac{D_{N_j}(I, \alpha, \cdot)}{j}$ *converges to* $\mathfrak{N}$.

Instead of considering subsequences, it is possible to randomize N.

THEOREM 8. *Let α be a quadratic surd.*

*(a) ([**10**]) If (x, a, l) is uniformly distributed on $\mathbb{T}^3$ then* $\dfrac{D_{[aN]}([0, l], \alpha, x)}{\sqrt{\ln N}}$ *converges to $\mathfrak{N}(\sigma^2)$ for some $\sigma^2 \neq 0$.*

*(b) ([**11**]) If M is uniformly distributed on $[1, N]$ and l is rational then there are constants $C(\alpha, l), \sigma(\alpha, l)$ such that* $\dfrac{D_M([0, l], \alpha, 0) - C(\alpha, l) \ln N}{\sqrt{\ln N}}$ *converges to $\mathfrak{N}(\sigma^2(\alpha, l))$.*

Note that even though we have normalized the discrepancy by subtracting the expected value an additional normalization is required in Theorem 8(b). The reason for this is explained at the end of Section 10.5.

So if one wants to have a unique limit distribution for all N one needs to allow random α.

The case when $d = 1$ was studied by Kesten. Define

$$V(u, v, w) = \sum_{k=1}^{\infty} \frac{\sin(2\pi u) \sin(2\pi v) \sin(2\pi w)}{k^2}.$$

If (r, q) are positive integers let

$$\theta(r, q) = \frac{\mathrm{Card}(j : 0 \leq j \leq q - 1 : gcd(j, r, q) = 1)}{\mathrm{Card}(j, k : 0 \leq j, k \leq q - 1 : gcd(j, k, q) = 1)}.$$

Finally let

$$c(r) = \begin{cases} \frac{\pi^3}{12} \left[\sum_{r=0}^{q-1} \theta(p, q) \int_0^1 \int_0^1 V(u, \frac{rp}{q}, v) du dv \right]^{-1} & \text{if } r = \frac{p}{q} \text{ and } gcd(p, q) = 1 \\ \frac{\pi^3}{12} \left[\int_0^1 \int_0^1 \int_0^1 V(u, r, v) du dr dv \right]^{-1} & \text{if } r \text{ is irrational.} \end{cases}$$

THEOREM 9. *([**61, 62**]) If (α, x) is uniformly distributed on $\mathbb{T}^2$ then* $\frac{D_N([0,l],\alpha,x)}{c(l) \ln N}$ *converges to the Cauchy distribution $\mathfrak{C}$.*

Note that the normalizing factor is discontinuous as a function of the length of the interval at rational values.

A natural question is to extend Theorem 9 to higher dimensions. The first issue is to decide which sets Ω to consider instead of intervals.

It appears that a quite flexible assumption is that Ω is semialgebraic, that is, it is defined by a finite number of algebraic inequalities.

QUESTION 15. *Suppose that Ω is semialgebraic then there is a sequence $a_N = a_N(\Omega)$ such that for a random translation of a random torus $\dfrac{D_N(\Omega, \mathbb{R}^d/L, \alpha, x)}{a_N}$ converges in distribution as $N \to \infty$.*

By random translation of a random torus, we mean a translation of random angle α on a torus $\mathbb{R}^d/L$ where $L = A\mathbb{Z}^d$ and the triple (α, x, A) has a smooth density on $\mathbb{T}^d \times \mathbb{T}^d \times \mathrm{GL}(\mathbb{R}, d)$. Notice that comparing to Kesten's result of Theorem 9, Question 15 allows for additional randomness, namely, the torus is random. In particular, for $d = 1$, the study of the discrepancy of visits to $[0, l]$ on the torus $\mathbb{R}/\mathbb{Z}$ is equivalent to the study of the discrepancy of visits to $[0, 1]$ on the torus $\mathbb{R}/(l^{-1}\mathbb{Z})$. Thus the purpose of the extra randomness is to avoid the irregular dependence on parameters observed in Theorem 9 (cf. also [**109, 110**]).

So far Question 15 has been answered for two classes of sets: strictly convex sets and (tilted) boxes, which includes the two natural counterparts to intervals in higher dimension that are balls and boxes.

Given a convex body Ω, we consider the family Ω_r of bodies obtained from Ω by rescaling it with a ratio $r > 0$ (we apply to Ω the homothety centered at the origin with scale r). We suppose $r < r_0$ so that the rescaled bodies can fit inside the unit cube of $\mathbb{R}^d$. We define

$$D_N(\Omega, r, \alpha, x) = \sum_{n=0}^{N-1} \chi_{\Omega_r}(x + n\alpha) - N\mathrm{Vol}(\Omega_r) \tag{6}$$

THEOREM 10. *([**28**]) If (r, α, x) is uniformly distributed on $X = [a, b] \times \mathbb{T}^d \times \mathbb{T}^d$ then $\frac{D_N(\Omega, r, \alpha, x)}{r^{(d-1)/2} N^{(d-1)/2d}}$ has a limit distribution as $N \to \infty$.*

The form of the limiting distribution is given in Theorem 18 in Section 7.

In the case of boxes we recover the same limit distribution as in Kesten but with a higher power of the logarithm in the normalization.

THEOREM 11. *([**29**]) In the context of Question 15, if Ω is a box, then $\dfrac{D_N}{c \ln^d N}$ converges to $\mathfrak{C}$ as $N \to \infty$.*

Alternatively, one can consider gilded boxes, namely: for $u = (u_1, \ldots, u_d)$ with $0 < u_i < 1/2$ for every i, we define a cube on the d-torus by $C_u = [-u_1, u_1] \times \ldots [-u_d, u_d]$. Let $\eta > 0$ and MC_u be the image of C_u by a matrix $M \in \mathrm{SL}(d, \mathbb{R})$ such that

$$M = (a_{ij}) \in G_\eta = \{|a_{i,i} - 1|, \text{ for every } i \text{ and } |a_{i,j}| < \eta \text{ for every } j \neq i\}.$$

For a point $x \in \mathbb{T}^d$ and a translation frequency vector $\alpha \in \mathbb{T}^d$ we denote $\xi = (u, M, \alpha, x)$ and define the following discrepancy function

$$D_N(\xi) = \#\{1 \leq m \leq N : (x + m\alpha) \bmod 1 \in MC_u\} - 2^d \left(\Pi_i u_i\right) N.$$

Fix d segments $[v_i, w_i]$ such that $0 < v_i < w_i < 1/2 \forall i = 1, \ldots, d$. Let

$$X = (u, \alpha, x, (a_{i,j})) \in [v_1, w_1] \times \ldots [v_d, w_d] \times \mathbb{T}^{2d} \times G_\eta \tag{7}$$

We denote by $\mathbb{P}$ the normalized restriction of the Lebesgue $\times$ Haar measure on X. Then, the precise statement of Theorem 11 is

THEOREM 12. *([**29**]) Let $\rho = \frac{1}{d!}\left(\frac{2}{\pi}\right)^{2d+2}$. If ξ is distributed according to λ then $\frac{D_N(\xi)}{\rho(\ln N)^d}$ converges to $\mathfrak{C}$ as $N \to \infty$.*

QUESTION 16. *Are Theorems 10–12 valid if*
(a) L is fixed; (b) x is fixed?

QUESTION 17. *Describe large deviations for D_N. That is, given $b_N \gg a_N$ where a_N is the same as in Question 15, study the asymptotics of $\mathbb{P}(D_N \geq b_N)$. One can study this question in the annealed setting when all variables are random or in the quenched setting where some of them are fixed.*

QUESTION 18. *Does a local limit theorem hold? That is, is it true that given a finite interval J we have*

$$\lim_{N\to\infty} a_N \mathbb{P}(D_N \in J) = c|J|?$$

4. Poisson regime

The results presented in the last section deal with the so called CLT regime. This is the regime when, since the target set Ω is *macroscopic* (having volume of order 1), if T was sufficiently mixing, one would get the Central Limit Theorem for the ergodic sums of χ_Ω. In this section we discuss Poisson (*microscopic*) regime, that is, we let $\Omega = \Omega_N$ shrink so that $\mathbb{E}(D_N(\Omega_N, \alpha, x))$ is constant. In this case, the sum in the discrepancy consists of a large number of terms each one of which vanishes with probability close to 1 so that typically only finitely many terms are non-zero.

THEOREM 13. *([**88**]) Suppose that Ω is bounded set whose boundary has zero measure.*

If (α, x) is uniformly distributed on $\mathbb{T}^d \times \mathbb{T}^d$ then both $D_N(N^{-1/d}\Omega, \alpha, x)$ and $D_N(N^{-1/d}\Omega, \alpha, 0)$ converge in distribution.

Note that in this case the result is less sensitive to the shape of Ω than in the case of sets of unit size.

We will see later (Theorem 16 in Section 7) that one can also handle several sets at the same time.

COROLLARY 14. *If (α, x) is uniformly distributed on $\mathbb{T}^d \times \mathbb{T}^d$ then the following random variables have limit distributions*

(a) $N^{1/d} \min\limits_{0 \leq n < N} d(x + n\alpha, \bar{x})$ where $\bar{x}$ is a given point in $\mathbb{T}^d$;

(b) $N^{2/d} \min\limits_{0 \leq n < N} [A(x+n\alpha) - A(\bar{x})]$ where A is a Morse function with minimum at $\bar{x}$.

PROOF. To prove (a) note that $N^{1/d} \min\limits_{0 \leq n < N} d(x + n\alpha, \bar{x}) \leq s$ iff the number of points of the orbit of x of length N inside $B(\bar{x}, sN^{-1/d})$ is zero.

To prove (b) note that if A is a Morse function and x is close to $\bar{x}$ then $A(x) \approx A(\bar{x}) + (D^2 A)(\bar{x})(x - \bar{x}, x - \bar{x})$. $\square$

There are two natural ways to extend this result.

QUESTION 19. *If $S \subset \mathbb{T}^d$ is an analytic submanifold of codimension q find the limit distribution of $N^{1/q} \min\limits_{0 \leq n < N} d(x + n\alpha, S)$.*

QUESTION 20. *Given a typical analytic function find a limit distribution of $N^{\frac{1}{d}} \min\limits_{0 \leq n < N} |A(x + n\alpha)|$.*

As we shall see in Section 7.2 this question is closely related to Question 5.

5. Poisson processes

In order to explain ideas of the proof of Theorems described above we need some preliminary information. In this section we recall some facts about the Poisson processes referring the reader to [**93**, Section 11] or [**65**] for more details. The next section contains preliminaries from homogenous dynamics.

Recall that a random variable N has Poisson distribution with parameter λ if $\mathbb{P}(N = k) = e^{-\lambda} \frac{\lambda_k}{k!}$. Now an easy combinatorics shows the following facts

(I) If $N_1, N_2 \dots N_m$ are independent random variables and N_j have Poisson distribution with parameters λ_j, then $N = \sum_{j=1}^m N_j$ has Poisson distribution with parameter $\sum_{j=1}^m \lambda_j$.

(II) Conversely, take N points distributed according to a Poisson distribution with parameter λ and color each point independently with one of m colors where color j is chosen with probability p_j. Let N_j be the number of points of color j. Then N_j are independent and N_j has Poisson distribution with parameter $\lambda_j = p_j \lambda$.

Now let (Ω, μ) be a measure space. By a Poisson process on this space we mean a random point process on X such that if $\Omega_1, \Omega_2 \dots \Omega_m$ are disjoint sets and N_j is the number of points in Ω_j then N_j are independent Poisson random variables with parameters $\mu(\Omega_j)$ (note that this definition is consistent due to (I)). We will write $\{x_j\} \sim \mathfrak{P}((X, \mu))$ to indicate that $\{x_j\}$ is a Poisson process with parameters (X, μ). If $(X, \mu) = (\mathbb{R}, c\mathrm{Leb})$ we shall say that X is a Poisson process with intensity c. The following properties of the Poisson process are straightforward consequences of (I) and (II) above.

PROPOSITION 1. *(a) If $\{x'_j\} \sim \mathfrak{P}(X, \mu')$ and $\{x''_j\} \sim \mathfrak{P}(X, \mu'')$ are independent then $\{x'_j\} \cup \{x''_j\} \sim \mathfrak{P}(X, \mu' + \mu'')$.*

(b) If $\{x_j\} \sim \mathfrak{P}(X, \mu)$ and $f : X \to Y$ is a measurable map then $\{f(x_j)\} \sim \mathfrak{P}(Y, f^{-1}\mu)$.

(c) Let $X = Y \times Z$, $\mu = \nu \times \lambda$ where λ is a probability measure on Z. Then $\{(y_j, z_j)\} \sim \mathfrak{P}(X, \mu)$ iff $\{y_j\} \sim \mathfrak{P}(Y, \nu)$ and z_j are random variables independent from $\{y_j\}$ and each other and distributed according to λ.

Next recall [**40**, Chapter XVII] that the Cauchy distribution is unique (up to scaling) symmetric distribution such that if Z, Z' and Z'' are independent random variables with that distribution then $Z' + Z''$ has the same distribution as $2Z$. We have the following representation of the Cauchy distribution.

PROPOSITION 2. *(a) If $\{x_j\}$ is a Poisson process with constant intensity then $\sum_j \frac{1}{x_j}$ has Cauchy distribution. (the sum is understood in the sense of principle value).*

(b) If $\{x_j\}$ is a Poisson distribution with constant intensity and ξ_j are random variables with finite expectation independent from $\{x_j\}$ and from each other then

$$\sum_j \frac{\xi_j}{x_j} \tag{8}$$

has Cauchy distribution.

To see part (a) let $\{x'_j\}, \{x''_j\}$ and $\{x_j\}$ are independent Poisson processes with intensity c. Then

$$\sum \frac{1}{x'_j} + \sum \frac{1}{x''_j} = \sum_{y \in \{x'_j\} \cup \{x''_j\}} \frac{1}{y}$$

and by Proposition 1 (a) and (b) both $\{x'_j\} \cup \{x''_j\}$ and $\{\frac{x_j}{2}\}$ are Poisson processes with intensity $2c$.

To see part (b) note that by Proposition 1(b) and (c), $\{\frac{x_j}{\xi_j}\}$ is a Poisson process with constant intensity.

6. Uniform distribution on the space of lattices

In order to describe ideas of the proofs from Sections 2, 3, and 4 we will first go over some preliminaries. By a **random d-dimensional lattice** (centered at 0) we mean a lattice $L = Q\mathbb{Z}^d$ where Q is distributed according to Haar measure on $G = SL_d(\mathbb{R})/SL_d(\mathbb{Z})$.

By a **random d-dimensional affine lattice** we mean an affine lattice $L = Q\mathbb{Z}^d + b$ where (Q, b) is distributed according to a Haar measure on $\bar{G} = (SL_d(\mathbb{R}) \ltimes \mathbb{R}^d)/(SL_d(\mathbb{Z}) \ltimes \mathbb{Z}^d)$. Here $\bar{G}$ is equipped with the multiplication rule $(A, a)(B, b) = (AB, a + Ab)$.

We denote by g_t the diagonal action on G given by

$$g^t = \begin{pmatrix} e^{t/d} & \dots & 0 & 0 \\ & & & \\ 0 & \dots & e^{t/d} & 0 \\ 0 & \dots & 0 & e^{-t} \end{pmatrix}$$

and for $\alpha \in \mathbb{R}^{d-1}$ we denote by Λ_α the horocyclic action

$$\Lambda_\alpha = \begin{pmatrix} 1 & \dots & 0 & \alpha_1 \\ & & & \\ 0 & \dots & 1 & \alpha_{d-1} \\ 0 & \dots & 0 & 1 \end{pmatrix}. \tag{9}$$

The action of g^t on the space of affine lattices G is partially hyperbolic and unstable manifolds are orbits of Λ_α where $\alpha \in \mathbb{R}^{d-1}$.

Similarly the action by $g_t := (g_t, 0)$, defined on the space of affine lattices is partially hyperbolic and unstable manifolds are orbits of $(\Lambda_\alpha, \bar{x})$ where $(\alpha, x) \in (\mathbb{R}^{d-1})^2$ and

$$\bar{x} = \begin{pmatrix} x_1 \\ \\ x_{d-1} \\ 0 \end{pmatrix}.$$

For convenience, here and below we will use the notation $\bar{x} = (x_1, \dots, x_{d-1}, 0)$ for the column vector $x \in \mathbb{R}^{d-1}$.

We can also equip $SL_d(\mathbb{R}) \ltimes (\mathbb{R}^d)^r$ with the multiplication rule

$$(A, a_1, \dots, a_r)(B, b_1, \dots, b_r) = (AB, a_1 + Ab_1, \dots, a_r + Ab_r), \tag{10}$$

and consider the space of periodic configurations in d-dimensional space $\hat{G} = SL_d(\mathbb{R}) \ltimes (\mathbb{R}^d)^r / SL_d(\mathbb{Z}) \ltimes (\mathbb{Z}^d)^r$.

The action of $g_t := (g^t, 0, \dots, 0)$ on $\hat{G}$ is partially hyperbolic and unstable manifolds are orbits of $(\Lambda_\alpha, \bar{x}_1, \dots, \bar{x}_r)$ where $\alpha, \bar{x}_j \in \mathbb{R}^{d-1}$.

We will denote these unstable manifolds by $n_+(\alpha)$ or $n_+(\alpha, \bar{x})$ or $n_+(\alpha, \bar{x}_1, \dots, \bar{x}_r)$. Note also that $n_+(\alpha, \bar{x})$ or $n_+(\alpha, \bar{x}_1, \dots, \bar{x}_r)$ for fixed $\bar{x}, \bar{x}_1, \dots, \bar{x}_r$ form positive codimension manifolds inside the full unstable leaves of the action of g_t.

We will often use the uniform distribution of the images of unstable manifolds for partially hyperbolic flows (see e.g. [**33**]) to assert that $g_t(\Lambda_\alpha)$ or $g_t(\Lambda_\alpha, \bar{x})$ or $g_t(\Lambda_\alpha, \bar{x}_1, \dots, \bar{x}_r)$ becomes uniformly distributed in the corresponding lattice spaces according to their Haar measures as $\alpha, \bar{x}, \bar{x}_1, \dots, \bar{x}_r$ are independent and distributed according to any absolutely continuous measure on $\mathbb{R}^{d-1}$. In fact, if the original measure has smooth density, then one has exponential estimate for the rate of equidistribution (cf. [**66**]). The explicit decay estimates play an important role in proving limit theorems by martingale methods. For example, such estimates are helpful in proving Theorem 9 in Section 3 and Theorems 26, 27 in Section 8.

Below we shall also encounter a more delicate situation when all or some of $\bar{x}, \bar{x}_j$ are fixed so we have to deal with positive codimension manifolds inside the full unstable horocycles. In this case one has to use Ratner classification theory for unipotent actions. Examples of unipotent actions are $n_+(\cdot)$ or $n_+(\cdot, \bar{x})$ or $n_+(\alpha, \cdot)$. The computations of the limiting distribution of the translates of unipotent orbits proceeds in two steps (cf. [**93**]). For several results described in the previous sections we need the limit distribution of $g_t \Lambda_\alpha w$ inside $X = \mathbf{G}/\mathbf{\Gamma}$ where X can be any of the sets $G, \bar{G}, \hat{G}$ described above and $w \in \mathbf{G}$. In fact, the identity $g_t \Lambda_\alpha w = w(w^{-1} g_t \Lambda_\alpha w)$ allows us to assume that $w = id$ at the cost of replacing the action of $SL_d(\mathbb{R})$ by right multiplication by a twisted action $\phi_w(M)u = w^{-1} M w u$. So we are interested in $\phi_w(g_t \Lambda_\alpha)\mathrm{Id}$ for some fixed $w \in X$. The first step in the analysis is to use Ratner Orbit Closure Theorem [**112**] to find a closed connected subgroup, that depends on w, $\mathbf{H} \subset \mathbf{G}$ such that $\overline{\phi_w(SL_d(\mathbb{R}))\Gamma} = \mathbf{H}\Gamma$ and $\mathbf{H} \cap \Gamma$ is a lattice in $\mathbf{H}$. The second step is to use Ratner Measure Classification Theorem [**111**] to conclude that the sets in question are uniformly distributed in $\mathbf{H}\Gamma/\Gamma$. Namely, we have the following result (see [**118**, Theorem 1.4] or [**93**, Theorem 18.1]).

Theorem 15. *(a) For any bounded piecewise continuous functions $f : X \to \mathbb{R}$ and $h : \mathbb{R}^{d-1} \to \mathbb{R}$ the following holds*

$$\lim_{t\to\infty}\int_{\mathbb{R}^{d-1}} f\left(\varphi_w(g_t\Lambda_\alpha)\right) h(\alpha)d\alpha = \int_X f d\mu_{\mathbf{H}} \int_{\mathbb{R}^{d-1}} h(\alpha)d\alpha \tag{11}$$

where $\mu_{\mathbf{H}}$ denotes the Haar measure on $\mathbf{H}\Gamma/\Gamma$.

(b) In particular, if $\phi(SL_d(\mathbb{R}))\Gamma$ is dense in X then

$$\lim_{t\to\infty}\int_{\mathbb{R}^{d-1}} f\left(\varphi_w(g_t\Lambda_\alpha)\right) h(\alpha)d\alpha = \int_X f d\mu_{\mathbf{G}} \int_{\mathbb{R}^{d-1}} h(\alpha)d\alpha$$

where $\mu_{\mathbf{G}}$ denotes the Haar measure on X.

To apply this Theorem one needs to compute $\mathbf{H} = \mathbf{H}(w)$. Here we provide an example of such computation based on [**93**, Sections 17-19], [**89**, Theorem 5.7], [**95**, Sections 2 and 4] and [**32**, Section 3].

Proposition 3. *Suppose that $d = 2$ and $w = \Lambda(I, \bar{x}_1 \dots \bar{x}_r)$.*

(a) If the vector $(x_1, \dots, x_r)$ is irrational then $\mathbf{H}(w) = SL_2(\mathbb{R}) \ltimes (\mathbb{R}^2)^r$.

(b) If the vector $(x_1, \dots, x_k)$ is irrational and for $j > k$

$$x_j = q_j + \sum_{i=1}^{k} q_{ij} x_j$$

where q_j and q_{ij} are rational numbers then $\mathbf{H}(w)$ is isomorphic to $SL_2(\mathbb{R}) \ltimes (\mathbb{R}^2)^k$.

Proof. (a) Denote

$$(M, 0) = \left(M, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \end{pmatrix}\right), \quad U_x = (I, \bar{x}) = \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \bar{x}_1, \dots, \bar{x}_r\right).$$

We need to show that $U_x^{-1}(M, 0)U_x(\gamma, n)$ is dense in $SL_d(\mathbb{R}) \ltimes (\mathbb{R}^2)^r$ as M, γ, $n = (n_1, \dots, n_r)$ vary in $\mathrm{SL}(2, \mathbb{R})$, $\mathrm{SL}(2, \mathbb{Z})$, and $(\mathbb{Z}^2)^r$ respectively. It is of course sufficient to prove the density of

$$(M, 0)U_x(\gamma, n) = (M\gamma, M\bar{x}_1 + Mn_1, \dots, M\bar{x}_r + Mn_r)$$

which in turn follows from the density in $(\mathbb{R}^2)^r$ of

$$\left(\gamma^{-1}(\bar{x}_1 + n_1), \dots, \gamma^{-1}(\bar{x}_r + n_r)\right).$$

To prove this last claim fix $\epsilon > 0$ and any $v \in (\mathbb{R}^2)^r$. Let $z = (x_1, \dots, x_r)$. Since $\{1, x_1 \dots x_r\}$ are linearly independent over $\mathbb{Q}$, the T_z orbit of 0 is dense in $\mathbb{T}^r$ and hence there exists $a \in \mathbb{N}$ and a vector $m_1 \in \mathbb{Z}^r$ such that

$$|ax_i - v_{i,1} - m_{i,1}| < \epsilon.$$

Since the T_{az} orbit of z is dense in $\mathbb{T}^r$ there exists $c \in \mathbb{N}$ such that

$$c \equiv 1 \bmod a \text{ and } |cx_i - v_{i,2} - m_{i,2}| < \epsilon.$$

Since $a \wedge c = 1$ we can find $b, d \in \mathbb{Z}$ such that $ad - bc = 1$. Let $\gamma^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $n_i = -\gamma m_i$ so that $\left|(\gamma^{-1}(\bar{x}_i + n_i))_j - v_{i,j}\right| < \varepsilon$ for every $i = 1, \dots, r$ and $j = 1, 2$. This finishes the proof of density completing the proof of part (a).

(b) Suppose first that q_j and q_{ij} are integers. In this case a direct inspection shows that $\phi(SL_2(\mathbb{R}))$ is contained in the orbit of $H = SL_2(\mathbb{R}) \times V$ where

$$V = \{(z_1, z_2, \dots, z_r) : z_j = q_j + \sum_i q_{ij} z_i\}$$

and that the orbit of H is closed. Hence $\mathbf{H}(w) \subset H$. To prove the opposite inclusion it suffices to show that $\dim(H) = \dim(\mathbf{H})$. To this end we note that since the action of $SL_2(\mathbb{R})$ is a skew product, $\mathbf{H}$ projects to $SL_2(\mathbb{R})$. On the other hand the argument of part (a) shows that the closure of $\phi(SL_2(\mathbb{R}))$ contains the elements of the form (Id, v) with $v \in V$. This proves the result in case q_i and q_{ij} are integers.

In the general case where $q_j = \frac{p_j}{Q}$ and $q_{ij} = \frac{p_{ij}}{Q}$ where Q, p_j and p_{ij} are integers, the foregoing argument shows that

$$\overline{\phi_w(SL_2(\mathbb{R}))(SL_2(\mathbb{Z}) \ltimes (\mathbb{Z}/Q)^{2r})} = H(SL_2(\mathbb{Z}) \ltimes (\mathbb{Z}/Q)^{2r}).$$

Accordingly, the orbit of Id is contained in a finite union of H-orbits and intersects one of these orbits by the set of measure at least $(1/Q)^r$. Again the dimensional considerations imply that $\mathbf{H}(w) = H$. $\square$

7. Ideas of the Proofs

We are now ready to explain some of the ideas behind the proofs of the Theorems of Sections 2, 3 and 4 following [**28, 29**]. Applications to similar techniques to the related problems could be found for example in [**12, 32, 33, 66, 67, 91, 92**]. We shall see later that the same approach can be used to prove several other limit theorems.

7.1. The Poisson regime. Theorem 13 is a consequence of the following more general result.

THEOREM 16. *([**88, 94**]) (a) If (α, x) is uniformly distributed on $\mathbb{T}^d \times \mathbb{T}^d$ then $N^{1/d}\{x + n\alpha\}$ converges in distribution to*

$$\{X \in \mathbb{R}^d \text{ such that for some } Y \in [0,1] \text{ the point } (X,Y) \in L\}$$

where $L \in \mathbb{R}^{d+1}$ is a random affine lattice.

(b) The same result holds if x is a fixed irrational vector and α is uniformly distributed on $\mathbb{T}^d$.

(c) If α is uniformly distributed on $\mathbb{T}^d$ then $N^{1/d}\{n\alpha\}$ converges in distribution to

$$\{X \in \mathbb{R}^d \textit{ such that for some } Y \in [0,1] \textit{ the point } (X,Y) \in L\}$$

where $L \in \mathbb{R}^{d+1}$ is a random lattice centered at 0.

Here the convergence in, say part (a), means the following. Take a collection of sets $\Omega_1, \Omega_2 \dots \Omega_r \subset \mathbb{R}^d$ whose boundary has zero measure and let $\mathcal{N}_j(\alpha, x, N) = \mathrm{Card}(0 \le n \le N-1 : N^{1/d}\{x + n\alpha\} \in \Omega_j$. Then for each $l_1 \dots l_m$

$$\lim_{N\to\infty} \mathbb{P}(\mathcal{N}_1(\alpha, x, N) = l_1, \dots, \mathcal{N}_r(\alpha, x, N) = l_r) = \tag{12}$$

$$\mu_G(L : \mathrm{Card}(L \cap (\Omega_1 \times [0,1])) = l_1, \dots, \mathrm{Card}(L \cap (\Omega_r \times [0,1])) = l_r)$$

where μ_G is the Haar measure on $G = SL_d(\mathbb{R})$, or $G = SL_d(\mathbb{R}) \ltimes \mathbb{R}^d$ respectively.

The sets appearing in Theorem 16 are called **cut-and-project sets.** We refer the reader to Section 11.2 of the present paper as well as to [**93**, Section 16] for more discussion of these objects.

Remark. Random (quasi)-lattices provide important examples of random point processes in the Euclidean space having a large symmetry group. This high symmetry explains why they appear as limit processes in several limit theorems (see discussion in [**93**, Section 20]). Another point process with large symmetry group is a Poisson process discussed in Section 5. Poisson processes will appear in Theorem 19 below.

A variation on Theorem 16 is the following.

Let $G = SL_2(\mathbb{R}) \ltimes (\mathbb{R}^2)^r$ equipped with the multiplication rule defined in Section 6 and consider $X = G/\Gamma$, $\Gamma = SL_2(\mathbb{Z}) \ltimes (\mathbb{Z}^2)^r$.

Theorem 17. *Assume $(\alpha, z_1 \dots z_r) \in [0,1]^{r+1}$. For any collection of sets $\Omega_{i,j} \subset \mathbb{R}^d$, $i = 1, \dots, r$ and $j = 1, \dots, M$ whose boundary has zero measure let*

$$\mathcal{N}_{i,j}(\alpha, N) = \mathrm{Card}\left(0 \le n \le N-1 : N\{z_i + n\alpha\} \in \Omega_{i,j}\right).$$

(a) If $(\alpha, z_1 \dots z_r)$ are uniformly distributed on $[0,1]^{r+1}$ or if α is uniformly distributed on $[0,1]$ and $(z_1, \dots, z_r)$ is a fixed irrational vector then for each $l_{i,j}$

$$\lim_{N\to\infty} \mathbb{P}\left(\mathcal{N}_{i,j}(\alpha, N) = l_{i,j}, \forall i,j\right) \tag{13}$$
$$= \mu_G\left((L, a_1, \dots, a_r) \in G : \mathrm{Card}(L + a_i \cap (\Omega_{i,j} \times [0,1])) = l_{i,j}, \forall i,j\right)$$

where we use the notation $\mathbb{P}$ for Haar measure on $\mathbb{T}$ (in the case of fixed vector $(z_1, \dots, z_r)$) as well as on $\mathbb{T}^{r+1}$ (in the case of random vector $(z_1, \dots, z_r)$), and μ_G is the Haar measure on.

(b) For arbitrary $(z_1, \dots, z_r)$ *there is a subgroup* $\mathbf{H} \subset G$ *such that for each* $l_{i,j}$

$$\begin{aligned}(14) \quad & \lim_{N\to\infty} \mathbb{P}\left(\mathcal{N}_{i,j}(\alpha, N) = l_{i,j}, \forall i, j\right) \\ &= \mu_{\mathbf{H}}\left((L, a_1, \dots, a_r) \in G : \mathrm{Card}(L + a_i \cap (\Omega_{i,j} \times [0,1])) = l_{i,j}, \forall i, j\right)\end{aligned}$$

where $\mu_{\mathbf{H}}$ *denotes the Haar measure on the orbit of* $\mathbf{H}$.

Proof of Theorem 16. (a) We provide a sketch of proof referring the reader to [**93**, Section 13] for more details.

Fix a collection of sets $\Omega_1, \Omega_2 \dots \Omega_r \subset \mathbb{R}^d$ and $l_1, \dots, l_r \in \mathbb{N}$. We want to prove (12).

Consider the following functions on the space of affine lattices $\bar{G} = (SL_{d+1}(\mathbb{R}) \ltimes \mathbb{R}^{d+1})/(SL_{d+1}(\mathbb{Z}) \ltimes \mathbb{Z}^{d+1})$

$$f_j(L) = \sum_{e\in L} \chi_{\Omega_j\times[0,1]}(e)$$

and let $\mathcal{A} = \{L : f_j(L) = l_j\}$. By definition, the right hand side of (12) is $\mu_{\bar{G}}(L : L \in \mathcal{A})$.

On the other hand, the Dani correspondence principle states that

$$(15) \quad \mathcal{N}_1(\alpha, x, N) = l_1, \dots, \mathcal{N}_r(\alpha, x, N) = l_r \text{ iff } g^{\ln N}(\Lambda_\alpha \mathbb{Z}^{d+1} + \bar{x}) \in \mathcal{A}$$

where g_t is the diagonal action $(e^{t/d}, \dots, e^{t/d}, e^{-t})$ and $\bar{x} = (x_1, \dots, x_d, 0)$ are defined as in Section 6. To see this, fix j and suppose that $\{x + n\alpha\} \in N^{-1/d}\Omega_j$ for some $n \in [0, N]$. Then

$$\{x + n\alpha\} = (x_1 + n\alpha_1 + m_1(n), \dots, x_d + n\alpha_d + m_d(n))$$

with $m_i(n)$ uniquely defined so that $x_i + n\alpha_i + m_i \in (-1/2, 1/2]$, and the vector $v = (m_1(n), \dots, m_d(n), n)$ is such that

$$\chi_{\Omega_j\times[0,1]}\left(g^{\ln N}(\Lambda_\alpha \mathbb{Z}^{d+1} + \bar{x})v\right) = 1.$$

The converse is similarly true, namely that any vector that counts in the right hand side of (15) corresponds uniquely to an n that counts in the left hand side visits.

Now (15), and thus (12), follow if we prove that

$$(16) \quad \lim_{N\to\infty} \mathbb{P}\left((\alpha, x) \in \mathbb{T}^d \times \mathbb{T}^d : g^{\ln N}(\Lambda_\alpha \mathbb{Z}^{d+1} + \bar{x}) \in \mathcal{A}\right) = \mu_G(L \in \bar{G} : L \in \mathcal{A})$$

Finally, (16) holds due to the uniform distribution of the images of unstable manifolds $n_+(\alpha, x)$ for partially hyperbolic flows.

(b) Following the same arguments as above, we see that in order to prove Theorem 16(b) we need to show that $(g^{\ln N}, 0).(\Lambda_\alpha, \bar{x})$ becomes equidisitributed with respect to Haar measure on

$$G = (SL_{d+1}(\mathbb{R}) \ltimes \mathbb{R}^{d+1})/(SL_{d+1}(\mathbb{Z}) \ltimes \mathbb{Z}^{d+1})$$

if α is random and x is a fixed irrational vector. This can be derived from Theorem 15(b) using a generalization of Proposition 3(a).

The argument for part (c) is the same as in part (a) but we use the space of lattices rather than the space of affine lattices. □

PROOF OF THEOREM 17. As in the proof of Theorem 16 we use Dani's correspondence principle to identify the left hand side in (13) with

$$\mathbb{P}\left(\mathrm{Card}((g_{\ln N}\Lambda_\alpha + g_{\ln N}\bar{z}_i) \cap (\Omega_{i,j} \times [0,1])) = l_{i,j}, \forall i, j\right)$$

where $\bar{z}_i = \begin{pmatrix} z_i \\ 0 \end{pmatrix}$.

Now (13) follows if we have that $(g_{\ln N}, 0).(\Lambda_\alpha, \bar{z}_1, \dots, \bar{z}_r))$ is distributed according to the Haar measure in $X = G/\Gamma$ with $G = SL_2(\mathbb{R}) \ltimes (\mathbb{R}^2)^r$ and $\Gamma = SL_2(\mathbb{Z}) \ltimes (\mathbb{Z}^2)^r$. But this last statement follows from Theorem 15(b) and Proposition 3(a).

Likewise (14) follows from Theorem 15(a). □

7.2. Application of the Poisson regime theorems to the ergodic sums of smooth functions with singularities. The microscopic or Poisson regime theorems are useful to treat the ergodic sums of smooth functions with singularities since the main contribution to these sums come from the visits to small neighborhoods of the singularity.

PROOF OF THEOREM 3. Due to Corollary 1 we may assume that $\tilde{A} = 0$.

Let R be a large number and denote by S'_N the sum of terms with $|x + n\alpha - x_0| > R/N$ and by S''_N the sum of terms with $|x + n\alpha - x_0| < R/N$. Then

$$\mathbb{E}(|S'_N/N^a|) \le \frac{C}{N^a}\mathbb{E}(|\xi|^{-a}\chi_{|\xi|>R/N}) = O(R^{1-a}).$$

On the other hand, by Theorem 16(a) if (α, x) is random $\dfrac{S''_N}{N^2}$ converges as $N \to \infty$ to

$$\sum_{(X,Y)\in L: Y\in[0,1], |X|<R} \frac{c_-\chi_{X<0} + c_+\chi_{X>0}}{|X|^a},$$

where L is a random affine lattice in $\mathbb{R}^2$. Letting $R \to \infty$ we get

$$\frac{S_N}{N^2} \Rightarrow \sum_{(X,Y)\in L: Y\in[0,1]} \frac{c_- \chi_{X<0} + c_+ \chi_{X>0}}{|X|^a}.$$

The case of fixed irrational x is dealt with similarly using Theorem 16(b). □

Sketch of proof of Theorem 5. Consider first the case when the highest pole has order $m > 1$. Then the argument given in the proof of Theorem 3 shows that for large R, A_N/N^m can be well approximated by

$$\frac{A_N}{N^m} \sim \sum_{j=1}^{r} \sum_{|x+n\alpha-x_j|<R/N} \frac{c_j}{(N(x+n\alpha-x_j))^m} \tag{17}$$

where $x_1 \dots x_r$ are all poles of order m and $c_1 \dots c_r$ are the corresponding Laurent coefficients.

We use Theorem 17 to analyze this sum. Namely, let $G = SL_2(\mathbb{R}) \ltimes (\mathbb{R}^2)^r$ with the multiplication rule defined in Section 6 and consider $X = G/\Gamma$, $\Gamma = SL_2(\mathbb{Z}) \ltimes (\mathbb{Z}^2)^r$. Consider the functions $\Phi : G/\Gamma \to \mathbb{R}$ given by

$$\Phi_R(A, a_1, \dots a_r) = \sum_{j=1}^{r} \sum_{e\in A\mathbb{Z}^2+a_j} \frac{c_j}{x(e)^m} \chi_{[-R,R]\times[0,1]}(e) \tag{18}$$

which as $R \to \infty$ will be distributed as

$$\Phi(A, a_1 \dots a_r) = \sum_{j=1}^{r} \sum_{e\in A\mathbb{Z}^2+a_j} \frac{c_j}{x(e)^m} \chi_{[0,1]}(y(e)). \tag{19}$$

Now Theorem 5 follows from Theorem 17. Namely part (a) of Theorem 5 follows from Theorem 17(a). To get part (c) we let $z_j = x - x_j$, we observe that for almost every x the vector $(z_1, \dots, z_r)$ is irrational. Hence Theorem 17 applies and gives us that (17) when $\alpha \in \mathbb{T}$ is random has the same distribution as (18) as $(A, a_1 \dots a_r)$ is random in G. Finally Theorem 5(b) follows from Theorem 17 (b).

The proof in case when all poles are simple is the same except that the proof that A_N/N is well approximated by (17) is more involved since one cannot use L^1 bounds. We refer to [**120**] for more details. □

7.3. Limit theorems for discrepancy. The proofs of Theorems 10–12 use a similar strategy as Theorems 3 and 5 of first localizing the important terms and then reducing their contribution to lattice counting problems. However the analysis in that case is more complicated, in particular, because the argument is carried over in the set of frequencies of the Fourier series of the discrepancy rather than in the phase space.

Let us describe the main steps in the proof of Theorem 10. Consider first the case where Ω is centrally symmetric. We start with Fourier series of the discrepancy

$$D_N(\Omega,\alpha,x) = r^{(d-1)/2} \sum_{k\in\mathbb{Z}^d-0} c_k(r) \frac{\cos(2\pi(k,x)+\pi(N-1)(k,\alpha))\sin(\pi N(k,\alpha))}{\sin(\pi(k,\alpha))}$$

where $r^{(d-1)/2}c_k(r)$ are Fourier coefficients of $\chi_{r\Omega}$ which have the following asymptotics for large k (see [**50**])

$$c_k(r) \approx \frac{1}{\pi|k|^{(d+1)/2}} K^{-1/2}(k/|k|) \sin\left(2\pi\left(rP(k)-\frac{d-1}{8}\right)\right).$$

Here $K(\xi)$ is the Gaussian curvature of $\partial\Omega$ at the point where the normal to $\partial\Omega$ is equal to ξ and $P(t)=\sup_{x\in\Omega}(x,t)$.

The proof consists of three steps. First, using elementary manipulations with Fourier series one shows that the main contribution to the discrepancy comes from k satisfying

$$\varepsilon N^{1/d} < |k| < \varepsilon^{-1} N^{1/d} \tag{20}$$

$$k^{(d+1)/2}|\{(k,\alpha)\}| < \frac{1}{\varepsilon N^{(d-1)/2d}}. \tag{21}$$

To understand the above conditions note that (20) and (21) imply that $\{(k,\alpha)\}$ is of order $1/N$ so the sum $\sum_{n=0}^{N-1} e^{i2\pi(k,(x+n\alpha))}$ is of order N, that is, it is as large as possible. Next the number of terms with $|k| \ll N^{1/d}$ is too small (much smaller than N) so for typical α, we have that $|\{k,\alpha)\}| \gg 1/N$ for such k ensuring the cancelations in ergodic sums. For the higher modes $k \gg N^{1/d}$, using L^2 norms and the decay of c_k one sees that their contribution is negligible

The second step consists in showing using the same argument as in the proof of Theorem 16 that if (α,x) is uniformly distributed on $\mathbb{T}^d\times\mathbb{T}^d$ then the distribution of

$$\left(\frac{k}{N^{1/d}}, (k,\alpha)|k|^{(d+1)/2}N^{(d-1)/2d}\right)$$

converge as $N\to\infty$ to the distribution of

$$(X(e), Z(e)|X(e)|^{(d+1)/2})_{e\in L}$$

where L is a random lattice in $\mathbb{R}^{d+1}$ centered at 0.

Finally the last step in the proof of Theorem 10 is to show that if we take prime k satisfying (20) and (21) then the phases (k, x), $N(k, \alpha)$ and $rP(k)$ are asymptotically independent of each other and of the numerators. For (k, x) and $rP(k)$ the independence comes from the fact that (20) and (21) do not involve x or r, while $N(k, \alpha)$ has wide oscillation due to the large prefactor N.

The argument for non-symmetric bodies is similar except that the asymptotics of their Fourier coefficients is slightly more complicated.

The foregoing discussion explains the form of the limit distribution which we now present. Let $\mathcal{M}_{2,d}$ be the space of quadruples (L, θ, b, b') where $L \in X$, the space of lattices in $\mathbb{R}^{d+1}$, and (θ, b, b') are elements of $\mathbb{T}^X$ satisfying the conditions

$$\theta_{e_1+e_2} = \theta_{e_1} + \theta_{e_2}, \quad b_{me} = mb_e \text{ and } b'_{me} = mb'_e.$$

Let $\mathcal{M}_d$ be the subset of $M_{2,d}$ defined by the condition $b = b'$. Consider the following function on $\mathcal{M}_{2,d}$

$$\mathcal{L}_\Omega(L, \theta, b, b') = \frac{1}{\pi^2} \sum_{e \in L} \kappa(e, \theta, b, b') \frac{\sin(\pi Z(e))}{|X(e)|^{\frac{d+1}{2}} Z(e)} \tag{22}$$

with

$$\begin{aligned} \kappa(e, \theta, b, b') = {} & K^{-\frac{1}{2}}(X(e)/|X(e)|) \sin(2\pi(b_e + \theta_e - (d-1)/8)) \\ & + K^{-\frac{1}{2}}(-X(e)/|X(e)|) \sin(2\pi(b'_e - \theta_e - (d-1)/8)). \end{aligned} \tag{23}$$

It is shown in [**28**] that this sum converges almost everywhere on $\mathcal{M}_{2,d}$ and $\mathcal{M}_d$. Now the limit distribution in Theorem 10 can be described as follows

THEOREM 18. *(a) If Ω is symmetric then $\dfrac{D_N(\Omega, r, \alpha, x)}{N^{(d-1)/2d} r^{(d-1)/2}}$ converges to $\mathcal{L}_\Omega(L, \theta, b, b')$ where $\mathcal{L}_\Omega(L, \theta, b, b')$ is uniformly distributed on $\mathcal{M}_d$.*

(b) If Ω is non-symmetric then $\dfrac{D_N(\Omega, r, \alpha, x)}{N^{(d-1)/2d} r^{(d-1)/2}}$ converges to $\mathcal{L}_\Omega(L, \theta, b, b')$ where $\mathcal{L}_\Omega(L, \theta, b, b')$ is uniformly distributed on $\mathcal{M}_{2,d}$.

QUESTION 21. *Study the properties of the limiting distribution in Theorem 18, in particular its tail behavior.*

The next question is inspired by Theorem 49 from Section 11.

QUESTION 22. *Consider the case where Ω is the standard ball. Thus in (23) $K \equiv 1$. Study the limit distribution of $\mathcal{L}$ when the dimension of the torus $d \to \infty$.*

Next we describe the idea of the proof of Theorem 12. Note that (8) looks similar to (22). The main ingredient in the proof of Theorem 12 involves a result on the distribution of small divisors of multiplicative form $\Pi|k_i|\|(k,\alpha)\|$. Namely, a harmonic analysis of the discrepancy's Fourier series related to boxes allows to bound the frequencies that have essential contributions to the discrepancy and show that they must be resonant with α. The main step is then to establish a Poisson limit theorem for the distribution of small denominators and the corresponding numerators. With the notation introduced before the statement of Theorem 12 let $\bar{k}_i = a_{i,1}k_1 + \cdots + a_{i,d}k_d$. Then we have

THEOREM 19. *([**29**]) Let $\xi \in X$ be distributed according to the normalized Lebesgue measure λ. Then as $N \to \infty$ the point process*

$$\left\{\left((\ln N)^d \Pi_i \bar{k}_i \|(k,\alpha)\|, N(k,\alpha)\mathrm{mod}(2), \{\bar{k}_1 u_1\}, \dots \{\bar{k}_d u_d\},\right)\right\}_{k\in Z(\xi,N)}$$

where

$$\begin{aligned} Z(\xi,N) = \big\{k \in \mathbb{Z}^d : |\bar{k}_i| \geq 1, |\Pi_i \bar{k}_i| < N, \bar{k}_1 > 0, \\ |\Pi_i \bar{k}_i| \|(k,\alpha)\| \leq \frac{1}{\varepsilon(\ln N)^d}, \\ \exists m \in \mathbb{Z} \text{ such that } k_1 \wedge \ldots \wedge k_d \wedge m = 1 \textit{ and } \|(k,\alpha)\| = (k,\alpha)+m\big\} \end{aligned}$$

converges to a Poisson process on $\mathbb{R} \times \mathbb{R}/(2\mathbb{Z}) \times (\mathbb{R}/\mathbb{Z})^d$ with intensity $2^{d-1}\mathbf{c}_1/d!$.

Comparing this result with the proof of Theorem 10 discussed above we see that Theorem 19 comprises analogies of both step 2 and 3 in the former proof. Namely, it shows both that the small denominators contributing most to the discrepancy have asymptotically Poisson distribution and that the numerators are asymptotically independent of the denominators (cf. Proposition 1(c)).

We note that Theorem 19 is interesting in its own right since it describes the number of solutions to Diophantine inequalities

$$\Pi_i|\bar{k}_i|\|(k,\alpha)\| < \frac{c}{\ln^d N}, \quad |\bar{k}_i| > 1, \quad \prod_i |\bar{k}_i| < N.$$

QUESTION 23. *What happens if in Theorem 19 $\ln^d N$ is replaced by $\ln^a N$ with $a \in (0,d)$?*

QUESTION 24. *Is Theorem 19 still valid if the distribution of ξ is concentrated on a submanifold of X? For example, one can take $\alpha = (s, s^2)$.*

A special case of Question 24 is when the matrix $(a_{i,j})$ is fixed equal to Identity. This case is directly related to Question 16(a).

The proof of Theorem 19 proceeds by martingale approach (see [**26, 27**]) which requires good mixing properties in the future conditioned to the past. In the present setting, to apply this method it suffices to prove that most orbits of certain unipotent subgroups are equidisitributed at a polynomial rate. Under the conditions of Theorem 19 one can assume (after an easy reduction) that the initial point has smooth density with respect to Haar measure. Then the required equidistribution follows easily form polynomial mixing of the unipotent flows. In the setting of Question 24 (as well as Question 53 in Section 11) the initial point is chosen from a positive codimension submanifold so one cannot use the mixing argument. The problem of estimating the rate of equidistribution for unipotent orbits starting from submanifolds interpolates between the problem of taking a random initial condition with smooth density which is solved and the problem of taking fixed initial condition which seems very hard.

8. Shrinking targets

Another classical result in probability theory is the Borel-Cantelli Lemma which says that if A_j are independent sets and $\sum_j \mathbb{P}(A_j) = \infty$ then $\mathbb{P}$-almost every point belongs to infinitely many sets. A yet stronger conclusion is given by the *strong Borel-Cantell Lemma* claiming that the number of A_j which happen up to time N is asymptotic to $\sum_{j=1}^N \mathbb{P}(A_j)$. In the context of ergodic dynamical systems (T, X, μ), the law of large numbers is reflected in the Birkhoff theorem of almost sure converge in average of the ergodic means associated to a measurable observable, for example the characteristic function of a measurable set $A \subset X$. In a similar fashion one can study the so called dynamical Borel-Cantelli properties of the system (X, T, μ) by considering instead of a fixed stet A a sequence of "target" sets $A_j \in X$ such that $\sum \mu(A_j) = \infty$. We then say that the dynamical Borel-Cantelli property is satisfied by $\{A_j\}$ if for almost every x, $T^j(x)$ belongs to A_j for infinitely many j.

In the context of a dynamical system (T, X, μ) on a metric space X it is natural to assume that the sets in question have nice geometric structure, since it is always possible for any dynamical system (with a non atomic invariant measure) to construct sets with divergent sum of measures that are missed after a certain iterate by the orbits of almost every point [**21**, Proposition 1.6]. The simplest assumption is that the sets be balls. The dynamical Borel-Cantelli property for balls is a common feature for deterministic systems displaying hyperbolicity features (see [**51, 108, 27**] and references therein).

Due to strong correlations among iterates of a toral translation the dynamical Borel-Cantelli properties are more delicate in the quasi-periodic context.

8.1. Dynamical Borel-Cantelli lemmas for translations. For toral translations one needs also to assume that the sets are nested since otherwise one can take $A_j \subset (A_0 + j\alpha)$ for some fixed set A_0 ensuring that the points from the compliment of A_0 do not visit any A_j at time j. This motivates the following definition (see [**51, 21, 38**]).

Given $T : (X, \mu) \to (X, \mu)$ let $V_N(x, y) = \sum_{n=1}^{N} \chi_{B(y,r_n)}(T^n x)$. We say that T has the **shrinking target property** (STP) if for any y, $\{r_n\}$ such that $\sum_n \mu(B(y, r_n)) = \infty$, it holds that $V_N(x, y) \to \infty$ for almost all x, i.e. the targets sequence $(B(y, r_n))$ satisfies the Borel-Cantelli property for T. We say that T has the **monotone shrinking target property** (MSTP) if for any $y, \{r_n\}$ such that $\sum_n \mu(B(r_n)) = \infty$ and r_n is non-increasing $V_N(x, y) \to \infty$ for almost all x.

In the case of translations, we can always assume without loss of generality that $y = 0$ (replace x by $x - y$). We then use the notation $V_N(x)$ for $V_N(x, y)$. We also use the notation $B(r)$ for the ball $B(0, r)$. Another interesting choice is to take $y = x$ in which case we study the rate of return rather than the rate of approach to 0. Note that if $V_N(x, x)$ does not depend on x and so the number of close returns depends only on α. We shall write $U_N(\alpha) = \sum_{n=0}^{N-1} \chi_{B(r_n)}(T^n 0)$.

The following is a straightforward consequence of the fact that toral translations are isometries.

THEOREM 20. *([**38**]) Toral translations do not have STP.*

It turns out that the following Diophantine condition is relevant to this problem. Let

$$\mathcal{D}^*(\sigma) = \{\alpha : \forall k \in \mathbb{Z} - 0, \max_{i\in[1,d]} \|k\alpha_i\| \geq C|k|^{-(1+\sigma)/d}\}. \tag{24}$$

THEOREM 21. *([**80**]) A toral translation T_α has the MSTP iff $\alpha \in \mathcal{D}^*(0)$.*

A simple proof of Theorem 21 can be found in [**38**]. Recall that $\mathcal{D}^*(0)$ has zero Lebesgue measure. Hence, the latter result shows that one has to further restrict the targets if one wants that typical translations display the dynamical Borel-Cantelli property relative to these targets.

One possible restriction on the targets is to impose a certain growth rate on the sum of their volumes. This actually allows to further distinguish among distinct Diophantine classes as it is shown in the following result. We say that T has **s-(M)STP** if for any $\{r_n\}$ such that $\sum_n r_n^{ds} = \infty$ (and r_n is non-increasing) $V_N(x) \to \infty$ for almost all x. We then have the following.

THEOREM 22. *([**124**])*
a) If $\alpha \notin \mathcal{D}^*(sd-d)$, then the toral translation T_α does not have the s-MSTP.
b) A circle rotation T_α has the s-MSTP iff $\alpha \in \mathcal{D}^*(s-1)$.

QUESTION 25. *Is this true that the toral translation T_α has the s-MSTP iff $\alpha \in \mathcal{D}^*(sd-d)$?*

Another possible direction is to study specific sequences, asking for example that $r_n = cn^{-\gamma}$, or that nr_n^d be decreasing, in which case the sequence r_n is coined a **Khinchin sequence**. The case $r_n = cn^{-1/d}$ in dimension d is very particular, but important. Indeed a vector $\alpha \in \mathbb{T}^d$ is said to be **badly approximable** if for some $c > 0$, the sequence $\lim_{N\to\infty} U_N(\alpha, \{cn^{-1/d}\}) < \infty$. It is known that the set of badly approximable vectors has zero measure. By contrast, vectors α such that $\lim_{N\to\infty} U_N(\alpha, \{cn^{-(1/d+\varepsilon)}\}) = \infty$ for some $\varepsilon > 0$ are called **very well approximated, or VWA**. The obvious direction of the Borel-Cantelli lemma implies that almost every $\alpha \in \mathbb{T}^d$ is not very well approximated (cf. [**19**, Chap. VII]). The latter facts are particular cases of a more general result, the Khintchine–Groshev theorem on Diophantine approximation which gives a very detailed description of the sequences such that $U_N(x, \{r_n\})$ diverges for almost all α. We refer the reader to [**13**] for a nice discussion of that theorem and its extensions, and to Section 11.1 below.

Khinchin sequences also display BC property much more likely than general sequences. For example, compare Theorem 23(b) below with Theorem 21 which shows that the set of vectors having mSTP has zero measure.

If a shrinking target property holds it is natural to investigate the asymptotics of the number of target hits. This makes the following definition natural. We say that a given sequence of targets $\{A_n\}$ is **sBC or strong Borel-Cantelli** for (T, X, μ) if for almost every x

$$\lim_{N\to\infty} \frac{\sum_{n=1}^N \chi_{A_n}(T^n x)}{\sum_{n=1}^N \mu(A_n)} = 1.$$

THEOREM 23. [**20**] *(a) For every α such that its convergents satisfy $a_n \le Cn^{\frac{7}{6}}$ the sequence $\{B(\frac{c}{n})\}$ is sBC for T_α.*

(b) For almost every $\alpha \in \mathbb{T}$, any Khinchin sequence is sBC for T_α.

(c) For any $\alpha \in \mathcal{D}(1)$, and any decreasing sequence $\{r_n\}$ such that $\sum r_n = \infty$, $\{B(r_n)\}$ is sBC for T_α.

Observe that the condition in (a) has full measure. On the other hand, it is not hard to see that if $a_n(\alpha) \sim n^{2+\varepsilon}$ for every n then the sequence $(B(\frac{1}{n}))$ does not have the sBC for T_α. Indeed, if

$$x \in \left[\frac{k}{q_n} - \frac{1}{2nq_n}, \frac{k}{q_n} + \frac{1}{2nq_n}\right]$$

then since $\|q_n\alpha\| \le \frac{1}{q_{n+1}} \le \frac{2}{n^{2+\varepsilon}q_n}$ and $\ln q_n \le Cn \ln n$

$$\sum_{l=q_n}^{n^{1+\varepsilon/2}q_n} \chi_{B(\frac{1}{n})}(x + l\alpha) \ge n^{1+\varepsilon/2} \gg \sum_{l=1}^{n^{1+\varepsilon/2}q_n} \frac{1}{n}.$$

But it is easy to see that a.e. x belongs to infinitely many intervals of the form $[\frac{k}{q_n} - \frac{1}{2nq_n}, \frac{k}{q_n} + \frac{1}{2nq_n}]$.

In higher dimensions, it was proved in [**117**] that

THEOREM 24. *If $\sum_n r_n^d = \infty$ then for almost every vector $\alpha \in \mathbb{T}^d$, the sequence $(B(r_n))$ is sBC for the translation T_α.*

8.2. On the distribution of hits. Theorems 23 and 24 motivate the study of the error terms

$$\Delta_N(c, \alpha, x) = V_N(\alpha, x) - \sum_{n=1}^{N} \mathrm{Vol}(B_{r_n}) \text{ and } \bar{\Delta}_N(c, \alpha) = U_N(\alpha) - \sum_{n=1}^{N} \mathrm{Vol}(B_{r_n}).$$

One can for example try to give lower and upper asymptotic bounds on the growth of Δ_N as a function of the arithmetic properties of α in the spirit of Kintchine-Beck Theorem 6 and Questions 6–8. Here we will be interested in the distribution of $\Delta_N(c, \alpha, x)$ after adequate normalization when α or x or (α, x) are random.

THEOREM 25. *([**9, 90**]) Let $r_n = cn^{-1/d}$. Suppose that x is uniformly distributed on $\mathbb{T}^d$. For any $c > 0$, if $\alpha \in \mathcal{D}^*(0)$, there is a constant K such that all limit points of $\dfrac{\Delta_N(c, \alpha, x)}{\sqrt{\ln N}}$ are $\mathfrak{N}(\sigma^2)$ with $\sigma^2 \le K$.*

In the case of random (α, x) we have

THEOREM 26. *Let* $r_n = cn^{-1/d}$. *([**30**]) There is* $\Sigma(c,d) > 0$ *such that if* (α, x) *is uniformly distributed on* $\mathbb{T}^d \times \mathbb{T}^d$ *then* $\dfrac{\Delta_N(c,\alpha,x)}{\sqrt{\ln N}}$ *converges to* $\mathfrak{N}(\Sigma(c,d))$.

There is an analogous statement for the return times.

THEOREM 27. *([**107, 114, 30**]) Let* $r_n = cn^{-1/d}$. *There is* $\bar{\Sigma}(c,d) > 0$ *such that if* α *is uniformly distributed on* $\mathbb{T}^d$ *then* $\dfrac{\bar{\Delta}_N(c,\alpha)}{\sqrt{b_N}}$ *converges in distribution to* $\mathfrak{N}(\bar{\Sigma}(c,d))$ *where*

$$b_N = \begin{cases} \ln N \ln\ln N & \text{if } d = 1 \\ \ln N & \text{if } d \geq 2. \end{cases}$$

The case $d = 1$ was obtained in [**107**, Theorem 3.1.1 on page 44] (see also [**114**]), based on the metric theory of the continued fractions. In fact, one can handle more general sequences. Namely, let $\phi(k)$ satisfy the following conditions

(i) $\phi(k) \searrow 0$, but $\sum_k \phi(k) = +\infty$,
(ii) There exists $0 < \delta < 1/2$ such that $\sum_{k=1}^n \frac{\phi(k)}{k^\delta} \leq C\sqrt{\sum_{k=1}^n \phi(k)}$
(iii) $\sum_{k=1}^n \phi^2(k) \leq C\sqrt{\sum_{k=1}^n \phi(k)}$.

THEOREM 28. *([**44**]) If* $r_n = \frac{\phi(\ln n)}{n}$ *and* α *is uniformly distributed on* $\mathbb{T}$ *then* $\dfrac{\bar{\Delta}_N(c,\alpha)}{\sqrt{F(n)\ln F(n)}}$ *converges in distribution to* $\mathfrak{N}(\boldsymbol{\Sigma}(c))$ *where* $F(n) = \sum_{k=1}^n \frac{\phi(\ln k)}{k}$.

The higher dimensional case is obtained *via* ergodic theory of homogeneous flows and martingale methods in [**30**].

QUESTION 26. *Study the limiting distribution of* U_N *and* V_N *in case* $r_n = \frac{c}{n^\gamma}$ *with* $\gamma < \frac{1}{d}$.

QUESTION 27. *Do Theorems 24, 26 and 27 hold when the random vector* α *is taken from a proper submanifold of* $\mathbb{T}^d$, *for example* $\alpha = (s, s^2, \dots, s^d)$.

One motivation for this question comes from *Diophantine approximation on manifolds* (see [**13**] and references wherein), another is multidimensional extension of Kesten Theorem (cf. Question 24).

8.3. Proofs outlines. First we sketch a proof of Theorem 24 in case $r_n = cn^{-1/d}$. Consider the number $N_m(\alpha, x)$ of solutions to

$$x + n\alpha \in B(0, cn^{-1/d}), \quad e^m < n < e^{m+1}.$$

The argument used to prove Theorem 16 shows that

$$N_m(\alpha, x) = f(g_m(\Lambda_\alpha \mathbb{Z}^{d+1} + x))$$

where f is the function on the space of affine lattices given by

$$f(L) = \sum_{v \in L} \chi_{B(0,c)\times[1,e]}(v).$$

Thus

$$\sum_{m=1}^{\ln N} N_m(\alpha, x) \sim \sum_{m=1}^{\ln N} f(g_m(\Lambda_\alpha \mathbb{Z}^{d+1} + x)) \tag{25}$$

and Theorem 24 for $r_n = cn^{-1/d}$ reduces to the study of ergodic sums (25) under the assumption that the initial condition has a density on $n_+(\alpha, x)$. In fact, a standard argument allows to reduce the problem to the case when the initial condition has density on the space of lattices. Namely, it is not difficult to check that the ergodic sums of f do not change much if we move in the stable or neutral direction in the space of lattices. After this reduction, the sBC property follows from the Ergodic Theorem.

The relation (25) also allows to reduce Theorem 26 to a Central Limit Theorem for ergodic sums of g_m which can be proven, for example, by a martingale argument (see [**81**]. We refer the reader to [**26**] for a nice introduction to the martingale approach to limit theorems for dynamical systems.)

The proof of Theorem 27 is similar but one needs to work with lattices centered at 0 rather than affine lattices.

In particular, the non-standard normalization in case $d = 1$ is explained by the fact that f in this case is not in L^2 and the main contribution comes from the region where f is large (in fact, the analysis is similar to [**46**, Section 4]).

9. Skew products. Random walks.

9.1. Basic properties. The properties of ergodic sums along toral translations are crucial to the study of some classes of dynamical systems, such as skew products or special flows. In this section we consider the skew products. Special flows are the subject of Section 10.

Skew products above T_α will be denoted $S_{\alpha,A} : \mathbb{T}^d \times \mathbb{T}^r \to \mathbb{T}^d \times \mathbb{T}^r$ They are given by $S_{\alpha,A}(x, y) = (x + \alpha, y + A(x) \bmod 1)$. **Cylindrical cascades** above T_α will be denoted $W_{\alpha,A} : \mathbb{T}^d \times \mathbb{R}^r \to \mathbb{T}^d \times \mathbb{R}^r$. They are given by $W_{\alpha,A}(x, y) = (x + \alpha, y + A(x))$. Note that

$$W^N_{\alpha,A}(x, y) = (x + N\alpha, y + A_N(x))$$

(the same formula holds for $S_{\alpha,A}$ but the second coordinate has to be taken mod 1). If A takes integer values then $W_{\alpha,A}$ preserves $\mathbb{T}^d \times \mathbb{Z}^r$ and it is natural to restrict the dynamics to this subset. Thus cylindrical cascades define random walks on $\mathbb{R}^r$ or $\mathbb{Z}^r$ driven by the translation T_α.

If α is Diophantine and A is smooth then the so called linear cohomological equation similar to (3)

$$A(x) - \int_{\mathbb{T}^d} A(u)du = -B(x+\alpha) + B(x) \tag{26}$$

has a smooth solution B, thus $S_{\alpha,A}$ and $W_{\alpha,A}$ are respectively smoothly conjugated to the translations $S_{\alpha,\int_{\mathbb{T}^d} A}$ and $W_{\alpha,\int_{\mathbb{T}^d} A}$ via the conjugacy $(x,y) \mapsto (x, y - B(x))$.

Hence the ergodic properties of the skew products and the cascades with smooth A are interesting to study only in the Liouville case. The following is a convenient ergodicity criterion for skew products.

Proposition 4. [**78**] *$S_{\alpha,A}$ is ergodic iff for any $\lambda \in \mathbb{Z}^r - \{0\}$, $\langle \lambda, A\rangle$ is not a measurable multiplicative coboundary above T_α, that is, iff there does not exist $\lambda \in \mathbb{Z}^r - \{0\}$ and a measurable solution $\psi : \mathbb{T}^d \to \mathbb{C}$ to*

$$e^{i2\pi\langle \lambda, A(x)\rangle} = \psi(x+\alpha)/\psi(x). \tag{27}$$

This ergodicity criterion can be simply derived from the observation that the spaces V_λ of functions of the form

$$\phi(x)e^{i2\pi\langle \lambda, y\rangle} \tag{28}$$

are invariant under $S_{\alpha,A}$. It then follows that the existence or nonexistence of an invariant function φ by $S_{\alpha,A}$ is determined by the existence or nonexistence of a solution to (27). We refer the reader to Section 10 for further discussion concerning (27).

When A is not a linear coboundary, i.e. (26) does not have a solution, it is very likely and often easy to prove that (27) does not have a solution either. For example, it suffices to show that the sums A_{N_n} do not concentrate on a subgroup of lower dimension for a sequence N_n such that $T_\alpha^{N_n} \to \mathrm{Id}$. Indeed, if a solution to (27) exists then $|\psi|$ is constant by ergodicity of the base translation. Therefore by Lebesgue Dominated Convergence Theorem

$$\lim_{n\to\infty} \int_{\mathbb{T}^d} e^{i2\pi\langle \lambda, A_{N_n}(x)\rangle} dx = \lim_{n\to\infty} \int_{\mathbb{T}^d} \psi(x+N_n\alpha)/\psi(x)dx = 1$$

which means that $A_{N_n}(x)$ is concentrated near the set

$$\{u \in \mathbb{R}^r : \langle \lambda, u\rangle \in \mathbb{Z}\}.$$

In particular it was shown, in [**35**], that for every Liouville translation vector $\alpha \in \mathbb{R}^d$, the generic smooth function A does not admit a solution to (27) for any $\lambda \in \mathbb{R}^d - \{0\}$. Hence the generic smooth skew product above a Liouville translation is ergodic (cf. Section 9.3 and Theorem 42 in Section 10).

It is known that ergodic skew products $S_{\alpha,A}$ are actually uniquely ergodic (see [**100**]). On the other hand, skew products above translations are never weak mixing since they have the translation as a factor. However, the same ideas as the ones used to prove ergodicity of the skew products often prove that all eigenfunctions come from that factor (see [**45, 42, 58, 59, 128**]).

If one considers skew products on $\mathbb{T} \times \mathbb{T}$ with smooth increasing functions on $(0,1)$ having a jump discontinuity at 0 then the corresponding skew product will even be mixing in the fibers, that is, the correlations between functions that depend only on the fiber coordinate tends to 0. A classical example is given by the skew shift $(x, y) \mapsto (x + \alpha, y + x)$. The mixing in the fibers can be easily derived from the invariance of V_λ defined by (28) and the fact that, by the Ergodic Theorem, $\frac{\partial A_n}{\partial x} = \left(\frac{\partial A}{\partial x}\right)_n \to +\infty$. A similar phenomenon can occurs for analytic skew products that are homotopic to identity but over higher dimensional tori $\mathbb{T}^d \times \mathbb{T} \ni (x, y) \mapsto (x+\alpha, y+\phi(x))$, with α and ϕ as in Theorem 47 below (see [**37**]). This mechanism can also be used to establish ergodicity of cylindrical cascades (see [**102**]). A fast decay of correlations in the fibers can be responsible for the existence of non trivial invariant distributions for these skew products similarly to what occurs for the skew shift $(x, y) \mapsto (x + \alpha, x + y)$ (see [**60**]).

The deviations of ergodic sums for skew products, that is the behavior of the sums

$$\sum_{n=0}^{N-1} B(S^n_{\alpha,A}(x, z)) - N \int_{\mathbb{T}^d} \int_{\mathbb{T}^r} B(x, z) dx dz$$

is poorly understood. The only cases where some results are available have significant extra symmetry [**60, 91, 41**].

9.2. Recurrence. Our next topic are cylindrical cascades. As it was mentioned above they are sometimes called deterministic random walks. So the first question one can ask is if the walk is **recurrent** (that is, A_N returns to some bounded region infinitely many times) or **transient**. We will assume in this section that A has zero mean since otherwise $W_{\alpha,A}$ is transient by the ergodic theorem. If $r = 1$ this condition is also sufficient. In fact, the next result is valid for skew products over arbitrary ergodic transformations (in fact, there is a multidimensional version of this result, see Theorem 32).

THEOREM 29. *([**5**]) If $r = 1$, A is integrable and has zero mean then W is recurrent.*

9.2.1. *Recurrence and the Denjoy Koksma Property.* Next we note that if the base dimension $d = 1$ and A has bounded variation then W is recurrent for all r and for all $\alpha \in \mathbb{R} - \mathbb{Q}$ due to the Denjoy-Koksma inequality stating that

$$\max_{x\in\mathbb{T}} |A_{q_n} - q_n \int_{\mathbb{T}} A(y)dy| \leq 2V \tag{29}$$

for every denominator of the convergence of α, where V is the total variation of A.

More generally we say that A (not necessarily of zero mean) has **the Denjoy-Koksma property** (DKP) if there exist constants $C, \delta > 0$ and a sequence $n_k \to \infty$ such that

$$\mathbb{P}(|A_{n_k} - n_k \int_{\mathbb{T}^d} A(y)dy| \leq C) \geq \delta. \tag{30}$$

We say that A has **the strong Denjoy-Koksma property** (sDKP) if (30) holds with $\delta = 1$.

Note that if DKP holds and A has zero mean then the set of points where $\liminf_{n\to\infty} |A_n| \leq C$ has positive measure and so by ergodicity of the base map $W_{\alpha,A}$ is recurrent.

Later, we will also see how the DKP can be very helpful in proving ergodicity of the cylindrical cascades as well as weak mixing of special flows.

The situation with DKP for translations on higher dimensional tori is delicate. Of course it holds for almost all α and for every smooth function by the existence of smooth solutions to the linear cohomological equation (3). But the DKP also holds above most translations even from a topological point of view.

THEOREM 30. *([**36**]) There is a residual set of vectors in $\alpha \in \mathbb{R}^d$ such that DKP holds above T_α for every function that is of class C^4.*

In fact, it is non-trivial to construct rotation vectors and smooth functions that do not have the DKP. The first construction is due to Yoccoz and it actually provides examples of non recurrent analytic cascades.

THEOREM 31. *([**129**, Appendix]) For $d = 2$ there exists an uncountable dense set Y of translation vectors and a real analytic function $\mathcal{A} : \mathbb{T}^2 \to \mathbb{C}$ with zero mean such that W is not recurrent.*

Denote the translation vector by (α', α''). The main ingredient in the construction of [**129**] is that the denominators, q'_n and q''_n of the convergents of α' and α'' are alternated, and more precisely, they are such that the sequence $...q'_n, q''_n, q'_{n+1}, q''_{n+1}...$ increases exponentially. We will see later that the same construction can be used to create examples of mixing special flows with an analytic ceiling function.

Let Y be the set of couples $(\alpha', \alpha'') \in \mathbb{R}^2 - \overline{\mathbb{Q}}^2$, whose sequences of best approximations q'_n and q''_n satisfy, for any $n \geq n_0(\alpha', \alpha'')$

$$q''_n \geq e^{3q'_n}, \quad q'_{n+1} \geq e^{3q''_n}.$$

Then [**129**] constructs a real analytic function $\mathcal{A} : \mathbb{T}^2 \to \mathbb{C}$ with zero integral such that for almost every $(x, y) \in \mathbb{T}^2$ $|A_n(x, y)| \to \infty$, hence $W_{\alpha,\mathcal{A}}$ is not recurrent. Note that the set Y as defined above is uncountable and dense in $\mathbb{R}^2$.

9.2.2. *Indicator functions.* Now we specify the study of $W_{\alpha,A}$ to the case where

$$A = (\chi_{\Omega_j} - \mathrm{Vol}(\Omega_j))_{j=1,\dots,r} \text{ where } \Omega_j \subset X = \mathbb{T}^d \text{ are regular sets.}$$

If $d > 1$ then the DKP does not seem to be well adapted for proving recurrence in this case (see Questions 6–10).

QUESTION 28. *Show that DKP does not hold when $d > 1$ and $A = (\chi_{\Omega_j} - \mathrm{Vol}(\Omega_j))_{j=1,\dots,r}$ and the $\Omega_j \subset X$ are balls or boxes.*

There is however another criterion for recurrence which is valid for arbitrary skew products.

THEOREM 32. *Given a sequence $\delta_n = o(n^{1/r})$ the following holds.*

*(a) ([**22**]) Consider the map $T : X \to X$ preserving a measure μ. Let $W(x, y) = (Tx, y + A(x))$. If there exists a sequence k_n such that $\lim_{n\to\infty} \mu(x : A_{k_n}(x) \leq \delta_n) = 1$ then W is recurrent.*

(b) Consider a parametric family of maps $T_\alpha : X \to X$, $\alpha \in \mathfrak{A}$. Assume that T_α preserves a measure μ_α. Let (α, x) be distributed according to a measure $\boldsymbol{\lambda}$ on $\mathfrak{A} \times X$ such that $d\boldsymbol{\lambda} = d\nu(\alpha) d\mu_\alpha(x)$ for some measure ν on $\mathfrak{A}$. If $\dfrac{\sum_{n=0}^{N-1} A(T_\alpha^n x)}{\delta_N}$ has a limiting distribution as $N \to \infty$ then $W_{\alpha,A}$ is recurrent for ν-almost all α.

Note that T is not required to be ergodic. On the other hand if T is ergodic, $r = 1$ and A has zero mean, then by the Ergodic Theorem

$\mu(|A_n/n| > \varepsilon) \to 0$ for any ε so one can take $k_n = n$ and $\delta_n = \varepsilon_n n$ where $\varepsilon_n \to 0$ sufficiently slowly. Therefore Theorem 32 implies Theorem 29.

PROOF. (a) Suppose B is a wondering set (that is, $W^k B$ are disjoint) of positive measure which is contained in $\{|z| < C\}$. Let

$$B_n = \{(x, z) \in B : A_{k_n}(x) \leq \delta_n\}.$$

Then $\mu(B_n) \to \mu(B)$ as $n \to \infty$ so for large n

$$\mu(\cup_{1\leq i\leq n} W^{k_i}(B_{k_i})) \geq n\frac{\mu(B)}{2}.$$

On the other hand, by assumption $W^{k_i}(B) \subset E_i := \{y \leq 2C+\delta_i\} \subset E_n$ if $i \in [1, n]$. Hence $\mu(\cup_{1\leq i\leq n} W^{k_i}(B_{k_i})) \leq \delta_n^r = o(n)$, a contradiction.

(b) follows from (a) applied to the map $\mathcal{T} : (\mathfrak{A} \times X) \times \mathbb{R}^r$ given by $\mathcal{T}(\alpha, x, y) = (\alpha, W_{\alpha,A}(x, y))$. $\square$

Combining Theorems 10 and 32(b) we obtain

COROLLARY 33. *If $\{\Omega_j\}_{j=1,\dots,r}$ are real analytic and strictly convex and $\frac{(d-1)}{2d} < \frac{1}{r}$ then W is recurrent for almost all α.*

Note that the proof of Theorem 32 is not constructive.

QUESTION 29. *(a) Construct α and $\{\Omega_j\}_{j=1,\dots,r}$ for which the corresponding W is non recurrent.*

(b) Find explicit arithmetic conditions which imply recurrence.

THEOREM 34. *([**22**]) (a) If $\{\Omega_j\}_{j=1,\dots,r}$ are polyhedra then W is recurrent for almost all α.*

(b) There are polyhedra $\{\Omega_j\}_{j=1,\dots,r}$ and α in $\mathbb{T}^2$ such that W is transient.

Here part (a) follows from Theorem 32 and a control on the growth of the ergodic sums. Namely it is proven in [**22**] that given any polyhedron $\Omega \subset \mathbb{T}^d$ then for any $\gamma > 0$, it holds that for almost every $\alpha \in \mathbb{R}^d$, $\|A_n\|_2 = O(n^\gamma)$ where $A = \chi_\Omega - \mathrm{Vol}(\Omega)$, the sums are considered above the translation T_α and the L_2 norm is considered with respect to the Haar measure on $\mathbb{T}^d$. In the case of boxes, the latter naturally follows from the power log control given by Beck's Theorem (see Section 3.1).

The proof of part (b) proceeds by extending the method of [**129**] discussed in Section 9.2.1 to the case of indicator functions.

QUESTION 30. *Is it true that for a generic choice of Ω_j as in Question 33, W is transient for almost all α when $\frac{(d-1)}{2d} > \frac{1}{r}$?*

An affirmative answer to Question 18 (Local Limit Theorem) would give evidence that Question 18 may be true due to Borel-Cantelli Lemma. (More precisely, to answer Question 30 we need a joint Local Limit Theorem for ergodic sums of indicators of several sets.)

QUESTION 31. *Let α be as in Theorem 34 (a) or Question 33. Does there exist x such that* $\lim_{N\to\infty} ||A_N(x)|| = \infty$*?*

Note that this is only possible if $d > 1$ due to the Denjoy-Koksma inequality. On the other hand in any dimension one can have orbits which stay in a half space. Such orbits have been studied extensively (see [**103**] and the references wherein).

Another case where recurrence is not easy to establish is that of skew products over circle rotations with functions having a singularity such as the examples discussed in Section 2. We will come back to this question in the next section.

9.3. Ergodicity. Next we discuss the ergodicity of cylindrical cascades. Here one has to overcome both problems of recurrence discussed in Section 9.2 and issues of non-arithmeticity appearing in the study of ergodicity of $S_{\alpha,A}$.

The ergodicity of $W_{\alpha,A}$ is usually established using the fact that the sums A_{N_n} are increasingly well distributed on $\mathbb{R}^r$ when considered above any small scale balls in the base and for some *rigidity* sequence N_n, i.e. such that $\|N_n\alpha\| \to 0$. More precisely, usual methods of proving their ergodicity take into consideration a sequence of distributions

$$(A_{n_k})_* (\mu), \; k \geq 1 \tag{31}$$

along some rigidity sequence $\{n_k\}$ as probability measures on $\bar{\mathbb{R}}^r$ where $\bar{\mathbb{R}}$ is the one-point compactification of $\mathbb{R}$. As shown in [**85**] each point in the topological support of a limit measure of (31) is a so called *essential value* for $W_{\alpha,A}$. Following [**115**] $a \in \mathbb{R}^r$ is called an **essential value** of A if for each $B \in \mathbb{T}^d$ of positive measure, for each $\epsilon > 0$ there exists $N \in \mathbb{Z}$ such that

$$\mu(B \cap T^{-N} B \cap [|A_N(\cdot) - a| < \epsilon]) > 0.$$

Denote by $E(A)$ the set of essential values of A. Then the essential value criterion states as follows

THEOREM 35. *(*[**115**],[**1**]*)*
(a) $E(A)$ is a closed subgroup of $\mathbb{R}^r$.
(b) $E(A) = \mathbb{R}^r$ iff $W_{\alpha,A}$ is ergodic.
(c) If A is integer valued and $E(A) = \mathbb{Z}^r$ then $W_{\alpha,A}$ is ergodic on $\mathbb{T}^d \times \mathbb{Z}^r$.

Hence if the supports of the probability measures in (31) are increasingly dense on $\mathbb{R}^r$ then $W_{\alpha,A}$ is ergodic.

The case where $d = r = 1$ is the most studied although there are still some open questions in this context. For $d = r = 1$ ergodicity is often proved using the Denjoy Koksma Property. Indeed, if A is not cohomologous to a constant then $A_N - N \int A$ are not bounded. Let q_n be a best denominator for the base rotation. Pick K_n which is large but not too large. Then Kq_n is still a rigidity time for the translation but A_{Kq_n} have sufficiently large albeit controlled oscillations to yield that a given value a in the fibers is indeed an essential value.

This method is actually well adapted to A whose Fourier transform satisfies $\hat{A}(n) = \mathrm{O}(1/|n|)$, since they display a DKP (see [**84**]). Example of such functions are functions of bounded variation and functions smooth everywhere except for a log symmetric singularity.

Ergodicity also holds in general for characteristic functions of intervals.

THEOREM 36. *(a)* [**36**] *If α is Liouville, there is a residual set of smooth functions A with zero integral such that the skew product $W_{\alpha,A}$ is ergodic.*

*(b) ([**43**]) If A has symmetric logarithmic singularity then $W_{\alpha,A}$ is ergodic for all irrational α.*

*(c) ([**24**]) If $A = \chi_{[0,1/2]} - \chi_{[1/2,1]}$ and α is irrational then $W_{\alpha,A}$ is ergodic on $\mathbb{T}^1 \times \mathbb{Z}$.*

*(d) ([**97**]) If $A = \chi_{[0,\beta]} - \beta$ then $W_{\alpha,A}$ is ergodic iff $1, \alpha$ and β are rationally independent.*

*(e) ([**102**]) If A is piecewise absolutely continuous, $\int_{\mathbb{T}^1} A(x)dx = 0$, A' is Riemann integrable and $\int_{\mathbb{T}^1} A'(x)dx \neq 0$ then $W_{\alpha,A}$ is ergodic for all $\alpha \in \mathbb{R} - \mathbb{Q}$.*

*(f) ([**23**]) If $A : \mathbb{T} \to \mathbb{R}^r = (A_1, \dots, A_r)$ with $A_j = \sum c_{j,i}\chi_{I_{j,i}} - \beta_j$ with $I_{j,i}$ a finite family of intervals, $c_{j,i} \in \mathbb{Z}$ and β_j is such that $\int_{\mathbb{T}} A_j(x)dx = 0$ and if the sequence $(\{q_n\beta_1\}, \dots, \{q_n\beta_r\})$ is equidisitributed on $\mathbb{T}^r$ as $n \to \infty$, where q_n is the sequence of denominators of α, then $W_{\alpha,A}$ is ergodic. In the case $r = 1$, it is sufficient to ask that $(\{q_n\beta\})$ has infinitely many accumulation points, then $W_{\alpha,A}$ is ergodic.*

For further results on the ergodicity of cascades defined over circle rotations with step functions as in (f), we refer to the recent work [**25**].

The proofs of (a) and (b) are based on DKP and progressive divergence of the sums as explained above. (c)–(e) are treated differently since the ergodic sums take discrete values. For example, the proof of (e) in the case $r = 1$ is based on the fact that A_{q_n} is bounded by

DKP and then the hypothesis on $\{q_n\beta\}$ implies that the set of essential values is not discrete, hence it is all of $\mathbb{R}$, and the ergodicity follows.

The cases of slower decay of the Fourier coefficients of A are more difficult to handle. We have nevertheless a positive result in the particular situation of log singularities.

THEOREM 37. [**39**] *If A has (asymmetric) logarithmic singularity then $W_{\alpha,A}$ is ergodic for almost every α.*

The delicate point in Theorem 37 is that the DKP does not hold. Indeed, it was shown in [**119**] that the special flow above T_α and under a function that has asymmetric log singularity is mixing for a.e. α. But, as we will see in the next section, mixing of the special flow is not compatible with the DKP. *A contrario* special flows under functions with symmetric logarithmic singularities are not mixing [**72, 84**] because of the DKP.

In the proof of Theorem 37, one first shows that the DKP (30) holds if the constant δ is replaced by a sequence δ_n which decays sufficiently slowly and then uses this to push through the standard techniques under appropriate arithmetic conditions.

The case of general angles for the base rotation or the case of stronger singularities are harder and all questions are still open.

QUESTION 32. *Are there examples of ergodic cylindrical cascades with smooth functions having power like singularities?*

Conversely, we may ask the following

QUESTION 33. *Are there examples of non ergodic cylindrical cascades with smooth functions having non symmetric logarithmic or (integrable) power singularities?*

The study of ergodicity when $d > 1$ and $r > 1$ is more tricky essentially because of the absence of DKP.

For smooth observable, only the Liouville frequencies are interesting. The ergodic sums above such frequencies tend to stretch at least along a subsequence of integers. And this stretch usually occurs gradually and independently in all the coordinates of A hence a positive answer to the following question is expected.

QUESTION 34. *Show that for any Liouville vector α, there is a residual set of smooth functions A with zero integral such that the skew product $W_{\alpha,A}$ is ergodic.*

As we discussed in the proof of Theorem 37, the cylindrical cascade on $\mathbb{T} \times \mathbb{R}$ with a function A having an asymmetric logarithmic singularity is ergodic for almost every α although the ergodic sums A_N above T_α concentrate at infinity as $N \to \infty$. The slow divergence of these sums that compare to $\ln N$ (see Question 1) plays a role in the proof of ergodicity. The logarithmic control of the discrepancy relative to a polyhedron (see Theorems 6, 11 and 34) motivates the following question.

QUESTION 35. *Is it true that for (almost) every polyhedra $\Omega_j \subset \mathbb{T}^d$, $j = 1, \dots, r$, and $A = (\chi_{\Omega_j} - \mathrm{Vol}(\Omega_j))_{j=1,\dots,r}$, the cascades $W_{\alpha,A}$ are ergodic?*

We note that the answer is unknown even for boxes with $d = 2$ and $r = 1$.

9.4. Rate of recurrence. Section 9.3 described several situations where $W_{\alpha,A}$ is ergodic. However for infinite measure preserving transformations the (ratio) ergodic theorem does not specify the growth of ergodic sums. Rather it shows that for any L^1 functions $B_1(x, y), B_2(x, y)$ with $B_2 > 0$ we have

$$\frac{\sum_{n=0}^{N-1} B_1(W^n_{\alpha,A}(x, y))}{\sum_{n=0}^{N-1} B_2(W^n_{\alpha,A}(x, y))} \to \frac{\iint B_1(x, y)dxdy}{\iint B_2(x, y)dxdy}. \tag{32}$$

In fact ([**1**]) there is no sequence a_N such that

$$\frac{\sum_{n=0}^{N-1} B_1(W^n_{\alpha,A}(x, y))}{a_N} \tag{33}$$

converges to 1 almost surely. On the other hand, one can try to find a_N such that (33) converges in distribution. By (32) it suffices to do it for one fixed function B. For example one can take $B = \chi_{B(0,1)}$. This motivates the following question.

QUESTION 36. *Let α be as in Theorem 34 (a) or Question 33. How often is $||W^N|| \le R$?*

So far this question has been answered only in a special case. Namely, let $d = r = 1$, $Z_N(x) = \sum_{n=0}^{N-1} \left[\chi_{[0,1/2]}(x + n\alpha) - 1/2\right]$. Denote $\mathbb{L}_N = \mathrm{Card}(n \le N : Z_n = 0)$.

THEOREM 38. [**6**] *If α is a quadratic surd then there exists a constant $c = c(\alpha)$ such that $\frac{\sqrt{\ln N}}{cN}\mathbb{L}_N$ converges to $e^{-\mathfrak{N}^2/2}$.*

Similar results have been previously obtained by Ledrappier-Sarig for abelian covers of compact hyperbolic surfaces ([**82**]). The fact that the correct normalization is $N/\sqrt{\ln N}$ was established in [**2**].

Question 37. *Extend Theorem 38 to the case when* $1/2$ *is replaced by*

(a) any rational number;

(b) any irrational number, (in which case one needs to replace $\{A_N = 0\}$ *by* $\{|A_N| \le 1\}$*).*

Question 38. *What happens for typical* α?

Note that in contrast to Theorem 38, $\mathbb{L}_N \sim \dfrac{N}{\ln N}$ (rather than $\mathbb{L}_N \sim \dfrac{N}{\sqrt{\ln N}}$) is expected in view of Kesten's Theorem 9. Ideas of the proof of Theorem 38 will be described in Section 10.5.

10. Special flows.

10.1. Ergodic integrals. In this section we consider **special flows** above T_α which will be denoted $T^t_{\alpha,A}$. Here $A(\cdot) > 0$ is called the **ceiling function** and the flow is given by

$$\begin{aligned} \mathbb{T}^d \times \mathbb{R}/\sim \;&\to\; \mathbb{T}^d \times \mathbb{R}/\sim \\ (x,s) \;&\to\; (x, s+t), \end{aligned}$$

where $\sim$ is the identification

$$(x, s + A(x)) \sim (T_\alpha(x), s). \tag{34}$$

Equivalently the flow is defined for $t + s \ge 0$ by

$$T^t(x,s) = (x + n\alpha, t + s - A_n(x))$$

where n is the unique integer such that

$$A_n(x) \le t + s < A_{n+1}(x). \tag{35}$$

Since T_α preserves a unique probability measure μ then the special flow will preserve a unique probability measure that is the normalized product measure of μ on the base and the Lebesgue measure on the fibers.

Special flows above ergodic maps are always ergodic for the product measure constructed as above. The interesting feature of special flows is that they can be more "chaotic" then the base map, displaying properties such as weak mixing or mixing even if the base map does not have them. Actually any map of a very wide class of zero entropy measure theoretic transformations, so called Loosely Bernoulli maps, are isomorphic to sections of special flows above any irrational rotation of the circle with a continuous ceiling function (see [**98**]).

If $A = \beta$ is constant then $T_{\alpha,A}$ is the linear flow on $\mathbb{T}^{d+1}$ with frequency vector (α, β). Thus special flows $T^t_{\alpha,A}$ can be viewed as time changes of translation flows on $\mathbb{T}^{d+1}$. In particular, if we consider the linear flow on $\mathbb{T}^{d+1}$ and multiply the velocity vector by a smooth *non-zero* function ϕ we get a special flow with a smooth ceiling function A.

10.2. Smooth time change. We recall that a translation flow frequency $v \in \mathbb{R}^d$ is said to be Diophantine if there exists $\sigma, \tau > 0$ such that $||(k, v)|| \geq C|k|^{-\sigma}$ for every $k \in \mathbb{Z}^d$. Hence a translation vector $(1, v) \in \mathbb{R}^{d+1}$ is Diophantine (homogeneous Diophantine or Diophantine in the sense of flows) if and only if v is a Diophantine vector in the sense of (2).

THEOREM 39. [**76**] *Smooth non vanishing time changes of translation flows with a Diophantine frequency vector are conjugated to translation flows.*

PROOF. Let v be a constant vector field on $\mathbb{T}^{d+1}$. We suppose WLOG that $v = (1, \alpha)$. Let $u(x)$ be a smooth function on the torus and $\dot{x} = u(x)v$. Then, making a change of variables $y = T_v^{\phi(x)}(x)$ we obtain the equation $\dot{y} = (\phi + \partial_v \phi)(y)u(y)v$. The equation for y is linear if $\phi + \partial_v \phi = \dfrac{c}{u}$. Passing to Fourier series, this equation can be solved if $c = \int \phi(x)dx \left(\int \dfrac{dx}{u(x)} \right)^{-1}$ and v is such that $|1 + (k, v)|| \geq C|k|^{-\sigma}$ for every $k \in \mathbb{Z}^{d+1}$ which is equivalent to α Diophantine as in (2).

One can also see this fact at the level of the special flow $T_{\alpha,A}$ associated to $\dot{x} = u(x)v$. Namely, making a change of variables $(y, s) = T_{\alpha,A}^{B(x,t)}(x, t)$ transforms $T_{\alpha,A}$ to $T_{\alpha,D}$ with

$$D(x) = A(x) + B(x + \alpha, 0) - B(x, 0)$$

so one can make the LHS constant provided α is Diophantine. Finally, the similarity between linear and nonlinear flows in the Diophantine case is also reflected in (35) since for Diophantine vectors α $A_n = n \int A(x)dx + O(1)$. $\square$

An interesting question is that of deviations of ergodic sums above time changed linear flows. In fact, the case of linear flows is already non trivial and can be studied by the methods described in Section 7.3. More precisely, as for translations the interesting case occurs when the function under consideration has singularities, for example, for indicator functions.

Namely, given a set Ω let

$$\mathbf{D}_\Omega(r, v, x, T) = \int_0^T \chi_{\Omega_r}(T_v^t x)dt - T\mathrm{Vol}(\Omega_r) \tag{36}$$

where T_v^t denotes the linear flow with velocity v.

We assume that (x, v, r) are distributed according to a smooth density.

THEOREM 40. *([**28, 29**]). Suppose that Ω is analytic and strictly convex.*

(a) If $d = 2$ then $\mathbf{D}_\Omega(r, v, x, T)$ converges in distribution.

(b) If $d = 3$ then $\frac{\mathbf{D}_\Omega(r,v,x,T)}{\ln T}$ converges to a Cauchy distribution.

(c) If $d \geq 4$ then $\frac{\mathbf{D}_\Omega(r,v,x,T)}{r^{\frac{d-1}{2}} T^{\frac{d-3}{2(d-1)}}}$ has limiting distribution as $T \to \infty$.

(d) For any $d \in \mathbb{N}$, if Ω is a box then $\mathbf{D}_\Omega(r, v, x, T)$ converges in distribution.

The proof of Theorem 40 is similar to the proofs of Theorems 10 and 11 and Corollary 1.

COROLLARY 41. *Theorem 40 remains valid for time changes $T_{u(x)v}$ where $u(x)$ is fixed smooth positive function and v is random as in Theorem 40.*

PROOF. To fix our ideas let us consider the case where Ω is analytic and strictly convex. Note that $T_{uv}^t x = T_v^{\tau(x,t)}$ where the by the above discussion the function τ satisfies

$$\tau(t, x) = at + \varepsilon(t, x, v) \text{ where } a = \left(\int \frac{dx}{u(x)}\right)^{-1}$$

and $\varepsilon(t, x, v)$ is bounded for almost all v uniformly in x and t. Accordingly it suffices to see how much time is spend inside Ω_r for the linear segment of length at.

Next if the linear flow stays inside Ω_r during the time $[t_1, t_2]$ then the time spend in Ω_r by the orbit of T_{uv} equals to $\int_{t_1}^{t_2} \frac{dt}{u(x(t))}$. Thus we need to control the following integral for linear flow

$$\tilde{\mathbf{D}}_\Omega(r, v, x, T) = \int_0^T \frac{\chi_{\Omega_r}(T_v^t x)}{u(T_v^t x)} dt - T \int_{\Omega_r} \frac{dx}{u(x)}.$$

However the Fourier transform of $\frac{\chi_{\Omega_r}(x)}{u(x)}$ has a similar asymptotics at infinity as the Fourier transform of $\chi_{\Omega_r}(x)$ (see [**123**]) so the proof of the Corollary proceeds in the same way as the proof of Theorem 40 in [**28**]. $\square$

Up to now, we were interested in smooth time change of linear flows with typical frequencies. We will further discuss smooth time changes for special frequencies in Section 10.4 devoted to mixing properties.

10.3. Time change with singularities. If the time changing function of an irrational flow has zeroes then the ceiling function of the corresponding special flow has poles. In this case the smooth invariant measure is infinite. In the case of a unique singularity, we have that the time changed flow is uniquely ergodic with the Dirac mass at the singularity the unique invariant probability measure:

PROPOSITION 5. *Consider a flow T^t given by a smooth time change of an irrational linear flow obtained by multiplying the constant vector field by a function which is smooth and non zero everywhere except for one point x_0, then for any continuous function b and any x*

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t b(T^u x)du = b(x_0).$$

PROOF. To simplify the notation we assume that the time change preserves the orientation of the flow. We use the representation as a special flow $T_{\alpha,A}$ with A having a pole.

It suffices to prove this statement in case b equals to 0 in a small neighborhood of x_0. In that case we have

$$\int_0^t b(T^u_{\alpha,A}(x,s))du = B_{n(t)}(x) + O(1) \tag{37}$$

where $B(x) = \int_0^{A(x)} b(x,s)ds$ and $n(t)$ is defined by (35). If b vanishes in a small neighborhood of x_0 then B is bounded and so $|B_{n(t)}| \leq Cn(t)$. Therefore it suffices to show that $\frac{n(t)}{t} \to 0$ which is equivalent to $\frac{A_n}{n} \to \infty$. Let $\tilde{A}$ be a continuous function which is less or equal to A everywhere. Then

$$\liminf \frac{A_n}{n} \geq \lim_{n\to\infty} \frac{\tilde{A}_n}{n} = \int \tilde{A}(x)dx.$$

Since $\int A(x)dx = \infty$ we can make $\int \tilde{A}(x)dx$ as large as possible proving our claim. $\square$

QUESTION 39. *In the setting of Proposition 5 describe the deviations of ergodic integrals from $b(x_0)$.*

QUESTION 40. *Consider the case where the time change has finite number of zeroes $x_1, x_2, \dots, x_m$. In that case all limit measures are of the form $\sum_{j=1}^m p_j\delta_{x_j}$. Which p_j describe the behavior of Lebesgue-typical points?*

In view of the relation (37) these questions are intimately related to Theorems 3 and 5 and Questions 2, 4 and 5 from Section 2.

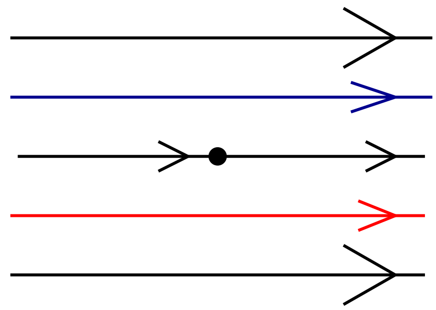

Figure 2. Kocergin Flow is topologically equivalent to the area preserving flow shown on Figure 1 with separatrix loop removed. The rest point is responsible for the shear along the orbits.

If one is interested in flows with singularities but that preserve a finite non-atomic measure then the simplest example can be obtained by plugging (by smooth surgery) in the phase space of the minimal two dimensional linear flow an isolated singularity coming from a Hamiltonian flow in $\mathbb{R}^2$ (see Figure 2). The so called Kochergin flows thus obtained preserve besides the Dirac measure at the singularity a measure that is equivalent to Lebesgue measure [**71**]. As it was explained in Section 2 Kochergin flows model smooth area preserving flows on $\mathbb{T}^2$. These flows still have $\mathbb{T}$ as a global section with a minimal rotation for the return map, but the slowing down near the fixed point produces a singularity for the return time function above the last point where the section intersects the incoming separatrix of the fixed point. The strength of the singularity depends on how abruptly the linear flow is slowed down in the neighborhood of the fixed point. A mild slowing down, or mild *shear*, is typically represented by the logarithm while stronger singularities such as $x^{-a}, a \in (0, 1)$ are also possible. Power-like singularities appear naturally in the study of area preserving flows with degenerate fixed points. We shall see below that dynamical properties of the special flows are quite different for logarithmic and power like singularities.

Question 41. *What can be said about the deviations of the ergodic sums above Kocergin flows?*

10.4. Mixing properties. We give first a classical criterion for weak mixing of special flows. Its proof is similar to the proof of the ergodicity criterion for skew products given by Proposition 4.

Proposition 6. *(*[**126**]*) $T_{\alpha,A}$ is weak mixing iff for any $\lambda \in \mathbb{R}^*$, there are no measurable solutions to the multiplicative cohomological* *equation*

$$e^{i2\pi\lambda A(x)} = \psi(x+\alpha)/\psi(x). \tag{38}$$

Indeed if $h(x,t)$ is the eigenfunction when for almost all x the function $h(x,t)e^{-\lambda t}$ takes the same value $\psi(x)$ for almost almost all t. Then (38) follows from the identification (34).

THEOREM 42. *(*[**35**]*) If the vector $\alpha \in \mathbb{R}^d$ is not β-Diophantine then there exists a dense G_δ for the $C^{\beta+d}$ topology, of functions $\varphi \in C^{\beta+d}(\mathbb{T}^d, \mathbb{R}_+^*)$, such that the special flow constructed over T_α with the ceiling function φ is weak mixing.*

This result is optimal since smooth time changes of linear flows with Diophantine vectors α, are smoothly conjugated to the linear flow and, hence, are not weak mixing.

Mixing of special flows is more delicate to establish since one needs to have uniform distribution on increasingly large scales in $\mathbb{R}^+$ of the sums A_N for *all* integers $N \to \infty$, and this above arbitrarily small sets of the base space. Indeed mixing of special flows above non mixing base dynamics is in general proved as follows: if the ergodic sums A_N become as $N \to \infty$ uniformly stretched (well distributed inside large intervals of $\mathbb{R}_+$) above small sets, the image by the special flow at a large time T of these small sets decomposes into long strips that are well distributed in the fibers due to uniform stretch and well distributed in projection on the base because of ergodicity of the base dynamics (see Figure 3).

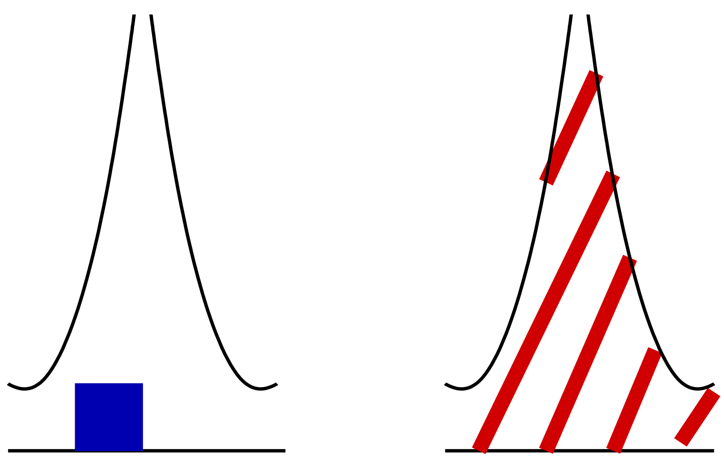

FIGURE 3. Mixing mechanism for special flows: the image of a rectangle is a union of long narrow strips which fill densely the phase space.

The delicate point however is to have uniform stretch for all integers $N \to \infty$. In particular the following result has been essentially proven in [**70**].

THEOREM 43. *If A has DKP then $T^t_{\alpha,A}$ is not mixing.*

PROOF. If A has the DKP then there is a set Ω of positive measure on which (30) holds for positive density of n_k. By passing to a subsequence we can find a set I of positive measure, a sequence $\{t_k\}$ and a vector β such that on Ω $|A_{n_k} - t_k| < C$ and $\alpha n_k \to \beta$. Pick a small η.

$$\Omega_i = \cup_{0 \le t \le \eta} T^t_{\alpha,A}[I \times \{0\}], \quad \Omega_f = \cup_{0 \le |t| \le C+\eta} T^t_{\alpha,A}[(I+\beta) \times \{0\}].$$

By decreasing I if necessary we obtain that those sets have measures strictly between 0 and 1. On the other hand it is not difficult to see from the definition of the special flow that $\mu(T^{t_k}_{\alpha,A}\Omega_i \cap \Omega_f) \to \mu(\Omega_i)$ contradicting the mixing property. □

In particular the flows with ceiling functions A of bounded variation or functions with symmetric log singularities are not mixing.

In fact, since the sDKP holds for any minimal circle diffeomorphism, it follows from (35) and (37) that any smooth flow on $\mathbb{T}^2$ without cycles or fixed points is not topologically mixing. We leave this as an exercise for the reader.

The first positive result about mixing of special flows is obtained in [**71**].

THEOREM 44. *If $\alpha \in \mathbb{R} - \mathbb{Q}$ and A has (integrable) power singularities then $T_{\alpha,A}$ is mixing.*

The reason why the case of power singularities is easier than the logarithmic case (corresponding to non-degenerate flows on $\mathbb{T}^2$) is the following. The standard approach for obtaining the stretching of ergodic sums is to control $\frac{\partial A_n}{\partial x} = \left(\frac{\partial A}{\partial x}\right)_n$ For A as in theorem 44, $\frac{\partial A}{\partial x}$ has singularities of the type x^{-a} with $a > 1$. In this case the main contribution to ergodic sums comes from the closest encounter with the singularity (cf. Theorem 3) making the control of the stretch easier. Moreover, the strength of the singularity allows to obtain speed of mixing estimates.

THEOREM 45. *([**34**]) If α is Diophantine and A has a (integrable) power singularity then $T^t_{\alpha,A}$ is power mixing.*

More precisely, there exists a constant $\beta = \beta(\alpha)$ such that if R_1, R_2 are rectangles in $\mathbb{T} \times \mathbb{R}$ then

$$\left|\mu(R_1 \cap T^t R_2) - \mu(R_1)\mu(R_2)\right| \le C t^{-\beta}. \tag{39}$$

The exponent β in [**34**] seems to be non optimal.

QUESTION 42. *For α Diophantine find the asymptotics of the LHS of* (39).

It is interesting to surpass the threshold $\beta = 1/2$. In particular, one would like to answer the following question.

QUESTION 43. [**83**] *Can a smooth area preserving flow on $\mathbb{T}^2$ have Lebesgue spectrum?*

On the other hand for logarithmic singularities there might be cancelations in ergodic sums of $\frac{\partial A}{\partial x}$, making the question of mixing more tricky.

THEOREM 46. *Let A be as in Question 1.*
*(a) ([**71**]) If $\sum_j a_j^+ = \sum_j a_j^-$ then $T^t_{\alpha,A}$ is not mixing for any $\alpha \in \mathbb{R} - \mathbb{Q}$.*
*(b) ([**119, 72**]) If $\sum_j a_j^+ \neq \sum_j a_j^-$ then $T^t_{\alpha,A}$ is mixing for almost every $\alpha \in \mathbb{R} - \mathbb{Q}$.*
*(c) ([**73**]) If $a_j^+ - a_j^-$ has the same sign for all j then $T^t_{\alpha,A}$ is mixing for each $\alpha \in \mathbb{R} - \mathbb{Q}$.*

QUESTION 44. *([**74**]) Does the condition that $\sum_j a_j^+ \neq \sum_j a_j^-$ imply $T^t_{\alpha,A}$ is mixing* **for every** *$\alpha \in \mathbb{R} - \mathbb{Q}$?*

QUESTION 45. *([**74**]) Under the conditions of Theorems 44 and 46 is $T^t_{\alpha,A}$ mixing of all orders?*

In higher dimensions much less is known. Note that for smooth ceiling functions Theorems 30, 39 and 43 precludes mixing for a set of rotation vectors of full measure that also contains a residual set.

The following was shown in [**36**]. Recall the definition of the set Y used in Theorem 31. Define the following real analytic complex valued function on $\mathbb{T}^2$:

$$\mathcal{A}(x, y) = \left(\sum_{k=2}^{\infty} \frac{e^{i2\pi kx}}{e^k} + \sum_{k=2}^{\infty} \frac{e^{i2\pi ky}}{e^k} \right).$$

THEOREM 47. *For any $(\alpha', \alpha'') \in Y$, the special flow constructed over the translation $T_{\alpha',\alpha''}$ on $\mathbb{T}^2$, with the ceiling function $1 + \mathrm{Re}\mathcal{A}$ is mixing.*

Because of the disposition of the best approximations of α' and α'' the ergodic sums φ_m of the function φ, for any m sufficiently large, will be always stretching (i.e. have big derivatives), in one or in the other of the two directions, x or y, depending on whether m is far from q'_n or far from q''_n. And this stretch will increase when m goes to infinity. So when time goes from 0 to t, t large, the image of a small typical interval

J from the basis $\mathbb{T}^2$ (depending on t the intervals should be taken along the x or the y axis) will be more and more distorted and stretched in the fibers' direction, until the image of J at time t will consist of a lot of almost vertical curves whose projection on the basis lies along a piece of a trajectory under the translation $T_{\alpha',\alpha''}$. By unique ergodicity these projections become more and more uniformly distributed, and so will $T^t(J)$. For each t, and except for increasingly small subsets of it (as function of t), we will be able to cover the basis with such "typical" intervals. Besides, what is true for J on the basis is true for $T^s(J)$ at any height s on the fibers. So applying Fubini Theorem in two directions, first along the other direction on the basis (for a time t all typical intervals are in the same direction), and second along the fibers, we will obtain the asymptotic uniform distribution of any measurable subset, which is, by definition, the mixing property.

QUESTION 46. *Are the flows obtained in Theorem 47 mixing of all orders?*

QUESTION 47. *For which vectors $\alpha \in \mathbb{R}^d$, there exist special flows above T_α with smooth functions A such that $T_{\alpha,A}$ is mixing?*

The foregoing discussion demonstrates that both ergodicity of cylindrical cascades and mixing of special flows require a detailed analysis of ergodic sums (1). However, the estimates needed in those two cases are quite different and somewhat conflicting. Namely, for ergodicity we need to bound from below the probability that ergodic sums hit certain intervals, while for mixing one needs to rule out too much concentration. For this reason it is difficult to construct functions A such that $W_{\alpha,A}$ is ergodic while $T^t_{\alpha,c+A}$ is mixing. In fact, so far this has only been achieved for smooth functions with asymmetric logarithmic singularities. However, it seems that in higher dimensions there is more flexibility so such examples should be more common.

QUESTION 48. *Is it true that for (almost) every polyhedron $\Omega \in \mathbb{T}^d$, $d \geq 2$, and almost every $a > 0$, and almost every $\alpha \in \mathbb{T}^d$, the special flow above α and under the function $a + \chi_\Omega$ is mixing?*

Note that a positive answer to both this question and Question 35 will give a large class of interesting examples where ergodicity of $W_{\alpha,A}$ and mixing for $T_{\alpha,c+A}$ (for any c such that $c + A > 0$) hold simultaneously.

QUESTION 49. *Answer Questions 35 and 48 in the case Ω is a strictly convex analytic set.*

10.5. An application. Here we show how the geometry of special flows above cylindrical cascades can be used to study the ergodic sums.

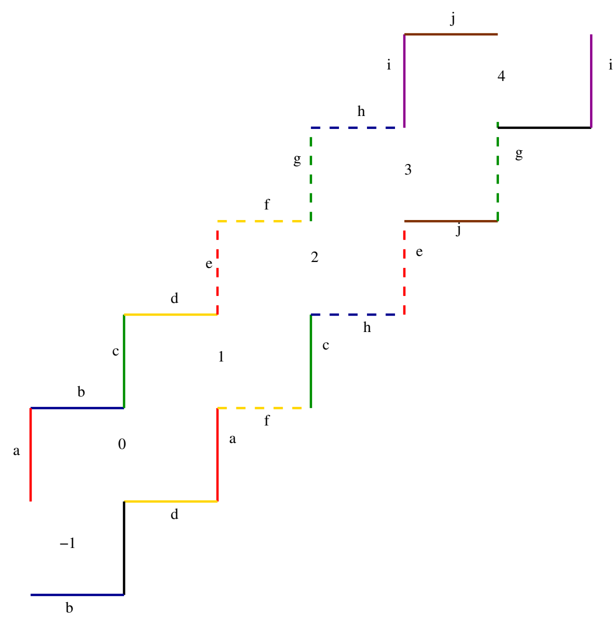

FIGURE 4. Staircase surfaces. The sides marked by the same symbol are identified.

PROOF OF THEOREM 38. The proof uses the properties of the staircase surface St shown on Figure 4. The staircase is an infinite pile of 2×1 rectangles so that the left bottom corner of the next rectangle is attached to the center of the top of the previous one. The sides which are differ by two units in either horizontal or vertical direction are identified. We number all the rectangles from $-\infty$ to $+\infty$ as shown on Figure 4. There is a translational symmetry given by $G(x, y) = (x + 1, y + 1)$ and St/G is a torus. We shall use coordinates $\bar{p} = (p, z)$ on the staircase where p are coordinates on the torus which is the identified with rectangle zero and $z \in \mathbb{Z}$ is the index of rectangle. Thus we have $(p, z) = G^z(p, 0)$.

The key step in the proof is an observation of [**53**] that St is a Veech surface. Namely, given $A \in SL_2(\mathbb{Z})$ such that $A \equiv I \mod 2$ there exists unique automorphism ϕ_A of St which commutes with G, fixes the singularities of St, has derivative A at the non-singular points and has drift 0. That is, in our coordinates

$$\phi(p, z) = (Ap, z + \tau(p)) \tag{40}$$

and the drift condition means that $\int_{\mathbb{T}^2} \tau(p)dp = 0$.

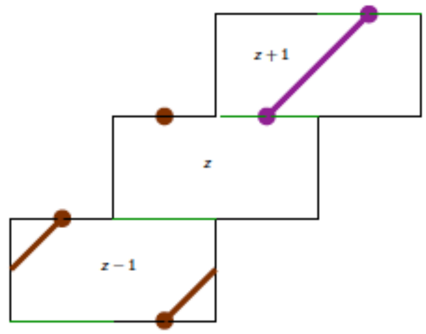

FIGURE 5. Poincare map for a linear flow on the staircase. Orbits starting from $[1/2, 1]$ go up while orbits starting from $[0, 1/2)$ have to go down due to the gluing conditions.

Consider the linear flow on St with slope θ which is locally given by $T^t(x, y) = (x + t\cos\theta, y + t\sin\theta)$. Let Π be the union of the top sides of the rectangles in St. We identify Π with $\mathbb{T} \times \mathbb{Z}$ using the map $\eta : \mathbb{T}\times\mathbb{Z} \to \Pi$ such that $\eta(x, z)$ is the point on the top side of rectangle z at the distance $2x$ from the left corner. It is easy to check (see Figure 5) that under this identification the Poincare map for T^t takes form

$$(x, z) = (x + \alpha, z + \chi_[1/2, 1](x) - \chi_{[0,1/2)}(z)) \text{ where } \alpha = \frac{\tan\theta + 1}{2}.$$

Now suppose that α and hence $\tan\theta$ is a quadratic surd. By Lagrange theorem there is $A \in SL_2(\mathbb{Z})$ such that $A\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix} = \lambda\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}$. By replacing A by A^k for a suitable (positive or negative) k we may assume that $A \equiv I \mod 2$ and that $\lambda < 1$. Let $\Gamma_N(x)$ be the ray starting from $\eta(x, 0)$ having slope θ and length $\frac{N}{\sin\theta}$.

$$\mathbb{L}_N = \frac{\text{mes}(\bar{p} \in \Gamma_N(x) : z(\bar{p}) = 0)}{\text{length}(\Gamma_N(x)} = \frac{\text{mes}(\bar{q} \in \tilde{\Gamma}(x) : z(\phi_A^{-m}) = 0)}{\text{length}(\tilde{\Gamma})}$$
$$= \mathbb{P}_x(z(\phi_A^{-m}\bar{q}) = 0)$$

where $\tilde{\Gamma} = \phi_A^m\Gamma_N(x)$ and $\mathbb{P}_x$ is computed under the assumption that $\bar{q}$ is uniformly distributed on $\tilde{\Gamma}$. Choose m to be the smallest number such that $\text{length}(\phi_A^m\Gamma_N(x)) = \lambda^m\frac{N}{\sin\theta} \le 1$. Note that $m \approx \frac{\ln N}{\ln\lambda}$. By our choice of m, $\tilde{\Gamma}$ is either contained in a single rectangle or intersects two of them. Let us consider the first case, the second one is similar. So we assume that $\tilde{\Gamma}$ is in the rectangle with index a so that $\bar{q} = (q, a)$. Due to (40) $z(\phi^{-m}\bar{q}) = a - \sum_{j=1}^{m} \tau(\phi_A^{-j} q)$. Thus

$$\mathbb{P}_x \left(z(\phi_A^{-m}\bar{q}) = 0 \right) = \mathbb{P}_x \left(\sum_{j=1}^{m} \tau(\phi_A^{-j} q = a) \right).$$

Now we apply the Local Limit Theorem for linear toral automorphisms (see [**101**, Section 4] or [**46**]) which says that there is a constant σ^2

$$\mathbb{P}_x \left(\sum_{j=1}^{n} \tau(\phi_A^{-j} q) \right) \approx \frac{1}{\sqrt{2\pi m}\sigma} e^{-a^2/2\sigma^2 m}.$$

It remains to note that

$$a(x) = \sum_{j=0}^{m-1} \tau(\phi_A^j \eta(x, 0))$$

so applying the Central Limit Theorem for linear toral automorphisms we see that if x is uniformly distributed on $\mathbb{T}^1$ then $\frac{a(x)}{\sqrt{m}}$ is approximately normal with zero mean and variance σ^2. $\square$

Next we discuss the proof of Theorem 8(b) in case $l = \frac{1}{2}$. The proof proceeds the same way as the proof of Theorem 38 with the following changes.

(I) Instead of estimating the probability that $z(\phi_A^{-m}\bar{q}) = 0$ we need to estimate the probability that $z(\phi_A^{-m}\bar{q})$ belongs to an interval of length $\sqrt{m}$ so we use the Central Limit Theorem instead of the Local Limit Theorem.

(II) Instead of taking x random we take x fixed at the origin. Note that the origin is fixed by A so $\tau(\phi_A^m(0,0)) = Cm$. (More precisely τ is multivalued at the origin since it belong to several rectangles so by $\tau(\phi_A^m(0,0))$ we mean the limit of $\tau(\phi_A^m(\bar{p}))$ as $\bar{p}$ approaches the origin inside $\Gamma_N(0)$.)

11. Higher dimensional actions

QUESTION 50. *Generalize the results presented in Sections 2-10 to higher dimensional actions.*

The orbits of commuting shifts $T^n x = x + \sum_{j=1}^{q} n_j \alpha_j$ are much less studied than their one-dimensional counterparts. We expect that some of the results of Sections 2-10 admit straightforward extensions while in other cases significant new ideas will be necessary. Below we discuss two areas of research where multidimensional actions appear naturally.

11.1. Linear forms. Statements about orbits of a single translation can be interpreted as results about joint distribution of fractional part of inhomogenuous linear forms of one variable evaluated over $\mathbb{Z}$. From the point of view of Number Theory it is natural to study linear forms of several variables evaluated over $\mathbb{Z}^d$. Let

$$l_i(n) = x_i + \sum_{j=1}^{q} \alpha_{ij} n_j, \quad i = 1 \dots d.$$

Thus it is of interest to study the discrepancy

$$\mathbb{D}_N(\Omega, \alpha, x) =$$
$$\mathrm{Card}(0 \leq n_j < N, j = 1, \dots, q : (\{l_1(n)\}, \dots \{l_d(n)\}) \in \Omega) - N^q \mathrm{Vol}(\Omega).$$

The latter problem is a classical subject in Number Theory, and there are several important results related to it. In particular, the Poisson regime is well understood ([**88**]). The following result generalizes Theorem 16 and can be proven by a similar argument.

Theorem 48. *Let (α, x) be uniformly distributed on $\mathbb{T}^{d(q+1)}$. Then the distribution of*

$$\mathrm{Card}(n : \frac{n}{N} \in \Sigma \text{ and } (\{l_1(n)\}, \dots \{l_d(n)\}) \in N^{-q/d}\Omega)$$

converges as $N \to \infty$ to

$$\mathcal{N}(\Omega, \Sigma) := \mathrm{Card}(e \in L, e = (x, y) : x(e) \in \Omega, y(e) \in \Sigma)$$

where L is a random affine lattice in $\mathbb{R}^{d+q}$.

Thus the Poisson regime for the rotations exhibits more regular behavior comparing to standard Poisson processes. However then the number of rotations becomes large the limiting distribution approaches the Poisson. Namely, the following is the special case of the result proven in [**125**].

Theorem 49. *If Σ_q are unit cubes in $\mathbb{R}^q$ then $\Omega \to \mathcal{N}(\Omega, \Sigma_q)$ converge as $q \to \infty$ to the Poisson measure $\boldsymbol{\mu}(\Omega) = \mathrm{Card}(\mathfrak{P} \cap \Omega)$ there $\mathfrak{P}$ is a Poisson process on $\mathbb{R}^d$ with constant intensity.*

Next we present extensions of Theorems 25, 24, 26 and 27 to the context of homogeneous and inhomogeneous linear forms. Let again $l_i(n) = x_i + \sum_{j=1}^{q} \alpha_{ij} n_j$, $i = 1 \dots d$. Consider

$$V_N(\alpha, x, c) = \mathrm{Card}(0 \leq n_i < N : (\{l_1(n)\}, \dots \{l_d(n)\}) \in B(c|n|^{-q/d})).$$

More generally given a function $\psi : \mathbb{R}^+ \to \mathbb{R}^+$ define

$$V_N^{\psi}(\alpha, x) = \mathrm{Card}(0 \leq n_i < N : (\{l_1(n)\}, \dots \{l_d(n)\}) \in B(\psi(|q|)).$$

We also let $U_N(\alpha, c) = V_N(0, \alpha, c)$ and $U_N^\psi(\alpha) = V_N^\psi(0, \alpha)$ be the quantities measuring the rate of recurrence.

In particular we call the matrix α **badly approximable** if there exists $c > 0$ such that for, $V_N(0, \alpha, c)$ is bounded. On the other hand, if $U_N^\psi(\alpha) \to \infty$ where $\psi(r) = r^{-(d/q+\varepsilon)}$ then α is called **very well approximable (VWA)**.

The following result is known as Khinchine–Groshev Theorem. Almost sure is considered relative to Lebesgue measure on the space of matrices $\alpha \in \mathbb{T}^{dq}$.

THEOREM 50. [**64, 47, 31, 116, 14, 7**] *(a) If* $\sum_{\mathbb{Z}^q} |\psi^d(|n|) < \infty$ *then* U_N^ψ *is bounded almost surely.*

(b) If $\sum_{\mathbb{Z}^q} \psi^d(|n|) = +\infty$ *and either* ψ *is decreasing or* $dq > 1$ *then* $\lim_{N\to\infty} U_N^\psi(\alpha) = +\infty$ *almost surely.*

(c) For $d = q = 1$ *there exists* ψ *such that* $\sum_{n\in Z} \psi(|n|) = +\infty$ *but* U_N^ψ *is bounded almost surely.*

(d) If ψ *is decreasing and* $\sum_{n\in Z^d} \psi^d(|n|) = +\infty$ *then* $V_N^\psi(\alpha, x) \to \infty$ *almost surely.*

In particular, both badly approximable and very well approximable αs have zero measure.

When the number of hits is infinite, it is natural to consider the question of the sBC property.

THEOREM 51. [**116**] *(a) For almost all* α

$$U_N^\psi(\alpha) = \mathbb{E}(U_N^\psi) + O\left(\sqrt{\Gamma(N) \ln^3 \Gamma(N)}\right)$$

where

$$\Gamma(N) = \sum_{|n|\leq N} \psi(|n|)^d D(gcd(n_1 \dots n_q))$$

and D *denotes the number of divisors.*

(b) $\Gamma(N) \leq C\mathbb{E}(U_N^\psi)$ *if either* $q > 3$ *or* $q = 2$ *and* $n\psi^2(n)$ *is decreasing.*

(c) If $q = 1$ *and* $\psi(n)$ *is decreasing then for each* δ

$$U_N^\psi(\alpha) = \mathbb{E}(U_N^\psi) + O\left(\sqrt{\tilde{\Gamma}(N)\mathbb{E}(U_N^\psi)} \ln^{2+\delta}(\mathbb{E}(U_N^\psi))\right)$$

where

$$\tilde{\Gamma}(N) = \sum_{n=1}^N \frac{\psi(n)}{n}.$$

QUESTION 51. *Does a similar formula as that of Theorem 51 hold for* V^ψ*?*

Some partial results are obtained in [**117**].

It follows from the same arguments as the proof of Theorem 24 sketched in Section 8.3 that the sBC property holds for $\psi(r) = r^{-(d/q)}$ for almost every (α, x), that is

$$\lim_{N\to\infty} \frac{V_N(\alpha, x)}{\mathbb{E}(V_N(\alpha, x))} = 1.$$

For badly approximable α we have the following.

THEOREM 52. [**90**] *Let* x *be uniformly distributed on* $\mathbb{T}^d$*. If* α *is badly approximable, there exists a constant* K *such that all limit points of* $\frac{V_N - \mathbb{E}(V_N)}{\sqrt{\ln N}}$ *are normal random variables with zero mean and variance* σ^2 *where* $0 \leq \sigma^2 \leq K$*.*

QUESTION 52. *(a) Show that there exist a constant* $\bar{\sigma}^2 > 0$ *that for almost all* α $\dfrac{V_N - \mathbb{E}(V_N)}{\sqrt{\ln N}}$ *converges to* $\mathfrak{N}(\bar{\sigma}^2)$*.*

(b) Does there exist α *such that* $\liminf_{N\to\infty} \dfrac{V_N - \mathbb{E}(V_N)}{\sqrt{\ln N}} = 0$ *(that is,* $\sqrt{\ln N}$ *is not a correct normalization for such* α*)?*

For random α we have the following.

THEOREM 53. *(*[**30**]*) There exists* σ *such that If* $\alpha_1, \dots, \alpha_r$ *and* $x_1, \dots, x_d$ *are randomly distributed on* $\mathbb{T}^{dr+d}$ *then* $\frac{V_N - \mathbb{E}(V_N)}{\sqrt{\ln N}}$ *converges in distribution to a normal random variables with zero mean and variance* σ^2*. A similar convergence holds if* $d + r > 2$ *and* $(x_1, \dots, x_r) = (0, \dots, 0)$ *and only the* α_i*'s are random.*

Still there are many open questions. We provide several examples.

QUESTION 53. *Extend Theorems 10 and 11 to the case* $q > 1$*.*

We note that in the case of Theorem 11, even the case $d = 1$ seems quite difficult. One can attack this question using the method of [**29**] but it runs into the problem of lack of parameters described after Question 24.

QUESTION 54. *Let* $l, \hat{l} : \mathbb{R}^d \to \mathbb{R}$*, be linear forms with random coefficients,* $Q : \mathbb{R}^d \to \mathbb{R}$ *be a positive definite quadratic form. Investigate limit theorems, after adequate renormalization, for the number of solutions to*

(a) $\{l(n)\}Q(n) \leq c, |n| \leq N$*;*

(b) $\{l(n)\}|\hat{l}(n)| \leq c, |n| \leq N$;
(c) $|l(n)Q(n)| \leq c, |n| \leq N$;
(d) $|l(n)\hat{l}(n)| < c, |n| \leq N$.

While (a) and (b) have obvious interpretation as shrinking target problems for toral translations, such interpretation for (c) and (d) is less straightforward. Consider for example (c). Let $l(n) = \sum_{j=1}^{q} \alpha_j n_j$. Dividing the distribution of α into thin slices we may assume that α_d is almost constant. If $\alpha_d \approx a$ then we can compare our problem with $|\left(\sum_{j=1}^{q-1} \tilde{\alpha}_j n_j\right) + n_q|Q(n) < \tilde{c}$ where $\tilde{\alpha}_j = \alpha_j/a, \tilde{c} = c/a$. Since $|l(n)|$ should be small we must have $|\left(\sum_{j=1}^{q-1} \tilde{\alpha}_j n_j\right) + n_q| = \{\sum_{j=1}^{q} \tilde{\alpha}_j\}$ in which case $Q(n_1, \dots, n_{q-1}, n_q)$ is well approximated by

$$Q(n_1, \dots, n_{q-1}, -\sum_{j=1}^{q-1} \tilde{\alpha}_j n_j)$$

so we have a shrinking target problem in lower dimensions. In fact as we saw in Section 7 typically the proof proceeds in the opposite direction by getting rid of fractional part at the expense of increasing dimension since problems (c) and (d) have more symmetry and so should be easier to analyze.

We note that part (d) deals with degenerate quadratic form. The case of non-degenerate forms is discussed in [**32**, Sections 5 and 6].

11.2. Cut-and-project sets. Cut-and-project sets are used in physics literature to model quasicrystals. To define them we need the following data: a lattice in $\mathbb{R}^d$, a decomposition $\mathbb{R}^d = E_1 \oplus E_2$ and a compact set (a window) $\mathcal{W} \subset E_2$. Let P_1 and P_2 be the projections to E_1 and E_2 respectively. The cut-and-project set is defined by

$$\mathcal{P} = \{P_1(e), e \in L \text{ and } P_2(e) \in \mathcal{W}\}.$$

We suppose in the following discussion that

$$\overline{E_1 + L} = \mathbb{R}^d \text{ and } L \cap E_2 = \emptyset.$$

Then $\mathcal{P}$ is a discrete subset of E_1 sharing many properties of lattices but having a more complicated structure. Note that the limiting distributions in Theorems 16 and 48 are described in terms of cut-and-project sets. We refer the reader to [**93**, Sections 16 and 17] for more discussion of cut-and-project set. Here we only mention the fact that such sets have asymptotic density. Let $\mathcal{P}_R = \{t \in \mathcal{P} : |t| \leq R\}$.

THEOREM 54. *Suppose that $\mathcal{W}$ is an open subset of E_2 with a piecewise smooth boundary. Then*

$$\lim_{R\to\infty} \frac{\mathrm{Card}(\mathcal{P}_R)}{\mathrm{Vol}(B(0,R))} = \frac{\mathrm{Vol}(\mathcal{W})}{\mathrm{covol}(L)} \frac{\mathrm{Vol}_{\mathbb{R}^d}}{\mathrm{Vol}_{E_1}\mathrm{Vol}_{E_2}}.$$

PROOF. (Following [**52**]). Note that $t \in \mathcal{P}$ iff there exists $e \in L$ such that $-t+e \in \mathcal{W}$, that is $-t \in \mathcal{W}$ mod L. Consider the action of E_1 on $\mathbb{R}^d/L$ given by $T^t(x) = x+t$. Then $\mathcal{P}_R$ counts the number of intersections of the orbit of the origin of size R with $\mathcal{W}$. Pick a small δ and let $\mathcal{W}_\delta = \{\mathcal{W}+t, |t| \leq \delta\}$. Then $\mathcal{W}_\delta$ is a subset of $\mathbb{R}^d/L$ and for small δ

$$\mathrm{Vol}(\mathcal{W}_\delta) = \mathrm{Vol}(\mathcal{W})\mathrm{Vol}(B(0,\delta)) \frac{\mathrm{Vol}_{\mathbb{R}^d}}{\mathrm{Vol}_{E_1}\mathrm{Vol}_{E_2}}. \tag{41}$$

Next,

$$\int_{|t|<R} \chi_{\mathcal{W}_\delta}(T^t 0)dt = \mathrm{Vol}(B(0,\delta))\mathrm{Card}(\mathcal{P}_R) + O(R^{q-1}) \tag{42}$$

where $q = \dim(E_1)$ and the second term represents boundary contribution. On the other hand by unique ergodicity of T^t

$$\int_{|t|<R} \chi_{\mathcal{W}_\delta}(T^t 0)dt = \mathrm{Vol}(B(0,R)) \frac{\mathrm{Vol}(\mathcal{W}_\delta)}{\mathrm{covol}(L)} + o(R^q). \tag{43}$$

Combining (41), (42) and (43) we get the result. □

QUESTION 55. *Describe the error term in the asymptotics of Theorem 54.*

If $q=1$ then the error term in (42) is negligible and so (42) can be used to describe the deviations (see [**28**] for the case where $\mathcal{W}$ is convex). If $q>1$ more work is needed to control both the LHS and the RHS of (42).

While methods of [**28, 29**] deal with the case where $\dim(E_1)$ is as small as possible, the most classical case is the opposite one when $\dim(E_2)$ as large as possible, that is, studying lattice points in large regions. Here we can not attempt to survey this enormous topic, so we refer the reader to the specialized literature on the subject ([**56, 57, 77**]. We just mention that the limit theorems similar in spirit to the results discussed in this paper are obtained in [**49, 16, 17, 104**]. More generally, instead of considering large balls one can count the number of lattice points in RD where D is a fixed regular set. As in Section 2 the order of the error term is sensitive to the geometry of D (see e.g. [**15, 78, 86, 96, 105, 109, 110, 121, 122**] and references wherein). In fact, one can also consider the varying shapes RD_R which includes both the Poisson regime where $\mathrm{Vol}(RD_R)$ does not grow (see [**18**] and references wherein) and the intermediate regime where $\mathrm{Vol}(RD_R)$ grows but at the rate slower than R^d (see [**54, 127**]).

This motivates the following question

Question 56. *Extend the results of the above mentioned papers to cut-and-project sets.*

References

[1] Aaronson J. *An Introduction to Infinite Ergodic Theory*, Math. Surveys and Monographs **50**, Amer. Math. Soc. 1997.

[2] Aaronson J., Keane M. *The visits to zero of some deterministic random walks,* Proc. London Math. Soc. **44** (1982) 535–553.

[3] Arnold V. I. *Topological and ergodic properties of closed* 1*-forms with incommensurable periods,* Funkts. Anal. Prilozhen. **25** (1991) 1–12.

[4] Athreya J. *Logarithm laws and shrinking target properties*, Proceedings–Math. Sciences **119** (2009) 541–557.

[5] Atkinson G. *Recurrence of co-cycles and random walks,* J. London Math. Soc. **13** (1976) 486–488.

[6] Avila A., Dolgopyat D., Duriev E., Sarig O. *The visits to zero of a random walk driven by an irrational rotation,* to appear in Israel Math. J.

[7] Badziahin D., Beresnevich V., Velani S. *Inhomogeneous theory of dual Diophantine approximation on manifolds,* Adv. Math. **232** (2013) 1–35.

[8] Beck J. *Probabilistic Diophantine approximation. I. Kronecker sequences*, Ann. of Math. **140** (1994) 449–502.

[9] Beck J., *Diophantine approximation and quadratic fields* in *Number theory (Eger, 1996),* (edited by K. Gyory, A. Petho, V. T. Sos) de Gruyter, Berlin, 1998, pp. 55–93.

[10] Beck J., *From probabilistic Diophantine approximation to quadratic fields,* Random and quasi-random point sets, 148, Springer Lecture Notes in Statist., **138** (1997) 1–48.

[11] Beck J., *Randomness of the square root of 2 and the giant leap,* Period. Math. Hungar. Part I: **60** (2010) 137–242; Part II: **62** (2011) 127–246.

[12] Bekka M. B., Mayer M. *Ergodic theory and topological dynamics of group actions on homogeneous spaces,*London Math. Soc. Lect. Note Ser. **269** Cambridge University Press, Cambridge, 2000. x+200 pp.

[13] V. Beresnevich, V. Bernik, D. Kleinbock, G. A. Margulis, *Metric Diophantine approximation: The Khintchine–Groshev theorem for non-degenerate manifolds*, Moscow Math. J. **2** (2002), 203–225.

[14] Beresnevich V., Velani S. *Classical metric Diophantine approximation revisited: the Khintchine-Groshev theorem,* IMRN (2010) no. 1, 69–86.

[15] Bleher P. *On the distribution of the number of lattice points inside a family of convex ovals,* Duke Math. J. **67** (1992) 461–481.

[16] Bleher P., Bourgain, J. *Distribution of the error term for the number of lattice points inside a shifted ball,* Progr. Math. **138** (1996) 141–153 Birkhauser Boston.

[17] Bleher P. M., Cheng Z., Dyson F. J., Lebowitz J. L. *Distribution of the error term for the number of lattice points inside a shifted circle,* Comm. Math. Phys. **154** (1993) 433–469.
[18] Bleher P. M., Lebowitz J. L. *Variance of number of lattice points in random narrow elliptic strip,* Ann. Inst. H. Poincare Probab. Statist. **31** (1995) 27–58.
[19] J. W. S. Cassels, *An introduction to Diophantine approximation*, Cambridge Tracts in Math., vol. 45, Cambridge Univ. Press, Cambridge, 1957.
[20] Chaika J, Constantine D. *Quantitative Shrinking Target Properties for rotations, interval exchanges and billiards in rational polygons*, arXiv preprint 1201.0941.
[21] Chernov N., Kleinbock D. *Dynamical Borel-Cantelli lemmas for Gibbs measures*, Israel J. Math. **122** (2001) 1–27.
[22] Chevallier N., Conze J.–P. *Examples of recurrent or transient stationary walks in $\mathbb{R}^d$ over a rotation of $\mathbb{T}^2$*, Contemp. Math **485** (2009) 71–84.
[23] Conze J.-P., *Recurrence, ergodicity and invariant measures. for cocycles over a rotation*, Contemp. Math. **485** (2009) 45–70.
[24] Conze J.-P., Keane, M. *Ergodicite d'un flot cylindrique,* Seminaire de Prob.-I, Univ. Rennes, Rennes, 1976, Exp. No. 5, 7 pp.
[25] Conze J.-P., Piekniewska A., *On multiple ergodicity of affine cocycles over irrational rotations*, arXiv:1209.3798
[26] de Simoi J., Liverani C. *The martingale approach after Varadhan and Dolgopyat,* this volume.
[27] Dolgopyat D. *Limit theorems for partially hyperbolic systems,* Trans. AMS **356** (2004) 1637–1689.
[28] Dolgopyat D., Fayad B. *Deviations of ergodic sums for toral translation-I: Convex bodies,* GAFA **24** (2014) 85–115.
[29] Dolgopyat D., Fayad B. *Deviations of ergodic sums for toral translations-II: Boxes,* arXiv preprint 1211.4323.
[30] Dolgopyat D., Fayad B., Vinogradov I., *Central limit theorems for simultaneous Diophantine approximations,* in preparation.
[31] Duffin R. J., Schaeffer A.C. *Khintchine's problem in metric Diophantine approximation,* Duke Math. J. **8** (1941) 243–255.
[32] Eskin A. *Unipotent flows and applications. Homogeneous flows, moduli spaces and arithmetic,* Clay Math. Proc. **10** (2010) 71–129, AMS, Providence, RI.
[33] Eskin A., McMullen C. *Mixing, counting, and equidistribution in Lie groups,* Duke Math. J. **71** (1993) 181–209.
[34] Fayad B. *Polynomial decay of correlations for a class of smooth flows on the two torus*, Bull. SMF **129** (2001) 487–503.
[35] Fayad B. *Weak mixing for reparametrized linear flows on the torus*, Erg. Th., Dynam. Sys. **22** (2002) 187–201.
[36] Fayad B. *Analytic mixing reparametrizations of irrational flows*, Erg. Th., Dynam. Sys. **22** (2002) 437–468.
[37] Fayad B. *Skew products over translations on $\mathbb{T}^d$, $d \geq 2$*, Proc. Amer. Math. Soc. **130** (2002) 103–109.
[38] Fayad B. *Mixing in the absence of the shrinking target property,* Bull. London Math. Soc. **38** (2006) 829–838.
[39] Fayad B., Lemańczyk M. *On the ergodicity of cylindrical transformations given by the logarithm,* Mosc. Math. J. **6** (2006) 657–672, 771–772.

[40] Feller W. *An introduction to probability theory and its applications,* Vol. **II** 2d ed. John Wiley & Sons, New York-London-Sydney, 1971, 669 pp.

[41] Flaminio L., Forni G. *Equidistribution of nilflows and applications to theta sums,* Erg. Th. Dyn. Sys. **26** (2006) 409–433.

[42] Frączek K. *Some examples of cocycles with simple continuous singular spectrum,* Studia Math. **146** (2001) 1–13.

[43] Frączek K., Lemańczyk M. *On symmetric logarithm and some old examples in smooth ergodic theory*, Fundamenta Math. **180** (2003), 241–255.

[44] Fuchs M. *On a problem of W. J. LeVeque concerning metric Diophantine approximation,* Trans. AMS **355** (2003) 1787–1801.

[45] Gabriel P., Lemańczyk M., Liardet P. *Ensemble d'invariants pour les produits croiss de Anzai,* Mem. SMF **47** (1991) 102 pp.

[46] Gouezel S. *Limit theorems in dynamical systems using the spectral method,* this volume.

[47] Groshev A. *A theorem on a system of linear forms,* Doklady Akademii Nauk SSSR **19** (1938) 151–152.

[48] Hardy G. H., Littlewood J. E. *Some problems of diophantine approximation: a series of cosecants,* Bull. Calcutta Math. Soc. **20** (1930) 251–266.

[49] Heath-Brown D. R. *The distribution and moments of the error term in the Dirichlet divisor problem,* Acta Arith. **60** (1992) 389–415.

[50] Herz C. S. *Fourier transforms related to convex sets,* Ann. Math. **75** (1962) 81–92.

[51] Hill B., Velani S. *The ergodic theory of shrinking targets*, Invent. Math. **119** (1995) 175–198.

[52] Hof A. *Uniform distribution and the projection method,* Fields Inst. Monogr. **10** (1998) 201–206, AMS, Providence, RI.

[53] Hooper P., Hubert P. Weiss B. *Dynamics on the infinite staircase,* Discr., Cont. Dyn. Syst. A **33** (2013) 4341–4347.

[54] Hughes C. P., Rudnick Z. *On the distribution of lattice points in thin annuli,* Int. Math. Res. Not. **13** (2004) 637–658.

[55] Huveneers F. *Subdiffusive behavior generated by irrational rotations,* Erg Th. Dyn. Sys. **29** (2009) 1217–1233.

[56] Huxley M. N. *Area, lattice points, and exponential sums,* LMS Monographs **13** (1996) 494 pp., Oxford Univ. Press, New York, 1996.

[57] Ivic A., Kratzel E., Kuhleitner M., Nowak W. G. *Lattice points in large regions and related arithmetic functions: recent developments in a very classic topic,* Schriften der Wissenschaftlichen Gesellschaft an der Johann Wolfgang Goethe-Universitt Frankfurt am Main **20** (2006) 89–128.

[58] Iwanik A. *Anzai skew products with Lebesgue component of infinite multiplicity,* Bull. London Math. Soc. **29** (1997) 195–199.

[59] Iwanik A., Lemańczyk M., Rudolph D. *Absolutely continuous cocycles over irrational rotations,* Israel J. Math. **83** (1993) 73–95.

[60] A.B. Katok, *Combinatorial constructions in Ergodic Theory and Dynamics*, University Lecture Series, **30**, 2003.

[61] Kesten H. *Uniform distribution* $\bmod 1$*,* Ann. of Math. **71** (1960) 445–471.

[62] Kesten H. *Uniform distribution* mod1*, II,* Acta Arith. **7** (1961/1962) 355–380.

[63] Keynes H. B., Newton D. *The structure of ergodic measures for compact group extensions,* Israel J. Math. **18** (1974) 363–389.

[64] A. Khintchine *Ein Satz ber Kettenbrche, mit arithmetischen Anwendungen,* Mathematische Zeitschrift **18** (1923) 289–306.
[65] Kingman J. F. C. Poisson processes, *Oxford Studies in Probability* **3** (1993) viii+104 pp, Oxford University Press, New York.
[66] Kleinbock D. Y., Margulis G. A. *Bounded orbits of nonquasiunipotent flows on homogeneous spaces,* AMS Transl. **171** (1996) 141–172.
[67] Kleinbock D. Y., Margulis G. A. *Flows on homogeneous spaces and Diophantine approximation on manifolds,* Ann. of Math. **148** (1998) 339–360.
[68] Knill O. *Selfsimilarity in the Birkhoff sum of the cotangent function,* arXiv preprint 1206.5458.
[69] Knill O., Tangerman F. *Self-similarity and growth in Birkhoff sums for the golden rotation,* Nonlinearity **24** (2011) 3115–3127.
[70] Kocergin A. V. *The absence of mixing in special flows over a rotation of the circle and in flows on a two-dimensional torus*, Dokl. Akad. Nauk SSSR, **205** (1972) 515–518.
[71] Kocergin A. V. *Mixing in special flows over a rearrangement of segments and in smooth flows on surfaces,* Mat. Sbornik **96** (1975) 471–502, 504.
[72] Kocergin A. V. *Nondegenerate saddles, and the absence of mixing,* Mat. Zametki **19** (1976) 453–468.
[73] Kochergin A. V. *Nondegenerate fixed points and mixing in flows on a two-dimensional torus* I: Sb. Math. **194** (2003) 1195–1224; II: Sb. Math. **195** (2004) 317–346.
[74] Kocergin A. V. *Causes of stretching of Birkhoff sums and mixing in flows on surfaces,* MSRI Publ. **54** (2007) 129–144, Cambridge Univ. Press, Cambridge.
[75] J. F. Koksma, *Diophantische Approximationen*, Ergebnisse der Mathematik, Band IV, Heft 4, Berlin, 1936.
[76] A. N. Kolmogorov, *On dynamical systems with an integral invariant on the torus*, Doklady Akad. Nauk SSSR, **93** (1953) 763–766.
[77] Kratzel E. *Lattice points* Math., Appl. **33** (1988) 320 pp., Kluwer Dordrecht.
[78] Kratzel E., Nowak W. G. *The lattice discrepancy of certain three-dimensional bodies,* Monatsh. Math. **163** (2011) 149–174.
[79] L. Kuipers, H. Niederreiter *Uniform Distribution of Sequences*, Dover Books on Math., 2006, 416pp.
[80] Kurzweil, J. *On the metric theory of inhomogeneous diophantine approximations,* Studia Math. **15** (1955) 84–112.
[81] Le Borgne S. *Principes d'invariance pour les flots diagonaux sur SL(d,R)/SL(d,Z),* Ann. Inst. H. Poincar Prob. Stat. **38** (2002) 581–612.
[82] Ledrappier F., Sarig O. *Fluctuations of ergodic sums for horocycle flows on* $\mathbb{Z}^d$*-covers of finite volume surfaces,* Disc. Cont. Dyn. Syst. **22** (2008) 247–325.
[83] Lemańczyk M., *Spectral Theory of Dynamical Systems* Encyclopedia of Complexity and Systems Science, Springer, (2009).
[84] Lemańczyk M. *Sur l'absence de mlange pour des flots spciaux au-dessus d'une rotation irrationnelle,* Colloq. Math. **84/85** (2000) 29–41.
[85] Lemańczyk M., Parreau F., Volný D. *Ergodic properties of real cocycles and pseudo-homogenous Banach spaces*, Trans. AMS **348** (1996) 4919–4938.
[86] Levin M. B., *On the Gaussian limiting distribution of lattice points in a parallelepiped,* arXiv preprint 1307.2076.

[87] Liu L., *Limit laws of quasi-periodic ergodic sums for observables with singularities*, in preparation.
[88] Marklof J. *The n-point correlations between values of a linear form,* Erg. Th., Dynam. Sys. **20** (2000) 1127–1172.
[89] Marklof J. *Pair correlation densities of inhomogeneous quadratic forms,* Ann. Math. **158** (2003) 419–471.
[90] Marklof J. *An application of Bakhtin Central Limit Theorem for Hyperbolic sequences to Beck's probabilistic Diophantine approximation,* preprint (2003).
[91] Marklof J. *Energy level statistics, lattice point problems and almost modular functions,* in *Frontiers in Number Theory, Physics and Geometry,* v. 1, pp. 163–181 (Les Houches, 9-21 March 2003), Springer.
[92] Marklof J. *Distribution modulo one and Ratner's theorem,* in *Equidistribution in number theory, an introduction,* NATO Sci. Ser. II Math. Phys. Chem., **237** (2007) 217–244, Springer, Dordrecht.
[93] Marklof J. *Kinetic limits of dynamical systems,* this volume.
[94] Marklof J., Strömbergsson A. *The distribution of free path lengths in the periodic Lorentz gas and related lattice point problems,* Annals of Mathematics **172** (2010) 1949–2033.
[95] Marklof J., Strömbergsson A. *Free path lengths in quasicrystals,* Comm. Math. Phys. **330** (2014) 723–755.
[96] Muller W., Nowak W.G. *On lattice points in planar domains,* Math. J. Okayama Univ. **27** (1985) 173–184.
[97] Oren I. *Ergodicity of cylinder flows arising from irregularities of distribution,* Israel J. Math. **44** (1983) 127–138.
[98] Ornstein D. S., Rudolph D. J., Weiss B. *Equivalence of measure preserving transformations,* Mem. Amer. Math. Soc. **37** (1982) no. 262, xii+116 pp.
[99] Ostrowski A. *Bemerkungen zur Theorie der Diophantischen Approximationen,* Abhandlungen aus dem Math. Seminar Univ. Hamburg **1** (1922) 77–98.
[100] W. Parry, *Introduction to ergodic theory*, *Topics in ergodic theory,* Cambridge Tracts in Math. **75** Cambridge Univ. Press, Cambridge-New York, 1981. x+110 pp.
[101] Parry W., Pollicott M. *Zeta functions and the periodic orbit structure of hyperbolic dynamics,* Asterisque **187–188** (1990) 268 pp.
[102] Pask D. A. *Skew products over the irrational rotation,* Israel J. Math. **69** (1990) 65–74.
[103] Peres Yu., Ralston D. *Heaviness in toral rotations,* Israel J. Math. **189** (2012), 337–346.
[104] Peter M. *Almost periodicity and the remainder in the ellipsoid problem,* Michigan Math. J. **49** (2001) 331–351.
[105] Peter M. *Lattice points in convex bodies with planar points on the boundary,* Monatsh. Math. **135** (2002) 37–57.
[106] K. Peterson, *On a series of cosecants related to a problem in ergodic theory,* Compositio Mathematica, **26** (1973), 313–317.
[107] W. Philipp, *Mixing Sequences of Random Variables and Probabilistic Number Theory*, AMS Memoirs **114**, 1971.
[108] W. Philipp, *Some metrical theorems in number theory,* Pacic J. Math. **20** (1967), p. 109–127.

[109] Randol B. *On the Fourier transform of the indicator function of a planar set,* Trans. Amer. Math. Soc. **139** (1969) 271–278.
[110] Randol B. *On the asymptotic behavior of the Fourier transform of the indicator function of a convex set,* Trans. AMS **139** (1969) 279–285.
[111] Ratner M. *On Raghunathan's measure conjecture,* Ann. of Math. **134** (1991) 545–607.
[112] Ratner M. *Raghunathan's topological conjecture and distributions of unipotent flows,* Duke Math. J. **63** (1991) 235–280.
[113] Roth K. F. *On irregularities of distribution,* Mathematika **1** (1954) 73–79.
[114] Samur J., *A functional central limit theorem in Diophantine approximation*, Proc. Amer. Math. Soc. **111** (1991) 901–911.
[115] Schmidt K. *Cocycles of Ergodic Transformation Groups*, Lect. Notes in Math. Vol. 1, Mac Millan Co. India, 1977.
[116] Schmidt W. M. *A metrical theorem in diophantine approximation,* Canadian J. Math. **12** (1960) 619–631.
[117] Schmidt W. *Metrical theorems on fractional parts of sequences*, Trans. Amer. Math. Soc. **110** (1964) 493–518.
[118] Shah N. A. *Limit distributions of expanding translates of certain orbits on homogeneous spaces,* Proc. Indian Acad. Sci. Math. Sci. **106** (1996) 105–125.
[119] Sinai, Ya. G., Khanin K. M. *Mixing of some classes of special flows over rotations of the circle,* Funct. Anal. Appl. **26** (1992) 155–169.
[120] Sinai Ya. G., Ulcigrai C. *A limit theorem for Birkhoff sums of non-integrable functions over rotations,* Contemp. Math. **469** (2008) 317–340.
[121] Skriganov M. M. *Ergodic theory on SL(n), Diophantine approximations and anomalies in the lattice point problem,* Invent. Math. **132** (1998) 1–72.
[122] Skriganov M. M., Starkov A. N. *On logarithmically small errors in the lattice point problem,* Erg. Th. Dynam. Sys. **20** (2000) 1469–1476.
[123] Stein E. M., Weiss G. *Introduction to Fourier analysis on Euclidean spaces,* Princeton Math. Ser. **32** (1971) 297pp, Princeton Univ. Press, Princeton, N.J.
[124] Tseng J. *On circle rotations and the shrinking target properties,* Discrete Contin. Dyn. Syst. **20** (2008) 1111–1122.
[125] Vinogradov I. *Limiting distribution of visits of sereval rotations to shrinking intervals,* arXiv preprint 1005.1622
[126] Von Neumann J., *Zur Operatorenmethode in der klassischen Mechanik*, Ann. of Math. (2) **33** (1932), 587–642.
[127] Wigman I. *The distribution of lattice points in elliptic annuli,* Q. J. Math. **57** (2006) 395–423, **58** (2007) 279–280.
[128] Wysokinska M. *A class of real cocycles over an irrational rotation for which Rokhlin cocycle extensions have Lebesgue component in the spectrum,* Topol. Methods Nonlinear Anal. **24** (2004) 387–407.
[129] Yoccoz J.-C. *Petits diviseurs en dimension 1,* Asterisque **231** (1995).

Dmitry Dolgopyat Department of Mathematics University of Maryland College Park MD 20814 USA
E-mail address: dmitry@math.umd.edu

Bassam Fayad IMJ-PRG, CNRS, Paris, France
E-mail address: bassam.fayad@imj-prg.fr