hexagon:2611.00976v1

LEOPARD SPOTS, CM POINTS, AND CLASS NUMBERS

DANNY CALEGARI

Abstract. Basmajian's identity writes the area of the totally geodesic boundary of a hyperbolic manifold as a sum over its orthogeodesics. Behind it is a tiling of the boundary, up to measure zero, by round disks, one for each orthogeodesic. We study this tiling in arithmetic examples. Cut the Bianchi orbifold $\mathbb{H}^3/\mathrm{PSL}(2, O_k)$ open along its totally geodesic modular surface. We show that the disks are centred at CM points of the modular surface. More generally, consider all orthogeodesics leaving one side of the surface and returning to it, now allowed to cross it. For n satisfying a congruence condition, the number of them with cosh-length n is the Hurwitz class number $H(4(n^2-1)/|D_k|)$. Combined with the counting of common perpendiculars, this gives, for $k \neq \mathbb{Q}(i)$,

$$\textstyle\sum_{1<n\leq Y} H\big(4(n^2-1)/|D_k|\big) \sim \frac{2\pi}{3}\, Y^2/\big(|D_k|^{3/2} L(2, \chi_{D_k})\big),$$

the sum running over the admissible n. This is a three-dimensional analogue of $\sum_{q\leq Q} \varphi(q) \sim 3Q^2/\pi^2$, which expresses Euler's $\zeta(2) = \pi^2/6$. We also explain, heuristically, how Basmajian's identity for the modular orbifold, opened up into an orbifold with geodesic boundary, already contains Euler's identity.

1. Introduction

1.1. **Basmajian's identity.** Let M be a compact hyperbolic manifold whose boundary S is totally geodesic. An *orthogeodesic* is a geodesic arc in M that meets S perpendicularly at both of its endpoints. Basmajian [Bas93] proved that the volume of S is a sum over the orthogeodesics of M of a function of their lengths alone. When M is a surface, so that S is a union of closed geodesics,

$$\mathrm{length}(S) = \sum_\gamma 2 \log \coth \frac{\ell(\gamma)}{2}, \tag{1}$$

and when M is a 3-manifold,

$$\mathrm{area}(S) = \sum_\gamma \frac{4\pi}{e^{2\ell(\gamma)} - 1}. \tag{2}$$

Both sums run over orthogeodesics γ with a chosen direction, so that each arc appears twice, once leaving from each of its endpoints.

The proof is a picture [Bas93, Cal10]. Lift M to the cover in which one boundary component π becomes a single plane. The other boundary components lift to disjoint planes. Each of them casts a *shadow* on π: the set of points of π whose perpendicular geodesic ray meets it. Thurston called these shadows *leopard spots*; they arise in the definition of the skinning map [Cal10]. They are round disks, they are disjoint, and they fill π up to a set of measure zero, because the limit set has measure zero. Each spot belongs to one orthogeodesic, and its size depends only on the length ℓ: its radius r satisfies $\sinh r = 1/\sinh \ell$. Summing the areas of the spots gives (1) and (2). Decomposing the unit tangent sphere at each point of M instead, according to the orthogeodesic that each geodesic arc to the boundary is homotopic to, gives the volume identities of Bridgeman and Kahn [Bri11, BK10, Cal10]. The identities of McShane and

Mirzakhani [McS98, Mir07] decompose a cusp or a boundary geodesic in a similar spirit. There is a survey by Bridgeman and Tan [BT16], and many relatives [Vla15, BT14, He18, He22, BPT].

1.2. **Arithmetic manifolds.** Arithmetic hyperbolic manifolds tend to contain many totally geodesic hypersurfaces. Their orthospectra are arithmetic too: the numbers $\cosh \ell(\gamma)$ lie in a fixed number field, often in $\frac{1}{2}\mathbb{Z}$. (Doan and Le study surfaces whose orthospectrum is integral in this sense [DL].) So in an arithmetic manifold, Basmajian's identity should be an identity of arithmetic. Its left side is a volume, which is usually a special value of a zeta function. Its right side is a sum over orthogeodesics, and one may hope that the orthogeodesics of a given length are counted by an arithmetic function.

This paper carries out that hope in the simplest family of examples, the *Bianchi orbifolds* $M_k = \mathbb{H}^3/\mathrm{PSL}(2, O_k)$, where O_k is the ring of integers of an imaginary quadratic field k. Each M_k contains a copy F of the modular surface $\mathbb{H}^2/\mathrm{PSL}(2,\mathbb{Z})$, as the image of the plane π over the real axis (for $k = \mathbb{Q}(i)$, the image is a quotient of the modular surface by an involution). We cut M_k open along F and look at the leopard spots on F. Before that, we look at the modular surface itself.

1.3. **The modular surface and Euler's identity.** The modular surface has a cusp rather than a geodesic boundary. But it is the limit of a family of orbifolds O_s with one geodesic boundary component, of length $L \to 0$. In §3 we write Basmajian's identity for O_s explicitly. Its terms are indexed by fractions p/q, and each term is L^2/q^2 to leading order. So, as the cusp closes up, the identity turns into a regularization of the divergent sum $\sum_{p/q} 1/q^2$ of the diameters of Ford circles. The rate of divergence is forced to be exactly $1/L$. Locating the cutoff in two ways, by counting fractions and by following the geodesic flow, gives the same answer exactly when $\zeta(2) = \pi^2/6$. This part of the paper is heuristic, and we support it numerically.

1.4. **Bianchi orbifolds: results.** The circles $g(\widehat{\mathbb{R}})$, for $g \in \mathrm{PSL}(2, O_k)$, form the *Schmidt arrangement* of k [Sch75, Sta18b]. They are the circles at infinity of the lifts of F. For $k \neq \mathbb{Q}(\sqrt{-3})$ no two of them cross, so F is embedded in M_k (Proposition 4.1). The leopard spots on F are the disks bounded by the outermost circles of the arrangement in the upper half-plane. For $k = \mathbb{Q}(i)$ they form the Apollonian strip packing (Figure 3).

For a circle c of the arrangement in the upper half-plane, let $n(c) = \cosh \ell$, where ℓ is the distance between π and the hemisphere over c. For a circle of centre $x + iy$ and radius r, $n(c) = y/r$. Let D be the discriminant of k, and let $H(N)$ denote the Hurwitz class number: the number of $\mathrm{SL}(2,\mathbb{Z})$-classes of positive definite binary quadratic forms of discriminant $-N$, with the classes of multiples of $x^2 + y^2$ and $x^2 + xy + y^2$ counted with weights $1/2$ and $1/3$.

Theorem A (Theorem 5.5)**.** *Let k be any imaginary quadratic field and $n > 1$. Count the circles c of the Schmidt arrangement in the upper half-plane with $n(c) = n$, up to the action of* $\mathrm{PSL}(2,\mathbb{Z})$*, each with weight* $1/|\mathrm{Stab}(c)|$*. If n is admissible (that is, $2n \in \mathbb{Z}$, $4(n^2-1)/|D| \in \mathbb{Z}$, and a congruence condition holds; Lemma 5.2), the number is*

$$H\Big(\frac{4(n^2-1)}{|D|}\Big),$$

and otherwise it is 0*. For each such circle, the hyperbolic centre of its disk is the root in $\mathbb{H}^2$ of a positive definite form of discriminant $-4(n^2-1)/|D|$, and every such form arises from exactly one circle.*

In other words, the leopard spots of the arrangement, and all the circles nested inside them, are disks centred at CM points; the radius of a disk records the discriminant of the (not

necessarily primitive) form attached to it. When $|D|$ is a prime or 8, every $n \in \frac{1}{2}\mathbb{Z}$ with $4(n^2-1)/|D| \in \mathbb{Z}$ is admissible; for $k = \mathbb{Q}(i)$, the odd integers are. The remaining n are not lost. They belong to other totally geodesic surfaces F_ξ in M_k, isometric to F. There are 2^{t-1} surfaces F_ξ, including F, where t is the number of primes dividing D: one for each genus of forms of discriminant D. For $k = \mathbb{Q}(i)$ there are two (Theorem 5.7).

Theorem A counts all circles, nested or not. The asymptotic count of common perpendiculars, due to Parkkonen and Paulin [PP17a] (see also [DRS93, EM93, OS13]), then gives:

Theorem B (Corollary 6.1)**.** *For every imaginary quadratic field $k \neq \mathbb{Q}(i)$,*

$$\sum_{\substack{1<n\le Y \\ n \text{ admissible}}} H\Big(\frac{4(n^2-1)}{|D|}\Big) \sim \frac{\operatorname{area}(F)^2}{4\pi \operatorname{vol}(M_k)}\, Y^2 = \frac{2\pi}{3} \cdot \frac{Y^2}{|D|^{3/2} L(2,\chi_D)},$$

and for $k = \mathbb{Q}(i)$, $\sum_{1<n\le Y,\, n \text{ odd}} H(n^2-1) \sim \pi Y^2/(24\,G)$, with G Catalan's constant. Here n runs over $\frac{1}{2}\mathbb{Z}$; it is an integer when D is even.

The value $L(2,\chi_D)$ has no known closed form. Euler's $\zeta(2) = \pi^2/6$ enters through $\operatorname{area}(F) = \pi/3$, and the Dedekind zeta value $\zeta_k(2) = \zeta(2)L(2,\chi_D)$ through Humbert's formula for the volume of M_k. In §6 we explain this constant geometrically: around each point of F the circles of the arrangement are nested, and $\log n$ grows along the nest at the rate $2\operatorname{vol}(M_k)/\operatorname{area}(F)$ on average. We conjecture that this rate holds at almost every point (Conjecture 6.2, Figure 6). This is the three-dimensional analogue of Lévy's constant for continued fractions. Sullivan used the geometry of Bianchi orbifolds in the same way, to study Diophantine approximation by imaginary quadratic numbers and the excursions of geodesics into the cusps [Sul82].

1.5. **Origin and credit.** This project grew out of some older unpublished ideas of mine about the arithmetic meaning of Basmajian's identity, which I discussed with Yan Mary He; see also her paper [He18] on Basmajian-type identities. The main arithmetic investigations reported here were carried out by Claude, an AI model made by Anthropic, working in conversation with me. Claude worked out the analysis of §3, found and proved the results of §§4–6 (several of them in answer to my questions), did all of the computer experiments, and made all of the figures. I, Danny Calegari, consider Claude at minimum to be a full coauthor of this paper. Only the rules of the site on which it is posted bar us from formalizing this in the author byline.

2. Leopard spots

We fix conventions for the picture behind Basmajian's identity, in the upper half-space model. Points of $\mathbb{H}^3$ are (z,t) with $z \in \mathbb{C}$ and $t > 0$. Planes are vertical half-planes and hemispheres orthogonal to $\mathbb{C}$, so a plane is determined by its circle at infinity. Isometries are the Möbius maps $z \mapsto (az+b)/(cz+d)$ of $\mathbb{C} \cup \{\infty\}$, extended to $\mathbb{H}^3$.

Let π be the vertical plane over $\mathbb{R}$, and identify π with the upper half-plane by $(x,t) \mapsto x+it$. The geodesics perpendicular to π are the quarter circles in the planes $x = \text{const}$ that join $(x,0,t) \in \pi$ to the point $x+it$ of $\mathbb{C}$. In other words, they rotate the upper half of $\mathbb{C}$ onto π, about the real axis. So the shadow on π of the hemisphere P over a circle c in the upper half-plane is *the disk bounded by c itself*, read in π (Figure 1). Its hyperbolic centre is the foot of the common perpendicular from π to P. Its hyperbolic radius r satisfies $\sinh r = 1/\sinh \ell$, where ℓ is the length of that perpendicular, and

$$\cosh \ell = n(c) = \frac{\operatorname{Im}(\text{centre of } c)}{\text{radius of } c}. \tag{3}$$

Its area is $2\pi(\cosh r - 1) = 2\pi\big(n/\sqrt{n^2-1} - 1\big) = 4\pi/(e^{2\ell} - 1)$.

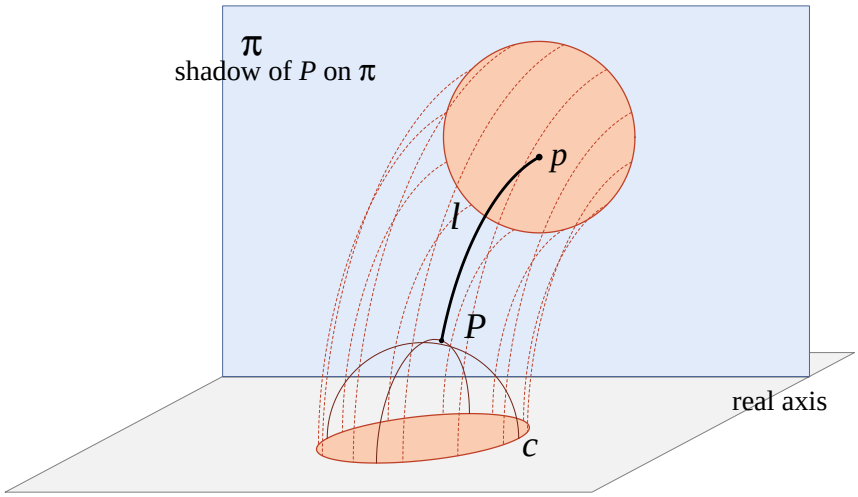


FIGURE 1. The plane π over the real axis, a circle c in $\mathbb{C}$ and the hemisphere P over it. The geodesics perpendicular to π (dashed) rotate $\mathbb{C}$ onto π, so the shadow of P on π is the disk of c rotated into π. The dashed arcs from its boundary to c sweep out the side of the *chimney* over the shadow. The bold arc is the common perpendicular, of length ℓ with $\cosh \ell = n$; its foot p is the hyperbolic centre of the shadow.

If M has finite volume and cusps, two boundary planes may be asymptotic, touching at a cusp. Then the shadow is a horodisk, and its contribution to Basmajian's identity is the area of a cusp neighbourhood of the boundary rather than a function of an orthogeodesic.

3. The modular surface and Euler's identity

This section is a warm-up in dimension two. It is heuristic, and supported by numerical evidence. We open the cusp of the modular orbifold into a short geodesic and watch Basmajian's identity as the geodesic shrinks back.

3.1. **Opening the cusp.** For $s > 1$ let G_s be the group generated by

$$T(z) = sz, \qquad S(z) = \frac{z - x}{z - 1}, \qquad x = s - 1 + \frac{1}{s}.$$

Then $S^2 = (ST)^3 = 1$. So S is the rotation by π about a point p, and ST the rotation by $2\pi/3$ about a point q. Let b be the geodesic through p and q, and a and c the geodesics through p and q making angles $\pi/2$ and $\pi/3$ with b, chosen so that $S = r_a r_b$ and $ST = r_b r_c$, where r_x is the reflection in x. Then $T = r_a r_c$ is hyperbolic, so a and c do not meet, and G_s is the orientation-preserving subgroup of the group generated by r_a, r_b, r_c. The region bounded by a, b and c has angles $\pi/2$ and $\pi/3$, so by Poincaré's polygon theorem (see [Bea83, §9.8]) it is a fundamental domain for that reflection group, which is therefore discrete. So $O_s = \mathbb{H}^2/G_s$ is an orbifold with cone points of orders 2 and 3 and one geodesic boundary component, the image of the axis $(0, \infty)$ of T, of length $L = \log s$. Its area is $\pi/3$. As $s \to 1$, after conjugation in $\mathrm{PSL}(2, \mathbb{R})$, it converges to the modular orbifold $\mathbb{H}^2/\mathrm{PSL}(2, \mathbb{Z})$, with the boundary shrinking into the cusp.

The rule $z \mapsto z+1 \ \rightsquigarrow T$, $z \mapsto -1/z \ \rightsquigarrow S$ identifies G_s with $\mathrm{PSL}(2,\mathbb{Z})$ as an abstract group. Under it, the lifts of the boundary correspond to the fractions p/q, just as the cusps of $\mathrm{PSL}(2,\mathbb{Z})$ do. For $p/q \in [0,1)$ the lift is a semicircle over an interval $[u,v] \subset [1,s]$, and these intervals are the leopard spots on the axis, rotated onto $\mathbb{R}$ (Figure 2). Basmajian's identity (1) becomes

$$\sum_{p/q\in[0,1)} \log \frac{v_{p/q}}{u_{p/q}} = \log s. \tag{4}$$

The endpoints are rational functions of s, so (4) is an infinite product $\prod v/u = s$ of rational functions. At $s=2$, for example, the first factors come from $[1, \frac{3}{2}]$, $[\frac{5}{3}, \frac{9}{5}]$ and $[\frac{11}{7}, \frac{21}{13}]$, for $0/1$, $1/2$ and $1/3$.

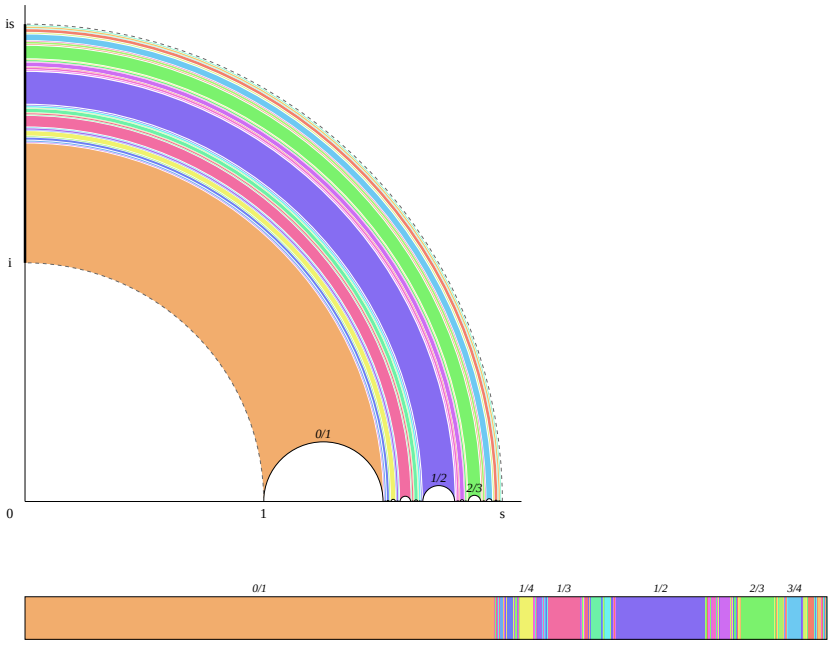


FIGURE 2. The orbifold O_s at $s=2$. Top: one period $1 \le |z| \le s$ of the cover in which the boundary is the axis $(0,\infty)$, tiled by chimneys, one for each fraction $p/q \in [0,1)$. Bottom: the boundary period in hyperbolic length, cut into its leopard spots.

3.2. **Closing the cusp.** As $s \to 1$, every term of (4) tends to 0 while their sum is L. In fact $G_{1/s} = G_s$, so the identity for $-L$ is the identity for L, with the same terms, while the sum is $|L|$. So the identity cannot be expanded term by term at $L=0$. The size of the terms is easy to find.

Proposition 3.1. *For $p/q = 1/m$, as $L \to 0$ with $\lambda = mL$ fixed,*

$$\log \frac{v_{1/m}}{u_{1/m}} = \frac{L^2}{m^2}\Big(\frac{\lambda/2}{\sinh(\lambda/2)}\Big)^2\big(1+O(L^2)\big).$$

Proof. The lift of $1/m$ is $ST^{-m}S(0,\infty)$, with endpoints $u = (x-E)/(1-E)$ and $v = x(1-E)/(1-xE)$, where $E = s^{-m} = e^{-\lambda}$. Put $\delta = x - 1 = 4\sinh^2(L/2)$. Expanding in δ, $\log(v/u) = \delta^2 E/(1-E)^2 + O(\delta^3)$, and $E/(1-E)^2 = 1/(4\sinh^2(\lambda/2))$. □

The factor $(\lambda/2)/\sinh(\lambda/2)$ suppresses a fraction once its partial quotient m exceeds about $1/L$. Numerically, the same is true for every fraction. For $p/q = [0; a_1, \ldots, a_k]$, the term is $(L^2/q^2)\prod_i w(a_i L)$ to leading order, where $w(\lambda) = ((\lambda/2)/\sinh(\lambda/2))^2$. So (4), divided by L^2, reads

$$\sum_{p/q\in[0,1)} \frac{1}{q^2}\, W_{p/q}(L) = \frac{1}{L}, \qquad W_{p/q}(L) \to 1 \text{ as } L \to 0. \tag{5}$$

The numbers $1/q^2$ are the diameters of the Ford circles [For38], and their sum diverges. (Their three-dimensional analogues, the disjoint horoballs of a Bianchi group, are the disjoint spheres that Sullivan used to study approximation of complex numbers by quotients of integers of an imaginary quadratic field [Sul82].) The deformation regularizes it, with divergence rate exactly $1/L$.

3.3. **Where Euler's identity appears.** Which fractions carry the mass in (5)? Weight each p/q by its share $\log(v/u)/L$ of the boundary. Two heuristics predict that, under this weighting, $\log q$ is approximately exponentially distributed with mean $\pi^2/(6L)$.

Heuristic (counting). Since $\sum_q \varphi(q) q^{-s} = \zeta(s-1)/\zeta(s)$ has residue $1/\zeta(2)$ at $s = 2$, the shares L/q^2 add up to $(L/\zeta(2))\log Q$ over the fractions with $q \le Q$. If the cutoff has the memoryless form $q^{-\kappa}$, then (5) forces $\zeta(1+\kappa)/\zeta(2+\kappa) \sim 1/L$, that is $\kappa \sim L/\zeta(2)$. So the mean of $\log q$ is $\zeta(2)/L$.

Heuristic (dynamics). Continued fractions code the geodesic flow on the modular surface [Ser85], and $\log q$ grows like half the flow time. By the factor w, a boundary point is captured by the convergent just before the first excursion of its geodesic above height about $1/L$ in the cusp. The geodesic flow enters that region at rate $L/(\pi A)$, where $A = \pi/3$ is the area. So the mean of $\log q$ is $\pi A/(2L) = \pi^2/(6L)$. In symbolic language: the factor $w(a_j L)$ removes a partial quotient a_j with probability about

$$\sum_a \frac{1 - w(aL)}{a^2 \log 2} \approx \frac{L}{\log 2} \int_0^\infty \frac{1 - w(\lambda)}{\lambda^2}\, d\lambda = \frac{L}{2\log 2},$$

using the Gauss measure $\approx 1/(a^2 \log 2)$ of large partial quotients and $\int_0^\infty (\mu^{-2} - \sinh^{-2}\mu)\, d\mu = 1$; and $\log q_j$ grows like $j\,\pi^2/(12\log 2)$, by Lévy's theorem [Lév29, Khi64].

The two heuristics agree exactly when $\zeta(2) = \frac{\pi}{2}\,\mathrm{area}(\mathbb{H}^2/\mathrm{PSL}(2,\mathbb{Z})) = \pi^2/6$. In this sense Basmajian's identity, read at the cusp, contains Euler's identity. A Monte Carlo test samples boundary points of O_s and finds the leopard spot containing each one exactly. With 4000 samples at each of $L = 0.4, 0.2, 0.1, 0.05, 0.025$, the mean of $\log q$ fits $a/L + b + cL$ with $a = 1.600 \pm 0.024$ (prediction $\pi^2/6 = 1.645$, about two standard errors away). The mean continued-fraction depth fits with $a = 1.396 \pm 0.023$ (prediction $2\log 2 = 1.386$). The distribution of $\log q$ approaches an exponential law.

Remark 3.2. The same reasoning suggests that the limit set of G_s has Hausdorff dimension δ_s with $1 - \delta_s \sim 3L/\pi^2$. This matches Hensley's asymptotic $1 - \dim E_N \sim 6/(\pi^2 N)$ for continued fractions with partial quotients at most N [Hen92] (see also [McM98]), with $N = 2/L$.

4. The modular surface inside a Bianchi orbifold

4.1. **Setup.** Let $k = \mathbb{Q}(\sqrt{-m})$, with $m > 0$ squarefree, be an imaginary quadratic field. Let D be its discriminant ($D = -m$ if $m \equiv 3 \bmod 4$, and $D = -4m$ otherwise), $O_k = \mathbb{Z}[\omega]$ its ring of integers, where $\omega = \sqrt{-m}$ if $D = -4m$ and $\omega = (1+\sqrt{-m})/2$ if $D = -m$ (so $\operatorname{Im}\omega = \sqrt{|D|}/2$), and $\mathfrak{d} = \sqrt{D}\,O_k$. The group $\Gamma = \mathrm{PSL}(2, O_k)$ acts on $\mathbb{H}^3$ with quotient the Bianchi orbifold M_k, of finite volume

$$\mathrm{vol}(M_k) = \frac{|D|^{3/2}}{4\pi^2}\zeta_k(2) = \frac{|D|^{3/2}}{24}L(2, \chi_D)$$

by Humbert's formula [Hum19]; a gap in Humbert's proof was filled by Grunewald and Kühnlein [GK98], and the formula is stated, for example, in [Şen14, §2] (see also [EGM98, MR03]). The stabilizer of the plane π over $\mathbb{R}$ is $\mathrm{PSL}(2, \mathbb{Z})$, except that for $k = \mathbb{Q}(i)$ it also contains $z \mapsto -z$, which exchanges the two sides of π. (An element of Γ preserving $\widehat{\mathbb{R}}$ is λA with A real. If $\det A < 0$ then λ is imaginary, so the entries of λA lie in $\sqrt{-m}\,\mathbb{Z}$, and its determinant lies in $m\mathbb{Z}$; this is possible only for $m = 1$.) So the image of π is a totally geodesic surface F: a two-sided copy of the modular surface, of area $\pi/3$, if $k \neq \mathbb{Q}(i)$, and a one-sided surface of area $\pi/6$, covered twice by the modular surface, if $k = \mathbb{Q}(i)$. Totally geodesic surfaces in Bianchi orbifolds are classified by binary Hermitian forms [MR91, JM96, MR03].

The lifts of F are the planes over the circles $g(\widehat{\mathbb{R}})$, $g \in \Gamma$. This is the Schmidt arrangement $\mathcal{S}_k$, studied by A. L. Schmidt in connection with continued fractions over k [Sch75] and by Stange [Sta18b, Sta18a]. Stange showed that the circles of $\mathcal{S}_k$ correspond to ideal classes in orders of k, the conductor of the order being a fixed multiple of the curvature of the circle, so that they can be counted by class numbers [Sta18b]. That count is by curvature, a distance from the point ∞. Below we count instead by distance from the circle $\widehat{\mathbb{R}}$, and other class numbers appear.

4.2. **Circles as Hermitian matrices.** A circle in $\mathbb{C}\cup\{\infty\}$ is the zero set of $A|z|^2 + 2\operatorname{Re}(\bar{B}z) + C$, for a Hermitian matrix $H = \left(\begin{smallmatrix} A & B \\ \bar{B} & C \end{smallmatrix}\right)$ with $\det H < 0$, which we normalize by $\det H = -1$. If H is the matrix of a circle, then $g^* H g$ is the matrix of its image under g^{-1}. The real line has $H_{\widehat{\mathbb{R}}} = \left(\begin{smallmatrix} 0 & i \\ -i & 0 \end{smallmatrix}\right)$. For a circle in the upper half-plane, with $A > 0$, the quantity n of (3) is $-\operatorname{Im} B$. For $g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}(2, O_k)$, a direct computation gives

$$g^* H_{\widehat{\mathbb{R}}} g = \begin{pmatrix} \sqrt{|D|}\,\alpha & iX \\ -i\bar{X} & \sqrt{|D|}\,\gamma \end{pmatrix}, \qquad X = \bar{a}d - \bar{c}b, \quad \alpha, \gamma \in \mathbb{Z}. \tag{6}$$

Since $ad - bc = 1$ and $x \equiv \bar{x} \bmod \mathfrak{d}$ for $x \in O_k$, we also have $X \equiv 1 \bmod \mathfrak{d}$. Taking real parts, $\operatorname{Re} X \equiv 1 \bmod 2m$ when $D = -4m$, and $2\operatorname{Re} X$ is an integer $\equiv 2 \bmod m$ when $D = -m$. For a circle with matrix (6), $|\operatorname{Re} X|$ is its inversive product with $\widehat{\mathbb{R}}$: the cosine of the angle of intersection if the two circles cross, and cosh of the distance between the two planes if they do not. Two circles cross exactly when the inversive product is less than 1.

Proposition 4.1. *If $k \neq \mathbb{Q}(\sqrt{-3})$, no two circles of $\mathcal{S}_k$ cross. So F is an embedded totally geodesic surface in M_k.*

Proof. The group Γ acts transitively on $\mathcal{S}_k$ and preserves inversive products, so it suffices to show that no circle of $\mathcal{S}_k$ crosses $\widehat{\mathbb{R}}$. When $D = -4m$, $\operatorname{Re} X$ is an integer $\equiv 1 \bmod 2m$, so it is not 0. When $D = -m$, $2\operatorname{Re} X$ is an integer $\equiv 2 \bmod m$; it can lie in $\{-1, 0, 1\}$ only if $m \mid 3$. So $|\operatorname{Re} X| \geq 1$ unless $m = 3$. □

4.3. **The leopard spots.** In this subsection $k \neq \mathbb{Q}(\sqrt{-3})$, so that F is embedded. Cut M_k open along F. The other boundary planes facing the upper side of π are the planes over the *outermost* circles of $\mathcal{S}_k$ in the upper half-plane, those not contained in any other. By §2 their leopard spots are the disks they bound, read in π. These fill π up to measure zero, as in §1.1. Indeed, let R be the region between π and the planes over the outermost circles, and Γ' its stabilizer in Γ, so that R/Γ' is the cut-open orbifold, of finite volume. The limit set Λ' of Γ' contains $\widehat{\mathbb{R}}$ and every outermost circle, since the stabilizer in Γ' of each of these is a lattice in the isometries of its plane. A point ζ of the upper half-plane outside every closed outermost disk lies in Λ' too. Otherwise a small round disk $D \ni \zeta$ would miss Λ', hence miss every outermost circle and lie outside every outermost disk. The half-space over D would then lie in R, and, after shrinking D, meet only finitely many of its translates by Γ', so R/Γ' would have infinite volume. Finally, the convex hull of Λ' lies in R. The faces of R fall into finitely many Γ'-orbits, whose quotients make up the boundary of the cut-open orbifold, copies of F of finite area; so the ε-collars beyond them have finite volume modulo Γ', and every bounded neighbourhood of the convex core of Γ' has finite volume; since the orders of the finite subgroups of Γ' are bounded, Γ' is geometrically finite, in each of the equivalent senses of [Bow93]. By Ahlfors's theorem [Ahl66], Λ', which misses the lower half-plane, has measure zero. The spots tangent to $\mathbb{R}$ are horodisks: the orbit of the half-plane above the line $\operatorname{Im} z = \operatorname{Im} \omega$. (A circle of $\mathcal{S}_k$ tangent to $\widehat{\mathbb{R}}$ has $n = 1$, so the form q_H, defined as in Lemma 5.1, has discriminant 0, and the point of tangency $u/2\alpha$ is rational. Moving it to ∞ by $\mathrm{PSL}(2,\mathbb{Z})$ turns the circle into a line of $\mathcal{S}_k$ parallel to $\mathbb{R}$, that is, one of the lines $\operatorname{Im} z = j \operatorname{Im} \omega$, $j \in \mathbb{Z}$, by the end of the proof of Theorem 5.4. Only $j = 1$ gives an outermost circle.) This orbit contributes area $1/\operatorname{Im}\omega = 2/\sqrt{|D|}$ to the upper side of F, whose area is $\pi/3$. Basmajian's identity therefore reads:

Proposition 4.2. *For $k \neq \mathbb{Q}(\sqrt{-3})$,*

$$\sum_c \frac{1}{|\mathrm{Stab}(c)|} \cdot \frac{1}{e^{2\ell(c)} - 1} = \frac{1}{12} - \frac{1}{2\pi\sqrt{|D|}},$$

where c runs over the outermost circles of $\mathcal{S}_k$ in the upper half-plane, not tangent to $\mathbb{R}$, up to the action of $\mathrm{PSL}(2,\mathbb{Z})$, and $\cosh \ell(c) = n(c)$.

The term $\frac{1}{12} = -\chi(\mathrm{PSL}(2,\mathbb{Z}))/2$ is the cusp-free value [Cal10]; the second term is the cusp correction. Numerically, summing the spots of radius at least $1/640$ and extrapolating gives the right side to within 0.2–1% for $k = \mathbb{Q}(i), \mathbb{Q}(\sqrt{-2}), \mathbb{Q}(\sqrt{-7})$ and $\mathbb{Q}(\sqrt{-11})$.

For $k = \mathbb{Q}(i)$ the outermost circles form the Apollonian strip packing [GLM$^+$03, GLM$^+$05, KO11, Sta18a]. For other k they form other circle packings of the same kind [Sta18a, KN19, FSZ19]. In every case the group of the packing has infinite index in Γ, and so the sum in Proposition 4.2 is over a sparse subset of the circles. The arithmetic appears when we count all the circles of $\mathcal{S}_k$, nested ones included.

5. The orthospectrum is arithmetic

5.1. **From circles to quadratic forms.** Every circle of $\mathcal{S}_k$ in the upper half-plane is recorded by two numbers: how far it is from $\mathbb{R}$, and where its shadow is centred. The next lemma shows that both are arithmetic.

Lemma 5.1. *Let H be as in (6), with $\alpha > 0$, and write $X = -n + \frac{\sqrt{D}}{2}u$ with $u \in \mathbb{Z}$. Suppose $n > 1$, that is, the circle lies in the upper half-plane and does not touch $\mathbb{R}$. Put $q_H(x,y) = \alpha x^2 - uxy + \gamma y^2$ and $N = 4(n^2-1)/|D|$. Then q_H is a positive definite integral*

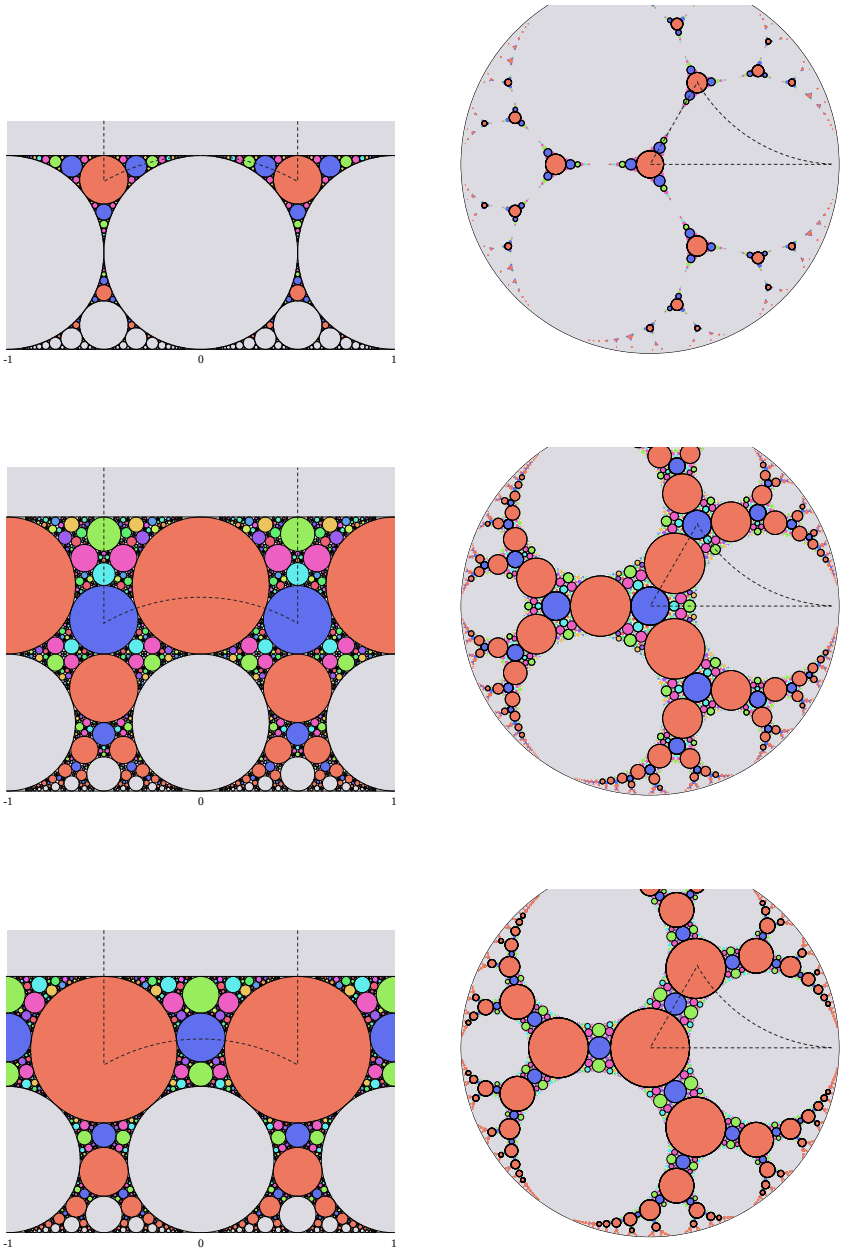


FIGURE 3. Leopard spots on the modular surface $F \subset M_k$, for $k = \mathbb{Q}(i)$, $\mathbb{Q}(\sqrt{-2})$ and $\mathbb{Q}(\sqrt{-7})$ (top to bottom). Left: the upper half-plane, with the standard fundamental domain of $\mathrm{PSL}(2,\mathbb{Z})$ dashed. Right: the same packing in the Poincaré disk, centred at $\rho = e^{i\pi/3}$. Grey: the horodisk $\operatorname{Im} z > \operatorname{Im} \omega$ and its images. Other spots are coloured by $n = \cosh \ell$. The largest are $n = 7$ at ρ (weight $1/3$) for $\mathbb{Q}(i)$; $n = 3$ at i and $n = 5$ at ρ for $\mathbb{Q}(\sqrt{-2})$; and $n = 5/2$ at ρ for $\mathbb{Q}(\sqrt{-7})$.

form of discriminant $-N$. Its root $(u+i\sqrt{N})/2\alpha$ in $\mathbb{H}^2$ is the hyperbolic centre of the shadow of the circle of H, and n is cosh *of the length of the perpendicular. The circle is determined by (n, q_H), and $h \in \mathrm{SL}(2,\mathbb{Z})$ acts by $q \mapsto q \circ h$, fixing n.*

Proof. Write Nm for the norm from k to $\mathbb{Q}$. Then $\det H = -1$ says $\mathrm{Nm}(X) = |D|\alpha\gamma + 1$, that is $n^2 + \frac{|D|}{4}u^2 = |D|\alpha\gamma + 1$, which is $u^2 - 4\alpha\gamma = -N$. The real part of H is $\sqrt{|D|}$ times the matrix of q_H, and its imaginary part is $-n$ times that of $H_{\widehat{\mathbb{R}}}$, which $\mathrm{SL}(2,\mathbb{Z})$ preserves. The centre of the circle is $-iX/\sqrt{|D|}\alpha$ and its radius is $1/\sqrt{|D|}\alpha$, so the hyperbolic centre of its disk is $\frac{u}{2\alpha} + i\frac{\sqrt{n^2-1}}{\sqrt{|D|}\alpha}$, the root of q_H. □

So the disk of every circle of $\mathcal{S}_k$ is centred at a CM point, a point of $\mathbb{H}^2$ that satisfies a quadratic equation over $\mathbb{Q}$. Its CM field is $\mathbb{Q}(\sqrt{D(n^2-1)})$. Moreover $e^\ell = \left(\nu + \sqrt{\nu^2-4}\right)/2$ with $\nu = 2n \in \mathbb{Z}$, a unit of norm 1. Compare Sarnak's count of the closed geodesics of the modular surface: those of trace ν are counted by class numbers of discriminant $\nu^2 - 4$ [Sar82]. Here the orthogeodesics of cosh-length $\nu/2$ will be counted by class numbers of discriminant $-(\nu^2-4)/|D|$.

5.2. **Which forms occur.** Not every (n, q) need come from a circle of $\mathcal{S}_k$. Let Λ be the set of matrices of the shape (6), with $X \in O_k$, $\det H = -1$ and $X \equiv \pm 1 \bmod \mathfrak{d}$. By the computation above, the matrices of the circles of $\mathcal{S}_k$ lie in Λ. Moreover Λ is preserved by $\mathrm{SL}(2, O_k)$. Indeed, let $H \in \Lambda$ and $g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}(2, O_k)$. The diagonal entries of $g^* H g$ are $\sqrt{|D|}(\alpha|a|^2 + \gamma|c|^2) - 2\,\mathrm{Im}(X\bar{a}c)$ and the same with b, d; they lie in $\sqrt{|D|}\,\mathbb{Z}$ because $2\,\mathrm{Im}\, w \in \sqrt{|D|}\,\mathbb{Z}$ for $w \in O_k$. The off-diagonal entry is iX' with

$$X' = X\bar{a}d - \bar{X}\bar{c}b - \sqrt{D}(\alpha\bar{a}b + \gamma\bar{c}d) \in O_k,$$

and $X' \equiv X(\bar{a}d - \bar{c}b) \equiv X \bmod \mathfrak{d}$, since $\bar{X} \equiv X$ and $\bar{a}d - \bar{c}b \equiv ad - bc = 1$.

Lemma 5.2. *Given $n > 1$ with $2n \in \mathbb{Z}$, with $N = 4(n^2-1)/|D|$ a positive integer, and a form q of discriminant $-N$, the number $X = -n + \frac{\sqrt{D}}{2}u$ lies in O_k, and its class modulo $\mathfrak{d}$ depends only on n. Call n* admissible *if this class is ± 1. Explicitly, n is admissible iff $N \equiv 0 \bmod 4$ and $n \equiv \pm 1 \bmod 2m$ (when $D = -4m$), or $2n \equiv \pm 2 \bmod m$ (when $D = -m$). If $|D|$ is a prime, or 8, every such n is admissible; if $D = -4$, the odd ones.*

Proof. If $D = -4m$, then $n \in \mathbb{Z}$: if $2n$ were odd, $4(n^2-1) = (2n)^2 - 4$ would be odd and N not an integer. So $X = -n + u\sqrt{-m}$ and $O_k/\mathfrak{d}$ records $(-n \bmod 2m,\ u \bmod 2)$, with $u \equiv N \bmod 2$ because $u^2 \equiv -N \bmod 4$. If $D = -m$, then $X = (-2n + u\sqrt{-m})/2$, and $2n \equiv u \bmod 2$ holds automatically because $u^2 \equiv -N \bmod 4$ and $mN = (2n)^2 - 4$; the map $(r + s\sqrt{-m})/2 \mapsto r/2$ identifies $O_k/\mathfrak{d}$ with $\mathbb{Z}/m$ and sends X to $-n$. For $|D|$ prime, $N \in \mathbb{Z}$ gives $(2n)^2 \equiv 4$, hence $2n \equiv \pm 2$. For $D = -8$ or -4, $N \in \mathbb{Z}$ (resp. $N \equiv 0 \bmod 4$) forces n odd, and then $n^2 - 1 \equiv 0 \bmod 8$ and $n \equiv \pm 1 \bmod 2m$. □

The heart of the matter is to show that every matrix in Λ comes from a circle of $\mathcal{S}_k$. Every circle in Λ passes through points of k. The key point is that it passes through a point that Γ can move to ∞, *whatever the class number of k.*

Lemma 5.3. *Let $H \in \Lambda$ with $\alpha \neq 0$, choose $\varepsilon = \pm 1$ with $X \equiv \varepsilon \bmod \mathfrak{d}$, and put $a_0 = (X - \varepsilon)/\sqrt{D} \in O_k$. Then $z^* = a_0/\alpha$ lies on the circle of H, and the ideal (a_0, α) of O_k is generated by a rational integer.*

Proof. That z^* lies on the circle is a one-line computation using $\mathrm{Nm}(X) = |D|\alpha\gamma + 1$. Let $F_H = H/\sqrt{|D|}$ and $v = (a_0, \alpha)^T$. A direct computation gives $v^* F_H e_1 = \varepsilon\alpha/\sqrt{D}$ and

$v^*F_H e_2 = -\varepsilon a_0/\sqrt{D}$, so $v^*F_H\, O_k^2 = \mathfrak{d}^{-1}(a_0, \alpha)$. On the other hand $\sqrt{D}\, F_H$ has entries in O_k and determinant 1, so $F_H O_k^2 = \mathfrak{d}^{-1} O_k^2$ and $v^*F_H\, O_k^2 = \mathfrak{d}^{-1}(\bar{a}_0, \alpha)$. Hence the ideal $\mathfrak{a} = (a_0, \alpha)$ equals its conjugate. So it is a rational ideal times a product of distinct ramified primes. It remains to see that no ramified prime $\mathfrak{p}$ occurs to an odd power. Put $\kappa = 2(1+\varepsilon n)/|D|$. Then $a_0 = (u + \varepsilon\kappa\sqrt{D})/2$, $\mathrm{Nm}(a_0) = \alpha\gamma + \kappa$ (so $\kappa \in \mathbb{Z}$), and $4\alpha\gamma - u^2 = \kappa(\kappa|D| - 4)$. For odd p the valuation of a_0 at $\mathfrak{p}$ is $\min(2v_p(u), 2v_p(\kappa) + 1)$. If it were odd and at most $2v_p(\alpha)$, then $v_p(\kappa) < v_p(u), v_p(\alpha)$. That would give $v_p(4\alpha\gamma - u^2) > v_p(\kappa) = v_p(\kappa(\kappa|D| - 4))$, a contradiction. At the prime above 2, when $D = -4m$, write $u = 2u'$, so that $a_0 = u' + \varepsilon\kappa\sqrt{-m}$ and $\alpha\gamma - u'^2 = \kappa(m\kappa - 1)$. If m is even, the same argument applies. If m is odd, then $m \equiv 1 \bmod 4$, so $x + y\sqrt{-m}$ has odd norm if exactly one of the integers x, y is odd, and norm $\equiv 2 \bmod 4$ if both are. Hence the valuation of a_0 is odd exactly when u' and κ have the same 2-adic valuation j, and it is then $2j + 1$. If also $2^{j+1} \mid \alpha$, then $\alpha\gamma - u'^2$ is odd while $\kappa(m\kappa - 1)$ is even (if $j = 0$), or $v_2(\alpha\gamma - u'^2) > j = v_2(\kappa(m\kappa - 1))$ (if $j \geq 1$). Either way we have a contradiction. □

Theorem 5.4. *For every imaginary quadratic field k, the matrices in Λ are exactly the matrices $\pm g^* H_{\widehat{\mathbb{R}}} g$, $g \in \mathrm{SL}(2, O_k)$.*

Proof. Let $H \in \Lambda$. If $\alpha \neq 0$, then by Lemma 5.3, $z^* = a/c$ with $a, c \in O_k$ generating O_k. Complete $\left(\begin{smallmatrix} a & * \\ c & * \end{smallmatrix}\right) \in \mathrm{SL}(2, O_k)$. This moves z^* to ∞, and Λ is preserved, so we may assume that $\alpha = 0$, that is, the circle of H is a line. Then $\mathrm{Nm}(X) = 1$, so X is a unit $\equiv \pm 1 \bmod \mathfrak{d}$. Hence $X = \pm 1$: for $k = \mathbb{Q}(i)$ the units $\pm i$ are not $\equiv \pm 1$ modulo $\mathfrak{d} = (2)$. (For $k = \mathbb{Q}(\sqrt{-3})$, first apply a rotation $z \mapsto \zeta z$, $\zeta^3 = 1$, which multiplies X by a cube root of unity, and change the sign of H if needed.) Now $H = \pm\left(\begin{smallmatrix} 0 & i \\ -i & \sqrt{|D|}\gamma \end{smallmatrix}\right)$, which is the translate of $H_{\widehat{\mathbb{R}}}$ by $z \mapsto z - \gamma\omega$. □

Theorem 5.5. *Theorem A holds.*

Proof. By Theorem 5.4 the circles of $\mathcal{S}_k$ are those of Λ. By Lemma 5.1 they inject into pairs (n, q), equivariantly. By Lemma 5.2 every form occurs when n is admissible, and none occurs when it is not. An element of $\mathrm{PSL}(2, \mathbb{Z})$ fixes a circle iff it fixes its form, so the stabilizers have orders 2, 3 or 1 exactly as in the definition of $H(N)$. □

Figure 4 shows Theorem A at work. Geometrically, Theorem A counts the orthogeodesics of M_k that leave F from the upper side of π and return to F, now allowed to cross F in their interior. Those that do not cross F give the leopard spots of §4.

5.3. **The other genera.** When n is not admissible, the forms of discriminant $-N$ still define circles. Their matrices lie in the larger set Λ_{tot} of matrices (6) with $X \in O_k$ and $\det H = -1$, but have $X \equiv \xi \bmod \mathfrak{d}$ with $\xi^2 = 1$ and $\xi \neq \pm 1$. These circles are the lifts of other totally geodesic surfaces, isometric to F.

Proposition 5.6. *Let $\mathfrak{a}$ be a product of distinct ramified primes, and let $\alpha_\mathfrak{a}$ be the product of the rational primes below them, so that $\mathfrak{a}^2 = (\alpha_\mathfrak{a})$ with $\alpha_\mathfrak{a}$ a positive integer. Choose $M \in M_2(\mathfrak{a})$ with $\det M = \alpha_\mathfrak{a}$. Then $\sigma = M/\sqrt{\alpha_\mathfrak{a}}$ normalizes $\mathrm{SL}(2, O_k)$ and preserves Λ_{tot}. It multiplies the class of X by a square root $\xi_\mathfrak{a}$ of 1 modulo $\mathfrak{d}$. The local component of $\xi_\mathfrak{a}$ at a ramified prime is nontrivial (it is -1 at odd primes) exactly when the prime divides $\mathfrak{a}$.*

Proof. If e, f generate $\mathfrak{a}$, then $e\mathfrak{a} + f\mathfrak{a} = \mathfrak{a}^2 \ni \alpha_\mathfrak{a}$, so $eh_2 - fh_1 = \alpha_\mathfrak{a}$ for some $h_1, h_2 \in \mathfrak{a}$, and $M = \left(\begin{smallmatrix} e & f \\ h_1 & h_2 \end{smallmatrix}\right)$ will do. For $g \in \mathrm{SL}(2, O_k)$, $\sigma g \sigma^{-1} = Mg\,\mathrm{adj}(M)/\alpha_\mathfrak{a}$ has entries in $\mathfrak{a}^2/\alpha_\mathfrak{a} = O_k$, and the same holds for σ^{-1}. Write $\mathrm{adj}\, M = \left(\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right)$, with entries in $\mathfrak{a}$. The matrix of σ applied to the circle of $H \in \Lambda_{\mathrm{tot}}$ is $H' = \mathrm{adj}(M)^* H\, \mathrm{adj}(M)/\alpha_\mathfrak{a}$, of determinant -1. Its diagonal entries

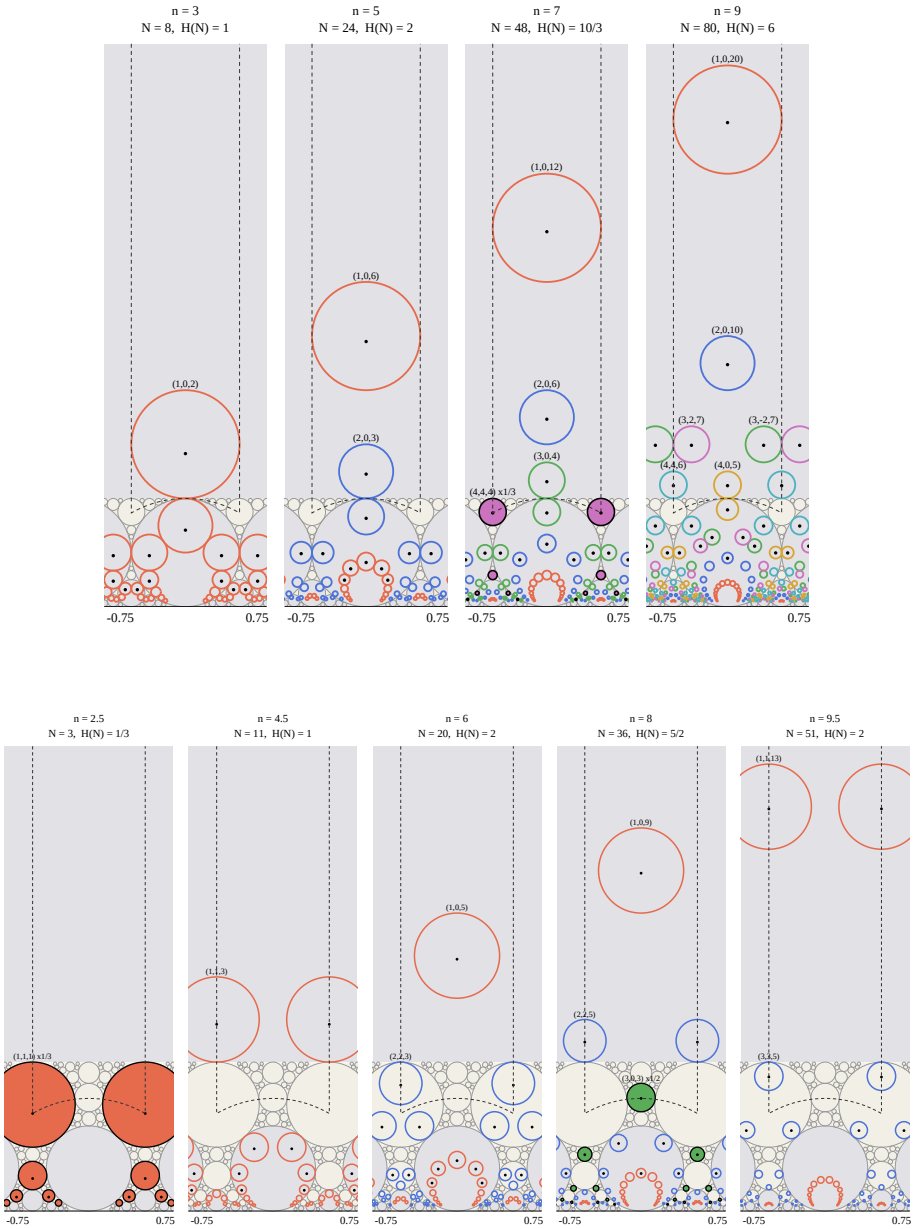


FIGURE 4. Theorem A for $k = \mathbb{Q}(i)$ (top, $N = n^2 - 1$) and $k = \mathbb{Q}(\sqrt{-7})$ (bottom, $N = (4n^2 - 4)/7$). Each panel shows all circles with one value of n, as hyperbolic disks of radius $\operatorname{arcsinh}(1/\sqrt{n^2 - 1})$ about CM points. The disks centred in the fundamental domain are labelled by their reduced forms (a, b, c), and their weighted number is $H(N)$. For $n = 7$ and $k = \mathbb{Q}(i)$: $H(48) = 10/3$, from $(1, 0, 12)$, $(2, 0, 6)$, $(3, 0, 4)$ and $(4, 4, 4)$ with weight $1/3$. Filled disks are leopard spots; outlined ones are nested inside larger circles, often inside the cusp horodisk (grey). Faint background: the leopard-spot packing.

have the form $(\sqrt{|D|}\,\alpha|x|^2 + \sqrt{|D|}\,\gamma|y|^2 + 2\,\mathrm{Re}(iX\bar{x}y))/\alpha_{\mathfrak{a}}$ with $x, y \in \mathfrak{a}$; since $|x|^2, |y|^2 \in \alpha_{\mathfrak{a}}\mathbb{Z}$ and $\bar{x}y \in \alpha_{\mathfrak{a}}O_k$, they lie in $\sqrt{|D|}\,\mathbb{Z}$. Its off-diagonal entry is iX' with

$$X' = \big(X\bar{a}'d' - \bar{X}\bar{c}'b' - \sqrt{D}(\alpha\bar{a}'b' + \gamma\bar{c}'d')\big)/\alpha_{\mathfrak{a}} \in O_k.$$

So $H' \in \Lambda_{\mathrm{tot}}$. Since $\bar{c}'b'/\alpha_{\mathfrak{a}} \in O_k$ and $\bar{X} \equiv X \bmod \mathfrak{d}$, we get $X' \equiv X\xi_{\mathfrak{a}} \bmod \mathfrak{d}$ with $\xi_{\mathfrak{a}} = (\bar{a}'d' - \bar{c}'b')/\alpha_{\mathfrak{a}}$. At a ramified prime $\mathfrak{p} \mid \mathfrak{a}$ with uniformizer $\pi_{\mathfrak{p}}$, this is congruent to $\bar{\pi}_{\mathfrak{p}}/\pi_{\mathfrak{p}}$, which is -1 for odd p (take $\pi_{\mathfrak{p}} = \sqrt{-m}$) and the non-trivial square root of 1 at 2. At $\mathfrak{p} \nmid \mathfrak{a}$, $\alpha_{\mathfrak{a}}$ is a unit at $\mathfrak{p}$ and $\bar{x} \equiv x$ modulo the $\mathfrak{p}$-part of $\mathfrak{d}$, so $\xi_{\mathfrak{a}} \equiv \det(\mathrm{adj}\, M)/\alpha_{\mathfrak{a}} = 1$. □

Theorem 5.7. *Let t be the number of primes dividing D, and R the set of square roots of 1 in $(O_k/\mathfrak{d})^\times$ modulo ± 1; it has 2^{t-1} elements, the number of genera of forms of discriminant D (and 2 elements if $D = -4$). For each $\xi \in R$ the matrices with $X \equiv \xi$ are the lifts of one totally geodesic surface $F_\xi = \sigma(F)$, isometric to F. For every $n > 1$ with $2n \in \mathbb{Z}$ and $N = 4(n^2-1)/|D| \in \mathbb{Z}$, the forms of discriminant $-N$ (whose weighted number is $H(N)$) correspond to the common perpendiculars, of length $\operatorname{arccosh} n$, leaving π on its upper side, from F to the surface F_ξ, where ξ is the class of n.*

Proof. Choose $\mathfrak{a}$ with $\xi_{\mathfrak{a}} = \xi$, by Proposition 5.6; the local components show that every ξ occurs. Then σ carries $\Lambda = \Gamma \cdot H_{\widehat{\mathbb{R}}}$ into the class ξ part of Λ_{tot}, and σ^{-1}, which is of the same kind, carries it back since $\xi^2 = 1$, and since it normalizes Γ it induces an isometry of M_k carrying F to F_ξ. The rest is Lemma 5.1. □

When $\mathfrak{a}$ is not principal, F_ξ runs out into a cusp of M_k of class $[\mathfrak{a}]$. For $k = \mathbb{Q}(\sqrt{-5})$, the even n belong to the surface through the cusp at $(1+\sqrt{-5})/2$.

6. Counting, and Euler's identity in dimension three

6.1. **The count.** In dimension two, the fractions p/q index the common perpendiculars from a cusp neighbourhood of the modular surface to itself. Their count $\sum_{q \le Q} \varphi(q) \sim 3Q^2/\pi^2$ is Euler's $\zeta(2) = \pi^2/6$ in disguise, and it is a special case of a general theorem: the number of common perpendiculars of length at most t between two finite-volume totally geodesic or horospherical submanifolds of a hyperbolic d-manifold is asymptotic to an explicit constant times $e^{(d-1)t}$ [PP17a, PP17b, BAPP19]. Applying the same theorem to $F \subset M_k$, and counting the common perpendiculars by Theorem A, gives the three-dimensional analogue.

Corollary 6.1. *For every imaginary quadratic field k and every $\xi \in R$, with n running over $\frac{1}{2}\mathbb{Z}$,*

$$\sum_{\substack{1<n\le Y \\ n \text{ of class } \xi}} H\Big(\frac{4(n^2-1)}{|D|}\Big) \sim c_k Y^2, \qquad c_k = \frac{\mathrm{area}(F)\ \mathrm{area}(F^+)}{4\pi\,\mathrm{vol}(M_k)}.$$

Here F^+ is the upper side of F, of area $\pi/3$. So $c_k = 2\pi/(3|D|^{3/2}L(2,\chi_D))$ when $k \neq \mathbb{Q}(i)$, where F is two-sided of area $\pi/3$, and $c_{\mathbb{Q}(i)} = \pi/(24G)$, where F is one-sided of area $\pi/6$. In particular the sum over all $n \in \frac{1}{2}\mathbb{Z}$ with $4(n^2-1)/|D| \in \mathbb{Z}$ is $\sim |R|\, c_k Y^2$.

Proof. By Theorems 5.5 and 5.7, the sum counts the common perpendiculars from F^+ to F_ξ with cosh-length at most Y, with orbifold weights. Parkkonen and Paulin [PP17a, Cor. 21] (we use the numbering of the arXiv version, arXiv:1305.1332v4) show that the number of common perpendiculars of length at most t between finite-volume totally geodesic suborbifolds A^- and A^+ of a hyperbolic 3-orbifold M is asymptotic to

$$\frac{\mathrm{area}(A^-)\,\mathrm{area}(A^+)}{8\pi\,\mathrm{vol}(M)}\,e^{2t}$$

when no nontrivial element of the group fixes $A^{\pm}$ pointwise, as is the case here. Their count runs over the lifts of A^+ disjoint from a fixed lift π of A^-, modulo the stabilizer of π, with orbifold weights; so here it counts the circles of class ξ in both half-planes. Here $e^{2t} \sim 4Y^2$. For $k \neq \mathbb{Q}(i)$, F is two-sided with $\mathrm{area}(F) = \pi/3$, the stabilizer of π is $\mathrm{PSL}(2,\mathbb{Z})$ (§4.1), and $z \mapsto \bar{z}$ matches the circles in the two half-planes, so their count is twice ours. For $k = \mathbb{Q}(i)$, F is one-sided, of area $\pi/6$, and the stabilizer of π exchanges the two half-planes, so their count is ours. Finally $\mathrm{vol}(M_k) = |D|^{3/2} L(2, \chi_D)/24$. □

Numerically, the ratio of the sum to $c_k Y^2$ is 1.000 at $Y = 801$ for $\mathbb{Q}(\sqrt{-2})$, and within about 3% for $\mathbb{Q}(i)$, $\mathbb{Q}(\sqrt{-3})$, $\mathbb{Q}(\sqrt{-5})$, $\mathbb{Q}(\sqrt{-7})$ and $\mathbb{Q}(\sqrt{-15})$ at $Y \approx 400$–800, class by class. The ratio oscillates by a few percent rather than converging monotonically. Parkkonen and Paulin also give a power-saving error term in the arithmetic case [PP17a, Cor. 21].

6.2. **A geometric reading: Basmajian at every depth.** The constant c_k also comes out of Basmajian's identity, used at every level of nesting. This is the three-dimensional analogue of the dynamical heuristic of §3. In this subsection $k \neq \mathbb{Q}(\sqrt{-3})$, so that the circles of $\mathcal{S}_k$ do not cross (Proposition 4.1).

Fix a point z of π. The circles of $\mathcal{S}_k$ whose disks contain z form a nested chain $c_1(z) \supset c_2(z) \supset \cdots$. They are exactly the lifts of F crossed by the geodesic ray perpendicular to π at z. Write $c_j(z) = g(\widehat{\mathbb{R}})$ with $g \in \Gamma$. Then g^{-1} carries the circles nested directly inside $c_j(z)$ to the outermost circles in a half-plane, whose disks tile that half-plane up to measure zero (§4.3). So the disks of the circles nested directly inside $c_j(z)$ tile the disk of $c_j(z)$, and the circles of each depth together tile F^+ (Figure 5). Hence the total area of the disks with $1 < n \leq Y$ is $\mathrm{area}(F^+)$ times the average number of nesting levels with $1 < n \leq Y$. (The circles with $n = 1$ are the horodisks; only finitely many of them contain a given point.)

Now a typical geodesic in M_k crosses F at average rate $\mathrm{area}(F)/(2\,\mathrm{vol}(M_k))$, by the Crofton formula. If the perpendicular rays are typical, then $\log n_j = \log n(c_j(z))$ grows at the rate $2\,\mathrm{vol}(M_k)/\,\mathrm{area}(F)$, and the disks with $1 < n \leq Y$ have total area about

$$\frac{\mathrm{area}(F^+)\,\mathrm{area}(F)}{2\,\mathrm{vol}(M_k)} \log Y.$$

On the other hand, by Theorem A that area is

$$\sum_{\substack{1<n\leq Y \\ n \text{ admissible}}} H(N) \cdot 2\pi\Big(\frac{n}{\sqrt{n^2-1}} - 1\Big) \approx \sum_{\substack{1<n\leq Y \\ n \text{ admissible}}} \frac{\pi H(N)}{n^2},$$

and Corollary 6.1 makes it $2\pi c_k \log Y$. The two agree. The averaged statement follows from Corollary 6.1. The statement for almost every z is still a conjecture:

Conjecture 6.2. *Let $k \neq \mathbb{Q}(\sqrt{-3})$. For almost every $z \in \pi$,*

$$\lim_{j\to\infty} \frac{1}{j} \log n(c_j(z)) = \frac{2\,\mathrm{vol}(M_k)}{\mathrm{area}(F)}.$$

The right side is about 1.166, 1.917 and 1.698 for $\mathbb{Q}(i)$, $\mathbb{Q}(\sqrt{-2})$ and $\mathbb{Q}(\sqrt{-7})$. By self-similarity, the chains can be computed exactly. Reduce z into the fundamental domain of $\mathrm{PSL}(2,\mathbb{Z})$, find the leopard spot $g(\widehat{\mathbb{R}})$ containing it, pull its inside back to the upper half-plane by g^{-1}, and repeat. Figure 6 shows the result for 1500 random points. The measured slopes lie 3–4% below the predictions. We expect this to be the bias from discarding the chains that enter spots too small to resolve, which would have had large increments. The chains also show parabolic runs, where n grows in arithmetic progression, as in Figure 5; these are the analogues

of large partial quotients. They are excursions of the perpendicular ray into a cusp of M_k. For a typical geodesic of M_k the depth of such excursions is governed by Sullivan's logarithm law [Sul82]; the perpendicular rays from F form a family of measure zero, so statements about typical geodesics do not apply to them directly. This is also the difficulty in Conjecture 6.2.

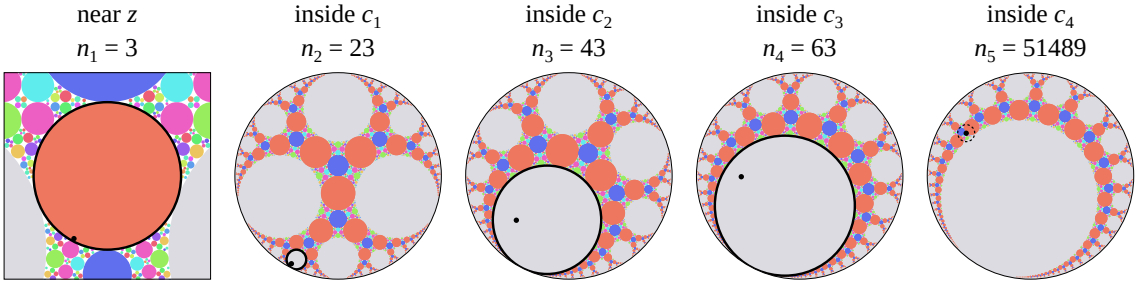


FIGURE 5. The nested chain around $z = 0.42 + 0.38i$ for $k = \mathbb{Q}(\sqrt{-2})$, with $n = 3, 23, 43, 63, 51489$. Each disk panel shows the inside of c_j, rescaled; by self-similarity it is a Möbius copy of the leopard-spot packing. In each panel the bold circle is the next circle of the chain, labelled by its value of n, and the dot is z. Panels 2–4 show a parabolic run: c_2, c_3, c_4 are internally tangent at one point, and n grows by 20 each time.

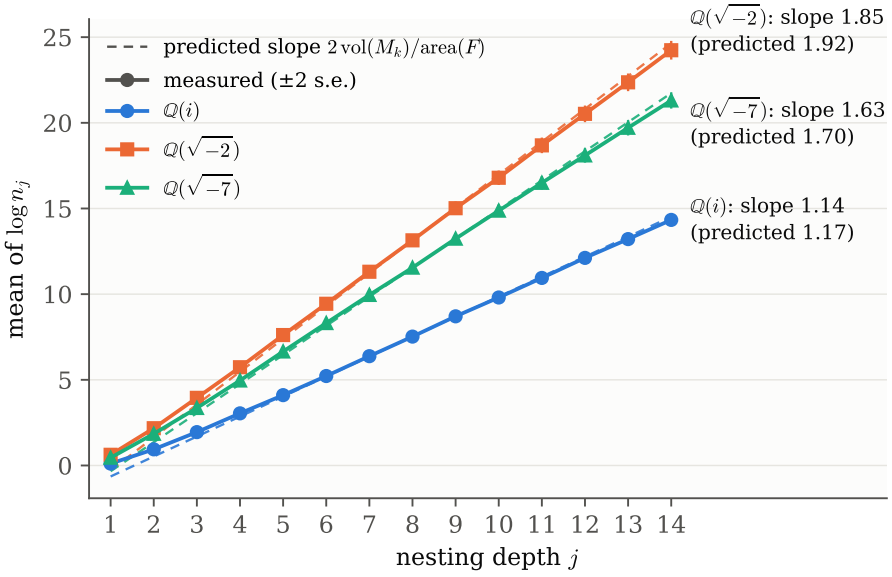


FIGURE 6. Mean of $\log n_j$ against the nesting depth j, over 1500 random points of F^+ for each field (chains resolved down to radius $1/1280$). Dashed: the slopes $2\,\mathrm{vol}(M_k)/\,\mathrm{area}(F)$ of Conjecture 6.2, drawn through the data at $j = 8$.

7. Questions

(1) Basmajian's identity (Proposition 4.2) runs over the outermost circles, a thin set, while Theorem A counts all circles. Each nesting level tiles F^+. Is there an exact identity, rather than an asymptotic one, relating the two counts?
(2) The Bridgeman–Kahn identity [BK10] for M_k cut along F expresses $\mathrm{vol}(M_k)$, and so $\zeta_k(2)$, as a sum over the same orthospectrum, with correction terms from the cusps. What are those corrections, and what does the identity say about class numbers?
(3) Let $Z(s) = \sum_n H(4(n^2-1)/|D|)\, n^{-s}$, over admissible n. Then $(s-2)Z(s) \to 2c_k$ as $s \to 2^+$. The generating function of Hurwitz class numbers is Zagier's mock modular form of weight 3/2 [Zag75], so this series is a shifted convolution. Does Z continue meromorphically, and what are its poles?
(4) Theorem A is a statement about one arithmetic group acting on circles, and the proof is the arithmetic of binary Hermitian forms. Arithmetic hyperbolic manifolds of simplest type contain totally geodesic hyperplanes for every rational vector of positive norm, and their orthospectra are representation numbers of quadratic forms. Is there an analogous formula in every dimension, presumably an instance of Siegel's mass formula?
(5) Prove Conjecture 6.2, perhaps by effective equidistribution of the rays normal to F, using the spectral gap of M_k.

References

[Ahl66] Lars V. Ahlfors, *Fundamental polyhedrons and limit point sets of Kleinian groups*, Proc. Nat. Acad. Sci. U.S.A. **55** (1966), no. 2, 251–254.

[BAPP19] Anne Broise-Alamichel, Jouni Parkkonen, and Frédéric Paulin, *Equidistribution and counting under equilibrium states in negative curvature and trees. Applications to non-Archimedean Diophantine approximation*, Progress in Mathematics, vol. 329, Birkhäuser/Springer, 2019.

[Bas93] Ara Basmajian, *The orthogonal spectrum of a hyperbolic manifold*, Amer. J. Math. **115** (1993), no. 5, 1139–1159.

[Bea83] Alan F. Beardon, *The geometry of discrete groups*, Graduate Texts in Mathematics, vol. 91, Springer, New York, 1983.

[BK10] Martin Bridgeman and Jeremy Kahn, *Hyperbolic volume of manifolds with geodesic boundary and orthospectra*, Geom. Funct. Anal. **20** (2010), 1210–1230.

[Bow93] B. H. Bowditch, *Geometrical finiteness for hyperbolic groups*, J. Funct. Anal. **113** (1993), 245–317.

[BPT] Ara Basmajian, Hugo Parlier, and Ser Peow Tan, *Prime orthogeodesics, concave cores and families of identities on hyperbolic surfaces*, arXiv:2006.04872.

[Bri11] Martin Bridgeman, *Orthospectra of geodesic laminations and dilogarithm identities on moduli space*, Geom. Topol. **15** (2011), 707–733.

[BT14] Martin Bridgeman and Ser Peow Tan, *Moments of the boundary hitting function for the geodesic flow on a hyperbolic manifold*, Geom. Topol. **18** (2014), no. 1, 491–520.

[BT16] ———, *Identities on hyperbolic manifolds*, Handbook of Teichmüller theory. Vol. V, IRMA Lect. Math. Theor. Phys., vol. 26, Eur. Math. Soc., Zürich, 2016, pp. 19–53.

[Cal10] Danny Calegari, *Chimneys, leopard spots, and the identities of Basmajian and Bridgeman*, Algebr. Geom. Topol. **10** (2010), no. 3, 1857–1863.

[DL] Nhat Minh Doan and Khanh Le, *Some arithmetic aspects of ortho-integral surfaces*, arXiv:2504.09403.

[DRS93] W. Duke, Z. Rudnick, and P. Sarnak, *Density of integer points on affine homogeneous varieties*, Duke Math. J. **71** (1993), 143–179.

[EGM98] Jürgen Elstrodt, Fritz Grunewald, and Jens Mennicke, *Groups acting on hyperbolic space. Harmonic analysis and number theory*, Springer Monographs in Mathematics, Springer, 1998.

[EM93] Alex Eskin and Curt McMullen, *Mixing, counting, and equidistribution in Lie groups*, Duke Math. J. **71** (1993), 181–209.

[For38] L. R. Ford, *Fractions*, Amer. Math. Monthly **45** (1938), 586–601.

[FSZ19] Elena Fuchs, Katherine E. Stange, and Xin Zhang, *Local-global principles in circle packings*, Compos. Math. **155** (2019), no. 6, 1118–1170.

[GK98] Fritz Grunewald and Stefan Kühnlein, *On the proof of Humbert's volume formula*, Manuscripta Math. **95** (1998), no. 4, 431–436.

[GLM+03] Ronald L. Graham, Jeffrey C. Lagarias, Colin L. Mallows, Allan R. Wilks, and Catherine H. Yan, *Apollonian circle packings: number theory*, J. Number Theory **100** (2003), 1–45.

[GLM+05] ———, *Apollonian circle packings: geometry and group theory I. The Apollonian group*, Discrete Comput. Geom. **34** (2005), 547–585.

[He18] Yan Mary He, *Basmajian-type identities and hausdorff dimension of limit sets*, Ergodic Theory Dynam. Systems **38** (2018), 2224–2244.

[He22] ———, *Identities for hyperconvex Anosov representations*, J. Topol. Anal. **16** (2022), no. 4, 617–640.

[Hen92] Doug Hensley, *Continued fraction Cantor sets, Hausdorff dimension, and functional analysis*, J. Number Theory **40** (1992), 336–358.

[Hum19] Georges Humbert, *Sur la mesure des classes d'Hermite de discriminant donné dans un corps quadratique imaginaire, et sur certains volumes non euclidiens*, C. R. Acad. Sci. Paris **169** (1919), 448–454.

[JM96] Donald G. James and Colin Maclachlan, *Fuchsian subgroups of Bianchi groups*, Trans. Amer. Math. Soc. **348** (1996), no. 5, 1989–2002.

[Khi64] A. Ya. Khinchin, *Continued fractions*, University of Chicago Press, Chicago, 1964.

[KN19] Alex Kontorovich and Kei Nakamura, *Geometry and arithmetic of crystallographic sphere packings*, Proc. Natl. Acad. Sci. USA **116** (2019), no. 2, 436–441.

[KO11] Alex Kontorovich and Hee Oh, *Apollonian circle packings and closed horospheres on hyperbolic 3-manifolds*, J. Amer. Math. Soc. **24** (2011), 603–648.

[Lév29] Paul Lévy, *Sur les lois de probabilité dont dépendent les quotients complets et incomplets d'une fraction continue*, Bull. Soc. Math. France **57** (1929), 178–194.

[McM98] Curtis T. McMullen, *Hausdorff dimension and conformal dynamics III: computation of dimension*, Amer. J. Math. **120** (1998), 691–721.

[McS98] Greg McShane, *Simple geodesics and a series constant over Teichmüller space*, Invent. Math. **132** (1998), 607–632.

[Mir07] Maryam Mirzakhani, *Simple geodesics and Weil–Petersson volumes of moduli spaces of bordered Riemann surfaces*, Invent. Math. **167** (2007), 179–222.

[MR91] Colin Maclachlan and Alan W. Reid, *Parametrizing Fuchsian subgroups of the Bianchi groups*, Canad. J. Math. **43** (1991), 158–181.

[MR03] ———, *The arithmetic of hyperbolic 3-manifolds*, Graduate Texts in Mathematics, vol. 219, Springer, 2003.

[OS13] Hee Oh and Nimish A. Shah, *Equidistribution and counting for orbits of geometrically finite hyperbolic groups*, J. Amer. Math. Soc. **26** (2013), no. 2, 511–562.

[PP17a] Jouni Parkkonen and Frédéric Paulin, *Counting common perpendicular arcs in negative curvature*, Ergodic Theory Dynam. Systems **37** (2017), no. 3, 900–938.

[PP17b] ———, *A survey of some arithmetic applications of ergodic theory in negative curvature*, Ergodic theory and negative curvature, Lecture Notes in Math., vol. 2164, Springer, 2017, pp. 293–326.

[Sar82] Peter Sarnak, *Class numbers of indefinite binary quadratic forms*, J. Number Theory **15** (1982), 229–247.

[Sch75] Asmus L. Schmidt, *Diophantine approximation of complex numbers*, Acta Math. **134** (1975), 1–85.

[Şen14] Mehmet Haluk Şengün, *Arithmetic aspects of Bianchi groups*, Computations with Modular Forms (Gebhard Böckle and Gabor Wiese, eds.), Contrib. Math. Comput. Sci., vol. 6, Springer, Cham, 2014, pp. 279–315.

[Ser85] Caroline Series, *The modular surface and continued fractions*, J. London Math. Soc. (2) **31** (1985), 69–80.

[Sta18a] Katherine E. Stange, *The Apollonian structure of Bianchi groups*, Trans. Amer. Math. Soc. **370** (2018), no. 9, 6169–6219.

[Sta18b] ———, *Visualising the arithmetic of imaginary quadratic fields*, Int. Math. Res. Not. IMRN (2018), no. 12, 3908–3938.

[Sul82] Dennis Sullivan, *Disjoint spheres, approximation by imaginary quadratic numbers, and the logarithm law for geodesics*, Acta Math. **149** (1982), no. 3-4, 215–237.

[Vla15] Nicholas G. Vlamis, *Moments of a length function on the boundary of a hyperbolic manifold*, Algebr. Geom. Topol. **15** (2015), no. 4, 1909–1929.

[Zag75] Don Zagier, *Nombres de classes et formes modulaires de poids* 3/2, C. R. Acad. Sci. Paris Sér. A-B **281** (1975), A883–A886.

Department of Mathematics, University of Chicago, Chicago, Illinois, 60637