Limiting distribution of visits of several rotations to shrinking intervals

Ilya Vinogradov

May 11, 2010

Abstract

We show that given \( n \) normalized intervals on the unit circle, the numbers of visits of \( d \) random rotations to these intervals have a joint limiting distribution as lengths of trajectories tend to infinity. If \( d \) then tends to infinity, then the numbers of points in different intervals become asymptotically independent unless an arithmetic obstruction arises. This is a generalization of earlier results of J. Marklof.

The following question arises from two results of Marklof about gap distribution for rotations. Fix a point \( \xi \) in \([0,1)\) and let \( B_N = (0, N^{1/(d-1)}) \oplus \mathbb{Z}^{d} \subset \mathbb{R}^{d} \). What is the limit behavior of the number of points of the form

\[
\left\{ \sum_{i=1}^{d-1} m_i \alpha_i \pmod{1} : m_i \in \left[1, N^{1/(d-1)}\right] \cap \mathbb{Z}, 1 \leq i \leq d-1 \right\},
\]

for \( \alpha_i \in [0,1) \) that land in \((\xi - \frac{\alpha}{N}, \xi + \frac{\alpha}{N})\) as \( N \to \infty \)? In [1], J. Marklof showed that

\[
\text{leb} \left\{ (\alpha, \xi) \in [0,1)^{d-1} \times [0,1) : \#\{m \in B_N \cap \mathbb{Z}^{d} : \xi + \sum_{i=1}^{d-1} m_i \alpha_i + m_d \in (-\frac{\alpha}{N}, \frac{\alpha}{N})\} = A \right\} \to P^{(d)}(A)
\]

as \( N \to \infty \) and found its decay as \( A \to \infty \). His main tool was the mixing property of a diagonal flow on \( \text{SL}(d,\mathbb{R})/\text{SL}(d,\mathbb{Z}) \) that had been proved by Moore [3]. In a later note Marklof remarked that for one variable (that is, \( d = 2 \)), a stronger result is true due to a Theorem of Shah [6]. Namely, for fixed \( \xi \in [0,1) \setminus \mathbb{Q} \),

\[
\text{leb} \left\{ \alpha \in [0,1) : \#\{m \in \{1, \ldots, N\} : \xi + m\alpha \pmod{1} \in (-\frac{\alpha}{N}, \frac{\alpha}{N})\} = A \right\} \to P^{(2)}(A).
\]

This result uses Ratner’s Theorem on measures invariant under unipotent flows [4]. We will generalize the theorems mentioned above to joint limiting probability distributions for several intervals and study their large \( d \) limits.
1 Notation and results

We will use the following notation.

- $N$, $n$, and $d$ denote positive integers with $d \geq 2$;
- upper indices (usually $j$) run from 1 to $n$ and lower indices (usually $i$) run from 1 to $d$ unless stated otherwise;
- $m = (m_1, \ldots, m_d)$ is a vector of $d$ integers;
- $\sigma = (\sigma^1, \ldots, \sigma^n)$ is a positive vector ($\sigma^j > 0$ for all $j$);
- $\alpha = (\alpha_1, \ldots, \alpha_{d-1}) \in (\mathbb{R}/\mathbb{Z})^{d-1}$;
- $\xi = (\xi^1, \ldots, \xi^n) \in (\mathbb{R}/\mathbb{Z})^n$;
- $\tau = (\tau^1, \ldots, \tau^n)$ is a real vector;
- $\text{Pois}_\sigma$ denotes the Poisson distribution with parameter $\sigma$.

If $n = 1$, we will write $\sigma$ instead of $\sigma^1$ and similarly for other variables. Let $B_N = (0, N^{1/(d-1)}) \oplus (d-1) \oplus \mathbb{R} \subset \mathbb{R}^d$ as before. For a measurable set $S$ define random variables $X^{N,d}_{\xi,S} : [0,1) \to \mathbb{Z}$ by

$$X^{N,d}_{\xi,S} = \# \left\{ m \in \mathbb{Z}^d \cap B_N : \sum_{i=1}^{d-1} m_i \alpha_i + m_d \in \xi + S N \right\}.$$

We will usually suppress the upper indices on $X_{\xi,S}$. For a vector $\xi \in T^n$, the set $\Xi \subset T^n$ is the closure of the orbit of rotation by $\xi$ on the torus: $\Xi = \overline{\{ k \xi : k \in \mathbb{Z} \}}$; it is the smallest closed Lie subgroup of $T^n$ that contains $\xi$.

Our results for limiting distributions of $X_{\xi,S}$ are as follows.

**Theorem 1.** Fix any absolutely continuous probability measure on $[0,1)$. With notation as above, the distribution of

$$X_{\xi,\tau,\sigma} = (X^{\xi_1,\tau^1,\sigma^1}, \ldots, X^{\xi_n,\tau^n,\sigma^n})$$

has a weak limit as $N \to \infty$; we denote it by $\mathbb{D}^{(d)}_{n,\sigma,\Xi/\mathbb{Z}}$.

In other words, the numbers of points in shrinking segments $(\xi^j + \frac{\tau^j}{N}, \xi^j + \frac{\sigma^j + \tau^j}{N})$, $1 \leq j \leq n$, with fixed “centers” $\xi^j$ have a joint limiting distribution as $N$ tends to infinity. The limiting distribution depends on $d$, $n$, $\sigma$, $\Xi$, and $\tau$ modulo $\Xi$. In particular, if $\Xi = T^n$, then the distribution is independent of $\tau$.

**Remark 1.** Jens Marklof proved special cases of this theorem. He proved the case $n = 1$, $d = 2$ in [2] and the case $n = 1$ and arbitrary $d$ with average over $\xi$ in [1].
Theorem 2. Let $\mathbf{P}_{n,\sigma,\Xi}^{(d)}$ be the distribution from Theorem 1. Then, $\mathbf{P}_{n,\sigma,\Xi}^{(d)}$ has a weak limit as $d \to \infty$. Furthermore,

$$\mathbf{P}_{n,\sigma,\Xi}^{(d)} \Rightarrow (\text{Pois} \sigma_1, \ldots, \text{Pois} \sigma_n)$$

as $d \to \infty$ iff

$$(\tau^j, \tau^j + \sigma^j) \cap (\tau^{j'}, \tau^{j'} + \sigma^{j'}) = \emptyset$$

whenever $\xi^j = \xi^{j'}$.

In effect, this Theorem says that as the number of rotations tends to infinity, the gap lengths exhibit random behavior. However, for every finite $d$ and $\Xi$ (with $n \geq 2$) we have that $\mathbf{P}_{n,\sigma,\Xi}^{(d)}$ is dependent.

Acknowledgements. I would like to thank Ya. Sinai, E. Lindenstrauss, J. Marklof, F. Cellarosi, A. Salehi Golsefidy, and Z. Wang for helpful discussions.

2 Large $N$ limit

Proof of Theorem 1. We reformulate the problem in the language of homogeneous spaces. Let $L = SL(d, \mathbb{R}) \times (\mathbb{R}^d)^{\oplus n}$ and let $\Lambda = SL(d, \mathbb{Z}) \times (\mathbb{Z}^d)^{\oplus n} \subset L$. Multiplication law on $L$ is given by

$$(M, v^1, \ldots, v^n)(N, w^1, \ldots, w^n) = (MN, v^1 + Mw^1, \ldots, v^n + Mw^n).$$

It is well-known that $\Lambda \subset L$ is a non-cocompact lattice. The homogeneous space $L/\Lambda$ is a bundle over $SL(d, \mathbb{R})/SL(d, \mathbb{Z})$ with fiber $(\mathbb{T}^d)^{\oplus n}$.

Given a set of vectors $v^1, \ldots, v^n \in \mathbb{T}^d$, let

$$L_V = \{(1, v^1, \ldots, v^n)^{-1}(M, 0, \ldots, 0)(1, v^1, \ldots, v^n) \mid M \in SL(d, \mathbb{R})\} \subset L;$$

it is of course isomorphic to $SL(d, \mathbb{R})$. Also define $\hat{L}_V$ to be the smallest group containing $L_V$ that is defined over $\mathbb{Q}$. Dimension of $\hat{L}_V$ depends on $v^j$. If all vectors $v^j$ have rational coordinates, then $\hat{L}_V = L_V$. Otherwise the fiber over the identity in $\hat{L}_V$ is the smallest $\mathbb{Q}$-vector space containing the identity fiber for $L_V$. This construction can be carried to other points. Finally set $\hat{L}_V = L_V \cap \Lambda$ which is a lattice in $\hat{L}_V$ by construction.

For our purposes fix $v^j = (0, \ldots, 0, \xi^j)^T$. For this choice of $v^j$ we get the homogeneous space $\hat{L}_V/\hat{\Lambda}_V$. We have constructed $\hat{L}_V$ so that $\pi(\hat{L}_V) = \pi(L_V)$, where $\pi: L \to L/\Lambda$ is the canonical projection. The structure of this space depends on $\Xi = \mathbb{Z}\xi \subset \mathbb{T}^n$. It is a subbundle of $L/\Lambda$: the base is still $SL(d, \mathbb{R})/SL(d, \mathbb{Z})$ but the fiber is $\mathbb{Z}^d$ after reordering coordinates.

We define $f_{\tau,\sigma}: \hat{L}_V/\hat{\Lambda}_V \to \mathbb{R}^n$ by

$$f_{\tau,\sigma}(M, v^1, \ldots, v^n) = (g_{\tau^1,\sigma^1}(M, v^1), \ldots, g_{\tau^n,\sigma^n}(M, v^n)),$$

where

$$g_{\tau,\sigma}(M, v) = \sum_{m \in \mathbb{Z}^d \setminus \{0\}} \chi_1(\tilde{m}_1) \cdots \chi_1(\tilde{m}_{d-1}) \chi_{(\tau,\tau+\sigma)}(\tilde{m}_d),$$

$$\chi_\sigma(x) = \begin{cases} 1 & x \in (0, \sigma) \\ 0 & \text{otherwise}, \end{cases}$$
\[ \hat{m} = Mm + v. \]

It is easily seen that \( f \) is \( \hat{\Lambda}_V \) invariant and hence well-defined on the quotient.

We need to show that \( \text{leb}\{ f_{\tau, \sigma} = (A^1, \ldots, A^n) \} \rightarrow P_{n, \sigma, \tau}^{(d)}(A^1, \ldots, A^n) \) as \( N \rightarrow \infty \). To this end we use a theorem of Shah (Theorem 1.4 in \([6]\)). The form we need is the following:

**Theorem 3** (Shah). Let

\[ U_\alpha = \begin{pmatrix} 1 & & & \cdot \cdot \cdot & \cdot \cdot \cdot & 1 \\ & \cdot \cdot \cdot & \cdot \cdot \cdot & 1 \\ \alpha_1 & \cdot \cdot \cdot & \alpha_{d-1} & 1 \end{pmatrix} \quad \text{and} \quad \Phi_t = \begin{pmatrix} e^{-t} & & & & & \\ & \cdot \cdot \cdot & & & \cdot \cdot \cdot & \cdot \cdot \cdot \\ & & \cdot \cdot \cdot & \cdot \cdot \cdot & \cdot \cdot \cdot & \cdot \cdot \cdot \\ & & & & & e^{(d-1)t} \end{pmatrix}. \]

Let \( L \) be a Lie group, \( \Lambda \subset L \) a lattice, \( \varphi : SL(d, \mathbb{R}) \rightarrow L \) an embedding. If the image of \( \varphi \) is dense when projected to \( L/\Lambda \), then for any bounded continuous \( \eta \)

\[ \lim_{t \rightarrow \infty} \int_{\mathbb{R}^{d-1}} \eta(\varphi(\Phi_t U_\alpha))d\nu(\alpha) = \int_{L/\Lambda} \eta(M)d\mu(M), \quad (1) \]

where \( \nu \) is any absolutely continuous probability measure on \( U_\alpha \) and \( \mu \) is the Haar probability measure on \( L/\Lambda \).

**Remark 2.** In effect, the Theorem says that the unstable manifold \( U_\alpha \) is equidistributed in the larger homogeneous space \( L/\Lambda \) provided the density assumption is satisfied.

We set \( N = e^{(d-1)t} \) and apply the Theorem 3 with \( L = \hat{L}_V \), \( \Lambda = \hat{\Lambda}_V \), and

\[ \varphi : M \mapsto (1, v^1, \ldots, v^n)^{-1}(M, 0, \ldots, 0)(1, v^1, \ldots, v^n). \]

This ensures density after projecting to \( \hat{L}_V/\hat{\Lambda}_V \) by construction. We now use the following elementary Lemma to construct appropriate functions \( \eta \).

**Lemma 4.** Let \( N_0 = N \cup \{0\} \). Then \( p_2 : N_0 \times N_0 \rightarrow N_0 \) given by

\[ (x, y) \mapsto \left( \frac{x + y + 2}{2} \right) - (y + 1) \]

is a bijection.

By induction, there exists a polynomial bijection between \( N_0^n \) and \( N_0 \) for each \( n \); call it \( p_n \). Set

\[ h_n(M, x^1, \ldots, x^n) = p_n(g_{\tau, \sigma^1}(M, x^1), \ldots, g_{\tau, \sigma^n}(M, x^n)) \]

and apply Shah’s Theorem to functions

\[ \eta_A(N, y^1, \ldots, y^n) = \chi(h_n(M, x^1, \ldots, x^n) = p_n(A^1, \ldots, A^n))(1, v^1, \ldots, v^n)(N, y^1, \ldots, y^n) \]
for all nonnegative integers $A^j$. These functions are not continuous, but are indicators of nice sets. Using a standard approximation argument we can apply the Theorem to them as well. For $M$ of the form

$$M = \begin{pmatrix} N^{-1/(d-1)} & \cdots & N^{-1/(d-1)} \\ \vdots & \ddots & \vdots \\ N & \cdots & N \\ \alpha_1 & \cdots & \alpha_{d-1} \end{pmatrix},$$

we recover the numbers of points in the $n$ segments. In fact, for all $M$

$$\eta_A((1, v^1, \ldots, v^n)^{-1}(M, 0, \ldots, 0)(1, v^1, \ldots, v^n)) = \chi_{\{h_n(M, x^1, \ldots, x^n) = p_n(A)\}}(M, Mv^1, \ldots, Mv^n)$$

and

$$g_{\tau, \sigma}(M, M\nu) = \sum_{m \in \mathbb{Z}^{d}\setminus\{0\}} \chi_1 \cdots \chi_1 \chi_{(\tau, \tau + \sigma)}(M(m + \nu)),$$

and for the particular choice of $M$ above,

$$M(m + \nu) = \begin{pmatrix} m_1 / N^{-1/(d-1)} \\ \vdots \\ m_{d-1} / N^{-1/(d-1)} \\ (m_1 \alpha_1 + \cdots + m_{d-1} \alpha_{d-1} + \xi)N \end{pmatrix}.$$

Thus, $\eta_A(\varphi(M)) = 1$ if and only if there are exactly $A^1, \ldots, A^n$ visits to the segments around $\xi^1, \ldots, \xi^n$ for the given $\alpha$ and $N$. Hence the form of the limiting distribution is

$$P^{(d)}_{n, \sigma, \Xi, \tau, x}(A^1, \ldots, A^n) = \mu_{\hat{L}_{\nu}} \{ f_{\tau, \sigma}(M, x^1, \ldots, x^n) = (A^1, \ldots, A^n) \}.$$

### 3 Large $d$ Limit

In this section we consider the large $d$ limit of the distributions from the previous sections and prove Theorem 2. Before proving the theorem, we will need basic information about the Poisson distribution.

Poisson distribution with parameter $\sigma$ weighs each non-negative integer $k$ with weight $e^{-\sigma} \sigma^k / k!$. We will denote Poisson distribution with parameter $\sigma$ by $\text{Pois} \sigma$. Its moments have the form

$$\sum_{k=0}^{\infty} k^n e^{-\sigma} \frac{\sigma^k}{k!} = e^{-\sigma} \left( \frac{d}{d\sigma} \right)^n e^\sigma = \sum_{k=1}^{n} S(n, k) \sigma^k,$$

where $S(n, k)$ is the Stirling number of the second kind. As can be easily seen from the above equality, the Stirling number is the number of partitions of a set of $n$ elements into $k$ nonempty sets. The first few moments of the Poisson distribution are $\sigma$, $\sigma^2 + \sigma$, $\sigma^3 + 3\sigma^2 + \sigma$. These correspond to partitions $\{1\}$; $\{12\}$; $\{11\}$; $\{123\}$; $\{112\}$; $\{121\}$; $\{211\}$; $\{111\}$.
To further study the limiting distributions we have obtained, we will need the following generalization of a proposition of Marklof from [1] which goes back to a theorem of Rogers [5]. Let $\text{Gr}(n, l) = O(n)/(O(l) \times O(n - l))$ denote the Grassmannian of $l$-planes in $\mathbb{R}^n$; we assume that the $l$-planes are embedded in $\mathbb{R}^n$ with respect to the standard basis. Let $\text{Gr}(n, l)(\mathbb{Q}) = \big\{ \pi \in \text{Gr}(n, l) \mid \pi \subset \mathbb{R}^n \text{ is defined over } \mathbb{Q} \big\} = \big\{ \pi \in \text{Gr}(n, l) \mid \pi \cap \mathbb{Z}^n \text{ is a lattice in } \pi \big\}$. For $\pi \in \text{Gr}(n, l)(\mathbb{Q})$, we write $\text{covol}_\pi \mathbb{Z}$ for the covolume of the lattice $\pi \cap \mathbb{Z}^n$ in $\pi$. We also set $G = \text{SL}(d, \mathbb{R})$, $\Gamma = \text{SL}(d, \mathbb{Z})$, and fix $\mu$ to be the Haar probability measure on $G/\Gamma$.

**Theorem 5.** Let $F : (\mathbb{R}^d)^\otimes r \to \mathbb{R}$ be a bounded piecewise continuous function with compact support. Let $f : G/\Gamma \to \mathbb{R}$ be defined by

$$f(M) = \sum_{m^1, \ldots, m^r \in \mathbb{Z}^d} F(Mm^1, \ldots, Mm^r)$$

with $r < d$ a positive integer. Then, the first moment of $f$ is given by the following expression:

$$\int_{G/\Gamma} f(M) d\mu(M) = \sum_{l=0}^r \sum_{\pi' \in \text{Gr}(r, l)(\mathbb{Q})} \int_{\pi'} F(x) \frac{dx}{(\text{covol}_\pi \mathbb{Z})^d},$$

(3)

where $\pi' \in \text{Gr}(rd, ld)(\mathbb{Q})$ is the image of $\pi$ under the embedding

$$(x^1, \ldots, x^r) \mapsto (x^1, \ldots, x^1, \ldots, x^r, \ldots, x^r)$$

and the measure $dx$ is the Lebesgue measure on $\pi'$ that should be interpreted as the delta measure at the origin when $l = 0$.

**Remark 3.** If in the sum over $m^j$ we omit the terms where any of the $m^j$ are 0, then in the sum over $\pi \in \text{Gr}(r, l)(\mathbb{Q})$ we omit planes that are generated by subsets of the standard basis. This follows from the fact that such subsets of $(\mathbb{Z}^d)^r$ are $\text{SL}(d, \mathbb{Z})$-invariant.

**Lemma 6.** With notation as in the Theorem, we have

$$\int_{G/\Gamma} \sum_{\text{linearly indep.} \ m^1, \ldots, m^r \in \mathbb{Z}^d} F(Mm^1, \ldots, Mm^r) d\mu(M) = \int_{\mathbb{R}^d} F(x^1, \ldots, x^r) dx^1 \ldots dx^r.$$

**Proof.** First note that the integral is well-defined since linearly independent sets of vectors are preserved by $\Gamma$. Further renormalize $\mu$ so that $\mu(G/\Gamma) = \prod_{k=2}^d \zeta(k)$ for $d \geq 2$ and write $d\mu(M)/\mu(G/\Gamma)$ in the integral; this normalization will be useful later. Write

$$M = \begin{pmatrix} x_{11} & \cdots & x_{1d} \\ \vdots & \ddots & \vdots \\ x_{d1} & \cdots & x_{dd} \end{pmatrix} \in G/\Gamma.$$
Then for $1 \leq r < d$ we have
\[
M = \begin{pmatrix}
x_{11} & \ldots & x_{1r} \\
\vdots & \ddots & \vdots \\
x_{r1} & \ldots & x_{rr} \\
x_{r+1,1} & \ldots & x_{r+1,r} \\
x_{r+2,1} & \ldots & x_{r+2,r} \\
\vdots & \ddots & \vdots \\
x_{d1} & \ldots & x_{dr}
\end{pmatrix}
\begin{pmatrix}
0_{r \times (d-r)}
\end{pmatrix}
\begin{pmatrix}
\text{Id}_{r \times r} \\
\vdots \\
\vdots \\
0_{(d-r) \times r}
\end{pmatrix}
\begin{pmatrix}
z_{11} & \ldots & z_{1,d-r} \\
\vdots & \ddots & \vdots \\
z_{r1} & \ldots & z_{r,d-r} \\
y_{11} & \ldots & y_{1,d-r} \\
\vdots & \ddots & \vdots \\
y_{d-r,1} & \ldots & y_{d-r,d-r}
\end{pmatrix}
\]
where $(y_{ij}) \in \text{SL}(d-r, \mathbb{R})$. In these coordinates
\[
d\mu = \prod_{i=r}^{d-1} dx_{ij} \prod_{i=r}^{d-1} dz_{ij} \cdot \delta(1 - \det(y_{ij})) \prod_{i,j=r}^{d-1} dy_{ij}.
\] (4)

The last factor is the Haar measure on $\text{SL}(d-r, \mathbb{R})$ normalized to $\zeta(2) \ldots \zeta(d-r)$ (or simply 1 in case $d-r = 1$).

For $j = 1, \ldots, r$ let $t^j = \text{gcd} \ m^j$. Writing $m^j/t^j$ for a column vector, we can find $N \in \text{SL}(d, \mathbb{Z})$ such that
\[
\begin{pmatrix}
m_1^1 & \ldots & m_r^r \\
t_1^1 & \ldots & t_r^r
\end{pmatrix} = N
\begin{pmatrix}
a_{11} & a_{12} & \cdots & a_{1r} \\
0 & a_{22} & \cdots & a_{2r} \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & a_{rr}
\end{pmatrix} = NA.
\]

$A$ is a matrix with integer entries uniquely determined by the following conditions:

- $a_{ij}, 1 \leq i \leq j$ are relatively prime for any fixed $j \in \{1, \ldots, r\}$ (in particular, $a_{11} = 1$);
- $0 \leq a_{1j}, \ldots, a_{j-1,j} < a_{jj}$.

The first condition is due to relative primality of $m^j/t^j$, and the second comes from applying row operations. Given $a_{11} = 1, a_{22}, \ldots, a_{rr}$, the number of possible matrices $A$ of this form is
\[
\prod_{j=1}^{r} \varphi_{j-1}(a_{jj}),
\]
where $\varphi_k$ is the number-theoretic function defined by
\[
\varphi_k(p^r) = p^{rk} \left(1 - \frac{1}{p^k}\right)
\]
for \( k \geq 1 \) and \( \varphi_0 \) is identically 1; \( \varphi_1 \) is Euler totient function. The function \( \varphi_k(n) \) counts the number of \( k \)-tuples \( (n_1, \ldots, n_k) \in \{0, \ldots, n-1\}^k \) such that \( \gcd(n_1, \ldots, n_k) = 1 \).

Let us compute the stabilizer of a fixed matrix \( A \):

\[
\Gamma_A = \{ \gamma \in \Gamma \mid \gamma A = A \} = \left( \begin{array}{c|c}
\text{Id}_{r \times r} & Z_{r \times (d-r)} \\
0_{(d-r) \times r} & \text{SL}(d-r, \mathbb{Z})
\end{array} \right).
\]

Thus we get

\[
\frac{1}{\mu(G/\Gamma)} \int_{G/\Gamma} \sum_{m^j \perp 1} F(Mm^1, \ldots, Mm^r)d\mu(M) =
\]

\[
\frac{1}{\mu(G/\Gamma)} \sum_{t_1^1, \ldots, t_r^r = 1_{G/\Gamma}} \int_{\mathbb{N} \in \Gamma/\Gamma_A} \sum F(MN \begin{pmatrix} a_{11} \\ 0 \\ \vdots \\ 0 \\ \vdots \\ \vdots \\ \vdots \end{pmatrix}, MN \begin{pmatrix} a_{12} \\ 0 \\ \vdots \\ 0 \\ \vdots \\ \vdots \end{pmatrix}, \ldots, MN \begin{pmatrix} a_{1r} \\ 0 \\ \vdots \\ 0 \\ \vdots \\ \vdots \end{pmatrix})d\mu(M) =
\]

\[
\frac{1}{\mu(G/\Gamma)} \sum_{t^j} \int_{G/\Gamma_A} F(M \begin{pmatrix} a_{11} \\ 0 \\ \vdots \\ 0 \\ \vdots \end{pmatrix}, M \begin{pmatrix} a_{12} \\ 0 \\ \vdots \\ 0 \\ \vdots \end{pmatrix}, \ldots, M \begin{pmatrix} a_{1r} \\ 0 \\ \vdots \\ 0 \\ \vdots \end{pmatrix})d\mu(M). \tag{5}
\]

Using \( \mathcal{H} \) to change the measure we get

\[
\frac{1}{\zeta(d) \ldots \zeta(d-r+1)} \sum_{t^j} \sum_{A \in (\mathbb{R}^d)^r} \int F(t_1^{1}x^1 + t_2^{2}a_{12}x^2 + \ldots + t_r^{r}a_{1r}x^r)dx^1 \ldots dx^r.
\]

Now we do a linear change of variables and get

\[
\frac{1}{\zeta(d) \ldots \zeta(d-r+1)} \sum_{t^j=1}^{\infty} \frac{1}{(t_1^{1} \ldots t_r^{r})^d} \sum_{a_{22}, \ldots, a_{rr}=1} \frac{\varphi_1(a_{22}) \ldots \varphi_1(a_{rr})}{a_{22}^{a_2} \ldots a_{rr}^{a_r}} \int_{(\mathbb{R}^d)^r} F(x^1, \ldots, x^r)dx^1 \ldots dx^r. \tag{6}
\]

It is easy to see that

\[
\sum_{n \geq 1} \frac{\varphi_k(n)}{n^d} = \frac{\zeta(d-k)}{\zeta(d)},
\]

whence the constant in front of the integral in \( \mathcal{H} \) is 1, as desired.

\[\square\]

**Proof of Theorem**. Rewrite the integral we are evaluating as

\[
\frac{1}{\mu(G/\Gamma)} \int_{G/\Gamma} \sum_{l=0}^{r} \sum_{\lambda(k(m^1, \ldots, m^r)) = l} F(Mm^1, \ldots, Mm^r)d\mu(M).
\]
Here we normalize $\mu$ like in the Lemma and the inner sum runs over those $r$-tuples of vectors whose $\mathbf{R}$-span has dimension $l$. Since the set $\{\text{rk } \mathbf{m} = l\}$ is $\text{SL}(d, \mathbf{Z})$-invariant for each $l$, we can pass the sum over $l$ through the integral sign. To prove the Theorem, it suffices to show that corresponding terms in the expression above and in (3) match for each $l$. Now observe that

$$\{\text{rk } \mathbf{m} = l\} = \bigcup_{\pi \in \text{Gr}(l,r)(\mathbf{Q})} \{r\text{-tuples of vectors from } (\pi \cap \mathbf{Z}^r)^d \text{ with rank } l\}.$$  

In fact, the sets whose union we are taking constitute an $\text{SL}(d, \mathbf{Z})$-invariant partition. Thus we need to parametrize linearly independent vectors in $(\pi \cap \mathbf{Z}^r)^d$ for each $\pi$. Let $B: \mathbf{R}^l \to \mathbf{R}^r$ be a linear map with image $\pi$ such that $B(\mathbf{Z}^l) = \pi \cap \mathbf{Z}^r$. There can be many of these; any one will do. Using the standard basis, $B = (b^i_j)$ with $1 \leq i \leq l$ and $1 \leq j \leq r$, and we obtain $\mathbf{m}^j = \sum b^i_j \mathbf{n}^i$, where $\mathbf{n}^i \in \mathbf{Z}^d$ form a linearly independent set. Thus the integral becomes

$$\frac{1}{\mu(G/\Gamma)} \int_{G/\Gamma} \sum_{\pi \in \text{Gr}(l,r)(\mathbf{Q})} \sum_{\mathbf{n}^1, \ldots, \mathbf{n}^l \in \mathbf{Z}^d \text{ linearly indep.}} F(M \sum b^i_1 \mathbf{n}^i, \ldots, M \sum b^i_r \mathbf{n}^i) d\mu(M).$$

The quantity in brackets is $\Gamma$-invariant, so the sum over $\pi$ can be interchanged with the integral.

For each $\pi$ and $B$ we can now apply the Lemma. It gives

$$\sum_{\pi \in \text{Gr}(l,r)(\mathbf{Q})} \int_{(\mathbf{R}^d)^r} F(\sum b^i_1 \mathbf{x}^i, \ldots, \sum b^i_r \mathbf{x}^i) d\mathbf{x}^1 \ldots d\mathbf{x}^l.$$  

Since $B(\mathbf{Z}^l) = \pi \cap \mathbf{Z}^r$, the Jacobian of $B$ is the covolume of $\pi$. The statement of the Theorem follows after a linear change of variables.

\[ \square \]

**Proposition 7.** For any $\Xi$ and $\tau$ we have

$$P_{1,\sigma,\Xi,\tau/\Xi}^{(d)} \rightarrow \text{Pois } \sigma$$

as $d \to \infty$.

**Proof.** From (2), all we need to show is that moments of $f_{\tau,\sigma}$ are Poissonian for large $d$. First consider that case when $\xi \not\in \mathbf{Q}$. Without loss of generality we set $\tau = 0$. Then we need to find

$$\lim_{d \to \infty} \int_{M \in G/\Gamma} \int_{v \in \mathbf{Z}^d \setminus \{0\}} \left[ \sum_{m \in \mathbf{Z}^d \setminus \{0\}} (\chi_1 \cdot \ldots \cdot \chi_1 \cdot \chi_{\sigma})(M \mathbf{m} + v) \right]^k d\mu(M) dv$$

for $k = 0, 1, \ldots$. Taking the integral over $v$ inside the sum, we clear the way for Theorem 5 applied to

$$F(\mathbf{x}_1, \ldots, \mathbf{x}_d) = G_1(\mathbf{x}_1) \ldots G_1(\mathbf{x}_{d-1}) G_{\sigma}(\mathbf{x}_d)$$
where
\[ G_t(z) = \int_{y=0}^{1} \chi_t(z_1 + y) \ldots \chi_t(z_k + y) dy. \]

For any plane \( \pi' \) as in the Theorem, we have that
\[
\int_{\pi'} F(x) dx / (\text{covol } \pi)^d = \left( \frac{\int_{\pi} G_1(x_1) dx_1}{\text{covol } \pi} \right)^d \cdot \frac{\int G_\sigma(x_d) dx_d}{\int G_1(x_d) dx_d}. \tag{8}
\]

It is elementary to see that the quantity raised to the power \( d \) is at most one: the numerator is the volume of \( \pi \) “lying” inside the “crystal” shape, and \( \text{covol } \pi_Z \) is the volume of a fundamental domain. To wit, consider first the case when \( G_t(z) \) is replaced by the indicator of \([0, 1]^k\). Since vertices of any fundamental domain for \( \pi_Z \) have integer coordinates, it can completely cover the part of the plane inside the cube. Furthermore, the quantity in parentheses can equal one only when \( \pi \cap [0, 1]^k \) constitutes a fundamental domain for \( \pi_Z \). This means that there exists a \( Z \)-basis \( \{e_i\}_1^k \) for \( \pi_Z \) such that

- \( e_i \in \{0, 1\}^k, 1 \leq i \leq k; \)
- \( e_i + e_{i'} \in \{0, 1\}^k, 1 \leq i, i' \leq k. \)

Hence two distinct \( e_i, e_{i'} \) cannot take on the value 1 in the same coordinate. The same argument extends to other cubes of the form \([-s, 1-s]^k\) for \( s \in [0, 1] \) and so, too, for the original \( G_t(z) \) as it is an average over cubes of this kind.

The above argument shows that the limit as \( d \to \infty \) exists for each moment and that rate of convergence is exponential. To understand this limit, we focus on the terms with \( \int_{\pi} G(x_1) dx_1 / \text{covol } \pi = 1 \). Since in (7) we omit the terms in which any of \( m^l = 0 \), the only terms that survive after taking the limit are the ones with \( \pi \) generated by \( e_i \) for which \( \sum_{i=1}^k e_i = (1, \ldots, 1) \) (no zero coordinates). For planes \( \pi \) of fixed dimension \( l \) the number of possibilities is the number of partitions of a set of \( k \) elements into \( l \) non-empty subsets, which is exactly \( S(k, l) \). Finally observing that the last factor in (8) is \( \sigma^{\text{dim } \pi} = \sigma^l \), we find that the \( k \)-th moment tends to
\[
\sum_{l=1}^k S(k, l) \sigma^l,
\]
which is the corresponding moment of the Poisson distribution with parameter \( \sigma \).

In the case when \( \xi = p/q \in \mathbb{Q} \) we can modify the above proof. The integral over \( v \) becomes a finite sum, and we let \( G_t(z) = \frac{1}{q} \sum_{r=0}^{l-1} \chi_t(z_1 + r/q) \ldots \chi_t(z_d + r/q) \); the variable \( \tau \) appears in an equation similar to (8) and doesn’t enter the definition of \( G_t(z) \). The statements from the continuous version are true for this function as well (since it is also an average over cubes), and the proof is complete.

Generalizing this proposition we can obtain the statement of Theorem 2.
Proof of Theorem 2. What we need to show is that
\[
\int \int_{M \in \mathbb{E}^d} \sum_{V \in \mathbb{Z}^d \setminus \{0\}} \prod_{j=1}^{n} \chi_1 \cdot \ldots \cdot \chi_1 \cdot \chi_{(r^j, r^j + \tau^j)} (Mm^j + v^j) d\mu(M) dV
\]

has a limit as \( d \to \infty \) for every choice of \( k^1, \ldots, k^j \). If \( Y^{(d)}_{n, \sigma, \Xi, \tau / \Xi} = (Y^1, \ldots, Y^n) \) is distributed according to \( \mathbf{P}^{(d)}_{m, \sigma, \Xi, \tau / \Xi} \), then this expression is nothing more than the moment of order \( (k^1, \ldots, k^j) \).

Now we make a simplifying observation: we can assume that \( k^j = 1 \) for all \( j \) without loss of generality since taking all possible \( n \) and \( \Xi \) and computing \( \mathbf{E} \prod Y^j \) produces all the moments \( \mathbf{E} \prod (Y^j)^{k_j} \). That is, duplicating the random variable \( Y^j \) \( k^j \) times allows us to assume that \( k^j \) is 1. So we need to analyze
\[
\int \int_{M \in \mathbb{E}^d} \sum_{V \in \mathbb{Z}^d \setminus \{0\}} \prod_{j=1}^{n} \chi_1 \cdot \ldots \cdot \chi_1 \cdot \chi_{(r^j, r^j + \tau^j)} (Mm^j + v^j) d\mu(M) dV,
\]

which by Theorem 5 is
\[
\sum_{\pi \in \text{Gr}(r, l)(\mathbf{Q})} \int \int_{M \in \mathbb{E}^d} \frac{d\pi}{(\text{covol } \pi \mathbb{Z})^d} \prod_{j=1}^{n} \chi_1 \cdot \ldots \cdot \chi_{(r^j, r^j + \sigma^j)} (x^j + v^j) d\pi dV. \tag{9}
\]

Since \( m^j \) are non-zero, we exclude the “coordinate planes” as in Remark 3; this is denoted by the prime in the formula above.

We need to account for planes \( \pi \in \text{Gr}(r, l)(\mathbf{Q}) \) that will contribute in the limit \( d \to \infty \). By the argument from the previous proposition we have that
\[
\int_{\pi'} \prod_{j=1}^{n} \chi_1 \ldots \chi_{(r^j, r^j + \sigma^j)} (x^j + v^j) dx \leq (\text{covol } \pi \mathbb{Z})^d. \tag{10}
\]

Since \( \mathbb{E}^d \) is normalized to have measure 1, it suffices to study the integrand for fixed \( V \in \mathbb{E}^d \). If we can find \( V \) and \( \pi \) for which strict inequality is true in (10), then by continuity we have strict inequality for the integral over \( V \in \mathbb{E}^d \) and thus conclude that \( \pi \) doesn’t contribute in the limit. We will do this for \( V = 0 \) first. A plane that will contribute in the limit \( d \to \infty \) must satisfy the property that \( \pi \cap [0, 1]^r \) is a fundamental domain for \( \pi \cap \mathbb{Z}^r \) as in the previous proposition. For each of these planes we can try to find another \( V \) that gives strict inequality in (10). If \( V \in \pi \), we are translating the cube along the plane and thus getting the same cross-sectional area. So suppose \( V \in \mathbb{E} \setminus \pi \); this corresponds to cutting the cube with a plane parallel to \( \pi \). It is easy to see that for such planes the section will always have smaller area than the one for \( V \in \pi \). Thus it must be the case that \( \mathbb{E} \subset \pi \).

To summarize, a plane \( \pi \) contributes to the limit only if \( \pi \cap [0, 1]^r \) is a fundamental domain for \( \pi \cap \mathbb{Z}^r \) and \( \mathbb{E} \subset \pi \). This means that \( \mathbf{P}^{(d)}_{m, \sigma, \Xi, \tau / \Xi} \) has a limit as \( d \to \infty \) because all moments
exist. If we write $\xi = (\xi^1, \ldots, \xi^1, \xi^2, \ldots, \xi^2, \ldots, \xi^n, \ldots, \xi^n')$ reordering as necessary, then $\pi$ must a product of admissible planes for $(\xi^1, \ldots, \xi^1), (\xi^2, \ldots, \xi^2), \ldots, (\xi^n, \ldots, \xi^n')$. Hence the moment will split as the product of moments over distinct $\xi^i$. Using this observation and the previous proposition we see that in the case of distinct $\xi^i$ in Theorem 2 the limiting distribution is the product of independent Poisson distributions. If $\xi^j = \xi^j'$ but $(\tau^j, \tau^j + \sigma^j) \cap (\tau^{j'}, \tau^{j'} + \sigma^{j'}) = \emptyset$, then such $\xi^j$ and $\xi^{j'}$ behave as if they were unequal since the factor

$$\prod \chi_{(\tau^j, \tau^j + \sigma^j)}(x^j + v^j)$$

from (10) vanishes. It is evident that if $(\tau^j, \tau^j + \sigma^j) \cap (\tau^{j'}, \tau^{j'} + \sigma^{j'}) \neq \emptyset$ for some $j, j'$, then the limiting distribution cannot be a product of independent distributions. This concludes the proof of Theorem 2. 

References

[1] Jens Marklof. The $n$-point correlations between values of a linear form. *Ergodic Theory and Dynamical Systems*, 20(4):1127–1172, 2000. With an appendix by Zeév Rudnick.

[2] Jens Marklof. Distribution modulo one and Ratner’s theorem. In *Equidistribution in number theory, an introduction*, volume 237 of NATO Sci. Ser. II Math. Phys. Chem., pages 217–244. Springer, Dordrecht, 2007.

[3] Calvin C Moore. Ergodicity of flows on homogeneous spaces. *American Journal of Mathematics*, 88:154–178, 1966.

[4] Marina Ratner. On Raghunathan’s measure conjecture. *Annals of Mathematics. Second Series*, 134(3):545–607, 1991.

[5] C. A Rogers. The moments of the number of points of a lattice in a bounded set. *Philosophical Transactions of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 248:225–251, 1955.

[6] Nimish A Shah. Limit distributions of expanding translates of certain orbits on homogeneous spaces. *Indian Academy of Sciences. Proceedings. Mathematical Sciences*, 106(2):105–125, 1996.