

A higher moment formula for the Siegel–Veech transform over quotients by Hecke triangle groups

Samantha Fairchild

Abstract. We compute higher moments of the Siegel–Veech transform over quotients of $SL(2, \mathbb{R})$ by the Hecke triangle groups. After fixing a normalization of the Haar measure on $SL(2, \mathbb{R})$ we use geometric results and linear algebra to create explicit integration formulas which give information about densities of k -tuples of vectors in discrete subsets of \mathbb{R}^2 which arise as orbits of Hecke triangle groups. This generalizes work of W. Schmidt on the variance of the Siegel transform over $SL(2, \mathbb{R})/SL(2, \mathbb{Z})$.

Mathematics Subject Classification (2010). Primary: 11H60; Secondary: 28C10, 52C23, 30F30.

Keywords. Hecke triangle group, Siegel–Veech transform.

1. Introduction

The Siegel–Veech transform maps a function on \mathbb{R}^2 to a function on sets of translation surfaces. This powerful transformation gives information about the asymptotic density of saddle connections [17] and cusp excursions [3]. On connected strata of translation surfaces, the Siegel–Veech transform is integrable [16] and in L^2 with respect to the Masur–Veech measure [2]. In [16], Veech also showed that the Siegel–Veech transform is integrable over closed $SL(2, \mathbb{R})$ orbits of Veech surfaces with respect to the induced Haar measure. Building on work of Siegel, Schmidt, and Rogers [15, 14, 13] we compute higher moments of the Siegel–Veech transform over sets of surfaces with the Hecke triangle groups as their stabilizer group.

The Hecke triangle group H_q for integers $q \geq 3$ is the discrete subgroup of $SL(2, \mathbb{R})$ generated by

$$S = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad T = \begin{bmatrix} 1 & \lambda_q \\ 0 & 1 \end{bmatrix}, \quad \text{where } \lambda_q = 2 \cos\left(\frac{\pi}{q}\right).$$

Note $H_3 = \mathrm{SL}(2, \mathbb{Z})$ and for all $q \geq 3$, H_q has finite co-volume in $\mathrm{SL}(2, \mathbb{R})$. For more information on Hecke triangle groups see [10].

Let V_q be the discrete subset of \mathbb{R}^2 defined by

$$V_q = H_q \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

which corresponds to a subset of saddle connections of a translation surface when q is odd (see Section 2.3). Define $Y_q = \mathrm{SL}(2, \mathbb{R})/H_q$ with corresponding Haar probability measure μ . Let $B_c((\mathbb{R}^2)^k)$ be the set of bounded measurable functions with compact support on $(\mathbb{R}^2)^k$.

Definition 1.1. For $f \in B_c((\mathbb{R}^2)^k)$ and, by abuse of notation $g = [g] \in Y_q$, we define the *Siegel–Veech transform* by

$$\hat{f}(g) = \sum_{(v_1, \dots, v_k) \in V_q^k} f(g(v_1, \dots, v_k)).$$

In the above definition $k = 1$ is the classical Siegel–Veech transform, and for particular $f \in B_c((\mathbb{R}^2)^k)$ of the form $f(x_1, \dots, x_k) = h(x_1) \cdots h(x_k)$ for $h \in B_c(\mathbb{R}^2)$, \hat{f} corresponds to the k th power of the classical Siegel–Veech Transform of h on \mathbb{R}^2 . Veech proved that the classical Siegel–Veech transform is integrable with the following formula from Section 16 of [17].

Theorem 1.2. For $f \in B_c(\mathbb{R}^2)$,

$$\int_{Y_q} \hat{f}(g) d\mu(g) = \frac{1}{c(q)} \int_{\mathbb{R}^2} f(x) dx,$$

where the Siegel–Veech constant is given by

$$c(q) \stackrel{\text{def}}{=} \pi \left(\pi - \frac{\pi}{q} - \frac{\pi}{2} \right).$$

We will first prove the following theorem which computes the square of the classical Siegel–Veech transform on $B_c(\mathbb{R}^2)$. To state the theorem, we introduce the following two definitions:

Definition 1.3 (Set of non-vanishing determinants). Let

$$\begin{aligned} N_q &\stackrel{\text{def}}{=} \{n \in \mathbb{Z}[\lambda_q] \setminus \{0\}: \text{there exists } v_1, v_2 \in V_q \text{ with } \det(v_1 v_2) = n\} \\ &= \left\{ n \in \mathbb{Z}[\lambda_q] \setminus \{0\}: \text{there exists } 0 \leq m < \lambda_q |n| \text{ with } \begin{bmatrix} m \\ n \end{bmatrix} \in V_q \right\}, \end{aligned} \quad (1)$$

where (1) will be proved in Lemma 3.5.

Definition 1.4 (q -geometric Euler totient function). For $b \in \mathbb{Z}[\lambda_q]$ define

$$\varphi_q(b) = \#\left\{1 \leq a \leq \lambda_q |b| : \begin{bmatrix} a \\ b \end{bmatrix} \in V_q\right\}.$$

Note that $\lambda_3 = 1$ and $V_3 = \mathrm{SL}(2, \mathbb{Z}) \cdot e_1$ so φ_3 reduces to the standard Euler totient function.

Theorem 1.5. Let $f \in B_c(\mathbb{R}^2 \times \mathbb{R}^2)$, N_q be the set of non-vanishing determinants, and φ_q the q -geometric Euler totient function. Then,

$$\int_{Y_q} \hat{f} d\mu(g) = \sum_{n \in N_q} \frac{\varphi_q(n)}{c(q)} \int_{\mathrm{SL}(2, \mathbb{R})} f(gJ_n) d\eta + \frac{1}{c(q)} \int_{\mathbb{R}^2} (f(x, -x) + f(x, x)) dx \quad (2)$$

where $J_n = \begin{bmatrix} 1 & 1 \\ 0 & n \end{bmatrix}$, μ is the Haar probability measure on Y_q , η is Haar measure on $\mathrm{SL}(2, \mathbb{R})$ normalized so $\eta(Y_3) = \frac{\pi^2}{6}$, and dx is the Lebesgue measure on \mathbb{R}^2 normalized so the area of the unit square is 1.

Note \hat{f} is uniformly bounded by Lemma 16.10 of [17], so both sides of (2) are finite. The proof of Theorem 1.5 will use Schmidt's outline of proof (see Section 2.1). It is a useful exercise to consider this proof in the case of Schmidt with $q = 3$. That is where $N_3 = \mathbb{Z} \setminus \{0\}$ and the constant $c(3) = \frac{\pi^2}{6} = \zeta(2)$. In Section 5 we will see how the formula in Theorem 1.5 allows us to understand the asymptotic densities of saddle connections of translation surfaces with Veech group H_q for q odd. Theorem 1.5 is in fact a special case of the main theorem, which calculates the k th moment of the classical Siegel–Veech transform.

Theorem 1.6. Let $f \in B_c((\mathbb{R}^2)^k)$ and define

$$J_{n,m} = \begin{bmatrix} 1 & m \\ 0 & n \end{bmatrix}.$$

Then

$$\begin{aligned} & \int_{Y_q} \hat{f} d\mu(g) \\ &= \sum_{\substack{\lambda \in \mathbb{R}^k \\ \lambda = (1, \pm 1, \dots, \pm 1)}} \frac{1}{c(q)} \int_{\mathbb{R}^2} f(\lambda x) dx \\ &+ \sum_{n \in N_q} \sum_{\substack{0 \leq m < \lambda_q |n| \\ (m,n)^T \in V_q}} \sum_{1 \leq j < k} \sum_{\lambda, \alpha, \beta} \frac{1}{c(q)} \int_{\mathrm{SL}(2, \mathbb{R})} f\left(\lambda g \begin{bmatrix} 1 \\ 0 \end{bmatrix}, gJ_{n,m} \begin{bmatrix} \alpha \\ \beta \end{bmatrix}\right) d\eta(g), \end{aligned}$$

where for each $1 \leq j < k$ we have $\lambda \in \mathbb{R}^j$ is of the form $(1, \pm 1, \dots, \pm 1)$ and $\alpha = (0, \alpha_2, \dots, \alpha_{k-j})$ and $\beta = (1, \beta_2, \dots, \beta_{k-j})$ where for each $2 \leq i \leq k-j$ we have

$$\begin{bmatrix} \alpha_i \\ \beta_i \end{bmatrix} \in J_{n,m}^{-1} V_q. \quad (3)$$

1.1. Outline. In Section 2 we give an overview of the history of the problem, followed by the necessary background on translation surfaces, Veech groups, and the geometric Euler totient function. In Section 3 we prove Theorem 1.5, followed by Section 4 where we prove Theorem 1.6. Finally in Section 5 we explain how we found numerical evidence for the result.

Acknowledgments. I thank Jayadev Athreya for proposing the project and many useful discussions. I thank Bianca Viray and Claire Burrin for useful comments and discussions about the totient function. Thanks to Anthony Sanchez for useful comments on the paper. Finally I would like to thank Kimberly Bautista, Maddy Brown, and Andrew Lim of the Washington Experimental Mathematics Lab for their contributions to my numerical experiments and discussion in generalizing to higher moments.

2. Background and history

We first give a summary of previous related results in the geometry of numbers, followed by background on translation surfaces, Veech groups, and the q -geometric Euler totient function.

2.1. Geometry of numbers. We will first focus on the mean and variance of the primitive Siegel transform, which is a special case of the Siegel–Veech transform defined in the previous section. First we set up some notation and definitions, then state the theorems of Siegel, Rogers, and Schmidt computing the mean and variance of the primitive Siegel transform.

Consider $f \in B_c(\mathbb{R}^d)$. We aim to understand f evaluated on *visible* lattice points in \mathbb{R}^d , where a point $v = (v_1, \dots, v_d)^T \in \mathbb{Z}^d$ is *primitive* or *visible* if $\gcd(v_1, \dots, v_d) = 1$. We denote the set of primitive vector points by $\mathbb{Z}_{\text{prim}}^d$, which one can show $\mathbb{Z}_{\text{prim}}^d = \text{SL}(d, \mathbb{Z}) \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Define $X_d = \text{SL}(d, \mathbb{R}) / \text{SL}(d, \mathbb{Z})$. By abuse of notation, for an equivalence class $g = [g] \in X_d$, we define the *primitive Siegel transform* by

$$\hat{f}(g) = \sum_{v \in \mathbb{Z}_{\text{prim}}^d} f(gv).$$

In 1945, Siegel [15], Sections 5–6, showed

$$\int_{X_d} \hat{f}(g) d\mu(g) = \frac{1}{\zeta(d)} \int_{\mathbb{R}^d} f(x) dx \quad (4)$$

where the standard Lebesgue measure on \mathbb{R}^d is dx , ζ is the Riemann zeta function, and μ is probability Haar measure on X_d .

In order to understand higher moments of \hat{f} , we split into the cases where $d = 2$ and $d > 2$. We address the latter case first.

For understanding higher moments of \hat{f} , C. A. Rogers' 1955 paper [13], Theorem 5 solved the case for \hat{f}^k with $d > 2$ and $k < d$. For simplicity, we will only consider the case $k = 2$ of Rogers' result. Recall for $f \in B_c(\mathbb{R}^d)$, and defining $h \in B_c(\mathbb{R}^2 \times \mathbb{R}^2)$ by $h(x, y) = f(x)f(y)$ we have

$$\sum_{v_1, v_2 \in \mathbb{Z}_{\text{prim}}^d} h(gv_1, gv_2) = \sum_{v_1, v_2 \in \mathbb{Z}_{\text{prim}}^d} f(gv_1)f(gv_2) = \left[\sum_{v \in \mathbb{Z}_{\text{prim}}^d} f(gv) \right]^2 = (\hat{f})^2.$$

Rogers showed that for $f \in B_c(\mathbb{R}^d)$, and $h(x, y) = f(x)f(y)$, the second moment of f is given by

$$\begin{aligned} \int_{X_d} (\hat{f})^2(g) d\mu(g) &= \int_{X_d} \sum_{v_1, v_2 \in \mathbb{Z}_{\text{prim}}^d} h(gv_1, gv_2) d\mu(g) \\ &= \frac{1}{\zeta(d)^2} \int_{\mathbb{R}^d \times \mathbb{R}^d} h(x, y) dx dy + \frac{1}{\zeta(d)} \int_{\mathbb{R}^d} [h(x, x) + h(x, -x)] dx. \end{aligned} \quad (5)$$

For a modern proof of (5), see Section 4 of [3].

For $k \geq d > 2$ and $f \in B_c(\mathbb{R}^d)$, the function \hat{f}^k is not integrable (Proposition 7.1 of [9]). However when $d = 2$ we have \hat{f} is bounded on X_2 , and thus \hat{f}^k integrable for any $k \geq 1$. So we now exclusively study the case $d = 2$. Rogers had a mistake in his paper claiming (5) held for $d = 2$, which we can see does not work by setting h_0 to be the characteristic function of the set given by

$$\{(v_1, v_2): \max(|v_1|, |v_2|) \leq R, \det(v_1 v_2) \notin \mathbb{Z}\}.$$

Applying (5) to h_0 , the left hand side of (5) will be identically zero as for any $v_1, v_2 \in \mathbb{Z}_{\text{prim}}^2$

$$\det(gv_1 gv_2) = \det(g) \det(v_1 v_2) \in \mathbb{Z},$$

and the right hand side will be nonzero as the vectors with integer determinant are a Lebesgue measure zero subset of $\mathbb{R}^2 \times \mathbb{R}^2$.

In correction to Rogers, Schmidt addressed the case where $d = 2$ (see [14], Section 6).

Theorem 2.1. *Let $f \in B_c(\mathbb{R}^2 \times \mathbb{R}^2)$. Then*

$$\int_{Y_3} \hat{f} d\mu(g) = \sum_{n \in \mathbb{Z} \setminus \{0\}} \frac{\varphi(n)}{\zeta(2)} \int_{\mathrm{SL}(2, \mathbb{R})} f(gJ_n) d\eta + \frac{1}{\zeta(2)} \int_{\mathbb{R}^2} f(x, -x) + f(x, x) dx \quad (6)$$

where φ is the standard Euler totient function.

Note this formula does not look exactly like the formula in Schmidt [14] as we have a different normalization of the Haar measure η . Note also that Theorem 2.1 is a special case of Theorem 1.5.

2.2. Translation surfaces. A *translation surface* is a surface formed by taking a finite number of polygons in the plane and gluing opposite sides by translation, where surfaces are equivalent up to cutting and pasting of these polygons via translation. Equivalently a translation surface is a closed Riemann surface X with a nonzero holomorphic 1-form ω . This section will focus on examples relevant to this paper. For more background see [11], [8], and [5].

Given $A \in \mathrm{SL}(2, \mathbb{R})$ and (X, ω) a translation surface, we produce a new translation surface $A \cdot (X, \omega)$, which is the surface with charts of (X, ω) composed with A acting linearly on \mathbb{R}^2 . The *Veech group* is the stabilizer subgroup of this action

$$\mathrm{SL}(X, \omega) \stackrel{\mathrm{def}}{=} \{A \in \mathrm{SL}(2, \mathbb{R}) : A \cdot (X, \omega) = (X, \omega)\}.$$

The Veech group is always discrete and in fact trivial for almost every translation surface [6].

Saddle connections on a translation surface (X, ω) are geodesics which start and end at zeros of the 1-form ω on X . For each saddle connection γ , there is an associated *holonomy vector* $v_\gamma = \int_\gamma \omega \in \mathbb{R}^2$ which records the length and direction of γ .

2.3. Hecke triangle groups as Veech groups. We will consider surfaces whose Veech group is given by $\mathrm{SL}(X, \omega) = H_q$ for $q \geq 3$. When $q = 3$, $H_3 = \mathrm{SL}(2, \mathbb{Z})$ which is the Veech group for the square torus. In general given a translation surface (X, ω) where we glue two regular $(2n + 1)$ -gons and then identify opposite sides, Veech showed in [16] that $\mathrm{SL}(X, \omega) = H_{2n+1}$. For even Hecke triangle groups, Bouw and Möller [4] followed by a constructive proof of Hooper [7] were able to show that there exists a translation surface (X, ω) with $\mathrm{SL}(X, \omega)$ conjugate to an index 2 subgroup of H_{2n} , but there is no translation surface with Veech group containing H_{2n} .

Notice the set of holonomy vectors for the square torus are

$$\mathbb{Z}_{\mathrm{prim}}^2 = \mathrm{SL}(2, \mathbb{Z}) \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} = H_3 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

The characterization is not as clean for other surfaces, but if (X, ω) is a translation surface with $\text{SL}(X, \omega)$ a lattice, then the set of holonomy vectors will always be given as a finite union of $\text{SL}(X, \omega)$ -orbits [16], 5th paragraph, Section 3. By studying the Siegel–Veech transform over V_q we will be able to understand asymptotic density of saddle connections for a class of translation surfaces [17].

2.4. Geometric Euler totient function. Recall we define the q -geometric Euler totient function by

$$\varphi_q(b) = \#\left\{1 \leq a \leq \lambda_q |b| : \begin{bmatrix} a \\ b \end{bmatrix} \in V_q\right\},$$

where $\varphi_3 = \varphi$ is the standard Euler totient function. Since V_q is discrete and thus φ_q is finite and well defined. Though φ_q generalizes the standard Euler totient function, φ_q does *not* agree with the more standard Euler totient function defined for the ring of integers over a number field in terms of the product formula over prime ideals.

Following [10], we can define a *greatest common q -divisor* denoted $(a, b)_q$ for $a, b \in \mathbb{Z}[\lambda_q]$ using a Euclidean pseudo-algorithm. This greatest common q -divisor has many similar properties to the gcd function, including for any $t \neq 0$,

$$(ta, tb)_q = t \cdot (a, b)_q. \tag{7}$$

With this definition we also have the following useful characterization of elements of H_q as proved in Proposition 3.7 of [10].

Proposition 2.2. *A matrix $\begin{bmatrix} a & c \\ b & d \end{bmatrix} \in \text{SL}(2, \mathbb{Z}[\lambda_q])$ is in H_q if and only if $(a, b)_q = (c, d)_q = 1$. In fact if $(a, b)_q < 1$ or $|b| < 1$, then $(a, b)^T$ cannot be a column of a matrix in H_q .*

3. Orbits and integrals

The goal of this section is to prove Theorem 1.5.

Let $f \in B_c(\mathbb{R}^2 \times \mathbb{R}^2)$, and define \hat{f} as in Definition 1.1. Consider the map

$$f \mapsto \int_{Y_q} \hat{f} d\mu.$$

This mapping is a positive linear functional which is $\text{SL}(2, \mathbb{R})$ -invariant, where $\text{SL}(2, \mathbb{R})$ acts diagonally by $g \cdot (v_1, v_2) = (gv_1, gv_2)$ for $(v_1, v_2) \in \mathbb{R}^2 \times \mathbb{R}^2$. Hence by the Riesz representation theorem, there exists a measure ν so that

$$\int_{\mathbb{R}^2 \times \mathbb{R}^2} f d\nu = \int_{Y_q} \hat{f} d\mu \quad \text{for any } f \in B_c(\mathbb{R}^2 \times \mathbb{R}^2).$$

Since ν is $\mathrm{SL}(2, \mathbb{R})$ -invariant, we can write ν as a combination of measures on $\mathrm{SL}(2, \mathbb{R})$ orbits of $\mathbb{R}^2 \times \mathbb{R}^2$. So to understand ν we need to understand our integral over $\mathrm{SL}(2, \mathbb{R})$ orbits.

The outline of the proof is as follows. In Section 3.1 we split $\mathbb{R}^2 \times \mathbb{R}^2$ into $\mathrm{SL}(2, \mathbb{R})$ orbits under the diagonal action and find the possible $\mathrm{SL}(2, \mathbb{R})$ -invariant measures on these subsets. In Section 3.2 we will reduce the uncountable number of orbits which occur in our setting to two linearly dependent orbits, and a countable number of linearly independent orbits. After setting up notation in Section 3.3, in Section 3.4 we reduce the linearly dependent case to Theorem 1.2, finally addressing the linearly independent case in Section 3.5.

3.1. Decomposition into orbits. Let $\mathbb{R}_0^n = \mathbb{R}^n \setminus \{0\}$, similarly $\mathbb{Z}_0^n = \mathbb{Z}^n \setminus \{0\}$, and $\mathbb{Z}_0[\lambda_q] = \mathbb{Z}[\lambda_q] \setminus \{0\}$.

Lemma 3.1. *The following decomposes $\mathbb{R}^2 \times \mathbb{R}^2$ into disjoint $\mathrm{SL}(2, \mathbb{R})$ orbits:*

$$\mathbb{R}^2 \times \mathbb{R}^2 = \left(\bigsqcup_{n \in \mathbb{R}_0} D_n \right) \sqcup \left(\bigsqcup_{t \in \mathbb{R}_0} \mathrm{LD}_t \right) \sqcup H \sqcup V \sqcup \{0\},$$

where we have the linearly independent determinants,

$$D_n = \{(v, w) \in \mathbb{R}^2 \times \mathbb{R}^2 : \det(v \ w) = n\},$$

the linearly dependent subsets

$$\mathrm{LD}_t = \{(v, tv) : v \in \mathbb{R}_0^2\},$$

and two special cases of linearly dependent vectors: horizontal and vertical

$$H = \{(v, 0) : v \in \mathbb{R}_0^2\} \quad V = \{(0, v) : v \in \mathbb{R}_0^2\}.$$

Proof. We will realize each subset as an orbit of $\mathrm{SL}(2, \mathbb{R})$ under the diagonal action on $\mathbb{R}^2 \times \mathbb{R}^2$. Since $g \cdot \{0\} = 0$ for all $g \in \mathrm{SL}(2, \mathbb{R})$, the point $\{0\}$ is an entire orbit.

Now notice that $g \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \mathbb{R}_0^2$. Using this fact, for any $t \in \mathbb{R}_0$,

$$\mathrm{SL}(2, \mathbb{R}) \cdot \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} t \\ 0 \end{bmatrix} \right) = \{(v, tv) : v \in \mathbb{R}_0^2\} = \mathrm{LD}_t.$$

Similarly for H and V , it suffices to see that they are both given by

$$H = \mathrm{SL}(2, \mathbb{R}) \cdot \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right) \quad V = \mathrm{SL}(2, \mathbb{R}) \cdot \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right).$$

Finally, for $n \neq 0$, since $n\mathbb{R}_0^2 = \mathbb{R}_0^2$,

$$\mathrm{SL}(2, \mathbb{R}) \cdot \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ n \end{bmatrix} \right) = \{(v, nu) : \det(v \ u) = 1\} = D_n.$$

Thus we have shown each of these subsets is an $SL(2, \mathbb{R})$ orbit. Finally, since every pair of elements in \mathbb{R}^2 is either linearly independent and thus have a nonzero determinant or linearly dependent and thus are scalar multiples we conclude every element $(v_1, v_2) \in \mathbb{R}^2 \times \mathbb{R}^2$ is contained in one of the given sets. Thus we have a decomposition of $\mathbb{R}^2 \times \mathbb{R}^2$ into $SL(2, \mathbb{R})$ orbits. \square

The last task of this subsection is to determine the possible measures on each of our subsets. We will freely use the fact that Haar measure is unique up to scaling. In this section we will fix a particular scaling of Haar measure for each measure, and then by taking a linear combination of these different measures we can obtain ν .

On $\{0\}$, there is only one probability measure given by δ_0 , which is trivially $SL(2, \mathbb{R})$ invariant.

On H, V , and LD_t for $t \in \mathbb{R}_0$, we have a copy of \mathbb{R}_0^2 . Notice Lebesgue measure m_2 on \mathbb{R}^2 is $SL(2, \mathbb{R})$ -invariant. So we will fix the standard Lebesgue measure giving the unit square $[0, 1]^2$ volume 1 on each of the subsets H, V , and LD_t for $t \in \mathbb{R}_0$. Since $\{(0, 0)\}$ is a measure zero subset, without loss of generality we can write integrals with respect to m_2 over all of \mathbb{R}^2 . To see this measure is the unique $SL(2, \mathbb{R})$ -invariant measure (up to scaling), consider the induced Haar measure under the quotient of $SL(2, \mathbb{R})/N \cong \mathbb{R}_0^2$, where $N = \left\{ \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} : t \in \mathbb{R} \right\}$.

To find a Haar measure on D_n , we will first find a Haar measure on $SL(2, \mathbb{R})$, then we will show how this can be viewed as a Haar measure on D_n . To construct a Haar measure on $SL(2, \mathbb{R})$, consider $SL(2, \mathbb{R})$ as a subset of (\mathbb{R}^4, m_4) where m_k is Lebesgue measure on \mathbb{R}^k . As a result, for measurable $A \subseteq SL(2, \mathbb{R})$, we can define the cone measure

$$\eta(A) = m_4(C(A)) \text{ where } C(A) = \{\alpha g : \alpha \in (0, 1], g \in A\}.$$

Under matrix multiplication, m_4 is $SL(2, \mathbb{R})$ invariant. Hence η is an $SL(2, \mathbb{R})$ invariant measure on $SL(2, \mathbb{R})$. Under this measure, the set of matrices with a zero in the top left corner is a null set. Thus we can write the measure $d\eta = da db ds$ under the coordinates

$$\begin{bmatrix} 1 & 0 \\ s & 1 \end{bmatrix} \cdot \begin{bmatrix} a & b \\ 0 & a^{-1} \end{bmatrix}.$$

With this normalization, in the quotient by $SL(2, \mathbb{Z})$, we can compute the push-forward defined in terms of the projection map π and fundamental domain F [1]

$$(\eta)_*(Y_3) = \eta(\pi^{-1}(Y_3) \cap F) = \frac{\pi^2}{6} = \zeta(2).$$

With this fixed normalization, η gives the Poincaré volume. This means that we in fact have

$$(\eta)_*(Y_q) = c(q).$$

Now having fixed Haar measure on D_1 , for D_n with $n \neq 0$, we identify D_n with $D_1 = \mathrm{SL}(2, \mathbb{R})$ as $D_n = D_1 J_n$. Since we can write $D_n = D_1 J_n$, we choose the coordinates on D_n to be the same as those on D_1 . In this manner, we have η is the Haar measure we will choose as our normalization of Haar measure on D_n .

We've now decomposed $\mathbb{R}^2 \times \mathbb{R}^2$ into $\mathrm{SL}(2, \mathbb{R})$ orbits, and fixed a normalization of Haar measure on each of these orbits.

Since Haar measure is unique up to scaling, we can now write our $\mathrm{SL}(2, \mathbb{R})$ invariant measure on \mathbb{R}^4 as

$$v = a\delta_0 + \sum_{t \in \mathbb{R} \cup \{\infty\}} b_t m_2 + \sum_{n \in \mathbb{R}_0} c_n \eta$$

for some constants a, b_t, c_n . where b_∞ corresponds to V and b_0 corresponds to H .

3.2. Reduction to visible determinants and removal of zero term. We have shown

$$\begin{aligned} \int_{Y_q} \hat{f} d\mu &= \int_{\mathbb{R}^2 \times \mathbb{R}^2} f dv \\ &= af(0, 0) + \int_{t \in \mathbb{R} \cup \{\infty\}} b_t \int_{\mathbb{R}^2} f(x, tx) dx + \int_{n \in \mathbb{R}_0} c_n \int_{\mathrm{SL}(2, \mathbb{R})} f(gJ_n) d\eta, \end{aligned} \quad (8)$$

where we define $(x, \infty x) = (0, x)$.

The purpose of this section is to prove the following.

Lemma 3.2. *In (8), $a = 0$, $t \in \{\pm 1\}$, and $n \in N_q$.*

Proof. To see that $t \in \{\pm 1\}$, consider the function f supported on LD_t for $t \in \mathbb{R} \cup \{\infty\}$ where $\mathrm{LD}_0 = H$ and $\mathrm{LD}_\infty = V$. That is, for some large R and $B(0, R)$ denoting the Euclidean ball in \mathbb{R}^4 , let

$$f_{R,t}(x, y) = \chi_{B(0,R) \setminus \{0\}}(x, y) \chi_{\mathrm{LD}_t}(x, y) \quad \text{for } x, y \in \mathbb{R}^2.$$

On the left hand side of (8), notice

$$\begin{aligned} \hat{f}_{R,t}([g]) &= \sum_{v_1, v_2 \in V_q} f_{R,t}(g v_1, g v_2) \\ &= \#\{v_1, v_2 \in V_q \cap B(0, R): g v_1 = t g v_2\} \\ &= \#\{v_1, v_2 \in V_q \cap B(0, R): v_1 = t v_2\}. \end{aligned}$$

If $\begin{bmatrix} a \\ c \end{bmatrix} \in V_q$ by Proposition 2.2, we have $(a, c)_q = 1$, and thus (7), $(ta, tc)_q = t$.

So by Proposition 2.2 [t_a^c] cannot be an element of V_q unless $t = \pm 1$. Or more geometrically since V_q are the set of vectors visible from the origin, tv is never visible from the origin unless $t = \pm 1$. Hence we've shown

$$\hat{f}_{R,t} = \begin{cases} 0 & \text{if } t \neq \pm 1, \\ \#\{v \in V_q \cap B(0, R)\} & \text{if } t = \pm 1. \end{cases}$$

On the right hand side of (8), the only nonzero term will be the coefficient of b_t for if $x \in B(0, \frac{R}{t})$, then $tx \in B(0, R)$. Thus,

$$\int_{\mathbb{R}^2} f_{R,t}(x, tx) dx \geq m_2\left(B_{\mathbb{R}^2}\left(0, \frac{R}{t}\right)\right) > 0.$$

But when $t \neq \pm 1$, the left hand side of (8) is zero since $\hat{f}_R = 0$. Hence $b_t = 0$ for $t \neq \pm 1$.

We now want to show that the set of possible determinants is N_q . For the determinant n loci ($n \neq 0$), we similarly define

$$f_{R,n}(x, y) = \chi_{B(0,R)}(x, y) \chi_{D_n}(x, y) \quad \text{for } x, y \in \mathbb{R}^2.$$

We compute

$$\begin{aligned} \hat{f}_{R,n}([g]) &= \sum_{v_1, v_2 \in V_q} f_{R,n}(g v_1, g v_2) \\ &= \#\{v_1, v_2 \in V_q \cap B(0, R): \det(v_1 v_2) = \det(g v_1, g v_2) = n\}. \end{aligned}$$

Since N_q is the set of determinants that can arise as the determinant of two elements in V_q , we can write

$$\hat{f}_{R,n} = \begin{cases} \#\{v_1, v_2 \in V_q \cap B(0, R): \det(v_1 v_2) = n\} & \text{if } n \in N_q, \\ 0 & \text{if } n \notin N_q. \end{cases}$$

On the right hand side of (8), the only nonzero term corresponds to c_n , and

$$\int_{\text{SL}(2, \mathbb{R})} f_R(g J_n) d\eta > 0$$

since $D_n \cap B(0, R)$ has positive cone measure. In order to match the left hand side of (8) for $\hat{f}_{R,n}$, we conclude $c_n = 0$ for all $n \notin N_q$.

We conclude this proof by showing $a = 0$. To see this, consider the characteristic function over the set $\{(0, 0)\} \subseteq \mathbb{R}^2 \times \mathbb{R}^2$. That is set $f_0(x, y) = \chi_{\{(0,0)\}}(x, y)$. Then on the right hand side of (8), we have $f_0(0, 0) = 1$, all other integrals are zero since $\{(0, 0)\}$ is a measure zero subset of \mathbb{R}^2 , and cannot show up in $\text{SL}(2, \mathbb{R}) J_n$ for any n . Thus the right hand side of (8) for f_0 is a . On the left hand side of (8), $(0, 0)$ is not a pair of visible vectors since $(0, 0)$ cannot be the first column of a matrix in H_q , so the left hand side is zero. Thus we conclude $a = 0$. \square

To summarize, in this section we reduced our equation (8) to

Corollary 3.3.

$$\int_{Y_q} \hat{f} d\mu = \int_{\mathbb{R}^2} b_1 f(x, x) + b_{-1} f(x, -x) dx + \sum_{n \in N_q} c_n \int_{\text{SL}(2, \mathbb{R})} f(gJ_n) d\eta.$$

3.3. Notation and division into smaller lemmas. In the proceeding sections, we will compute the values for b_1, b_{-1} , and c_n for $n \in N_q$. In order to do this, we introduce the following notation: for D a discrete subset of $(\mathbb{R}^2)^k$ which is V_q -invariant under the diagonal action, define $f_D: Y_q \rightarrow \mathbb{R}$ by

$$f_D([g]) = \sum_{v \in D^k} f(gv)$$

In a similar manner define the functional $T_D: B_c((\mathbb{R}^2)^k) \rightarrow \mathbb{R}$ by

$$T_D(f) = \int_{Y_q} f_D([g]) d\mu([g]).$$

We now define the following sets:

$$D_n^V = D_n \cap (V_q \times V_q) = \{(v, w) \in V_q \times V_q : \det(v w) = n\},$$

$$\text{LD}_{\pm 1}^V = \{(v, \pm v) : v \in V_q\}.$$

Then we can rewrite the left hand side of Corollary 3.3 as

$$\begin{aligned} & \int_{Y_q} \left[f_{\text{LD}_1^V} + f_{\text{LD}_{-1}^V} + \sum_{n \in N_q} f_{D_n^V} \right] d\mu \\ &= T_{\text{LD}_1^V}(f) + T_{\text{LD}_{-1}^V}(f) + \sum_{n \in N_q} T_{D_n^V}(f). \end{aligned}$$

Thus finding the coefficients in Corollary 3.3 is reduced to finding coefficients individually in each of these equations:

$$T_{\text{LD}_{\pm 1}^V}(f) = b_{\pm 1} \int_{\mathbb{R}^2} f(x, \pm x) dx, \quad (9)$$

and for each $n \in N_q$

$$T_{D_n^V}(f) = c_n \int_{\text{SL}(2, \mathbb{R})} f(gJ_n) d\eta. \quad (10)$$

3.4. Reducing to Siegel–Veech formula in linearly dependent case. In this section, we will prove that the coefficients b_1 and b_{-1} in (9) are given by $b_1 = b_{-1} = \frac{1}{c(q)}$ by reducing to the Siegel–Veech Primitive Integral Formula (Theorem 1.2). That is, we will prove the following:

Lemma 3.4. *For any $f \in B_c(\mathbb{R}^2 \times \mathbb{R}^2)$,*

$$T_{\text{LD}_{\pm 1}^V}(f) = \frac{1}{c(q)} \int_{\mathbb{R}^2} f(v, \pm v) dv$$

where $c(q)$ is the Poincaré volume of the unit tangent bundle over \mathbb{H}^2/H_q .

Proof. Given $f \in B_c(\mathbb{R}^2 \times \mathbb{R}^2)$, define $\bar{f} \in B_c(\mathbb{R}^2)$ by

$$\bar{f}_{\pm}(u) = f(u, \pm u).$$

So we now compute

$$\begin{aligned} T_{\text{LD}_{\pm 1}^V}(f) &= \int_{Y_q} \sum_{v \in V_q} f(gu, \pm gu) d\mu([g]) \\ &= \int_{Y_q} \sum_{u \in V_q} \bar{f}_{\pm}(gu) d\mu([g]) \\ &= \frac{1}{c(q)} \int_{\mathbb{R}^2} \bar{f}_{\pm}(x) dx && \text{(by Theorem 1.2)} \\ &= \frac{1}{c(q)} \int_{\mathbb{R}^2} f(x, \pm x) dx. \end{aligned}$$

This concludes the proof of the lemma. □

We've now shown $b_{\pm 1} = c(q)^{-1}$, in the next section, we address the coefficients c_n for $n \in N_q$.

3.5. Coefficients on loci with fixed determinant. The goal of this section is to prove that each $c_n = c(q)^{-1} \varphi_q(n)$ for $n \in N_q$. We will first decompose D_n^V into H_q orbits under the diagonal action, showing there are $\varphi_q(n)$ orbits which each contribute equally to $T_{D_n^V}$. After showing this, we will find the value over a single orbit.

Lemma 3.5. *Let $n \in \mathbb{Z}_0[\lambda_q]$. There exists $v_1, v_2 \in V_q$ with $\det(v_1 v_2) = n$ if and only if there exists $m \in \mathbb{Z}[\lambda_q]$ with $0 \leq m < \lambda_q |n|$ and $\begin{bmatrix} m \\ n \end{bmatrix} \in V_q$.*

In particular, the equality in (1) for N_q holds.

Proof. First, suppose there exists $m \in \mathbb{Z}[\lambda_q]$ with $0 \leq m < \lambda_q |n|$ and $\begin{bmatrix} m \\ n \end{bmatrix} \in V_q$. Set $v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, which is in V_q since $Tv_1 = v_1$, and set $v_2 = \begin{bmatrix} m \\ n \end{bmatrix}$. Then $v_1, v_2 \in V_q$ with determinant n , so $n \in N_q$.

Conversely suppose $v_1, v_2 \in V_q$ with $\det(v_1 v_2) = n$. Let $g \in H_q$ so that $ge_1 = v_1$. Then,

$$g^{-1}[v_1 \ v_2] = [e_1 \ g^{-1}v_2].$$

Since the determinant is n ,

$$g^{-1}v_2 = \begin{bmatrix} \ell \\ n \end{bmatrix}$$

for some $\ell \in \mathbb{Z}[\lambda_q]$. Applying the matrix T to the left j times for some $j \in \mathbb{Z}$ gives

$$T^j g^{-1}[v_1 \ v_2] = [e_1 \ T^j g^{-1}v_2]$$

where

$$T^j g^{-1}v_2 = \begin{bmatrix} \ell + jn\lambda_q \\ n \end{bmatrix}.$$

Thus we can find $j \in \mathbb{Z}$ so that $m = \ell + jn\lambda_q$ satisfies

$$0 \leq m < \lambda_q |n|,$$

and the proof is complete. □

Lemma 3.6. *For $n \in N_q$ The subset D_n^V is the union of $\varphi_q(n)$ different orbits*

$$D_n^V = \bigsqcup_{\substack{0 \leq m < |n|\lambda_q \\ (m,n)^T \in V_q}} E_n^{(m)},$$

where

$$E_n^{(m)} = \left\{ \gamma \cdot \begin{bmatrix} 1 & m \\ 0 & n \end{bmatrix} : \gamma \in H_q \right\}.$$

Proof. We will first show that the decomposition of every element in D_n^V can be written as an element $E_n^{(m)}$ for some m .

Let

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in D_n^V$$

From the proof of Lemma 3.5, there exists a matrix $h \in H_q$ with

$$h \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & m \\ 0 & n \end{bmatrix} \quad \text{for } 0 \leq m < \lambda_q |n|.$$

Since $\begin{bmatrix} b \\ d \end{bmatrix} \in V_q$, we also have $\begin{bmatrix} m \\ n \end{bmatrix} = h\begin{bmatrix} b \\ d \end{bmatrix} \in V_q$. We have now shown every element in D_n^V is in $E_n^{(m)}$ for some m with $0 \leq m < \lambda_q|n|$ and $\begin{bmatrix} m \\ n \end{bmatrix} \in V_q$.

To see that we have no duplicate representatives of our orbits, let $0 < m_1, m_2 \leq \lambda_q|n|$ with $m_1 \neq m_2$ and $\begin{bmatrix} m_1 \\ n \end{bmatrix}, \begin{bmatrix} m_2 \\ n \end{bmatrix} \in V_q$. Without loss of generality suppose $m_2 > m_1$. If the representatives

$$\begin{bmatrix} 1 & m_1 \\ 0 & n \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & m_2 \\ 0 & n \end{bmatrix}$$

were in the same H_q orbit, there would exist an element $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in H_q$ such that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & m_1 \\ 0 & n \end{bmatrix} = \begin{bmatrix} 1 & m_2 \\ 0 & n \end{bmatrix}.$$

This implies $a = 1, c = 0, d = 1$, and $b = \frac{m_2 - m_1}{n}$. Since $0 < m_2 - m_1 < \lambda_q n$, we have $0 < b < \lambda_q$. This is a parabolic element with upper right entry smaller than the generating matrix T , and thus not in H_q . Therefore we conclude that $E_n^{(m)}$ are $\varphi_q(n)$ distinct H_q orbits whose union is all of D_n^V . \square

Lemma 3.7. *For a fixed m with $0 \leq m < \lambda_q|n|$ and $\begin{bmatrix} m \\ n \end{bmatrix} \in V_q$,*

$$T_{E_n^{(m)}}(f) = \frac{1}{c(q)} \int_{\mathrm{SL}(2, \mathbb{R})} f(gJ_n) d\eta_2.$$

Proof. Let $\pi: \mathrm{SL}(2, \mathbb{R}) \rightarrow Y_q$ be the projection map $g \mapsto [g]$. Recall we normalize η so that $\pi_*(\eta)(Y_q) = c(q)$. Hence $\pi_*(c(q)^{-1}\eta) = \mu$. Moreover, to push a function from $\mathrm{SL}(2, \mathbb{R})$ to a function on X_2 , we have to sum over the orbits H_q . Thus,

$$\frac{1}{c(q)} \int_{\mathrm{SL}(2, \mathbb{R})} f\left(g \begin{bmatrix} 1 & m \\ 0 & n \end{bmatrix}\right) d\eta = \int \sum_{Y_q \gamma \in H_q} f\left(g \cdot \gamma \cdot \begin{bmatrix} 1 & m \\ 0 & n \end{bmatrix}\right) d\mu = T_{E_n^{(m)}}(f).$$

For the last part of the lemma, we compute the following:

$$\begin{aligned} & \frac{1}{c(q)} \int_{\mathrm{SL}(2, \mathbb{R})} f\left(g \begin{bmatrix} 1 & m \\ 0 & n \end{bmatrix}\right) d\eta \\ &= \frac{1}{c(q)} \int_{\mathrm{SL}(2, \mathbb{R})} f\left(g \begin{bmatrix} 1 & \frac{1-m}{n} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & m \\ 0 & n \end{bmatrix}\right) d\eta \left(g \begin{bmatrix} 1 & \frac{1-m}{n} \\ 0 & 1 \end{bmatrix}\right) \\ &= \frac{1}{c(q)} \int_{\mathrm{SL}(2, \mathbb{R})} f\left(g \begin{bmatrix} 1 & 1 \\ 0 & n \end{bmatrix}\right) d\eta(g), \end{aligned}$$

where the last equality follows from the fact that $\mathrm{SL}(2, \mathbb{R})$ is a unimodular group, so the Haar measure η is both left and right invariant under the action of $\mathrm{SL}(2, \mathbb{R})$. \square

Lemma 3.7 shows that $T_{E_n^{(m)}}(f)$ is constant with respect to m . Hence we conclude

$$T_{D_n^V} = \sum_{\substack{0 \leq m < \lambda_q |n| \\ (m,n)^T \in V_q}} T_{E_n^{(m)}} = \frac{\varphi_q(n)}{c(q)} \int_{\mathrm{SL}(2, \mathbb{R})} f(gJ_n) d\eta(g).$$

In conclusion, we've now shown that

$$T_{D_n^V} = \frac{\varphi_q(n)}{c(q)} \int_{\mathrm{SL}(2, \mathbb{R})} f(gJ_n) d\eta(g)$$

As well as

$$T_{\mathrm{LD}_{\pm 1}} = \frac{1}{c(q)} \int_{\mathbb{R}^2} f(x, \pm x) dx.$$

Putting these results together with Corollary 3.3, we have now shown Theorem 1.5 holds.

4. Higher moments

We will prove Theorem 1.6 which is the generalization of Theorem 1.5 which corresponds to higher moments of the classical Siegel–Veech Transform on \mathbb{R}^2 .

4.1. Decomposition into orbits. We first decompose $(\mathbb{R}^2)^k$ into $\mathrm{SL}(2, \mathbb{R})$ orbits. Given a point in $(\mathbb{R}^2)^k$, either all the terms are linearly dependent, or there exist two terms in the k -tuples which are linearly dependent.

Lemma 4.1. *The following decomposes $(\mathbb{R}^2)^k$ into disjoint $\mathrm{SL}(2, \mathbb{R})$ orbits:*

$$(\mathbb{R}^2)^k = \left(\bigsqcup_{\lambda} \mathrm{LD}_{\lambda} \right) \sqcup \left(\bigsqcup_{n, \lambda, \alpha, \beta} D_{n, \lambda, \alpha, \beta} \right).$$

In the linearly dependent case,

$$\mathrm{LD}_{\lambda} = \{\lambda x : x \in \mathbb{R}^2\}$$

for $\lambda \in \mathbb{R}^k$ with first nonzero entry (if it exists) given by 1.

In the linearly independent case,

$$D_{n,\lambda,\alpha,\beta} \stackrel{\text{def}}{=} \{(\lambda x, \alpha x + \beta y) : \det(x y) = n\},$$

where $n \in \mathbb{R}_0$ is the determinant of the first nonzero vector with the first linearly independent vector. For $0 \leq j < k$ we have $\lambda \in \mathbb{R}^j$ where the first nonzero entry is 1 and $\alpha, \beta \in \mathbb{R}^{k-j}$ where $\alpha = (0, \alpha_2, \dots, \alpha_{k-j})$ and $\beta = (1, \beta_2, \dots, \beta_{k-j})$.

Proof. We first claim that LD_λ and $D_{n,\lambda,\alpha,\beta}$ can be written as $\text{SL}(2, \mathbb{R})$ orbits. Indeed since $\text{SL}(2, \mathbb{R})$ acts transitively and linearly by matrix multiplication on $\mathbb{R}^2 \setminus \{0\}$ we can write

$$\text{SL}(2, \mathbb{R}) \cdot \lambda \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \text{LD}_\lambda.$$

Similarly since $\text{SL}(2, \mathbb{R})$ acts transitively on determinant n subsets as proved in Lemma 3.1 and linearly on \mathbb{R}^2 , we can write

$$\text{SL}(2, \mathbb{R}) \cdot \left(\lambda \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \alpha \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ n \end{bmatrix} \right) = D_{n,\lambda,\alpha,\beta}.$$

Next we show that the union of the orbits in fact covers all of $(\mathbb{R}^2)^k$. To do this consider a vector $v = (v_1, v_2, \dots, v_k) \in (\mathbb{R}^2)^k$. If all $v_i = 0$ then $v \in \text{LD}_0$. Otherwise there is some first nonzero v_i which we will call x . If $\dim(\text{span}(v_1, \dots, v_k)) = 1$, then every other element will be a linear multiple of x . Hence $v \in \text{LD}_\lambda$ where λ has first nonzero entry is 1 and all remaining entries are real numbers.

If however $\dim(\text{span}(v_1, \dots, v_k)) = 2$, then set y to be the first vector after x which is linearly independent of x . For all v_i which occur after y , v_i can be written as a linear combination of x and y , thus written as $\alpha_i x + \beta_i y$ for some real numbers. Since LD_λ and $D_{n,\lambda,\alpha,\beta}$ are subsets of $(\mathbb{R}^2)^k$ we conclude that

$$(\mathbb{R}^2)^k = \left(\bigcup_{\lambda} \text{LD}_\lambda \right) \cup \left(\bigcup_{n,\lambda,\alpha,\beta} D_{n,\lambda,\alpha,\beta} \right).$$

Finally we finish the proof by proving each of these orbits is distinct. Since all pairs of entries in LD_λ have determinant 0 and $\text{SL}(2, \mathbb{R})$ preserves determinants, we know that the LD_λ and $D_{n,\lambda,\alpha,\beta}$ must be disjoint.

Now suppose that $\text{LD}_\lambda = \text{LD}_{\lambda'}$. Since the first nonzero vector must have a coefficient of 1, if $\lambda_j = 1$ is the first nonzero element, then $\lambda'_j = 1$ as well. Now every vector after the first vector is linearly dependent on the first nonzero vector, so there is a unique coefficient and $\lambda = \lambda'$.

Similarly, since we choose a coefficient of 1 for the first nonzero vector, and a coefficient of 1 for the first vector which is linearly independent, we have a unique representation of the linear combinations $\alpha_i x + \beta_i y$. Hence the $D_{n,\lambda,\alpha,\beta}$ are also all disjoint. This completes the proof. \square

4.2. Reduction to smaller lemmas. Using the notation of Section 3.3 we will rewrite $\int_{Y_q} \hat{f} d\mu$, reducing the proof of Theorem 1.6 to smaller lemmas.

Define

$$D_{n,\lambda,\alpha,\beta}^V = D_{n,\lambda,\alpha,\beta} \cap (V_q)^k$$

and

$$\text{LD}_\lambda^V = \text{LD}_\lambda \cap (V_q)^k.$$

Moreover for $1 \leq m \leq |n|$ with $(m, n)^T \in V_q$ define

$$E_{n,\lambda,\alpha,\beta}^{(m)} = H_q \cdot \left(\lambda \begin{bmatrix} 1 \\ 0 \end{bmatrix}, J_{m,n} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \right)$$

where $\begin{bmatrix} \alpha \\ \beta \end{bmatrix}$ is a $2 \times (k - j)$ matrix and

$$J_{n,m} = \begin{bmatrix} 1 & m \\ 0 & n \end{bmatrix}$$

corresponds to $J_n = J_{n,1}$ in the $k = 2$ case. By Lemma 3.6, we can write

$$D_{n,\lambda,\alpha,\beta} = \bigsqcup_{\substack{0 \leq m < \lambda_q |n| \\ (m,n)^T \in V_q}} E_{n,\lambda,\alpha,\beta}^{(m)}.$$

Lemma 4.2. *We have*

$$\int_{Y_q} \hat{f} d\mu = \sum_{\lambda} T_{\text{LD}_\lambda^V} + \sum_{n \in N_q} \sum_{\substack{0 \leq m < \lambda_q |n| \\ (m,n)^T \in V_q}} \sum_{\lambda,\alpha,\beta} T_{E_{n,\lambda,\alpha,\beta}^{(m)}},$$

where in the linearly dependent case $\lambda \in \mathbb{R}^k$, where the first element of λ must be 1, and any remaining elements of λ must be ± 1 .

In the linearly independent case, given $n \in N_q$, there exists a unique $0 \leq m < \lambda_q |n|$ so that the two vectors lie in the H_q orbit of $J_{m,n}$. Given n and m , we have $\lambda \in \mathbb{R}^j$ with first entry 1 and all remaining elements ± 1 . Moreover $\alpha = (0, \alpha_2, \dots, \alpha_{k-j})$ and $\beta = (1, \beta_2, \dots, \beta_{k-j})$ where α and β satisfy (3) for each $2 \leq i \leq k - j$.

Proof. By Lemma 4.1, given $v \in (V_q)^k$, we must have $v \in \text{LD}_\lambda^V$ or $D_{n,\lambda,\alpha,\beta}^V$ for some λ or $(n, \lambda, \alpha, \beta)$.

If $v \in \text{LD}_\lambda^V$, first note that the zero vector is not in V_q . Hence the first entry in λ must be 1. Since the vectors are in V_q and must be constant multiples of the first vector, all other vectors must be $\pm v_1$. Thus λ must have the specified form.

Moving onto the linearly independent case, if $v \in D_{n,\lambda,\alpha,\beta}^V$, then we can write $v = (\lambda v_1, \alpha v_1 + \beta v_j)$ for some $1 < j \leq k$ where v_j is the first vector after v_1 which is not co-linear with v_1 . Since all the vectors in λv_1 must be in V_q , λ must have first entry 1 and all other entries ± 1 .

Setting $\det(v_1 v_j) = n$ by the definition of N_q we have $n \in N_q$.

Finally we need to determine the criterion for $\alpha = (0, \alpha_2, \dots, \alpha_{k-j})$ and $\beta = (1, \beta_2, \dots, \beta_{k-j})$. So that for all $1 \leq i \leq k-j$ we have $\alpha_i v_1 + \beta_i v_j \in V_q$.

By Lemma 3.6, there exists $\gamma \in H_q$ and $0 \leq m < \lambda_q |n|$ with $(m, n)^T \in V_q$ so that

$$\gamma \cdot (v_1 | v_j) = \begin{bmatrix} 1 & m \\ 0 & n \end{bmatrix}.$$

So we have

$$\gamma \cdot (\alpha_i v_1 + \beta_i v_j) = \begin{bmatrix} \alpha_i + m\beta_i \\ n\beta_i \end{bmatrix}.$$

Since H_q acts transitively on V_q , $\alpha_i v_1 + \beta_i v_j \in V_q$ is equivalent to

$$\begin{bmatrix} 1 & m \\ 0 & n \end{bmatrix} \begin{bmatrix} \alpha_i \\ \beta_i \end{bmatrix} = \begin{bmatrix} \alpha_i + m\beta_i \\ n\beta_i \end{bmatrix} \in V_q.$$

Multiplying by the inverse of $J_{n,m}$ we see α and β must satisfy (3).

So for each n , there exists m with $0 \leq m < \lambda_q |n|$ so that $(m, n)^T \in V_q$. Given this m , we already have the requirement for λ , and the α and β must satisfy (3). This concludes the proof. \square

Now that we have decomposed $\int_{Y_q} \hat{f} d\mu$, the higher moments case is complete once we prove the following two lemmas.

Lemma 4.3. *Given the restrictions of λ in Lemma 4.2,*

$$T_{LD_\lambda^V} = \frac{1}{c(q)} \int_{\mathbb{R}^2} f(\lambda v) dv.$$

Proof. The proof strategy is identical to the strategy in Lemma 3.4. \square

Lemma 4.4. *Given the restrictions of $n, m, \lambda, \alpha, \beta$ in Lemma 4.2'*

$$T_{E_{n,\lambda,\alpha,\beta}^{(m)}} = \frac{1}{c(q)} \int_{\text{SL}(2,\mathbb{R})} f\left(\lambda g \begin{bmatrix} 1 \\ 0 \end{bmatrix}, g J_{n,m} \begin{bmatrix} \alpha \\ \beta \end{bmatrix}\right) d\eta(g).$$

Proof. This proof strategy is identical to the proof in Lemma 3.7. Note we cannot use the change of variables to get equal contribution for each $T_{E_n^{(m)}}$ as in the case of Lemma 3.7 because the criterion for which α and β can occur depends on m . \square

Combining Lemma 4.2 with Lemma 4.3 and Lemma 4.4, we have now concluded the proof of Theorem 1.6.

5. Numerical evidence

This section discusses how to interpret Theorem 1.6 in terms of a counting problem. We will focus on the case $k = 2$, that is Theorem 1.5. The following proposition is from Section 16 of [17].

Proposition 5.1. *For $B(0, R)$ the ball of radius R in \mathbb{R}^2 ,*

$$\lim_{R \rightarrow \infty} \frac{\#\{V_q \cap B(0, R)\}}{\pi R^2} = \frac{1}{c(q)}.$$

We will use the notation

$$\#\{V_q \cap B(0, R)\} \sim \frac{\pi R^2}{c(q)},$$

where

$$f(R) \sim g(R) \iff \lim_{R \rightarrow \infty} f(R)/g(R) = 1.$$

In the case $q = 3$ this can be interpreted as the probability a randomly chosen integer vector is primitive is $\frac{1}{\zeta(2)}$. See Figure 1 and Figure 2 for visualization of the points of V_q for $q = 3, 4, 5$.

To construct the set V_q , we used a Farey tree construction in the first quadrant and then used the 4-fold symmetry of V_q , that is $(a, b) \in V_q$ implies $(a, -b), (-a, b), (-a, -b) \in V_q$. The generalization of the Farey tree construction as found in [10] begins with the vectors

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} \in V_q \quad \text{and} \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} \in V_q.$$

Then for $i = 2, \dots, q - 1$ we add the vectors

$$\begin{bmatrix} a_i \\ b_i \end{bmatrix} = \begin{bmatrix} \lambda_q a_{i-1} - a_{i-2} \\ \lambda_q b_{i-1} - b_{i-2} \end{bmatrix},$$

where

$$\begin{bmatrix} a_1 \\ b_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} a_0 \\ b_0 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}.$$

Iterating this step between each pair of adjacent vectors we obtain the elements of V_q .

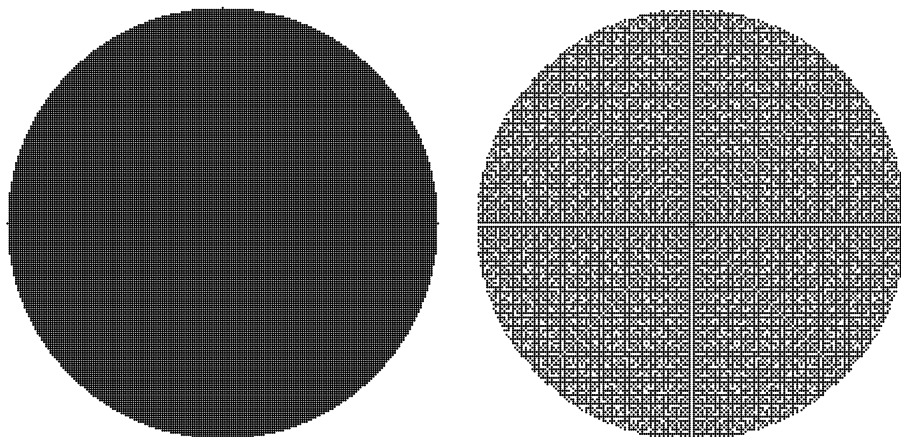


Figure 1. For $R = 100$, the left hand plot includes all integer pairs (a, b) and the right hand plot includes all primitive integers pairs (a, b) with $\gcd(a, b) = 1$ where $a^2 + b^2 \leq 100^2$. This demonstrated the expectations that for large R , the number of points on the left should be approximately $\pi 100^2$, but on the right the number of points should be $\frac{\pi 100^2}{\zeta(2)}$.

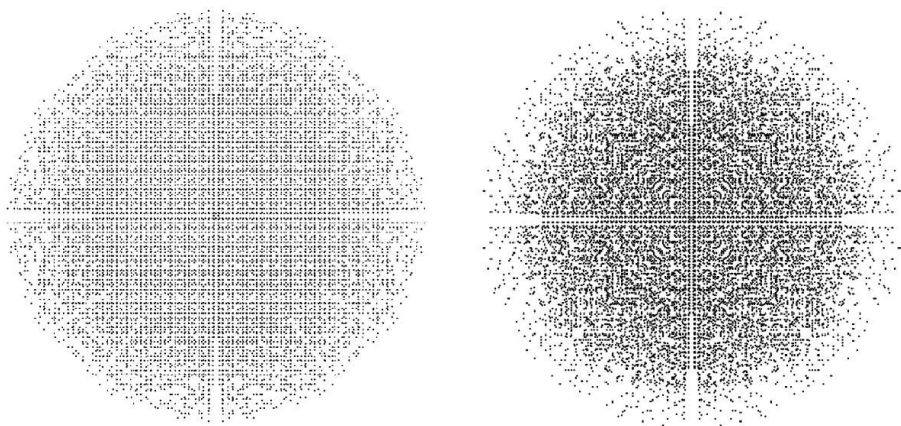


Figure 2. On the left is a plot of the vectors V_4 , and on the right is a plot of the vectors V_5 . These plots were generated using the Farey Tree construction.

Now that we've generated plots for V_q , we can now count pairs of elements in V_q corresponding to the square of the Siegel–Veech transform. Specifically for $f = \chi_{B_{\mathbb{R}^4}(0,R)}$ the characteristic function of the Euclidean ball in \mathbb{R}^4 , we want to understand $T_{D_n^V}(f)$ which will asymptotically grow like the function

$$\text{Count}_q(R, n) \stackrel{\text{def}}{=} \# \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in D_n^V : a^2 + b^2 + c^2 + d^2 \leq R^2 \right\}.$$

Theorem 1.5 states that

$$\frac{\text{Count}_q(R, n)}{\int_{\text{SL}(2, \mathbb{R})} f(gJ_n) d\eta} \sim \frac{\varphi_q(n)}{c(q)},$$

which is not useful for understanding data without more knowledge about the integral $\int_{\text{SL}(2, \mathbb{R})} f(gJ_n) d\eta$.

Newman [12] showed that $\text{Count}_3(R, 1) \sim 6R^2$. In particular combining with Theorem 1.5, we obtain

$$\int_{\text{SL}(2, \mathbb{R})} f(gJ_1) d\eta \sim \pi^2 R^2.$$

Next using the result of Schmidt [14], we can extend this result to the fact that when $q = 3$,

$$\int_{\text{SL}(2, \mathbb{R})} f(gJ_n) d\eta \sim \frac{\pi^2}{n} R^2.$$

Thus we deduce that for any $q \geq 3$,

$$\frac{\text{Count}_q(R, n)}{R^2} \sim \frac{1}{c(q)} \cdot \varphi_q(n) \cdot \frac{\pi^2}{n} = \frac{\varphi_q(n) \cdot \pi^2}{n \cdot c(q)}.$$

Indeed in our numerical experiments we obtained the desired results. In Figure 3, we show the convergence for $k = 1, 2, 3, 4$. Recall D_n^V can be decomposed into $\varphi_q(n)$ orbits $E_n^{(m)}$ where $0 \leq m < \lambda_q |n|$, and on each orbit we were able to verify we had density asymptotic to $\frac{\pi^2}{n \cdot c(q)}$ as desired and graphed in Figure 4. Finally in Figure 5, we provide a visualization for pairs of elements in V_q for $q = 3$ and $q = 5$.

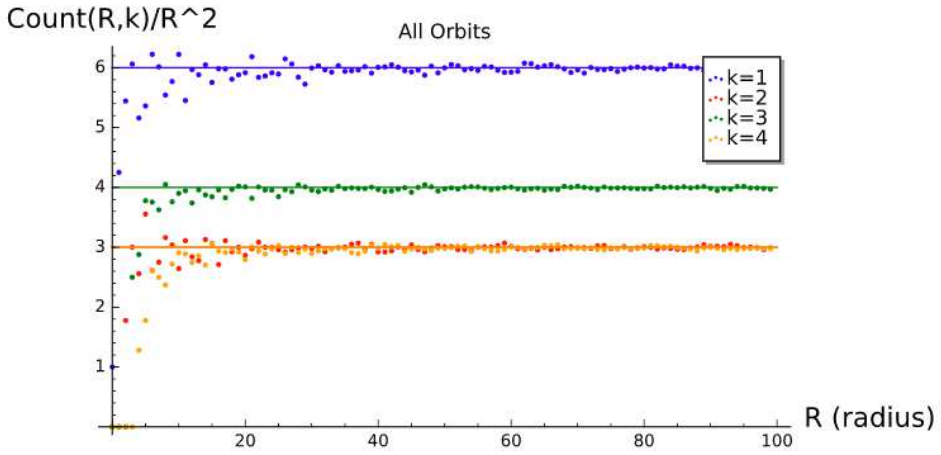


Figure 3. For $q = 3$ this is a plot of R on the horizontal axis, and a plot of $\frac{\text{Count}(R,k)}{R^2}$ on the vertical axis. Notice that $k = 2$ and $k = 4$ both converge to 3 since $\frac{6\varphi(4)}{4} = \frac{6\varphi(2)}{2} = 3$.

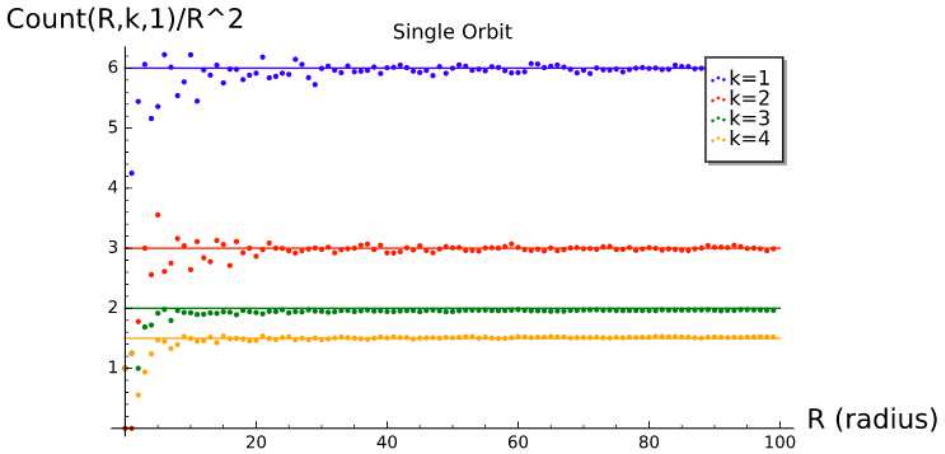


Figure 4. For $q = 3$, a plot of R on the horizontal axis, and a plot of $\frac{\text{Count}(R,n,1)}{R^2}$ on the vertical axis, where $\text{Count}(R,n,1)$ is number of elements within the ball of radius R , which are in the orbit $\text{SL}(2, \mathbb{Z}) \cdot J_n$.

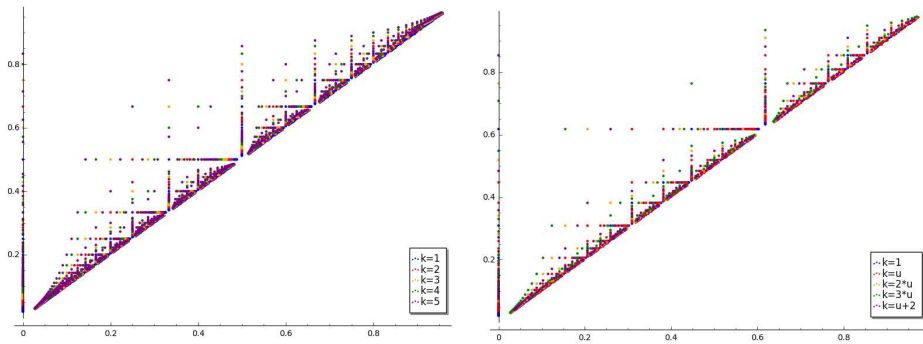


Figure 5. For each element (a, b) in V_q , we can visualize these points by considering a/b , which for simplicity we will consider the pairs with $a/b \in [0, 1]$. These pictures plot the points in D_k^V by placing a point if the point a/b on the x -axis and the point c/d on the y -axis have the corresponding pairs (a, b) and (b, c) in D_k^V . On the left are pairs of primitive vectors, and on the right are pairs of vectors in V_5 where in the legend $u = \phi$, the golden ratio.

References

- [1] J. S. Athreya and Y. Cheung, A Poincaré section for the horocycle flow on the space of lattices. *Int. Math. Res. Not. IMRN* (2014), no. 10, 2643–2690. [Zbl 1351.37145](#) [MR 3214280](#)
- [2] J. S. Athreya, Y. Cheung, and H. Masur, Siegel–Veech transforms are in L^2 . With an appendix by J. S. Athreya and R. Rühr. *J. Mod. Dyn.* **14** (2019), 1–19. [Zbl 1432.32011](#) [MR 3959354](#)
- [3] J. S. Athreya and G. A. Margulis, Logarithm laws for unipotent flows. I. *J. Mod. Dyn.* **3** (2009), no. 3, 359–378. [Zbl 1184.37007](#) [MR 2538473](#)
- [4] I. I. Bouw and M. Möller, Teichmüller curves, triangle groups, and Lyapunov exponents. *Ann. of Math. (2)* **172** (2010), no. 1, 139–185. [Zbl 1203.37049](#) [MR 2680418](#)
- [5] A. Eskin, Counting problems in moduli space. In B. Hasselblatt and A. Katok (eds.), *Handbook of dynamical systems*. Vol. 1B. Elsevier B. V., Amsterdam, 2006, 581–595. [Zbl 1130.37300](#) [MR 2186249](#)
- [6] E. Gutkin and C. Judge, The geometry and arithmetic of translation surfaces with applications to polygonal billiards. *Math. Res. Lett.* **3** (1996), no. 3, 391–403. [Zbl 0865.30060](#) [MR 1397686](#)
- [7] W. P. Hooper, Grid graphs and lattice surfaces. *Int. Math. Res. Not. IMRN* (2013), no. 12, 2657–2698. [Zbl 1333.37047](#) [MR 3071661](#)
- [8] P. Hubert and T. A. Schmidt, An introduction to Veech surfaces. In B. Hasselblatt and A. Katok (eds.), *Handbook of dynamical systems*. Vol. 1B. Elsevier B. V., Amsterdam, 2006, 501–526. [Zbl 1130.37367](#) [MR 2186246](#)

- [9] D. Y. Kleinbock and G. A. Margulis, Logarithm laws for flows on homogeneous spaces. *Invent. Math.* **138** (1999), no. 3, 451–494. [Zbl 0934.22016](#) [MR 1719827](#)
- [10] C. L. Lang and M. L. Lang, Arithmetic and geometry of the Hecke groups. *J. Algebra* **460** (2016), 392–417. [Zbl 1417.11032](#) [MR 3510408](#)
- [11] H. Masur, Ergodic theory of translation surfaces. In B. Hasselblatt and A. Katok (eds.), *Handbook of dynamical systems*. Vol. 1B. Elsevier B. V., Amsterdam, 2006, 527–547. [Zbl 1130.37313](#) [MR 2186247](#)
- [12] M. Newman, Counting modular matrices with specified Euclidean norm. *J. Combin. Theory Ser. A* **47** (1988), no. 1, 145–149. [Zbl 0643.10020](#) [MR 924457](#)
- [13] C. A. Rogers, Mean values over the space of lattices. *Acta Math.* **94** (1955), 249–287. [Zbl 0065.28201](#) [MR 0075243](#)
- [14] W. M. Schmidt, A metrical theorem in geometry of numbers. *Trans. Amer. Math. Soc.* **95** (1960), 516–529. [Zbl 0101.27904](#) [MR 0117222](#)
- [15] C. L. Siegel, A mean value theorem in geometry of numbers. *Ann. of Math. (2)* **46** (1945), 340–347. [Zbl 0063.07011](#) [MR 0012093](#)
- [16] W. A. Veech, Teichmüller curves in moduli space, Eisenstein series and an application to triangular billiards. *Invent. Math.* **97** (1989), no. 3, 553–583. [Zbl 0676.32006](#) [MR 1005006](#)
- [17] W. A. Veech, Siegel measures. *Ann. of Math. (2)* **148** (1998), no. 3, 895–944. [Zbl 0922.22003](#) [MR 1670061](#)

Received January 29, 2019

Samantha Fairchild, Department of Mathematics, University of Washington, Box 354350, Seattle, WA 98195-4350, USA

e-mail: skayf@uw.edu