ACTA ARITHMETICA
LXVI.1 (1994)

On consecutive Farey arcs II

by

R. R. Hall (York)

1. Introduction. Let $\mathcal{F}_N = \{x_r : 1 \le r \le R\}$ denote the Farey sequence of order N, that is, the positive irreducible fractions ≤ 1, with denominators $\le N$, arranged in increasing order. We have

$$R = R(N) = \varphi(1) + \ldots + \varphi(N) = \frac{3}{\pi^2}N^2 + O(N \log N)$$

where φ is Euler's function. We set $\ell_r = x_r - x_{r-1}$, $2 \le r \le R$, $\ell_1 = x_1$, $\ell_{r+R} = \ell_r$ for all r.

In our previous paper [2] Tenenbaum and I gave an asymptotic formula for the sum

$$T_N(\alpha, \beta) := \sum_{r=1}^{R} \ell_r^\alpha \ell_{r+1}^\beta \tag{1}$$

for (α, β) belonging to the set $\mathcal{D}_1 \cup \mathcal{D}_2$ in the plane: $\mathcal{D}_1 = \{(\alpha, \beta) : \alpha, \beta, \alpha + \beta < 2\}$, $\mathcal{D}_2 = \{(\alpha, \beta) : \alpha > 0, \beta > 0, \alpha + \beta \ge 2\}$. There is a *threshold* across the line $\alpha + \beta = 2$. The term threshold was defined in our later paper [3]: it applies to any asymptotic formula containing one or more parameters when

(i) the main term is a discontinuous function of the parameters, and

(ii) the main term has a simple shape in one domain and a much more complicated shape in another domain.

In the case of $T_N(\alpha, \beta)$ these domains are respectively $\mathcal{D}_1$ and $\mathcal{D}_2$. Our weakest error term was on the boundary, $\alpha + \beta = 2$. We showed that for $0 < \alpha < 2$,

$$T_N(\alpha, 2 - \alpha) = \frac{6}{\pi^2}N^{-2} \log N + O_\alpha(N^{-2})\,. \tag{2}$$

I now show that in the special case $\alpha = 1$, this formula may be substantially improved. I write $T_N := T_N(1, 1)$.

Theorem. *We have*

$$T_N = \frac{6}{\pi^2} N^{-2} \log N + A N^{-2} + O\left(\frac{\log N}{N^2\sqrt{N}}\right) \tag{3}$$

where

$$A = \frac{6}{\pi^2}\left(\gamma - \frac{\zeta'(2)}{\zeta(2)} + B\right), \tag{4}$$

γ *is Euler's constant, and*

$$B = \frac{1}{2} + \log 2 + 2\sum_{h=1}^{\infty} \frac{\zeta(2h)-1}{2h-1} = 2.546277\ldots \tag{5}$$

The method is elementary and depends on the particular choice of α and β: I have not identified the second main term in (2) in the general case. Some of the complications encountered in $\mathcal{D}_2$ remain, finally resolving themselves into the constant B. The formula should be compared with one of those given by Kanemitsu, Sita Rama Chandra Rao and Siva Rama Sarma [4], viz.

$$\begin{aligned} T_N(2,0) &= \sum_{r=1}^{R} \ell_r^2 = \sum_{r=1}^{R} (x_{r+1} - x_r)^2 \\ &= \frac{12}{\pi^2} N^{-2}\left(\log N + \gamma - \frac{\zeta'(2)}{\zeta(2)} + \frac{1}{2}\right) \\ &\quad + O_\varepsilon(N^{-3}\log^{5/3} N(\log\log N)^{1+\varepsilon})\,. \end{aligned} \tag{6}$$

We may combine (3) and (6) to obtain

$$\begin{aligned} &\sum_{r=1}^{R} (\ell_{r+1} - \ell_r)^2 \\ &\quad = \frac{12}{\pi^2} N^{-2}\left(\log N + \gamma - \frac{\zeta'(2)}{\zeta(2)} + 1 - B\right) + O(N^{-5/2}\log N) \end{aligned} \tag{7}$$

and

$$\begin{aligned} &\sum_{r=1}^{R} (x_{r+2} - x_r)^2 \\ &\quad = \frac{12}{\pi^2} N^{-2}\left(3\log N + 3\gamma - 3\frac{\zeta'(2)}{\zeta(2)} + 1 + B\right) + O(N^{-5/2}\log N)\,. \end{aligned} \tag{8}$$

These results suggest the conjecture that for each fixed h there exist constants $C(h)$ and $D(h)$ such that as $N \to \infty$,

$$\sum_{r=1}^{R} (x_{r+h} - x_r)^2 = C(h) N^{-2}\log N + D(h) N^{-2} + o(N^{-2})\,.$$

I am grateful to the referee of an earlier version of this paper and to Martin Huxley who each supplied a partial result in this direction. Huxley's is

$$\sum_{r=1}^{R}(x_{r+h}-x_r)^2 = \frac{12}{\pi^2}(2h-1)N^{-2}\log N + O\left(\frac{h^2\log h}{N^2}\right) \tag{9}$$

and the referee had the better error term $O(h^2N^{-2})$. The main terms must change for large h: the sum on the left of (9) is clearly not less than $h^2/R(N)$ and on the Riemann Hypothesis we obtain, via a result of Franel [1],

$$\sum_{r=1}^{R}(x_{r+h}-x_r)^2 = h^2R^{-1} + O_\varepsilon(N^{-1+\varepsilon})$$

uniformly for all h. We may deduce (9) from the following result.

PROPOSITION. *Uniformly for $j \ge 2$, we have*

$$G_j(N) = \sum_{r \pmod R} \ell_r\ell_{r+j} \ll N^{-2}\log j\,.$$

The sum in (9) is

$$hG_0(N) + 2\sum_{j=1}^{h-1}(h-j)G_j(N)$$

and of course we know $G_0(N)$ and $G_1(N)$ from (6) and (3). I will just sketch a proof of the proposition here.

First, if $x_s = a/c$ and $x_t = b/d$ are distinct elements of $\mathcal{F}_N$ then we have

$$|s-t| \gg \frac{N^2}{(c+d)^2}$$

because $\mathcal{F}_N$ contains all the (distinct) intermediate fractions

$$x = \frac{ua+vb}{uc+vd}, \qquad (u,v)=1,\ u,v \le N/(c+d)\,.$$

If ℓ_i is large, one of its end-points has a small denominator. It follows that provided $j \ge 2$, we have

$$\min\{\ell_r, \ell_{r+j}\} \ll \frac{\sqrt{j}}{N^2}\,. \tag{10}$$

Uniformly for $0 \le \alpha < 2 < \beta$ we have both

$$\sum_{r \pmod R} \ell_r^\alpha \ll (2-\alpha)^{-1}N^{2-2\alpha}$$

and

$$\sum_{\substack{r \ (\mathrm{mod}\ R) \\ \ell_r \le \lambda/N^2}} \ell_r^\beta \ll \left(\frac{\beta}{\beta-2}\right) N^{2-2\beta}\lambda^{\beta-2}$$

and we estimate $G_j(N)$ by applying (10) and Hölder's inequality with exponents α and $\beta = 2 + (\log j)^{-1}$, choosing $\lambda = c\sqrt{j}$ in the last sum with c big enough for (10). This proves the proposition and Huxley's formula (9) is a corollary.

2. Proof of the theorem. Our starting point is Lemma 2 of [2] which gives (for $\alpha = \beta = 1$)

$$T_N = \sum_{s=1}^{N} s^{-2} \sum_{\substack{r=N-s+1 \\ (r,s)=1}}^{N} r^{-1}t^{-1} \tag{11}$$

where

$$t = t(r,s,N) = s\left[\frac{N+r}{s}\right] - r\,. \tag{12}$$

For $k = 2, 3, \ldots$ we set $s_k = (2N+1)/k$, and we split the sum (11) into two parts U_N and V_N according as $s \le s_K = z$ or not. We choose

$$K = [N^{1/4}\log^{-1/2} N]\,. \tag{13}$$

We set

$$k(s) = \left[\frac{2N+1}{s}\right], \qquad 2N+1 = sk(s) + a(s)\,, \tag{14}$$

so that $k(s) = k$ for $s_{k+1} < s \le s_k$. We have

$$\left[\frac{N+r}{s}\right] = \begin{cases} k(s)-1, & N-s+1 \le r \le N-a(s)\,, \\ k(s), & N-a(s) < r \le N\,. \end{cases} \tag{15}$$

Notice that for each s, $r+t$ takes just two values (one value if $s|(2N+1)$), determined by (12) and (15). We consider the sum U_N. Put $r = N - r'$, $t = N - t'$ so that $0 \le r', t' < s \le z$ and

$$\frac{1}{rt} = \frac{1}{N^2} + \frac{r'+t'}{N^3} + O\left(\frac{s^2}{N^4}\right). \tag{16}$$

Hence

$$U_N = \frac{1}{N^2}\sum_{s\le z}\frac{\varphi(s)}{s^2} + E_N + O\left(\frac{z^2}{N^4}\right) \tag{17}$$

where

$$E_N = \frac{1}{N^3}\sum_{s\le z}\frac{1}{s^2}\sum_{\substack{N-s+1\le r\le N\\(r,s)=1}}(r'+t')\,.$$

We have

$$\sum_{s\le z}\frac{\varphi(s)}{s^2} = \frac{6}{\pi^2}\left(\log z+\gamma-\frac{\zeta'(2)}{\zeta(2)}\right)+O\left(\frac{\log z}{z}\right). \tag{18}$$

Next, by (12) and (15),

$$r'+t' = 2N-r-t = 2N-s\left[\frac{N+r}{s}\right] = 2N-sk(s)+s^* \tag{19}$$

where * denotes that this term counts if and only if $N-s<r\le N-a(s)$. Now

$$\sum_{\substack{x<r\le y\\(r,s)=1}}1 = \frac{\varphi(s)}{s}(y-x)+O(\tau(s)) \tag{20}$$

where τ is the divisor function. It follows that

$$\begin{aligned}\sum_{\substack{N-s+1\le r\le N\\(r,s)=1}}&(2N-sk(s)+s^*) \\ &= \varphi(s)(2N-sk(s))+\varphi(s)(s-a(s))+O(s\tau(s)) = \varphi(s)s+O(s\tau(s))\,,\end{aligned} \tag{21}$$

by (14). Hence

$$E_N = \frac{1}{N^3}\sum_{s\le z}\left(\frac{\varphi(s)}{s}+O\left(\frac{\tau(s)}{s}\right)\right) = \frac{6}{\pi^2}N^{-3}z+O(N^{-3}\log^2 z)\,. \tag{22}$$

We combine (17), (18) and (22) to obtain

$$U_N = \frac{6}{\pi^2}N^{-2}\left(\log z+\gamma-\frac{\zeta'(2)}{\zeta(2)}+\frac{z}{N}\right)+O\left(\frac{z^2}{N^4}+\frac{\log z}{N^2 z}+\frac{\log^2 z}{N^3}\right). \tag{23}$$

The error terms on the right are within that appearing in (3).

We turn our attention to V_N. We begin by writing the inner sum in (11) as

$$\begin{aligned}\sum_{\substack{r=N-s+1\\(r,s)=1}}^{N} r^{-1}t^{-1} &= \sum_{\substack{r=N-s+1\\(r,s)=1}}^{N}\frac{1}{r+t}\left(\frac{1}{r}+\frac{1}{t}\right) \\ &= 2s^{-1}\sum_{\substack{r=N-s+1\\(r,s)=1}}^{N}\left[\frac{N+r}{s}\right]^{-1}\frac{1}{r}\,,\end{aligned} \tag{24}$$

noticing the symmetry in r and t, and using (12).

We employ (15) and we obtain

$$V_N = \sum_{z<s\le N} \frac{2}{s^3}\left\{\frac{1}{k(s)-1}\sum_{\substack{r=N-s+1\\(r,s)=1}}^{N}\frac{1}{r} - \frac{1}{k(s)(k(s)-1)}\sum_{\substack{r=sk(s)-N\\(r,s)=1}}^{N}\frac{1}{r}\right\}, \tag{25}$$

the right-hand inner sum being empty if $s\,|\,(2N+1)$. For positive integers u, v $(u\le v)$ we have

$$\sum_{\substack{r=u\\(r,s)=1}}^{v}\frac{1}{r} = \frac{\varphi(s)}{s}\log\frac{v}{u} + O\Big(\frac{\tau(s)}{u}\Big) \tag{26}$$

and we apply this in (25). The error term is

$$\ll \sum_{z<s\le N}\frac{\tau(s)}{s^2N(N-s+1)} \ll N^{-2}\sum_{s>z}\frac{\tau(s)}{s^2} + N^{-3}\sum_{N/2<s\le N}\frac{\tau(s)}{N-s+1}$$
$$\ll N^{-2}z^{-1}\log z + c(\varepsilon)N^{-3+\varepsilon}.$$

This is (substantially) smaller than the error term in (3). We therefore have to consider the sum

$$\sum_{z<s\le N}\frac{2}{s^4}\left\{\frac{\varphi(s)}{k(s)-1}\log\Big(\frac{N}{N-s+1}\Big) - \frac{\varphi(s)}{k(s)(k(s)-1)}\log\Big(\frac{N}{sk(s)-N}\Big)\right\} \tag{27}$$

and we split this into ranges $(s_{k+1}, s_k]$, $2\le k<K$, in which $k(s)=k$. We employ the formula

$$\sum_{s\le x}\frac{\varphi(s)}{s} = \frac{6}{\pi^2}x + O(\log x) \tag{28}$$

and partial summation to obtain

$$\sum_{s_{k+1}<s\le s_k}\frac{\varphi(s)}{s^4}\log\Big(\frac{N}{N-s+1}\Big) = \frac{6}{\pi^2}\int_{s_{k+1}}^{s_k}\log\Big(\frac{N}{N-s+1}\Big)\frac{ds}{s^3} + O\Big(\frac{\log^2 N}{Ns_k^2}\Big), \tag{29}$$

$$\sum_{s_{k+1}<s\le s_k}\frac{\varphi(s)}{s^4}\log\Big(\frac{N}{sk-N}\Big) = \frac{6}{\pi^2}\int_{s_{k+1}}^{s_k}\log\Big(\frac{N}{sk-N}\Big)\frac{ds}{s^3} + O\Big(\frac{k\log^2 N}{Ns_k^2}\Big). \tag{30}$$

Hence

$$V_N = \frac{12}{\pi^2} \int_{s_K}^{N+1/2} \left\{ \frac{1}{k(s)-1} \log\left(\frac{N}{N-s+1}\right) - \frac{1}{k(s)(k(s)-1)} \log\left(\frac{N}{sk(s)-N}\right) \right\} \frac{ds}{s^3} + O(K^2 N^{-3} \log^2 N)\,, \tag{31}$$

the error term here absorbing the previous ones; it is contained in that given in (3). Let us denote the first term on the right of (31) by I_N. We substitute $s = (2N+1)/x$ to obtain

$$(2N+1)^2 I_N = \frac{12}{\pi^2} \int_2^K \left\{ \frac{x}{[x]-1} \log\left(\frac{Nx}{(N+1)x-2N-1}\right) + \frac{x}{[x]([x]-1)} \log\left(\frac{(2N+1)[x]-Nx}{Nx}\right) \right\} dx\,. \tag{32}$$

We have

$$\log\left(\frac{Nx}{(N+1)x-2N-1}\right) = \log\left(\frac{x}{x-2}\right) + O\left(\frac{1}{Nx}\right) \qquad (x \ge 3) \tag{33}$$

and

$$\log\left(\frac{(2N+1)[x]}{Nx} - 1\right) = \log\left(2\frac{[x]}{x} - 1\right) + O\left(\frac{1}{N}\right) \qquad (x \ge 2)\,. \tag{34}$$

We insert (33) and (34) into the right-hand side of (32). There remains an integral over the interval $2 \le x \le 3$ which may be evaluated explicitly. This yields

$$I_N = \frac{3}{\pi^2 N^2} \int_2^K \left(f(x) + \frac{2}{x}\right) dx + O\left(\frac{\log N}{N^3}\right) \tag{35}$$

in which

$$f(x) = -\frac{2}{x} + \frac{x}{[x]-1} \log\left(\frac{x}{x-2}\right) + \frac{x}{[x]([x]-1)} \log\left(2\frac{[x]}{x} - 1\right). \tag{36}$$

Let us define

$$B = \frac{1}{2} \int_2^\infty f(x)\, dx, \qquad B(K) = \frac{1}{2} \int_K^\infty f(x)\, dx\,. \tag{37}$$

We may put (31), (35) and (37) together to obtain

$$V_N = \frac{6}{\pi^2} N^{-2} \left(\log\frac{K}{2} + B - B(K)\right) + O(N^{-5/2} \log N) \tag{38}$$

and it remains to simplify B (which we do in the next section) and to estimate $B(K)$. A calculation shows that if $k \geq 3$ then

$$f(x) = \frac{4}{x^2} + \left(\frac{20}{3} + 4\theta(1-\theta)\right)\frac{1}{x^3} + O\left(\frac{1}{x^4}\right), \tag{39}$$

where $\theta = x - [x]$. The Bernoulli function $B_2(x) = \theta^2 - \theta + \frac{1}{6}$ has mean value 0 and the second mean value theorem gives

$$\int_K^\infty B_2(x)\,\frac{dx}{x^3} = O\left(\frac{1}{K^3}\right). \tag{40}$$

We may assume that $K \geq 3$ by (13). From (39) and (40) we obtain

$$B(K) = \frac{2}{K} + \frac{11}{6K^2} + O\left(\frac{1}{K^3}\right). \tag{41}$$

We insert this into (38) (it is more precise than we need in the present analysis) and add the result to (23). This gives (3), subject to a proof that (5) and (37) are equivalent.

3. The formula for B. It remains to show that the rather awkward expression for B given in (36) and (37) may be simplified. Let

$$a_m = \frac{1}{2}\int_m^{m+1} f(x)\,dx \tag{42}$$

so that by (39), $a_m \ll 1/m^2$, moreover

$$B = \sum_{m=2}^{\infty} a_m\,. \tag{43}$$

A computation gives

$$\begin{aligned} a_m = {} & \frac{(m-1)^2}{4m}\log(m+1) - \frac{m^2-8m+4}{4(m-1)}\log m \\ & - \frac{m^2+6m+1}{4m}\log(m-1) \\ & + \frac{m^2-4}{4(m-1)}\log(m-2) \qquad (m \geq 2) \end{aligned} \tag{44}$$

(where it is understood that when $m = 2$ the last term on the right is interpreted as O). We consider the partial sum $a_2 + a_3 + \ldots + a_n$, bracketing together the terms in this sum involving $\log l$, for $l = 2, 3, \ldots, n+1$. After some simplification we find that

$$a_2 + a_3 + \ldots + a_n = 2\sum_{l=2}^{n} \frac{\log l}{l^2-1} + b_n \tag{45}$$

where

$$(46)\quad b_n = -\frac{(n-1)(n+3)}{4n}\log(n-1)+\frac{n+2}{n+1}\log n+\frac{(n-1)^2}{4n}\log(n+1)$$
$$= \frac{-\log n}{n(n+1)}-\frac{(n-1)(n+3)}{4n}\log\left(1-\frac{1}{n}\right)+\frac{(n-1)^2}{4n}\log\left(1+\frac{1}{n}\right)$$
$$= \frac{1}{2}+O\left(\frac{1}{n}\right).$$

It follows from (43), (45) and (46) that

$$(47)\qquad B = \frac{1}{2}+2\sum_{l=2}^{\infty}\frac{\log l}{l^2-1}$$
$$= \frac{1}{2}+\log 2+2\sum_{n=2}^{\infty}\frac{1}{n}\log\left(\frac{n+1}{n-1}\right)$$
$$= \frac{1}{2}+\log 2+2\left(\zeta(2)-1+\frac{\zeta(4)-1}{3}+\ldots\right)$$

as required.

References

[1] J. Franel, *Les suites de Farey et le problème des nombres premiers*, Göttinger Nachr. 1924, 198–201.

[2] R. R. Hall and G. Tenenbaum, *On consecutive Farey arcs*, Acta Arith. 44 (1984), 397–405.

[3] —, —, *The set of multiples of a short interval*, in: Number Theory, New York Seminar 1989–90, D. V. Chudnovsky, G. V. Chudnovsky, H. Cohn and M. B. Nathanson (eds.), Springer, 1991, 119–128.

[4] S. Kanemitsu, R. Sita Rama Chandra Rao and A. Siva Rama Sarma, *Some sums involving Farey fractions* 1, J. Math. Soc. Japan 34 (1982), 125–142.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF YORK
YORK Y01 5DD
ENGLAND

Received on 14.2.1992
and in revised form on 27.7.1992 (2227)