Ergodic Deviations of Degenerate Multidimensional Actions - Convex Bodies

Hao Wu

Abstract. We prove that the ergodic deviation of a degenerate $\mathbb{Z}^2$-action on the torus $\mathbb{T}^2$ relative to a symmetric, strictly convex body can be decomposed into two parts, and that each part admits a limit distribution after choosing a suitable normalizer. Specifically, the first part is similar to an ergodic sum of smooth observables after being normalized by $N$, and the second part is similar to the case of a random toral translation, i.e., the $\mathbb{Z}$-action, but with a normalizer of $N^{\frac{1}{2}}$. The key difference is that we employ the product flow on the product space of $\mathbb{Z}^2$ lattices for the multidimensional action.
## Contents

1 INTRODUCTION  

2 LIMIT DISTRIBUTIONS  
   2.1 Limit Distribution for the case $\bar{d} = 1$.  
   2.2 Limit Distribution for the case $\bar{d} = 2$.  

3 FOURIER SERIES AND PROOF OF PROPOSITION 2.1  
   3.1 Fourier series for convex bodies.  
   3.2 Proof for the limit distribution when $\bar{d} = 1$.  

4 NON-RESONANT TERMS.  

5 GEOMETRY OF THE SPACE OF LATTICES.  

6 OSCILLATING TERMS.  

References
1 INTRODUCTION

In a $d$-dimensional torus, given a translation vector $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}^d$, we can consider the dynamical system $(\mathbb{T}^d, T_\alpha, \mu)$, where $\mu$ is the Haar measure on $\mathbb{T}^d$, and $T_\alpha$ is the translation from $\mathbb{T}^d \to \mathbb{T}^d$ defined by $T(x) = x + \alpha$, in the sense of modulo 1 for each coordinate. In this dynamical system, ergodic theory states that for every irrational translation, the number of visits inside a measurable set $C$ before time $N$ has a ratio converging to the measure of the set $\text{Vol}(C)$. One object of interest is the discrepancy function defined as the difference of the actual visits before time $N$ and the expected visits $N\text{Vol}(C)$. In dimension 1, Kesten [6, 7] proved that the discrepancy function for the circle rotation relative to an interval converges to a standard Cauchy distribution after being normalized by $\log N$.

There are different ways to extend this result to higher dimensions, one way is to study the random toral translation relative to higher dimensional counterparts of the interval, such as balls (analytic convex bodies) and boxes, both of which were studied by Dolgopyat and Fayad in [3, 4]. They showed that $d$-dimensional boxes behave similarly to 1-dimensional intervals, i.e., the discrepancy function also converges to a Cauchy distribution after normalized by $(\log N)^d$. As for balls, they showed that the discrepancy function converges to a distribution function defined over the product space of infinite tori and the homogeneous space $\text{SL}(d+1, \mathbb{R})/\text{SL}(d+1, \mathbb{Z})$ with the normalizer $N^{(d-1)/2d}$. Their proof consists of a combination of harmonic analysis of the discrepancy’s Fourier series, probability, and an important ingredient is the equidistribution of locally unstable submanifolds over the whole space of unimodular lattices.

In this paper, we follow a similar approach as Dolgopyat and Fayad, but instead of translations, we will consider the degenerate $\mathbb{Z}^2$ action in dimension $d = 2$ (see the definition below), we restrict the set to be strictly convex, symmetric, and analytic bodies $C$. Given a convex body $C$, we denote $C_r$ the rescaled bodies with ratio $r > 0$ by the homothety centered at the origin, where $r < r_0$ so that $C_r$ can fit inside the unit cube of $\mathbb{R}^2$. Let $\alpha = (\alpha_1, \alpha_2) \in \mathbb{T}^2$ be the action vector, we consider the following discrepancy function:

$$D_C(r, x, \alpha; N) = \sum_{0 \leq n_1, n_2 \leq N-1} \chi_{C_r}(x_1 + n_1 \alpha_1, x_2 + n_2 \alpha_2) - N^2 \text{Vol}(C_r)$$ (1.1)

where $\chi_{C_r}$ is the indicator function of the set $C_r$.

We will show that by decomposing the discrepancy function into 2 components, each
component would admit a limit distribution after a suitable normalization, specifically:

\[ D_C(r, x, \alpha; N) = \sum_{d=1}^{2} D_{C,d}(r, x, \alpha; N) \quad (1.2) \]

where \( D_{C,d} \) represents the part of the Fourier series of \( D_C \) with coefficients of \( d \) non-zero coordinate(s), whose definitions will be clearer after we introduce the Fourier series of \( D_C \) in Section 3.

Our main result is the following:

**Theorem 1.** Let \( C \) be a symmetric, strictly convex analytic body that fits inside the unit cube of \( \mathbb{R}^2 \), and \( D_C, D_{C,d} \) defined as in (1.1) and (1.2), there exists a limit distribution for each \( D_{C,d}(r, x, \alpha) \) after a suitable normalization, specifically, we have 2 distinct cases:

(a) For \( \bar{d} = 1 \), assume that \((x, \alpha)\) are uniformly distributed in \( T_2 \times T_2 \), then for every fixed \( r \), there exists a function \( D_{C,1}(x, \alpha, \beta) \) such that as \( N \to \infty \),

\[ \frac{D_{C,1}(r, x, \alpha; N)}{N} \Rightarrow D_{C,1}(x, \alpha, \beta) \]

in distribution, where \((x, \alpha, \beta)\) is uniformly distributed on \( (T_2)^3 \).

(b) For \( \bar{d} = 2 \), assume that \((r, x, \alpha)\) are uniformly distributed in \( X = [a, b] \times T_2 \times T_2 \), and denote \( \lambda \) the normalized Lebesgue measure on \( X \), then there exists a distribution function \( D_{C,2}(z) : \mathbb{R} \to [0, 1] \) such that for any \( b > a > 0 \), we have

\[ \lim _{N \to \infty} \lambda \{ (r, x, \alpha) \in [a, b] \times T_2 \times T_2 \mid \frac{D_{C,2}(r, x, \alpha; N)}{r^{1/2} N^{1/2}} \leq z \} = D_{C,2}(z). \quad (1.3) \]

The explicit forms of \( D_{C,d} \) will be given in Proposition 2.2 of Section 2.

**Remark.** With the same reason as in [3,4], \( r \) needs to be random in part (b) in Theorem 1 in order to help prove independence between variables. While in part (a), the function behaves like a smooth function, and \( r \) does not need to be random.

This paper is organized as follows: Section 2 will present the explicit form of the distribution functions. In Section 3 we will prove the limit distribution of the easier part of the discrepancy function \( D_{C,1} \). Sections 4 to 6 are devoted to the general d-dimensional counterpart of the sum \( D_{C,2} \), we give a detailed description of the sum in terms of short vectors of the lattice spaces, and how the variables become independent as \( N \to \infty \). Section 4 obtains the main part of the sum that contributes to the discrepancy by using harmonic analysis. Section 5 introduces the space of lattices and express the
discrepancy in the language of lattices. Section 6 shows the variables in the expression of Section 5 become independent as $N \to \infty$.

2 LIMIT DISTRIBUTIONS

2.1 Limit Distribution for the case $\bar{d} = 1$.

Proposition 2.1. If $C$ is an analytic symmetric strictly convex body in $\mathbb{R}^2$, then we have

$$D_C(r, x, \alpha, \beta) = B_C(\alpha, \beta) - B_C(\alpha, x),$$

where

$$B_C(\alpha, x) = \sum_{k \neq 0} \frac{a_k(r)}{e^{2\pi i k x} - 1} e^{2\pi i k r},$$

$a_k(r) = 0$ when $k_1 k_2 \neq 0$ and $a_k(r) = \hat{\chi}_C(k)$ when $k_1 k_2 = 0$, where $\hat{\chi}_C(k)$ represent the $k$th Fourier coefficient of $\chi_C$, the specific form of which is shown in (3.1).

2.2 Limit Distribution for the case $\bar{d} = 2$.

Notations. Let $M = SL(2, \mathbb{R})/SL(2, \mathbb{Z})$ denote the space of 2-dimensional unimodular lattices of $\mathbb{R}^2$. $M^2 = \prod_2$ copies $M$. Given $L = (L_1, L_2) \in M^2$ we denote by $e_1(L_i)$ the shortest vector in $L_i$, and $e_2(L_i)$ the shortest vector in $L_i$ among those who have the shortest nonzero projection on the orthocomplement of the line generated by $e_1(L_i)$. Clearly the vectors $e_1(L_i), e_2(L_i)$ are well defined outside a set of Haar measure 0. In fact, these vectors generate the lattice (see [1]). We denote $e(L_i) = (e_1(L_i), e_2(L_i))$.

Let $Z$ be the set of prime vectors $m \in \mathbb{Z}^2$ (i.e. the coordinates are coprime) with positive first nonzero coordinate. For later usage in Section 4 and 5, we define $Z^2 = \{m = (m^1, m^2), m^i \in Z\}$, and let

$$T_2^\infty = (T^2)^2 \times T^2 \times Z^2$$

We denote elements of $T_2^\infty$ by $(\theta, b)$, where $\theta = (\theta^1, \theta^2), \theta^i \in T^2$, and $b = (b_{p, m})_{(p, m) \in \mathbb{Z} \times \mathbb{Z}^2}$. For $m = (m^1, m^2) \in Z^2$ and $L = (L_1, L_2) \in M^2$, we denote by $(X_{m^i, Z_m^i}) = (m^i, e(L_i))$ the vector $m^i e_1(L_i) + m^2 e_2(L_i)$. Given a prime vector $p = (p_1, p_2) \in Z$, we denote $X_{p, m} = (p_1 X_{m^1}, p_2 X_{m^2})$ and $R_{p, m} = \|X_{p, m}\|$ the Euclidean norm of $X_{p, m}$.

Limit distribution. Let $C$ be a strictly convex body with smooth boundary. For each vector $\xi \in S^1$, we denote by $K(\xi)$ the gaussian curvature of $\partial C$ at the unique point
$x(\xi) \in \partial C$ where the unit outer normal vector is $\xi$.

Denote

$$\mathcal{M}_2 = M^2 \times T^\infty_2,$$

and let $\mu$ be the Haar measure on $\mathcal{M}_2$. Define the following function on $\mathcal{M}_2$

$$L_C(L, \theta, b) = \frac{1}{\pi^3} \sum_{|\bar{p}|=1}^{\infty} \sum_{p \in \mathbb{Z}} \sum_{m \in \mathbb{Z}_2} K^{-\frac{1}{2}} \left( \frac{X_{p,m}}{R_{p,m}} \right) \cos \left( 2\pi \bar{p} \left( \sum_{i=1}^{2} \left( p_i (m^i, \theta^i) \right) \right) \right) \sin \left( 2\pi \left( \bar{p} b_{p,m} - \frac{1}{8} \right) \right) \prod_{i=1}^{2} \sin \left( \pi \bar{p}_i Z_{m^i} \right) \prod_{i=1}^{2} (p_i Z_{m^i}). \tag{2.1}$$

The distribution $D_{C,2}$ of Theorem 1 can now be described by the level set of the function above:

**Proposition 2.2.** If $C$ is an analytic, symmetric, strictly convex body in $\mathbb{R}^2$, then for any $z \in \mathbb{R}$ we have

$$D_{C,2}(z) = \mu\{(L, \theta, b) \in \mathcal{M}_2 : L(L, \theta, b) \leq z\}.$$

### 3 FOURIER SERIES AND PROOF OF PROPOSITION 2.1

In this section, we first introduce the Fourier series of the discrepancy function as our main object of study, and then give a proof for the limit distribution of the first part of the discrepancy function.

In following sections, $\epsilon > 0$ is fixed and can be arbitrarily small. The constants $C$ may vary between inequalities but it does not depend on any variables other than the dimension $d$, which is fixed to 2 in our case. Though we only treat the special case $d = 2$ for the sum $D_C$, it will be clear from the proof that higher dimensional counterparts $D_{C,d < \frac{d+1}{2}}$ can be treated in the exact same way as $D_{C,1}$, and higher dimensional counterparts $D_{C,d > \frac{d+1}{2}}$ is similar to the case of $D_{C,2}$. The sum $D_{C,d \frac{d+1}{2}}$ exhibits distinctly different behavior and we conjecture that it is similar to the case of toral translations relative to boxes and admits a limit Cauchy distribution.

#### 3.1 Fourier series for convex bodies.

We introduce the Fourier series for the smooth strictly convex body $C$ by using the asymptotic formula obtained in $[5]$. For each vector $t \in \mathbb{R}^2$, define its maximal projection
on \(\partial C\) by \(P(t) = \sup_{x \in \partial C}(t, x)\), if \(C\) is of class \(C^{2,2}\), then we have the following formula for the Fourier node:

\[
(2\pi i|t|)\hat{\chi}_C(t) = \rho(C, t) - \bar{\rho}(C, -t)
\]

with

\[
\rho(C, t) = |t|^{-\frac{1}{2}}K^{-\frac{1}{2}}(t/|t|)e^{i2\pi(P(t)-\frac{1}{8})} + O(|t|^{-\frac{3}{2}}).
\]

By a change of variable we have \(\hat{\chi}_C(r) = r\hat{\chi}_C(1)\), and by grouping the corresponding positive and negative terms in the Fourier series we get that for a symmetric body:

\[
\chi_C(x) - \text{Vol}(C) = r^{\frac{1}{2}} \sum_{k \in \mathbb{Z}^2-\{0\}} c_k(r) \cos(2\pi(k, x)),
\]

\[
c_k(r) = d_k(r) + O(|k|^{-\frac{2}{3}}),
\]

\[
d_k(r) = \frac{1}{\pi} \frac{g(k, r)}{|k|^{\frac{3}{2}}},
\]

\[
g(k, r) = K^{-\frac{3}{4}}(k/|k|) \sin(2\pi(rP(k)) - \frac{1}{8}).
\]

### 3.2 Proof for the limit distribution when \(\bar{d} = 1\).

We will show that after being normalized by \(N\), \(D_{\mathcal{C}, 1}\), the part of the Fourier series that consists of nodes

\(k = (k_1, k_2) \neq (0, 0)\) and \(k_1k_2 = 0\),

will behave like the ergodic sum of a smooth function.

First, we define

\[
A_{\mathcal{C}, r}(x) = \sum_{k_1 \neq 0} \hat{\chi}_{\mathcal{C}, r}(k_1, 0)e^{i2\pi k_1 x_1} + \sum_{k_2 \neq 0} \hat{\chi}_{\mathcal{C}, r}(0, k_2)e^{i2\pi k_2 x_2} =: \sum_{k \in \mathbb{Z}^2-0} a_k(r)e^{i2\pi(k, x)},
\]

where \(a_k(r) = 0\) when \(k_1k_2 \neq 0\) and \(a_k(r) = \hat{\chi}_{\mathcal{C}, r}(k)\) when \(k_1k_2 = 0\), then \(D_{\mathcal{C}, 1}\) takes the following form:

\[
D_{\mathcal{C}, 1}(r, x, \alpha; N) = N \sum_{n=0}^{N-1} A_{\mathcal{C}, r}(x + n\alpha),
\]

Proposition 2.1 will follow if we could prove the following:
Lemma 3.1. For almost every $\alpha \in \mathbb{T}^2$, the series defined by:

$$B_{C_{r}}(\alpha, x) = \sum_{k \neq 0} \frac{a_k(r)}{e^{2\pi i k(\alpha, x)} - 1} e^{2\pi i (k, x)},$$

is convergent in $L^2(x)$, and we have

$$A_{C_{r}}(x + n\alpha) = B_{C_{r}}(\alpha, x + (n + 1)\alpha) - B_{C_{r}}(\alpha, x + n\alpha).$$

Proof. The identity is obtained by direct calculation. We will focus on the convergence of the series $B_{C_{r}}(\alpha, x)$. Note that

$$\int_{\mathbb{T}^2}|B_{C_{r}}(\alpha, x)|^2 dx \leq C \left( \sum_{k_1 \neq 0} \frac{1}{|k_1|^3 |e^{2\pi i k_1 \alpha_1} - 1|^2} + \sum_{k_2 \neq 0} \frac{1}{|k_2|^3 |e^{2\pi i k_2 \alpha_2} - 1|^2} \right)$$

Therefore it suffices to prove that the series

$$\sum_{k_i \neq 0} \frac{1}{|k_i|^3 \|k_i \alpha_i\|^2}$$

is convergent for almost every $\alpha_i \in \mathbb{T}, i = 1, 2$.

By standard application of Borel-Cantelli Lemma, we have for almost every $\alpha_i \in \mathbb{T}$, every $k_i > 0$ and every $\delta > 0$ we have

$$\|k_i \alpha_i\| \geq \frac{C(\alpha_i, \delta)}{|k_i||\ln |k_i||^{1+\delta}},$$

which gives

$$|\ln \|k_i \alpha_i\|| \leq C \ln |k_i|$$

where for convenience, $\ln 1$ is defined as 1. Therefore by taking $\delta$ small, and let the constant $C(\alpha, \delta)$ vary from line to line,

$$\sum_{k_i \neq 0} \frac{1}{|k_i|^3 \|k_i \alpha_i\|^2} \leq C(\alpha, \delta) \sum_{k_i \neq 0} \frac{(\ln |k_i|)^{1+\delta}}{|k_i|^2 \|k_i \alpha_i\|}$$

$$\leq C(\alpha, \delta) \sum_{k_i \neq 0} \frac{1}{|k_i||(\ln |k_i||)^{2+2\delta} \|k_i \alpha_i\|}$$

$$\leq C(\alpha, \delta) \sum_{k_i \neq 0} \frac{1}{|k_i|(\ln |k_i||)^{1+\delta} \|k_i \alpha_i\| \ln \|k_i \alpha_i\||^{1+\delta}}.$$
Note that the integral
\[ J(k_i) = \int_T \frac{1}{\|k_i\alpha_i\| \ln(\|k_i\alpha_i\|)^{1+\delta}} d\alpha_i \]
is convergent and the value is the same for all \( k_i \), thus for almost every \( \alpha_i \in T \),
\[ \sum_{k_i \neq 0} \frac{1}{|k_i|(\ln |k_i|)^{1+\delta} \|k_i\alpha_i\| \ln \|k_i\alpha_i\|)^{1+\delta}} \]
is also convergent. Then the \( L^2 \) convergence of \( B_{C_r}(\alpha, x) \) follows from the convergence of \( (3.3) \) through \( (3.6) \). \( \square \)

4 NON-RESONANT TERMS.

This section is devoted to highlight the nodes with main contributions in the Fourier series \( D_{C,2} \), the final goal is to arrive at the sum \( (4.1) \) as an equivalent expression for our Fourier series in terms of limit distributions. Throughout Section 4, we will use the formula \( (3.2) \) since we restrict ourselves to the case symmetric shapes.

For \( k = (k_1, k_2) \) and \( \alpha = (\alpha_1, \alpha_2) \), we use the notation \( \{k_i\alpha_i\} := k_i\alpha_i + l_i \) where \( l_i \) is the unique integer such that \(-1/2 < k_i\alpha_i + l_i \leq 1/2\). To evaluate \( D_{C,2} \), we sum up term by term in the Fourier expansion \( (3.2) \) of \( \chi_{C_r} \) for \( n_1, n_2 \), and we will simplify by using the summation formula
\[ \sum_{n=0}^{N-1} \cos(A + nB) = \frac{\cos(A + \frac{N-1}{2}B) \sin(\frac{N}{2}B)}{\sin \frac{B}{2}}. \]

Then the normalized term for the node \( k \) becomes
\[ f(r, x, \alpha; N, k) = c_k(r) \frac{\cos(2\pi(k, x) + \pi(N - 1)(\sum_{i=1}^{2} \{k_i\alpha_i\}) \prod_{i=1}^{2} \sin(\pi N \{k_i\alpha_i\}))}{N^{\frac{1}{2}} \prod_{i=1}^{2} \sin(\pi \{k_i\alpha_i\})} \]
where \( N^{\frac{1}{2}} \) is the normalizer.

Since the sum \( D_{C,2} \) consists of all-non-zero coordinates nodes, it becomes the following:
\[ \Delta(r, x, \alpha; N) = \sum_{k \in \mathbb{Z}^2; \prod_{i=1}^{2} k_i \neq 0} f(r, x, \alpha; N, k) \]

**Step 1.** This step shows that the nodes outside the circle of radius \( N/\epsilon \) have a negligible
combined contribution. Given a set \( S \), for function \( h \) defined on \( (\mathbb{T}^2)^2 \times S \), we denote by \( \|h\|_2 \) the supremum of the \( L^2 \) norms \( \|h(\cdot, s)\| \) over all \( s \in S \). Let

\[
\Delta_1(r, x, \alpha; N) = \sum_{k \in \mathbb{Z}^2: \forall 1 \leq i \leq 2, ~ 0 < |k_i| < N} f(r, x, \alpha; N, k)
\]

**Lemma 4.1.** We have

\[
\|\Delta - \Delta_1\|_2 \leq C\epsilon^{1/2}
\]

**Proof.** Since

\[
\int_{\mathbb{T}} \left( \frac{\sin(\pi N(k_i \alpha_i))}{\sin(\pi(k_i \alpha_i))} \right)^2 \, d\alpha_i = \int_{\mathbb{T}} |\frac{e^{i\pi Nk_i \alpha_i} - e^{-i\pi Nk_i \alpha_i}}{e^{i\pi k_i \alpha_i} - e^{-i\pi k_i \alpha_i}}|^2 \, d\alpha_i
\]

\[
= \int_{\mathbb{T}} \left| \frac{e^{i2\pi Nk_i \alpha_i} - 1}{e^{i2\pi k_i \alpha_i} - 1} \right|^2 \, d\alpha_i = \int_{\mathbb{T}} \left| \sum_{n=0}^{N-1} e^{i2\pi nk_i \alpha_i} \right|^2 \, d\alpha_i,
\]

we have for every \( 1 \leq i \leq d \),

\[
\int_{\mathbb{T}} \left( \frac{\sin(\pi N(k_i \alpha_i))}{\sin(\pi(k_i \alpha_i))} \right)^2 \, d\alpha_i \leq N.
\]

Since in the integral only the square terms have non zero contributions, and \( |d_r(k)| = O(|k|^{-\frac{3}{2}}) \), we get that

\[
\|\Delta - \Delta_1\|_2^2 \leq CN^2 \frac{1}{N} \sum_{|k| \geq \frac{1}{N}} \frac{1}{|k|^3} \leq CN \frac{1}{\epsilon} = C\epsilon.
\]

**Step 2.** We show that, within the range of \( |k| < N/\epsilon \), by taking out a small measure set of \( \alpha \), the divisors admit a lower bound such that \( N^{\frac{1}{3}}|k_i|^{\frac{1}{3}}|\{k_i \alpha_i\}| > \epsilon^{1/2} \), for every \( 1 \leq i \leq 2 \). Therefore we can further restrict our sum in the set of small divisors \( S(N, \alpha) \) (see \( \text{(4.3)} \)).

Let

\[
E_N = \bigcup_{1 \leq |n| \leq \frac{N}{\epsilon}} \left\{ \alpha \in \mathbb{T}^2: \exists 1 \leq i \leq 2, ~ |n|^{\frac{3}{2}} |\{n\alpha_i\}| < \frac{\epsilon^{\frac{1}{3}}}{N^{\frac{1}{3}}} \right\}.
\]
Note that
\[ |E_N| \leq d \sum_{n=1}^{N} \frac{\epsilon^2}{\epsilon N^\frac{1}{4}} \leq d \epsilon^\frac{1}{4}. \]

Outside the 2\epsilon^\frac{1}{4} measure set \( E_N \), we have for \( 0 < |k_i| < N/\epsilon, N^\frac{1}{4}|k_i|^\frac{1}{4}|\{k_i \alpha_i\}| > \epsilon^{1/2} \), for every \( 1 \leq i \leq 2 \). This is how we apply the short vector argument in the next section.

Let \( S(N, \alpha) = \{ k \in \mathbb{Z}^2 : \forall 1 \leq i \leq 2, \ 0 < |k_i| < N/\epsilon, |k_i|^\frac{1}{4}|\{k_i \alpha_i\}| < \frac{1}{\epsilon^2 N^\frac{1}{4}} \}. \)

\[ \Delta_2(r, x, \alpha; N) = \sum_{k \in S(N, \alpha)} f(r, x, \alpha; N, k), \]  

We have

**Lemma 4.2.**
\[ \| \Delta - \Delta_2 \|_{L^2(T^2 \times (T^2 - E_N))} \leq C\epsilon^{1/2} \]

**Proof.** By (4.2) it is sufficient to show that \( \| \Delta_1 - \Delta_2 \|_{L^2(T^2 \times (T^2 - E_N))} \leq C\epsilon \). We have

\[ \| \Delta_1 - \Delta_2 \|_{L^2(T^2 \times (T^2 - E_N))} \leq \frac{C}{N} \sum_{|k| < \frac{N}{\epsilon}} A_k \]

with

\[ A_k = c_k^2 \int_{T^2} \frac{1}{\prod_{i=1}^2 |k_i \alpha_i|} \chi_{\exists 1 \leq i \leq 2, |k_i|^\frac{1}{4}|\{k_i \alpha_i\}| \leq \frac{1}{\epsilon^2 N^\frac{1}{4}}} d\alpha. \]

We have

\[ A_k \leq c_k^2 \sum_{j=1}^{2} A(k, j), \]

where \( A(k, j) \) denote the part when the \( j \)-coordinate violates the condition in \( S(N, \alpha) \):

\[ A(k, j) = \prod_{i \neq j} \sum_{p_i \geq 1} \int_{T} \frac{1}{|\{k_i \alpha_i\}|^2} \chi_{\{p_i \epsilon^{\frac{1}{2}} \leq N^\frac{1}{4} \mid k_i \mid^\frac{1}{4} \{k_i \alpha_i\} \leq (p_i + 1) \epsilon^{\frac{1}{2}}} \left( k_j \alpha_j \right) \}
\] \times \sum_{p_j \geq 1} \int_{T} \frac{1}{|\{k_j \alpha_j\}|^2} \chi_{\{p_j \epsilon^{\frac{1}{4}} \leq N^\frac{1}{4} \mid k_j \mid^\frac{1}{4} \{k_j \alpha_j\} \leq (p_j + 1) \epsilon^{\frac{1}{4}}} \left( k_j \alpha_j \right) \}
\] \[ =: \prod_{i \neq j} \sum_{p_i \geq 1} A(k, i, p_i) \sum_{p_j \geq 1} \bar{A}(k, j, p_j) \]
For $p_i \geq 1$ we define
\[ B(k, i, p_i) = \left\{ \alpha_i \in \mathbb{T} : p_i \epsilon^2 \leq N^\frac{1}{p} \abs{k^i} \abs{\{k_i \alpha_i\}} \leq (p_i + 1) \epsilon^2 \right\}, \]
and for $p_j \geq 1$, define
\[ \bar{B}(k, j, p_j) = \left\{ p_j \epsilon^2 \leq N^\frac{1}{p} \abs{k_j} \abs{\{k_j \alpha_j\}} \leq \frac{(p_j + 1) \epsilon^2}{\epsilon^2} \right\}. \]
Then
\[ |B(k, i, p_i)| \leq \epsilon^2 N^\frac{1}{p} \abs{k^i} \epsilon^2 \]
and
\[ |ar{B}(k, j, p_j)| \leq \frac{1}{\epsilon^2 N^\frac{1}{p} \abs{k_j} \epsilon^2}. \]
Thus
\[ A(k, i, p_i) \leq \frac{\epsilon^2 (N^\frac{1}{p} \abs{k^i})^2}{(\epsilon^2)^2 p_i^2 N^\frac{1}{p} \abs{k^i} \epsilon^2} \leq \frac{N^\frac{1}{p} \abs{k^i} \epsilon^2}{\epsilon^2 p_i^2}, \tag{4.7} \]
similarly,
\[ \bar{A}(k, j, p_j) \leq \frac{(\epsilon^2)^2 (N^\frac{1}{p} \abs{k_j} \epsilon^2)^2}{\epsilon^2 p_j^2 N^\frac{1}{p} \abs{k_j} \epsilon^2} \leq \epsilon^2 N^\frac{1}{p} \abs{k_j} \epsilon^2, \]
By using $c_k = O \left( \frac{1}{\abs{k}^2} \right)$, we obtain
\[ A_k \leq C \frac{1}{\abs{k}^3} \epsilon^2 \prod_{i=1}^{2} \left( N^\frac{1}{p} \abs{k^i} \epsilon^2 \right) \leq C \epsilon^2 \frac{N^\frac{1}{p}}{\abs{k}^2} \]
Summing over $k$, we get
\[ \sum_{|k| \leq \frac{N}{\epsilon}} A_k \leq C \epsilon^2 N^\frac{1}{p} \sum_{|k| \leq \frac{N}{\epsilon}} \frac{1}{\abs{k}^3} \leq C \epsilon N, \]
and the claim follows. \[ \square \]

**Step 3.** In fact, with the bounded range of $\{k_i \alpha_i\}$ in Step 2, we can show that the main contribution of the Fourier series comes from the nodes of coordinates of order $N$. Let
\[ \hat{S}(N, \alpha) = \left\{ k \in \mathbb{Z}^2 : \forall 1 \leq i \leq 2, \quad N \epsilon^3 < \abs{k^i} < \frac{N}{\epsilon}, \quad \abs{k^i} \abs{\{k_i \alpha_i\}} < \frac{1}{\epsilon^2 N^\frac{1}{3}}, \right\}, \]
\[ \Delta_3(r, x, \alpha; N) = \sum_{k \in \hat{S}(N, \alpha)} f(r, x, \alpha; N, k), \tag{4.8} \]
We have
Lemma 4.3.

\[ \| \Delta - \Delta_3 \|_{L^2(T^2 \times (T^2 - E_N))} \leq C \epsilon^{1/2} \]

Proof. By (4.5) it is sufficient to show that \( \| \Delta_3 - \Delta_2 \|_{L^2(T^2 \times (T^2 - E_N))} \leq C \epsilon \). We have

\[ \| \Delta_3 - \Delta_2 \|_{L^2(T^2 \times (T^2 - E_N))} \leq \frac{C}{N} \sum_{|k| < N \epsilon^3} \hat{A}_k \]

with

\[ \hat{A}_k = c_k^2 \prod_{i=1}^{2} \int_{T^2} \frac{1}{\{k_i \alpha_i \}} \chi \left\{ \left| k_i \right|^{3/4} \left| \{k_i \alpha_i \} \right| \geq \frac{\epsilon}{N^{1/4}} \right\} d\alpha. \]

Repeating the argument in the Lemma 4.2 by replacing \( \bar{A}(k, j, p_j) \) in (4.6) with \( A(k, i, p_i) \), and using the inequality (4.7) we obtain

\[ \hat{A}_k \leq C \frac{1}{|k|^{3/2}} \epsilon^{-1/2} \prod_{i=1}^{2} \left( N^{1/2} \left| k_i \right|^{3/4} \right) \leq C \epsilon^{-1/2} N^{1/2} \left| k \right|^{-1/2} \]

Summing over \( |k| \leq N \epsilon^3 \), we get

\[ \sum_{|k| < N \epsilon^3} \hat{A}_k \leq C \epsilon^{-1/2} N^{1/2} \sum_{|k| < N \epsilon^3} \frac{1}{|k|^{1/2}} \leq C \epsilon N, \]

and the claim follows. \( \square \)

**Step 4.** Now the error terms in the Fourier series can be safely removed. Introduce

\[ \hat{f}(r, x; \alpha; N, k) = \frac{d_k(r)}{c_k(r)} f(r, x; \alpha; N, k) \]

and let

\[ \hat{\Delta}(r, x; \alpha; N) = \sum_{k \in S(N, \alpha)} \hat{f}(r, x; \alpha; N, k). \]  

(4.9)

Since \( |c_k - d_k| = O(|k|^{-5/2}) \) and \( \epsilon \) is fixed,

\[ \| \Delta - \hat{\Delta} \|_{L^2(T^2 \times (T^2 - E_N))} \leq \sum_{\epsilon^3 N < |k| < N} \frac{C}{|k|^{d+3}} N^2 \leq O(N^{-1}). \]  

(4.10)

Therefore \( \hat{\Delta} \) and \( \hat{\Delta} \) admit the same limit distribution if there exists one.
Step 5. Observe that when $\epsilon$ is fixed, the sum in (4.9) is limited to large $k_i$ and small $\prod_{i=1}^{2} |\{k_i\alpha_i\}|$. We can replace $\hat{f}$ and $\hat{\Delta}$ by the following

$$g(r, x, \alpha; N, k) = \frac{d_k(r) \cos(2\pi(k, x) + \pi(N - 1)\sum_{i=1}^{2} \{k_i\alpha_i\}) \prod_{i=1}^{2} \sin(\pi(N\{k_i\alpha_i\}))}{\pi^dN^{\frac{1}{2}} \prod_{i=1}^{2} \{k_i\alpha_i\}}.$$  

Then it suffices to prove that

$$\lim_{N \to \infty} \lambda\{(r, x, \alpha) \in [a, b] \times \mathbb{T}^d \times \mathbb{T}^d \mid \Delta'(r, x, \alpha; N) \leq z\} = D(z) \quad (4.11)$$

where

$$\Delta' = \sum_{k \in U(N, \alpha)} g(r, x, \alpha; N, k) \quad (4.12)$$

and $U(N, \alpha)$ is any subset of $\mathbb{Z}^2$ that contains $\hat{S}(N, \alpha)$.

5 GEOMETRY OF THE SPACE OF LATTICES.

5.1. Following [2], Section 2, and [4], Section 4, we show that the set $\hat{S}(N, \alpha)$ corresponds to a set of short vectors in lattices in $M^2 = M \times M$, where the lattice takes the form $L_1 \times L_2$, and $L_i \in M = SL(2, \mathbb{R})/SL(2, \mathbb{Z})$. Then the discrepancy function $\Delta'$ can be seen as a function on the homogeneous space $M^2$.

Let

$$g_T = \begin{pmatrix} e^{-T} & 0 \\ 0 & e^T \end{pmatrix}, \quad \Lambda_{\alpha_i} = \begin{pmatrix} 1 & 0 \\ \alpha_i & 1 \end{pmatrix}.$$  

Consider the product lattice $L(N, \alpha) = L(N, \alpha_1) \times L(N, \alpha_2)$, where $L(N, \alpha_i) = g_{inN\Lambda_{\alpha_i}}\mathbb{Z}^2$.

For each $k = (k_1, k_2) \in \mathbb{Z}^2$, we associate the vectors $k_i = k_i(k_i) = (k_i, l_i)$, where $l_i$ is the unique interger such that $-\frac{1}{2} < k_i\alpha_i + l_i \leq \frac{1}{2}$. We then denote

$$(X_i, Z_i) = (k_i/N, N\{k_i\alpha_i\}) = g_{inN\Lambda_{\alpha_i}}k_i \quad (5.1)$$

We have $k \in \hat{S}(N, \alpha)$ if and only if :

$$\epsilon^3 < |X_i| < \frac{1}{\epsilon}, \quad |X_i|^2|Z_i| < \frac{1}{\epsilon^2} \quad (5.2)$$

Recall the definition of the shortest vectors $\{e_1(N, \alpha_i), e_2(N, \alpha_i)\}$ of $L(N, \alpha_i)$ in Section 2. We will prove a version of Lemma 4.1 in [4] that works for our product lattice space.

Lemma 5.1. For every $\epsilon > 0$ there exists $K(\epsilon) > 0$ such that for $\alpha$ outside $E_N$, each
Proof. From (\ref{eq:1}) norms. For every \(i = 1, 2\), \(\|m^i\| \leq K(\epsilon)\) and
\[
g_{n N} \Lambda, k_i = m_i^1 e_1(N, \alpha_i) + m_i^2 e_2(N, \alpha_i).
\]

Conversely, for \(\epsilon > 0\) fixed and \(N\) large enough, \(\alpha \notin E_N\) implies that for each pair of vectors \((m^1, m^2) \in \mathbb{Z}^2 \times \mathbb{Z}^2\), where \(\|m^i\| \leq K(\epsilon), i = 1, 2\), there exists a unique \(k = (k_1, k_2) \in \mathbb{Z}^2\) such that for \(i = 1, 2\),
\[
g_{n N} \Lambda, k_i = (m^i, e(N, \alpha_i)) = m_i^1 e_1(N, \alpha_i) + m_i^2 e_2(N, \alpha_i).
\]

Denote \(U(N, \alpha, \epsilon)\) the set of \(k = (k_1, k_2) \in \mathbb{Z}^2\) that corresponds to the set of pairs of vectors \(\{(m^1, m^2) \in \mathbb{Z}^2 \times \mathbb{Z}^2 \mid \|m^i\| \leq K(\epsilon), i = 1, 2\}\)

Then for every \(L = L_1 \times L_2\), there exists \(K(L)\), such that if \(m^i \in \mathbb{Z}^2\) satisfies \(\|m^i\| \geq K(L)\), we have \(\|(m^i, e(L_i))\| \geq R(\epsilon)\). Now we show that the choice of \(K(L)\) can be uniform for the set of lattices \(\{L \in \mathbb{M}^2 \mid \prod_{i=1}^{2} L(N, \alpha_i), \alpha \notin E_N\}\). Therefore it is enough to show that the set
\[
\{\prod_{i=1}^{2} L(N, \alpha_i), \alpha \notin E_N\}
\]
is precompact, since we can write the set as \(\prod_{i=1}^{2} \{L(N, \alpha_i), \alpha \notin E_N(i)\}\), we prove that each component is precompact. By the bound \((\ref{eq:2})\) for \(X_i\) and \(Z_i\), when \(\alpha_i \notin E_N(i)\), if \(|X_i| < \epsilon^3\), then \(|Z_i| \geq \frac{1}{\epsilon^2}\). For any \(l \in \mathbb{Z}\), \(|N(k_i \alpha_i + l)| > |N\{\{k_i, \alpha_i\}\} = |Z_i|\), so \(|N(k_i \alpha_i + l)|\) has a lower bound, therefore all vectors in \(L_i\) are longer than some \(\delta\). Then by the Mahler compactness criterion for lattices \([9]\), the set \((\ref{eq:3})\) is precompact.

For the converse, when we fix \(\epsilon\) and let \(N\) be sufficiently large, if \(\|m^i\| \leq K(\epsilon)\), we have that \(\|(m_i, e(N, \alpha_i))\|\) is much smaller than \(N\) because of the equivalence between the two norms. For every \(m^i \in \mathbb{Z}^2\), \((m_i, e(N, \alpha_i)) = g_{n N} \Lambda, k_i\) for some unique \(k_i = (k_i, \tilde{l}_i) \in \mathbb{Z}^2\), we need to show that \(\tilde{l}_i = l_i\), where \(l_i\) allows \(-\frac{1}{2} < k_i \alpha_i + l_i \leq \frac{1}{2}\). When \(\tilde{l}_i\) is not equal to \(l_i\), then \(|N(k_i \alpha_i + \tilde{l}_i)| \geq N/2\), contradicting the fact that \(\|g_{n N} \Lambda, k_i\|\) is much smaller than \(N\). Therefore \(\tilde{l}_i = l_i\), and \(k_i = k_i\).

For each pair of vectors \(\{(m^1, m^2) \in \mathbb{Z}^2 \times \mathbb{Z}^2 \mid \|m^i\| \leq K(\epsilon), i = 1, 2\}\), we have the corresponding set of \(\{k_i\}_{i=1,2}\), this gives us a unique vector \(k = (k_1, k_2) \in \mathbb{Z}^2\). Therefore the second statement follows. \(\square\)
5.2. For \( m^i \in \mathbb{Z}^2 \), \( \alpha_i \in \mathbb{T} \), we define the coordinates of the corresponding vector as

\[
(m^i, e(N, \alpha_i)) = (X_{m^i}, Z_{m^i})
\]

(5.4)

and define \( m = (m^1, m^2) \in \mathbb{Z}^2 \times \mathbb{Z}^2 \), \( X_m = (X_{m^1}, X_{m^2}) \in \mathbb{R}^2 \) and \( R_m = \|X_m\| \). We introduce a corresponding function to \( g(r, x, \alpha; N, k) \) with \( m \) as a variable:

\[
h(r, x, \alpha; N, m) = d_r(N, m) \cos(2\pi N(X_m, x) + \pi \frac{N-1}{N}(\sum_{i=1}^2 Z_{m^i})) \prod_{i=1}^2 \sin(\pi Z_{m^i})
\]

\[
R_m \prod_{i=1}^2 Z_{m^i}
\]

with

\[
d_r(N, m) = \frac{1}{\pi \frac{N}{2}} \left( \frac{X_m}{R_m} \right) \sin(2\pi (r NP(X_m) - \frac{1}{8})).
\]

From Section 4.1. we see that for \( \alpha \notin \mathbb{E}_N \),

\[
\sum_{m \in \mathbb{Z}^2 \times \mathbb{Z}^2, \|m\| \leq K(\epsilon), i=1,2} h(r, x, \alpha; N, m) = \sum_{k \in U(N, \alpha, \epsilon)} g(r, x, \alpha; N, k),
\]

where \( U(N, \alpha, \epsilon) \supset \hat{S}(N, \alpha) \).

When we restricted ourselves to prime vectors \( m \in \mathbb{Z}^2 \times \mathbb{Z}^2 \), the variables \( r NP(X_m) \mod 1 \) will become independent random variables that are uniformly distributed on \( \mathbb{T}^1 \). In fact, \( m \) could be rewritten as \( \hat{p}(p_1 m^1, p_2 m^2) \), where \( \hat{p} \) is the signed greatest common divisor such that the first coordinate of \( (p_1 m^1, p_2 m^2) \) is positive, and \( m^1, m^2 \) and \( p = (p_1, p_2) \) are in \( \mathbb{Z} \). Since all the vectors are multiples of the prime ones, we introduce

\[
X_{p,m} = (p_1 X_{m^1}, p_2 X_{m^2})
\]

(5.5)

and \( R_{p,m} = \|X_{p,m}\| \).

Introduce

\[
q(r, x, \alpha; N, m, p, \hat{p}) =
\]

\[
d_r(N, m, p, \hat{p}) \cos \left( 2\pi \hat{p} \left( \sum_{i=1}^2 (p_i (m_1, \gamma_i (\alpha, x, N))) \right) + \pi \frac{N-1}{N} \hat{p} \left( \sum_{i=1}^2 (p_i Z_{m^i}) \right) \right) \prod_{i=1}^2 \sin(\pi \hat{p} p_i Z_{m^i})
\]

\[
|\hat{p}| R_{p,m} \prod_{i=1}^2 (p_i Z_{m^i})
\]

(5.6)

where

\[
d_r(N, m, p, \hat{p}) = \frac{1}{\pi \frac{N}{2}} \left( \frac{X_{p,m}}{R_{p,m}} \right) \sin(2\pi (\hat{p} r NP(X_{p,m}) - \frac{1}{8}))
\]

and

\[
\gamma_i(\alpha, x, N) = N x_i (e_{11}(N, \alpha_i), e_{21}(N, \alpha_i)),
\]

(5.7)
where $e_{ij}$ is the $j$th coordinate of the short vector $e_i$.

Recall the definition of $Z$ in Section 2. Remind that $Z^2 = \{ m = (m^1, m^2) \in \mathbb{Z}^2 \times \mathbb{Z}^2 : m^i \in \mathbb{Z}; i = 1, 2 \}$ and define $Z_\varepsilon = \{ m \in Z^2 : \|m^i\| \leq K(\varepsilon); i = 1, 2 \}$. Since the set $Z$ only consists of the primitive vectors with positive first coordinate, we need to add all the positive and negative $\bar{p}$’s, the summation of $q$ above becomes a summation of $q$, note that the summation over $p$ and $m$ is finite, and the sum over $\bar{p}$ is finite due to its large power:

$$
\sum_{|\bar{p}|=1}^{\infty} \sum_{p \in Z_\varepsilon} \sum_{m \in Z_\varepsilon} q(r, x, \alpha; N, m, p, \bar{p}).
$$ (5.8)

Essentially we have reformulated our discrepancy function as a function defined on the lattice space, and from Step 5 of Section 4 we have the following proposition:

**Proposition 5.1.** Assume that $(r, x, \alpha)$ are uniformly distributed in $X = [a, b] \times T^2 \times T^2$, then the sum (5.8) and the normalized discrepancy function $D_{C, 2}$ in (1.5) of Theorem 1 converge to the same law in distribution as $N \to \infty$ and then $\varepsilon \to 0$, if there exists a limit law for the sum (5.8).

**5.3. Uniform distribution of unstable submanifold $\Lambda_\alpha$.** Since for each $i$, $\Lambda_{\alpha_i}$ is the unstable submanifold under the geodesic flow $g_T$ and will become equidistributed over the whole manifold $M$, naturally the same uniform distribution law of holds in the finite product space $M^2$. We have the following proposition(see [5], Theorem 5.3):

**Proposition 5.2.** Denote by $\mu$ the Haar measure on $M^2$. If $\Phi : (\mathbb{R}^2 \times \mathbb{R}^2)^2 \times \mathbb{R}^2 \to \mathbb{R}$ is a bounded continuous function, then

$$
\lim_{N \to \infty} \int_{T^2} \Phi(e(L(N, \alpha_1)), e(L(N, \alpha_2)), \alpha)d\alpha = \int_{M^2 \times T^2} \Phi(e(L_1), e(L_2), \alpha)d\mu(L_1 \times L_2)d\alpha
$$ (5.9)

**6 OSCILLATING TERMS.**

In this section we will prove that typical variables appeared in the sum (5.8) will behave like independent uniformly distributed random variables. We denote by $\mu_2$ the distribution of $e(L_1) \times e(L_2)$ when $L = L_1 \times L_2$ is distributed according to Haar measure on $M^2 = \prod_2 \text{copies } SL(2, \mathbb{R})/SL(2, \mathbb{Z})$. We denote by $\lambda_{2, \varepsilon}$ the Haar measure on $(T^2)^2 \times \mathbb{T}^2 \times \mathbb{Z}$.

The main result of this section is the following, from which the main theorem follows:
Proposition 6.1. Assume that \((x, \alpha, r)\) are uniformly distributed on \(\mathbb{T}^2 \times \mathbb{T}^2 \times [a, b]\), then the following random variables

\[ e(N, \alpha_1), \ldots, e(N, \alpha_2), \{\gamma_j\}_{j=1}^2, \{A_{p,m}\}_{p \in \mathbb{Z}, m \in \mathbb{Z}}, \]

where \(A_{p,m} = rNP(X_{p,m})\), converge in distribution as \(N \to \infty\) to \(\mu_2 \times \lambda_{2,\epsilon}\).

In order to prove Proposition 6.1 in Section 6.2, we will first prove that for different vectors \((p^{(1)}, m^{(1)}), \ldots, (p^{(K)}, m^{(K)})\) in \(\mathbb{Z} \times \mathbb{Z}^2\), \(\{P(X_{p^{(i)}, m^{(i)}})\}_{i=1}^K\) are typically independent over \(\mathbb{Q}\).

6.1. Exceptionally in this subsection we use the lower index for \(m_i\) to represent a vector in \(\mathbb{Z}^2\), not to be confused with the coordinates in the Notations in section 2. For \(m = (m_1, m_2) \in \mathbb{Z}^2 \times \mathbb{Z}^2\) with \(m_i \in \mathbb{Z}^2\), and \(p = (p_1, p_2) \in \mathbb{Z}^2\), \(p_i \geq 1\), define the function \(Q_{p,m} : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R} : (z_1, z_2) \mapsto P((pm_1, z_1), (p_2m_2, z_2))\), where \(z_i = (z_{i1}, z_{i2}) \in \mathbb{Z}^2\) is a vector, and the bracket means euclidean inner product.

Proposition 6.2. For different vectors \((p^{(1)}, m^{(1)}), \ldots, (p^{(K)}, m^{(K)})\) in \(\mathbb{Z} \times \mathbb{Z}^2\), if \(l_1, \ldots, l_K\) are such that \(\sum_{i=1}^K l_i Q_{p^{(i)}, m^{(i)}} \equiv 0\), then \(l_i = 0\) for \(i = 1, \ldots, K\).

Proof. The proof follows the same line of reasoning of Lemma 5.2 and Proposition 5.4 in [4]. Similarly to Lemma 5.2 in [4], it is easy to see that the functions \(Q_{p,m}\) and \(P\) are real analytic and not equal to a polynomial in their variables. First we prove that the sum of terms with the same \(m^{(i)}\), (therefore with different \(p^{(i)}\)'s) must be zero.

For the first part \(m_1^{(i)}\) of \(m^{(i)}\), let \(\alpha_1 = (\alpha_{11}, \alpha_{12})\), \(\beta_1 = (\beta_{11}, \beta_{12})\), \(z_{ij} = \delta \alpha_{ij} + \theta \beta_{ij}\) and \(z_2 = (0, 0)\), we have

\[ Q_{p^{(i)}, m^{(i)}}(z_1, z_2) = p_1^{(i)}(m_1^{(i)}, \alpha_1)P\left(\delta + \frac{(m_1^{(i)}, \beta_1)}{(m_1^{(i)}, \alpha_1)}\theta, 0\right), \quad (6.1) \]

develop \(Q_{p^{(i)}, m^{(i)}}\) with respect to \(\theta\) and consider the coefficient of \(\theta^2\) in the sum for \(l_i Q_{p^{(i)}, m^{(i)}}\), we have:

\[ \sum_{i=1}^K l_i \left| p_1^{(i)}(m_1^{(i)}, \beta_1)\right|^2 Q'_{p^{(i)}, m^{(i)}}(\delta) = 0. \quad (6.2) \]

With \(m_1^{(i)}\)'s all being prime vectors in \(\mathbb{Z}^2\), we can choose a special \(\alpha_1\) such that \(|(m_1^{(i)}, \alpha_1)|\) is arbitrarily small for one \(i\) while all the other \(|(m_1^{(j)}, \alpha_1)|\) that are distinct from \((m_1^{(i)}, \alpha_1)\) have a uniform lower bound, then the sum of \(l_i Q_{p^{(i)}, m^{(i)}}\) with identical \(m_1^{(i)}\) must be zero. Repeat this procedure for \(m_2^{(i)}\), then the sum of \(l_i Q_{p^{(i)}, m^{(i)}}\) with identical \(m^{(i)}\) must be
Next, we can assume that all \( m^{(i)} \) are the same. For the sake of simplicity, we assume that \( m^{(1)}_1 = (1, 0) \) and \( m^{(2)}_2 = (1, 0) \). First we suppose that \( z_{21} = 0 \), choose \( j \) such that \( p^{(j)}_1 \) is the greatest among all \( p^{(i)}_1 \), then

\[
Q_{p^{(i)}, m^{(i)}}(z_1, 0) = P(p^{(i)}_1 z_{11}, 0) = p^{(j)}_1 P\left(\frac{p^{(i)}_1}{p^{(j)}_1} z_{11}, 0\right),
\]

Consider the \( n \)-th partial derivative of \( Q_{p^{(i)}, m^{(i)}} \) with respect to \( z_{11} \), then

\[
\frac{\partial^n}{\partial z_{11}^n} Q_{p^{(i)}, m^{(i)}}(z_1, 0) = \left(\frac{p^{(i)}_1}{p^{(j)}_1}\right)^n \frac{\partial^n}{\partial z_{11}^n} P\left(\frac{p^{(i)}_1}{p^{(j)}_1} z_{11}, 0\right).
\]

Since \( \left(\frac{p^{(i)}_1}{p^{(j)}_1}\right)^n < 1 \) for all \( i \neq j \), we can take \( n \) sufficiently large by analyticity of \( P \), then \( l_j \) becomes the dominant coefficient in the linear combination of \( n \)-th derivatives, we must have the linear combination of terms of identical maximal \( p^{(i)}_1 \) is zero. By repeating this procedure for \( p^{(i)}_2 \), we can deduce that the coefficient \( l_i \) in front the term that has the greatest \( p^{(2)}_2 \) among those having the greatest \( p^{(1)}_1 \) is zero. Inductively, all coefficients \( l_i \) are zero. 

By Proposition 6.2, we can deduce the following: if we take a lattice \( L \in M^2 \) and let \( z_i = (e_{11}(L_i), e_{21}(L_i)), i = 1, 2, \) then \( P(X_{p,m}(L)) = P((p_1 m_1, z_1), (p_2 m_2, z_2)) = Q_{p,m}(z_1, z_2) \). By analyticity, for any different \( (p^{(1)}, m^{(1)}), \ldots, (p^{(K)}, m^{(K)}) \) in \( Z \times Z^2 \),

\[
\mu \left( L : \sum_{i=1}^{K} l_i P(X_{p,m}(L)) = 0 \right) = 0. \tag{6.3}
\]

Now by Proposition 5.2 we have that

\[
\text{mes} \left( \alpha \in \mathbb{T}^2 : \left| \sum_{i=1}^{K} l_i P(X_{p^{(i)}, m^{(i)}}(L(N, \alpha))) \right| < \epsilon \right) \to 0 \tag{6.4}
\]

as \( \epsilon \to 0, N \to \infty \).

6.2. Proof of Proposition 6.1. Take integers \( n_{ij}, n_{21}, n_{12}, n_{22}, \{l_{p,m}\}_{p \in \mathbb{Z}^2, m \in \mathbb{Z}^2} \) and
a function $\Phi : (\mathbb{R}^2 \times \mathbb{R}^2)^2 \to \mathbb{R}$ of compact support. It remains to show that as $N \to \infty$
\[
\iiint \Phi(e(L(N, \alpha_1)), e(L(N, \alpha_2))) \exp \left[ 2\pi i \left( \sum_{j=1}^{2} (n_j \gamma_{j1} + n_j \gamma_{j2}) + \sum_{p \in \mathbb{Z}^2, m \in \mathbb{Z}^2} l_{p,m} A_{p,m} \right) \right] dxd\alpha dr \to \int_{M^2} \Phi(e(L_1), e(L_2)) d\mu(L) \int_{\mathbb{T}^{2d}} e^{2\pi i \sum_j (n_j \gamma_{j1} + n_j \gamma_{j2})} d\gamma \int_{\mathbb{T}^{2\times 2}} e^{2\pi i \sum_{p,m} l_{p,m} A_{p,m}} dA,
\]
(6.5)
as $N \to \infty$.

Proof. This proof is very close to the proof of Proposition 5.1 in [4], it suffice to rewrite the original proof with the new variables and use Proposition 6.2 and (6.4). If for all $j$ and $p, m$, $n_{j1} \equiv 0$, $n_{j2} \equiv 0$ and $l_{p,m} \equiv 0$, (6.5) is a special case of (5.3). Then it suffice to prove (6.5) in the case that at least some $n_j$ or some $l_{p,m}$ are non-zero, then the right-hand side of (6.5) is zero, and it reduces to the following:
\[
\iiint \Phi(e(L(N, \alpha_1)), e(L(N, \alpha_2))) \exp \left[ 2\pi i \left( \sum_{j=1}^{2} (n_j \gamma_{j1} + n_j \gamma_{j2}) + \sum_{p \in \mathbb{Z}^2, m \in \mathbb{Z}^2} l_{p,m} A_{p,m} \right) \right] dxd\alpha dr \to 0
\]
(6.6)
If $n_{j1} \neq 0$ for at least one $j$, recall that $\gamma_j(\alpha, x, N) = N x_j (e_{11}(N, \alpha_j), e_{21}(N, \alpha_j))$, then the coefficient in front of $x_j$ in $\sum_j (n_{j1} \gamma_{j1} + n_{j2} \gamma_{j2})$ is $N(n_{j1} e_{11}(N, \alpha_j) + n_{j2} e_{21}(N, \alpha_j))$. Note that the coordinates $e_{11}(N, \alpha_j)$ and $e_{21}(N, \alpha_j)$ are typically $\mathbb{Z}$-independent outside a zero measure set of $\alpha_j$. Hence (5.3) implies that
\[
\text{mes} \left( \alpha_j \in \mathbb{T} : |n_{j1} e_{11}(N, \alpha_j) + n_{j2} e_{21}(N, \alpha_j)| < \frac{1}{\sqrt{N}} \right) \to 0
\]
(6.7)
as $N \to \infty$. This limit states that most $\alpha_j$ will not allow the coefficient in front of $x_j$ to be too small, then the integral of (6.3) can be decomposed into two parts, $LHS = I_1 + I_2$, where $I_1$ corresponds to the part of integral for $\alpha_j$ with $|n_{j1} e_{11}(N, \alpha_j) + n_{j2} e_{21}(N, \alpha_j)| < \frac{1}{\sqrt{N}}$ and $I_2$ the part for $\alpha_j$ with $|n_{j1} e_{11}(N, \alpha_j) + n_{j2} e_{21}(N, \alpha_j)| \geq \frac{1}{\sqrt{N}}$. Then
\[
|I_1| \leq \text{Const}(\Phi) \text{mes} \left( \alpha_j \in \mathbb{T} : |n_{j1} e_{11}(N, \alpha_j) + n_{j2} e_{21}(N, \alpha_j)| < \frac{1}{\sqrt{N}} \right)
\]
so it can be arbitrarily small as $N \to \infty$ by (6.7). For $I_2$, since the coefficient of $x_j$ is not too small, we use integrate by parts with respect to $x_j$ to achieve the following
\[ |I_2| \leq \frac{\text{Const}(\Phi)}{\sqrt{N}}. \]

Therefore this proves the case where not all \( n_{j1} \) vanish, the case where not all \( n_{j2} \) vanish is the same.

Similarly, if there exists some \((p, m)\), such that \( l_{p,m} \) is non-zero, we can use (6.4) and integrate with respect to \( r \) to obtain (6.3), using the same decomposition integration technique.

6.3. Proof of Theorem 1(b). Combining Proposition 5.1 and Proposition 6.1, by letting \( N \to \infty \) and then \( \epsilon \to 0 \), we can substitute the variables in (5.8) by uniformly distibuted random variables on the infinite tori, thus we obtain Theorem 1(b) and Proposition 2.2.

References

[1] Vladimir Igorevich Arnol’d. *Mathematical methods of classical mechanics*, volume 60. Springer Science & Business Media, 2013.

[2] Shrikrishna Gopal Dani. Divergent trajectories of flows on homogeneous spaces and diophantine approximation. 1985.

[3] Dmitry Dolgopyat and Bassam Fayad. Deviations of ergodic sums for toral translations ii. boxes. *arXiv preprint arXiv:1211.4323*, 2012.

[4] Dmitry Dolgopyat and Bassam Fayad. Deviations of ergodic sums for toral translations i. convex bodies. *Geometric and Functional Analysis*, 24(1):85–115, 2014.

[5] CS Herz. Fourier transforms related to convex sets. *Annals of Mathematics*, pages 81–92, 1962.

[6] H Kesten. Uniform distribution mod 1. *Annals of Mathematics*, pages 445–471, 1960.

[7] H Kesten. Uniform distribution mod 1 (ii). *Acta Arithmetica*, 7(4):355–380, 1962.

[8] Jens Marklof and Andreas Strömbergsson. The distribution of free path lengths in the periodic lorentz gas and related lattice point problems. *Annals of Mathematics*, pages 1949–2033, 2010.

[9] Madabusi Santanam Raghunathan. *Discrete subgroups of Lie groups*, volume 3. Springer, 1972.