AN ERROR TERM IN THE CENTRAL LIMIT THEOREM FOR SUMS OF DISCRETE RANDOM VARIABLES.

DMITRY DOLGOPYAT AND KASUN FERNANDO

ABSTRACT. We consider sums of independent identically distributed random variables those distributions have $d+1$ atoms. Such distributions never admit an Edgeworth expansion of order d but we show that for almost all parameters the Edgeworth expansion of order $d-1$ is valid and the error of the order $d-1$ Edgeworth expansion is typically of order $n^{-d/2}$.

1. INTRODUCTION.

Let X a random variable with zero mean and variance σ^2. Let $S_n = \sum_{j=1}^{n} X_j$ where X_j are independent identically distributed and have the same distribution as X. The Central Limit Theorem says that for each z

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{S_n}{\sigma\sqrt{n}} \le z\right) = \mathfrak{N}(z)$$

where

$$\mathfrak{N}(z) = \int_{-\infty}^{z} \mathfrak{n}(y)dy \text{ and } \mathfrak{n}(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}.$$

A classical problem in probability theory is computing higher order approximations to $\mathbb{P}\left(\frac{S_n}{\sigma\sqrt{n}} \le z\right)$. In particular, the order r Edgeworth series of S_n is an expression of the form

$$\mathcal{E}_r(z) = \mathfrak{N}(z) + \mathfrak{n}(z) \sum_{k=1}^{r} \frac{P_k(z)}{n^{k/2}}$$

2000 *Mathematics Subject Classification.* Primary 60F05; Secondary 11J13, 37A50.

there P_k are polynomials such that the characteristic function $\phi(t) = \mathbb{E}(e^{itX})$ and the Fourier transform $\hat{\mathcal{E}}_r$ of $\mathcal{E}_r$ satisfy

$$\phi\left(\frac{t}{\sigma\sqrt{n}}\right)^n - \hat{\mathcal{E}}_r(t) = o\left(n^{-r/2}\right).$$

In particular

$$\mathcal{E}_1(z) = \mathfrak{N}(z) + \mathfrak{n}(z)\frac{\mathbb{E}(X^3)}{6\sigma^3\sqrt{n}}(1-z^2),$$

$$\begin{aligned}\mathcal{E}_2(z) = \mathfrak{N}(z) + \mathfrak{n}(z)\Bigg[\frac{\mathbb{E}(X^3)}{6\sqrt{n}\sigma^3}(1-z^2) &+ \frac{\mathbb{E}(X^4)-3\sigma^4}{24n\sigma^4}(3z-z^3)\\ &-\frac{\mathbb{E}(X^3)^2}{72n\sigma^6}(15z-10z^3+z^5)\Bigg].\end{aligned}$$

We say that S_n admits an order r Edgeworth expansion if for all z

$$\lim_{n\to\infty} n^{r/2}\left[\mathbb{P}\left(\frac{S_n}{\sigma\sqrt{n}} \le z\right) - \mathcal{E}_r(z)\right] = 0. \tag{1.1}$$

It is known [6] that S_n admits an order 1 Edgeworth expansion if and only if X is non-lattice. The problem of higher order expansion is more complicated. For example, a sufficient condition for S_n to admit the order r Edgeworth expansion is that $\mathbb{E}(|X|^{r+2}) < \infty$ and X has a density. But this condition is far from necessary. We refer the reader to [7, Chapter XVI] for discussion of these and related results. We also note that [1] discusses a weak Edgeworth expansion where the the LHS of (1.1) is convolved with smooth compactly supported functions.

In this paper we consider a case which is opposite to X having a density, namely we suppose that X has a discrete distribution with $d+1$ atoms where $d \ge 2$. $d = 2$ is the simplest non-trivial case since the distributions with two atoms are lattice and as a result they do not admit even the first order Edgeworth expansion.

Thus we suppose that X takes values $a_1, \dots, a_{d+1}$ with probabilities $p_1, \dots, p_{d+1}$ respectively. Since X should have zero mean we suppose that our $2(d+1)-$tuple $(\mathbf{a}, \mathbf{p})$ belongs to the set

$$\Omega = \{p_i > 0, \quad p_1 + \dots + p_{d+1} = 1, \quad p_1 a_1 + \dots + p_{d+1} a_{d+1} = 0\}.$$

It is easy to see that S_n never admits the order d Edgeworth expansion. Indeed

$$\mathbb{P}_{\mathbf{a},\mathbf{p}}(S_n \le z) = \sum_{\substack{m_i\ge 0,\ \sum m_i = n\\ \sum m_i a_i \le z}} \frac{n!}{m_1! \dots m_{d+1}!} p_1^{m_1} \dots p_{d+1}^{m_{d+1}}. \tag{1.2}$$

Applying the Local Central Limit Theorem to the time homogeneous $\mathbb{Z}^d$-random walk which jumps to $\mathbf{e}_i$ from the origin $\mathbf{0}$ with probability p_i for $i = 1, \dots, d$ and stays at $\mathbf{0}$ with probability p_{d+1} we conclude that if

$$\sum m_i a_i = n \sum a_i p_i + \mathcal{O}(\sqrt{n})$$

then

$$n^{d/2} \frac{n!}{m_1! \dots m_{d+1}!} p_1^{m_1} \dots p_{d+1}^{m_{d+1}}$$

is uniformly bounded from below. Accordingly $\mathbb{P}_{\mathbf{a},\mathbf{p}}(S_n \leq z)$ has jumps of order $n^{-d/2}$. On the other hand $\mathcal{E}_d(z)$ is a smooth function of z. So, it can not approximate both $\mathbb{P}_{\mathbf{a},\mathbf{p}}(S_n \leq z - 0)$ and $\mathbb{P}_{\mathbf{a},\mathbf{p}}(S_n \leq z + 0)$ at the points of jumps.

In this paper we show that for typical $(\mathbf{a}, \mathbf{p})$ the order d Edgeworth expansion just barely fails. We present two results in this direction. For the first result let

$$b_j = a_j - a_1, \text{ for } j = 2 \dots d+1.$$

Set

$$d(s) = \max_{j \in \{2, \dots d+1\}} \operatorname{dist}(b_j s, 2\pi\mathbb{Z}).$$

We say that $\mathbf{a}$ is β-Diophantine if there is a constant K such that for $|s| > 1$

$$d(s) \geq \frac{K}{|s|^\beta}.$$

It is easy to see ([10, 14]) that almost all $\mathbf{a}$ is β-Diophantine provided that $\beta > (d-1)^{-1}$.

Theorem 1. *If* $\mathbf{a}$ *is* β*-Diophantine and*

$$2\left(R - \frac{1}{2}\right)\beta < 1 \tag{1.3}$$

then

$$\lim_{n\to\infty} n^R \left[\mathbb{P}_{\mathbf{a},\mathbf{p}}\left(\frac{S_n}{\sigma\sqrt{n}} \leq z\right) - \mathcal{E}_{d-1}(z)\right] = 0.$$

Thus for almost every $\mathbf{a}$ the order $(d-1)$ Edgeworth expansion approximates the distribution of $\frac{S_n}{\sigma\sqrt{n}}$ with error $\mathcal{O}(n^{\varepsilon - d/2})$ for any ε.

Note that Theorem 1 applies for all βs, in particular for βs which are much larger than $(d-1)^{-1}$. However if β is large, then the statement of the theorem can simplified. Namely, let r be the integer such that

$r < 2R \le r+1$. (Note that (1.3) can be rewritten as $2R < \dfrac{1}{\beta} + 1$ so provided that $2R$ is suffciently close to $\dfrac{1}{\beta}+1$ we can take $r = \left\langle \dfrac{1}{\beta} \right\rangle + 1$ where $\langle s \rangle$ denotes the largest integer which is strictly smaller than s.) Then,

$$\begin{aligned}\mathbb{P}_{\mathbf{a},\mathbf{p}}\left(\frac{S_n}{\sigma\sqrt{n}} \le z\right) &= \mathcal{E}_{d-1}(z) + \mathcal{O}\left(\frac{1}{n^R}\right)\\ &= \mathcal{E}_r(z) + \mathcal{O}\left(\frac{1}{n^R}\right) + \mathcal{O}\left(\mathcal{E}_{d-1}(z) - \mathcal{E}_r(z)\right).\end{aligned}$$

Since $\dfrac{r+1}{2} > R$ the first term dominates the second and we obtain the following result.

Corollary 1.1.

$$\lim_{n\to\infty} n^R \left[\mathbb{P}_{\mathbf{a},\mathbf{p}}\left(\frac{S_n}{\sigma\sqrt{n}} \le z\right) - \mathcal{E}_r(z)\right] = 0$$

provided that $\mathbf{a}$ *is* β*-Diophantine,* $r = 1 + \left\langle \dfrac{1}{\beta} \right\rangle$, *and* $r < 2R < \dfrac{1}{\beta} + 1$.

Theorem 1 shows that for almost every $\mathbf{a}$ and for $r \in \{1, \dots, d-1\}$, the order r Edgeworth expansion is valid. Our next results show that

$$\mathbb{P}_{\mathbf{a},\mathbf{p}}\left(\frac{S_n}{\sigma\sqrt{n}} \le z\right) - \mathcal{E}_d(z) \tag{1.4}$$

is typically of order $\mathcal{O}(n^{-d/2})$ but the $\mathcal{O}(n^{-d/2})$ terms has wild oscillations. To formulate this result precisely we suppose that our $2(d+1)$-tuple is chosen at random according to an absolutely continuous distribution $\mathbf{P}$ on Ω. Thus (1.4) becomes a random variable.

Theorem 2. *There exists a smooth function* $\Lambda(\mathbf{a},\mathbf{p})$ *such that for each* z *the random variable*

$$e^{z^2/2}\frac{n^{d/2}}{\Lambda(\mathbf{a},\mathbf{p})}\left[\mathcal{E}_d(z) - \mathbb{P}_{\mathbf{a},\mathbf{p}}\left(\frac{S_n}{\sigma\sqrt{n}} \le z\right)\right]$$

converges in law to a non-trivial random variable $\mathcal{X}$.

More precisely we have

$$\Lambda(\mathbf{a},\mathbf{p}) = \frac{|a_{d+1} - a_1|}{2^d \pi^{d+\frac{1}{2}} \sqrt{\det(D_{\mathbf{a},\mathbf{p}})}\ \ \sigma(\mathbf{a},\mathbf{p})} \tag{1.5}$$

where $D_{\mathbf{a},\mathbf{p}}$ is a $(d-1)\times(d-1)$ matrix defined by equations (5.10)–(5.11) of Section 5, $\sigma(\mathbf{a},\mathbf{p})$ denotes the standard deviation of the distribution of the random variable taking value a_j with probability p_j and $\mathcal{X}$ is defined as follows. Let $\mathbb{M}$ be the space of pairs $(\mathcal{L},\chi)$ where $\mathcal{L}$ is a unimodular lattice in $\mathbb{R}^d$ and χ is a homeomorphism $\chi:\mathcal{L}\to\mathbb{T}$. In the formulas below we identify $\mathbb{T}$ with segment $[0,1)$ equipped with addition modulo one. Given a vector $\mathbf{w}\in\mathbb{R}^d$ we denote by $y(\mathbf{w})$ its first coordinate and by $\mathbf{x}(\mathbf{w})$ its last $d-1$ coordinates.

Lemma 1.2. *For almost every pair $(\mathcal{L},\chi)\in\mathbb{M}$ with respect to the Haar measure the following limit exists*

$$\mathcal{X}(\mathcal{L},\chi)=\lim_{R\to\infty}\sum_{\mathbf{w}\in\mathcal{L}\setminus\{\mathbf{0}\},\ ||\mathbf{w}||\leq R}\frac{\sin(2\pi\chi(\mathbf{w}))}{y(\mathbf{w})}e^{-||\mathbf{x}(\mathbf{w})||^2}. \tag{1.6}$$

In order to simplify the notation we will abbreviate expressions such as (1.6) by

$$\mathcal{X}(\mathcal{L},\chi)=\sum_{\mathbf{w}\in\mathcal{L}\setminus\{\mathbf{0}\}}\frac{\sin(2\pi\chi(\mathbf{w}))}{y(\mathbf{w})}e^{-||\mathbf{x}(\mathbf{w})||^2}. \tag{1.7}$$

The Haar measure on $\mathbb{M}$ can be defined in two equivalent ways. First, note that χ is of the form $\chi(\mathbf{w})=e^{i\tilde{\chi}(\mathbf{w})}$ for some linear functional $\tilde{\chi}\in(\mathbb{R}^d)^*$. $SL_d(\mathbb{R})$ acts on $\mathbb{R}^d\oplus(\mathbb{R}^d)^*$ by the formula

$$A(\mathbf{w},\tilde{\chi})=(A\mathbf{w},\tilde{\chi}A^{-1}).$$

Observe that if $A(\mathbf{w},\tilde{\chi})=(\hat{\mathbf{w}},\hat{\chi})$ then

$$\tilde{\chi}(\mathbf{w})=\hat{\mathbf{w}}(\hat{\chi}). \tag{1.8}$$

The above action of $SL_d(\mathbb{R})$ induces the following action of $SL_d(\mathbb{R})\ltimes(\mathbb{R}^d)^*$ on $\mathbb{M}$

$$(A,\tilde{\chi})(\mathcal{L},\chi)=(A\mathcal{L},e^{2\pi it\tilde{\chi}}\cdot(\chi\circ A^{-1})).$$

This action is transitive because $SL_d(\mathbb{R})$ acts transitively on unimodular lattices and $(\mathbb{R}^d)^*$ acts transitively on characters. This allows to identify $\mathbb{M}$ with $(SL_d(\mathbb{R})\ltimes\mathbb{R}^d)/(SL_d(\mathbb{Z})\ltimes\mathbb{Z}^d)$ and so $\mathbb{M}$ inherits the Haar measure from $SL_d(\mathbb{R})\ltimes\mathbb{R}^d$.

The second way to define the Haar measure is to note that the space $\mathcal{M}$ of unimodular lattices is naturally identified with $SL_d(\mathbb{R})/SL_d(\mathbb{Z})$ and so it inherits the Haar measure from $SL_d(\mathbb{R})$. Next for a fixed $\mathcal{L}$ the set of homeomorphisms $\chi:\mathcal{L}\to\mathbb{T}$ is a d dimensional torus so it comes with its own Haar measure. Now if we want to compute the average of a function $\Phi(\mathcal{L},\chi)$ with respect to the Haar measure then we can first compute its average $\bar{\Phi}(\mathcal{L})$ in each fiber and then integrate the result with respect to the Haar measure on the space of lattices. In the proof of Lemma 1.2 given in Section 9 the averaging inside a fiber will be denoted by $\mathbf{E}_\chi$ and the averaging with respect to the Haar measure on the space of lattices will be denoted by $\mathbf{E}_{\mathcal{L}}$.

If we assume that the pair $(\mathcal{L}, \chi)$ is distributed according to the Haar measure on $\mathbb{M}$ then $\mathcal{X}$ becomes a random variable. This is the variable mentioned in Theorem 2.

Note that the distribution of $\mathcal{X}$ depends neither on $\mathbf{P}$ nor on z.

Using the second representation of the Haar measure we can also describe $\mathcal{X}$ as follows. Let $\mathbf{w}_1, \dots, \mathbf{w}_d$ be the shortest spanning set of $\mathcal{L}$. That is $\mathbf{w}_1$ is the shortest non zero vector in $\mathcal{L}$ and, for $j > 1$, $\mathbf{w}_j$ is the shortest vector which is linearly independent of $\mathbf{w}_1, \dots, \mathbf{w}_{j-1}$. Given $\mathbf{m} = (m_1, \dots, m_d) \in \mathbb{Z}^d$ let $(y, \mathbf{x})(\mathbf{m})$, $y \in \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^{d-1}$, denote the point

$$m_1\mathbf{w}_1 + \cdots + m_d\mathbf{w}_d \in \mathcal{L}. \tag{1.9}$$

Let $\theta_j = \chi(\mathbf{w}_j)$. Then θ_j are uniformly distributed on $\mathbb{T}$ and independent of each other. Set $\theta(\mathbf{m}) = m_1\theta_1 + \cdots + m_d\theta_d$. Now $\mathcal{X}$ can be rewritten as

$$\mathcal{X} = \sum_{\mathbf{m}\in\mathbb{Z}^d\setminus\{\mathbf{0}\}} \frac{\sin(2\pi\theta(\mathbf{m}))}{y(\mathbf{m})} e^{-||\mathbf{x}(\mathbf{m})||^2} \tag{1.10}$$

where $\mathcal{L}$ is uniformly distributed on the space of lattices, $(y, \mathbf{x})(\mathbf{m})$ is defined by (1.9), and $(\theta_1, \dots \theta_d)$ is uniformly distributed on $\mathbb{T}^d$ and independent of $\mathcal{L}$.

Theorems 1 and 2 have analogues in case we are interested in probability that S_n belongs to a finite interval. In particular our results have applications to the Local Limit Theorem.

Theorem 3. *Let $z_1(n)$ and $z_2(n)$ be two uniformly bounded sequences such that $|z_1(n) - z_2(n)|n^{d/2} \to \infty$. Then the random vector*

$$\frac{n^{d/2}}{\Lambda(\mathbf{a},\mathbf{p})}\left(e^{z_1^2/2}\left[\mathcal{E}_d(z_1) - \mathbb{P}_{\mathbf{a},\mathbf{p}}\left(\frac{S_n}{\sigma\sqrt{n}} \le z_1\right)\right], e^{z_2^2/2}\left[\mathcal{E}_d(z_2) - \mathbb{P}_{\mathbf{a},\mathbf{p}}\left(\frac{S_n}{\sigma\sqrt{n}} \le z_2\right)\right]\right) \tag{1.11}$$

converges in law to a random vector $(\mathcal{X}(\mathcal{L}, \chi_1), \mathcal{X}(\mathcal{L}, \chi_2))$ where $\mathcal{X}(\mathcal{L}, \chi)$ is defined by (1.7) *and the triple $(\mathcal{L}, \chi_1, \chi_2)$ is uniformly distributed on $(SL_d(\mathbb{R})/SL_d(\mathbb{Z})) \times \mathbb{T}^d \times \mathbb{T}^d$.*

Here and below the uniform distribution of $(\mathcal{L}, \chi_1, \chi_2)$ means that $\mathcal{L}$ is uniformly distributed on the space of lattices and for a given lattice

χ_1 and χ_2 are chosen independently and uniformly from the space of characters.

Theorem 4. *Let $z_1(n) < z_2(n)$ be two uniformly bounded sequences such that $l_n = z_2(n) - z_1(n) \to 0$.*

(a) If $l_n \geq Cn^{\varepsilon - d/2}$ for some $\varepsilon > 0$ then

$$\frac{\mathbb{P}_{\mathbf{a},\mathbf{p}}(z_1 < \frac{S_n}{\sigma\sqrt{n}} < z_2)}{l_n \mathfrak{n}(z_1)} \to 1$$

almost surely.

(b) If $l_n n^{d/2} \to \infty$ then

$$\frac{\mathbb{P}_{\mathbf{a},\mathbf{p}}(z_1 < \frac{S_n}{\sigma\sqrt{n}} < z_2)}{l_n \mathfrak{n}(z_1)} \Rightarrow 1$$

(here and below "$\Rightarrow$" denotes the convergence in law).

(c) If $l_n = \dfrac{c|a_{d+1} - a_1|}{\sigma(\mathbf{a}, \mathbf{p}) n^{d/2}}$ then

$$2^{d-\frac{3}{2}} \pi^d \sqrt{\det(D_{\mathbf{a},\mathbf{p}})} \left[\frac{\mathbb{P}_{\mathbf{a},\mathbf{p}}(z_1 < \frac{S_n}{\sigma\sqrt{n}} < z_2)}{l_n \mathfrak{n}(z_1)} - 1 \right] \Rightarrow \mathcal{Y}$$

where

$$\mathcal{Y}(\mathcal{L}, \chi, c) = \sum_{\mathbf{w} \in \mathbb{Z}^d \setminus \{\mathbf{0}\}} \frac{\sin(2\pi[\chi(\mathbf{w}) - cy(\mathbf{w})]) - \sin(2\pi\chi(\mathbf{w}))}{y(\mathbf{w})} e^{-||\mathbf{x}(\mathbf{w})||^2}$$

and $\mathcal{L}, \chi$ are as in Theorem 2.

The intuition behind this result is the following. Call y_n δ-plausible if $\mathbb{P}(S_n = y_n) \geq \delta n^{-d/2}$. The discussion following (1.2) shows that for each δ there are about $C(\delta) n^{d/2}$ δ-plausible values. Therefore, if $l_n \ll n^{-d/2}$ then the interval $[z_1(n), z_2(n)]$ would typically contain no plausible values. Hence, we should not expect the LLT to hold on that scale. Theorem 4 shows that as soon as interval $[z_1(n), z_2(n)]$ contains many plausible values then Local Limit Theorem typically holds for this interval.

Recall that

$$\mathbb{P}_{\mathbf{a},\mathbf{p}}(S_n \in [z_1, z_2]) = \sum_{\substack{m_i \geq 0,\, \sum m_i = n \\ z_1 \leq \sum m_i a_i \leq z_2}} \frac{n!}{m_1! \dots m_{d+1}!} p_1^{m_1} \dots p_{d+1}^{m_{d+1}}$$

so in Theorem 4 we just count the number of visits of a random linear form $\sum m_i a_i$ to a finite interval with weights given by multinomial coefficients. It is also interesting to consider counting with equal weight. In this case the analogue of Theorem 4(c) is obtained in [11] while for longer intervals only partial results are available, see [4, 8].

The layout of the paper is the following. Theorem 1 is proven in Section 2. The proof is a minor modification of the arguments of [7, Chapter XVI]. The bulk of the paper is devoted to the proof of Theorem 2. In Section 3 we provide an equivalent formula for $\mathcal{X}$. This formula looks more complicated than (1.7) but it is easier to identify with the limit of the error term. Section 4 contains preliminary reductions. We show that the density ρ on Ω could be assumed smooth and the integration in the Fourier inversion formula could be restricted to a finite domain. In Section 5 we show that main contribution to the error term comes from resonances where characteristic function of S_n is close to 1 in absolute value. The proof relies on several technical estimates which are established in Section 6. In Section 7 we use dynamics on homogenuous spaces in order to show that the contribution of resonances converges to (1.7) completing the proof of Theorem 2. The proofs of Theorems 3 and 4 are similar to the proof of Theorem 2. The necessary modifications are explained in Section 8. Finally, Section 9 contains the proof of Lemma 1.2.

2. Edgeworth Expansion under Diophantine conditions.

Theorem 1 is a consequence of Theorem 5 below and the fact that in our case there is a positive constant c such that

$$|\phi(s)| \leq 1 - cd(s)^2. \tag{2.1}$$

(2.1) follows from inequality (5.8) proven in Section 5.

Theorem 5. *If the distribution of X has $d+2$ moments and its characteristic function satisfies*

$$|\phi(s)| \leq 1 - \frac{K}{|s|^\gamma} \tag{2.2}$$

and $R < \dfrac{d}{2}$ is such that

$$\left(R - \frac{1}{2}\right)\gamma < 1 \tag{2.3}$$

then

$$\lim_{n\to\infty} n^R \left[\mathbb{P}_{\mathbf{a},\mathbf{p}} \left(\frac{S_n}{\sigma\sqrt{n}} \leq z \right) - \mathcal{E}_{d-1}(z) \right] = 0.$$

Theorem 5 follows easily from the estimates in [7, ChapterXVI] but we provide the proof here for completeness.

Proof. Denoting

$$\bar{\Delta}_n(\mathbf{a},\mathbf{p}) = \mathbb{P}_{\mathbf{a},\mathbf{p}} \left(\frac{S_n}{\sigma\sqrt{n}} \leq z \right) - \mathcal{E}_{d-1}(z)$$

we get by [7, Chapter XVI] that for each T

$$|\bar{\Delta}_n(\mathbf{a},\mathbf{p})| \leq \frac{1}{\pi} \int_{-\frac{T}{\sigma\sqrt{n}}}^{\frac{T}{\sigma\sqrt{n}}} \left| \frac{\phi^n(s) - \hat{\mathcal{E}}_{d-1}(s\sigma\sqrt{n})}{s} \right| ds + \frac{C}{T}. \tag{2.4}$$

Choose $T = Bn^R$ with $B = \dfrac{C}{\varepsilon}$. Then, $\dfrac{C}{T} = \dfrac{\varepsilon}{n^R}$. Take a small δ and split RHS of (2.4) into two parts.

$$\frac{1}{\pi} \int_{-\delta}^{\delta} \left| \frac{\phi^n(s) - \hat{\mathcal{E}}_{d-1}(s\sigma\sqrt{n})}{s} \right| ds + \frac{1}{\pi} \int_{\delta<|s|<Bn^{R-1/2}/\sigma} \left| \frac{\phi^n(s) - \hat{\mathcal{E}}_{d-1}(s\sigma\sqrt{n})}{s} \right| ds \tag{2.5}$$

Again by [7, Chapter XVI], we have that the first integral of (2.5) is $\mathcal{O}\left(n^{-d/2}\right)$. Also, $\displaystyle\int_{|s|>\delta} \left| \frac{\hat{\mathcal{E}}_{d-1}(s\sigma\sqrt{n})}{s} \right| ds$ has exponential decay as $n \to \infty$. Put $J = \{s : \ \delta < |s| < Bn^{R-1/2}/\sigma\}$. Thus we only need to approximate

$$\int_J \left| \frac{\phi^n(s)}{s} \right| ds \leq \frac{1}{\delta} \int_J |\phi^n(s)|\, ds \leq \frac{C}{\delta} \int_J \exp\left(-\bar{c}\, n^{1-\left(R-\frac{1}{2}\right)\gamma} \right) ds \tag{2.6}$$

where the last inequality is due to (2.2). By (2.3) the integral decay faster than any power of n. Because $R < \dfrac{d}{2}$ the contribution of $|s| \leq \delta$ is also under control. Hence, the result follows. □

3. Change of variables.

Here we deduce Theorem 2 from:

Theorem 2* *For each z the random variable*

$$n^{d/2}\left[\mathcal{E}_d(z)-\mathbb{P}_{\mathbf{a},\mathbf{p}}\left(\frac{S_n}{\sigma\sqrt{n}}\leq z\right)\right]$$

converges in law to $\hat{\mathcal{X}}$ where

(3.1)

$$\hat{\mathcal{X}}(\mathfrak{a},\mathfrak{p},\mathcal{L},\chi)=e^{-z^2/2}\frac{|\mathfrak{a}_{d+1}-\mathfrak{a}_1|}{2\sigma(\mathfrak{a},\mathfrak{p})\sqrt{\pi^3}}\sum_{\mathbf{w}\in\mathcal{L}\setminus\{\mathbf{0}\}}\frac{\sin 2\pi\chi(\mathbf{w})}{y(\mathbf{w})}e^{-4\pi^2\mathbf{x}(\mathbf{w})^T D_{\mathfrak{a},\mathfrak{p}}\mathbf{x}(\mathbf{w})}$$

$\mathfrak{a}=(\mathfrak{a}_1,\dots,\mathfrak{a}_{d+1})$, $\mathfrak{p}=(\mathfrak{p}_1,\dots,\mathfrak{p}_{d+1})$ *and* $(\mathfrak{a},\mathfrak{p})\in\Omega$ *are distributed according to the density* ρ *and* $D_{\mathfrak{a},\mathfrak{p}}$ *and* $\sigma(\mathfrak{a},\mathfrak{p})$ *are defined immediately after* (1.5).

In order to deduce Theorem 2 from Theorem 2* we need to show that $e^{z^2/2}\dfrac{\hat{\mathcal{X}}}{\Lambda(\mathfrak{a},\mathfrak{p})}$ has the same distribution as $\mathcal{X}$. To this end we rewrite the sum in (3.1) as

$$\frac{1}{(2\pi)^{d-1}\det(\sqrt{D_{\mathfrak{a},\mathfrak{p}}})}\sum_{\mathbf{w}\in\mathcal{L}-0}\frac{\sin(2\pi\chi(\mathbf{w}))}{y(\mathbf{w})/((2\pi)^{d-1}\det(\sqrt{D_{\mathfrak{a},\mathfrak{p}}}))}e^{-4\pi^2||(\sqrt{D_{\mathfrak{a},\mathfrak{p}}}\mathbf{x}(\mathbf{w}))||^2}. \tag{3.2}$$

Let A be the linear map such that

$$A(y,\mathbf{x})=\left(\frac{y}{(2\pi)^{d-1}\sqrt{\det(D_{\mathfrak{a},\mathfrak{p}})}},\ 2\pi\sqrt{D_{\mathfrak{a},\mathfrak{p}}}\ \ \mathbf{x}\right).$$

Put $(\bar{\mathcal{L}},\bar{\chi})=A(\mathcal{L},\chi)$. Then, using (1.8), (3.2) can be rewritten as:

$$\frac{1}{(2\pi)^{d-1}\det(\sqrt{D_{\mathfrak{a},\mathfrak{p}}})}\sum_{\bar{\mathbf{w}}\in\bar{\mathcal{L}}-0}\frac{\sin(2\pi\bar{\chi}(\bar{\mathbf{w}}))}{y(\bar{\mathbf{w}})}e^{-||\mathbf{x}(\bar{\mathbf{w}}))||^2}.$$

Since $\det(A)=1$, the pair $(\bar{\mathcal{L}},\bar{\chi})$ is distributed according to the Haar measure on $\mathbb{M}$ proving our formula for $\mathcal{X}$.

Sections 4–7 are devoted to the proof of Theorem 2*. Note that similarly to (1.10) we have

$$\hat{\mathcal{X}}=e^{-z^2/2}\frac{|\mathfrak{a}_{d+1}-\mathfrak{a}_1|}{2\sigma(\mathfrak{a},\mathfrak{p})\sqrt{\pi^3}}\sum_{\mathbf{m}\in\mathbb{Z}^d\setminus\{\mathbf{0}\}}\frac{\sin 2\pi\theta(\mathbf{m})}{y(\mathbf{m})}e^{-4\pi^2\mathbf{x}(\mathbf{m})^T D_{\mathfrak{a},\mathfrak{p}}\mathbf{x}(\mathbf{m})}.$$

The statements of Theorems 2 and 2* look similar, however, there is an important distinction. Namely the proof of Theorem 2* is constructive. In the course of the proof given n, $\mathbf{a}$ and z we construct a lattice $\mathcal{L}(\mathbf{a}, n)$ and a character $\chi(\mathbf{a}, \mathbf{p}, n, z)$ such that the expression $n^{-d/2}\hat{\mathcal{X}}(\mathbf{a}, \mathbf{p}, \mathcal{L}(\mathbf{a}, n), \chi(\mathbf{a}, \mathbf{p}, n, z))$ well-approximates the error in the Edgeworth expansion. We believe that such a construction could be made for more general distributions where the Edgeworth expansion fails, and this will be a subject of a future investigation. So the difference between Theorems 2 and 2* is that in the first case we have only an approximation in law while in the second case we are able to obtain an approximation in probability.

4. Cut off.

4.1. **Density.** Here we show that it is enough to prove Theorem 2 under the assumption that $\mathbf{P}$ has smooth density supported on a subset

$$\Omega_\kappa = \{(\mathbf{a}, \mathbf{p}) \in \Omega : \forall i \; p_i \geq \kappa \quad \text{and} \quad \forall i \neq j \; |a_i - a_j| \geq \kappa\}$$

for some $\kappa > 0$. Indeed suppose that the theorem is true for such densities. Let $p(\mathbf{a}, \mathbf{p})$ the original density of $\mathbf{P}$. Let ϕ be a bounded continuous test function. Given ε we can find a smooth density $\tilde{p}(\mathbf{a}, \mathbf{p})$ supported on some Ω_κ such that $||p - \tilde{p}||_{L^1} \leq \varepsilon$. In Section 7 we prove that

$$\int \phi(n^{d/2}\Delta_n)\tilde{p}\, d\mathbf{a}\, d\mathbf{p} \to \iint \phi(\hat{\mathcal{X}}(\mathbf{a}, \mathbf{p}, \mathcal{L}, \boldsymbol{\theta}))\tilde{p}\, d\mathbf{a}\, d\mathbf{p}\, d\mu(\mathcal{L}, \boldsymbol{\theta}) \tag{4.1}$$

where $\Delta_n = \mathcal{E}_d(z) - \mathbb{P}\left(\frac{S_n}{\sigma\sqrt{n}} \leq z\right)$ and μ is the Haar measure on $(SL_d(\mathbb{R})/SL_d(\mathbb{Z})) \times \mathbb{T}^d$. Let $p_m(\mathbf{a}, \mathbf{p})$ be the smooth density supported on Ω_κ corresponding to $\varepsilon = m^{-1}$. Passing to subsequence, $p_m \to p$ almost surely. Because $|p_m\phi| \leq \|\phi\||p_m| \in L^1$ and $|p\phi| \leq \|\phi\||p| \in L^1$ and $\|\phi\||p_m| \to \|\phi\||p|$ almost surely, Lebesgue Dominated Convergence Theorem gives

$$\begin{aligned} &\iint \phi(\hat{\mathcal{X}}(\mathbf{a}, \mathbf{p}, \mathcal{L}, \boldsymbol{\theta}))p_m\, d\mathbf{a}\, d\mathbf{p}\, d\mu(\mathcal{L}, \boldsymbol{\theta}) \\ &\qquad\qquad \to \iint \phi(\hat{\mathcal{X}}(\mathbf{a}, \mathbf{p}, \mathcal{L}, \boldsymbol{\theta}))p\, d\mathbf{a}\, d\mathbf{p}\, d\mu(\mathcal{L}, \boldsymbol{\theta}). \end{aligned} \tag{4.2}$$

Combining (4.1) and (4.2) we have that,

$$\begin{aligned}(4.3)\quad \int \phi(n^{d/2}\Delta_n)p\,d\mathbf{a}\,d\mathbf{p} &= \int \phi(n^{d/2}\Delta_n)p_m\,d\mathbf{a}\,d\mathbf{p} + \mathcal{O}(m^{-1}\|\phi\|)\\ &\xrightarrow{n\to\infty} \iint \phi(\hat{\mathcal{X}}(\mathbf{a},\mathbf{p},\mathcal{L},\boldsymbol{\theta}))p_m\,d\mathbf{a}\,d\mathbf{p}\,d\mu(\mathcal{L},\boldsymbol{\theta}) + \mathcal{O}(m^{-1}\|\phi\|)\\ &\xrightarrow{m\to\infty} \iint \phi(\hat{\mathcal{X}}(\mathbf{a},\mathbf{p},\mathcal{L},\boldsymbol{\theta}))p\,d\mathbf{a}\,d\mathbf{p}\,d\mu(\mathcal{L},\boldsymbol{\theta}).\end{aligned}$$

4.2. **Fourier transform.** As in the previous section let

$$\Delta_n = \mathcal{E}_d(z) - F_n(z) \quad \text{where} \quad F_n(z) = \mathbb{P}_{\mathbf{a},\mathbf{p}}\left(\frac{S_n}{\sigma\sqrt{n}} \le z\right).$$

Denote by $v_T(x) = \dfrac{1}{\pi}\cdot\dfrac{1-\cos Tx}{Tx^2}$ and let $\mathcal{V}(s,T) = \left(1 - \dfrac{|s|}{T}\right)\mathbb{1}_{|s|\le T}$ be its Fourier transform. Using the approach of [7, Section XVI.3] we let $T_2 = n^{2d+6}$ and decompose

$$(4.4)\quad \Delta_n = [\mathcal{E}_d(z) - F_n(z)] \star v_{T_2} - [F_n - F_n \star v_{T_2}](z) + [\mathcal{E}_d - \mathcal{E}_d \star v_{T_2}](z).$$

To estimate the last term we split

$$\begin{aligned}(4.5)\qquad [\mathcal{E}_d - \mathcal{E}_d \star v_{T_2}](z) &= \int_{|x|\le 1} [\mathcal{E}_d(z) - \mathcal{E}_d(z-x)]\, v_{T_2}(x)dx\\ &+ \int_{|x|\ge 1} [\mathcal{E}_d(z) - \mathcal{E}_d(z-x)]\, v_{T_2}(x)dx.\end{aligned}$$

Since v_T is even the first integral in (4.5) equals to

$$\begin{aligned}&\int_{|x|\le 1} \mathcal{E}_d'(z)x v_{T_2}(x)dx + \int_{|x|\le 1} \frac{\mathcal{E}_d''(y(z,x))}{2}x^2 v_{T_2}(x)dx\\ &= \int_{|x|\le 1} \frac{\mathcal{E}_d''(y(z,x))}{2}\frac{1-\cos T_2 x}{T_2}dx = \mathcal{O}\left(\frac{1}{T_2}\right).\end{aligned}$$

Since both $\mathcal{E}_d$ and cosine are bounded the second integral in (4.5) is bounded by

$$C\int_{|x|\ge 1} \frac{dx}{T_2 x^2} = \frac{C}{T_2}.$$

Thus the last term in (4.4) is $\mathcal{O}\left(T_2^{-1}\right)$. To estimate the second term in (4.4) we split the integral in $F_n \star v_{T_2}$ into regions $\{|x| \ge 1/\sqrt{T_2}\}$

and $\{|x| \le 1/\sqrt{T_2}\}$. The contribution of $\{|x| \ge 1/\sqrt{T_2}\}$ is bounded by

$$C\int_{1/\sqrt{T_2}}^{\infty} \frac{dx}{T_2 x^2} = \frac{C}{\sqrt{T_2}}.$$

On the other hand

$$\int_{|x|\le 1/\sqrt{T_2}} \left[F_n(z) - F_n(z-x)\right] V_{T_2}(x)dx = 0$$

unless there is a point of increase of F_n inside $\left[z - 1/\sqrt{T_2}, z + 1/\sqrt{T_2}\right]$. The probability that such a point exists is bounded by

(4.6)
$$\sum_{m_1+\dots+m_{d+1}=n} \mathbb{P}_{\mathbf{a},\mathbf{p}}\left(m_1 a_1 + \dots + m_{d+1}a_{d+1} \in \left[z - 1/\sqrt{T_2}, z + 1/\sqrt{T_2}\right]\right).$$

Note that for each fixed $(m_1, \dots, m_{d+1})$ the random variable

$$m_1 a_1 + \dots + m_{d+1}a_{d+1}$$

has a bounded density with respect to the uniform distribution on the segment of length $\mathcal{O}\left(\sqrt{m_1^2 + \dots + m_{d+1}^2}\right)$ and so

$$\mathbb{P}(\mathbf{m}.\mathbf{a} \in J) = \mathcal{O}\left(\frac{1}{\|\mathbf{m}\||J|}\right)$$

for any interval J. Hence each term in (4.6) is $\mathcal{O}\left(\frac{1}{n\sqrt{T_2}}\right)$ and so the sum is $\mathcal{O}\left(\frac{n^d}{n\sqrt{T_2}}\right)$. Thus with probability $1 - \mathcal{O}\left(\frac{1}{n^4}\right)$ we have that $\Delta_n = \Delta_{n,2} + o\left(T_2^{-1/2}\right)$ where

$$\begin{aligned}
\Delta_{n,2} &= \frac{1}{2\pi}\int_{-T_2}^{T_2} \frac{\left[\phi^n\left(\frac{t}{\sqrt{n}}\right) - \hat{\mathcal{E}}_d(t)\right]}{it}\mathcal{V}(t, T_2)e^{-itz}dt \\
&= \frac{1}{2\pi}\int_{-\frac{T_2}{\sigma\sqrt{n}}}^{\frac{T_2}{\sigma\sqrt{n}}} e^{-isz\sigma\sqrt{n}}\ \frac{\phi^n(s) - \hat{\mathcal{E}}_d(s\sigma\sqrt{n})}{is}\mathcal{V}(s, n, T_2)ds\,,
\end{aligned}$$

$\mathcal{V}(s, n, T) = 1 - \left|\frac{s\sigma\sqrt{n}}{T}\right|$ and $\phi(s)$ is the characteristic function given by

$$\phi(s) = p_1 e^{isa_1} + \dots + p_{d+1}e^{isa_{d+1}}.$$

Let $T_1 = K_1 n^{d/2}$ and define

$$\Delta_{n,1} = \frac{1}{2\pi} \int_{-\frac{T_1}{\sigma\sqrt{n}}}^{\frac{T_1}{\sigma\sqrt{n}}} e^{-isz\sigma\sqrt{n}} \ \frac{\phi^n(s) - \hat{\mathcal{E}}_d(s\sigma\sqrt{n})}{is} \ \mathcal{V}(s,n,T_2)\, ds.$$

Let $\Gamma_n = \Delta_{n,2} - \Delta_{n,1}$. Put

$$\tilde{\Gamma}_n = \frac{1}{2\pi} \int_{|s| \in \left[T_1/(\sigma\sqrt{n}), T_2/(\sigma\sqrt{n})\right]} e^{-isz\sigma\sqrt{n}} \ \frac{\phi^n(s)}{is} \ \mathcal{V}(s,n,T_2)\, ds.$$

Then, we have $\Gamma_n = \tilde{\Gamma}_n + \mathcal{O}\left(e^{-\varepsilon T_1^2}\right)$ due to the exponential decay of $\hat{\mathcal{E}}_d$.

The main result of Subsection 4.2 is the following.

Proposition 4.1.

$$\left\| \tilde{\Gamma}_n \right\|_{L^2} \leq \frac{C}{\sqrt{T_1 n^d}}. \tag{4.7}$$

Proof.

$$\mathbf{E}(\tilde{\Gamma}_n^2) = \iint \mathbf{E}\left(e^{-i(s_1+s_2)z\sigma\sqrt{n}} \phi^n(s_1)\phi^n(s_2)\mathcal{V}(s_1,n,T_2)\mathcal{V}(s_2,n,T_2)\right) \frac{ds_1}{s_1}\frac{ds_2}{s_2}.$$

We split this integral into two parts.

(1) In the region where $|s_1 + s_2| \leq 1$ we use Corollary 5.2 proven in Section 5 to estimate the integral by

$$\mathcal{O}\left(\int_{|s| \in \left[T_1/(\sigma\sqrt{n}), T_2/(\sigma\sqrt{n})\right]} \frac{1}{\sqrt{n} s_1^2} \mathbf{E}\left(|\phi^n(s_1)|\right)\, ds_1 \right). \tag{4.8}$$

The next result will be proven in Section 6.

Lemma 4.2.

$$\mathbf{E}\left(|\phi^n(s_1)|\right) \leq \frac{C}{n^{d/2}}.$$

Plugging the estimate of Lemma 4.2 into (4.8) and integrating we see that the contribution of the first region to $\mathbf{E}(\tilde{\Gamma}_n^2)$ is $\mathcal{O}\left(\frac{1}{T_1 n^{d/2}}\right)$.

(2) Consider now the region where $|s_1 + s_2| \geq 1$. Denote

$$b_{d+1} = a_{d+1} - a_1, \ \ldots, \ b_2 = a_2 - a_1.$$

Then

$$\phi(s) = e^{isa_1}\psi(s) \quad \text{where} \quad \psi(s) = p_1 + p_2 e^{isb_2} + \cdots + p_{d+1} e^{isb_{d+1}}.$$

Denote $\boldsymbol{\nu} = (p_1, \dots, p_d, b_2, \dots, b_d)$. Then there exists a compactly supported density $\rho = \rho(a_1, \boldsymbol{\nu})$ such that the contribution of the second region is

$$\iint_{|s_1+s_2|\geq 1} \left(\int e^{-i(s_1+s_2)z\sigma\sqrt{n}} e^{in(s_1+s_2)a_1} \psi^n(s_1)\psi^n(s_2)\mathcal{V}(s_1)\mathcal{V}(s_2)\rho \, da_1 \, d\boldsymbol{\nu} \right) \frac{ds_1}{s_1}\frac{ds_2}{s_2}.$$

We are able to use a $2d$-dimensional coordinate system because on Ω

$$p_1 + \cdots + p_{d+1} = 1, \quad \text{and} \quad p_1 a_1 + \cdots + p_{d+1} a_{d+1} = 0. \tag{4.9}$$

To estimate this integral we integrate by parts with respect to a_1. We use that

$$e^{isna_1} da_1 = \left[\frac{1}{isn} \frac{d}{da_1} \right]^k de^{isna_1}$$

for some large k (for example we can take $k = 2d + 1$). The integration by parts amounts to applying $\left(\frac{d}{da_1}\right)^k$ to $\left(e^{isz\sigma\sqrt{n}} \rho[\psi(s_1)\psi(s_2)]^n\right)$ which leads to the terms

$$\left\{ \left(\frac{d}{da_1}\right)^{k_1} \left[e^{i(s_1+s_2)z\sigma\sqrt{n}} \right] \right\} \left\{ \left(\frac{d}{da_1}\right)^{k_2} [\rho] \right\} \left\{ \left(\frac{d}{da_1}\right)^{k_3} [[\psi(s_1)\psi(s_2)]^n] \right\}$$

where $k_1+k_2+k_3 = k$. (Note that both σ and ψ depend on a_1 implicitly due to the second equation in (4.9)). Thus, the contribution of the above term to the integral is bounded by

$$C \iint_{\substack{|s_1|,|s_2|\in[T_1/\sigma\sqrt{n}, T_2/\sigma\sqrt{n}] \\ |s_1+s_2|\geq 1}} \frac{(s_1+s_2)^{k_1} n^{(k_1/2)+k_3}}{(s_1+s_2)^k n^k} \mathbf{E}\left(|\phi^n(s_1)|\right) \frac{ds_1}{s_1}\frac{ds_2}{s_2}.$$

Using Lemma 4.2 again we can estimate the above integral by

$$\begin{cases} \dfrac{C}{n^{k/2}} & \text{if} \quad k_1 \geq k-2 \\ \dfrac{C}{T_1 n^{k+d/2-k_1/2-k_3}} & \text{otherwise.} \end{cases}$$

Thus the main contribution comes from $k_1 = k_2 = 0$, $k_3 = k$ proving Proposition 4.1. $\square$

Proposition 4.1 shows that the contribution from $\tilde{\Gamma}_n$ to the L^2-limit of $n^{d/2}\Delta_n$ can be made arbitrarily small by choosing K_1 large. Also, on $|s| \leq T_1/\sigma\sqrt{n}$ we have

$$\mathcal{V}(s, n, T_2) = \left(1 - \frac{s\sigma\sqrt{n}}{T_2}\right) \mathbb{1}_{|s|<T_2/\sigma\sqrt{n}} = 1 - \frac{s\sigma}{n^{2d+\frac{11}{2}}}.$$

Hence $\Delta_{n,1} = \hat{\mathbf{\Delta}}_n + o(n^{-2d})$ where

$$\hat{\mathbf{\Delta}}_n := \frac{1}{2\pi} \int_{|s|\leq T_1/\sigma\sqrt{n}} \frac{\phi^n(s) - \hat{\mathcal{E}}_d(s\sigma\sqrt{n})}{is} e^{-isz\sigma\sqrt{n}} ds$$

approximates well $\Delta_{n,1}$ and hence, Δ_n too. Also, the error from this approximation of $n^{d/2}\Delta_n$ converges to 0 in L^2. Hence, we only need to analyze $n^{d/2}\hat{\mathbf{\Delta}}_n$ for large n.

5. Simplifying the error.

Denote

$$s_k = \frac{2\pi k}{|b_{d+1}|}$$

and let I_k be the segment of length $\dfrac{2\pi}{|b_{d+1}|}$ centered at s_k. Put $K_2 \gg K_1$. Due to the results of the previous section it is sufficient to study

$$\hat{\mathbf{\Delta}}_n = \sum_{|k|\leq K_2\sqrt{n}} \tilde{\mathcal{I}}_k$$

where

$$\tilde{\mathcal{I}}_k = \frac{1}{2\pi i} \int_{I_k} e^{-isz\sigma\sqrt{n}} \frac{\phi^n(s) - \hat{\mathcal{E}}_d(s\sigma\sqrt{n})}{s} ds.$$

$\tilde{\mathcal{I}}_0 = \mathcal{O}(n^{-(d+1)/2})$ due to [7, Section XVI.2]. Next, $\hat{\mathcal{E}}_d(s\sigma\sqrt{n})$ decays exponentially with respect to n outside of I_0. So, its contribution to $\tilde{\mathcal{I}}_k$ is negligible for $k \neq 0$. Accordingly,

$$\hat{\mathbf{\Delta}}_n = \sum_{0<|k|\leq K\sqrt{n}} \mathcal{I}_k + \mathcal{O}\left(\frac{1}{n^{(d+1)/2}}\right)$$

where

$$\mathcal{I}_k = \frac{1}{2\pi i} \int_{I_k} e^{-isz\sigma\sqrt{n}} \frac{\phi^n(s)}{s} \mathbb{1}_{|s|\leq T_1/\sigma\sqrt{n}} \, ds.$$

Introduce the following notation

$$\bar{s}_k = \arg\max_{s\in I_k} |\phi(s)|, \quad \phi(\bar{s}_k) = r_k e^{i\phi_k}.$$

The following lemma is similar to the results of [2, Section 5.2].

Lemma 5.1. *Suppose that*

$$r_k^n \geq n^{-100} \tag{5.1}$$

and

$$\pm\frac{T_1}{\sigma\sqrt{n}} \notin I_k. \tag{5.2}$$

Then

$$\mathcal{I}_k = \frac{1}{i\sqrt{\pi n}\sigma}\frac{r_k^n}{\bar{s}_k}e^{-z^2/2}\ e^{in\phi_k - i\bar{s}_k z\sigma\sqrt{n}}(1+o_{n\to\infty}(1)).$$

Proof. Let $e^{i\bar{s}_k a_j} = e^{i(\phi_k+\beta_j(k))}$. Then

$$r_k = \sum_{j=1}^{d+1} p_j \cos(\beta_j(k)) \tag{5.3}$$

and

$$\sum_{j=1}^{d+1} p_j \sin(\beta_j(k)) = 0. \tag{5.4}$$

Since (5.1) implies that $r_k \geq 1 - \dfrac{C\ln n}{n}$, (5.3) shows that $|\beta_j(k)| \leq C\sqrt{\dfrac{\ln n}{n}}$ and so (5.4) gives

$$\sum_{j=1}^{d+1} p_j\beta_j(k) = \mathcal{O}\left(\frac{\ln^{3/2} n}{n^{3/2}}\right). \tag{5.5}$$

Now we use Taylor expansion

$$e^{i(\bar{s}_k+\delta)a_j} = e^{i\phi_k}\left(1 + i\beta_j(k) - \frac{\beta_j(k)^2}{2}\right)\left(1 + ia_j\delta - \frac{a_j^2\delta^2}{2}\right) + \mathcal{O}\left(\frac{\ln^{3/2} n}{n^{3/2}} + \delta^3\right). \tag{5.6}$$

Thus,

$$\begin{aligned} \phi(\bar{s}_k+\delta) = &\quad e^{i\phi_k}\sum_{j=1}^{d+1} p_j\left(\cos(\beta_j(k)) - \frac{a_j^2\delta^2}{2}\right) + \mathcal{O}\left(\frac{\ln^{3/2} n}{n^{3/2}} + \delta^3\right) \\ = &\quad r_k e^{i\phi_k}\left(1 - \frac{\sigma^2\delta^2}{2}\right) + \mathcal{O}\left(\frac{\ln^{3/2} n}{n^{3/2}} + \delta^3\right) \end{aligned} \tag{5.7}$$

where we have used (5.5) as well as

$$p_1a_1 + \cdots + p_{d+1}a_{d+1} = 0, \quad p_1a_1^2 + \cdots + p_{d+1}a_{d+1}^2 = \sigma^2.$$

Hence for large n, the main contribution to $\mathcal{I}_k$ equals to

$$\frac{r_k^n}{2\pi i \bar{s}_k} e^{i(n\phi_k - \sqrt{n}\sigma z \bar{s}_k)} \int \left(1 - \frac{\sigma^2\delta^2}{2}\right)^n e^{-i\sigma z \delta\sqrt{n}} d\delta$$
$$\approx \frac{ir_k^n}{2\pi i \bar{s}_k} e^{i(n\phi_k - \sqrt{n}\sigma z \bar{s}_k)} \int e^{-\sigma^2\delta^2 n/2 - i\sigma\delta\sqrt{n} z} d\delta.$$

Making the change of variables $\sigma\delta\sqrt{n/2} = t$ we evaluate the last integral as $\dfrac{2\sqrt{\pi}e^{-z^2/2}}{\sigma\sqrt{n}}$. □

Corollary 5.2. *If I is a finite interval of order 1. Then*

$$\int_I |\phi^n(s)| \mathbb{1}_{|s| \le T_1/\sigma\sqrt{n}}\, ds = \mathcal{O}\left(\frac{1}{\sqrt{n}}\right).$$

Proof. We can cover I by a finite number of intervals I_k. The intervals where $r_k \le \dfrac{1}{n^{100}}$ contribute $\mathcal{O}\left(\dfrac{|I|}{n^{100}}\right)$ while the contribution of the intervals where $r_k \ge \dfrac{1}{n^{100}}$ is $\mathcal{O}\left(\dfrac{1}{\sqrt{n}}\right)$ due to Lemma 5.1. □

Because $r_k \approx 1$, $r_k = |\psi(\bar{s}_k)| = \left| p_1 + \sum_{j=2}^{d+1} p_j e^{ib_j \bar{s}_k} \right| \approx \sum p_j$. Therefore, $a_j \bar{s}_k \approx a_1 \bar{s}_k \mod 2\pi$ for all $j \ge 2$. Thus, $\dfrac{2\pi k b_j}{b_{d+1}} \approx 0 \pmod{2\pi}$ for all $2 \le j \le d$ and hence, $\phi(s_k) \approx 1$ which means s_k and $\bar{s}_k$ are close. Define, $\xi_k = \bar{s}_k - s_k$, $\eta_{j,k} = \dfrac{2\pi k b_j}{b_{d+1}} + 2\pi l_{j,k}$, $j = 1, \dots, d$ where $l_{j,k}$ is the unique integer such that $\dfrac{2\pi k b_j}{b_{d+1}} + 2\pi l_{j,k} \approx 0$. Then,

$$r_k^2 = \sum_{j=1}^{d+1} p_j^2 + 2 \sum_{l>j, j\neq 1} p_l p_j \cos[(b_l - b_j)\xi_k + \eta_{l,k} - \eta_{j,k}] + 2p_{d+1}p_1 \cos b_{d+1}\xi_k + 2\sum_{j=2}^{d} p_j p_1 \cos(b_j \xi_k + \eta_{j,k}). \tag{5.8}$$

Therefore

$$r_k^2 = 1 - \sum_{l>j, j\neq 1} p_l p_j [(b_l - b_j)\xi_k + \eta_{l,k} - \eta_{j,k}]^2 - p_{d+1} p_1 b_{d+1}^2 \xi_k^2 \\ - \sum_{j=2}^{d} p_j p_1 (b_j \xi_k + \eta_{j,k})^2 + \mathcal{O}\left(\xi_k^3 + \sum_{l=1}^{d} \eta_{l,k}^3\right).$$

Taking $\eta_{1,k} = b_1 = 0$ we can write the above as,

$$r_k^2 = -\xi_k^2 \sum_{l>j} p_l p_j (b_l - b_j)^2 - 2\xi_k \sum_{\substack{l>j \\ (l,j)\neq(d,1)}} p_l p_j (b_l - b_j)(\eta_{l,k} - \eta_{j,k}) \\ + 1 - \sum_{\substack{l>j \\ (l,j)\neq(d,1)}} p_l p_j (b_l - b_j)(\eta_{l,k} - \eta_{j,k})^2 + \mathcal{O}\left(\xi_k^3 + \sum_{l=1}^{d} \eta_{l,k}^3\right).$$

Since we have r_k^2 approximated by a quadratic polynomial of ξ_k (the unknown) we can approximate ξ_k by determining the maximizer of $r_k^2(\xi_k)$, obtaining

$$\xi_k = -\frac{\sum_{\substack{l>j \\ (l,j)\neq(d,1)}} p_l p_j (b_l - b_j)(\eta_{l,k} - \eta_{j,k})}{\sum_{l>j} p_l p_j (b_l - b_j)^2} + \mathcal{O}\left(\|\boldsymbol{\eta}_k|^2\right). \tag{5.9}$$

Substituting back we find r_k in terms of $\eta_{j,k}$ only. Ignoring higher order terms we compute the maximum to be:

$$r_k^2 = 1 - \sum_{\substack{l>j \\ (l,j)\neq(d,1)}} p_l p_j (b_l - b_j)(\eta_{l,k} - \eta_{j,k})^2 \\ + \frac{\left[\sum_{\substack{l>j \\ (l,j)\neq(d,1)}} p_l p_j (b_l - b_j)(\eta_{l,k} - \eta_{j,k})\right]^2}{\sum_{l>j} p_l p_j (b_l - b_j)^2} + \mathcal{O}\left(\sum_{l=1}^{d} \eta_{l,k}^3\right)$$

Put $R = \left[\sum_{l>j} p_l p_j (b_l - b_j)^2\right]^{-1}$. Then,

$$r_k^2 = 1 + \sum_{\substack{l>j\\(l,j)\neq(d,1)}} p_l p_j (b_l - b_j)\left[p_l p_j (b_l - b_j)R - 1\right](\eta_{l,k} - \eta_{j,k})^2$$

$$+ \sum_{\substack{l>j,m>n\\ l\neq m, j\neq n\\ (l,j),(m,n)\neq(d,1)}} p_l p_j p_m p_n (b_l - b_j)(b_m - b_n) R (\eta_{l,k} - \eta_{j,k})(\eta_{m,k} - \eta_{n,k}) + \mathcal{O}\left(\sum_{l>j} \eta_{l,j}^3\right)$$

$$:= 1 - 2\sum_{l,j=2}^{d} D_{l,j}(\mathbf{a},\mathbf{p})\eta_{l,k}\eta_{j,k} + \mathcal{O}\left(\sum_{l>j}\eta_{l,j}^3\right). \tag{5.10}$$

Thus,

$$r_k = 1 - \sum_{l,j=2}^{d} D_{l,j}(\mathbf{a},\mathbf{p})\eta_{l,k}\eta_{j,k} + \mathcal{O}\left(\sum_{l>j}\eta_{l,j}^3\right) = 1 - \boldsymbol{\eta}_k^T D_{\mathbf{a},\mathbf{p}} \boldsymbol{\eta}_k + \mathcal{O}(\|\boldsymbol{\eta}_k\|^3)$$

where $D_{\mathbf{a},\mathbf{p}}$ is a $(d-1)\times(d-1)$ matrix with

$$[D_{\mathbf{a},\mathbf{p}}]_{i,j} = D_{i,j}(\mathbf{a},\mathbf{p}) \tag{5.11}$$

and $\boldsymbol{\eta}_k^T = (\eta_{2,k}, \dots, \eta_{d,k})$. From this we have,

$$\mathcal{I}_k = \frac{e^{-z^2/2}}{i\sqrt{\pi n}\sigma} \frac{(1 - \boldsymbol{\eta}_k^T D_{\mathbf{a},\mathbf{p}}\boldsymbol{\eta}_k + \mathcal{O}(\|\boldsymbol{\eta}_k\|^3))^n}{s_k} e^{in\phi_k - i\tilde{s}_k z\sigma\sqrt{n}}(1 + o(1)).$$

Let $\mathcal{B}(\mathbf{a},\mathbf{p})$ be the contribution of the boundary terms $\pm\dfrac{T_1}{\sigma\sqrt{n}} \in I_k$.

Lemma 5.3.

$$\mathbf{E}(|\mathcal{B}|) \leq \frac{C}{n^{(2d-1)/2}}.$$

Lemma 5.4. *Let*

$$\mathcal{I}_{k,l} = \mathcal{I}_k \mathbb{1}_{|k|^\alpha n^{1/4}\|\boldsymbol{\eta}_k\| \in [2^l, 2^{l+1}]}.$$

with $\alpha = [2(d-1)]^{-1}$. *Then there is a constant* $\tilde{c}$ *such that*

$$\mathbf{E}\left(\sum_{0<|k|<Kn^{(d-1)/2}} \sum_{l>K} |\mathcal{I}_{k,l}|\right) = \mathcal{O}\left(\frac{1}{n^{d/2}} 2^K \exp(-\tilde{c}2^{2K})\right).$$

Lemmas 5.3 and 5.4 will be proven in Section 6.

Next we prove a lemma that would allow us to further simplify $\hat{\boldsymbol{\Delta}}_n$.

Lemma 5.5. *(a)* $\overline{s}_k = s_k + \boldsymbol{\omega}^T \boldsymbol{\eta}_k + \mathcal{O}(\|\eta\|_k^2)$ *where* $\boldsymbol{\omega} = \boldsymbol{\omega}(\mathbf{a}, \mathbf{p})$ *is a* $1 \times (d-1)$ *vector.*
(b) If $\|\boldsymbol{\eta}\| = \mathcal{O}\left(\frac{\ln n}{\sqrt{n}}\right)$ *then* $n\phi_k = ns_k a_1 + np_2\eta_{2,k} + \cdots + np_d\eta_{d,k} + o(1)$.

Proof. Since $\overline{s}_k - s_k = \zeta_k$ part (a) follows by (5.9).

Next, by (5.7)

$$\phi_k = \arg\, \phi(s_k) + \mathcal{O}\left(\delta^3 + \frac{\ln^{3/2} n}{n^{3/2}}\right)$$

Note that,

$$\phi(s_k) = e^{is_k a_1}(p_1 + p_2 e^{i\eta_{2,k}} + \cdots + p_d e^{i\eta_{d,k}} + p_{d+1}).$$

Thus,

$$\begin{aligned}\arg(\phi(s_k)) &= s_k a_1 + \tan^{-1}\left(\frac{p_2 \sin\eta_{2,k} + \cdots + p_d \sin\eta_{d,k}}{p_1 + p_2\cos\eta_{2,k} + \cdots + p_d\cos\eta_{d,k} + p_{d+1}}\right) \\ &= s_k a_1 + \sum_{l=2}^{d} p_l \eta_{l,k} + \mathcal{O}(\|\boldsymbol{\eta}_k\|^3)\end{aligned}$$

since the denominator in the first line is $1 + \mathcal{O}(||\boldsymbol{\eta}||^2)$. Now part (b) follows easily. $\square$

Now, we continue the analysis of the leading term in $\hat{\boldsymbol{\Delta}}_n$. Pick a small δ and define

$$A_1 = \{(\mathbf{a}, \mathbf{p}) |\, \mathcal{I}_{k,l} = 0\ \forall k, l \text{ s.t. } |k| < \delta n^{(d-1)/2} \text{ and } l < K\}.$$

Then

$$A_1^c = \{(\mathbf{a}, \mathbf{p}) |\, \exists |k| < \delta n^{(d-1)/2},\ |k|^\alpha n^{1/4}\|\eta_k\| \le 2^K\}.$$

Thus,

$$\mathbb{P}(A_1^c) = \sum_{|k| < \delta n^{(d-1)/2}} \frac{C 2^K}{|k|^{(d-1)\alpha} n^{(d-1)/4}} = \mathcal{O}(\sqrt{\delta} 2^K)$$

if $\alpha = \dfrac{1}{2(d-1)}$. Hence, for very large K and δ such that $\sqrt{\delta}2^K$ is very small, we can approximate Δ_n by the sum of $\mathcal{I}_k$'s with $\delta \le \dfrac{|k|}{n^{(d-1)/2}} \le K$ and $|k|^\alpha n^{1/4}\|\boldsymbol{\eta}_k\| \le 2^K$.

We define the random vector $X_k = \sqrt{n}\boldsymbol{\eta}_k$ and $Y_k = \dfrac{k}{n^{(d-1)/2}}$. Then, for large n, combining terms corresponding to k and $-k$ we obtain the following approximation to the distribution of Δ_n

$$\frac{|b_{d+1}|e^{-z^2/2}}{n^{d/2}\sigma\sqrt{\pi^3}} \sum_{k\in S(n,\delta,K)} \frac{\sin(n\phi_k - \overline{s}_k z\sigma\sqrt{n})}{Y_k} e^{-X_k^T D_{\mathbf{a},\mathbf{p}} X_k}$$

where $S(n,\delta,K) = \{k>0 | \delta < Y_k < K, \quad |Y_k|^\alpha \|X_k\| < 2^K\}$.

Define $\mathbf{q} = (p_2, \ldots, p_d)$. Then, Lemma 5.5 shows that

$$\begin{aligned} n\phi_k - \overline{s}_k z\sigma\sqrt{n} &= s_k(na_1 - z\sigma\sqrt{n}) + n\mathbf{q}^T\boldsymbol{\eta}_k - z\sigma\sqrt{n}\boldsymbol{\omega}^T\boldsymbol{\eta}_k + o(1) \\ &= \frac{2\pi n^{d/2}}{|b_{d+1}|}(\sqrt{n}a_1 - z\sigma)Y_k + (\sqrt{n}\mathbf{q} - z\sigma\boldsymbol{\omega})^T X_k + o(1). \end{aligned}$$

Therefore, for large n and K and δ such that $\sqrt{\delta}2^K$ is very small, the distribution of Δ_n is well approximated by

$$\tilde{\Delta}_n(\delta,K) = \frac{|b_{d+1}|e^{-z^2/2}}{n^{d/2}\sigma\sqrt{\pi^3}} \sum_{k\in S(n,\delta,K)} \frac{\sin\left(\frac{2\pi n^{d/2}}{|b_{d+1}|}(\sqrt{n}a_1 - z\sigma)Y_k + (\sqrt{n}\mathbf{q} - z\sigma\boldsymbol{\omega})^T X_k\right)}{Y_k} e^{-X_k^T D_{\mathbf{a},\mathbf{p}} X_k}.$$

6. Expectation of characteristic function.

Proof of Lemma 4.2. Recall that $d(s) = \max\limits_{2\le j\le d+1} d(b_j s, 0)$ where the distance is computed on the torus $\mathbb{R}/(2\pi\mathbb{Z})$. Formula (5.8) shows that there are positive constants C, c such that

$$\frac{1}{C} \le \frac{|\phi^n(s)|}{e^{-cnd(s)^2}} < C. \tag{6.1}$$

To prove the lemma we decompose $\mathbf{E}\left(e^{-cnd(s)^2}\right)$ into the pieces where $d(s)\sqrt{n}$ is of order 2^l for some $l \le (\log_2 n)/2$. Thus

$$\begin{aligned} \mathbf{E}\left(\phi^n(s)\right) &\le C\mathbf{P}\left(d(s) < \frac{1}{\sqrt{n}}\right) + C\sum_{l=0}^{(\log_2 n)/2} \mathbf{P}\left(d(s)\sqrt{n} \in [2^l, 2^{l+1}]\right) e^{-c4^l} \\ &\le \frac{C}{n^{d/2}} + C\sum_{l=0}^{(\log_2 n)/2} \frac{4^l}{n^{d/2}} e^{-c4^l} \le \frac{C}{n^{d/2}} \end{aligned}$$

completing the proof. □

Proof of Lemma 5.3. Let k be such that $\dfrac{T_1}{\sigma\sqrt{n}} \in I_k$. Then

$$\mathcal{I}_k = \int_{\pi(2k-1)/|b_3|}^{T_1/\sigma\sqrt{n}} e^{-isz\sigma\sqrt{n}} \frac{\phi^n(s)}{s}\, ds.$$

Because $T_1 = K_1 n^{d/2}$ and $s \in \left[\dfrac{\pi(2k-1)}{|b_3|}, \dfrac{T_1}{\sigma\sqrt{n}}\right]$ we have $s \approx n^{(d-1)/2}$. Thus

$$\mathbf{E}(|\mathcal{I}_k|) \leq \frac{C}{n^{(d-1)/2}} \mathbf{E}\left(\int_{\pi(2k-1)/|b_3|}^{T_1/\sigma\sqrt{n}} |\phi^n(s)|\, ds \right).$$

We claim that for all fixed b_d,

$$\iint e^{-cnd(s)^2}\, ds\, db_2 \dots\ db_{d-1} \leq \frac{C}{n^{d/2}}. \tag{6.2}$$

If this is true then using that ρ is a smooth compactly supported density of b_d we have that,

$$\begin{aligned}
\mathbf{E}\bigg(\int_{\pi(2k-1)/|b_3|}^{T_1/\sigma\sqrt{n}} |\phi^n(s)|\, ds \bigg) &= \iiint \int_{\pi(2k-1)/|b_3|}^{T_1/\sigma\sqrt{n}} |\phi^n(s)|\, ds\, db_d\, db_{d-1} \dots\ db_2 \\
&\leq C \iiint \int_{\pi(2k-1)/|x|}^{T_1/\sigma\sqrt{n}} e^{-cnd(s)^2} \rho(x)\, ds\, dx\, db_{d-1} \dots\ db_2 \\
&\leq C \int \iint e^{-cnd(s)^2}\, ds\, db_{d-1} \dots\ db_2\, \rho(x)\, dx \\
&\leq \frac{C}{n^{d/2}} \int \rho(x)\, dx = \mathcal{O}\left(\frac{1}{n^{d/2}}\right).
\end{aligned}$$

Thus

$$\mathbf{E}(|\mathcal{I}_k|) \leq \frac{C}{n^{(2d-1)/2}}. \tag{6.3}$$

Similarly, if $-\dfrac{T}{\sigma\sqrt{n}} \in \mathcal{I}_k$, then(6.3) holds. Hence, $\mathbf{E}(|\mathcal{B}|) \leq \dfrac{C}{n^{(2d-1)/2}}$ as required. To prove (6.2) we decompose it into pieces where $d(s)\sqrt{n}$ is of order 2^l. Taking μ to be the product measure $ds\, db_{d-1} \dots\ db_2$ from

(6.1) we have:

$$\iint e^{-cnd(s)^2}\, ds\, db_{d-1}\dots db_2 \le C\mu\{(s,b_2,\dots,b_{d-1})|d(s)<1/n\}$$
$$+C\sum_{l=0}^{(\log_2 n)/2} \mu\{(s,b_2,\dots,b_{d-1})|d(s)\sqrt{n}\in[2^l,2^{l+1}]\}e^{-c4^l}$$
$$\le \frac{C}{n^{d/2}}+C\sum_{l=0}^{(\log_2 n)/2}\frac{4^l}{n^{d/2}}e^{-c4^l}\le\frac{C}{n^{d/2}}$$

as required. □

Proof of Lemma 5.4. Because

$$r_k = 1-\boldsymbol{\eta}_k^T D_{\mathbf{a},\mathbf{p}}\boldsymbol{\eta}_k+\mathcal{O}(\|\boldsymbol{\eta}_k\|^3) \text{ and } |k|^\alpha n^{1/4}\|\boldsymbol{\eta}_k\|\in[2^l,2^{l+1}]$$

we can write

$$r_k = 1-c\frac{4^l}{|k|^{2\alpha}\sqrt{n}}+\mathcal{O}(n^{-3/4}).$$

Accordingly

$$r_k^n\le Ce^{-\frac{c2^{2l}\sqrt{n}}{|k|^{2\alpha}}}.$$

Also

$$\mathbf{P}(|k|^\alpha n^{1/4}\|\boldsymbol{\eta}\|\in[2^l,2^{l+1}])\le\frac{C2^l}{\sqrt{|k|}n^{(d-1)/4}}.$$

Hence,

$$\mathbf{E}(\mathcal{I}_{k,l})\le\frac{Ce^{-\frac{c2^{2l}\sqrt{n}}{|k|^{2\alpha}}}}{\sqrt{n}|k|}\frac{2^l}{\sqrt{|k|}n^{(d-1)/4}}=\frac{C2^le^{-\frac{c2^{2l}\sqrt{n}}{|k|^{2\alpha}}}}{|k|^{3/2}n^{(d+1)/4}}.$$

Thus

$$\sum_{l>K}\mathbf{E}(\mathcal{I}_{k,l})\le\frac{C2^Ke^{-\frac{c2^{2K}\sqrt{n}}{|k|^{2\alpha}}}}{|k|^{3/2}n^{(d+1)/4}}.$$

Therefore we need to estimate

$$\sum_{0<|k|<Kn^{(d-1)/2}}\frac{C2^Ke^{-\frac{c2^{2K}\sqrt{n}}{|k|^{2\alpha}}}}{|k|^{3/2}n^{(d+1)/4}}=$$

$$\frac{C}{n^{d/2}}\sum_{0<|k|<Kn^{(d-1)/2}}\frac{1}{|k|}\sqrt{\frac{2^{2K}n^{(d-1)/2}}{|k|}}e^{-\frac{c2^{2K}\sqrt{n}}{|k|^{2\alpha}}}. \tag{6.4}$$

Split the sum over

$$|k| \in \left[\frac{Kn^{(d-1)/2}}{2^{s+1}}, \frac{Kn^{(d-1)/2}}{2^s}\right] \tag{6.5}$$

for $s \in \mathbb{N}$. Then, for a fixed s we have

$$|k|^{2\alpha} = \mathcal{O}\left(\frac{K^{\frac{1}{d-1}}\sqrt{n}}{2^{\frac{s}{d-1}}}\right),$$

so each term in the sum (6.4) is of order

$$\frac{2^{K+(3s/2)}}{K^{3/2}n^{(d-1)/2}} \exp\left(-\frac{c2^{2K+\frac{s}{d-1}}}{K^{\frac{1}{d-1}}}\right).$$

But the number of such terms is of order $\dfrac{n^{(d-1)/2}}{2^s}$. Hence, the sum over k in (6.5) is $\mathcal{O}\left(\dfrac{2^{K+s/2}}{K^{3/2}} \exp\left(-\dfrac{c2^{2K+\frac{s}{d-1}}}{K^{\frac{1}{d-1}}}\right)\right)$. Summing over s we obtain the result. □

7. Relation to homogeneous flows.

Given $\mathbf{u} \in \mathbb{R}^{d-1}, v \in \mathbb{R}$ consider the following function on space $\mathcal{M}$ of unimodular lattices in $\mathbb{R}^d$:

$$\mathcal{Z}(L) = \sum_{(y,\mathbf{x})\in L\setminus\{\mathbf{0}\}} \frac{\sin 2\pi(\mathbf{u}^T\mathbf{x} + vy)}{y}\, e^{-4\pi^2\mathbf{x}^T D_{\mathbf{a},\mathbf{p}}\mathbf{x}}\, \mathbb{1}_{\{\delta<y<K, y^\alpha\|\mathbf{x}\|<2^K\}}. \tag{7.1}$$

Define $\boldsymbol{\gamma} = \dfrac{1}{k}\boldsymbol{\eta}$. Introduce the following matrices

$$H_{\boldsymbol{\gamma}} = \begin{pmatrix} 1 & \boldsymbol{\gamma} \\ \mathbf{0}^T & I_{d-1}\end{pmatrix}, \qquad G_t = \begin{pmatrix} e^{-(d-1)t} & \mathbf{0} \\ \mathbf{0}^T & e^t I_{d-1}\end{pmatrix}.$$

Then we get

$$n^{d/2}\tilde{\Delta}_n = -\frac{|b_{d+1}|e^{-z^2/2}}{\sigma\sqrt{\pi^3}}\mathcal{Z}(\mathbb{Z}^d\ H_{\boldsymbol{\gamma}}\ G_{\frac{\ln n}{2}}),$$

where

$$\mathbf{u} = \sqrt{n}\mathbf{q} - z\sigma\boldsymbol{\omega} \text{ and } v = \frac{n^{d/2}}{|b_{d+1}|}(\sqrt{n}a_1 - z\sigma)$$

and $\mathbf{q}$ and $\boldsymbol{\omega}$ are defined at the end of Section 5. Let $\mathcal{L}(n,\mathbf{a})$ be the unimodular lattice $\mathbb{Z}^d\ H_{\boldsymbol{\gamma}}\ G_{\frac{\ln(n)}{2}}$. Let

$$\mathbf{w}_j(n,\mathbf{a}) = (y_j(n,\mathbf{a}), \mathbf{x}_j(n,\mathbf{a})),\ j = 1, \ldots, d$$

with $y_j \in \mathbb{R}$ and $\mathbf{x}_j \in \mathbb{R}^{d-1}$ be the shortest spanning set of $\mathcal{L}$. Put,

$$\theta_j(n, (\mathbf{a}, \mathbf{p})) = \mathbf{u}^T \mathbf{x}_j(n, \mathbf{a}) + v y_j(n, \mathbf{a}), \ j = 1, \dots, d.$$

Proposition 7.1. *If $(\mathbf{a}, \mathbf{p})$ is distributed according to $\mathbf{P}$ then the distribution of the random vector*

$$((\mathbf{a}, \mathbf{p}), \mathcal{L}(n, \mathbf{a}), \boldsymbol{\theta}(n, (\mathbf{a}, \mathbf{p})))$$

converges to $\mathbf{P} \times \mu$ as $n \to \infty$, where μ is the Haar measure on

$$[SL_d(\mathbb{R})/SL_d(\mathbb{Z})] \times \mathbb{T}^d.$$

If we restrict our attention only to $((\mathbf{a}, \mathbf{p}), \mathcal{L}(n, \mathbf{a}))$ then the result is standard (see [12, Theorem 5.8], as well as [5, 9, 13]). The proof in the general case follows the approach of the proof of Proposition 5.1 in [3].

Proof. We need to show that for each bounded smooth test function f,

$$\int_\Omega f((\mathbf{a}, \mathbf{p}), \mathcal{L}(n, \mathbf{a}), \boldsymbol{\theta}) \, d\mathbf{P} \to \int_{\Omega \times \mathcal{M} \times \mathbb{T}^d} f((\mathbf{a}, \mathbf{p}), \mathcal{L}, \boldsymbol{\theta}) \, d\mathbf{P} \, d\mathcal{L} \, d\boldsymbol{\theta} \tag{7.2}$$

as $n \to \infty$. Write the Fourier series expansion of f:

$$f((\mathbf{a}, \mathbf{p}), \mathcal{L}(n, \mathbf{a}), \boldsymbol{\theta}) = \sum_{\mathbf{k}=(k_1,\dots,k_d) \in \mathbb{Z}^d} f_{\mathbf{k}}((\mathbf{a}, \mathbf{p}), \mathcal{L}(n, \mathbf{a})) \ e^{2\pi i \mathbf{k}^T \boldsymbol{\theta}}. \tag{7.3}$$

Then, it is enough to prove (7.2) for individual terms in (7.3).

If $\mathbf{k} = \mathbf{0}$ then by [12, Theorem 5.8] we can conclude that

$$\int_\Omega f_{\mathbf{0}}((\mathbf{a}, \mathbf{p}), \mathcal{L}(n, \mathbf{a})) \, d\mathbf{P} \to \int_{\Omega \times \mathcal{M} \times \mathbb{T}^d} f_{\mathbf{0}}((\mathbf{a}, \mathbf{p}), \mathcal{L}) \, d\mathbf{P} \, d\mathcal{L} \, d\boldsymbol{\theta}.$$

Now assume that $\mathbf{k} \neq \mathbf{0}$. Since Ω is $2d$ dimensional, we can use $(p_1, \dots, p_d, a_1, b_2, \dots, b_d)$ as local coordinates. In these coordinates $\mathcal{L}$ is independent of a_1. Hence, y_j's and $\mathbf{x}_j$'s are independent of a_1. Put $\boldsymbol{\nu} = (p_1, \dots, p_d, b_2, \dots, b_d)$. Then there exists a compactly supported density ρ such that,

$$\begin{aligned} J_{n,\mathbf{k}} &= \int f_{\mathbf{k}}((\mathbf{a}, \mathbf{p}), \mathcal{L}(n, \mathbf{a})) \ e^{2\pi i \mathbf{k}^T \boldsymbol{\theta}} \, d\mathbf{P} \\ &= \int f_{\mathbf{k}}((\mathbf{a}, \mathbf{p}), \mathcal{L}(n, \mathbf{a})) \exp 2\pi i \left(\sqrt{n} \sum k_j \mathbf{q}^T \mathbf{x}_j \right) \\ &\left[\int \rho(a_1, \boldsymbol{\nu}) \exp 2\pi i \left(\frac{n^{d/2}}{|b_{d+1}|} \left(\sqrt{n} a_1 - z\sigma \right) \sum y_j k_j - z\sigma \sum k_j \boldsymbol{\omega}^T \mathbf{x}_j \right) da_1 \right] d\boldsymbol{\nu}. \end{aligned} \tag{7.4}$$

Note that,

$$\int_{\mathbb{T}^d\times\Omega\times\mathcal{M}} f_{\mathbf{k}}((\mathbf{a},\mathbf{p}),\mathcal{L}(n,\mathbf{a}))\ e^{2\pi i\mathbf{k}^T\boldsymbol{\theta}}d\theta_1\dots\ d\theta_d\, d\mathbf{P}\, d\mathcal{L}=0$$

because

$$\int_{\mathbb{T}^d} e^{2\pi i\mathbf{k}^T\boldsymbol{\theta}}d\theta_1\dots d\theta_d=0.$$

Therefore, it is enough to prove that $J_{n,\mathbf{k}}$ converges to 0 as $n\to\infty$. To prove this we use integration by parts as follows. Put,

$$g(a_1,\boldsymbol{\nu})=\exp i\left(\frac{2\pi n^{(d+1)/2}\sum y_jk_j}{|b_{d+1}|}a_1\right)=\exp i\left(n^{(d+1)/2}\phi(\boldsymbol{\nu})a_1\right)$$

where $\phi(\boldsymbol{\nu})=\dfrac{2\pi\sum y_jk_j}{|b_{d+1}|}$ and,

$$h(a_1,\boldsymbol{\nu})=\rho(a_1,\boldsymbol{\nu})\exp\left[-i\left(\frac{2\pi n^{d/2}\sum y_jk_j}{|b_{d+1}|}+4\pi\sum k_j\boldsymbol{\omega}^T\mathbf{x}_j\right)z\sigma(a_1,\boldsymbol{\nu})\right]$$

Then, the inner integral in (7.4) is $\displaystyle\int g(a_1,\boldsymbol{\nu})h(a_1,\boldsymbol{\nu})\,da_1$. Let $\varepsilon>0$. On the set $Q_{\mathbf{k}}=\{\phi(\boldsymbol{\nu})>\varepsilon\}$ we can write

$$g(a_1,\boldsymbol{\nu})\,da_1=\frac{1}{i\phi(\boldsymbol{\nu})n^{(d+1)/2}}\,d\exp\left(ia_1n^{(d+1)/2}\phi(\boldsymbol{\nu})\right).$$

Integrating by parts on $Q_{\mathbf{k}}$ (note that h) has compact support, and using trivial bounds on $Q_{\mathbf{k}}^c$, we can conclude that

$$\begin{aligned}|J_{n,\mathbf{k}}|&\le\left|\int\frac{\exp\left(ia_1n^{(d+1)/2}\phi(\boldsymbol{\nu})\right)}{i\phi(\boldsymbol{\nu})n^{(d+1)/2}}\,h'(a_1,\boldsymbol{\nu})\,da_1\right|+C\mathbf{P}(\{\phi(\boldsymbol{\nu})\le\varepsilon\})\\&\le\frac{1}{\varepsilon n^{(d+1)/2}}\int|h'(a_1,\boldsymbol{\nu})|\,da_1+C\mathbf{P}(\{\phi(\boldsymbol{\nu})\le\varepsilon\})\end{aligned}$$

for small enough ε. But $h'(a_1,\boldsymbol{\nu})=\mathcal{O}(n^{d/2})$, hence the first term is $\mathcal{O}(1/\sqrt{n})$. Therefore, first taking $n\to\infty$ and then taking $\varepsilon\to 0$ we have the required result. □

Proposition 7.1 implies that as $n\to\infty$ the distribution of $n^{d/2}\tilde{\boldsymbol{\Delta}}_n(\delta,K)$ converges to the distribution of

$$e^{-z^2/2}\frac{|\mathfrak{a}_{d+1}-\mathfrak{a}_1|}{2\sigma(\mathfrak{a},\mathfrak{p})\sqrt{\pi^3}}\sum_{\mathbf{m}\in\mathbb{Z}^d\setminus\{\mathbf{0}\}}\frac{\sin 2\pi\theta(\mathbf{m})}{y(\mathbf{m})}e^{-4\pi^2\mathbf{x}^TD_{\mathfrak{a},\mathfrak{p}}\mathbf{x}}\mathbb{1}_{\{\delta<|y(\mathbf{m})|<K,\ |y(\mathbf{m})|^\alpha\|\mathbf{x}(\mathbf{m})\|<2^K\}}.\tag{7.5}$$

Next we let $\delta\to 0$ and $K\to\infty$ in such a way that $\sqrt{\delta}2^K\to 0$. Then, $\mathbb{1}_{\{\delta<|y(\mathbf{m})|<K,\ |y(\mathbf{m})|^\alpha|\mathbf{x}(\mathbf{m})|<2^K\}}\to 1$. Thus, (7.5) converges to $\hat{\mathcal{X}}$ proving Theorem 2*.

8. Finite Intervals.

The proofs of Theorems 3 and 4 are similar to the proof of Theorems 1 and 2 so we just explain the necessary changes leaving the details to the readers.

Proof of Theorem 3. The random vector (1.11) can be approximated by $(\mathcal{Z}^{(1)}, \mathcal{Z}^{(2)})$ where $\mathcal{Z}^{(i)}$ are defined as in (7.1) with $\mathbf{u}$ and v replaced by

$$\mathbf{u}^{(i)} = \sqrt{n}\mathbf{q} - z_i\sigma\boldsymbol{\omega} \quad \text{and} \quad v^{(i)} = \frac{n^{d/2}}{|b_{d+1}|}(\sqrt{n}a_1 - z_i\sigma)$$

respectively. Define $\boldsymbol{\theta}^{(i)}$ as in Proposition 7.1 but $\mathbf{u}$ and v replaced by $\mathbf{u}^{(i)}$ and $v^{(i)}$. To complete the proof we prove an analogue of Proposition 7.1. Namely that $((\mathbf{a},\mathbf{p}), \mathcal{L}(n,\mathbf{a}), \boldsymbol{\theta}^{(1)}(n,(\mathbf{a},\mathbf{p})), \boldsymbol{\theta}^{(2)}(n,(\mathbf{a},\mathbf{p})))$ converges to $\mathbf{P}\times\mu'$ as $n\to\infty$ where μ' is the Haar measure on $[SL_d(\mathbb{R})/SL_d(\mathbb{Z})]\times\mathbb{T}^d\times\mathbb{T}^d$.

As in the proof of Proposition 7.1 we prove that for individual terms in the Fourier series of a smooth function f on $[SL_d(\mathbb{R})/SL_d(\mathbb{Z})]\times\mathbb{T}^d\times\mathbb{T}^d$

$$\sum_{(\mathbf{k}_1,\mathbf{k}_2)\in\mathbb{Z}^d\times\mathbb{Z}^d} f_{\mathbf{k}_1,\mathbf{k}_2}((\mathbf{a},\mathbf{p}),\mathcal{L}(n,\mathbf{a}))\; e^{2\pi i[\mathbf{k}_1^T\boldsymbol{\theta}^{(1)}+\mathbf{k}_2^T(\boldsymbol{\theta}^{(1)}-\boldsymbol{\theta}^{(2)})]}$$

we have

$$J_{n,\mathbf{k}_1,\mathbf{k}_2} := \int_\Omega f_{\mathbf{k}_1,\mathbf{k}_2}((\mathbf{a},\mathbf{p}),\mathcal{L}(n,\mathbf{a}))e^{2\pi i[\mathbf{k}_1^T\boldsymbol{\theta}^{(1)}+\mathbf{k}_2^T(\boldsymbol{\theta}^{(1)}-\boldsymbol{\theta}^{(2)})]}\,d\mathbf{P}$$

$$\xrightarrow{n\to\infty} \int_{\Omega\times\mathcal{M}\times\mathbb{T}^d\times\mathbb{T}^d} f_{\mathbf{k}_1,\mathbf{k}_2}((\mathbf{a},\mathbf{p}),\mathcal{L})e^{2\pi i[\mathbf{k}_1^T\boldsymbol{\theta}_1+\mathbf{k}_2^T(\boldsymbol{\theta}_1-\boldsymbol{\theta}_2)]}\,d\mathbf{P}\,d\mathcal{L}\,d\boldsymbol{\theta}_1 d\boldsymbol{\theta}_2.$$

The case $\mathbf{k}_1=\mathbf{k}_2=0$ follows from [12, Theorem 5.8]. Note that

$$\mathbf{k}_2^T(\boldsymbol{\theta}^{(1)}-\boldsymbol{\theta}^{(2)})] = (z_2(n)-z_1(n))\left(\frac{2\pi n^{d/2}}{|b_{d+1}|}\sum y_j k_{2,j} + \sum k_{2,j}\boldsymbol{\omega}^T\mathbf{x}_j\right)\sigma.$$

If $\mathbf{k}_1=0$ choose appropriate local-coordinates in which σ is a coordinate. Integrating by parts with respect to $\sigma=\sigma(\mathbf{a},\mathbf{p})$ and using $|z_1(n)-z_2(n)|n^{d/2}\to\infty$ we see that $J_{n,\mathbf{0},\mathbf{k}_2}\to 0$ as $n\to\infty$.

If $\mathbf{k}_1\neq 0$ then using local coordinates $(a_1,\boldsymbol{\nu})$ as in the proof of Proposition 7.1 we can integrate by parts to conclude that $J_{n,\mathbf{k}_1,\mathbf{k}_2}\to 0$ as $n\to\infty$. The same proof follows through because the leading term of $\mathbf{k}_1^T\boldsymbol{\theta}^{(1)}+\mathbf{k}_2^T(\boldsymbol{\theta}^{(1)}-\boldsymbol{\theta}^{(2)})$ is still $n^{(d+1)/2}\phi(\boldsymbol{\nu})a_1$. $\square$

Proof of Theorem 4. To prove part (a) pick $\bar{\varepsilon} < \varepsilon$. Applying Theorem 1 we obtain that for almost every $(\mathbf{a}, \mathbf{p})$

$$\mathbb{P}_{(\mathbf{a},\mathbf{p})}\left(z_1 \leq \frac{S_n}{\sigma\sqrt{n}} \leq z_2\right) = \mathcal{E}_{d-1}(z_2) - \mathcal{E}_{d-1}(z_1) + \mathcal{O}\left(n^{-(d-\bar{\varepsilon})/2}\right)$$

$$= \mathfrak{n}(z_1) l_n + \mathcal{O}(l_n^2) + \mathcal{O}(l_n/\sqrt{n}) + \mathcal{O}\left(n^{-(d-\bar{\varepsilon})/2}\right).$$

According to the assumptions of part (a) the first term is much larger than the remaining terms proving the result.

The proof of part (b) is similar except that we apply Theorem 3 instead of Theorem 1 so we only get convergence in probability.

To prove part (c) we first prove the following analogue of Theorem 3 in case $z_2 = z_1 + \dfrac{c|a_{d+1} - a_1|}{n^{d/2}\sigma}$

$$\frac{n^{d/2}}{\Lambda(\mathbf{a},\mathbf{p})}\left(e^{z_1^2/2}\left[\mathcal{E}_d(z_1) - \mathbb{P}_{\mathbf{a},\mathbf{p}}\left(\frac{S_n}{\sigma\sqrt{n}} \leq z_1\right)\right], e^{z_2^2/2}\left[\mathcal{E}_d(z_2) - \mathbb{P}_{\mathbf{a},\mathbf{p}}\left(\frac{S_n}{\sigma\sqrt{n}} \leq z_2\right)\right]\right)$$

converges in law to a random vector $(\tilde{\mathcal{X}}_1, \tilde{\mathcal{X}}_2)(\mathcal{L}, \theta, c)$ where

$$(\tilde{\mathcal{X}}_1, \tilde{\mathcal{X}}_2)(\mathcal{L}, \theta, c) = \sum_{\mathbf{m} \in \mathbb{Z}^d \setminus \{\mathbf{0}\}} \frac{e^{-4\pi^2 ||\mathbf{x}(\mathbf{m})||^2}}{y(\mathbf{m})} \Big(\sin\theta(\mathbf{m}), \sin(\theta(\mathbf{m}) - cy(\mathbf{m}))\Big).$$

Once this convergence is established the proof of part (c) is the same as the proof of part (b). The proof of convergence is similar to the proof of Theorem 3 except that $\boldsymbol{\theta}^{(1)}$ and $\boldsymbol{\theta}^{(2)}$ are now not independent. Namely using the same notation as in the proof of Theorem 3 we have that $\mathbf{u}^{(2)} = \mathbf{u}^{(1)} + o(1)$, while $v^{(2)} = v^{(1)} - c + o(1)$. Following the same argument as in the proof of Proposition 7.1 we obtain that $(\mathcal{L}(n, \mathbf{a}), \boldsymbol{\theta}^{(1)}(n, \mathbf{a}), [\boldsymbol{\theta}^{(2)} - \boldsymbol{\theta}^{(1)}](n, \mathbf{a}))$ converges as $n \to \infty$ to $(\mathcal{L}^*, \boldsymbol{\theta}^*, \hat{\boldsymbol{\theta}}^*)$ where $(\mathcal{L}^*, \boldsymbol{\theta}^*)$ is distributed according to the Haar measure on $SL_d(\mathbb{R})/SL_d(\mathbb{Z}) \times \mathbb{T}^d$ and $\hat{\boldsymbol{\theta}}_j^* = \boldsymbol{\theta}_j^* - cy_j$. This justifies the formula for $(\tilde{\mathcal{X}}_1, \tilde{\mathcal{X}}_2)$. □

9. Convergence of $\mathcal{X}$.

We need some background information. Given a piecewise smooth function $g : \mathbb{R}^d \to \mathbb{R}$ of compact support its Siegel transform is a function on the space of lattices defined by

$$\mathcal{S}(g)(\mathcal{L}) = \sum_{\mathbf{w} \in \mathcal{L} \setminus \{\mathbf{0}\}} g(\mathbf{w}).$$

We need an identity of Rogers [11] saying that

$$\mathbf{E}_{\mathcal{L}}(\mathcal{S}(g)) = \int_{\mathbb{R}^d} g(\mathbf{w}) d\mathbf{w}. \tag{9.1}$$

In particular, if B is a set in $\mathbb{R}^d$ with piecewise smooth boundary not containing $\mathbf{0}$ then

$$\mathbf{P}_{\mathcal{L}}(\mathcal{L} \cap B \neq \emptyset) \leq \mathbf{P}(\mathcal{S}(\mathbb{1}_B)(\mathcal{L}) \geq 1) \leq \mathbf{E}_{\mathcal{L}}(\mathcal{S}(\mathbb{1}_B)) = \mathrm{Vol}(B). \tag{9.2}$$

Proof of Lemma 1.2. Let $\mathcal{L}^+ = \{\mathbf{w} \in \mathcal{L} : y(\mathbf{w}) > 0\}$. Since $\dfrac{\sin(2\pi\chi(\mathbf{w}))}{y(\mathbf{w})}$ is even it is enough to restrict the attention to $\mathbf{w} \in \mathcal{L}^+$.

Throughout the proof we fix two numbers $\varepsilon > 0, \tau < 1$ such that $\varepsilon \ll (1-\tau) \ll 1$.

It is easy to see using (9.2) and Borel-Cantelli Lemma that for almost every lattice $\mathcal{L}$, there exists C and β such that $y(\mathbf{w}) > \dfrac{C}{\|\mathbf{w}\|^\beta}$. It follows that

$$\sum_{\mathbf{w}\in\mathcal{L}^+:\, ||x(\mathbf{w})|| \geq ||\mathbf{w}||^\varepsilon} \frac{\sin 2\pi\chi(\mathbf{w})}{y(\mathbf{w})} e^{-||\mathbf{w}||^2} \leq \sum_{\mathbf{w}\in\mathcal{L}^+} C||\mathbf{w}||^\beta e^{-||\mathbf{w}||^\varepsilon}$$

converges absolutely. Hence it suffices to establish the convergence of

$$\bar{\mathcal{X}} := \sum_{\mathbf{w}\in\mathcal{L}^+:\, ||x(\mathbf{w})|| \leq ||\mathbf{w}||^\varepsilon < R^\varepsilon} \frac{\sin 2\pi\chi(\mathbf{w})}{y(\mathbf{w})} e^{-||\mathbf{w}||^2}.$$

Let $R_{j,k} = 2^k + j2^{\tau k}$, $j = 0, \dots 2^{1-\tau}k$. To prove the convergence of $\bar{\mathcal{X}}$ we will show that for all $\mathcal{L}$ almost all χ satisfy two estimates below

$$\forall \text{ sequence } \{j_k\}\ \bar{\mathcal{X}}_{R_{j_k,k}} \text{ converges as } k \to \infty, \tag{9.3}$$

$$\max_j \sup_{R_{j,k} \leq R \leq R_{j+1,k}} \left|\bar{\mathcal{X}}_R - \bar{\mathcal{X}}_{j,k}\right| \to 0 \text{ as } k \to \infty. \tag{9.4}$$

To prove (9.3) let

$$S_{j,k} = \sum_{\mathbf{w}\in\mathcal{L}^+:\, ||x(\mathbf{w})|| \leq ||\mathbf{w}||^\varepsilon,\, R_{j,k} \leq ||\mathbf{w}|| \leq R_{j+1,k}} \frac{\sin 2\pi\chi(\mathbf{w})}{y(\mathbf{w})} e^{-||\mathbf{w}||^2}.$$

Using that $\mathbf{E}_\chi(\sin(2\pi(\chi(\mathbf{w})))) = 0$ and for $\mathbf{w}_1 \neq \pm\mathbf{w}_2$ we have

$$\mathbf{E}_\chi(\sin(2\pi(\chi(\mathbf{w}_1))) \sin(2\pi(\chi(\mathbf{w}_2)))) = 0$$

we see that $\mathbf{E}_\chi(S_{j,k})=0$ and

$$\begin{aligned}\mathrm{Var}_\chi(S_{j,k}) &= \sum_{\mathbf{w}\in\mathcal{L}^+:\ ||x(\mathbf{w})||\leq ||\mathbf{w}||^\varepsilon, R_{j,k}\leq||\mathbf{w}||\leq R_{j+1,k}} \frac{e^{-2\|\mathbf{x}(\mathbf{w})\|^2}}{2y^2(\mathbf{w})}\\ &\leq \frac{1}{2^{2k+1}}\mathrm{Card}(\mathbf{w}: ||x(\mathbf{w})||\leq ||\mathbf{w}||^\varepsilon, R_{j,k}\leq||\mathbf{w}||\leq R_{j+1,k})\\ &\leq \frac{C(\mathcal{L})}{2^{2k}}\mathrm{Vol}(\mathbf{w}: ||x(\mathbf{w})||\leq ||\mathbf{w}||^\varepsilon, R_{j,k}\leq||\mathbf{w}||\leq R_{j+1,k})\\ &\leq C(\mathcal{L})2^{(\tau+\varepsilon(d-1)-2)k}.\end{aligned}$$

Hence by Chebyshev inequality for each j

$$\mathbf{P}_\chi\left(S_{j,k}\geq 2^{-(1-\tau)+\varepsilon)k}\right)\leq C(\mathcal{L})2^{(\varepsilon d-\tau)k}$$

and so

$$\mathbf{P}_\chi\left(\exists j:\ S_{j,k}\geq 2^{-(1-\tau+\varepsilon)k}\right)\leq C(\mathcal{L})2^{(1+\varepsilon d-2\tau)k}.$$

Thus if ε is sufficiently small and τ is sufficiently close to 1 then Borel-Cantelli Lemma shows that for almost every χ, if k is large enough, then for all j $S_{j,k}\leq 2^{-(1-\tau+\varepsilon)k}$ and thus $\sum_j S_{j,k}\leq 2^{-\varepsilon k}$ proving (9.3).

Likewise,

$$\begin{aligned}&\sup_{R_{j,k}\leq R\leq R_{j+1,k}}\left|\bar{\mathcal{X}}_R-\bar{\mathcal{X}}_{j,k}\right|\\ &\qquad\leq \sum_{\mathbf{w}\in\mathcal{L}^+:\ ||x(\mathbf{w})||\leq||\mathbf{w}||^\varepsilon, ||\mathbf{w}||\in[R_{j,k},R_{j+1,k}]}\frac{1}{|y(\mathbf{w})|}\ e^{-||x(\mathbf{w})||^2}\\ &\qquad\leq C(\mathcal{L})2^{-2k}\mathrm{Vol}(\mathbf{w}: ||x(\mathbf{w})||\leq ||\mathbf{w}||^\varepsilon, R_{j,k}\leq||\mathbf{w}||\leq R_{j+1,k})\\ &\qquad\leq \bar{C}(\mathcal{L})2^{\tau+\varepsilon(d-1)-1}\end{aligned}$$

proving (9.4). Lemma 1.2 is established. □

References

[1] Breuillard E. *Distributions diophantiennes et theoreme limite local sur* $\mathbb{R}^d$, Prob. Th. Related Fields **132** (2005) 39–73.

[2] Dolgopyat D. *A Local Limit Theorem for sum of independent random vectors,* Electronic J. Prob. 21 (2016) paper 39.

[3] Dolgopyat D., Fayad B. *Deviations of ergodic sums for toral translations: Convex bodies,* GAFA **24** (2014) 85–115.

[4] Dolgopyat D., Fayad B. *Limit theorems for toral translations,* Proc. Sympos. Pure Math **89** (2015) 227–277.

[5] Eskin A., McMullen C. *Mixing, counting, and equidistribution in Lie groups,* Duke Math. J. **71** (1993) 181–209.

[6] Esseen C.-G. *Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian law,* Acta Math. **77** (1945) 1–125.

[7] Feller W. *An introduction to probability theory and its applications,* Vol. II. 2nd ed. John Wiley & Sons, New York-London-Sydney (1971) xxiv+669 pp.

[8] H. Kesten *Uniform distribution* mod 1*,* part I: Ann. of Math. **71** (1960) 445–471, part II: Acta Arith. **7** (1961/1962) 355–380.

[9] Kleinbock D. Y., Margulis G. A. *Bounded orbits of nonquasiunipotent flows on homogeneous spaces,* AMS Transl. **171** (1996) 141–172.

[10] Kleinbock D. Y., Margulis G. A. *Logarithm laws for flows on homogeneous spaces,* Invent. Math. **138** (1999) 451–494.

[11] J. Marklof *The n-point correlations between values of a linear form,* Erg. Th., Dynam. Sys. **20** (2000) 1127–1172.

[12] Marklof J., Strombergsson A. *The distribution of free path lengths in the periodic Lorentz gas and related lattice point problems,* Ann. Math. **172** (2010) 1949–2033.

[13] Shah N. A. *Limit distributions of expanding translates of certain orbits on homogeneous spaces,* Proc. Indian Acad. Sci. Math. Sci. **106** (1996) 105–125.

[14] Sprindzuk V. G. *Metric theory of Diophantine approximations,* Scripta Series in Math. V. H. Winston & Sons, Washington, D.C. 1979, xiii+156 pp.