ACTA ET COMMENTATIONES UNIVERSITATIS TARTUENSIS DE MATHEMATICA
Volume 9, 2005

Asymptotic formulas for a class of sums over consecutive Farey fractions

ALEXANDRU ZAHARESCU

ABSTRACT. We provide asymptotic formulas, as $Q \to \infty$, for a class of sums over consecutive Farey fractions of order Q, which are defined in terms of a given quadratic form.

1. Introduction

A question on the distribution of Farey fractions which came up in the work of Galway [8] on the problem of enumerating primes in an interval with the use of a dissected sieve, is the following.

Denote by $\mathcal{F}_Q$ the Farey sequence of order Q, that is the set of rational numbers a/q in the interval $[0, 1]$ with a, q relatively prime and $1 \le q \le Q$. Given a 2×2 symmetric matrix A and its associated quadratic form Q_A, the problem asks to find an asymptotic formula, as $Q \to \infty$, for a sum of the form

$$S(Q) = \sum_{\substack{a/q \in \mathcal{F}_Q \\ a/q < M_A}} \frac{1}{F^3\left(\frac{a+a'}{q+q'}\right) qq'(q+q')} \tag{1.1}$$

where $0 < M_A \le 1$ is a constant depending on A, $F(t) = \sqrt{P(t)}$, with $P(t)$ a polynomial of degree 2 defined in terms of the quadratic form Q_A, and $\frac{a}{q} < \frac{a'}{q'}$ are consecutive elements of $\mathcal{F}_Q$.

Atkin and Bernstein [1] introduce an algorithm that computes the prime numbers up to N using $O(N/\log\log N)$ additions and $N^{1/2+o(1)}$ bits of memory. The algorithm enumerates representations of integers by certain binary quadratic forms.

Starting from the method of Atkin and Bernstein, and inspired by the classical work of Voronoi [11] on the Dirichlet divisor problem and of Sierpiński [10] on the circle problem, Galway [7], [8] introduces a Farey partition in the

Received November 15, 2003.
2000 *Mathematics Subject Classification.* 11B57.
Key words and phrases. Farey fractions, Kloosterman sums, binary quadratic forms.

algorithm of Atkin and Bernstein. The resulting dissected sieve algorithm uses somewhat more operations but only $N^{1/3+o(1)}$ bits of memory.

Galway [8] discusses the efficiency of his new algorithm, and he points out that this depends on the number of scanlines enumerated by the algorithm. This in turn depends on the behavior of the sum $S(Q)$ from (1.1) for large Q. In the case when the quadratic form Q_A is given by $Q_A(X,Y) = X^2 + Y^2$, $M_A = 1$ and $P(t) = t^2 + 1$, he found numerically that for $Q = 10$, 100, 1000, 10000 and 100000, the corresponding values of the quantity $QS(Q)/2$ are 0.29094833, 0.29778011, 0.29785085, 0.29795720 and 0.29796200 respectively. This led him to conjecture that $QS(Q)$ approaches a certain nonzero finite limit as $Q \to \infty$.

In a series of papers [2], [3], [4] and [5] various aspects of the distribution of Farey fractions have been investigated, and in particular suitable methods to deal with sums over consecutive Farey fractions have been devised. In the next section we describe such a method by which one indeed obtains an asymptotic result of the form

$$S(Q) \sim \frac{C(A)}{Q} \tag{1.2}$$

as $Q \to \infty$, where $C(A)$ is a nonzero constant depending on the given matrix A only. As we shall see below, the constant $C(A)$ can be computed in closed form. For instance, in the case when $Q_A(X,Y) = X^2 + Y^2$, $M_A = 1$ and $P(t) = t^2 + 1$, we find that $C(A) = (6\sqrt{2}\log 2)/\pi^2$. Here $C(A)/2 = 0.2979627...$, which confirms the trend apparent in the numerical data above.

As was pointed out by Atkin and Bernstein [1], and also by Galway [8], one can vary the binary quadratic forms used in their algorithms. In order to increase the efficiency of the algorithm, in that case one could perhaps use the asymptotic result (1.2), with the simple formula for $C(A)$ provided below, in order to estimate in advance, for various combinations of binary quadratic forms, the number of scanlines to be enumerated by the dissected sieve algorithm.

2. An asymptotic formula for $S(Q)$

In this section we explain how one can produce asymptotic results for sums of the form

$$S(Q) = \sum_{\substack{a/q \in \mathcal{F}_Q \\ a/q < M_A}} \frac{1}{F^3\left(\frac{a+a'}{q+q'}\right) qq'(q+q')} \tag{2.1}$$

where $F(t) = \sqrt{P(t)}$, $P(t)$ polynomial of degree 2, and $\frac{a}{q} < \frac{a'}{q'}$ are consecutive elements of the Farey series of order Q.

Since our main goal in this paper is to obtain (1.2), in particular to understand the constant $C(A)$, and in order to simplify the presentation, in what follows we will ignore the error terms that appear in our estimates, and we will only focus on the main term. Suffices to say that if one carefully takes into account the contribution of the error terms, one ends up with an asymptotic formula as in (1.2), with a very good upper bound for the error term. This is due to the powerful Weil type bounds for Kloosterman sums (see [12], [9], [6]), which are ultimately used in the proof of the result from [5] quoted below.

We now proceed to estimate the sum $S(Q)$. The first step in dealing with such a sum is to use the well known equality $a'q - aq' = 1$, which is valid for any two consecutive Farey fractions, in order to obtain a good approximation for a' as a function of a, q and q'. More precisely, we replace a' by $\frac{aq'}{q}$ in the sum $S(Q)$. One easily sees that the error produced by this approximation is very small, and we obtain an asymptotic relation

$$S(Q) \sim \sum_{\substack{a/q \in \mathcal{F}_Q \\ a/q < M_A}} \frac{1}{F^3\left(\frac{a}{q}\right) qq'(q+q')} \,. \tag{2.2}$$

The next step is to put the sum from the right side of (2.2) in the form

$$\sum_{1 \le q \le Q} \sum_{\substack{x \in I, y \in J \\ xy = -1 (mod q)}} f(x, y, q). \tag{2.3}$$

This would then clear the way for bringing Kloosterman sums into play. Here I, J are intervals and $f(x, y, q)$ will be a function given explicitly. In our case x stands for a and y stands for q'. From the equality $a'q - aq' = 1$ we see that $aq' = -1 \pmod q$, which explains why we impose the above condition $xy = -1 \pmod q$. Also, since we want $a/q < M_A$, we will define I to be the interval $(0, qM_A)$.

Now, from the equality $q + q' > Q$, which holds true for the denominators of any two consecutive Farey fractions in $\mathcal{F}_Q$, we know that q' belongs to the interval $(Q - q, Q]$. This is an interval of length q, so it will contain exactly one integer from each residue class modulo q. Only one of these integers satisfies the congruence $aq' = -1 \pmod q$. This is how one obtains uniquely q' as a function of q and a. In conclusion, we will take $J = (Q - q, Q]$, and then obtain

$$S(Q) \sim \sum_{1 \le q \le Q} \sum_{\substack{x \in I, y \in J \\ xy = -1 (mod q)}} f(x, y, q)\,, \tag{2.4}$$

where

$$f(x, y, q) = \frac{1}{F^3\left(\frac{x}{q}\right) qy(q+y)} \,.$$

Then Lemma 3.3 from [5] gives

$$\sum_{\substack{x\in I, y\in J\\ xy=-1 (mod q)}} f(x,y,q) \sim \frac{\varphi(q)}{q^2}\int_I\int_J f(x,y,q)dydx, \tag{2.5}$$

where φ denotes the Euler function. The lemma actually gives precise bounds for the error term in this asymptotic result. Next, in our case $f(x,y,q)$ is a product, and we have

$$\sum_{\substack{x\in I, y\in J\\ xy=-1 (mod q)}} f(x,y,q) \sim \frac{\varphi(q)}{q^3}\left(\int_I \frac{1}{F^3\left(\frac{x}{q}\right)}dx\right)\left(\int_J \frac{1}{y(q+y)}dy\right). \tag{2.6}$$

Both integrals can be computed in closed form, as functions of q. We would like to put the right side of (2.6) in the form $\big(\varphi(q)/q\big)g(q)$. The second integral above equals

$$\frac{1}{q}\int_{Q-q}^{Q}\left(\frac{1}{y}-\frac{1}{y+q}\right)dy = \frac{1}{q}\left(2\log Q - \log(Q+q) - \log(Q-q)\right).$$

This can also be written as

$$\int_J \frac{1}{y(q+y)}dy = -\frac{1}{q}\log\left(1-\frac{q^2}{Q^2}\right). \tag{2.7}$$

In the other integral we perform a change of variable to obtain

$$\int_I \frac{1}{F^3\left(\frac{x}{q}\right)}dx = qc_A, \tag{2.8}$$

where

$$c_A = \int_0^{M_A} \frac{1}{F^3(t)}dt \tag{2.9}$$

is a nonzero constant which only depends on the original matrix A, and can be computed in closed form. For instance, in the case when $Q_A(X,Y) = X^2+Y^2$, $M_A = 1$ and $P(t) = t^2+1$, after a straightforward computation one finds that

$$c_A = \int_0^1 \frac{1}{(t^2+1)^{3/2}}dt = \frac{1}{\sqrt{2}}.$$

From (2.6), (2.7) and (2.8) we obtain

$$\sum_{\substack{x\in I, y\in J \\ xy=-1 (mod q)}} f(x,y,q) \sim -c_A \frac{\varphi(q)}{q^3} \log\left(1-\frac{q^2}{Q^2}\right). \tag{2.10}$$

We set

$$g(x) = -\frac{\log\left(1-\frac{x^2}{Q^2}\right)}{x^2}$$

and then insert (2.10) in (2.4) to obtain

$$S(Q) \sim c_A \sum_{1\le q\le Q} \frac{\varphi(q)}{q} g(q). \tag{2.11}$$

Now Lemma 2.3 from [3] gives, again with good bounds for the error term,

$$S(Q) \sim \frac{6c_A}{\pi^2}\int_1^Q g(x)dx = -\frac{6c_A}{\pi^2}\int_1^Q \frac{\log\left(1-\frac{x^2}{Q^2}\right)}{x^2}dx. \tag{2.12}$$

Here we make a change of variable and find that

$$S(Q) \sim \frac{6c_A c_0}{\pi^2 Q} \tag{2.13}$$

where

$$c_0 = \int_0^1 -\frac{\log(1-t^2)}{t^2}dt.$$

After a short computation one obtains $c_0 = 2\log 2$.

In conclusion we have the following result.

Theorem. *Let $S(Q)$ be defined by (2.1). Then*

$$S(Q) \sim \frac{12 c_A \log 2}{\pi^2 Q}$$

as $Q\to\infty$, where c_A is given by (2.9).

In the particular case mentioned above, this asymptotic result has the following explicit form, which is consistent with the numerical data provided by Galway.

Corollary. *In the case when $M_A = 1$ and $P(t) = t^2+1$ one has*

$$S(Q) \sim \frac{6\sqrt{2}\log 2}{\pi^2 Q}$$

as $Q\to\infty$.

References

[1] A. O. L. Atkin and D. J. Bernstein, *Prime sieves using binary quadratic forms*, Math. Comp. **73** (2004), no. 246, 1023–1030 (electronic).

[2] V. Augustin, F. P. Boca, C. Cobeli and A. Zaharescu, *The h-spacing distribution between Farey points*, Math. Proc. Cambridge Philos. Soc. **131** (2001), no. 1, 23–38.

[3] F. P. Boca, C. Cobeli and A. Zaharescu, *Distribution of lattice points visible from the origin*, Comm. Math. Phys. **213** (2000), no. 2, 433–470.

[4] F. P. Boca, C. Cobeli and A. Zaharescu, *A conjecture of R. R. Hall on Farey points*, J. Reine Angew. Math. **535** (2001), 207–236.

[5] F. P. Boca, R. N. Gologan and A. Zaharescu, *The statistics of the trajectory of a certain billiard in a flat two-torus*, Comm. Math. Phys. **240** (2003), no. 1-2, 53–73.

[6] T. Estermann, *On Kloosterman's sums*, Mathematika **8** (1961), 83–86.

[7] W. F. Galway, *Dissecting a sieve to cut its need for space*; In: Algorithmic number theory (Leiden, 2000), 297–312, Lecture Notes in Comput. Sci., 1838, Springer, Berlin, 2000.

[8] W. F. Galway, *Analytic computation of the prime-counting function*, Ph.D. thesis, University of Illinois at Urbana-Champaign, 2001.

[9] C. Hooley, *An asymptotic formula in the theory of numbers*, Proc. London Math. Soc. **7** (1957), 396–413.

[10] W. Sierpiński, *Sur un problème du calcul des fonctions asymptotiques*, 1906, Oeuvres choisies, Tome I, PWN–Éditions Scientifiques de Pologne, Warsaw, 1974, 73–108.

[11] G. Voronoi, *Sur un problème du calcul des fonctions asymptotiques*, J. Reine Angew. Math. **126** (1903), 241–282.

[12] A. Weil, *On some exponential sums*, Proc. Nat. Acad. U.S.A. **34** (1948), 204–207.

Department of Mathematics, University of Illinois at Urbana-Champaign, 1409 W. Green Street, Urbana, IL, 61801, USA

E-mail address: zaharesc@math.uiuc.edu