On the distribution of Farey fractions and hyperbolic lattice points

Martin N. Huxley and Anatoly A. Zhigljavsky

Cardiff University

Abstract

We derive an asymptotic formula for the number of pairs of consecutive fractions a'/q' and a/q in the Farey sequence of order Q such that $a/q, q/Q$, and $(Q-q')/q)$ lie each in prescribed subintervals of the interval $[0,1]$. We deduce the leading term in the asymptotic formula for 'the hyperbolic lattice point problem' for the modular group $\mathrm{PSL}(2,\mathbb{Z})$, the number of images of a given point under the action of the group in a given circle in the hyperbolic plane.

Keywords: Farey fractions, hyperbolic geometry, lattice point problem, Selberg trace formula

1991 AMS Subject Classification: 11 B57, 11 F72.

1 Introduction

The Farey sequence of $\mathcal{F}_Q$ of order Q is the finite set of rational numbers a/q with $1 \le q \le Q,\ 0 \le a \le q-1$ and the highest common factor $(a,q)=1$. The extended Farey sequence of order Q is defined by dropping the condition $0 \le a \le q-1$. It consists of all translates of the Farey sequence by integers. In both cases the Farey sequence is ordered by the relation $x < y$ on real numbers. However to calculate one uses the lexicographic ordering, firstly by q, and secondly by a. Some well-known properties are that if a'/q' and a/q

are consecutive in $\mathcal{F}_Q$, then

$$aq' - a'q = 1, \quad \begin{pmatrix} a & a' \\ q & q' \end{pmatrix} \in \mathrm{SL}(2, \mathbb{Z}) \tag{1}$$

where $\mathrm{SL}(2, \mathbb{Z})$ is the group of integer matrices of determinant one. In $\mathcal{F}_Q$ the fraction a'/q' determines its successor a/q uniquely through (1) and the restrictions $q' \leq Q$, $q + q' > Q$. Each pair of denominators q, q' with highest common factor $(q, q') = 1$ and with $q + q' > Q$ occurs exactly once as denominators of consecutive fractions a'/q' and a/q in $\mathcal{F}_Q$.

So a pair of consecutive denominators in $\mathcal{F}_Q$ corresponds to a primitive integer vector in the plane. Proposition 1 follows immediately by Möbius inversion.

Proposition 1. *Let $\gamma_1, \gamma_2, \delta_1, \delta_2$ be real numbers with*

$$0 \leq \gamma_1 < \gamma_2 \leq 1, \;\; 0 \leq \delta_1 < \delta_2 \leq 1$$

and let $Q \geq 2$. Then the number of pairs of consecutive fractions a'/q', a/q in $\mathcal{F}_Q$ with

$$\delta_1 < \frac{q}{Q} \leq \delta_2, \;\; \gamma_1 < \frac{Q - q'}{q} \leq \gamma_2 \tag{2}$$

is

$$\frac{3}{\pi^2}(\gamma_2 - \gamma_1)(\delta_2^2 - \delta_1^2)Q^2 \; + \; O(Q \log Q) \tag{3}$$

uniformly in $\gamma_1, \gamma_2, \delta_1, \delta_2$.

Proposition 1, or results of the same type, has been applied in questions of uniform distribution, see Kargaev and Zhigljavsky [4], Zhigljavsky and Aliev [11]. It is natural to consider also the position of a/q in the interval $[0, 1)$.

Theorem 1. *Let $\beta_1, \beta_2; \gamma_1, \gamma_2; \delta_1, \delta_2$ be real numbers with*

$$0 \leq \beta_1 < \beta_2 \leq 1, \;\; 0 \leq \gamma_1 < \gamma_2 \leq 1, \;\; 0 \leq \delta_1 < \delta_2 \leq 1,$$

and let $Q \geq 2$. Then the number of pairs of consecutive fractions a'/q', a/q in $\mathcal{F}_Q$ with

$$\beta_1 < \frac{a}{q} \leq \beta_2, \;\; \delta_1 < \frac{q}{Q} \leq \delta_2, \;\; \gamma_1 < \frac{Q - q'}{q} \leq \gamma_2 \tag{4}$$

is

$$\frac{3}{\pi^2}(\beta_2-\beta_1)(\gamma_2-\gamma_1)(\delta_2^2-\delta_1^2)Q^2 \;+\; O(Q^{3/2}\log Q)\,. \tag{5}$$

The matrix group $\mathrm{SL}(2,\mathbb{Z})$ acts as rigid motions in hyperbolic geometry. The action is faithful when we factor out $\pm I$, where I is the identity matrix. The quotient is called the modular group $\mathrm{PSL}(2,\mathbb{Z})$.

Theorem 2. *Let z and z' be points in the hyperbolic plane, realised as the complex upper half-plane, with the modular group Γ acting by $z \longrightarrow (az+b)/(cz+d)$. Then the number of images of z under Γ in a circle with centre z', radius r, in the hyperbolic metric, counted according to multiplicity, is asymptotically*

$$3e^r + O\left(e^{\frac{(3+\theta)r}{4}+\epsilon r}\right)$$

as $r \to \infty$, for any $\epsilon > 0$, where θ is an exponent such that the Kloosterman sums $K(a,b;q)$ with highest common factor $(a,b,q)=1$ satisfy

$$K(a,b;q) = O(q^{\theta+\epsilon}) \tag{6}$$

for any $\epsilon > 0$, as $q \to \infty$.

The asymptotic formula of Theorem 2, with a better error term $O\left(\exp(\frac{3}{4}r)\right)$, is a well-known consequence of the Selberg trace formula, see Patterson [5], and the proof of the trace formula is based on the image-counting function, see Hejhal [2]. A technical difficulty is to construct the spectral Eisenstein series. For the modular group the Eisenstein series is essentially an Epstein zeta function, corresponding to counting primitive lattice points inside ellipses on the Euclidean plane. We sketch the proof of Theorem 2, which shows that the hyperbolic lattice point counting can also be done number theoretically. For the exponent θ in Theorem 2, we note that the estimate (6) with $\theta = \frac{2}{3}$ can be proved quite simply, see Davenport [1] and Salie [6, 7]. Of course Weil's bound [10] with $\theta = \frac{1}{2}$ is more difficult than the Selberg trace formula; the exponent $\theta = \frac{1}{2}$ can also be obtained by Stepanov's method, see Schmidt [8], akin to transcendence theory.

2 Proof of Theorem 1.

We recall standard number-theoretical notation. We use (m,n) for the highest common factor of m and n, $d(n), \omega(n), \phi(n), \mu(n)$ for the number of divisors of n, the number of distinct prime factors of n, the number of residue classes a mod n with $(a,n)=1$ (Euler's function), and the Möbius function. For real t, $e(t)=\exp(2\pi i t)$ is the complex exponential function, $[t]$ is the integer part of t and $\rho(t)=[t]-t+\frac{1}{2}$ is the row-of-teeth function. In sums $\sum^*$ modulo q, $*$ means that the variable (a say) has $(a,q)=1$, and $\bar{a}$ denotes the residue class d mod q with $ad\equiv 1$, which exists when $(a,q)=1$. The Kloosterman sum is

$$K(g,h;q)=\sum_{a \bmod q}^{*} e\left(\frac{ga+h\bar{a}}{q}\right).$$

We note that if a'/q', a/q are consecutive in $\mathcal{F}_Q$, then by (1) $q'\equiv\bar{a}$ (mod q). For fixed q, the number of solutions of

$$\beta_1<\frac{a}{q}\le\beta_2,\ \ \gamma_1<\frac{Q-\bar{a}}{q}\le\gamma_2 \tag{7}$$

is

$$\sum_{a \bmod q}^{*}\left(\left[\beta_2-\frac{a}{q}\right]-\left[\beta_1-\frac{a}{q}\right]\right)\left(\left[\gamma_2-\frac{Q-\bar{a}}{q}\right]-\left[\gamma_1-\frac{Q-\bar{a}}{q}\right]\right). \tag{8}$$

We have

$$\left[\beta_2-\frac{a}{q}\right]-\left[\beta_1-\frac{a}{q}\right]=\beta_2-\beta_1+\rho\left(\beta_2-\frac{a}{q}\right)-\rho\left(\beta_1-\frac{a}{q}\right) \tag{9}$$

and similarly for the second factor.

The main term is

$$(\beta_2-\beta_1)(\gamma_2-\gamma_1)\phi(q).$$

Since

$$\begin{aligned}\sum_{\delta_1 Q<q\le\delta_2 Q}\phi(q) &= \sum_{\substack{d,e:\\ \delta_1 Q<de\le\delta_2 Q}}\mu(d)e\\ &=\sum_{d\le Q}\mu(d)\left(\frac{\delta_2^2Q^2}{2d^2}-\frac{\delta_1^2Q^2}{2d^2}+O\left(\frac{Q}{d}\right)\right)\\ &=\frac{1}{2}(\delta_2^2-\delta_1^2)Q^2\left(\frac{6}{\pi^2}+O\left(\frac{1}{Q}\right)\right)+O(Q\log Q)\\ &=\frac{3}{\pi^2}(\delta_2^2-\delta_1^2)Q^2+O(Q\log Q)\end{aligned} \tag{10}$$

we obtain the main term in Theorem 1.

When we substitute (9) and its analogue involving γ_1 and γ_2 into (8), one cross-product term is

$$\Big(\gamma_2 - \gamma_1\Big) \sum_{a \bmod q}^{*} \left(\rho\Big(\beta_2 - \frac{a}{q}\Big) - \rho\Big(\beta_1 - \frac{a}{q}\Big)\right)$$

$$= \Big(\gamma_2 - \gamma_1\Big) \sum_{d|q} \mu(d) \sum_{\substack{a \bmod q \\ a \equiv 0 \, (\bmod d)}} \left(\rho\Big(\beta_2 - \frac{a}{q}\Big) - \rho\Big(\beta_1 - \frac{a}{q}\Big)\right)$$

$$= \Big(\gamma_2 - \gamma_1\Big) \sum_{d|q} \mu(d) \sum_{b \bmod q/d} \left(\rho\Big(\beta_2 - \frac{b}{q/d}\Big) - \rho\Big(\beta_1 - \frac{b}{q/d}\Big)\right)$$

$$= \Big(\gamma_2 - \gamma_1\Big) \sum_{d|q} \mu(d) \left(\rho\Big(\frac{q\beta_2}{d}\Big) - \rho\Big(\frac{q\beta_1}{d}\Big)\right) = O(d(q)) \,,$$

where we have used the functional equations for $\mu(d)$,

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{if } \ d = 1, \\ 0 & \text{if } \ d \geq 2, \end{cases}$$

and for $\rho(t)$,

$$\sum_{b \bmod n} \rho\left(t + \frac{b}{n}\right) = \rho(nt) \,.$$

We have

$$\sum_{\gamma_1 Q < q \leq \gamma_2 Q} d(q) \leq \sum_{q=1}^{Q} \sum_{d|q} 1 \leq \sum_{d \leq Q} \frac{Q}{d} \leq Q(\log Q + 1) \leq 3Q \log Q \,,$$

so the cross-product term gives an error term of the same size as in (10). Similarly, the other cross-product term is

$$(\beta_2 - \beta_1) \sum_{d|q} \mu(d) \left(\rho\Big(\frac{q\gamma_2 - Q}{d}\Big)\right) - \rho\left(\frac{q\gamma_1 - Q}{d}\right)\Big) = O\left(Q \log Q\right) \,.$$

The main error terms are four terms of the type

$$\pm \sum_{a \bmod q}^{*} \rho\Big(\beta - \frac{a}{q}\Big) \rho\Big(\alpha - \frac{\bar{a}}{q}\Big) \tag{11}$$

where $\beta = \beta_1$ or β_2 and $\alpha = \gamma_1 - Q/q$ or $\gamma_2 - Q/q$.

For each integer $H \geq 2$ there are approximations

$$\sum_{h=-H}^{H} b(h)e(ht) \leq \rho(t) \leq \sum_{h=-H}^{H} B(h)e(ht) \tag{12}$$

(see Vaaler [9], section 8 or Huxley [3], section 5.3), with

$$b(0),\ B(0) = O\left(\frac{1}{H}\right),$$

and for $h \neq 0$

$$b(h),\ B(h) = O\left(\frac{1}{|h|}\right).$$

We use these to estimate the sum (11) from above and below. So we estimate

$$\sum_{a \bmod q}^{*} \sum_{g=-H}^{H} c(g)e\left(g\beta - \frac{ag}{q}\right) \sum_{h=-H}^{H} C(h)e\left(h\alpha - \frac{\bar{a}h}{q}\right)$$

$$= \sum_{g=-H}^{H} c(g)e(g\beta) \sum_{h=-H}^{H} C(h)e(h\alpha)K(-g,-h;q)\,, \tag{13}$$

where $c(n), C(n)$ are each either the set of coefficients $b(n)$ or $B(n)$, and $K(-g,-h;q)$ is a Kloosterman sum, equal to $K(g,h;q)$ by symmetry.

Let $e = (g,h,q)$ be the highest common factor of g, h, and q. The well-known bound

$$|K(g,h;q)| \leq 2^{\omega(q/e)}\sqrt{eq} \leq d(q)\sqrt{eq},$$

incorporates the theorem of Weil [10] and earlier work of Salie [6] . Using this, we estimate (13) as follows:

$$O\left(\frac{q}{H^2} + \frac{1}{H}\sum_{e|q} d(q)\sqrt{eq} \sum_{\substack{h=1 \\ h\equiv 0 \,(\bmod e)}}^{H} \frac{1}{h} + \sum_{e|q} d(q)\sqrt{eq} \sum_{\substack{g=1 \\ g\equiv 0 \,(\bmod e)}}^{H} \frac{1}{g} \sum_{\substack{h=1 \\ h\equiv 0 \,(\bmod e)}}^{H} \frac{1}{h}\right)$$

$$= O\left(\frac{q}{H^2} + \frac{1}{H}\sum_{\substack{e|q \\ e\leq H}} d(q)\sqrt{\frac{q}{e}}\log H + \sum_{\substack{e|q \\ e\leq H}} \frac{d(q)\sqrt{eq}}{e^2}\log^2 H\right) = O\left(d(q)\sqrt{q}\log^2 Q\right),$$

when we pick $H = Q$. Finally we have

$$\sum_{q=1}^{Q} d(q)\sqrt{q}\log^2 Q \le \sqrt{Q}\log^2 Q \sum_{q=1}^{Q} d(q) = O\left(Q^{3/2}\log^3 Q\right)$$

which completes the proof of Theorem 1.

3 Theorem 2: Sketch of the Proof

Each element of the modular group is represented by a pair of matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $\begin{pmatrix} -a & -b \\ -c & -d \end{pmatrix}$. We choose the sign so that $c > 0$ if $|c| \ge |d|$, $d > 0$ if $|d| > |c|$. The two columns correspond to two fractions in the Farey sequence $\mathcal{F}_Q$ with $Q = \max(|c|, |d|)$.

The point $u + iv$ in the upper half plane goes, under the transformation $z \longrightarrow (az + b)/(cz + d)$, to $(H + iv)/G$ where

$$G(c,d) = (cu + d)^2 + c^2v^2, \quad H(a,b,c,d) = ac(u^2 + v^2) + (ad + bc)u + bd\,.$$

For points z, z', the hyperbolic distance D from z to z' is given by

$$4\mathrm{sh}^2\frac{D}{2} = \frac{|z - z'|^2}{\mathrm{Im}z\,\mathrm{Im}z'}$$

Let $R = 2\mathrm{sh}(r/2)$ where r is the radius in Theorem 2. Let $z = u + iv$, $z' = w + it$. Then in Theorem 2 we want to count representative matrices with

$$(H - Gw)^2 + (v - tG)^2 \le tvGR^2 \tag{14}$$

We can suppose by symmetry that $z = u + iv$, $z' = w + it$ are in the fundamental domain of the modular group, with $-\frac{1}{2} \le \mathrm{Re}z \le \frac{1}{2}$, $|z| \ge 1$, and that $t \ge v > 0$. Matrices with $c = d$ have $c = d = 1$, $b = a - 1$; there are $O(R)$ of these satisfying (14). Matrices with $d = -c$ have $c = 1$, $d = -1$, $b = -1 - a$. There are again $O(R)$ of these satisfying (14). The constants implied in these estimates $O(R)$ and in subsequent estimates depend on t and v.

In the case $c > |d|$ we eliminate b using the identity

$$cH - aG = -cu - d$$

to show that the number of solutions of (14) with $c > |d|$ lies between the numbers of solutions with $c > |d|$ of two inequalities of the form

$$\left(\left(\frac{a}{c} - w\right) + t^2\right) G \leq tvQ^2 \tag{15}$$

with Q taking two values Q_1, Q_2 of the form $R + O(1)$. Similarly in the case $d > |c|$ we eliminate a using the identity

$$dH - bG = c(u^2 + v^2) + du$$

to show that the number of solutions of (14) with $d > |c|$ lies between the numbers of solutions with $d > |c|$ of two inequalities of the form

$$\left(\left(\frac{b}{c} - w\right) + t^2\right) G \leq tvQ^2 \tag{16}$$

We can count the solutions of (15) or (16) using Theorem 1. There is a complication that a/c in (15) and b/c in (16) lie in the extended Farey sequence, and we use integer shifts to bring them into $[0, 1)$. It is as simple to argue directly. We sketch the calculation. The fraction a/c in (15), or b/c in (16), lies in a range

$$\alpha_1(G) \leq \frac{a}{c} \leq \alpha_2(G)\,. \tag{17}$$

For c, d fixed, the number of possible $a \equiv \bar{d}$ in (17) is

$$\alpha_2(G) - \alpha_1(G) + \rho\left(\alpha_2(G) - \frac{\bar{d}}{c}\right) + \rho\left(\frac{\bar{d}}{c} - \alpha_1(G)\right)$$

The main term now involves summing $\alpha_2(G) - \alpha_1(G)$ over primitive integer vectors (c, d) with $|d| < c$ and $G \leq vQ^2/t$ from (16). This is a weighted average of the number of primitive integer vectors in a fixed sector of an ellipse of fixed shape but variable size. The case $d > |c|$ similarly gives the sums of $\alpha_2(G) - \alpha_1(G)$ over primitive integer vectors (c, d) in a different sector. The two sectors fit together to form half the ellipse cut by a diameter. The Eisenstein series at the point $u + iv$ is also constructed by considering the primitive integer vectors in the semi-ellipse as the size varies. The powers of v and t in the main term cancel. The numerical factor 3 arises as $6/\pi^2$ from the condition $(c, d) = 1$ as in Theorem 1, times π from the weight

$\alpha_2(G) - \alpha_1(G)$, times $\pi/2$ from the area of the semi-ellipse. The main term thus simplifies to

$$3Q^2 + O(Q^{3/2} \log Q)$$

and $e^r = Q^2 + O(Q)$.

For the error term in the case $c > |d|$ we have a sum

$$\sum_{|d|<c}^{*} \rho \left(\alpha(G) - \frac{\bar{d}}{c} \right)$$

where $\alpha = \alpha_1$ or α_2, $G = G(c,d)$. The approximations (12) lead to sums

$$\sum_{|d|<c}^{*} e \left(h\alpha(G) - \frac{\bar{d}h}{c} \right) = O \left((V+1) \max_{I} \left| \sum_{d \in I}^{*} e \left(-\frac{\bar{d}h}{c} \right) \right| \right) ,$$

where V is the total variation of Re $e(h\alpha(G(c,x))$ plus the total variation of Im $e(h\alpha(G(c,x))$, both taken on the interval $x \in J = [1-c, c-1]$, so that

$$V = O \left(\frac{|h|Q}{c} \sqrt{\frac{t}{v}} \right)$$

and I runs through subintervals of J.

After taking out the highest common factor (c,h), we express the sum over I as a linear combination of Kloosterman sums which we estimate using (6). The trivial contribution of the term $h = 0$ in (12) is $O(c/H)$, and that of the terms $|h| = 1, \ldots, H$ in (12) is

$$O\left(Hc^{-1+\theta+\epsilon}\right) .$$

Choosing H to equalise these two contributions and summing over c gives the dominant error term in Theorem 2.

References

[1] Davenport H. (1933) On certain exponential sums, *Journal für Math.* 169, pp. 158-176.

[2] Hejhal D.A. (1976,1983) *The Selberg trace formula for* $PSL[2,\mathbb{R}]$, Lecture Notes in Mathematics, v. 548 (volume 1) and v. 1001 (volume 2). Berlin: Springer-Verlag, 516pp.

[3] Huxley M.N. (1996) *Area, lattice points, and exponential sums,* Oxford, New York: Oxford University Press.

[4] Kargaev P.P. and Zhigljavsky A.A. (1997) Asymptotic distribution of the distance function to the Farey points, *Journal of Number Theory,* 65, pp. 130 - 149.

[5] Patterson S.J. (1975) A lattice point problem in hyperbolic space, *Mathematika,* 22, pp. 81-88.

[6] Salie H. (1931) Über die Kloostermanschen Summen, *Math. Zeitschrift,* 34, pp. 91-109.

[7] Salie H. (1932) Zur Abschätzung der Fourierkoeffizienten ganzer Modulformen, *Math. Zeitschrift,* 36, pp. 263-278.

[8] Schmidt W. M. (1976) *Equations over Finite Fields: An Elementary Approach.* Berlin: Springer-Verlag.

[9] Vaaler J.D. (1985) Some extremal functions in Fourier analysis, *Bulletin of the Americal Mathematical Society,* 12, pp. 183-216.

[10] Weil A. (1948) On some exponential sums, *Proc. Nat. Acad. Sci. USA*, 34, pp. 204-207.

[11] Zhigljavsky A.A. and Aliev I. (1999) Weyl Sequences: Asymptotic Distributions of the Partition Lengths, *Acta Arithmetica,* LXXXVIII, pp. 351-361.

Address:

School of Mathematics, Cardiff University, Cardiff, CF24 4YH, UK
e-mails: Huxley@cardiff.ac.uk, ZhiglajvskyAA@cardiff.ac.uk