

## A note on Farey fractions with denominators in arithmetic progressions

by

DMITRY A. BADZIAHIN and ALAN K. HAYNES (York)

**1. Introduction.** For  $Q \in \mathbb{N}$  the *Farey fractions of order  $Q$*  are defined as

$$\mathcal{F}_Q = \left\{ \frac{a}{q} \in \mathbb{Q} : 1 \leq q \leq Q, 0 < a \leq q, (a, q) = 1 \right\},$$

where  $(a, q)$  denotes the greatest common divisor of  $a$  and  $q$ . They are assumed to be ordered in the natural way on the unit interval. For each  $d \in \mathbb{N}$  we define a subset  $\mathcal{F}_{Q,d} \subseteq \mathcal{F}_Q$  by

$$\mathcal{F}_{Q,d} = \left\{ \frac{a}{q} \in \mathcal{F}_Q : (q, d) = 1 \right\}.$$

Farey fractions play an important role in mathematics, computer science, and physics. For example there are many statements about the distribution of the Farey sequence which are known to be equivalent to the Riemann Hypothesis ([6], [14], [13]). Also knowledge about Farey fractions has led to significant progress in the study of billiards ([4]). By comparison, the restricted sets  $\mathcal{F}_{Q,d}$  are known to have applications to the Generalized Riemann Hypothesis ([12]) and to billiards in which the source is a point with non-zero rational coordinates ([2, Introduction]).

A fundamental problem is to understand the distribution of gaps between consecutive elements of the sets  $\mathcal{F}_{Q,d}$ . Toward this end for each  $k \in \mathbb{N}$  we define

$$N_{Q,d}(k) = \# \left\{ \frac{a}{q} < \frac{a'}{q'} \text{ consecutive in } \mathcal{F}_{Q,d} : qa' - aq' = k \right\}.$$

In this paper we prove the following theorem.

---

2010 *Mathematics Subject Classification*: Primary 11B57.  
*Key words and phrases*: Farey fractions.

THEOREM 1. *For any  $d > 1$  and  $k \in \mathbb{N}$  there exists a positive constant  $c(d, k)$  for which*

$$N_{Q,d}(k) = c(d, k)Q^2 + O_{d,k}(Q \log Q) \quad \text{as } Q \rightarrow \infty.$$

The case when  $d$  is a prime number has been previously studied in [9] and [10]. When  $d$  is composite the situation is somewhat more complicated. This is because when  $d$  is prime at least one out of any two consecutive fractions in  $\mathcal{F}_Q$  must also be an element of  $\mathcal{F}_{Q,d}$  (a crucial fact in the proofs of [9, Theorem 1] and [10, Theorem 1]). However when  $d$  is composite and when  $Q$  is large enough there are longer sequences of consecutive fractions in  $\mathcal{F}_Q$  which do not appear in  $\mathcal{F}_{Q,d}$ . In fact it is easy to see that by letting  $Q$  and  $d$  vary we can find sequences of this kind which are arbitrarily long. For a naive example of how this can happen, we could let  $d = Q!$  and then take  $Q$  to be as large as we like. Interestingly this appears not to have been an issue in the recent related papers [1] and [2].

We are able to overcome the phenomenon of large gaps by using properties of the  $\ell$ -index of the Farey sequence, introduced and studied in [11]. Let us write

$$\mathcal{F}_Q = \{\gamma_1, \gamma_2, \dots, \gamma_{N(Q)}\}$$

with  $1/Q = \gamma_1 < \gamma_2 < \dots < \gamma_{N(Q)} = 1$ . It is also natural to extend this sequence by requiring that  $\gamma_{i+N(Q)} = \gamma_i + 1$  for all  $i \in \mathbb{Z}$ . Then for each  $i$  we write  $\gamma_i = a_i/q_i$  with  $(a_i, q_i) = 1$  and  $q_i > 0$ . Given a positive integer  $\ell$  and a fraction  $\gamma_i$  in  $\mathcal{F}_Q$  the  $\ell$ -index of  $\gamma_i$  is defined to be the quantity

$$\nu_\ell(\gamma_i) = a_{i+\ell-1}q_{i-1} - a_{i-1}q_{i+\ell-1}.$$

To avoid confusion note that the definition of  $\nu_\ell$  is dependent on  $Q$ . The main theorem in [11] is an asymptotic formula for the average value of  $\nu_\ell$ , as  $Q \rightarrow \infty$ . In the course of proving that theorem it was also necessary to establish algebraic identities which allow one to express the  $\ell$ -index as a multivariable polynomial evaluated at the values taken by the 2-index. To be precise define  $K_0 \in \mathbb{Z}$  and  $K_1 \in \mathbb{Z}[x_1]$  by

$$K_0(\cdot) = 1 \quad \text{and} \quad K_1(x_1) = x_1,$$

and then for each  $\ell \geq 2$  define  $K_\ell \in \mathbb{Z}[x_1, \dots, x_\ell]$  by

$$K_\ell(x_1, \dots, x_\ell) = x_\ell K_{\ell-1}(x_1, \dots, x_{\ell-1}) + K_{\ell-2}(x_1, \dots, x_{\ell-2}).$$

These polynomials are known as *the convergent polynomials*, and they appear in the study of continued fractions (see [7, Section 6.7]). In [11, Theorem 1] it was proved that for any  $Q, \ell \in \mathbb{N}$  and for any  $\gamma_i \in \mathcal{F}_Q$  we have

$$(1) \quad \nu_\ell(\gamma_i) = \left(\frac{2\ell - 1}{2}\right) K_{\ell-1}(-\nu_2(\gamma_i), \nu_2(\gamma_{i+1}), \dots, (-1)^{\ell-1} \nu_2(\gamma_{i+\ell-2})),$$

where  $\left(\frac{\cdot}{2}\right)$  is the Kronecker symbol, defined by

$$\left(\frac{\ell}{2}\right) = \begin{cases} 0 & \text{if } 2 \mid \ell, \\ 1 & \text{if } \ell \equiv \pm 1 \pmod{8}, \\ -1 & \text{if } \ell \equiv \pm 3 \pmod{8}. \end{cases}$$

The  $\ell$ -index tells us what happens to the numerator of the difference of the endpoints of a consecutive  $(\ell + 1)$ -tuple in  $\mathcal{F}_Q$ , while the 2-index has a geometrical description in terms of an area-preserving surjection of a certain region in  $\mathbb{R}^2$ . In what follows we will show how these facts combine with (1) to give us much of what we need to prove Theorem 1.

**2. Notation and definitions.** In order to take advantage of identity (1) we will use a geometrical description of the Farey sequence and a convenient area-preserving map. These tools have been previously used by several authors (e.g. [3], [5], [11])

The *Farey triangle*  $\mathcal{T} \subseteq [0, 1]^2$  is defined by

$$\mathcal{T} = \{(x, y) \in [0, 1]^2 : x + y > 1\}$$

and the *Farey map*  $T : [0, 1]^2 \rightarrow [0, 1]^2$  is defined by

$$T(x, y) = \left(y, \left[\frac{1+x}{y}\right]y - x\right),$$

where  $[x]$  denotes the greatest integer less than or equal to  $x$ . The map  $T$  is a one-to-one area-preserving transformation of  $\mathcal{T}$  onto itself ([3]) and has the important property that

$$(2) \quad T\left(\frac{q_{j-1}}{Q}, \frac{q_j}{Q}\right) = \left(\frac{q_j}{Q}, \frac{q_{j+1}}{Q}\right).$$

Also for each positive integer  $k$  we let

$$\mathcal{T}_k = \left\{ (x, y) \in \mathcal{T} : \left[\frac{1+x}{y}\right] = k \right\}.$$

Note that the set  $\mathcal{T}$  is the disjoint union of the sets  $\mathcal{T}_k$  and that

$$T(x, y) = (y, ky - x) \quad \text{for all } (x, y) \in \mathcal{T}_k.$$

Moreover, it follows from [8, equation (1.4)] that given  $Q$  and  $i$ ,

$$(3) \quad \nu_2(\gamma_i) = k \quad \text{if and only if} \quad \left(\frac{q_{i-1}}{Q}, \frac{q_i}{Q}\right) \in \mathcal{T}_k.$$

In the next section we will need to know how frequently large values of the Farey index can be obtained. The following lemma answers this question by showing that if the 2-index of a particular fraction is large, then the 2-indices of the fractions close-by must be small.

LEMMA 1. *If  $Q \in \mathbb{N}$ ,  $k \geq 1$  and  $\nu_2(\gamma_i) \geq 4k + 1$  then*

$$\nu_2(\gamma_{i\pm 1}) = 1 \quad \text{and} \quad \nu_2(\gamma_{i\pm j}) = 2 \quad \text{for } 2 \leq j \leq k.$$

*Proof.* Since  $4k + 5 = 4(k + 1) + 1$ , it is enough to verify that the theorem is true when  $\nu_2(\gamma_i) = 4k + j$  with  $j = 1, 2, 3$ , or  $4$ .

Suppose that the fractions  $a'/q' < \gamma_i < a''/q''$  are consecutive in  $\mathcal{F}_{q_i}$  and for each positive integer  $m$  write

$$b_m = a' + ma_i, \quad c_m = a'' + ma_i, \quad r_m = q' + mq_i, \quad s_m = q'' + mq_i.$$

Then  $(b_m, r_m) = (c_m, s_m) = 1$  for all  $m$ .

First we consider the case when  $\nu_2(\gamma_i) = 4k + 1$ . In this case we claim that the fractions

$$\frac{b_k}{r_k} < \dots < \frac{b_{2k-1}}{r_{2k-1}} < \frac{b_{2k}}{r_{2k}} < \frac{a_i}{q_i} < \frac{c_{2k}}{s_{2k}} < \frac{c_{2k-1}}{s_{2k-1}} < \dots < \frac{c_k}{s_k}$$

are consecutive in  $\mathcal{F}_Q$ . To see this first let  $M$  and  $N$  be the largest positive integers for which  $r_M, s_N \leq Q$ . Then we must certainly have  $r_M, s_N > Q - q_i$  and thus  $|r_M - s_N| < q_i$  and  $|M - N| \leq 1$ . By a well known formula for the 2-index ([8, equation (1.2)]) we also have

$$\nu_2(\gamma_i) = \frac{r_M + s_N}{q_i} = M + N + 1$$

and since  $\nu_2(\gamma_i) = 4k + 1$  this implies that  $M \equiv N \pmod{2}$  and we deduce that  $M = N = 2k$ .

Now notice that the fractions  $b_{2k}/r_{2k} < a_i/q_i$  are consecutive in  $\mathcal{F}_{r_{2k}}$  and by the mediant property of Farey fractions the fraction with smallest denominator which lies between them is  $b_{2k+1}/r_{2k+1}$ . Since we have already checked that  $r_{2k+1} > Q$  this shows that  $b_{2k}/r_{2k}$  and  $a_i/q_i$  are consecutive in  $\mathcal{F}_Q$ . Similarly for  $0 \leq i \leq k - 1$  the fractions  $b_{k+i}/r_{k+i} < b_{k+i+1}/r_{k+i+1}$  are consecutive in  $\mathcal{F}_{r_{k+i+1}}$  and the fraction with smallest denominator which lies between them is

$$(b_{k+i} + b_{k+i+1}) / (r_{k+i} + r_{k+i+1}).$$

Since

$$r_{k+i} + r_{k+i+1} = 2q' + (2(k + i) + 1)q_i > r_{2k+1} > Q$$

this shows that  $b_{k+i}/r_{k+i}$  and  $b_{k+i+1}/r_{k+i+1}$  are consecutive in  $\mathcal{F}_Q$ . The analysis of the fractions to the right of  $a_i/q_i$  is identical and this verifies our above claim. The rest of the proof in this case follows from a straightforward calculation.

In the case when  $\nu_2(\gamma_i) = 4k + 2$  we apply the same arguments to deduce that either the fractions

$$\frac{b_{k+1}}{r_{k+1}} < \dots < \frac{b_{2k}}{r_{2k}} < \frac{b_{2k+1}}{r_{2k+1}} < \frac{a_i}{q_i} < \frac{c_{2k}}{s_{2k}} < \frac{c_{2k-1}}{s_{2k-1}} < \dots < \frac{c_k}{s_k}$$

or the fractions

$$\frac{b_k}{r_k} < \dots < \frac{b_{2k-1}}{r_{2k-1}} < \frac{b_{2k}}{r_{2k}} < \frac{a_i}{q_i} < \frac{c_{2k+1}}{s_{2k+1}} < \frac{c_{2k}}{s_{2k}} < \dots < \frac{c_{k+1}}{s_{k+1}}$$

are consecutive in  $\mathcal{F}_Q$ . Either way the lemma is verified, and the two remaining cases are dealt with in the same manner. ■

We will also need estimates for the number of lattice points in a convex region satisfying certain congruence constraints.

LEMMA 2. *Suppose that  $\Omega$  is a convex region contained in the square  $[0, Q] \times [0, Q]$ , that  $\mathcal{A}$  and  $\mathcal{B}$  are subsets of  $\{1, \dots, d\}$ , and that each element  $a$  of  $\mathcal{A}$  satisfies  $(a, d) = 1$ . Then for the quantity*

$$Z_{\mathcal{A}, \mathcal{B}, d}(\Omega) = \#\{(m, n) \in \Omega \cap \mathbb{Z}_{\text{vis}}^2 : (m, n) \equiv (a, b) \pmod{d} \text{ for some } (a, b) \in \mathcal{A} \times \mathcal{B}\}$$

we have the estimate

$$Z_{\mathcal{A}, \mathcal{B}, d}(\Omega) = \frac{6 \text{Area}(\Omega) |\mathcal{A}| |\mathcal{B}|}{\pi^2 d^2} \prod_{p|d} \left(1 - \frac{1}{p^2}\right)^{-1} + O\left(\frac{\text{Area}(\Omega)}{Q} + \text{Length}(\partial\Omega) \log Q\right),$$

where the product is taken over primes dividing  $d$ .

*Proof.* First of all by Möbius inversion we have

$$Z_{\mathcal{A}, \mathcal{B}, d}(\Omega) = \sum_{\substack{(m,n) \in \Omega \\ (m,n) \pmod{d} \in \mathcal{A} \times \mathcal{B}}} \sum_{e|m,n} \mu(e) = \sum_{e \leq Q} \mu(e) \sum_{\substack{(m,n) \in \Omega, e|m,n \\ (m,n) \pmod{d} \in \mathcal{A} \times \mathcal{B}}} 1.$$

Since the residue classes in  $\mathcal{A}$  are coprime to  $d$  the inner sum on the right hand side here will be zero if  $(e, d) > 1$ . Thus the right hand side is equal to

$$\sum_{\substack{e \leq Q \\ (e,d)=1}} \mu(e) \sum_{\substack{(m,n) \in \Omega \\ (m,n) \pmod{de} \in \mathcal{A}_e \times \mathcal{B}_e}} 1,$$

where  $\mathcal{A}_e, \mathcal{B}_e \subseteq \{1, \dots, de\}$  are uniquely determined (i.e. by the Chinese Remainder Theorem) and satisfy  $|\mathcal{A}_e| = |\mathcal{A}|$  and  $|\mathcal{B}_e| = |\mathcal{B}|$ . Now by using a standard estimate to count the lattice points in the innermost sum we find that the above expression equals

$$\sum_{\substack{e \leq Q \\ (e,d)=1}} \mu(e) \left( \frac{\text{Area}(\Omega) |\mathcal{A}| |\mathcal{B}|}{(de)^2} + O\left(\frac{\text{Length}(\partial(\Omega)) \max\{|\mathcal{A}|, |\mathcal{B}|\}}{de}\right) \right)$$

$$\begin{aligned}
 &= \frac{\text{Area}(\Omega)|\mathcal{A}||\mathcal{B}|}{d^2} \sum_{\substack{e \leq Q \\ (e,d)=1}} \frac{\mu(e)}{e^2} + O(\text{Length}(\partial(\Omega)) \log Q) \\
 &= \frac{6 \text{Area}(\Omega)|\mathcal{A}||\mathcal{B}|}{\pi^2 d^2} \prod_{p|d} \left(1 - \frac{1}{p^2}\right)^{-1} + O\left(\frac{\text{Area}(\Omega)}{Q} + \text{Length}(\partial\Omega) \log Q\right).
 \end{aligned}$$

The final equality here comes from the fact that

$$\sum_{\substack{e \leq Q \\ (e,d)=1}} \frac{\mu(e)}{e^2} = \sum_{e=1}^{\infty} \frac{\mu(e)}{e^2} + O\left(\frac{1}{Q}\right) = \zeta(2)^{-1} \prod_{p|d} \left(1 - \frac{1}{p^2}\right)^{-1} + O\left(\frac{1}{Q}\right). \blacksquare$$

**3. Proof of Theorem 1.** For each positive integer  $\ell$  let us define

$$\begin{aligned}
 N_{Q,d}(k, \ell) = \#\{1 \leq i \leq N(Q) : (q_{i-1}, d) = (q_{i+\ell-1}, d) = 1, \\
 (q_{i+j-1}, d) > 1 \text{ for } 1 \leq j < \ell, \nu_\ell(\gamma_i) = k\}.
 \end{aligned}$$

Then it is clear that

$$(4) \quad N_{Q,d}(k) = \sum_{\ell=1}^{N(Q)-1} N_{Q,d}(k, \ell).$$

Our proof will consist of the following three steps:

- i. We will show that there is an integer  $L = L(d)$  such that for any  $Q$  and  $i$  we have  $(q_{i+j-1}, d) = 1$  for some  $1 \leq j \leq L$ .
- ii. We will show that for each  $\ell$  and  $k$  there is an integer  $K = K(\ell, k)$  such that if  $\nu_\ell(\gamma_i) = k$  for some  $Q$  and  $i$  then the  $(\ell - 1)$ -tuple  $(\nu_2(\gamma_i), \dots, \nu_2(\gamma_{i+\ell-2}))$  must take one of at most  $K^{\ell-1}$  possible values.
- iii. We will use the information gathered in the first two steps to estimate (4) by counting visible lattice points in subregions of the Farey triangle which satisfy our congruence constraints.

STEP i. First suppose that  $Q$  and  $i_1 < i_2$  are chosen so that

$$(5) \quad \max\{q_{i_1}, q_{i_2}\} \leq q_j \quad \text{and} \quad (q_j, d) > 1 \quad \text{for all } i_1 \leq j \leq i_2.$$

It follows easily from this that  $\gamma_{i_1}$  and  $\gamma_{i_2}$  are consecutive in the Farey fractions of order  $\max\{q_{i_1}, q_{i_2}\}$ . Thus by the mediant property of the Farey sequence we have

$$\begin{aligned}
 (6) \quad \{q_j : i_1 < j < i_2\} \\
 = \{mq_{i_1} + nq_{i_2} : m, n \in \mathbb{N}, (m, n) = 1, mq_{i_1} + nq_{i_2} \leq Q\}.
 \end{aligned}$$

Now let  $d_1$  be the largest divisor of  $d$  which is coprime to  $q_{i_1}$  and choose  $m_1 \in \mathbb{N}$  so that  $m_1 \leq d_1$  and

$$m_1 \equiv (1 - q_{i_2})q_{i_1}^{-1} \pmod{d_1}.$$

Then for any prime  $p$  which divides  $d_1$  we have  $p \nmid m_1q_{i_1} + q_{i_2}$ . On the other hand since  $\gamma_{i_1}$  and  $\gamma_{i_2}$  are consecutive Farey fractions of some order we have  $(q_{i_1}, q_{i_2}) = 1$  and so if  $p \mid (d/d_1)$  then  $p \nmid q_{i_2}$ . This shows that  $(m_1q_{i_1} + q_{i_2}, d) = 1$  and it follows from (5) and (6) that  $Q < m_1q_{i_1} + q_{i_2} \leq dq_{i_1} + q_{i_2}$ . By repeating the same argument we also have  $Q < q_{i_1} + dq_{i_2}$  and so

$$(7) \quad \{q_j : i_1 < j < i_2\} \subseteq \{mq_{i_1} + nq_{i_2} : 1 \leq m, n < d\}.$$

From this it is apparent that if (5) is satisfied then we must have  $i_2 - i_1 < d^2$ .

Now suppose more generally that  $Q$ ,  $i$ , and  $L_0$  are chosen so that

$$(q_{i+j-1}, d) > 1 \quad \text{for all } 1 \leq j \leq L_0.$$

Choose an integer  $i_0 \in \{i, \dots, i + L_0 - 1\}$  such that

$$q_{i_0} = \min\{q_j : i \leq j \leq i + L_0 - 1\}.$$

This integer uniquely determines a pair of non-negative integers  $n$  and  $n'$  and a finite sequence  $i_{-n} < \dots < i_0 < \dots < i_{n'}$  with the properties that  $i_{-n} = i$ ,  $i_{n'} = i + L_0 - 1$ ,

$$q_{i_m} = \begin{cases} \min\{q_j : i_{m-1} < j \leq i + L_0 - 1\} & \text{for } 1 \leq m \leq n', \\ \min\{q_j : i \leq j \leq i_{m+1}\} & \text{for } -n \leq m \leq -1, \end{cases}$$

and

$$\max\{q_{i_m}, q_{i_{m+1}}\} < q_j \quad \text{for all } -n \leq m < n' \text{ and } i_m < j < i_{m+1}.$$

Our argument above shows that  $i_{m+1} - i_m < d^2$  for each  $-n \leq m < n'$ , so all that remains is to bound  $n$  and  $n'$  in terms of  $d$ .

If  $q_{i_m}$  were equal to  $q_{i_{m+1}}$  for some  $m$  then there would have to be a fraction with smaller denominator between  $a_{i_m}/q_{i_m}$  and  $a_{i_{m+1}}/q_{i_{m+1}}$ . Since our setup does not allow this we must have

$$q_{i_0} < q_{i_{-1}} < \dots < q_{i_{-n}} \quad \text{and} \quad q_{i_0} < q_{i_1} < \dots < q_{i_{n'}}.$$

Now for each integer  $0 \leq m < n'$  let  $b_m/r_m$  be the fraction which immediately follows  $a_{i_m}/q_{i_m}$  in the Farey fractions of order  $q_{i_{m+1}} - 1$ . It is evident that  $r_0 \leq r_1 \leq \dots \leq r_{n'-1}$  and  $q_{i_{m+1}} = q_{i_m} + r_m$  for each  $0 \leq m < n'$ . Furthermore it is not difficult to see that if  $r_m > r_{m-1}$  for some  $m$  then  $r_m > q_{i_m}$ . Suppose that  $r_0 = r_m$  for all  $m < m'$ . Then

$$q_{i_{m'}} = q_{i_0} + m'r_0$$

and making use of the argument used to prove (7) we find that  $m' < d$ . Thus if  $n' > d$  then for some  $1 \leq m \leq d$  we must have  $r_m > r_{m-1}$ . Then

since  $r_m > q_{i_m} > q_{i_0}$  and

$$q_{i_{n'}} \geq q_{i_m} + (n' - m)r_m$$

we have

$$q_{i_{n'}} > q_{i_0} + (n' - m)q_{i_1}.$$

Again by the argument used to prove (7) we have  $n' - m < d$ , which allows us to conclude that  $n' < 2d$ . Similarly we find that  $n < 2d$  and putting these results together shows that  $L_0 < L(d) = 4d^3$ .

STEP ii. In this step we are supposing that  $\ell$  and  $k$  are fixed and we want to show that if  $\nu_\ell(\gamma_i) = k$  for some  $Q$  and  $i$  then there are only finitely many possible values for the  $(\ell - 1)$ -tuple of integers  $(\nu_2(\gamma_i), \dots, \nu_2(\gamma_{i+\ell-2}))$ .

First suppose that  $K_0 > \ell$  and that  $\nu_2(\gamma_{i_0}) = 4K_0 + 1$  for some  $i \leq i_0 \leq i + \ell - 2$ . Then by Lemma 1 we know that  $\nu_2(\gamma_{i_0 \pm 1}) = 1$  and that  $\nu_2(\gamma_{i_0 \pm j}) = 2$  for  $2 \leq j \leq \ell$ . Furthermore following the proof of the lemma if we suppose that

$$\frac{a'}{q'} < \frac{a_{i_0}}{q_{i_0}} < \frac{a''}{q''}$$

are consecutive in  $\mathcal{F}_{q_{i_0}}$  then it is easy to see that

$$\gamma_{i-1} = \frac{a' + (2K_0 - (i_0 - i))a_{i_0}}{q' + (2K_0 - (i_0 - i))q_{i_0}}$$

and

$$\gamma_{i+\ell-1} = \frac{a'' + (2K_0 - (i + \ell - 2 - i_0))a_{i_0}}{q'' + (2K_0 - (i + \ell - 2 - i_0))q_{i_0}}.$$

Using the definition of  $\nu_\ell$  we have

$$(8) \quad \nu_\ell(\gamma_i) = 2K_0(q_{i_0}a'' + a_{i_0}q' - q_{i_0}a' - a_{i_0}q'') + C = 4K_0 + C,$$

where the constant  $C$  is an integer which depends only on  $a_{i_0}$ ,  $q_{i_0}$ , and  $i$ .

We know from identity (1) that  $\nu_\ell(\gamma_i)$  is a linear function of each of the variables  $\nu_2(\gamma_{i+j})$ ,  $0 \leq j \leq \ell - 2$ , and this together with Lemma 1 allows us to interpolate the values of  $\nu_2(\gamma_{i_0})$  between integers of the form  $4K_0 + 1$ . Since the right hand side of (8) tends to infinity with  $K_0$  this allows us to conclude that there is an integer  $K$  for which  $\nu_\ell(\gamma_i) > k$  whenever  $\nu_2(\gamma_{i_0}) \geq K$  for some  $i \leq i_0 \leq i + \ell - 2$ .

Notice that the right hand side of (1) can be evaluated just by knowing the value of the  $(\ell - 1)$ -tuple  $(\nu_2(\gamma_i), \dots, \nu_2(\gamma_{i+\ell-2}))$ . Once these integers have been determined no further information is needed about  $Q$  or  $i$  in order to specify the value of  $\nu_\ell(\gamma_i)$ . Thus the integer  $K$  which we found above can be chosen so that it depends only on  $\ell$  and  $k$ . This allows us to conclude that for any  $Q$  and  $i$  if  $\nu_\ell(\gamma_i) = k$  then the number of possibilities for the  $(\ell - 1)$ -tuple  $(\nu_2(\gamma_i), \dots, \nu_2(\gamma_{i+\ell-2}))$  is at most  $K^{\ell-1}$ .



STEP iii. Let us assume that  $Q$  is large enough that  $N(Q) > L(d)$ . Then (4) may be replaced by

$$N_{Q,d}(k) = \sum_{\ell=1}^L N_{Q,d}(k, \ell).$$

For each  $1 \leq \ell \leq L$  let  $\{\vec{x}(\ell, m)\}_{1 \leq m \leq n(\ell)}$  be the collection of possible  $(\ell - 1)$ -tuples of integers from Step ii which can appear as the 2-indices corresponding to an  $\ell$ -index which takes the value  $k$ . Then we can write  $N_{Q,d}(k, \ell)$  as

$$\sum_{m=1}^{n(\ell)} \#\{1 \leq i \leq N(Q) : (q_{i-1}, d) = (q_{i+\ell-1}, d) = 1, (q_{i+j-1}, d) > 1$$

$$\text{for } 1 \leq j < \ell, (\nu_2(\gamma_i), \dots, \nu_2(\gamma_{i+\ell-2})) = \vec{x}(\ell, m)\}.$$

If we write

$$\vec{x}(\ell, m) = (x_1(\ell, m), \dots, x_{\ell-1}(\ell, m))$$

then if  $(\nu_2(\gamma_i), \dots, \nu_2(\gamma_{i+\ell-2})) = \vec{x}(\ell, m)$  for some  $i$  we obtain

$$q_{i+j-1} = x_{j-1}(\ell, m)q_{i+j-2} - q_{i+j-3}$$

for each  $2 \leq j \leq \ell$ . Thus each  $(\ell - 1)$ -tuple  $\vec{x}(\ell, m)$  determines a pair of collections of residue classes  $\mathcal{A}(\ell, m)$  and  $\mathcal{B}(\ell, m)$  modulo  $d$  with the property that

$$\begin{cases} (q_{i-1}, d) = (q_{i+\ell-1}, d) = 1, & (q_{i+j-1}, d) > 1 \quad \text{for } 1 \leq j < \ell, \\ (\nu_2(\gamma_i), \dots, \nu_2(\gamma_{i+\ell-2})) = \vec{x}(\ell, m) \end{cases}$$

if and only if

$$(q_{i-1}, q_i) \bmod d \in \mathcal{A}(\ell, m) \times \mathcal{B}(\ell, m), \quad (\nu_2(\gamma_i), \dots, \nu_2(\gamma_{i+\ell-2})) = \vec{x}(\ell, m).$$

Furthermore it is obvious that  $(a, d) = 1$  for all  $a \in \mathcal{A}(\ell, m)$ . Next using (3) and (2) we see that

$$(\nu_2(\gamma_i), \dots, \nu_2(\gamma_{i+\ell-2})) = \vec{x}(\ell, m)$$

if and only if

$$(q_{i-1}, q_i) \in Q(\mathcal{T}_{x_1} \cap T^{-1}\mathcal{T}_{x_2} \cap \dots \cap T^{-(\ell-2)}\mathcal{T}_{x_{\ell-1}}) \cap \mathbb{Z}_{\text{vis}}^2.$$

Here we are implicitly using the well known fact that if  $(a, b) \in Q\mathcal{T} \cap \mathbb{Z}_{\text{vis}}^2$  then  $a = q_{i-1}$  and  $b = q_i$  for some  $i$ . Putting all this together gives us the formula

$$N_{Q,d}(k) = \sum_{\ell=1}^L \sum_{m=1}^{n(\ell)} Z_{\mathcal{A}(\ell,m), \mathcal{B}(\ell,m), d}(Q(\mathcal{T}_{x_1} \cap \dots \cap T^{-(\ell-2)}\mathcal{T}_{x_{\ell-1}})),$$

where for ease of notation we have suppressed the dependence on  $\ell$  and  $m$ . It is easy to check that each of the regions  $Q(\mathcal{T}_{x_1} \cap T^{-1}\mathcal{T}_{x_2} \cap \dots \cap T^{-(\ell-2)}\mathcal{T}_{x_{\ell-1}})$

is convex and so we may apply Lemma 2 to obtain the statement of the theorem with  $c(d, k)$  equal to

$$\frac{6}{\pi^2 d^2} \prod_{p|d} \left(1 - \frac{1}{p^2}\right)^{-1} \\ \times \sum_{\ell=1}^L \sum_{m=1}^{n(\ell)} |\mathcal{A}(\ell, m)| |\mathcal{B}(\ell, m)| \text{Area}(\mathcal{T}_{x_1} \cap \dots \cap T^{-(\ell-2)} \mathcal{T}_{x_{\ell-1}}).$$

Finally we note that some of the sets  $\mathcal{A}(\ell, m)$  and  $\mathcal{B}(\ell, m)$  and some of the regions  $\mathcal{T}_{x_1} \cap T^{-1} \mathcal{T}_{x_2} \cap \dots \cap T^{-(\ell-2)} \mathcal{T}_{x_{\ell-1}}$  may be empty, so to finish the proof we must show that  $c(d, k)$  is never zero. To see this note that the set  $\mathcal{T}_k$  is always non-empty. By Lemma 2 if  $Q$  is large enough there will be on the order of  $Q^2$  lattice points  $(a, b) \in Q\mathcal{T}_k \cap \mathbb{Z}_{\text{vis}}^2$  which satisfy  $a \equiv 1 \pmod{d}$  and  $b \equiv 0 \pmod{d}$ . For each of these points we will have  $(kb - a, d) = 1$  and they must therefore correspond to fractions  $\gamma_i \in \mathcal{F}_Q$  with  $(q_{i-1}, d) = (q_{i+1}, d) = 1$ ,  $(q_i, d) > 1$ , and  $\nu_2(\gamma_i) = k$ . Thus for  $Q$  large enough  $N_{Q,d}(k)$  is at least as large as a constant times  $Q^2$ . This precludes the possibility of having  $c(d, k)$  equal to zero, and completes our proof of Theorem 1. ■

**Acknowledgements.** Research of DB supported by EPSRC grant EP/E061613/1. Research of AH supported by EPSRC grant EP/F027028/1.

## References

- [1] E. Alkan, A. H. Ledoan, M. Văjăitu, and A. Zaharescu, *Discrepancy of fractions with divisibility constraints*, Monatsh. Math. 149 (2006), 179–192.
- [2] —, —, —, *Discrepancy of sets of fractions with congruence constraints*, Rev. Roumaine Math. Pures Appl. 51 (2006), 265–276.
- [3] 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.
- [4] F. P. Boca, R. N. Gologan, and A. Zaharescu, *The average length of a trajectory in a certain billiard in a flat two-torus*, New York J. Math. 9 (2003), 303–330.
- [5] F. P. Boca, A. Zaharescu, and R. N. Gologan, *On the index of Farey sequences*, Quart. J. Math. 53 (2002), 377–391.
- [6] J. Franel, *Les suites de Farey et le problème des nombres premiers*, Göttinger Nachr. 1924, 198–201.
- [7] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics*, Addison-Wesley, New York, 1994.
- [8] R. R. Hall and P. Shiu, *The index of a Farey sequence*, Michigan Math. J. 51 (2003), 209–223.
- [9] A. K. Haynes, *A note on Farey fractions with odd denominators*, J. Number Theory 98 (2003), 89–104.
- [10] —, *The distribution of special subsets of the Farey sequence*, *ibid.* 107 (2004), 95–104.

- [11] A. K. Haynes, *Numerators of differences of nonconsecutive Farey fractions*, Int. J. Number Theory 6 (2010), 655–666.
- [12] M. N. Huxley, *The distribution of Farey points I*, Acta Arith. 18 (1971), 281–287.
- [13] S. Kanemitsu and M. Yoshimoto, *Farey series and the Riemann hypothesis*, *ibid.* 75 (1996), 351–374.
- [14] E. Landau, *Bemerkungen zu der vorstehenden Abhandlung von Herrn Franel*, Göttinger Nachr. 1924, 202–206.

Dmitry A. Badziahin, Alan K. Haynes  
Department of Mathematics  
University of York  
Heslington, York, YO10 5DD, England  
E-mail: db528@york.ac.uk  
akh502@york.ac.uk

*Received on 27.4.2009*

(6012)

