J. Inst. Math. Jussieu (2024), **23**(5), 2081–2125
doi:10.1017/S147474802300035X

HIGHER MOMENT FORMULAE AND LIMITING DISTRIBUTIONS OF LATTICE POINTS

MAHBUB ALAM[1], ANISH GHOSH[2] AND JIYOUNG HAN[3]

[1] *Department of Mathematics, Uppsala University, Sweden*
https://sites.google.com/view/mahbubweb
(mahbub.dta@gmail.com; mahbub.alam@math.uu.se)

[2] *School of Mathematics, Tata Institute of Fundamental Research, Homi Bhabha Road, Colaba, Mumbai, India 400005*
(ghosh@math.tifr.res.in)

[3] *Korea Institute for Advanced Study (KIAS), Seoul, Republic of Korea*
(hanjiwind@gmail.com)

(Received 13 July 2022; revised 28 September 2023; accepted 2 October 2023; first published online 28 November 2023)

Abstract We establish higher moment formulae for Siegel transforms on the space of affine unimodular lattices as well as on certain congruence quotients of $\mathrm{SL}_d(\mathbb{R})$. As applications, we prove functional central limit theorems for lattice point counting for affine and congruence lattices using the method of moments.

Contents

Key words and phrases: Lattice point counting; Geometry of numbers; Brownian motion; Central Limit Theorem; Siegel transform; Higher moments

2020 Mathematics subject classification: Primary 11N45
Secondary 22F30

1. Introduction

Let X_d denote the space of unimodular lattices in $\mathbb{R}^d$ which can be naturally identified with $\mathrm{SL}_d(\mathbb{Z})\backslash\mathrm{SL}_d(\mathbb{R})$ and denote by μ the Haar measure on X_d normalised to be a probability measure. Let $f:\mathbb{R}^d\to\mathbb{R}$ be a bounded function of compact support. The Siegel transform $\mathcal{S}_1(f)$ of f is defined by

$$\mathcal{S}_1(f)(\Lambda)=\sum_{\mathbf{m}\in\Lambda}f(\mathbf{m}),\ \Lambda\in\mathrm{SL}_d(\mathbb{Z})\backslash\mathrm{SL}_d(\mathbb{R}).$$

In [20], Siegel proved that

$$\int_{X_d}\mathcal{S}_1(f)d\mu=\int_{\mathbb{R}^d}f(x)dx+f(\mathbf{0}).$$

This result, often referred to as *Siegel's mean value formula*, is a fundamental result in the geometry of numbers and has proved to be indispensable in homogeneous dynamics, especially in applications to Diophantine problems. Following Siegel's result, Rogers [13] established intricate formulae for the higher moments of Siegel transforms (see Theorem 2.2 in Section 2). These formulae have since become an important tool in a wide variety of Diophantine problems. It is of considerable interest to prove analogues of Siegel's and Rogers' formulae for other homogeneous spaces. In this paper, we will establish explicit higher moment formulae for analogues of the Siegel transform on the following two homogeneous spaces, which are equipped with natural invariant probability measures μ_Y and μ_q on Y and $Y_{\mathbf{p}/q}$, respectively (see Section 2).

- The space $Y:=\mathrm{ASL}_d(\mathbb{Z})\backslash\mathrm{ASL}_d(\mathbb{R})$.
- The space $Y_{\mathbf{p}/q}:=\left\{\left(\mathbb{Z}^d+\dfrac{\mathbf{p}}{q}\right)g:g\in\mathrm{SL}_d(\mathbb{R})\right\}$, where $\mathbf{p}\in\mathbb{Z}^d\setminus\{\mathbf{0}\}$ and $q\in\mathbb{N}_{\geq 2}$ with $\gcd(\mathbf{p},q)=1$.

There have been many developments since Rogers' work; among those pertinent to the present paper is the recent work [8] of the third named author where S-arithmetic versions of Rogers' theorems are established. Analogues of Siegel transforms for Y and $Y_{\mathbf{p}/q}$ have been considered, and, in fact, a second moment formula has been obtained in each case – in the affine case, by El-Baz, Marklof and Vinogradov [4], where they were used to study the distribution of gaps between lattice directions (see also [2]), and in the congruence case, by Ghosh, Kelmer and Yu [6], where they were used to study effective versions of an inhomogeneous version of Oppenheim's conjecture on quadratic forms. In fact, they have other applications as well. We refer the reader to [1] for an application of the congruence second moment formula to Diophantine approximation and to [5] for an S-arithmetic version of the congruence second moment formula with applications to quadratic forms.

The main results in the present paper are formulas computing all the higher moments of Siegel transforms for both the affine and congruence cases. We also obtain analogues of a modification to Rogers' formula, due to Strömbergsson and Södergren [22]. Our proof of the higher moment formulae owes a lot to the breakthrough work of Marklof and Strömbergsson [11]. As will become clear, we make significant use of the ideas in

Section 7 of their paper. Our formulas are explicit but, as is the case with Rogers' original formula, are heavy on notation and need some buildup to state. We therefore postpone stating them to the next section. The reader will find the higher moment formula for Siegel transforms on Y in Theorem 2.12, and the formula for Siegel transforms on $Y_{\mathbf{p}/q}$ in Theorem 2.13. If history is a reliable guide, then our higher moment formulae will find good uses in counting problems. In the present paper, we provide applications to limiting distributions in lattice point counting problems. We devote the remainder of the introduction to discussing these applications.

1.1. Applications to counting results

Our counting results are inspired by the work of Strömbergsson and Södergren [22]. Given $d \geq 2$, a lattice $L \in X_d$ and a real number $x \geq 0$, set

$$N_{d,L}(x) := \#\left\{ m \in L \backslash \{0\} \ : \ |m| \leq \left(\frac{x}{V_d}\right)^{1/d} \right\},$$

where V_d denotes the volume of the unit ball in $\mathbb{R}^d$. Further, let

$$R_{d,L}(x) := N_{d,L}(x) - x$$

be the error term in the Gauss circle problem. Strömbergsson and Södergren proved several interesting results regarding the behaviour of $R_{n,L}$, including the following central limit theorem for a random lattice L.

Theorem (Strömbergsson and Södergren [22]). Let $\phi : \mathbb{Z}_+ \to \mathbb{R}_+$ be any function satisfying $\lim_{n\to\infty} \phi(d) = \infty$ and $\phi(d) = O_\varepsilon(e^{\varepsilon d})$ for every $\varepsilon > 0$. Let $Z_d^{(B)}$ be the random variable

$$Z_d^{(B)} := \frac{1}{\sqrt{2\phi(d)}} R_{d,L}(\phi(d))$$

with L picked at random in (X_d, μ). Then

$$Z_d^{(B)} \to \mathcal{N}(0,1) \text{ as } d \to \infty$$

in distribution.

Earlier, Södegren [23] studied the distribution of lengths of lattice vectors in a random lattice of large dimension. Strömbergsson and Södergren used the central limit theorem above in conjunction with Södegren's theorem to establish the following theorem indicating Poissonian behaviour for sequences growing sub-exponentially with respect to the dimension.

Theorem (Strömbergsson and Södergren [22]). Let $\phi : \mathbb{Z}_+ \to \mathbb{R}_+$ be any function satisfying $\lim_{n\to\infty} \phi(d) = \infty$ and $\phi(d) = O_\varepsilon(e^{\varepsilon d})$ for every $\varepsilon > 0$. Let $\mathcal{N}(x)$ be a Poisson distributed random variable with expectation $x/2$. Then

$$\mathrm{Prob}_\mu(N_{d,L}(x) \leq 2N) - \mathrm{Prob}(\mathcal{N}(x) \leq N) \to 0 \text{ as } d \to \infty,$$

uniformly with respect to all $N, x \geq 0$, satisfying $\min(x,N) \leq \phi(d)$.

More generally, they considered the case of several pairwise disjoint subsets and studied the joint distribution of the normalised counting variables and obtained a functional central limit theorem.

In this paper, we are concerned with two natural variations on this theme. Namely, we will consider the lattice point counting problem where the lattice is chosen at random from the spaces $(Y,\ \mu_Y)$ and $(Y_{\mathbf{p}/q},\ \mu_q)$.

We refer to these as the *affine* lattice point counting problem and the *congruence* lattice point counting problem, respectively. We prove analogues of the results of Strömbergsson and Södergren in the affine and congruence setting, and also analogues of results of Rogers [16], Schmidt [18] and Södergren [23] on Poissonian behaviour of lengths of lattice vectors in a randomly chosen lattice; see also related work of Kim [9]. The main tool in [22] is a version of Rogers' formula; in fact, one needs all moments, not just the second moment. In an analogous fashion, Theorems 2.12 and 2.13 will play a starring role in the proofs of the results stated below.

1.2. Counting results

Our first two results are analogues of Södergren's results [23] in the affine and congruence setting, respectively. For each $d \geq 2$, let $\mathcal{S} = \mathcal{S}_d = \{S_t : t \geq 0\}$ be an increasing family of subsets of $\mathbb{R}^d$ with $\mathrm{vol}(S_t) = t$, and for $\Lambda \in Y = \mathrm{ASL}_d(\mathbb{Z}) \setminus \mathrm{ASL}_d(\mathbb{R})$, set

$$N_t(\Lambda) := \#(S_t \cap \Lambda).$$

Denote by $\{N^\lambda(t) : t \geq 0\}$ a Poisson process on the non-negative real line with intensity λ.

Theorem 1.1. *The stochastic process $\{N_t(\Lambda) : t \geq 0\}$ converges weakly to $\{N^1(t) : t \geq 0\}$ as d goes to infinity.*

Let $q \in \mathbb{N}_{\geq 2}$ be given. For each $d \geq 2$, consider $\mathcal{S} = \mathcal{S}_d = \{S_t : t > 0\}$, an increasing family of subsets of $\mathbb{R}^d$ and $\mathbf{p}/q \in \mathbb{Q}^d$ for some $\mathbf{p} = \mathbf{p}_d \in \mathbb{Z}^d$ coprime with q. By abuse of notation, set

$$N_t(\Lambda) = \#(S_t \cap \Lambda),$$

for $\Lambda \in (Y_{\mathbf{p}/q},\ \mu_q)$.

Theorem 1.2.

(i) *For $q \geq 3$, the stochastic process $\{N_t(\Lambda) : t > 0\}$ converges weakly to $\{N^1(t) : t > 0\}$ as d goes to infinity.*

(ii) *For $q = 2$, assume that S_t's are symmetric about origin, and let $\widetilde{N}_t = \frac{1}{2} N_t$. Then the stochastic process $\left\{\widetilde{N}_t(\Lambda) : t > 0\right\}$ converges weakly to $\{N^{1/2}(t) : t > 0\}$ as d goes to infinity.*

Next, we establish a central limit theorem for the normalised error term in the lattice point problem for a random affine lattice.

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

Theorem 1.3. *Let $\phi:\mathbb{N}\to\mathbb{R}_{>0}$ be a function for which*

$$\lim_{d\to\infty}\phi(d)=\infty \quad \textit{and} \quad \phi(d)=O_\varepsilon(e^{\varepsilon d}),\ \forall\varepsilon>0. \tag{1.1}$$

Consider a sequence $\{S_d\}_{d\in\mathbb{N}}$ of Borel sets $S_d\subseteq\mathbb{R}^d$ such that $\mathrm{vol}(S_d)=\phi(d)$. *Let*

$$Z_d^1=\frac{\#(\Lambda\cap S_d)-\phi(d)}{\sqrt{\phi(d)}}$$

be the random variable with $\Lambda\in(Y,\mu_Y)$. Then

$$Z_d^1\to\mathcal{N}(0,1) \ \textit{as}\ d\to\infty$$

in distribution.

We now turn to the space $Y_{\mathbf{p}/q}$ which can be viewed as a finite volume homogeneous space of $\mathrm{SL}_d(\mathbb{R})$ (see Section 2.2) and therefore inherits a natural finite Haar measure μ_q.

Theorem 1.4. *Let a function $\phi:\mathbb{N}\to\mathbb{R}_{>0}$ and a sequence $\{S_d\}$ of Borel sets be given as in Theorem 1.3. When $q=2$, we further assume that each S_d is symmetric with respect to the origin. Let*

$$Z_d^{\mathbf{p}/q}=\begin{cases}\dfrac{\#(\Lambda\cap S_d)-\phi(d)}{\sqrt{2\phi(d)}}, & \textit{if } q=2;\\[2ex] \dfrac{\#(\Lambda\cap S_d)-\phi(d)}{\sqrt{\phi(d)}}, & \textit{otherwise}\end{cases}$$

be a random variable associated with $\Lambda\in(Y_{\mathbf{p}/q},\mu_q)$. Then

$$Z_d^{\mathbf{p}/q}\to\mathcal{N}(0,1) \ \textit{as}\ d\to\infty$$

in distribution.

The next two theorems are functional central limit theorems in the affine and congruence case respectively.

Theorem 1.5. *Let a function $\phi:\mathbb{N}\to\mathbb{R}_{>0}$ be given as in Theorem 1.3. Consider a sequence $\{S_d\}_{d\in\mathbb{N}}$ of star-shaped Borel sets $S_d\subseteq\mathbb{R}^d$ centered at the origin such that* $\mathrm{vol}(S_d)=\phi(d)$. *Let us define the random function*

$$t\in[0,1]\mapsto Z_d^1(t):=\frac{\#\left(\Lambda\cap t^{1/d}S_d\right)-t\phi(d)}{\sqrt{\phi(d)}},$$

where Λ is a random affine lattice in (Y,μ_Y). Here, $tS=\{t\mathbf{v}\in\mathbb{R}^d:\mathbf{v}\in S\}$ for any $t\in\mathbb{R}_{\geq 0}$ and $S\subseteq\mathbb{R}^d$. Then $Z_d^1(t)$ converges in distribution to one-dimensional Brownian motion as d goes to infinity.

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

Theorem 1.6. *Let a function $\phi : \mathbb{N} \to \mathbb{R}_{>0}$ and a sequence $\{S_d\}_{d\in\mathbb{N}}$ of Borel sets be as in Theorem 1.5. When $q = 2$, we further assume that each S_d is symmetric with respect to the origin. Define the random function*

$$t \in [0,1] \mapsto Z_d^{\mathbf{p}/q}(t) := \begin{cases} \dfrac{\#\left(\Lambda \cap t^{1/d} S_d\right) - t\phi(d)}{\sqrt{2\phi(d)}}, & \text{if } q = 2; \\ \dfrac{\#\left(\Lambda \cap t^{1/d} S_d\right) - t\phi(d)}{\sqrt{\phi(d)}}, & \text{otherwise.} \end{cases}$$

Then $Z_d^{\mathbf{p}/q}(t)$ converges in distribution to one-dimensional Brownian motion.

Structure of the paper

In Section 2, we state and prove the moment formulae for the affine and congruence cases. In fact, we provide two approaches, one kindly suggested to us by the referee. Section 3 is devoted to the study of Poissonian behaviour. In particular, analogues of results of Södergren [23] and Rogers [14, 15] in the affine and congruence setting are established. These results might be of independent interest. Section 4 contains affine and congruence versions of the variation on Rogers' formula developed by Strömbergsson and Södergren. Finally, Section 5 is devoted to the proofs of the counting results.

2. Higher Moment Formulae

We define

$$\mathrm{ASL}_d(\mathbb{R}) := \left\{ \begin{pmatrix} g & 0 \\ \xi & 1 \end{pmatrix} : g \in \mathrm{SL}_d(\mathbb{R}),\, \xi \in \mathbb{R}^d \right\}$$

and denote by (ξ, g) an element of $\mathrm{ASL}_d(\mathbb{R})$. One can identify the space of affine unimodular lattices with

$$Y_d = Y = \mathrm{ASL}_d(\mathbb{Z})\backslash\mathrm{ASL}_d(\mathbb{R})$$

via the map

$$\mathrm{ASL}_d(\mathbb{Z})(\xi, g) \mapsto \mathbb{Z}^d g + \xi.$$

We denote by μ_Y the Haar measure on $\mathrm{ASL}_d(\mathbb{R})$ normalised so that

$$\mu_Y(\mathrm{ASL}_d(\mathbb{Z})\backslash\mathrm{ASL}_d(\mathbb{R})) = 1.$$

Let $F : (\mathbb{R}^d)^k \to \mathbb{R}$ be a bounded function of compact support. Define the transform $\mathcal{S}_k(F)$ of F by

$$\mathcal{S}_k(F)(\Lambda) = \sum_{\substack{\mathbf{m}_i \in \Lambda \\ 1 \le i \le k}} F(\mathbf{m}_1, \ldots, \mathbf{m}_k),\ \Lambda \in \mathrm{ASL}_d(\mathbb{Z})\backslash\mathrm{ASL}_d(\mathbb{R}).$$

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

By a mild abuse of notation, we will use $\mathcal{S}_k(F)$ to also denote the function induced by the natural inclusion

$$\mathrm{SL}_d(\mathbb{Z})\backslash\mathrm{SL}_d(\mathbb{R}) \hookrightarrow \mathrm{ASL}_d(\mathbb{Z})\backslash\mathrm{ASL}_d(\mathbb{R}).$$

Notation 2.1. We follow Rogers [13] in setting some notation and recalling the definition of *admissible* matrices.

(1) We will identify the k-th power $(\mathbb{R}^d)^k$ of $\mathbb{R}^d$ with $\mathrm{Mat}_{k,d}(\mathbb{R})$. For a matrix D, denote by $[D]^j$ the j-th column of D and $[D]_i$ the i-th row of D.

(2) For $u\in\mathbb{N}$ and $r\in\{1,\ldots,k\}$, the collection $\mathfrak{D}^k_{r,u}$ is the set of integral matrices $D=(d_{ij})\in\mathrm{Mat}_{k,r}(\mathbb{Z})$ such that the greatest common divisor of all elements of D is one and there are $1\le i_1<\ldots<i_r\le k$ with the following properties:
   (i) ${}^{\mathrm{t}}([D]_{i_1},\ldots,[D]_{i_r})=u\mathrm{Id}_r$;
   (ii) $d_{ij}=0$ for $1\le j\le r$ and $1\le i<i_j$.
   We say that D is admissible if D satisfies the above properties.

(3) For each $D\in\mathfrak{D}^k_{r,u}$,
   (a) set $I_D:=\{i_1<\ldots<i_r\}$, where $i_1<\ldots<i_r$ are as above;
   (b) let

$$\Phi^{(d)}(D,u)=\left\{\begin{pmatrix}\mathbf{n}_1\\ \vdots\\ \mathbf{n}_r\end{pmatrix}\in(\mathbb{Z}^d)^r : \frac{D}{u}\begin{pmatrix}\mathbf{n}_1\\ \vdots\\ \mathbf{n}_r\end{pmatrix}\in(\mathbb{Z}^d)^k \quad\text{and}\quad \begin{matrix}\mathbf{n}_1,\ldots,\mathbf{n}_r\text{ are}\\ \text{linearly independent}\end{matrix}\right\};$$

   (c) define $N(D,u)$ to be the number of vectors $\mathbf{v}\in\{0,1\ldots,u-1\}^r$ for which

$$\frac{1}{u}D\,{}^{\mathrm{t}}\mathbf{v}\in\mathbb{Z}^k.$$

We are now ready to state Rogers' famous integral formula for $\mathcal{S}_k(F)$ on $\mathrm{SL}_d(\mathbb{Z})\backslash\mathrm{SL}_d(\mathbb{R})$ introduced in [13].

Theorem 2.2 (Rogers [13]). *Let $F:(\mathbb{R}^d)^k\to\mathbb{R}_{\ge0}$, where $1\le k\le d-1$, be a bounded function of compact support. Then,*

$$\int_{X_d}\mathcal{S}_k(F)(\Lambda)\,\mathrm{d}\mu(\Lambda)=F\begin{pmatrix}\mathbf{0}\\ \vdots\\ \mathbf{0}\end{pmatrix}+\sum_{r=1}^{k}\sum_{u\in\mathbb{N}}\sum_{D\in\mathfrak{D}^k_{r,u}}\frac{N(D,u)^d}{u^{dr}}\int_{(\mathbb{R}^d)^r}F\left(\frac{D}{u}\begin{pmatrix}\mathbf{v}_1\\ \vdots\\ \mathbf{v}_r\end{pmatrix}\right)\mathrm{d}\mathbf{v}_1\cdots\mathrm{d}\mathbf{v}_r.$$

We note that Rogers did not comment on the nature of convergence of the RHS of the above equation. He did, however, mention [13 , second paragraph of page 279] that results in another paper of his [14, §9] imply absolute convergence for $d\ge[\frac{1}{4}k^2]+2)$. Schmidt [17] showed that in the case of a bounded compactly supported function $F:(\mathbb{R}^d)^k\to\mathbb{R}_{\ge0}$, the above sum is absolutely convergent; in other words, both sides of the above equation are

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

finite (and equal). Thus, Rogers' theorem holds also for a bounded compactly supported function $F:(\mathbb{R}^d)^k \to \mathbb{R}$, and both sides of the above equation are finite in this case (since Rogers' theorem holds for $|F|$, we have absolute convergence of the sum, and we can rearrange the terms in the sum).

Theorem 2.2 follows from the fact that

$$(\mathbb{Z}^d)^k = \{{}^{\mathrm{t}}(\mathbf{0},\ldots,\mathbf{0})\} \sqcup \bigsqcup_{r=1}^{k} \bigsqcup_{u\in\mathbb{N}} \bigsqcup_{D\in\mathfrak{D}^k_{r,u}} \frac{D}{u}\Phi^{(d)}(D,u)$$

and the following proposition.

Proposition 2.3 (Rogers [13]). *Let $F:(\mathbb{R}^d)^k \to \mathbb{R}$ be a bounded function of compact support. For each $D\in\mathfrak{D}^k_{r,u}$, we have*

$$\int_{X_d} \sum_{\substack{{}^{\mathrm{t}}(\mathbf{n}_1,\ldots,\mathbf{n}_r)\\ \in\Phi^{(d)}(D,u)}} F\left(\frac{D}{u}\begin{pmatrix}\mathbf{n}_1\\ \vdots\\ \mathbf{n}_r\end{pmatrix} g\right) \mathrm{d}\mu(g) = \frac{N(D,u)^d}{u^{dr}} \int_{(\mathbb{R}^d)^r} F\left(\frac{D}{u}\begin{pmatrix}\mathbf{v}_1\\ \vdots\\ \mathbf{v}_r\end{pmatrix}\right) \mathrm{d}\mathbf{v}_1\cdots\mathrm{d}\mathbf{v}_r.$$

2.1. Higher moment formulae for Y

In [4], El-Baz, Marklof and Vinogradov established a second moment formula for the Siegel transform on $Y = \mathrm{ASL}_2(\mathbb{Z})\backslash\mathrm{ASL}_2(\mathbb{R})$ which easily extends to the case when $d\geq 3$ (see [4, Appendix B]). We will generalise their result to higher moment formulae for the transform $\mathcal{S}_k(\cdot)$ on Y. It is well-known that

$$\bigcup_{g\in\mathcal{F}} \{(\xi,g) : \xi\in[0,1)^d g\}$$

is a fundamental domain for Y, where $\mathcal{F}$ is any fixed fundamental domain for $\mathrm{SL}_d(\mathbb{Z})\backslash\mathrm{SL}_d(\mathbb{R})$. Thus, one can take the probability $\mathrm{ASL}_d(\mathbb{R})$-invariant measure μ_Y on Y as the measure inherited from the product of the Haar measure μ on $\mathrm{SL}_d(\mathbb{R})$ and the Lebesgue measure on $\mathbb{R}^d$.

Theorem 2.4. *Let $F:(\mathbb{R}^d)^k \to \mathbb{R}$ be a bounded compactly supported function, and $d\geq 2$. We have the following:*

(i) *For $k=1$,*

$$\int_Y \mathcal{S}_1(F)(\Lambda)\,\mathrm{d}\mu_Y(\Lambda) = \int_{\mathbb{R}^d} F(\mathbf{y})\,\mathrm{d}\mathbf{y}. \tag{2.1}$$

(ii) *For $2 \le k \le d$,*

$$\int_Y \mathcal{S}_k(F)(\Lambda)\,\mathrm{d}\mu_Y(\Lambda) = \int_{(\mathbb{R}^d)^k} F\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_k \end{pmatrix} \mathrm{d}\mathbf{y}_1\,\mathrm{d}\mathbf{y}_2\cdots\mathrm{d}\mathbf{y}_k + \int_{\mathbb{R}^d} F\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_1 \end{pmatrix} \mathrm{d}\mathbf{y}_1$$
$$+\sum_{r=1}^{k-2}\sum_{u\in\mathbb{N}}\sum_{D\in\mathfrak{D}_{r,u}^{k-1}} \frac{N(D,u)^d}{u^{dr}} \int_{(\mathbb{R}^d)^{r+1}} F\left(D'\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_{r+1} \end{pmatrix}\right) \mathrm{d}\mathbf{y}_1\,\mathrm{d}\mathbf{y}_2\cdots\mathrm{d}\mathbf{y}_{r+1}, \tag{2.2}$$

where D' for $D \in \mathfrak{D}_{r,u}^{k-1}$ is $k\times(r+1)$ matrix defined by

$$D' = \left(\begin{array}{c|c} 1 & 0\cdots 0 \\ \hline 1 & \\ \vdots & \dfrac{D}{u} \\ 1 & \end{array}\right). \tag{2.3}$$

Here, as a convention, for $k=2$, let us assume that $\sum_{r=1}^{0}$ is the empty summation.

Finally, both sides of the equation (2.2) are finite.

Proof. We first remark that the $k=1$ case is classical and can be proved using the *folding and unfolding* argument. When $k=2$, the result can be deduced from [4, Appendix B], where the authors proved the second moment formula for $d=2$. However, their proof can be seen to work in full generality. We will therefore focus on the case when $k\ge 3$.

Fix any fundamental domain $\mathcal{F}$ for $\mathrm{SL}_d(\mathbb{Z})\backslash\mathrm{SL}_d(\mathbb{R})$. For each $g\in\mathcal{F}$, by the change of variables $\xi=\eta g$, we have

$$\int_Y \mathcal{S}_k(F)(\mathbb{Z}^d g+\xi)\,\mathrm{d}\mu(g)\,\mathrm{d}\xi = \int_Y \mathcal{S}_k(F)((\mathbb{Z}^d+\eta)g)\,\mathrm{d}\mu(g)\,\mathrm{d}\eta$$
$$= \int_{\mathcal{F}}\int_{[0,1)^d} \sum_{\substack{\mathbf{m}_i\in\mathbb{Z}^d \\ 1\le i\le k}} F\begin{pmatrix} (\mathbf{m}_1+\eta)g \\ (\mathbf{m}_2+\eta)g \\ \vdots \\ (\mathbf{m}_k+\eta)g \end{pmatrix} \mathrm{d}\eta\,\mathrm{d}\mu(g).$$

For each $g\in\mathcal{F}$ and $\mathbf{m}_1\in\mathbb{Z}^d$, put $\mathbf{y}_1=(\eta+\mathbf{m}_1)g$ and $\mathbf{m}_j'=\mathbf{m}_j-\mathbf{m}_1$ for $2\le j\le k$. Since $\bigcup_{\mathbf{m}_1\in\mathbb{Z}^d}\left(\mathbf{m}_1+[0,1)^d\right)=\mathbb{R}^d$, the above expression is

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

$$= \int_{\mathcal{F}} \int_{\mathbb{R}^d} \sum_{\substack{\mathbf{m}_j' \in \mathbb{Z}^d \\ 2 \le j \le k}} F \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_1 + \mathbf{m}_2' g \\ \vdots \\ \mathbf{y}_1 + \mathbf{m}_k' g \end{pmatrix} \mathrm{d}\mathbf{y}_1 \, \mathrm{d}\mu(g)$$

$$= \int_{\mathbb{R}^d} F \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_1 \end{pmatrix} \mathrm{d}\mathbf{y}_1 + \int_{(\mathbb{R}^d)^k} F \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_k \end{pmatrix} \mathrm{d}\mathbf{y}_1 \, \mathrm{d}\mathbf{y}_2 \cdots \mathrm{d}\mathbf{y}_k$$

$$+ \sum_{r=1}^{k-2} \sum_{u=1}^{\infty} \sum_{D \in \mathfrak{D}_{r,u}^{k-1}} \frac{N(D,u)^d}{u^{dr}} \int_{(\mathbb{R}^d)^{r+1}} F \left(D' \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_{r+1} \end{pmatrix} \right) \mathrm{d}\mathbf{y}_1 \, \mathrm{d}\mathbf{y}_2 \cdots \mathrm{d}\mathbf{y}_{r+1},$$

where D' is defined as in (2.3) In the last equality, we applied Theorem 2.2 to the function

$$F' : (\mathbf{y}_2, \ldots, \mathbf{y}_k) \mapsto \int_{\mathbb{R}^d} F(\mathbf{y}_1, \mathbf{y}_1 + \mathbf{y}_2, \ldots, \mathbf{y}_1 + \mathbf{y}_k) \, \mathrm{d}\mathbf{y}_1.$$

Observe that it is enough to prove finiteness for $F \ge 0$. Indeed, for general F, finiteness for $|F|$ proves the absolute convergence of the sum in the RHS of (2.2). We note that (for $F \ge 0$) F' is a compactly supported bounded positive function, and hence, invoking Schmidt [17, Theorem 2] for this function proves our claim. □

2.2. Higher moment formulae for $Y_{\mathbf{p}/q}$

Recall that for $\mathbf{p} \in \mathbb{Z}^d \setminus \{\mathbf{0}\}$ and $q \in \mathbb{N}_{\ge 2}$ such that $\gcd(\mathbf{p}, q) = 1$, we set

$$Y_{\mathbf{p}/q} := \left\{ \left(\mathbb{Z}^d + \frac{\mathbf{p}}{q} \right) g : g \in \mathrm{SL}_d(\mathbb{R}) \right\} \subseteq Y.$$

We remark that the space $Y_{\mathbf{p}/q}$ does not depend on $\mathbf{p}$ because $Y_{\mathbf{p}/q}$ is the space of all affine grids $L + \mathbf{v}$, where L is a unimodular lattice in $\mathbb{R}^d$ and $\mathbf{v} \in \mathbb{R}^d$ is a representative of a torsion point of order q in the torus $\mathbb{R}^d/L$. Indeed, one can see that for such $L + \mathbf{v}$, $\exists\, g \in \mathrm{SL}_d(\mathbb{R})$ such that $L = \mathbb{Z}^d g$, and since $q\mathbf{v} \in L$, we have $\mathbf{v} = \frac{\mathbf{w} g}{q}$, where $\mathbf{w} \in \mathbb{Z}^d$ and $\mathbf{w}$ is of order q in $(\mathbb{Z}/q\mathbb{Z})^d$ (since $\mathbf{v}$ is of order q). Therefore, $L + \mathbf{v} = \left(\mathbb{Z}^d + \frac{\mathbf{w}}{q} \right) g$. Since $\mathbf{p}$ is also of order q in $(\mathbb{Z}/q\mathbb{Z})^d$ and $\mathrm{SL}_d(\mathbb{Z})$ acts transitively on elements of order q in $(\mathbb{Z}/q\mathbb{Z})^d$, $\exists\, \gamma \in \mathrm{SL}_d(\mathbb{Z})$ such that $\mathbf{w} = \mathbf{p}\gamma$. Hence,

$$L + \mathbf{v} = \left(\mathbb{Z}^d + \frac{\mathbf{p}\gamma}{q} \right) g = \left(\mathbb{Z}^d + \frac{\mathbf{p}}{q} \right) \gamma g.$$

Let $\{\mathbf{e}_j\}$ be the canonical basis of $\mathbb{R}^d$. Define

$$\Gamma(q) = \{\gamma \in \mathrm{SL}_d(\mathbb{Z}) : \gamma \equiv \mathrm{Id}_d \mod q\},$$
$$\Gamma_1(q) = \{\gamma \in \mathrm{SL}_d(\mathbb{Z}) : \mathbf{e}_1 \gamma \equiv \mathbf{e}_1 \mod q\},$$

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

and $X_q = \Gamma(q)\backslash \mathrm{SL}_d(\mathbb{R})$. If we choose any $\gamma_{\mathbf{p}} \in \mathrm{SL}_d(\mathbb{Z})$ for which $\mathbf{p} = r\mathbf{e}_1\gamma_{\mathbf{p}}$, where $r = \gcd \mathbf{p}$, then $Y_{\mathbf{p}/q}$ can be identified with $\gamma_{\mathbf{p}}^{-1}\Gamma_1(q)\gamma_{\mathbf{p}}\backslash \mathrm{SL}_d(\mathbb{R})$ ([6, Lemma 3.1]). Denote by μ_q the Haar measure on $\mathrm{SL}_d(\mathbb{R})$ normalised so that $\mu_q(Y_{\mathbf{p}/q}) = 1$. More precisely, let $J_q = [\mathrm{SL}_d(\mathbb{Z}) : \Gamma_1(q)]$. We can see that $\mu_q = \frac{1}{J_q}\mu$, which is independent of the choice of $\mathbf{p}$.

Recall that we identify the k-tuple $(\mathbb{R}^d)^k$ of $\mathbb{R}^d$ with $\mathrm{Mat}_{k,d}(\mathbb{R})$. Let $\{E_{ij} : 1 \le i \le k, 1 \le j \le d\}$ be the standard basis for $(\mathbb{R}^d)^k$; that is, the (k,ℓ)-entry $[E_{ij}]_{k\ell} = 0$ except that $[E_{ij}]_{ij} = 1$.

The Lemma below essentially follows from the definition. However, we provide a proof since it is vital in setting up and proving moment formulas for congruence quotients.

Lemma 2.5. *For each $D \in \mathfrak{D}^k_{r,u}$, where $\mathfrak{D}^k_{r,u}$ is as in Notation 2.1, define*

$$\Lambda_D = \left\{ \begin{pmatrix} \ell_1 \\ \vdots \\ \ell_r \end{pmatrix} \in \mathbb{Z}^r : \frac{D}{u}\begin{pmatrix} \ell_1 \\ \vdots \\ \ell_r \end{pmatrix} \in \mathbb{Z}^k \right\}.$$

It follows that $\frac{D}{u} : \Lambda_D \to \frac{D}{u}\mathbb{R}^r$ is injective, and moreover,

$$\frac{D}{u}\Lambda_D = \frac{D}{u}\mathbb{R}^r \cap \mathbb{Z}^k.$$

In other words, the set $\frac{D}{u}\Lambda_D$ is a primitive sublattice of $\mathbb{Z}^k$ of rank r, which is given by intersecting with the rational subspace $\frac{D}{u}\mathbb{R}^r \subseteq \mathbb{R}^k$.

Proof. One direction as well as the injectivity is obvious. Let us show the other direction. Suppose that $\boldsymbol{\ell} \in \mathbb{R}^r$ satisfies that $\frac{D}{u}\boldsymbol{\ell} \in \mathbb{Z}^k$. Considering indices $1 \le i_1 < \ldots < i_r \le k$ in Notation 2.1 (2), we have that $\boldsymbol{\ell} = ([\frac{D}{u}\boldsymbol{\ell}]^{i_1}, \ldots, [\frac{D}{u}\boldsymbol{\ell}]^{i_r}) \in \mathbb{Z}^r$. This proves the lemma since $\Lambda_D = \mathbb{Z}^r \cap \left(\frac{D}{u}\right)^{-1}\mathbb{Z}^k$. $\square$

Notation 2.6. For each $D \in \mathfrak{D}^k_{r,u}$, since Λ_D defined as in Lemma 2.5 is primitive, one can find elements $\mathbf{b}_1, \ldots, \mathbf{b}_{k-r}$ in $\mathbb{Z}^k$ such that for any $\mathbb{Z}$-basis $\{\mathbf{b}_{k-r+1}, \ldots, \mathbf{b}_k\}$ of $\frac{D}{u}\Lambda$, it holds that

$$\mathbb{Z}^k = \mathbb{Z}\mathbf{b}_1 \oplus \cdots \oplus \mathbb{Z}\mathbf{b}_k.$$

Fix such a set $\{\mathbf{b}_1, \ldots, \mathbf{b}_{k-r}\}$ for each $D \in \mathfrak{D}^k_{r,u}$ and denote

$$\mathcal{R}(D) = \mathbb{Z}\mathbf{b}_1 \oplus \cdots \oplus \mathbb{Z}\mathbf{b}_{k-r}$$

so that $\mathbb{Z}^k = \bigsqcup_{\boldsymbol{\ell} \in \mathcal{R}(D)} \left(\boldsymbol{\ell} + \frac{D}{u}\Lambda_D\right)$. We also define the set $P_t(\mathcal{R}(D))$ for every $t \in \mathbb{N}$ with $\gcd(t,q) = 1$ as

$$P_t(\mathcal{R}(D)) = \{\boldsymbol{\ell} \in \mathcal{R}(D) : \gcd(\boldsymbol{\ell}, t) = 1\}.$$

We are now ready to formulate the higher moment formula for $Y_{\mathbf{p}/q}$, based on Notation 2.6. The formula in equation (2.4) below depends on a choice of $\mathcal{R}(D)$ for each $\mathfrak{D}^k_{r,u}$. We are very grateful to the anonymous referee for providing an alternative formulation which does not involve any ad hoc choices. This formulation can be found in Theorem 2.13. We have chosen to include both formulations because we believe that (2.4) is more "intrinsic" in some sense (i.e., more indicative of the proof); see, for instance, the similarity with the second moment formula proven in [6] (see also [11, Proposition 7.6]).

Theorem 2.7. *Let $d \geq 3$ and $1 \leq k \leq d-1$. Let $F : (\mathbb{R}^d)^k \to \mathbb{R}$ be bounded and compactly supported. Then*

(1) *For $k = 1$,*

$$\int_{Y_{\mathbf{p}/q}} \mathcal{S}_1(F)(\Lambda)\,\mathrm{d}\mu_q(\Lambda) = \int_{\mathbb{R}^d} F(\mathbf{y})\,\mathrm{d}\mathbf{y}.$$

(2) *For $2 \leq k \leq d-1$,*

$$\begin{aligned}
&\int_{Y_{\mathbf{p}/q}} \mathcal{S}_k(F)(\Lambda)\,\mathrm{d}\mu_q(\Lambda) = \int_{(\mathbb{R}^d)^k} F\left({}^{\mathrm{t}}(\mathbf{y}_1,\ldots,\mathbf{y}_k)\right)\mathrm{d}\mathbf{y}_1\cdots\mathrm{d}\mathbf{y}_k \\
&+\int_{\mathbb{R}^d} F\left({}^{\mathrm{t}}(\mathbf{y},\ldots,\mathbf{y})\right)d\mathbf{y} + \sum_{\substack{t\in\mathbb{N}\\(t,q)=1}}\ \sum_{\substack{\boldsymbol{\ell}\neq\mathbf{0}\\ \in\mathbb{Z}^{k-1}}} \int_{\mathbb{R}^d} F\left(\left(\begin{array}{c} t\mathbf{y}\\ (t+\ell_1 q)\mathbf{y}\\ \vdots\\ (t+\ell_{k-1}q)\mathbf{y}\end{array}\right)\right)d\mathbf{y}\\
&+\sum_{r=1}^{k-2}\sum_{u\in\mathbb{N}}\sum_{D\in\mathfrak{D}_{r,u}^{k-1}}\left[\frac{N(D,u)^d}{u^{dr}}\int_{(\mathbb{R}^d)^{r+1}} F\left(D'\left(\begin{array}{c}\mathbf{y}_1\\ \vdots\\ \mathbf{y}_{r+1}\end{array}\right)\right)\mathrm{d}\mathbf{y}_1\cdots\mathrm{d}\mathbf{y}_{r+1}\right.\\
&\qquad\left.\sum_{\substack{t\in\mathbb{N}\\(t,q)=1}}\ \sum_{\substack{\boldsymbol{\ell}\in\\ P_t(\mathcal{R}(D))}} \frac{N(D,u)^d}{t^d\cdot u^{dr}}\int_{(\mathbb{R})^{r+1}} F\left(D'_{t,\boldsymbol{\ell}}\left(\begin{array}{c}\mathbf{y}_1\\ \vdots\\ \mathbf{y}_{r+1}\end{array}\right)\right)d\mathbf{y}_1\cdots d\mathbf{y}_{r+1}\right],
\end{aligned} \tag{2.4}$$

where D' and $D'_{t,\boldsymbol{\ell}}$ for $D \in \mathfrak{D}_{r,u}^{k-1}$ and $\boldsymbol{\ell} = {}^{\mathrm{t}}(\ell_1,\ldots,\ell_{k-1}) \in P_t(\mathcal{R}(D))$ are $k\times(r+1)$ matrices defined as follows:

$$D' = \left(\begin{array}{c|c} 1 & 0\cdots 0\\ \hline 1 & \\ \vdots & \frac{1}{u}D\\ 1 & \end{array}\right) \quad \text{and} \quad D'_{t,\boldsymbol{\ell}} = \left(\begin{array}{c|c} t & 0\cdots 0\\ \hline t+\ell_1 q & \\ \vdots & \frac{1}{u}D\\ t+\ell_{k-1}q & \end{array}\right).$$

Here, if $k = 2$, we will consider $\sum_{m=1}^{0}$ as the empty summation.

Finally, both sides of the equation (2.4) are finite.

Notice that the right-hand side of the above expression does not depend on $\mathbf{p} \in \mathbb{Z}^d\setminus\{\mathbf{0}\}$, once $\gcd(\mathbf{p},q) = 1$.

We need several lemmas for the proof of Theorem 2.7. Let

$$H = \left\{\left(\begin{array}{cc} 1 & 0\\ {}^{\mathrm{t}}\mathbf{v}' & g'\end{array}\right) : \mathbf{v}' \in \mathbb{R}^{d-1} \text{ and } g' \in \mathrm{SL}_{d-1}(\mathbb{R})\right\}$$

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

and denote an element of H by $[\mathbf{v}',g']$. Let us identify $\mathrm{SL}_{d-1}(\mathbb{R})$ with the subgroup $\{[0,g']: g'\in \mathrm{SL}_{d-1}(\mathbb{R})\}$ of H. One can define the Haar measure μ_H on H by the product of μ' and the Lebesgue measure on $\mathbb{R}^{d-1}$, where μ' is the Haar measure such that $\mu'(X_{d-1})=1$.

Notice the difference between H and $\mathrm{ASL}_{d-1}(\mathbb{R})$. For instance, a fundamental domain of $(\mathrm{SL}_d(\mathbb{Z})\cap H)\backslash H$ is given by $[0,1)^{d-1}\times \mathcal{F}_{d-1}$, where $\mathcal{F}_{d-1}$ is a fundamental domain of $\mathrm{SL}_{d-1}(\mathbb{Z})\backslash \mathrm{SL}_{d-1}(\mathbb{R})$, whereas that of $\mathrm{ASL}_{d-1}(\mathbb{Z})\backslash \mathrm{ASL}_{d-1}(\mathbb{R})$ is given by

$$\left\{[\xi' g',g'] : g'\in \mathcal{F}_{d-1} \text{ and } \xi'\in [0,1)^{d-1}\right\}.$$

Proposition 2.8. *Let $F:(\mathbb{R}^d)^k\to \mathbb{R}_{\geq 0}$, where $d\geq 3$ and $1\leq k\leq d-2$, be a bounded and compactly supported function. Suppose that $\xi=(z_1,\xi')\in\mathbb{R}^d$ with $z_1\in\mathbb{R}$ and $\xi'\in\mathbb{Z}^{d-1}$. Then,*

$$\begin{aligned}
&\int_{\mathrm{SL}_d(\mathbb{Z})\cap H\backslash H} \mathcal{S}_k(F)\left((\mathbb{Z}^d+\xi)g\right)\,\mathrm{d}\mu_H(g)\\
&=\sum_{\ell_1,\ldots,\ell_k\in\mathbb{Z}} F\left(\sum_{i=1}^k (z_1+\ell_i)E_{i1}\right)\\
&\quad+\sum_{r=1}^k\sum_{u\in\mathbb{N}}\sum_{D\in\mathfrak{D}^k_{r,u}}\sum_{\substack{{}^{\mathrm{t}}(\ell_1,\ldots,\ell_k)\\ \in\mathcal{R}(D)}}\\
&\qquad\frac{N(D,u)^d}{u^{dr}}\int_{(\mathbb{R}^d)^r} F\left(\sum_{i=1}^k (z_1+\ell_i)E_{i1}+\frac{D}{u}\begin{pmatrix}\mathbf{x}_1\\ \vdots\\ \mathbf{x}_r\end{pmatrix}\right)\mathrm{d}\mathbf{x}_1\cdots\mathrm{d}\mathbf{x}_r.
\end{aligned}$$

Note that H is the isotropy subgroup of $\mathbf{e}_1$ in $\mathrm{SL}_d(\mathbb{R})$. We will compute the integral $\int_{\mathrm{SL}_d(\mathbb{Z})\cap H\backslash H}\mathcal{S}_k(F)\,\mathrm{d}\mu_H$ in two steps: we first process the integrals associated to the first column in $(\mathbb{R}^d)^k\simeq \mathrm{Mat}_{k,d}(\mathbb{R})$ and then apply Theorem 2.2 to the integrals associated to the remaining columns. For this, we need the lemma below which describes the relation between the primitive sublattice $\frac{D}{u}\Lambda_D$ of $\mathbb{Z}^k$ for $D\in\mathfrak{D}^k_{r,u}$ and its sublattice $\frac{C}{w}\Lambda_C$ for some $C\in\mathfrak{D}^k_{r-1,w}$.

Lemma 2.9. *Recall Notation 2.6. Let $D\in\mathfrak{D}^k_{r,u}$ with $r\geq 2$.*

(a) *For $1\leq j_0\leq r$ and $a_1,\ldots,a_{j_0-1}\in\mathbb{Q}$, define $C_1\in\mathrm{Mat}_{k,r-1}(\mathbb{Q})$ by*

$$[C_1]^j=\begin{cases}[D/u]^j+a_j[D/u]^{j_0} & \text{for } 1\leq j<j_0;\\ [D/u]^{j+1} & \text{for } j_0\leq j\leq r-1.\end{cases}$$

Let $w\in\mathbb{N}$ be the least common denominator of C_1 and $C:=wC_1\in\mathfrak{D}^k_{r-1,w}$. Let $\mathfrak{D}^k_{r-1,w}(D)$ be the collection of such matrices C. There is a one-to-one correspondence between

$$\bigcup_{w\in\mathbb{N}}\mathfrak{D}^k_{r-1,w}(D) \text{ and } \{(r-1)\text{-dimensional rational subspaces in } D\mathbb{R}^r\}.$$

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

(b) *For each* $C \in \mathfrak{D}^k_{r-1,w}(D)$*, define*

$$\Lambda_D(C) = \left\{ {}^{\mathrm{t}}(\ell_1,\ldots,\ell_r) \in \Lambda_D : a_1\ell_1 + \cdots + a_{j_0-1}\ell_{j_0-1} = \ell_{j_0} \right\}.$$

Then there is a natural isomorphism from Λ_C *to* $\Lambda_D(C)$ *so that*

$$\frac{D}{u}\Lambda_D(C) = \frac{C}{w}\Lambda_C \subseteq \frac{D}{u}\Lambda_D \subseteq \mathbb{Z}^k.$$

For each such pair (D,C)*, one can choose (and fix from now on) an element* $\mathbf{b}_D \in \Lambda_D - \Lambda_D(C)$ *so that if we let* $\mathcal{R}_D(C) = \mathbb{Z}\mathbf{b}_D \setminus \{\mathbf{0}\}$*, then it holds that*

$$\Lambda_D - \Lambda_D(C) = \bigsqcup_{\boldsymbol{\ell} \in \mathcal{R}_D(C)} (\boldsymbol{\ell} + \Lambda_D(C)).$$

(c) *For a given* $C \in \mathfrak{D}^k_{r-1,w}$*, let* $D_1 \in \mathfrak{D}^k_{r,u_1}$ *and* $D_2 \in \mathfrak{D}^k_{r,u_2}$ *be such that* $D_1 \neq D_2$ *and* $C \in \mathfrak{D}^k_{r-1,w}(D_1) \cap \mathfrak{D}^k_{r-1,w}(D_2)$*. Then*

$$\frac{D_1}{u_1}\mathcal{R}_{D_1}(C) \cap \frac{D_2}{u_2}\mathcal{R}_{D_2}(C) = \emptyset.$$

Hence, for any $\boldsymbol{\ell}_1 \in \frac{D_1}{u_1}\mathcal{R}_{D_1}(C)$ *and* $\boldsymbol{\ell}_2 \in \frac{D_2}{u_2}\mathcal{R}_{D_2}(C)$*, it follows that*

$$\left(\boldsymbol{\ell}_1 + \frac{C}{w}\Lambda_C\right) \cap \left(\boldsymbol{\ell}_2 + \frac{C}{w}\Lambda_C\right) = \emptyset.$$

(d) *For a given* $C \in \mathfrak{D}^k_{r-1,w}$*, one can choose* $\mathcal{R}(C)$ *in Notation 2.6 to be*

$$\mathcal{R}(C) = \left\{ {}^{\mathrm{t}}(0,\ldots,0) \right\} \sqcup \bigsqcup_{u \in \mathbb{N}} \bigsqcup_{\substack{D \in \mathfrak{D}^k_{r,u} \text{ s.t.} \\ C \in \mathfrak{D}^k_{r-1,w}(D)}} \frac{D}{u}\mathcal{R}_D(C)$$

and vice versa.

Proof. (a) One way is obvious from its construction. Suppose that $W \subsetneq D\mathbb{R}^r$ is a codimension-one rational subspace of $D\mathbb{R}^r$. Then there is $C \in \mathfrak{D}^k_{r-1,w}$ so that $\frac{C}{w}\mathbb{R}^{r-1} = W$. We want to show that $C \in \mathfrak{D}^k_{r-1,w}(D)$. Pick any ${}^{\mathrm{t}}(\mathbf{m}_1,\ldots,\mathbf{m}_k) \in (\mathbb{R}^d)^k \in \frac{C}{w}(\mathbb{R}^d)^k \cap (\mathbb{Z}^d)^k$ with rank $(r-1)$ and ${}^{\mathrm{t}}(\mathbf{m}'_1,\ldots,\mathbf{m}'_r)$ and ${}^{\mathrm{t}}(\mathbf{m}''_1,\ldots,\mathbf{m}''_{r-1})$ be such that

$$\frac{D}{u}\begin{pmatrix} \mathbf{m}'_1 \\ \vdots \\ \mathbf{m}'_r \end{pmatrix} = \begin{pmatrix} \mathbf{m}_1 \\ \vdots \\ \mathbf{m}_k \end{pmatrix} = \frac{C}{w}\begin{pmatrix} \mathbf{m}''_1 \\ \vdots \\ \mathbf{m}''_{r-1} \end{pmatrix}.$$

Let $I_D = \{i_1 < \ldots < i_r\}$ be as in Notation 2.1 (3). Since $\mathrm{rk}\,{}^{\mathrm{t}}(\mathbf{m}_1,\ldots,\mathbf{m}_k) = r-1$, and by definition of D and C, there is $1 \leq j_0 \leq r$ for which

$$\mathbf{m}'_1 = \mathbf{m}_{i_1} = \mathbf{m}''_1,\ \ldots,\ \mathbf{m}'_{j_0-1} = \mathbf{m}_{i_{j_0-1}} = \mathbf{m}''_{j_0-1};$$
$$\mathbf{m}'_{j_0} = a_1\mathbf{m}'_1 + \cdots + a_{j_0-1}\mathbf{m}'_{j_0-1} \text{ for some } a_1,\ldots,a_{j_0-1} \in \mathbb{Q};$$
$$\mathbf{m}'_{j_0+1} = \mathbf{m}_{i_{j_0+1}} = \mathbf{m}''_{j_0},\ \ldots,\ \mathbf{m}'_r = \mathbf{m}_{i_r} = \mathbf{m}''_{r-1}.$$

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

It is easily seen that C_1 constructed from D with j_0 and $a_1,\dots,a_{j_0-1}\in\mathbb{Q}$ in Property (a) is equal to C.

(b) It is obvious from the definition that $C\in\mathfrak{D}^k_{r-1,w}$. The map

$$^{\mathrm{t}}(\ell_1,\dots,\ell_{r-1})\mapsto{}^{\mathrm{t}}(\ell_1,\dots,\ell_{j_0-1},a_1\ell_1+\cdots+a_{j_0-1}\ell_{j_0-1},\ell_{j_0},\dots,\ell_{r-1})$$

gives an isomorphism from Λ_C to $\Lambda_D(C)$, and by definition,

$$\frac{C}{w}{}^{\mathrm{t}}(\ell_1,\dots,\ell_{r-1})=\frac{D}{u}{}^{\mathrm{t}}(\ell_1,\dots,\ell_{j_0-1},a_1\ell_1+\cdots+a_{j_0-1}\ell_{j_0-1},\ell_{j_0},\dots,\ell_{r-1})\in\mathbb{Z}^k.$$

Recall Lemma 2.5. Since $\frac{C}{w}\Lambda_C$ is primitive, there is an element $\mathbf{b}\in\frac{D}{u}\Lambda_D$ for which $\frac{D}{u}\Lambda_D=\frac{C}{w}\Lambda_C\oplus\mathbb{Z}\mathbf{b}$. Set $\mathbf{b}_D:=\left(\frac{D}{u}\right)^{-1}\mathbf{b}$.

(c) Let $\mathcal{R}_{D_i}(C)$ be generated by $\mathbf{b}_{D_i}$ for $i=1,2$. From the fact that $\frac{D_1}{u_1}\mathbb{R}^r\cap\frac{D_2}{u_2}\mathbb{R}^r=\frac{C}{w}\mathbb{R}^{r-1}$, in other words,

$$\frac{D_i}{u_i}\mathbb{R}^r=\frac{C}{w}\mathbb{R}^{r-1}\oplus\mathbb{R}\left(\frac{D_i}{u_i}\mathbf{b}_{D_i}\right)\ (i=1,2),$$

it is obvious that $\frac{D_1}{u_1}\mathcal{R}_{D_1}(C)\cap\frac{D_2}{u_2}\mathcal{R}_{D_2}(C)=\emptyset$. Moreover, for any $\boldsymbol{\ell}_i\in\frac{D_i}{u_i}$, where $i=1,2$,

$$\boldsymbol{\ell}_i+\frac{C}{w}\Lambda_C\subseteq\boldsymbol{\ell}_i+\frac{C}{w}\mathbb{R}^{r-1},$$

which are affine subspaces of $\frac{C}{w}\mathbb{R}^{r-1}$ lying on $\frac{D_i}{u_i}\mathbb{R}^r-\frac{C}{w}\mathbb{R}^{r-1}$ for $i=1,2$, respectively. Hence, they are disjoint.

To deduce (d) from (c), it suffices to show that

$$\frac{C}{w}\Lambda_C+\left(\left\{^{\mathrm{t}}(0,\dots,0)\right\}\sqcup\bigsqcup_{u\in\mathbb{N}}\ \bigsqcup_{\substack{D\in\mathfrak{D}^k_{r,u}\\ C\in\mathfrak{D}^k_{r-1,w}(D)}}\frac{D}{u}\mathcal{R}_D(C)\right)=\frac{C}{w}\Lambda_C+\mathcal{R}(C).$$

Let $\boldsymbol{\ell}\in\mathcal{R}(C)$ be given. Since $\frac{C}{w}\mathbb{R}^{r-1}\oplus\mathbb{R}\boldsymbol{\ell}$ is a rational subspace of rank r, there are $u\in\mathbb{N}$ and $D\in\mathfrak{D}^k_{r,u}$, which are uniquely determined, such that $\frac{C}{w}\mathbb{R}^{r-1}\oplus\mathbb{R}\boldsymbol{\ell}=\frac{D}{u}\mathbb{R}^r$. It is obvious that $\boldsymbol{\ell}\in\frac{D}{u}\Lambda_D-\Lambda_D(C)$; hence, there is $\boldsymbol{\ell}'\in\frac{D}{u}\mathcal{R}_D(C)$ so that

$$\boldsymbol{\ell}\in\boldsymbol{\ell}'+\frac{D}{u}\Lambda_D(C)=\boldsymbol{\ell}'+\frac{C}{w}\Lambda_C,$$

as asserted in the claim. □

Proof of Proposition 2.8. Fix a fundamental domain $\mathcal{F}'$ of $\mathrm{SL}_{d-1}(\mathbb{Z})\backslash\mathrm{SL}_{d-1}(\mathbb{R})$ so that $\mathcal{F}'\times[0,1)^{d-1}$ is a fundamental domain of $(\mathrm{SL}_d(\mathbb{Z})\cap H)\backslash H$.

Recall that $(\mathbb{Z}^d)^k\setminus\{^{\mathrm{t}}(\mathbf{0},\dots,\mathbf{0})\}$ is partitioned into $\bigsqcup_{r=1}^k\bigsqcup_{u\in\mathbb{N}}\bigsqcup_{D\in\mathfrak{D}^k_{r,u}}\frac{D}{u}\Phi^{(d)}(D,u)$, where $\mathfrak{D}^k_{r,u}$ and $\Phi^{(d)}(D,u)$ are as in Notation 2.1.

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

By taking $g=[\mathbf{v}',g']$ and from Rogers' formula, we have that

$$\begin{aligned}
&\int_{(\mathrm{SL}_d(\mathbb{Z})\cap H)\backslash H} \mathcal{S}_k(F)\left((\mathbb{Z}^d+\xi)[\mathbf{v}',g']\right)\,\mathrm{d}\mu_H([\mathbf{v}',g'])\\
&=\int_{\mathcal{F}'\times[0,1)^{d-1}} \sum_{\substack{\ell_i\in\mathbb{Z}\\ \mathbf{m}_i'\in\mathbb{Z}^{d-1}\\ 1\le i\le k}} F\left(\begin{array}{c|c} (\ell_1+z_1)+\mathbf{m}_1'\,{}^{\mathrm{t}}\mathbf{v}' & \mathbf{m}_1'g'\\ \vdots & \vdots\\ (\ell_k+z_1)+\mathbf{m}_k'\,{}^{\mathrm{t}}\mathbf{v}' & \mathbf{m}_k'g'\end{array}\right)\mathrm{d}\mathbf{v}'\,\mathrm{d}\mu'(g')\\
&=F\left(\begin{array}{c|c} z_1 & 0,\dots,0\\ \vdots & \vdots\\ z_1 & 0,\dots,0\end{array}\right)+\sum_{r=1}^{k}\sum_{u\in\mathbb{N}}\sum_{D\in\mathfrak{D}^k_{r,u}}\int_{\mathcal{F}'\times[0,1)^{d-1}}\sum_{\substack{{}^{\mathrm{t}}(\mathbf{n}_1,\dots,\mathbf{n}_r)\\ \in\Phi^{(d)}(D,u)}}\\
&\qquad F\left(\left.\begin{pmatrix} z_1\\ \vdots\\ z_1\end{pmatrix}+\frac{D}{u}\begin{pmatrix}\ell_1'\\ \vdots\\ \ell_r'\end{pmatrix}+\frac{D}{u}\begin{pmatrix}\mathbf{n}_1'\\ \vdots\\ \mathbf{n}_r'\end{pmatrix}{}^{\mathrm{t}}\mathbf{v}'\ \right|\ \frac{D}{u}\begin{pmatrix}\mathbf{n}_1'\\ \vdots\\ \mathbf{n}_r'\end{pmatrix}g'\right)\mathrm{d}\mathbf{v}'\,\mathrm{d}\mu'(g),
\end{aligned}$$

where $\mathbf{n}_j=(\ell_j',\mathbf{n}_j')$ for $1\le j\le r$.

Now, let $D\in\mathfrak{D}^k_{r,u}$ be given. For each ${}^{\mathrm{t}}((\ell_1',\mathbf{n}_1'),\dots,(\ell_r',\mathbf{n}_r'))\in\Phi^{(d)}(D,u)$, the rank of ${}^{\mathrm{t}}(\mathbf{n}_1',\dots,\mathbf{n}_r')$ is either r or $r-1$.

Assume that $r\ge 2$. It is easy to verify that

$$\begin{aligned}
&\Phi^{(d)}(D,u)=\left(\Lambda_D\times\Phi^{(d-1)}(D,u)\right)\sqcup\\
&\bigsqcup_{w\in\mathbb{N}}\bigsqcup_{C\in\mathfrak{D}^k_{r-1,w}(D)}(\Lambda_D-\Lambda_D(C))\times\left\{\begin{pmatrix}\mathbf{n}_1'\\ \vdots\\ \sum_{k=1}^{j_0-1}a_k\mathbf{n}_k'\\ \vdots\\ \mathbf{n}_{r-1}'\end{pmatrix}_{r\times(d-1)} : \begin{pmatrix}\mathbf{n}_1'\\ \vdots\\ \mathbf{n}_{r-1}'\end{pmatrix}\in\Phi^{(d-1)}(C,w)\right\},
\end{aligned}$$

where $\mathfrak{D}^k_{r-1,w}(D)$ and $\Lambda_D(C)$ are as in Lemma 2.9.

Let us first compute the following integral

$$\begin{aligned}
&\int_{\mathcal{F}'\times[0,1)^{d-1}}\sum_{\substack{{}^{\mathrm{t}}(\ell_1',\dots,\ell_r')\\ \in\Lambda_D}}\sum_{\substack{{}^{\mathrm{t}}(\mathbf{n}_1',\dots,\mathbf{n}_r')\\ \in\Phi^{(d-1)}(D,u)}}\\
&\quad F\left(\left.\begin{pmatrix} z_1\\ \vdots\\ z_1\end{pmatrix}+\frac{D}{u}\begin{pmatrix}\ell_1'\\ \vdots\\ \ell_r'\end{pmatrix}+\frac{D}{u}\begin{pmatrix}\mathbf{n}_1'\\ \vdots\\ \mathbf{n}_r'\end{pmatrix}{}^{\mathrm{t}}\mathbf{v}'\ \right|\ \frac{D}{u}\begin{pmatrix}\mathbf{n}_1'\\ \vdots\\ \mathbf{n}_r'\end{pmatrix}g'\right)\mathrm{d}\mathbf{v}'\,\mathrm{d}\mu'(g').
\end{aligned}\tag{2.5}$$

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

Fix $g' \in \mathcal{F}'$ and $N := {}^{\mathrm{t}}(\mathbf{n}'_1, \ldots, \mathbf{n}'_r) \in \Phi^{(d-1)}(D,u)$. Set $J_N = \{1 \leq j_1 < \ldots < j_r \leq d-1\}$ such that $N^{J_N} := \left([N]^{j_1}, \ldots, [N]^{j_r}\right)$ has a nonzero determinant. Denote by

$$N\,{}^{\mathrm{t}}\mathbf{v}' = N^{J_N}\,{}^{\mathrm{t}}\mathbf{v}'_{J_N} + N^{J_N^c}\,{}^{\mathrm{t}}\mathbf{v}'_{J_N^c},$$

where $\mathbf{v}'_{J_N} = (v_j)_{j \in J_N} \in \mathbb{R}^r$ and $\mathbf{v}'_{J_N^c} = (v_i)_{i \in J_N^c} \in \mathbb{R}^{(d-1)-r}$. Define

$$G\begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} := \sum_{\substack{{}^{\mathrm{t}}(\ell'_1, \ldots, \ell'_r) \\ \in \Lambda_D}} F\left(\begin{pmatrix} z_1 \\ \vdots \\ z_1 \end{pmatrix} + \frac{D}{u}\begin{pmatrix} \ell'_1 \\ \vdots \\ \ell'_r \end{pmatrix} + \frac{D}{u}\begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} \,\middle|\, \frac{D}{u} N g' \right).$$

Obviously, G is Λ_D-invariant so that it is $u\mathbb{Z}^r$-invariant, and $G(N\,\cdot)$ is $\mathbb{Z}^{d-1}$-invariant. We want to compute the integral

$$\begin{aligned}
&\int_{[0,1)^{d-1}} G\left(N\begin{pmatrix} v_1 \\ \vdots \\ v_{d-1} \end{pmatrix}\right) \mathrm{d}v_1 \cdots \mathrm{d}v_{d-1} \\
&\quad = \frac{1}{u^{d-1}} \int_{[0,u)^{d-1}} G\left(N^{J_N}\mathbf{v}^{J_N} + N^{J_N^c}\mathbf{v}^{J_N^c}\right) \mathrm{d}\mathbf{v}^{J_N}\,\mathrm{d}\mathbf{v}^{J_N^c}.
\end{aligned}$$

By the change of variables

$$\mathbf{v}^{J_N} \mapsto N^{J_N}\mathbf{v}^{J_N} + N^{J_N^c}\mathbf{v}^{J_N^c} = {}^{\mathrm{t}}(x_1, \ldots, x_r),$$

the above integral is

$$\begin{aligned}
&= \frac{1}{u^{d-1}} \int_{[0,u)^{d-1-r}} \int_{N^{J_N}[0,u)^r + N^{J_N^c}\mathbf{v}^{J_N^c}} G\begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} |\det N^{J_N}|^{-1}\,\mathrm{d}x_1 \cdots \mathrm{d}x_r\,\mathrm{d}\mathbf{v}^{J_N^c} \\
&= \frac{1}{u^{d-1}} \int_{[0,u)^{d-1-r}} \int_{N^{J_N}[0,u)^r} G\begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} |\det N^{J_N}|^{-1}\,\mathrm{d}x_1 \cdots \mathrm{d}x_r\,\mathrm{d}\mathbf{v}^{J_N^c} \\
&= \frac{1}{u^r} \int_{[0,u)^r} G\begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} \mathrm{d}x_1 \cdots \mathrm{d}x_r.
\end{aligned}$$

Let $\mathcal{F}_{\Lambda_D}$ be a fundamental domain for Λ_D in $[0,u)^r$. Since $[0,u)^r$ is an $N(D,u)$-covering of $\mathcal{F}_{\Lambda_D}$, it follows that

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

$$\int_{[0,1)^{d-1}} G\left(N\begin{pmatrix} v_1 \\ \vdots \\ v_{d-1}\end{pmatrix}\right) \mathrm{d}v_1 \cdots \mathrm{d}v_{d-1}$$
$$= \frac{N(D,u)}{u^r} \int_{\mathcal{F}_{\Lambda_D}} \sum_{\substack{{}^{\mathrm{t}}(\ell'_1,\ldots,\ell'_r) \\ \in \Lambda_D}} F\left(\left. \begin{pmatrix} z_1 \\ \vdots \\ z_1 \end{pmatrix} + \frac{D}{u}\begin{pmatrix} x_1+\ell'_1 \\ \vdots \\ x_r+\ell'_r \end{pmatrix} \right| \frac{D}{u}Ng' \right) \mathrm{d}x_1 \cdots \mathrm{d}x_r$$
$$= \frac{N(D,u)}{u^r} \int_{\mathbb{R}^r} F\left(\left. \begin{pmatrix} z_1 \\ \vdots \\ z_1 \end{pmatrix} + \frac{D}{u}\begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} \right| \frac{D}{u}Ng' \right) \mathrm{d}x_1 \cdots \mathrm{d}x_r.$$

Therefore, applying Proposition 2.3, the integral (2.5) is

$$\frac{N(D,u)}{u^r} \int_{\mathrm{SL}_{d-1}(\mathbb{Z})\backslash \mathrm{SL}_{d-1}(\mathbb{R})} \sum_{\substack{{}^{\mathrm{t}}(\mathbf{n}'_1,\ldots,\mathbf{n}'_r) \\ \in \Phi^{(d-1)}(D,u)}}$$
$$\int_{\mathbb{R}^r} F\left(\left. \begin{pmatrix} z_1 \\ \vdots \\ z_1 \end{pmatrix} + \frac{D}{u}\begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} \right| \frac{D}{u}\begin{pmatrix} \mathbf{n}'_1 \\ \vdots \\ \mathbf{n}'_r \end{pmatrix} g' \right) \mathrm{d}x_1 \cdots \mathrm{d}x_r \, \mathrm{d}\mu'(g')$$
$$= \frac{N(D,u)}{u^r} \cdot \frac{N(D,u)^{d-1}}{u^{(d-1)r}}$$
$$\int_{(\mathbb{R}^{d-1})^r} \int_{\mathbb{R}^r} F\left(\left. \begin{pmatrix} z_1 \\ \vdots \\ z_1 \end{pmatrix} + \frac{D}{u}\begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} \right| \frac{D}{u}\begin{pmatrix} \mathbf{x}'_1 \\ \vdots \\ \mathbf{x}'_r \end{pmatrix} \right) \mathrm{d}x_1 \cdots \mathrm{d}x_r \, \mathrm{d}\mathbf{x}'_1 \cdots \mathrm{d}\mathbf{x}'_r$$
$$= \frac{N(D,u)^d}{u^{dr}} \int_{(\mathbb{R}^d)^r} F\left(\sum_{i=1}^k z_1 E_{i1} + \frac{D}{u}\begin{pmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_r \end{pmatrix} \right) \mathrm{d}\mathbf{x}_1 \cdots \mathrm{d}\mathbf{x}_r.$$

Now, let us fix $C \in \mathfrak{D}^k_{r-1,w}(D)$ and let $\Lambda_D(C)$ and $\mathcal{R}_D(C)$ be as in Lemma 2.9. We want to compute

$$\int_{\mathcal{F}'\times[0,1)^{d-1}} \sum_{\substack{{}^{\mathrm{t}}(\ell'_1,\ldots,\ell'_r) \\ \in \Lambda_D - \Lambda_D(C)}} \sum_{\substack{{}^{\mathrm{t}}(\mathbf{n}'_1,\ldots,\mathbf{n}'_{r-1}) \\ \in \Phi^{(d-1)}(C,w)}}$$
$$F\left(\left. \begin{pmatrix} z_1 \\ \vdots \\ z_1 \end{pmatrix} + \frac{D}{u}\begin{pmatrix} \ell'_1 \\ \vdots \\ \ell'_r \end{pmatrix} + \frac{C}{w}\begin{pmatrix} \mathbf{n}'_1 \\ \vdots \\ \mathbf{n}'_{r-1} \end{pmatrix} {}^{\mathrm{t}}\mathbf{v}' \right| \frac{C}{w}\begin{pmatrix} \mathbf{n}'_1 \\ \vdots \\ \mathbf{n}'_{r-1} \end{pmatrix} g' \right) \mathrm{d}\mathbf{v}' \, \mathrm{d}\mu'(g'). \tag{2.6}$$

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

Since $\frac{D}{u}(\Lambda_D - \Lambda_D(C)) = \frac{D}{u}(\mathcal{R}_D(C) + \Lambda_D(C)) = \frac{D}{u}\mathcal{R}_D(C) + \frac{C}{w}\Lambda_C$ from Lemma 2.9 (a), the integral (2.6) is

$$\sum_{\boldsymbol{\ell}\in\mathcal{R}_D(C)} \int_{\mathcal{F}'\times[0,1)^{d-1}} \sum_{\substack{{}^{\mathrm{t}}(\ell'_1,\ldots,\ell'_{r-1})\\ \in \Lambda_C}} \sum_{\substack{{}^{\mathrm{t}}(\mathbf{n}'_1,\ldots,\mathbf{n}'_{r-1})\\ \in \Phi^{(d-1)}(C,w)}}$$
$$F\left(\left(\begin{pmatrix} z_1 \\ \vdots \\ z_1 \end{pmatrix} + \frac{D}{u}\boldsymbol{\ell} + \frac{C}{w}\begin{pmatrix} \ell'_1 \\ \vdots \\ \ell'_{r-1} \end{pmatrix} + \frac{C}{w}\begin{pmatrix} \mathbf{n}'_1 \\ \vdots \\ \mathbf{n}'_{r-1} \end{pmatrix} {}^{\mathrm{t}}\mathbf{v}' \;\middle|\; \frac{C}{w}\begin{pmatrix} \mathbf{n}'_1 \\ \vdots \\ \mathbf{n}'_{r-1} \end{pmatrix} g' \right)\right) \mathrm{d}\mathbf{v}'\,\mathrm{d}\mu'(g').$$

Repeating the same argument with the above, where now we put $N = {}^{\mathrm{t}}(\mathbf{n}'_1,\ldots,\mathbf{n}'_{r-1})$ with ${}^{\mathrm{t}}(\mathbf{n}'_1,\ldots,\mathbf{n}'_{r-1}) \in \Phi^{(d-1)}(C,w)$ and

$$G\begin{pmatrix} x_1 \\ \vdots \\ x_{r-1} \end{pmatrix} := \sum_{\substack{{}^{\mathrm{t}}(\ell'_1,\ldots,\ell'_{r-1})\\ \in \Lambda_C}} F\left(\left(\begin{pmatrix} z_1 \\ \vdots \\ z_1 \end{pmatrix} + \frac{D}{u}\boldsymbol{\ell} + \frac{C}{w}\begin{pmatrix} \ell'_1 \\ \vdots \\ \ell'_{r-1} \end{pmatrix} + \frac{C}{w}\begin{pmatrix} x_1 \\ \vdots \\ x_{r-1} \end{pmatrix} \;\middle|\; \frac{C}{w}Ng' \right)\right),$$

we have that the integral (2.6) is

$$\frac{N(C,w)^d}{w^{d(r-1)}} \sum_{\substack{{}^{\mathrm{t}}(\ell_1,\ldots,\ell_k)\\ \in \frac{D}{u}\mathcal{R}_D(C)}} \int_{(\mathbb{R}^d)^{r-1}} F\left(\sum_{i=1}^k (z_1+\ell_i)E_{i1} + \frac{C}{w}\begin{pmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_{r-1} \end{pmatrix}\right) \mathrm{d}\mathbf{x}_1\cdots\mathrm{d}\mathbf{x}_{r-1}.$$

If $r=1$ and $\mathrm{rk}\,{}^{\mathrm{t}}(\mathbf{n}'_1,\ldots,\mathbf{n}'_r) = 0$, that is, ${}^{\mathrm{t}}(\mathbf{n}'_1,\ldots,\mathbf{n}'_r) = {}^{\mathrm{t}}(\mathbf{0},\ldots,\mathbf{0})$ and the integral is

$$F\left(\sum_{i=1}^k (z_1+\ell_i)E_{i1}\right),$$

where ${}^{\mathrm{t}}(\ell_1,\ldots,\ell_k) \neq {}^{\mathrm{t}}(0,\ldots,0)$. Otherwise, they form $\Lambda_D \times \Phi^{(d-1)}(D,u)$, and one can proceed the same computation with the first case when $r \geq 2$.

Now the proposition follows from Lemma 2.9 (c) after rearranging the summation with respect to $C \in \mathfrak{D}^k_{r-1,w}$ for $1 \leq r-1 \leq k$ and $w \in \mathbb{N}$. □

For each $\mathbf{y} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$, define

$$X_q(\mathbf{y}) = \left\{\Gamma(q)g \in X_q : \mathbf{y} \in \left(\mathbb{Z}^d + \frac{\mathbf{p}}{q}\right) g\right\}.$$

It is known that for each $t \in \mathbb{N}$ with $\gcd(t,q) = 1$, there is $\mathbf{k}_t \in \mathbb{Z}^d + \mathbf{p}/q$ with $\gcd(q\mathbf{k}_t) = t$ so that we have the decomposition

$$X_q(\mathbf{y}) = \bigsqcup_{\substack{t\in\mathbb{N}\\ (t,q)=1}} X_q(\mathbf{k}_t,\mathbf{y}),$$

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

where $X_q(\mathbf{k}_t,\mathbf{y}) := \{\Gamma(q)g \in X_q : \mathbf{k}_t g = \mathbf{y}\}$ (see [11, Page 1993] for details). Note that the above decomposition holds for any such choice of $\mathbf{k}_t$. Moreover, if we put $g_t \in \mathrm{SL}_d(\mathbb{R})$ for each $t \in \mathbb{N}$ with $\gcd(t,q) = 1$ and $g_{\mathbf{y}} \in \mathrm{SL}_d(\mathbb{R})$, respectively, such that $\mathbf{e}_1 g_t = \mathbf{k}_t$ and $\mathbf{e}_1 g_{\mathbf{y}} = \mathbf{y}$, it follows that

$$X_q(\mathbf{k}_t,\mathbf{y}) \simeq g_t^{-1}\left((g_t\Gamma(q)g_t^{-1} \cap H)\backslash H\right) g_y \tag{2.7}$$

and one can define the probability measure $\nu_{\mathbf{y}}$ on $X_q(\mathbf{y})$ for which $\nu_{\mathbf{y}}|_{X_q(\mathbf{k}_t,\mathbf{y})}$ is the pull-back measure of $\frac{1}{I_q\zeta(d)}\mu_H$, where $I_q := [\mathrm{SL}_d(\mathbb{Z}) : \Gamma(q)]$, with respect to the above identification (see [11], especially (7.10)$\sim$(7.15) and Proposition 7.5).

Proposition 2.10. *Let $d \geq 3$ and $1 \leq k \leq d-1$. Suppose that $\mathbf{p} \in \mathbb{Z}^d \setminus \{\mathbf{0}\}$ and $q \in \mathbb{N}_{\geq 2}$ such that $\gcd(q,\mathbf{p}) = 1$. Let $P_t(\mathcal{R}(D))$ be as in Notation 2.6 after fixing $\mathcal{R}(D)$ for each $D \in \mathfrak{D}^k_{r,u}$.*

Let $F : (\mathbb{R}^d)^k \to \mathbb{R}_{\geq 0}$ be a bounded and compactly supported function. For any $\mathbf{y} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$, it follows that

$$\begin{aligned}
&\int_{X_q(\mathbf{y})} \mathcal{S}_k(F)\left(\left(\mathbb{Z}^d + \frac{\mathbf{p}}{q}\right) g\right) \mathrm{d}\nu_{\mathbf{y}}(g) \\
&= F\left({}^{\mathrm{t}}(\mathbf{y},\ldots,\mathbf{y})\right) + \sum_{\substack{t \in \mathbb{N} \\ (t,q)=1}} \frac{1}{t^d} \sum_{\substack{{}^{\mathrm{t}}(\ell_1,\ldots,\ell_k) \in \mathbb{Z}^k \\ (\ell_1,\ldots,\ell_k,t)=1}} F\left({}^{\mathrm{t}}\left(\frac{t+\ell_1 q}{t}\mathbf{y},\ldots,\frac{t+\ell_k}{t}\mathbf{y}\right)\right) \\
&+ \sum_{r=1}^{k-1} \sum_{u \in \mathbb{N}} \sum_{D \in \mathfrak{D}^k_{r,u}} \left[\frac{N(D,u)^d}{u^{dr}} \int_{(\mathbb{R}^d)^r} F\left(\begin{pmatrix} \mathbf{y} \\ \vdots \\ \mathbf{y} \end{pmatrix} + \frac{D}{u}\begin{pmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_r \end{pmatrix}\right) \mathrm{d}\mathbf{x}_1 \cdots \mathrm{d}\mathbf{x}_r \right. \\
&+ \sum_{\substack{t \in \mathbb{N} \\ (t,q)=1}} \sum_{\substack{\boldsymbol{\ell} \in \\ P_t(\mathcal{R}(D))}} \frac{N(D,u)^d}{t^d \cdot u^{dr}} \times \\
&\qquad \left. \int_{(\mathbb{R}^d)^r} F\left(\begin{pmatrix} \frac{t+\ell_1 q}{t}\mathbf{y} \\ \vdots \\ \frac{t+\ell_k q}{t}\mathbf{y} \end{pmatrix} + \frac{D}{u}\begin{pmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_r \end{pmatrix}\right) \mathrm{d}\mathbf{x}_1 \cdots \mathrm{d}\mathbf{x}_r \right] \\
&+ \int_{(\mathbb{R}^d)^k} F\left({}^{\mathrm{t}}(\mathbf{x}_1,\ldots,\mathbf{x}_k)\right) \mathrm{d}\mathbf{x}_1 \cdots \mathrm{d}\mathbf{x}_k.
\end{aligned}$$

Proof. Recall the definitions of g_t, $g_{\mathbf{y}}$ as in (2.7). If we let $a_{t/q} = \mathrm{diag}(t/q, q/t, 1, \ldots, 1)$, then one can further assume that $g_t = a_{t/q}\gamma_t$ for some $\gamma_t \in \mathrm{SL}_d(\mathbb{Z})$ ([11, Page 1993]).

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

By the definition of $\nu_{\mathbf{y}}$ on $X_q(\mathbf{y})$,

$$\int_{X_q(\mathbf{y})} \mathcal{S}_k(F)\left(\left(\mathbb{Z}^d+\frac{\mathbf{p}}{q}\right)g\right)\mathrm{d}\nu_{\mathbf{y}}(g)$$

$$=\frac{1}{I_q\zeta(d)}\sum_{\substack{t\in\mathbb{N}\\(t,q)=1}}\int_{(g_t\Gamma(q)g_t^{-1}\cap H)\backslash H}\sum_{\substack{\mathbf{m}_i\in\mathbb{Z}^d\\1\le i\le k}}F\left(\begin{pmatrix}\mathbf{m}_1+\mathbf{p}/q\\ \vdots\\ \mathbf{m}_k+\mathbf{p}/q\end{pmatrix}g_t^{-1}hg_{\mathbf{y}}\right)\mathrm{d}\mu_H(h)$$

$$=\frac{q^d}{I_q\zeta(d)}\sum_{\substack{t\in\mathbb{N}\\(t,q)=1}}\frac{1}{t^d}\int_{(\Gamma(q)\cap H)\backslash H}\sum_{\substack{\mathbf{m}_i\in\mathbb{Z}^d\\1\le i\le k}}F\left(\begin{pmatrix}\mathbf{m}_1+\mathbf{p}/q\\ \vdots\\ \mathbf{m}_k+\mathbf{p}/q\end{pmatrix}\gamma_t^{-1}ha_{t/q}^{-1}g_{\mathbf{y}}\right)\mathrm{d}\mu_H(h).$$

Note that $(\mathbb{Z}^d+\mathbf{p}/q)\gamma_t^{-1}=(\mathbb{Z}^d+\mathbf{k}_t)\gamma_t^{-1}=\mathbb{Z}^d+(t/q)\mathbf{e}_1$ and $(\Gamma(q)\cap H)\setminus H$ is a $(q^{d-1}I_q^{(d-1)})$-covering of $(\mathrm{SL}_d(\mathbb{Z})\cap H)\setminus H$, where $I_q^{(d-1)}:=[\mathrm{SL}_{d-1}(\mathbb{Z}):\mathrm{SL}_{d-1}(\mathbb{Z})\cap\Gamma(q)]$, and one can apply Proposition 2.8. Since $E_{i1}a_{t/q}^{-1}=(q/t)E_{i1}$, the above expression equals

$$\frac{q^{2d-1}I_q^{(d-1)}}{I_q\zeta(d)}\sum_{\substack{t\in\mathbb{N}\\(t,q)=1}}\frac{1}{t^d}\left[\sum_{\ell_1,\ldots,\ell_k\in\mathbb{Z}}F\left({}^{\mathrm{t}}\left(\frac{t+\ell_1q}{t}\mathbf{y},\ldots,\frac{t+\ell_kq}{t}\mathbf{y}\right)\right)\right.$$

$$+\sum_{r=1}^{k}\sum_{u\in\mathbb{N}}\sum_{D\in\mathfrak{D}_{r,u}}\sum_{\substack{\boldsymbol{\ell}={}^{\mathrm{t}}(\ell_1,\ldots,\ell_k)\\\in\mathcal{R}(D)}}\frac{N(D,u)^d}{u^{dr}}\times$$

$$\left.\int_{(\mathbb{R}^d)^r}F\left(\begin{pmatrix}\frac{t+\ell_1q}{t}\mathbf{y}\\ \vdots\\ \frac{t+\ell_kq}{t}\mathbf{y}\end{pmatrix}+\frac{D}{u}\begin{pmatrix}\mathbf{x}_1\\ \vdots\\ \mathbf{x}_r\end{pmatrix}a_{t/q}^{-1}g_{\mathbf{y}}\right)\mathrm{d}\mathbf{x}_1\cdots\mathrm{d}\mathbf{x}_r\right].$$

We will use the well-known fact that

$$\frac{I_q\zeta(d)}{q^{2d-1}I_q^{(d-1)}}=\sum_{\substack{t_1\in\mathbb{N}\\(t_1,q)=1}}\frac{1}{t_1^d}.$$

For the first summation, which is the case when $r=0$, put $t=t_1\cdot t_2$, where $t_1=\gcd(\ell_1,\ldots,\ell_k,t)$. By renaming $(\ell_1/t_1,\ldots,\ell_k/t_1)$ by $(\ell_1,\ldots,\ell_k)$, it follows that

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

$$\frac{q^{2d-1}I_q^{(d-1)}}{I_q\zeta(d)}\sum_{\substack{t\in\mathbb{N}\\(t,q)=1}}\frac{1}{t^d}\sum_{\ell_1,\ldots,\ell_k\in\mathbb{Z}}F\left({}^{\mathrm{t}}\left(\frac{t+\ell_1q}{t}\mathbf{y},\ldots,\frac{t+\ell_kq}{t}\mathbf{y}\right)\right)$$

$$=\frac{q^{2d-1}I_q^{(d-1)}}{I_q\zeta(d)}\sum_{\substack{t_1\in\mathbb{N}\\(t_1,q)=1}}\frac{1}{t_1^d}\left[F\left({}^{\mathrm{t}}(\mathbf{y},\ldots,\mathbf{y})\right)+\right.$$

$$\left.\sum_{\substack{t_2\in\mathbb{N}\\(t_2,q)=1}}\frac{1}{t_2^d}\sum_{\substack{\ell_1,\ldots,\ell_k\in\mathbb{Z}\\(\ell_1,\ldots,\ell_k,t_2)=1}}F\left({}^{\mathrm{t}}\left(\frac{t_2+\ell_1q}{t_2}\mathbf{y},\ldots,\frac{t_2+\ell_kq}{t_2}\mathbf{y}\right)\right)\right]$$

$$=F\left({}^{\mathrm{t}}(\mathbf{y},\ldots,\mathbf{y})\right)+\sum_{\substack{t\in\mathbb{N}\\(t,q)=1}}\frac{1}{t^d}\sum_{\substack{\ell_1,\ldots,\ell_k\in\mathbb{Z}\\(\ell_1,\ldots,\ell_k,t)=1}}F\left({}^{\mathrm{t}}\left(\frac{t+\ell_1q}{t}\mathbf{y},\ldots,\frac{t+\ell_kq}{t}\mathbf{y}\right)\right).$$

For the case when $r=k$, we only have $u=1$, $D=\mathrm{Id}_k$ and $\mathcal{R}(D)=\{{}^{\mathrm{t}}(0,\ldots,0)\}$. After a change of variables, the integral in this case is

$$\frac{q^{2d-1}I_q^{(d-1)}}{I_q\zeta(d)}\sum_{\substack{t\in\mathbb{N}\\(t,q)=1}}\frac{1}{t^d}\int_{(\mathbb{R}^d)^k}F\left(\begin{pmatrix}\mathbf{y}\\\vdots\\\mathbf{y}\end{pmatrix}+\begin{pmatrix}\mathbf{x}_1\\\vdots\\\mathbf{x}_k\end{pmatrix}a_{t/q}^{-1}g_{\mathbf{y}}\right)\mathrm{d}\mathbf{x}_1\cdots\mathrm{d}\mathbf{x}_k$$

$$=\frac{q^{2d-1}I_q^{(d-1)}}{I_q\zeta(d)}\sum_{\substack{t\in\mathbb{N}\\(t,q)=1}}\frac{1}{t^d}\int_{(\mathbb{R}^d)^k}F\begin{pmatrix}\mathbf{x}_1\\\vdots\\\mathbf{x}_k\end{pmatrix}\mathrm{d}\mathbf{x}_1\cdots\mathrm{d}\mathbf{x}_k$$

$$=\int_{(\mathbb{R}^d)^k}F\begin{pmatrix}\mathbf{x}_1\\\vdots\\\mathbf{x}_k\end{pmatrix}\mathrm{d}\mathbf{x}_1\cdots\mathrm{d}\mathbf{x}_k.$$

Suppose that for $1\le r\le k-1$ and $u\in\mathbb{N}$, $D\in\mathfrak{D}_{r,u}^k$ and $\mathcal{R}(D)$ are given. By rearranging the summation, it holds that

$$\frac{q^{2d-1}I_q^{(d-1)}}{I_q\zeta(d)}\sum_{\substack{t\in\mathbb{N}\\(t,q)=1}}\frac{1}{t^d}\sum_{\substack{\boldsymbol{\ell}={}^{\mathrm{t}}(\ell_1,\ldots,\ell_k)\\\in\mathcal{R}(D)}}\frac{N(D,u)^d}{u^{dr}}$$

$$\int_{(\mathbb{R}^d)^r}F\left(\begin{pmatrix}\frac{t+\ell_1q}{t}\mathbf{y}\\\vdots\\\frac{t+\ell_kq}{t}\mathbf{y}\end{pmatrix}+\frac{D}{u}\begin{pmatrix}\mathbf{x}_1\\\vdots\\\mathbf{x}_r\end{pmatrix}a_{t/q}^{-1}g_{\mathbf{y}}\right)\mathrm{d}\mathbf{x}_1\cdots\mathrm{d}\mathbf{x}_r$$

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

$$= \frac{q^{2d-1} I_q^{(d-1)}}{I_q \zeta(d)} \sum_{\substack{t_1 \in \mathbb{N} \\ (t_1,q)=1}} \frac{1}{t_1^d} \left[\frac{N(D,u)^d}{u^{dr}} \int_{(\mathbb{R}^d)^r} F\left(\begin{pmatrix} \mathbf{y} \\ \vdots \\ \mathbf{y} \end{pmatrix} + \frac{D}{u} \begin{pmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_r \end{pmatrix} \right) \mathrm{d}\mathbf{x}_1 \cdots \mathrm{d}\mathbf{x}_r \right.$$

$$\left. + \sum_{\substack{t \in \mathbb{N} \\ (t,q)=1}} \sum_{\substack{\boldsymbol{\ell} \in \\ P_t(\mathcal{R}(D))}} \frac{1}{t^d} \frac{N(D,u)^d}{u^{dr}} \int_{(\mathbb{R}^d)^r} F\left(\begin{pmatrix} \frac{t+\ell_1 q}{t}\mathbf{y} \\ \vdots \\ \frac{t+\ell_k q}{t}\mathbf{y} \end{pmatrix} + \frac{D}{u} \begin{pmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_r \end{pmatrix} \right) \mathrm{d}\mathbf{x}_1 \cdots \mathrm{d}\mathbf{x}_r \right]$$

$$= \frac{N(D,u)^d}{u^{dr}} \int_{(\mathbb{R}^d)^r} F\left(\begin{pmatrix} \mathbf{y} \\ \vdots \\ \mathbf{y} \end{pmatrix} + \frac{D}{u} \begin{pmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_r \end{pmatrix} \right) \mathrm{d}\mathbf{x}_1 \cdots \mathrm{d}\mathbf{x}_r$$

$$+ \sum_{\substack{t \in \mathbb{N} \\ (t,q)=1}} \sum_{\substack{\boldsymbol{\ell} \in \\ P_t(\mathcal{R}(D))}} \frac{N(D,u)^d}{t^d \cdot u^{dr}} \int_{(\mathbb{R}^d)^r} F\left(\begin{pmatrix} \frac{t+\ell_1 q}{t}\mathbf{y} \\ \vdots \\ \frac{t+\ell_k q}{t}\mathbf{y} \end{pmatrix} + \frac{D}{u} \begin{pmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_r \end{pmatrix} \right) \mathrm{d}\mathbf{x}_1 \cdots \mathrm{d}\mathbf{x}_r,$$

where for the first equality, as before, we put $t = t_1 t_2$ with $t_1 = \gcd(\boldsymbol{\ell}, t)$ and rename t_2 and $\boldsymbol{\ell}/t_1$ by t and $\boldsymbol{\ell}$, respectively. This completes the proof of the proposition. □

To prove Theorem 2.13, we need one more lemma which has appeared in [6, (3.6)] and also in [11, (7.25)].

Lemma 2.11 [6, (3.6)]. *For a Borel measurable function $\varphi : X_q \times \mathbb{R}^d \to \mathbb{R}_{\geq 0}$, we have*

$$\frac{1}{I_q} \int_{X_q} \sum_{\mathbf{m} \in \mathbb{Z}^d} \varphi\left(\Gamma(q)g, \left(\mathbf{m} + \frac{\mathbf{p}}{q} \right) g \right) \mathrm{d}\mu(g) = \int_{\mathbb{R}^d \setminus \{\mathbf{0}\}} \int_{X_q(\mathbf{y})} \varphi(\Gamma(q)g, \mathbf{y}) \, \mathrm{d}\nu_{\mathbf{y}}(g) \, \mathrm{d}\mathbf{y}.$$

Proof of Theorem 2.7. Let $F : (\mathbb{R}^d)^k \to \mathbb{R}_{\geq 0}$ be compactly supported bounded and

$$\varphi(\Gamma(q)g, \mathbf{y}) = \sum_{\substack{\mathbf{m}_i \in \mathbb{Z}^d \\ 1 \leq i \leq k-1}} F\left(\mathbf{y}, \left(\mathbf{m}_1 + \frac{\mathbf{p}}{q} \right) g, \ldots, \left(\mathbf{m}_{k-1} + \frac{\mathbf{p}}{q} \right) g \right)$$

$$= \mathcal{S}_{k-1}(F_{\mathbf{y}}) \left(\left(\mathbb{Z}^d + \frac{\mathbf{p}}{q} \right) g \right),$$

where $F_{\mathbf{y}} : (\mathbb{R}^d)^{k-1} \to \mathbb{R}_{\geq 0}$ is defined by

$$F_{\mathbf{y}}(\mathbf{y}_2, \ldots, \mathbf{y}_k) = F(\mathbf{y}, \mathbf{y}_2, \ldots, \mathbf{y}_k).$$

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

By Lemma 2.11, we have that

$$\frac{1}{J_q}\int_{Y_{\mathbf{p}/q}} \mathcal{S}_k(F)(\Lambda)\,\mathrm{d}\mu(\Lambda) = \frac{1}{I_q}\int_{X_q}\sum_{\mathbf{m}\in\mathbb{Z}^d}\varphi\left(\Gamma(q)g, \left(\mathbf{m}+\frac{\mathbf{p}}{q}\right)g\right)\mathrm{d}\mu(g)$$
$$= \int_{\mathbb{R}^d\setminus\{\mathbf{0}\}}\int_{X_q(\mathbf{y})}\varphi(\Gamma(q)g,\mathbf{y})\,\mathrm{d}\nu_{\mathbf{y}}(g)\,\mathrm{d}\mathbf{y}$$
$$= \int_{\mathbb{R}^d\setminus\{\mathbf{0}\}}\int_{X_q(\mathbf{y}_1)}\mathcal{S}_{k-1}(F_{\mathbf{y}_1})(\Gamma(q)g)\,\mathrm{d}\nu_{\mathbf{y}_1}(g)\,\mathrm{d}\mathbf{y}_1.$$

For the first equality, let us recall that $X_q = \Gamma(q)\setminus \mathrm{SL}_d(\mathbb{R})$ is a I_q/J_q-covering of $Y_{\mathbf{p}/q}$, where $I_q = [\mathrm{SL}_d(\mathbb{Z}) : \Gamma(q)]$ and $J_q = [\mathrm{SL}_d(\mathbb{Z}) : \Gamma_1(q)]$.

Thus, for $F \geq 0$, the equations in Theorem 2.7 immediately follow from Proposition 2.10, where we replace $\mathbf{y}_1$ by $\frac{1}{t}\mathbf{y}_1$.

Let us deal with the finiteness claim for $F \geq 0$: we will show that the LHS of (2.4) is finite. Define $\Phi : Y_{\mathbf{p}/q} \to X_d$ by $\Phi(\Lambda) := \Lambda - \Lambda$ for every $\Lambda \in Y_{\mathbf{p}/q}$. This map induces the measure $\Phi_*(\mu_q)$ on X_d, which is easily seen to be $\mathrm{SL}_d(\mathbb{R})$-invariant. Therefore, $\Phi_*(\mu_q)$ equals to μ up to a positive constant. In fact, Φ is the natural J_q-to-1 covering map from $Y_{\mathbf{p}/q}$ to X_d; thus, $\Phi_*(\mu) = J_q\mu$. For $\Lambda \in Y_{\mathbf{p}/q}$, we have $\Lambda \subseteq q^{-1}\Phi(\Lambda)$. Therefore, $\mathcal{S}_k(F)(\Lambda) \leq \mathcal{S}_k(F_q)(\Phi(\Lambda))$, where $F_q : (\mathbb{R}^d)^k \to \mathbb{R}_{\geq 0}$, $\mathbf{x} \mapsto F(q^{-1}\mathbf{x})$ is a compactly supported function. Hence,

$$\int_{Y_{\mathbf{p}/q}} \mathcal{S}_k(F)(\Lambda)\,\mathrm{d}\mu_q(\Lambda) \leq \int_{Y_{\mathbf{p}/q}} \mathcal{S}_k(F_q)(\Phi(\Lambda))\,\mathrm{d}\mu_q(\Lambda)$$
$$= J_q\int_{X_d}\mathcal{S}_k(F_q)(\Lambda)\,\mathrm{d}\mu(\Lambda) < \infty,$$

by Schmidt [17]. Thus, in the present case, the sum in the RHS of (2.4) is convergent.

We can now use classical techniques to prove Theorem 2.7 for a compactly supported bounded function $F : (\mathbb{R}^d)^k \to \mathbb{R}$. We first note that Theorem 2.7 holds for $F_+ := \max(F,0)$ and $F_- := \max(-F,0)$. Finiteness for the function $|F|$ implies that the sum with F in the RHS of (2.4) is absolutely convergent. Furthermore, $\mathcal{S}_k(F)(\Lambda) = \mathcal{S}_k(F_+)(\Lambda) - \mathcal{S}_k(F_-)(\Lambda)$ (for a.e. Λ); we can integrate and rearrange to see that Theorem 2.7 holds. □

2.3. Higher moment formulae revisited

For applications to Poisson distribution which are proved in the next section, we will need that the "admissible matrices" appearing in the higher moment formula for Y are contained in $\mathfrak{D}^{k'}_{r',u'}$ for some k', r' and u', which does not hold in Theorem 2.4 and Theorem 2.7. In the process of proving the needed variations of the higher moment formulae, we will define canonical sets of admissible matrices for each cases. In particular, we will see that the set of "congruence-admissible matrices" can be defined without using any choice of $\mathcal{R}(D)$ in Notation 2.6.

Let us first refine the higher moment formula for the space Y of affine lattices in $\mathbb{R}^d$.

Theorem 2.12. *Let $F : (\mathbb{R}^d)^k \to \mathbb{R}_{\geq 0}$ be bounded and compactly supported. For $d \geq 3$ and $3 \leq k \leq d$,*

$$\int_Y \mathcal{S}_k(F)(\Lambda)\, \mathrm{d}\mu_Y(\Lambda) = \sum_{m=1}^{k} \sum_{u \in \mathbb{N}} \sum_{\widetilde{D} \in \mathfrak{A}^k_{m,u}} \frac{N(\widetilde{D},u)^d}{u^{dm}} \int_{(\mathbb{R}^d)^m} F\left(\frac{\widetilde{D}}{u}\begin{pmatrix}\mathbf{y}_1\\ \vdots \\ \mathbf{y}_m\end{pmatrix}\right) \mathrm{d}\mathbf{y}_1 \cdots \mathrm{d}\mathbf{y}_m, \tag{2.8}$$

where $\mathfrak{A}^k_{m,u}$ is a subset of $\mathfrak{D}^k_{m,u}$ given by

$$\mathfrak{A}^k_{m,u} = \left\{ C \in \mathfrak{D}^k_{m,u} : \sum_{i=1}^{m} [C]^i = {}^{\mathrm{t}}(u,\ldots,u) \right\}.$$

Notice that when $m = 1$ and k, the only possible u is $u = 1$ and

$$\mathfrak{A}^k_{1,1} = \left\{ {}^{\mathrm{t}}(1,\ldots,1) \right\} \quad \text{and} \quad \mathfrak{A}^k_{k,1} = \{\mathrm{Id}_k\},$$

which corresponds to the first and second integrals of the RHS in (2.2), respectively.

Proof. Assume that $2 \leq m \leq k-1$ so that $1 \leq r := m-1 \leq k-2$. Recall the $k \times m$ matrix D' in Theorem 2.4 from $D \in \mathfrak{D}^{k-1}_{r,u}$.

Take the map

$$D \in \mathfrak{D}^{k-1}_{r,u} \quad \mapsto \quad uD' \quad \mapsto \quad \widetilde{D} \in \mathfrak{D}^k_{m,u}, \tag{2.9}$$

where we define $[\widetilde{D}]^1 = [uD']^1 - \sum_{j=2}^m [uD']^j$ and $[\widetilde{D}]^j = [uD']^j$ for $2 \leq j \leq m$. Clearly, the map is injective and $\widetilde{D} \in \mathfrak{A}^k_{m,u}$.

Conversely, for any $C \in \mathfrak{A}^k_{m,u}$, denote by D the right-bottom minor of C of the size $(k-1) \times r$. Then one can verify that $D \in \mathfrak{D}^{k-1}_{r,u}$ and $\widetilde{D} = C$.

Moreover, it is easy to show from their definitions that

$$\frac{N(\widetilde{D},u)^d}{u^{d(r+1)}} = \frac{N(D,u)^d}{u^{dr}},$$

and the map $uD' \mapsto \widetilde{D}$ is the simple change of variables $\mathbf{y}_j + \mathbf{y}_1 \mapsto \mathbf{y}'_j$ for $2 \leq j \leq m$:

$$\int_{(\mathbb{R}^d)^m} F\left(D'\begin{pmatrix}\mathbf{y}_1\\ \vdots \\ \mathbf{y}_m\end{pmatrix}\right) \mathrm{d}\mathbf{y}_1 \cdots \mathrm{d}\mathbf{y}_m = \int_{(\mathbb{R}^d)^m} F\left(\frac{\widetilde{D}}{u}\begin{pmatrix}\mathbf{y}_1\\ \vdots \\ \mathbf{y}_m\end{pmatrix}\right) \mathrm{d}\mathbf{y}_1 \cdots \mathrm{d}\mathbf{y}_m. \qquad \square$$

In contrast to the affine case, in the congruence case it is difficult and complicated to describe the subset of matrices $\widetilde{D} \in \mathfrak{D}^k_{m,u_0}$, for given $1 \leq m \leq k$ and $u_0 \in \mathbb{N}$, such that

$$\widetilde{D}\mathbb{R}^m = D'\mathbb{R}^m \text{ or } D'_{t,\boldsymbol{\ell}}\mathbb{R}^m$$

for some $t \in \mathbb{N}$ with $(t,q) = 1$ and $\boldsymbol{\ell} \in P_t(\mathcal{R}(D))$ appearing in Theorem 2.7.

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

For each $u \in \mathbb{N}$ and $D \in \mathfrak{D}^{k-1}_{m-1,u}$, once we fix $\mathcal{R}(D)$ in Notation 2.6, by the map

$$D \mapsto uD' \mapsto \widetilde{D} \text{ as in (2.9)}$$
$$(D,t,\boldsymbol{\ell}) \mapsto u_0 D'_{t,\boldsymbol{\ell}} = u_0 \left(\begin{array}{c|c} 1 & 0 \cdots 0 \\ \hline (t+\ell_1 q)/t & \\ \vdots & \frac{1}{u}D \\ (t+\ell_{k-1}q)/t & \end{array} \right) \mapsto \widetilde{D}, \tag{2.10}$$

where $\widetilde{D}$ is defined by

$$[\widetilde{D}]^1 = [u_0 D'_{t,\boldsymbol{\ell}}]^1 - \sum_{j=1}^{m-1} \frac{t+\ell_{i_j} q}{t} [u_0 D'_{t,\boldsymbol{\ell}}]^{j+1} \quad \text{and} \quad [\widetilde{D}]^j = [u_0 D'_{t,\boldsymbol{\ell}}]^j \ (j = 2, \ldots, m),$$

and $u_0 \in \mathbb{N}$ is taken such that $\widetilde{D} \in \mathrm{Mat}_{k,m}(\mathbb{Z})$ with $\gcd \widetilde{D} = 1$. Clearly, $\widetilde{D} \in \mathfrak{D}^k_{m,u_0}$.

Hence, one can attempt to define such a subset $\mathfrak{C}^k_{m,u_0}$ of $\mathfrak{D}^k_{m,u_0}$ by

$$\mathfrak{C}^k_{m,u_0} := \left\{ C \in \mathfrak{D}^k_{m,u_0} : \begin{array}{c} C = \widetilde{D} \text{ for some } D \in \mathfrak{D}^{k-1}_{m-1,u} \text{ or} \\ (D,t,\boldsymbol{\ell}) \in \mathfrak{D}^{k-1}_{m-1,u} \times \mathbb{N} \times \mathcal{R}(D) \text{ in Notation 2.6} \\ \text{defined as in (2.10)} \end{array} \right\} \tag{2.11}$$

and reformulate the higher moment formula using these $\mathfrak{C}^k_{m,u_0}$.

As things stand, $\mathfrak{C}^k_{m,u_0}$ seems to depend on an ad hoc choice of a set of representatives $\mathcal{R}(D)$. However, the anonymous referee has kindly provided us with an argument using the Riesz representation theorem which shows that the set $\mathfrak{C}^k_{m,u_0}$ is independent to the choice of $\mathcal{R}(D)$ regardless of its role in the construction. With this as background, we now provide a cleaner definition of the set $\mathfrak{C}^k_{m,u_0}$, meaning that we do not need *an ad hoc choice* of $\mathcal{R}(D)$ for each $D \in \mathfrak{D}^{k-1}_{m-1,r}$. This definition was also suggested by the referee.

Theorem 2.13. *Let $d \geq 3$ and $1 \leq k \leq d-1$. Let $F : (\mathbb{R}^d)^k \to \mathbb{R}_{\geq 0}$ be bounded and compactly supported. Then*

$$\begin{aligned} \int_{Y_{\mathbf{p}/q}} \mathcal{S}_k(F)(\Lambda)\, \mathrm{d}\mu_q(\Lambda) &= \int_{(\mathbb{R}^d)^k} F\left({}^{\mathrm{t}}(\mathbf{y}_1, \ldots, \mathbf{y}_k)\right) \mathrm{d}\mathbf{y}_1 \cdots \mathrm{d}\mathbf{y}_k \\ &+ \sum_{m=1}^{k-1} \sum_{u \in \mathbb{N}} \sum_{\widetilde{D} \in \mathfrak{C}^k_{m,u_0}} \frac{N(\widetilde{D},u_0)^d}{u_0^{dm}} \int_{(\mathbb{R}^d)^m} F\left(\frac{\widetilde{D}}{u_0} \begin{pmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_m \end{pmatrix} \right) \mathrm{d}\mathbf{y}_1 \cdots \mathrm{d}\mathbf{y}_m, \end{aligned} \tag{2.12}$$

where for $1 \leq m \leq k-1$ and $u_0 \in \mathbb{N}$,

$$\mathfrak{C}^k_{m,u_0} = \left\{ C \in \mathfrak{D}^k_{m,u_0} : \exists \mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_k \end{pmatrix} \in \frac{C}{u_0}\Lambda_C \ \textit{s.t.} \ \begin{array}{c} \gcd(v_1, q) = 1, \\ v_1 \equiv \cdots \equiv v_k \mod q, \textit{ and} \\ |v_1| = \min(\mathbb{N} \cap \{\mathbf{v}' \cdot \mathbf{e}_1 : \mathbf{v}' \in \frac{C}{u}\Lambda_C\}) \end{array} \right\}. \tag{2.13}$$

Here, $\mathbf{e}_1 = {}^{\mathrm{t}}(1,0,\ldots,0) \in \mathbb{R}^k$ and $\mathbf{v}_1 \cdot \mathbf{v}_2 = {}^{\mathrm{t}}\mathbf{v}_1\mathbf{v}_2$ is the standard dot product of $\mathbb{R}^k$.

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

Proof. We will consider the case when $m \geq 2$; then the case when $m = 1$ would be easily seen. Let us first show that the sets defined as in (2.11) and (2.13) are identical.

Assume that $C = \widetilde{D}$ is an element of the set in (2.11). Note that

$$\frac{C}{u_0}\Lambda_C = C\mathbb{R}^m \cap \mathbb{Z}^k = D'\mathbb{R}^m \cap \mathbb{Z}^k \text{ or } D'_{t,\boldsymbol{\ell}}\mathbb{R}^m \cap \mathbb{Z}^k,$$

where D' or $D'_{t,\boldsymbol{\ell}}$ is as in (2.10) for some $D \in \mathfrak{D}^{k-1}_{r,u}$ $(r := m-1)$, or $D \in \mathfrak{D}^{k-1}_{r,u}$, $t \in \mathbb{N}$ with $(t,q) = 1$, and $\boldsymbol{\ell} = {}^{\mathrm{t}}(\ell_1,\ldots,\ell_{k-1}) \in \mathcal{R}(D) \subset \mathbb{Z}^{k-1}$ settled in Notation 2.6, respectively. In particular,

$$\mathbf{v} := {}^{\mathrm{t}}(1,\ldots,1) \quad \text{or} \quad {}^{\mathrm{t}}(t, t+\ell_1 q, \ldots, t+\ell_{k-1}q) \in \frac{C}{u_0}\Lambda_C,$$

respectively.

It suffices to show that

$$\frac{C}{u_0}\Lambda_C = \mathbb{Z}\mathbf{v} \oplus {}^{\mathrm{t}}\left(0, \frac{D}{u}\Lambda_D\right), \tag{2.14}$$

where ${}^{\mathrm{t}}\big(0, \frac{D}{u}\Lambda_D\big)$ is the embedded image of $\frac{D}{u}\Lambda_D \subseteq \mathbb{R}^{k-1}$ into the last $(k-1)$ coordinates of $\mathbb{R}^k$, since then it gives the fact that $v_1 = \min(\mathbb{N} \cap \{\mathbf{v}' \cdot \mathbf{e}_1 : \mathbf{v}' \in \frac{C}{u}\Lambda_C\})$. The inclusion of the reverse direction is obvious.

Pick an arbitrary $\mathbf{w} \in \frac{C}{u_0}\Lambda_C$. Since $C\mathbb{R}^m = \mathbb{R}\mathbf{v} \oplus {}^{\mathrm{t}}(0, D\mathbb{R}^{r-1})$, one can take

$$\mathbf{w} = \frac{c_1}{c_2}\mathbf{v} + {}^{\mathrm{t}}(0, \mathbf{v}'),$$

where $c_1 \in \mathbb{Z}$, $c_2 \in \mathbb{N}$ with $\gcd(c_1, c_2) = 1$ and $\mathbf{v}' \in D\mathbb{R}^{r-1} \subseteq \mathbb{R}^{k-1}$. Since $\mathbf{w} \in \mathbb{Z}^k$, it holds that

$$\frac{c_1}{c_2}v_1 \in \mathbb{Z} \Leftrightarrow c_2 | v_1 \quad \text{and} \quad \frac{c_1}{c_2}q\boldsymbol{\ell} + \mathbf{v}' \in \mathbb{Z}^{k-1}. \tag{2.15}$$

If $v_1 = 1$, then automatically $c_2 = 1$ and $\mathbf{v}' \in \mathbb{Z}^{k-1} \cap D\mathbb{R}^r = \frac{D}{u}\Lambda_D$, which implies that $\mathbf{w} \in \mathbb{Z}\mathbf{v} \oplus {}^{\mathrm{t}}\big(0, \frac{D}{u}\Lambda_D\big)$.

Suppose that $v_1 = t \geq 2$ so that $\boldsymbol{\ell} \neq {}^{\mathrm{t}}(0,\ldots,0)$. Denote $\ell = \gcd(\boldsymbol{\ell})$ and $\widehat{\boldsymbol{\ell}} = \frac{1}{\ell}\boldsymbol{\ell}$, the primitive vector of the $\boldsymbol{\ell}$-direction. Following Notation 2.6, let $\mathbf{b}_{k-m}, \ldots, \mathbf{b}_{k-1}$ be the basis of $\frac{D}{u}\Lambda_D$. Then it follows from the definition of $\mathcal{R}(D)$ that $\{\widehat{\boldsymbol{\ell}}, \mathbf{b}_{k-m}, \ldots, \mathbf{b}_{k-1}\}$ is a primitive set; that is,

$$\mathbb{Z}\widehat{\boldsymbol{\ell}} \oplus \mathbb{Z}\mathbf{b}_{k-m} \oplus \cdots \oplus \mathbb{Z}\mathbf{b}_{k-1} = \left(\mathbb{R}\widehat{\boldsymbol{\ell}} \oplus \mathbb{R}\mathbf{b}_{k-m} \oplus \cdots \oplus \mathbb{R}\mathbf{b}_{k-1}\right) \cap \mathbb{Z}^{k-1}.$$

Hence, the second condition in (2.15) implies that

$$\frac{c_1}{c_2}q\ell \in \mathbb{Z} \Leftrightarrow c_2 | q\ell \quad \text{and} \quad \mathbf{v}' \in \frac{D}{u}\Lambda_D.$$

Since $c_2 | t$ from (2.15) as well and $\gcd(t, q\ell) = 1$, we obtain the fact that $c_2 = 1$ and $\mathbf{w} \in \mathbb{R}\mathbf{v} \oplus {}^{\mathrm{t}}\big(0, \frac{D}{u}\Lambda_D\big)$. And this shows one inclusion.

Conversely, let $C \in \mathfrak{D}^k_{m,u_0}$ be such that there is $\mathbf{v} \in \frac{C}{u_0}\Lambda_C$ satisfying three conditions in (2.13). One can easily extract $D \in \mathfrak{D}^k_{r,u}$ from the right-bottom $(k-1) \times r$-minor of C by making a primitive matrix which will be D, and u is the unique nonzero entry of the first nonzero row of D. Fix any $\mathcal{R}(D)$.

Notice that the third condition is equivalent to saying that

$$\frac{C}{u_0}\Lambda_C = \mathbb{Z}\mathbf{v} \oplus \left(\frac{C}{u_0}\mathbb{R}^m \cap {}^{\mathrm{t}}(0,\mathbb{Z}^{k-1})\right) = \mathbb{Z}\mathbf{v}^{\mathrm{t}}\left(0,\frac{D}{u}\Lambda_D\right).$$

Set $v_1 = t$ (if $v_1 < 0$, replace $\mathbf{v}$ by $-\mathbf{v}$). From the first and second conditions, $\gcd(t,q) = 1$ and $\mathbf{v} = {}^{\mathrm{t}}(t, t+\ell'_1 q, \ldots, t+\ell'_{k-1}q)$ for some $\boldsymbol{\ell}' = {}^{\mathrm{t}}(\ell'_1,\ldots,\ell'_{k-1}) \in \mathbb{Z}^{k-1}$. Since $\mathcal{R}(D) \simeq \mathbb{Z}^{k-1}/\frac{D}{u}\Lambda_D$, there is the unique $\boldsymbol{\ell} = {}^{\mathrm{t}}(\ell_1,\ldots,\ell_{k-1}) \in \mathcal{R}(D)$ for which $\boldsymbol{\ell} + \frac{D}{u}\Lambda_D = \boldsymbol{\ell}' + \frac{D}{u}\Lambda_D$ and

$$\frac{C}{u_0}\Lambda_C = \mathbb{Z}\begin{pmatrix} t \\ t+\ell_1 q \\ \vdots \\ t+\ell_{k-1}q \end{pmatrix} \oplus \begin{pmatrix} 0 \\ \\ \frac{D}{u}\Lambda_D \end{pmatrix}.$$

This shows that ${}^{\mathrm{t}}(t, t+\ell_1 q, \ldots, t+\ell_{k-1}q)$ is primitive. If $\boldsymbol{\ell} = {}^{\mathrm{t}}(0,\ldots,0)$, then it holds that $C = \widetilde{D}$ of the first type described in (2.11). If $\boldsymbol{\ell} \neq {}^{\mathrm{t}}(0,\ldots,0)$, then $\gcd(\boldsymbol{\ell},t) = 1$ so that $\boldsymbol{\ell} \in P_t(\mathcal{R}(D))$ and $C = \widetilde{D}$ defined from $(D,t,\boldsymbol{\ell})$.

Now, to establish the theorem, considering the change of variables on $\mathbf{y}_1$ in Theorem 2.7, it is left to show that

$$\frac{N(\widetilde{D},u_0)^d}{u_0^{dm}} = \frac{N(D,u)^d}{t^d \cdot u^{dr}},$$

where we put $t = 1$ when $\widetilde{D} \in \mathfrak{C}^k_{m,u_0}$ is of the first type in (2.11). Recall that $N(\widetilde{D},u_0)$ is the number of integral solutions $\mathbf{z} = {}^{\mathrm{t}}(z_1,\ldots,z_m) \in \mathbb{Z}^m$ modulo u_0 for which $\frac{\widetilde{D}}{u_0}\mathbf{z} \in \mathbb{Z}^k$. Equivalently, $N(\widetilde{D},u_0)$ is the number of integral solutions $\mathbf{z} \in \mathbb{Z}^m$ modulo u_0 for which $D'\mathbf{z} \in \mathbb{Z}^k$ or $D'_{t,\boldsymbol{\ell}}\mathbf{z} \in \mathbb{Z}^k$, respectively.

Based on (2.14), it follows that $z_1 \in t\mathbb{Z}$ and there are (u_0/t)-number of such $z_1 \in \mathbb{Z}$ modulo u_0. Moreover, as long as $z_1 \in t\mathbb{Z}$, $t[D']^1 \in \mathbb{Z}^k$ and we reduce that

$$\frac{D}{u}{}^{\mathrm{t}}(z_2,\ldots,z_k) \in \mathbb{Z}^{k-1}$$

modulo u_0, and the number of such ${}^{\mathrm{t}}(z_2,\ldots,z_k)$ is $(u_0/u)^r(N(D,u))$. Therefore,

$$\frac{N(\widetilde{D},u_0)^d}{u_0^{dm}} = \frac{(u_0/t)^d \cdot (u_0/u)^{dr} N(D,u)^d}{u_0^{dm}} = \frac{N(D,u)^d}{t^d \cdot u^{dr}}. \qquad \square$$

3. Poissonian Behaviour

3.1. Affine case

In this section, we prove Theorem 1.1. Recall that for each $d \geq 2$, we set $\mathcal{S} = \{S_t : t \geq 0\}$ to be an increasing family of subsets of $\mathbb{R}^d$ with $\mathrm{vol}(S_t) = t$, and for $\Lambda \in Y$, set

$$N_t(\Lambda) := \#(S_t \cap \Lambda).$$

Denote by $\{N^\lambda(t) : t \geq 0\}$ a Poisson process on the non-negative real line with intensity λ.

For $\Lambda \in Y$, we order the lengths of vectors in Λ as $0 \le \ell_1 \le \ell_2 \le \ell_3 \le \cdots$, and let $\mathscr{V}_i$ denote the volume of the closed ball of radius ℓ_i centered at origin. If we take $\mathcal{S} = \{B_t : t \ge 0\}$ to be the family of closed balls with $\mathrm{vol}(B_t) = t$ around origin, then

$$N_t(\Lambda) = \#\{i : \mathscr{V}_i \le t\}.$$

In this specific case, Theorem 1.1 is equivalent to the following:

Theorem 3.1. *For any fixed n, the n-dimensional random variable $(\mathscr{V}_1, \ldots, \mathscr{V}_n)$ converges in distribution to the distribution of the first n points of a Poisson process on the non-negative real line with intensity 1 as $d \to \infty$.*

In this form, the theorem determines the limit distribution of lengths of vectors in a random lattice as $d \to \infty$.

We will now prove the above general Theorem 1.1 by proving a joint moment formula for $N_t(\cdot)$. Let $k \ge 1$ and $0 \le V_1 \le \cdots \le V_k$. We use, by abuse of notation, $N_i(\cdot)$ to denote $N_{V_i}(\cdot)$. Note that $N_i(\cdot) = \widehat{\rho_i}(\cdot)$, where ρ_i is the characteristic function of S_{V_i}. We calculate, following Södergren [23], the “main term” of the joint moment of N_i's. In this regard, we apply Theorem 2.12 with $F = \prod_{i=1}^{k} \rho_i$ defined as

$$F\begin{pmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_k \end{pmatrix} = \prod_{i=1}^{k} \rho_i(\mathbf{y}_i).$$

We consider the sub-collection of the RHS of (2.8) consisting of terms corresponding to $m = 1$ and k, and terms from the sum corresponding to $u = 1$ and $\widetilde{D} \in \mathfrak{A}^k_{m,1}$ satisfying that $\widetilde{D}$ has exactly one nonzero entry in each row, with all nonzero entries of $\widetilde{D}$ being of modulus 1. The set of such matrices $\widetilde{D} \in \mathfrak{A}^k_{m,1}$ is $\mathfrak{M}^k$, where

$$\mathfrak{R}^k_1 = \bigcup_{2 \le m \le k-1} \left(\left(\bigcup_{u \ge 2} \mathfrak{A}^k_{m,u} \right) \cup \left\{ \widetilde{D} = \left(\widetilde{D}_{ij} \right) \in \mathfrak{A}^k_{m,1} : \exists |\widetilde{D}_{ij}| \ge 2 \right\} \right),$$

$$\mathfrak{R}^k_2 = \left\{ \widetilde{D} \in \left(\bigcup_{2 \le m \le k-1} \mathfrak{A}^k_{m,1} \right) \setminus \mathfrak{R}^k_1 : \begin{array}{c} \exists \text{ row such that at least} \\ \text{two entries are nonzero} \end{array} \right\},$$

$$\mathfrak{M}^k = \left(\bigcup_{2 \le m \le k-1} \mathfrak{A}^k_{m,1} \setminus \left(\mathfrak{R}^k_1 \cup \mathfrak{R}^k_2 \right) \right) \cup \left\{ \mathrm{Id}_k, \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \right\}.$$

Here, we want to mention that we will use the same notations $\mathfrak{R}^k_1$, $\mathfrak{R}^k_2$ and $\mathfrak{M}^k$ for the analogous (but different) sets in each subsection (see Subsection 3.2 and Section 5). This will hopefully cause no confusion.

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

We denote this sub-collection of the RHS of (2.8) as $M_{d,k}^{\text{affine}}$ and the rest of the terms as $R_{d,k}^{\text{affine}}$. That is,

$$\mathbb{E}\left(\prod_{i=1}^{k} N_i\right) = M_{d,k}^{\text{affine}} + R_{d,k}^{\text{affine}},$$

where

$$M_{d,k}^{\text{affine}} = \sum_{\widetilde{D} \in \mathfrak{M}^k} \int_{(\mathbb{R}^d)^m} \prod_{i=1}^{k} \rho_i \left(\widetilde{D} \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_m \end{pmatrix} \right) \mathrm{d}\mathbf{y}_1 \, \mathrm{d}\mathbf{y}_2 \cdots \mathrm{d}\mathbf{y}_m \tag{3.1}$$

and

$$R_{d,k}^{\text{affine}} = \sum_{\widetilde{D} \in \mathfrak{R}_1^k \cup \mathfrak{R}_2^k} \frac{N(\widetilde{D},u)^d}{u^{dm}} \int_{(\mathbb{R}^d)^m} \prod_{i=1}^{k} \rho_i \left(\frac{1}{u} \widetilde{D} \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_m \end{pmatrix} \right) \mathrm{d}\mathbf{y}_1 \, \mathrm{d}\mathbf{y}_2 \cdots \mathrm{d}\mathbf{y}_m. \tag{3.2}$$

Let (α,β) be a division of $\{1,\ldots,k\}$; that is, $\alpha = \{\alpha_1 < \cdots < \alpha_m\}$ and $\beta = \{\beta_1 < \cdots < \beta_{k-r}\}$ are complementary subsets of $\{1,\ldots,k\}$ with $\alpha \neq \varnothing$. Define

$$\mathfrak{M}_{\alpha,\beta}^{\text{affine}} := \left\{ \widetilde{D} \in \mathfrak{M}^k : I_{\widetilde{D}} = \alpha \right\}$$

and let $M_{\alpha,\beta}^{\text{affine}}$ denote the cardinality of $\mathfrak{M}_{\alpha,\beta}^{\text{affine}}$. We allow for the case $(\alpha,\beta) = (\{1,\ldots,k\},\varnothing)$, in which case $\mathfrak{M}_{\alpha,\beta}^k = \{\mathrm{Id}_k\}$. Thus, we can rewrite (3.1) as

$$M_{d,k}^{\text{affine}} = \sum_{(\alpha,\beta)} \sum_{\widetilde{D} \in \mathfrak{M}_{\alpha,\beta}^{\text{affine}}} \int_{(\mathbb{R}^d)^m} \prod_{i=1}^{k} \rho_i \left(\widetilde{D} \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_m \end{pmatrix} \right) \mathrm{d}\mathbf{y}_1 \, \mathrm{d}\mathbf{y}_2 \cdots \mathrm{d}\mathbf{y}_m, \tag{3.3}$$

where the outer sum is over all possible divisions (α,β) of $\{1,\ldots,k\}$.

Remark 3.2. It follows from the definition of $\widetilde{D}$ that for $\widetilde{D} \in \mathfrak{M}^k$, the nonzero entries of the matrix $\widetilde{D}$ can only be 1. Since $\widetilde{D} \notin \mathfrak{R}_1^k$, we already know that entries of $\widetilde{D} \in \{0,\pm 1\}$. The fact that -1 is not possible for entries of $\widetilde{D}$ comes from notations in Theorem 2.12. Suppose that there is a row having -1 in its entries. Let $(x_1,x_2,\ldots,x_m)$ be such a row. If $x_1 = -1$, since $x_1 = 1 - \sum_{\ell=2}^{m} x_\ell$, there should be at least one nonzero element in $(x_2,\ldots,x_m)$, which contradicts the fact that in each row, there is only one nonzero entry. One can also obtain a contradiction when one assumes that there is some $2 \leq i_0 \leq m$ for which $x_{i_0} = -1$.

Lemma 3.3. *With notations as above,*

$$M_{d,k}^{\textit{affine}} = \sum_{(\alpha,\beta)} M_{\alpha,\beta}^{\textit{affine}} \prod_{i=1}^{m} V_{\alpha_i}. \tag{3.4}$$

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

Proof. Consider any matrix $\widetilde{D} = \left(\widetilde{D}_{ij}\right) \in \mathfrak{M}^{\text{affine}}_{\alpha,\beta}$ and let λ_ℓ be such that $\widetilde{D}_{\beta_\ell,\lambda_\ell} = 1$, $1 \le \ell \le k-m$. Then, as V_i's are increasing, the following calculation finishes the proof:

$$\int_{(\mathbb{R}^d)^m} \prod_{i=1}^{k} \rho_i \left(\widetilde{D} \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_m \end{pmatrix} \right) \mathrm{d}\mathbf{y}_1\, \mathrm{d}\mathbf{y}_2 \cdots \mathrm{d}\mathbf{y}_m = \int_{(\mathbb{R}^d)^m} \prod_{i=1}^{k} \rho_i \left(\sum_{j=1}^{m} \widetilde{D_{ij}} \mathbf{y}_j \right) \mathrm{d}\mathbf{y}_1 \cdots \mathrm{d}\mathbf{y}_m$$

$$= \int_{(\mathbb{R}^d)^m} \prod_{i=1}^{m} \rho_{\alpha_i}(\mathbf{y}_i) \prod_{\ell=1}^{k-m} \rho_{\beta_\ell}(\mathbf{y}_{\lambda_\ell})\, \mathrm{d}\mathbf{y}_1 \cdots \mathrm{d}\mathbf{y}_m = \int_{(\mathbb{R}^d)^m} \prod_{i=1}^{m} \rho_{\alpha_i}(\mathbf{y}_i) = \prod_{i=1}^{m} V_{\alpha_i}. \qquad \square$$

We shall now mention some estimates regarding $R^{\text{affine}}_{d,k}$. These estimates are originally due to Rogers [14, 16], and they were generalised to Lemma 3.4 (below) by Södergren [23]. For $D \in \mathfrak{D}^k_{r,u}$ set

$$I(D,u) := \int_{(\mathbb{R}^d)^r} \prod_{i=1}^{k} \rho_i \left(\frac{1}{u} D \begin{pmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_r \end{pmatrix} \right) \mathrm{d}\mathbf{y}_1 \cdots \mathrm{d}\mathbf{y}_r.$$

Lemma 3.4 (Estimates from [14], [16] and [23]). *For $d > [k^2/4] + 2$,*

(i) $\displaystyle\sum_{r=1}^{k} \sum_{u=2}^{\infty} \sum_{D \in \mathfrak{D}^k_{r,u}} \frac{N(D,u)^d}{u^{dr}} \cdot I(D,u) = O\left(2^{-d}\right),$

(ii) $\displaystyle\sum_{r=1}^{k} \sum_{D \in \mathfrak{D}^{k,1}_{r,1}} I(D,u) = O\left(2^{-d}\right),$ *where $\mathfrak{D}^{k,1}_{r,1} \subseteq \mathfrak{D}^k_{r,1}$ contains matrices D such that* $\max|d_{ij}| \ge 2$,

(iii) $\displaystyle\sum_{r=1}^{k} \sum_{D \in \mathfrak{D}^{k,2}_{r,1}} I(D,u) = O\left((3/4)^{d/2}\right),$ *where $\mathfrak{D}^{k,2}_{r,1} \subseteq \mathfrak{D}^k_{r,1}$ contains matrices D such that* $\max|d_{ij}| = 1$ *and at least one row of D has at least two nonzero entries.*

Proof. The proof of this lemma can be found in [23, Proposition 2, Lemma 1 and Lemma 2]. The main ingredients in Södergren's proof are [14, §9] and the contents of [16, §4].

We remark that we only need the fact that $N(D,u)^d/u^{dr} \le 1/u^d$ to prove Property (i). Hence, one can use Lemma 3.4 for applications with the space $Y_{\mathbf{p}/q}$ as well as the space Y in Section 3 and Section 5. $\square$

Rogers' estimate shows the following:

Lemma 3.5.

$$R^{\textit{affine}}_{d,k} = O\left((3/4)^{d/2}\right).$$

Proof. It follows from (3.2) that $R^{\text{affine}}_{d,k}$ is less than the sum of LHS's in Lemma 3.4. $\square$

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

The lemmas above combine to give us the following theorem:

Theorem 3.6.

$$\mathbb{E}\left(\prod_{i=1}^{k} N_i\right) \to \sum_{(\alpha,\beta)} M_{\alpha,\beta}^{affine} \prod_{i=1}^{r} V_{\alpha_i} \tag{3.5}$$

as $d \to \infty$.

3.1.1. Proof of Theorem 1.1. This proof closely follows the proof of Theorem 1 in [23, §4]. Let us discuss the Poisson process $\{N^\lambda(t) : t \geq 0\}$. By definition, $N^\lambda(t)$ denotes the number of points falling in the interval $[0,t]$ and $N^\lambda(t)$ is Poisson distributed with expectation λt. By $0 \leq T_1 \leq T_2 \leq T_3 \leq \cdots$, let us denote the points of the Poisson process.

Lemma 3.7. *Let* $k \geq 1$ *and let* $\mathscr{P}(k)$ *denote the set of partitions of* $\{1,\ldots,k\}$. *For* $1 \leq i \leq k$, *let* $f_i : \mathbb{R}_{\geq 0} \to \mathbb{R}$ *be functions satisfying* $\prod_{i\in B} f_i \in L^1(\mathbb{R}_{\geq 0})$ *for every nonempty subset* $B \subseteq \{1,\ldots,k\}$. *Then*

$$\mathbb{E}\left(\prod_{i=1}^{k}\left(\sum_{\ell=1}^{\infty} f_i(T_\ell)\right)\right) = \sum_{P\in\mathscr{P}(k)} \lambda^{\#P}\left(\int_0^\infty \prod_{i\in B} f_i(x)\,\mathrm{d}x\right). \tag{3.6}$$

Proof. The proof of this lemma is similar to [23, Proposition 3]. □

We apply Lemma 3.7 with functions $f_i = \chi_i, 1 \leq i \leq k$, where χ_i is the characteristic function of the interval $[0,V_i]$. Thus, we get

$$\begin{aligned}\mathbb{E}\left(\prod_{i=1}^{k} N^\lambda(V_i)\right) &= \mathbb{E}\left(\prod_{i=1}^{k}\left(\sum_{\ell=1}^{\infty}\chi_i(T_\ell)\right)\right)\\ &= \sum_{P\in\mathscr{P}(k)} \lambda^{\#P}\prod_{B\in P}\left(\int_0^\infty \prod_{i\in B}\chi_i(x)\,\mathrm{d}x\right)\\ &= \sum_{P\in\mathscr{P}(k)} \lambda^{\#P}\prod_{B\in P} V_{i_B},\end{aligned} \tag{3.7}$$

where $i_B = \min_{i\in B} i$.

The following lemma helps us compare the RHS's of (3.5) and (3.7) for $\lambda = 1$.

Lemma 3.8 ([23], Lemma 3). *There is bijection* $g : \mathfrak{M}^k \to \mathscr{P}(k)$ *with the property that if* $\widetilde{D} \in \mathfrak{M}^k$ *is an* $k \times m$ *matrix and* $g(\widetilde{D}) = P = \{B_1,\ldots,B_{\#P}\}$, *then* $\#P = m$ *and* $\{\alpha_1 < \cdots < \alpha_m\} = \{i_{B_1} < \cdots < i_{B_m}\}$.

Proof. Other than switching the rows and columns of the matrices D, the proof of this lemma is the same as [23, Lemma 3]. □

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

Theorem 3.6, (3.7) and Lemma 3.8 imply the following result:

Theorem 3.9.

$$\mathbb{E}\left(\prod_{i=1}^{k} N_i\right) \to \mathbb{E}\left(\prod_{i=1}^{k} N^1(V_i)\right)$$

as $d \to \infty$.

Corollary 3.10. *Let* $\mathbf{V} = (V_1, \ldots, V_k)$ *and consider the random vectors*

$$\mathbf{N}(\Lambda, \mathbf{V}) = (N_1(\Lambda), \ldots, N_k(\Lambda))$$

and

$$\mathbf{N}(\mathbf{V}) = \big(N^1(V_1), \ldots, N^1(V_k)\big).$$

Then $\mathbf{N}(\Lambda, \mathbf{V})$ *converges in distribution to* $\mathbf{N}(\mathbf{V})$ *as* $d \to \infty$.

Proof. This proof follows a similar line of argument as [23, Corollary 1]. We omit it for the sake of brevity. □

Corollary 3.10 implies that all finite dimensional distributions coming from the process $\{N_t(\Lambda) : t \geq 0\}$ converge to the corresponding finite dimensional distributions of the Poisson process $\{N^1(t) : t \geq 0\}$ as $d \to \infty$. By [3, Theorem 12.6 and Theorem 16.7], we see that the process $\{N_t(\Lambda) : t \geq 0\}$ converges weakly to the process $\{N^1(t) : t \geq 0\}$ as $d \to \infty$.

Corollary 3.10, with $k = 1$, is a generalisation of [16, Theorem 3] to the affine case.

3.2. Congruence Case

In this section, we prove Theorem 1.2. We recall the notation. For $d \geq 2$, let $\mathcal{S} = \{S_t : t > 0\}$, an increasing family of subsets of $\mathbb{R}^d$ and $\mathbf{p}/q \in \mathbb{Q}^d$. For $\Lambda \in Y_{\mathbf{p}/q}$, set

$$N_t(\Lambda) = \#(S_t \cap \Lambda).$$

For $\Lambda \in Y_{\mathbf{p}/q}$, let us order the lengths of nonzero vectors in Λ as $0 < \ell_1 \leq \ell_2 \leq \ell_3 \leq \cdots$, and let $\mathscr{V}_i$ denote the volume of the closed ball of radius ℓ_i centered at origin. Taking $\mathcal{S} = \{B_t : t > 0\}$ to be the increasing family of closed balls with $\mathrm{vol}(B_t) = t$ around the origin, we see that

$$N_t(\Lambda) = \#\{i : \mathscr{V}_i \leq t\}.$$

Thus, Theorem 1.2 is equivalent to the following:

Theorem 3.11. *For any fixed* n*, the* n*-dimensional random variable* $(\mathscr{V}_1, \ldots, \mathscr{V}_n)$ *converges in distribution to the distribution of the first* n *points of a Poisson process on the non-negative real line with intensity*

$$\begin{cases} 1 & \text{if } q \geq 3, \\ \frac{1}{2} & \text{if } q = 2. \end{cases}$$

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

As in the affine case, we approach Theorem 1.2 via a joint moment formula for $N_t(\cdot)$. Let $k \geq 1$ and $0 < V_1 \leq V_2 \leq \cdots \leq V_k$. Define N_i's, ρ_i's and F similar to the affine case. We apply Theorem 2.13 to the function F. We first consider the sub-collection of the RHS of (2.12) denoted by $M_{d,k}^{\text{cong}}$, defined as

$$M_{d,k}^{\text{cong}} := \sum_{\widetilde{D} \in \mathfrak{M}^k} \int_{(\mathbb{R}^d)^m} \prod_{i=1}^{k} \rho_i \left(\widetilde{D} \begin{pmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_m \end{pmatrix} \right) \mathrm{d}\mathbf{y}_1 \cdots \mathrm{d}\mathbf{y}_m, \tag{3.8}$$

where

$$\mathfrak{R}_1^k = \bigcup_{1 \leq m \leq k-1} \left(\left(\bigcup_{u \geq 2} \mathfrak{C}_{m,u}^k \right) \cup \left\{ \widetilde{D} = \left(\widetilde{D}_{ij} \right) \in \mathfrak{C}_{m,1}^k : \exists |\widetilde{D}_{ij}| \geq 2 \right\} \right),$$

$$\mathfrak{R}_2^k = \left\{ \widetilde{D} \in \left(\bigcup_{1 \leq m \leq k-1} \mathfrak{C}_{m,1}^k \right) \setminus \mathfrak{R}_1^k : \begin{array}{c} \exists \text{ row such that at least} \\ \text{two entries are nonzero} \end{array} \right\},$$

$$\mathfrak{M}^k = \left(\bigcup_{1 \leq m \leq k-1} \mathfrak{C}_{m,1}^k \setminus \left(\mathfrak{R}_1^k \cup \mathfrak{R}_2^k \right) \right) \cup \{\mathrm{Id}_k\}.$$

The rest of the terms in (2.12) will be denoted as $R_{d,k}^{\text{cong}}$; that is,

$$\mathbb{E}\left(\prod_{i=1}^{k} N_i \right) = M_{d,k}^{\text{cong}} + R_{d,k}^{\text{cong}}.$$

Define $\mathfrak{M}_{\alpha,\beta}^{\text{cong}}$, for (α,β) a division of $\{1,\ldots,k\}$, similar to the affine case and let $M_{\alpha,\beta}^{\text{cong}}$ denote the cardinality of $\mathfrak{M}_{\alpha,\beta}^{\text{cong}}$. We can rewrite (3.8) as

$$M_{d,k}^{\text{cong}} = \sum_{(\alpha,\beta)} \sum_{\widetilde{D} \in \mathfrak{M}_{\alpha,\beta}^{\text{cong}}} \int_{(\mathbb{R}^d)^m} \prod_{i=1}^{k} \rho_i \left(\widetilde{D} \begin{pmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_m \end{pmatrix} \right) \mathrm{d}\mathbf{y}_1 \cdots \mathrm{d}\mathbf{y}_m, \tag{3.9}$$

where the outer sum is over all possible divisions (α,β) of $\{1,\ldots,k\}$.

Remark 3.12. For $q \geq 3$, it follows from the definition of $\widetilde{D}$ and similar arguments as in Remark 3.2 that for $\widetilde{D} \in \mathfrak{M}^k$, the nonzero entries of $\widetilde{D}$ can only be 1. But for $q = 2$, the nonzero entries can be ± 1. In particular, this is the reason why we need the condition that S_d is symmetric for the case when $q = 2$ (see the second last equality in (3.11) below).

Lemma 3.13. *For $q \geq 3$ and for $q = 2$ with S_t being symmetric around the origin, we have*

$$M_{d,k}^{cong} = \sum_{(\alpha,\beta)} M_{\alpha,\beta}^{cong} \prod_{i=1}^{m} V_{\alpha_i}. \tag{3.10}$$

Proof. For $q \geq 3$, the proof of this lemma is identical to that of Lemma 3.3. Hence, we only focus on the case when $q = 2$.

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

Consider any matrix $\widetilde{D} = \left(\widetilde{D}_{ij}\right) \in \mathfrak{M}^{\mathrm{cong}}_{\alpha,\beta}$ and let λ_ℓ be such that $\widetilde{D}_{\beta_\ell,\lambda_\ell} = 1$, $1 \le \ell \le k-m$. Then, as S_t's are symmetric and V_i's are increasing, the following calculation finishes the proof:

$$\int_{(\mathbb{R}^d)^m} \prod_{i=1}^{k} \rho_i \left(\widetilde{D} \begin{pmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_m \end{pmatrix} \right) \mathrm{d}\mathbf{y}_1 \cdots \mathrm{d}\mathbf{y}_m = \int_{(\mathbb{R}^d)^m} \prod_{i=1}^{k} \rho_i \left(\sum_{j=1}^{m} \widetilde{D_{ij}} \mathbf{y}_j \right) \mathrm{d}\mathbf{y}_1 \cdots \mathrm{d}\mathbf{y}_m \tag{3.11}$$

$$= \int_{(\mathbb{R}^d)^m} \prod_{i=1}^{m} \rho_{\alpha_i}(\mathbf{y}_i) \prod_{\ell=1}^{k-m} \rho_{\beta_\ell}(\pm \mathbf{y}_{\lambda_\ell})\, \mathrm{d}\mathbf{y}_1 \cdots \mathrm{d}\mathbf{y}_m = \int_{(\mathbb{R}^d)^m} \prod_{i=1}^{m} \rho_{\alpha_i}(\mathbf{y}_i) = \prod_{i=1}^{m} V_{\alpha_i}. \qquad \square$$

From Lemma 3.13 and Lemma 3.4, we find the following:

Theorem 3.14.

$$\mathbb{E}\left(\prod_{i=1}^{k} N_i \right) \to \sum_{(\alpha,\beta)} M^{cong}_{\alpha,\beta} \prod_{i=1}^{m} V_{\alpha_i}. \tag{3.12}$$

3.2.1. Proof of Theorem 1.2. For $q \ge 3$, the proof of Theorem 1.2 follows the proof of Theorem 1.1. We need a small modification in Lemma 3.8 for the case $q = 2$ because in this case, the entries of matrices in $\mathfrak{M}^k$ can be negative. From now on, we only focus on the case $q = 2$.

Let $\mathfrak{M}^{\mathrm{cong}}_{\alpha,\beta,+}$ denote the subset of $\mathfrak{M}^{\mathrm{cong}}_{\alpha,\beta}$ of matrices with positive entries, and similarly, let $\mathfrak{M}^k_+$ denote the subset of $\mathfrak{M}^k$ of matrices with positive entries. With $M^{\mathrm{cong}}_{\alpha,\beta,+} := \#(\mathfrak{M}^{\mathrm{cong}}_{\alpha,\beta,+})$, note that

$$M^{\mathrm{cong}}_{\alpha,\beta} = \#\left(\mathfrak{M}^{\mathrm{cong}}_{\alpha,\beta}\right) = 2^{k-\#\alpha} M^{\mathrm{cong}}_{\alpha,\beta,+}.$$

Thus, from (3.12), we find

$$\mathbb{E}\left(\prod_{i=1}^{k} \widetilde{N}_i \right) \to \sum_{(\alpha,\beta)} 2^{-\#\alpha} M^{\mathrm{cong}}_{\alpha,\beta,+} \prod_{i=1}^{m} V_{\alpha_i}, \tag{3.13}$$

where $\widetilde{N}_i = \frac{1}{2} N_i$ for $1 \le i \le k$, i.e., $\widetilde{N}_t = \frac{1}{2} N_t$.

With the following lemma, we can compare the RHS's of Theorem 3.13 and (3.7) for $\lambda = 1/2$.

Lemma 3.15. *There is bijection $g : \mathfrak{M}^k_+ \to \mathscr{P}(k)$ with the property that if $\widetilde{D} \in \mathfrak{M}^k_+$ is an $k \times m$ matrix and $g(\widetilde{D}) = P = \{B_1, \ldots, B_{\#P}\}$, then $\#P = m$ and $\{\alpha_1 < \cdots < \alpha_m\} = \{i_{B_1} < \cdots < i_{B_m}\}$.*

Proof. The same argument with Lemma 3.8 holds. $\square$

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

(3.7), (3.13) and Lemma 3.15 combine to show the following:

Theorem 3.16. *For $q=2$,*

$$\mathbb{E}\left(\prod_{i=1}^{k}\widetilde{N}_i\right)\to\mathbb{E}\left(\prod_{i=1}^{k}N^{1/2}(V_i)\right).$$

Corollary 3.17. *Let $q=2$, $\mathbf{V}=(V_1,\dots,V_k)$ and consider the random vectors*

$$\widetilde{\mathbf{N}}(\Lambda,\mathbf{V})=\left(\widetilde{N}_1(\Lambda),\dots,\widetilde{N}_k(\Lambda)\right)$$

and

$$\mathbf{N}(\mathbf{V})=\left(N^{1/2}(V_1),\dots,N^{1/2}(V_k)\right).$$

Then $\widetilde{\mathbf{N}}(\Lambda,\mathbf{V})$ converges in distribution to $\mathbf{N}(\mathbf{V})$ as $d\to\infty$.

Proof. This proof follows a similar line of argument as [23, Corollary 1]. We omit it for the sake of brevity. □

Corollary 3.17 implies that all finite dimensional distributions coming from the process $\{\widetilde{N}_t(\Lambda):t\geq 0\}$ converge to the corresponding finite dimensional distributions of the Poisson process $\{N^{1/2}(t):t\geq 0\}$ as $d\to\infty$. By [3, Theorem 12.6 and Theorem 16.7], we see that the process $\{N_t(\Lambda):t\geq 0\}$ converges weakly to the process $\{N^{1/2}(t):t\geq 0\}$ as $d\to\infty$.

Corollary 3.17, with $k=1$, is a generalisation of [16, Theorem 3] to the congruence case.

4. New Moment Formulae

In this section, we want to simplify Theorem 2.12 and Theorem 2.13 for the special case as considered by Strömbergsson and Södergren in [22]. Theorems 4.1 and 4.2 below will be used in Section 5.

For bounded and compactly supported functions $f_i:\mathbb{R}^d\to\mathbb{R}_{\geq 0}$ $(1\leq i\leq k)$, define

$$F_i(\mathbf{v})=f_i(\mathbf{v})-\int_{\mathbb{R}^d}f_i\,\mathrm{d}\mathbf{v}. \tag{4.1}$$

We want to compute the integrals of $\mathcal{S}_k(\prod_{i=1}^{k}F_i)=\prod_{i=1}^{k}\widehat{F}_i$ over Y and $Y_{\mathbf{p}/q}$.

We first observe that by applying Theorem 2.12,

$$\begin{aligned}\int_Y\prod_{i=1}^{k}\widehat{F}_i(\Lambda)\,\mathrm{d}\mu_Y(\Lambda)&=\int_Y\prod_{i=1}^{k}\left(\widehat{f}_i(\Lambda)-\int_{\mathbb{R}^d}f_i\,\mathrm{d}\mathbf{v}\right)\\&=\sum_{A\subseteq\{1,\dots,k\}}(-1)^a\left(\prod_{i''\in A}\int_{\mathbb{R}^d}f_{i''}\,\mathrm{d}\mathbf{v}\right)\int_Y\prod_{i\in A^c}\widehat{f}_i(\Lambda)\,\mathrm{d}\mu_Y(\Lambda)\end{aligned}$$

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

$$
\begin{aligned}
&= \sum_{A\subseteq\{1,\ldots,k\}} (-1)^a \prod_{i''\in A} \int_{\mathbb{R}^d} f_{i''}\,\mathrm{d}\mathbf{v} \times \\
&\left(\sum_{m=1}^{k-a} \sum_{u\in\mathbb{N}} \sum_{\widetilde{D}\in\mathfrak{A}^{k-a}_{m,u}} \frac{N(\widetilde{D},u)^d}{u^{dm}} \int_{(\mathbb{R}^d)^m} \left(\prod_{i\in A^c} f_i\right) \left(\frac{\widetilde{D}}{u} \begin{pmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_m \end{pmatrix} \right) \mathrm{d}\mathbf{w}_1\cdots\mathrm{d}\mathbf{w}_m \right),
\end{aligned} \tag{4.2}
$$

where $a=\#A$ and $A^c=\{1,\ldots,k\}-A$.

Note that for a given $A\subseteq\{1,\ldots,k\}$ and $\widetilde{D}\in\mathfrak{A}^{k-a}_{m,u}$, one can find a unique matrix $D''=D''(A,\widetilde{D})\in\mathfrak{D}^k_{m+a,u}$ (in fact, $\mathfrak{A}^k_{m+a,u}$) for which

$$
\begin{aligned}
&\left(\prod_{i''\in A}\int_{\mathbb{R}^d} f_{i''}\,\mathrm{d}\mathbf{v}\right)\cdot \int_{(\mathbb{R}^d)^m} \left(\prod_{i\in A^c} f_i\right) \left(\frac{\widetilde{D}}{u} \begin{pmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_m \end{pmatrix} \right) \mathrm{d}\mathbf{w}_1\cdots\mathrm{d}\mathbf{w}_m \\
&\qquad = \int_{(\mathbb{R}^d)^{m+a}} \left(\prod_{i=1}^{k} f_i\right) \left(\frac{D''}{u} \begin{pmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_{m+a} \end{pmatrix} \right) \mathrm{d}\mathbf{w}_1\cdots\mathrm{d}\mathbf{w}_{m+a}.
\end{aligned} \tag{4.3}
$$

Moreover, from the definitions of $N(D'',u)$ and $N(\widetilde{D},u)$ in Notation 2.1 (3), one can directly obtain the following equality:

$$
\frac{N(D'',u)^d}{u^{dn}} = \frac{N(\widetilde{D},u)^d}{u^{dm}}. \tag{4.4}
$$

We claim the following:

Theorem 4.1. *For $1\le i\le k$, let F_i be the function defined as in (4.1) for a bounded and compactly supported function $f_i:\mathbb{R}^d\to\mathbb{R}_{\ge 0}$ $(1\le i\le k)$. It follows that*

$$
\int_Y \prod_{i=1}^{k} \widehat{F}_i(\Lambda)\,\mathrm{d}\mu_Y(\Lambda) = \sum_{n=1}^{k-1}\sum_{u\in\mathbb{N}}\sum_{D''\in\mathfrak{S}^k_{n,u}} \frac{N(D'',u)^d}{u^{dn}} \int_{(\mathbb{R}^d)^n} \prod_{i=1}^{k} f_i \left(\frac{D''}{u} \begin{pmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_n \end{pmatrix} \right) \mathrm{d}\mathbf{w}_1\cdots\mathrm{d}\mathbf{w}_n, \tag{4.5}
$$

where $\mathfrak{S}^k_{n,u}\subseteq\mathfrak{A}^k_{n,u}$ is the set of D'' which is one of the following:

(a) *Each column of $[D'']$ has at least two nonzero elements.*

(b) *There are $0\le a\le n-2$ and $D\in\mathfrak{D}^{k-a-1}_{n-a-1,u}-\mathfrak{A}^{k-a-1}_{n-a-1,u}$ for which*

$$
D'' = \left(\begin{array}{c|cc}
u\mathrm{Id}_a & & \\ \hline
 & u & 0\cdots 0 \\ \cline{2-3}
 & 0 & \\
 & \vdots & D \\
 & 0 &
\end{array}\right),
$$

where each column of D has at least two nonzero elements.

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

Similarly, from Theorem 2.13, we have that $\int_{Y_{\mathbf{p}/q}}\prod_{i=1}^{k}\widehat{F}_i(\Lambda)d\mu_q$ is the sum of integrals given as in (4.2) with replacing $\mathfrak{A}_{m,u}^{k-a}$ by $\mathfrak{C}_{m,u}^{k-a}$. For $D''=D''(A,\widetilde{D})\in\mathfrak{D}_{m+a,u}^{k}$ defined using $A\subseteq\{1,\dots,k\}$ and $\widetilde{D}\in\mathfrak{C}_{m,u}^{k-a}$ as in (4.3), we will see that $D''\in\mathfrak{C}_{m+a,u}^{k}$. It is easily seen that the equality (4.4) holds in the congruence case.

Theorem 4.2. *For $1\le i\le k$, let F_i be the function defined as in (4.1) for a bounded and compactly supported function $f_i:\mathbb{R}^d\to\mathbb{R}_{\ge 0}$ $(1\le i\le k)$. It follows that*

$$\int_{Y_{\mathbf{p}/q}}\prod_{i=1}^{k}\widehat{F}_i(\Lambda)\,\mathrm{d}\mu_q(\Lambda)$$
$$=\sum_{n=1}^{k-1}\sum_{u\in\mathbb{N}}\sum_{D''\in\mathfrak{T}_{n,u}^{k}}\frac{N(D'',u)^d}{u^{dn}}\int_{(\mathbb{R}^d)^n}\prod_{i=1}^{k}f_i\left(\frac{D''}{u}\begin{pmatrix}\mathbf{w}_1\\ \vdots\\ \mathbf{w}_n\end{pmatrix}\right)\mathrm{d}\mathbf{w}_1\cdots\mathrm{d}\mathbf{w}_n,$$

where $\mathfrak{T}_{n,u}^{k}$ is a subset of $\mathfrak{C}_{n,u}^{k}$ collecting D'' which is one of the following:

(a) *Each column of D'' has at least two nonzero elements.*

(b) *There are $0\le a\le n-2$ and $\widetilde{D}\in\mathfrak{C}_{n-a,u}^{k-a}$ so that*

$$D''=\begin{pmatrix}u\mathrm{Id}_a & \\ & \widetilde{D}\end{pmatrix},$$

where $[\widetilde{D}]^1={}^{\mathrm{t}}(u,0,\dots,0)$ and any other columns of $\widetilde{D}$ have at least two nonzero elements. Moreover, the right-bottom minor of $[\widetilde{D}]$ with size $(k-a-1)\times(n-a-1)$ is not an element of $\mathfrak{C}_{n-a-1,u}^{k-a-1}$ (or any $\mathfrak{C}_{n-a-1,*}^{k-a-1}$).

Proof of Theorem 4.1 and Theorem 4.2. As described in (4.3), a possible matrix D'' among elements of $\mathfrak{D}_{n,u}^{k}$ is constructed by using $A\subseteq\{1,\dots,k\}$ and $\widetilde{D}\in\mathfrak{A}_{n-a,u}^{k-a}$. Conversely, we want to consider all possible pairs $(A,\widetilde{D})$ which give the same D''.

Let such a $D''=(d''_{ij})$ be given. Denote

$$B=\left\{1\le i''\le k:\begin{array}{l}1\le\exists j_0\le n\text{ for which}\\ \quad d''_{i''j}=0\text{ for all }j\text{ except }d''_{i''j_0}=u\text{ and}\\ \quad d''_{ij_0}=0\text{ for all }i\text{ except }d''_{i''j_0}=u\end{array}\right\}.$$

After changing the (last $(k-b_1)$) coordinates of $\mathbb{R}^k$, we may assume that

$$\frac{D''}{u}=\begin{pmatrix}\mathrm{Id}_{b_1} & & \\ & \dfrac{\widetilde{D_0}}{u} & \\ & & \mathrm{Id}_{b_2}\end{pmatrix},\tag{4.6}$$

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

where b_1 and b_2 could be 0 (then D''/u will be one- or two-block diagonal matrix) and $\widetilde{D}_0 \in \mathfrak{A}^{k-b_1-b_2}_{n-b_1-b_2,u}$ (or $\mathfrak{C}^{k-b_1-b_2}_{n-b_1-b_2,u}$, respectively) is the minimal size among possible $(A,\widetilde{D})$ for which $D''(A,\widetilde{D}) = D''$; that is,

$$\text{each column of } \widetilde{D}_0 \text{ except } [\widetilde{D_0}]^1 \text{ has at least two nonzero elements.}$$

Notice that any matrix constructed by choosing more than $k-b_1-b_2$ rows and more than $n-b_1-b_2$ columns from D''/u and having $\widetilde{D_0}/u$ as its minor is element of $\mathfrak{A}^*_{*,u}$ (or $\mathfrak{C}^*_{*,u}$, respectively). For example, $D'' \in \mathfrak{C}^k_{n,u}$ since D'' is constructed by $(\overline{D}, t, {}^{\mathrm{t}}(0,\ldots,0,\boldsymbol{\ell},0,\ldots,0))$ under the map in (2.10), where $\overline{D}$ is the right-bottom minor of uD'' with size $(k-1)\times(n-1)$, and $(t,\boldsymbol{\ell})$ is a pair used for defining $\widetilde{D}$.

Now let us check case by case. Denote by

$$B_1 = \{i \in B : i \le b_1 + 1\} \quad \text{and} \quad B_2 = \{k - b_2 + 1, \ldots, k\}$$

so that $B = B_1 \cup B_2$. Note that (b_1+1) could be not contained in B. Observe that possible A for constructing D'' is of the form $A_1 \cup A_2$, where $A_1 \subseteq B_1$ and $A_2 \subseteq B_2$. The difference between A_1 and A_2 is that A_1 may have an extra condition according to the given D'', but any subset of B_2 can be A_2.

We first assume that $B_2 \neq \emptyset$. Since

$$\begin{aligned}
\sum_{\substack{\text{“possible”}\\ A\subseteq B}} (-1)^{\#A} &= \sum_{\substack{\text{“possible”}\\ A_1\subseteq B_1}} (-1)^{\#A_1} \sum_{\forall A_2 \subseteq B_2} (-1)^{\#A_2} \\
&= \sum_{\substack{\text{“possible”}\\ A_1\subseteq B_1}} (-1)^{\#A_1} \cdot 0 = 0,
\end{aligned}$$

with the observation in (4.3), the partial sum

$$\sum_{\substack{A,\widetilde{D}\text{ s.t.}\\ D''(A,\widetilde{D}) = D''}} (-1)^{\#A} \frac{N(D'',u)^d}{u^{dr}} \int_{(\mathbb{R}^d)^n} \prod_{i=1}^{k} f_i \left(\frac{D''}{u} \begin{pmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_{m+a} \end{pmatrix} \right) \mathrm{d}\mathbf{w}_1 \cdots \mathrm{d}\mathbf{w}_{m+a} \tag{4.7}$$

associated with D'' in the right-hand side of (4.5) is zero.

Now let us assume that $b_2 = 0$. If $B = \emptyset$, that is,

$$\text{each column of } D'' \text{ has at least two nonzero vectors,}$$

and only possible $(A,\widetilde{D})$ is $(\emptyset, D'')$. This is the case (a) in the theorem.

Suppose that $|B| = b_1 \ge 1$. Equivalently, suppose that

$$[\widetilde{D_0}] \text{ as well as other columns of } \widetilde{D}_0 \text{ has at least two nonzero elements.}$$

Then any subset A of B is possible for defining D''; hence, the partial sum (4.7) is zero.

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

The only left case is when $|B| = b_1 + 1 \geq 2$. Notice that

$$[\widetilde{D_0}]^1 = {}^{\mathrm{t}}(u, 0, \ldots, 0), \text{and}$$

$$\text{the right-bottom minor of } \widetilde{D_0} \text{ with size } (k - b_1 - 1) \times (n - b_1 - 1)$$

$$\text{is not an element of } \mathfrak{A}^{k-b_1-1}_{n-b_1-1,u} (\mathfrak{C}^{k-b_1-1}_{n-b_1-1,u}, \text{ respectively}).$$

In this case, any subset of B except B itself can be possible A for defining D'', and this is the case (b) in the theorem. □

5. CLT and Brownian motion

As in Section 3, we will use the method of moments which is applicable with the normal distribution and Brownian motion, following [22]. Recall that the k-th moment of the normal distribution is 0 when k is odd and $(k-1)!!$ when k is even.

For Brownian motion, it suffices to show that the induced measure P^1_d and $P^{\mathbf{p}/q}_d$ from $Z^1_d(t)$ and $Z^{\mathbf{p}/q}_d(t)$, respectively, on the space $C[0,1]$ of continuous real-valued functions on $[0,1]$ weakly converge to Wiener measure as d goes to infinity.

Let $\phi : \mathbb{N} \to \mathbb{R}_{>0}$ be a function for which $\lim_{d\to\infty} \phi(d) = \infty$ and $\phi(d) = O_\varepsilon(e^{\varepsilon d})$ for every $\varepsilon > 0$. Let $\iota \in \mathbb{N}$ and $c_1, \ldots, c_\iota > 0$ be arbitrarily given. For each $d \in \mathbb{N}$, consider $S_{i,d} \in \mathbb{R}^d$ to be a Borel measurable set satisfying $\mathrm{vol}(S_{i,d}) = c_i \phi(d)$ for $1 \leq i \leq \iota$ and $S_{i,d} \cap S_{i',d} = \emptyset$ if $i \neq i'$. If we consider the case that $\Lambda \in Y_{\mathbf{p}/q}$ with $q = 2$, we further assume that $S_{i,d} = -S_{i,d}$ for $1 \leq i \leq \iota$ and $d \in \mathbb{N}$.

Let

$$Z^1_{i,d} := \frac{\#(\Lambda \cap S_{i,d}) - c_i\phi(d)}{\sqrt{\phi(d)}}, \qquad \Lambda \in Y \text{ and}$$

$$Z^{\mathbf{p}/q}_{i,d} := \begin{cases} \dfrac{\#(\Lambda \cap S_{i,d}) - c_i\phi(d)}{\sqrt{\phi(d)}}, & \text{if } q \neq 2; \\[2ex] \dfrac{\#(\Lambda \cap S_{i,d}) - c_i\phi(d)}{\sqrt{2\phi(d)}}, & \text{if } q = 2, \end{cases} \qquad \Lambda \in Y^{\mathbf{p}/q}.$$

Proposition 5.1. *Let $\diamondsuit = 1$ or $\mathbf{p}/q$. For any fixed $\mathbf{k} = (k_1, \ldots, k_\iota) \in \mathbb{N}^\iota$, it follows that*

$$\lim_{d\to\infty} \mathbb{E}\left((Z^\diamondsuit_{1,d})^{k_1} \cdots (Z^\diamondsuit_{\iota,d})^{k_\iota} \right) = \begin{cases} \prod_{i=1}^{\iota} c_i^{k_i/2}(k_i - 1)!!, & \textit{if } k_1, \ldots, k_\iota \textit{ are all even}, \\ 0, & \textit{otherwise}. \end{cases}$$

Proof. Let $k = k_1 + \cdots + k_\iota$ and consider $d > \lfloor k^2/4 \rfloor + 3$. For each $d \in \mathbb{N}$ and $1 \leq i \leq \iota$, let $f_{i,d}$ be the indicator function of $S_{i,d}$ and define

$$F_{i,d}(\Lambda) = \widehat{f_{i,d}}(\Lambda) - \int_{\mathbb{R}^d} f_{i,d}\, \mathrm{d}\mathbf{v} = \widehat{f_{i,d}}(\Lambda) - c_i\phi(d),\ \Lambda \in Y.$$

We will use Theorem 4.1 and Theorem 4.2.

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

Recall that we can divide $\bigcup_{1\le n\le k}\bigcup_{u\in\mathbb{N}}\mathfrak{D}^k_{n,u}$ as the union of $\mathfrak{R}^k_1$, $\mathfrak{R}^k_2$ and $\mathfrak{M}^k$, where

$$\mathfrak{R}^k_1=\bigcup_{1\le n\le k-1}\left(\left(\bigcup_{u\ge 2}\mathfrak{D}^k_{n,u}\right)\cup\left\{D=(d_{ij})\in\mathfrak{D}^k_{n,1}:\exists|d_{ij}|\ge 2\right\}\right);$$

$$\mathfrak{R}^k_2=\left\{D\in\Big(\bigcup_{1\le n\le k-1}\mathfrak{D}^k_{n,1}\Big)-\mathfrak{R}^k_1:\begin{array}{l}\exists\text{ column such that}\\ \text{at least two entries are nonzero}\end{array}\right\};$$

$$\mathfrak{M}^k=\Big(\bigcup_{1\le n\le k-1}\mathfrak{D}^k_{n,1}\Big)-\big(\mathfrak{R}^k_1\cup\mathfrak{R}^k_2\big).$$

(i) The space Y
By Theorem 4.1 and Theorem 3.4, one can deduce that

$$\begin{aligned}\mathbb{E}\left(\prod_{i=1}^{\iota}(Z^1_{i,d})^{k_i}\right)&=\frac{1}{\phi(d)^{k/2}}\int_Y\prod_{i=1}^{\iota}\widehat{F}^{k_i}_{i,d}(\Lambda)\,\mathrm{d}\mu_Y(\Lambda)\\
&=\frac{1}{\phi(d)^{k/2}}\sum_{n=1}^{k-1}\sum_{\substack{D''\in\\ \mathfrak{S}^k_{n,1}\cap\mathfrak{M}^k}}\int_{(\mathbb{R}^d)^n}\left(\prod_{i=1}^{\iota}f^{k_i}_{i,d}\right)\left(D''\begin{pmatrix}\mathbf{w}_1\\ \vdots\\ \mathbf{w}_n\end{pmatrix}\right)\mathrm{d}\mathbf{w}_1\cdots\mathrm{d}\mathbf{w}_n\\
&\quad+O\left(\left(\frac{\sqrt{3}}{2}\right)^d\phi(d)^{k/2-1}\right).\end{aligned}\tag{5.1}$$

Notice that if $D''\in\mathfrak{S}^k_{n,1}\cap\mathfrak{M}^k$, then D'' is of type (a) in Theorem 4.1. Hence, for each column of D'', there are at least two nonzero entries, and for each row of D'', there is exactly one nonzero entry. Moreover, as mentioned in Remark 3.2, entries of D'' are $\{0,1\}$.

We first claim that D'' for which the inner integral above is nontrivial is the block diagonal matrix of the form

$$\begin{pmatrix}D''_{k_1,n_1}&&&\\&D''_{k_2,n_2}&&\\&&\ddots&\\&&&D''_{k_\iota,n_\iota}\end{pmatrix},$$

where $n_1+\cdots+n_\iota=n$ and $n_i\ge 1$ for each $1\le i\le\iota$. Moreover, each $D''_{k_i,n_i}\in\mathfrak{S}^{k_i}_{n_i,1}\cap\mathfrak{M}^{k_i}$. Indeed, since the set $\{S_{i,n}\}_{1\le i\le\iota}$ is mutually disjoint, for each column, it is only possible that nontrivial entries are located between the $\left((\sum_{\ell=1}^{i-1}k_i)+1\right)$-th row and the $\left(\sum_{\ell=1}^{i}k_i\right)$-th row for some $1\le i\le\iota$. In other words, nontrivial entries are concentrated in rows which correspond to f_i. The fact that D'' is a block diagonal matrix comes from that $D''\in\mathfrak{D}^k_{n,u}$, especially, from the first property of Notation 2.1 (2).

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

It is not hard to show that each D''_{k_i,n_i} is in $\mathfrak{S}^{k_i}_{n_i,1} \cap \mathfrak{M}^{k_i}$ from the fact that $D'' \in \mathfrak{S}^{k}_{n,1} \cap \mathfrak{M}^{k}$. Hence, the main term in (5.1) is

$$\prod_{i=1}^{\iota} \frac{1}{\phi(d)^{k_i/2}} \sum_{n_i=1}^{\lfloor k_i/2 \rfloor} \sum_{\substack{D''_{k_i,n_i} \in \\ \mathfrak{S}^{k_i}_{n_i,1} \cap \mathfrak{M}^{k_i}}} \int_{(\mathbb{R}^d)^{n_i}} f_i^{k_i} \left(D''_{k_i,n_i} \begin{pmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_{n_i} \end{pmatrix} \right) \mathrm{d}\mathbf{w}_1 \cdots \mathrm{d}\mathbf{w}_{n_i}. \tag{5.2}$$

The next claim is that for each i, there is a one-to-one correspondence between $\mathfrak{S}^{k_i}_{n_i,1} \cap \mathfrak{M}^{k_i}$ and the set of partitions $\mathcal{P} = \{P_1, \ldots, P_{n_i}\}$ of $\{1, \ldots, k_i\}$ such that

$$|\mathcal{P}| = n_i \text{ and } |P_\ell| \geq 2 \text{ for } 1 \leq \ell \leq n_i.$$

Let $\mathcal{P}$ be such a partition. Reordering if necessary, we may assume that $\min P_1 < \ldots < \min P_{n_i}$. The corresponding element in $\mathfrak{S}^{k_i}_{n_i,1} \cap \mathfrak{M}^{k_i}$ is

$$\left[D''_{k_i,n_i}\right]_{\ell j} = \begin{cases} 1, & \text{if } \ell \in P_j; \\ 0, & \text{otherwise.} \end{cases} \tag{5.3}$$

It is obvious that from the first property of Notation 2.1 (2) and the definition of $\mathfrak{M}^{k_i}$, any element in $\mathfrak{S}^{k_i}_{n_i,1} \cap \mathfrak{M}^{k_i}$ is a matrix of the form (5.3) for some partition $\{P_1, \ldots, P_{n_i}\}$ of $\{1, \ldots, k_i\}$.

Let $N(k_i, n_i)$ be the number of such partitions. If $n_i < k_i/2$, since $\lim_{d \to \infty} \phi(d) = \infty$,

$$\begin{aligned} &\frac{1}{\phi(d)^{k_i/2}} \sum_{\substack{D''_{k_i,n_i} \in \\ \mathcal{S}^{k_i}_{n_i,1} \cap \mathfrak{M}^{k_i}}} \int_{(\mathbb{R}^d)^{n_i}} F_i^{k_i} \left(D''_{k_i,n_i} \begin{pmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_{n_i} \end{pmatrix} \right) \mathrm{d}\mathbf{w}_1 \cdots \mathrm{d}\mathbf{w}_{n_i} \\ &\qquad \leq \frac{c_i^{n_i} N(k_i, n_i)}{\phi(d)^{k_i/2 - n_i}} \longrightarrow 0 \text{ as } d \to \infty. \end{aligned} \tag{5.4}$$

If $n_i = k_i/2$, by the induction on $k_i/2$, one can show that

$$\begin{aligned} &\frac{1}{\phi(d)^{k_i/2}} \sum_{\substack{D''_{k_i,n_i} \in \\ \mathcal{S}^{k_i}_{n_i,1} \cap \mathfrak{M}^{k_i}}} \int_{(\mathbb{R}^d)^{n_i}} F_i^{k_i} \left(D''_{k_i,n_i} \begin{pmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_{n_i} \end{pmatrix} \right) \mathrm{d}\mathbf{w}_1 \cdots \mathrm{d}\mathbf{w}_{n_i} \\ &\qquad = c_i^{k_i/2} N(k_i, k_i/2) = c_i^{k_i/2} (k_i - 1)!!. \end{aligned} \tag{5.5}$$

The result follows from (5.2), (5.4) and (5.5).

(ii) The space $Y_{\mathbf{p}/q}$

The proof is similar to that of (i), where we use Theorem 4.2, Lemma 3.4. One can check that $D'' \in \mathfrak{T}^{k}_{n,1} \cap \mathfrak{M}^{k}$ is of type (a) in Theorem 4.2.

One difference from the affine case is when $q = 2$, $D'' \in \mathfrak{T}^{k}_{n,1} \cap \mathfrak{M}^{k}$, which permits to have -1 as its entries. More precisely, the rows corresponding to $I^c_{D''}$ can have ± 1 as their nonzero entries.

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

It follows that

$$\lim_{d\to\infty}\mathbb{E}\left(\prod_{i=1}^{\iota} Z_{i,d}^{k_i}\right)=\lim_{d\to\infty}\prod_{i=1}^{\iota}\frac{1}{(2\phi(d))^{k_i/2}}\times$$
$$\sum_{n_i=1}^{\lfloor k_i/2\rfloor}\sum_{\substack{D''_{k_i,n_i}\in\\ \mathfrak{T}^{k_i}_{n_i,1}\cap\mathfrak{M}^{k_i}}}\int_{(\mathbb{R}^d)^{n_i}} f_i^{k_i}\left(D''_{k_i,n_i}\begin{pmatrix}\mathbf{w}_1\\ \vdots\\ \mathbf{w}_{n_i}\end{pmatrix}\right)d\mathbf{w}_1\cdots d\mathbf{w}_{n_i}.$$

As in the affine case, the limit is nontrivial only if all k_i's are even and is determined by summation over $\mathfrak{T}^{k_i}_{k_i/2,1}\cap\mathfrak{M}^{k_i}$. Hence, if $q=2$, since $\#I^c_{D''_{k_i,k_i/2}}=k_i/2$, the number $\#\left(\mathfrak{T}^{k_i}_{k_i/2,1}\cap\mathfrak{M}^{k_i}\right)$ is $2^{k_i/2}N(k_i,k_i/2)$. Therefore,

$$\prod_{i=1}^{\iota}\frac{1}{(2\phi(d))^{k_i/2}}\sum_{\substack{D''_{k_i,n_i}\in\\ \mathfrak{T}^{k_i}_{n_i,1}\cap\mathfrak{M}^{k_i}}}\int_{(\mathbb{R}^d)^{n_i}} F_i^{k_i}\left(D''_{k_i,n_i}\begin{pmatrix}\mathbf{w}_1\\ \vdots\\ \mathbf{w}_{n_i}\end{pmatrix}\right)d\mathbf{w}_1\cdots d\mathbf{w}_{n_i}$$
$$=\prod_{i=1}^{\iota}\frac{1}{(2\phi(d))^{k_i/2}}(c_i\phi(d))^{k_i/2}\cdot 2^{k_i/2}N(k_i,k_i/2)=\prod_{i=1}^{\iota}c_i^{k_i/2}(k_i-1)!!.\qquad\square$$

Proofs of Theorem 1.3 and 1.4. As a corollary of Proposition 5.1 with $\iota=1$, for $\diamondsuit=1$ and $\mathbf{p}/q$, it follows that for any $k\in\mathbb{N}$,

$$\lim_{d\to\infty}\mathbb{E}\left((Z_d^{\diamondsuit})^k\right)=\begin{cases}(k-1)!!, & \text{if } k\in 2\mathbb{N};\\ 0, & \text{otherwise,}\end{cases}$$

which shows that $Z_d^{\diamondsuit}\to\mathcal{N}(0,1)$ as $d\to\infty$ in distribution by the method of moments. $\square$

Proofs of Theorem 1.5 and 1.6. For any $0<t_1<\ldots<t_\iota<1$, set

$$S_{i,d}=(t_i)^{1/d}S_d-(t_{i-1})^{1/d}S_d,\ 2\le i\le\iota$$

and $S_{1,d}=(t_1)^{1/d}S_d$. Since S_d is star-shaped, all $S_{i,d}$'s are mutually disjoint. By Proposition 5.1, for $\diamondsuit=1$ and $\mathbf{p}/q$, the random vector

$$\left(Z_d^{\diamondsuit}(t_1),Z_d^{\diamondsuit}(t_2)-Z_d^{\diamondsuit}(t_1),\ldots,Z_d^{\diamondsuit}(t_\iota)-Z_d^{\diamondsuit}(t_{\iota-1})\right)$$

converges weakly as finite-dimensional distributions to

$$(\mathcal{N}(0,t_1),\mathcal{N}(0,t_2-t_1),\ldots,\mathcal{N}(0,t_\iota-t_{\iota-1}))$$

by the method of moments.

The rest of the proof is to show the tightness. As in the proof of Theorem 1.6 in [22], by [3, Theorem 13.3 and (13.14)], it suffices to show that for any $0\le r\le s\le t\le 1$,

$$\mathbb{E}\left((Z_d^{\diamondsuit}(s)-Z_d^{\diamondsuit}(r))^2(Z_d^{\diamondsuit}(t)-Z_d^{\diamondsuit}(s))^2\right)\ll(\sqrt{t}-\sqrt{r})^2.$$

https://doi.org/10.1017/S147474802300035X Published online by Cambridge University Press

We omit the proof since it is almost the same as in the proof of [22, Theorem 1.6] (see especially equations from (4.4) to (4.9)), where the arguments are applicable to a star-shaped set $S_d \subseteq \mathbb{R}^d$ centered at the origin without any modification. Here, we want to remark that we need the argument in [22] only for the congruence case. For the affine case, since $\bigcup_{u\in\mathbb{N}} \mathfrak{S}^4_{1,u} \cap (\mathfrak{R}_1 \cup \mathfrak{R}_2) = \emptyset$, it is deduced directly from (4.5) in [22] that

$$\begin{aligned}
&\mathbb{E}\left((Z_d(s)-Z_d(r))^2(Z_d(t)-Z_d(s))^2\right)\\
&\ll (t-r)^2 + \max\left(\left(\frac{3}{4}\right)^{n/2}(t-r)^2, \left(\frac{3}{4}\right)^{n/2}(t-r)^3\phi(d)\right)\\
&\ll (t-r)^2 < (\sqrt{t}-\sqrt{r})^2. \qquad \square
\end{aligned}$$

Acknowledgements. We are very grateful to the anonymous referee for an extremely detailed report which pointed out several mistakes in an earlier version of the paper and also generously offered solutions to some of the issues. A. G. gratefully acknowledges support from a MATRICS grant from the Science and Engineering Research Board, a grant from the Infosys foundation and a Department of Science and Technology, Government of India, Swarnajayanti fellowship. J. H. was supported by a KIAS Individual Grant MG088401 at Korea Institute for Advanced Study. The authors were supported by the Department of Atomic Energy, Government of India, under project no.12-R&D-TFR-5.01-0500.

Competing interest. On behalf of all authors, the corresponding author states that there is no competing interest.

References

[1] M. Alam, A. Ghosh and S. Yu, Quantitative Diophantine approximation with congruence conditions, *J. Théor. Nombres Bordeaux* **33**(1) (2021), 261–271.

[2] J. S. Athreya, Random affine lattices, *Contemp. Math.* **639** (2015), 160–74.

[3] P. Billingsley, *Convergence of Probability Measures (Wiley Series in Probability and Statistics)*, second edn. (John Wiley & Sons Inc., New York, 1999).

[4] D. El-Baz, J. Marklof and I. Vinogradov, The distribution of directions in an affine lattice: Two-point correlations and mixed moments, *Int. Math. Res. Not. IMRN* 2015(5) (**2015**), 1371–1400.

[5] A. Ghosh and J. Han, Values of inhomogeneous forms at S-integral points, *Mathematika* **68**(2) (2022), 565–593.

[6] A. Ghosh, D. Kelmer and S. Yu, Effective density for inhomogeneous quadratic forms I: Generic forms and fixed shifts, *Int. Math. Res. Not. IMRN* 2022(6) (**2022**), 4682–4719.

[7] J. Han, S. Lim and K. Mallahi-Karai, Asymptotic distribution of values of isotropic quadratic forms at S-integral points, *J. Mod. Dyn.* **11** (2017), 501–550.

[8] J. Han, Rogers' mean value theorem for S-arithmetic Siegel transform and applications to the geometry of numbers, *J. Number Theory* **240** (2022), 74–106.

[9] S. Kim, Random lattice vectors in a set of size O(n), *Int. Math. Res. Not. IMRN* 2020(5) (**2020**), 1385–1416.

[10] J. Marklof, The n-point correlations between values of a linear form, *Ergodic Theory Dynam. Systems* **20** (2000), 1127–1172.

[11] J. Marklof and A. Strömbergsson, The distribution of free path lengths in the periodic Lorentz gas and related lattice point problems, *Ann. of Math. (2)* **172**(3) (2010), 1949–2033.
[12] G. Prasad, Volumes of S-arithmetic quotients of semi-simple groups, *Inst. Hautes Études Sci. Publ. Math.* **69** (1989), 91–117. With an appendix by Moshe Jarden and the author.
[13] C. Rogers, Mean values over the space of lattices, *Acta Math.* **94** (1955), 249–287.
[14] C. Rogers, The moments of the number of points of a lattice in a bounded set, *Phil. Trans. Roy. Soc. London. Ser. A.* **248** (1955), 225–251.
[15] C. Rogers, Two integral inequalities, *J. London Math. Soc.* **31** (1956), 235–238.
[16] C. Rogers, The number of lattice points in a set, *Proc. London Math. Soc. (3)* **6** (1956), 305–320.
[17] W. M. Schmidt, On the convergence of mean values over lattices, *Canad. J. Math.* **10** (1958), 103–110.
[18] W. M. Schmidt, Masstheorie in der Geometrie der Zahlen, *Acta Math.* **102** (1959), 159–224.
[19] W. M. Schmidt, 'A metrical theorem in geometry of numbers', *Trans. Amer. Math. Soc.* **95** (1960), 516–529.
[20] C. Siegel, A mean value theorem in geometry of numbers, *Ann. Math.* **46** (1945), 340–347.
[21] C. Siegel, *Lectures on the Geometry of Numbers* (Springer-Verlag Berlin Heidelberg GmbH, 1989). x+160 pp.
[22] A. Strömbergsson and A. Södergren, On the generalized circle problem for a random lattice in large dimension, *Adv. Math.* **345** (2019), 1042–1074.
[23] A. Södergren, On the Poisson distribution of lengths of lattice vectors in a random lattice, *Math. Z.* **269**(3–4) (2011), 945–954.