On the Distribution of Small Denominators in the Farey Series of Order N

C.L. Stewart*

In memory of Professor Herb Wilf

1 Introduction

Let N be a positive integer. The Farey series of order N is the sequence of rationals h/k with h and k coprime and $1 \leq h \leq k \leq N$ arranged in increasing order between 0 and 1, see [1]. There are $\varphi(k)$ rationals with denominator k in F_N and thus the number of terms in F_N is R where

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

(see Theorem 330 of [3]). Let

$$S(N) = \sum_{i=1}^{N} q_i$$

where q_i denotes the smallest denominator possessed by a rational from F_N which lies in the interval $\left(\frac{i-1}{N}, \frac{i}{N}\right]$. In [4] Kruyswijk and Meijer proved that

$$N^{3/2} \ll S(N) \ll N^{3/2} \qquad (2)$$

*Research supported in part by the Canada Research Chairs Program and by Grant A3528 from the Natural Sciences and Engineering Research Council of Canada.

C.L. Stewart (✉)
Department of Pure Mathematics, University of Waterloo, Waterloo, ON N2L 3G1, Canada
e-mail: cstewart@uwaterloo.ca

I.S. Kotsireas and E.V. Zima (eds.), *Advances in Combinatorics*,
DOI 10.1007/978-3-642-30979-3_15,

and they remarked that the function $S(N)$ is connected with a problem in combinatorial group theory. In particular, C. Schaap proved that for any prime p, $S(p) = p^2 - p + 1 - L(p)$ where $L = L(p)$ is the largest integer for which there is a sequence of integers $a_1, \ldots, a_L$ with $1 \leq a_1 \leq a_2 \leq \cdots \leq a_L \leq p - 1$ for which $a_1 + \cdots + a_j \not\equiv 0 (\bmod\ p)$ for $1 \leq j \leq L$. An examination of Kruyswijk and Meijer's proof shows that the implied constants in (2) may be made explicit and that $\frac{1}{\pi^2} N^{3/2} < S(N) < 96 N^{3/2}$ for N sufficiently large. They conjectured that $\lim_{N\to\infty} S(N)/N^{3/2}$ exists and is equal to $(\frac{4}{\pi})^2 = 1.62\ldots$. Numerical work seems to be in agreement with this conjecture. In the report [5] we gave an alternative proof of (2) and in fact showed that

$$1.20 N^{3/2} < S(N) < 2.33 N^{3/2}$$

for N sufficiently large. We are now able to refine this estimate.

Theorem 1. *For N sufficiently large*

$$1.35 N^{3/2} < S(N) < 2.04 N^{3/2}.$$

Our proof of Theorem 1 depends on two results of R.R. Hall [2] on the distribution and the second moments of gaps in the Farey series.

2 Preliminary Lemmas

Let N be a positive integer and let $F_N = \{x_1, \ldots, x_R\}$ where $0 < x_1 < \cdots < x_R = 1$. Put $\ell_1 = x_1$ and $\ell_r = x_r - x_{r-1}$ for $r = 2, \ldots, R$ so that the ℓ_i's correspond to gaps in the Farey series with the points 0 and 1 identified.

Lemma 1. *There is a positive number C_0 such that for $N \geq 2$,*

$$\sum_{r=1}^{R} \ell_r^2 < (C_0 \log N)/N^2.$$

Proof. This follows from Theorem 1 of [2]. □

For each positive real number t and each positive integer N we define $\sigma_N(t)$ to be the number of gaps ℓ_r for which $\ell_r > t/N^2$. Thus

$$\sigma_N(t) = \sum_{\substack{r=1 \\ t<N^2\ell_r}}^{R} 1.$$

We also define $\delta_N(t)$ by

$$\delta_N(t) = \sigma_N(t)/R(N).$$

Then $\delta_N(t)$ is a distribution function and Hall [2] proves that $\delta_N(t)$ tends to a limit as N tends to infinity.

Lemma 2. *If* $4 \le t \le N$ *and* $w = w(t)$ *is the smaller root of the equation* $w^2 = t(w-1)$ *then*

$$\delta_N(t) = 2t^{-1}(1 - w + 2\log w) + O(t^{-1}N^{-1}\log N + N^{-3/2}).$$

If $1 \le t \le 4$ *then*

$$\delta_N(t) = 2t^{-1}\left(1 + \log t - \frac{t}{2}\right) + O(N^{-1}\log N).$$

Proof. The first assertion follows from Theorem 4 of [2] together with (1). The second assertion follows from (1.2) of [2]. □

Let us define $f(t)$ for $1 \le t$ by

$$f(t) = \begin{cases} 2\left(1 + \log t - \frac{t}{2}\right) & \text{for } 1 \le t \le 4 \\ 2(1 - w + 2\log w) & \text{for } 4 < t \end{cases} \tag{3}$$

where

$$w = \frac{t}{2}\left(1 - \left(1 - \frac{4}{t}\right)^{1/2}\right) \quad \text{for } 4 < t.$$

Observe that

$$\lim_{t\to\infty} f(t)/(2/t) = 1. \tag{4}$$

Lemma 3. *For* $4 \le t \le N$ *we have*

$$\sigma_N(t) \le \frac{24(2\log 2 - 1)}{\pi^2}\left(\frac{N}{t}\right)^2 + O\left(\frac{N}{t}\log N + N^{1/2}\right).$$

Proof. Since $\sigma_N(t) = R(N)\delta_N(t)$ it suffices, by (1) and Lemma 2 to show that for $t \ge 4$, $g(t)$ is a decreasing function of t where

$$g(t) = t(2\log w(t) - (w(t) - 1)).$$

Since

$$w(t) = \left(t - t\,(1 - 4/t)^{1/2}\right)/2$$

we find that

$$g'(t) = 2\log w - (w-1) + ((2/w) - 1)tw'(t)$$

so

$$g'(t) = 2\log w - 2w + 2.$$

On observing that $\log(1+x) \le x$ for $x \ge 0$ and putting $x = w - 1$ we conclude that

$$g'(t) \le 2(w-1) - 2w + 2 = 0$$

whenever $w \ge 1$. Since, for $t > 4$,

$$w(t) = 1 + \frac{1}{t} + \frac{2}{t^2} + \cdots + \frac{c_n}{t^n} + \cdots$$

where the c_n are positive numbers we see that $w > 1$ for $t > 4$ hence for $t \ge 4$. Thus $g(t)$ is a decreasing function of t as required. □

3 Further Preliminaries

For each positive integer M we define $\theta(M)$ to be the number of q_i's in the sum giving $S(N)$ which are larger than M. Thus

$$\theta(M) = \sum_{\substack{i=1 \\ q_i > M}}^{N} 1.$$

For positive integers j and M let $\psi(j)$ $(= \psi_M(j))$ denote the number of gaps ℓ_r in F_M of size larger than $\frac{j}{N}$. Accordingly we have

$$\psi(j) = \sum_{\substack{r=1 \\ \ell_r > \frac{j}{N}}}^{R(M)} 1.$$

A gap ℓ_r in F_M with $\ell_r \le \frac{j+1}{N}$ properly contains at most j intervals $\left(\frac{h-1}{N}, \frac{h}{N}\right]$ with $1 \le h \le N$. $\theta(M)$ is the total number of intervals $\left(\frac{h-1}{N}, \frac{h}{N}\right]$ which are properly contained in gaps of F_M. Thus

$$\theta(M) \le \psi(1) + \psi(2) + \cdots.$$

Similarly a gap ℓ_r in F_M with $\ell_r > \frac{j+1}{N}$ properly contains at least j intervals of the form $\left(\frac{h-1}{N}, \frac{h}{N}\right]$. Therefore

$$\psi(2) + \psi(3) + \cdots \leq \theta(M).$$

Since $\psi(j) = \sigma_M\left(\frac{jM^2}{N}\right)$, it follows that

$$\sum_{j=2}^{v} \sigma_M\left(\frac{jM^2}{N}\right) \leq \theta(M) \leq \sum_{j=1}^{v} \sigma_M\left(\frac{jM^2}{N}\right), \tag{5}$$

where v $(= v(M))$ satisfies

$$v < \frac{N}{M} \leq v + 1. \tag{6}$$

Let u_1 be the number of rationals $\frac{h}{k}$ with $(h, k) = 1$ and $1 \leq h \leq k \leq \sqrt{N}$. Then by (1)

$$u_1 = \frac{3}{\pi^2} N + O(N^{1/2} \log N) \tag{7}$$

and the sum S_1 of the denominators of these rationals is

$$S_1 = \sum_{k \leq \sqrt{N}} k\varphi(k).$$

By Abel summation and (1) we find that

$$S_1 = \frac{2}{\pi^2} N^{3/2} + O(N \log N). \tag{8}$$

Observe that if q is an integer with $1 \leq q \leq \sqrt{N}$ then each rational p/q with p positive and coprime with q contributes a term q to $S(N)$. Thus S_1 is the sum of the u_1 smallest terms in the sum giving $S(N)$. Put

$$u_2 = N - u_1 \tag{9}$$

and let S_2 be the sum of the u_2 largest q's which appear in the sum for $S(N)$. Then

$$S(N) = S_1 + S_2. \tag{10}$$

4 The Upper Bound in Theorem 1

In order to establish an upper bound for $S(N)$ we shall establish an upper bound for S_2 and then appeal to (8) and (10).

For any positive integer M with $M \leq N$ we have

$$S_2 \leq Mu_2 + \theta(M) + \theta(M+1) + \cdots + \theta(N). \tag{11}$$

Put $\lambda = 1.38$ and $M_1 = [\lambda N^{1/2}]$. Since $\lambda(1 - 3/\pi^2) < 0.96054$ and $\theta(M_1) \leq N$, it follows from (7), (9) and (11) that

$$S_2 < 0.96054 N^{3/2} + \theta(M_1 + 1) + \theta(M_1 + 2) + \cdots + \theta(N) \tag{12}$$

for N sufficiently large. Next, put

$$S_3 = \sum_{M_1 < M < N^{3/5}} \theta(M) \quad \text{and} \quad S_4 = \sum_{N^{3/5} \leq M \leq N} \theta(M).$$

Thus, by (12),

$$S_2 < 0.96054\, N^{3/2} + S_3 + S_4. \tag{13}$$

Let us first estimate S_4. To that end recall that $\theta(M)$ is the number of q_i's in the sum $S(N)$ which are larger than M. Thus there are $\theta(M)$ intervals $\left(\frac{j-1}{N}, \frac{j}{N}\right]$ which contain no element of F_M. In particular there must exist differences $\ell_{r_1}, \ldots, \ell_{r_s}$ in F_M for which we can find positive integers $k_1, \ldots, k_s$ with $\ell_{r_i} \geq k_i/N$ for $i = 1, \ldots, s$ and such that $k_1 + \cdots + k_s \geq \theta(M)$. Thus we certainly have

$$\sum_{i=1}^{s} \ell_{r_i}^2 \geq \frac{\theta(M)}{N^2}. \tag{14}$$

On the other hand, by Lemma 1,

$$\sum_{r=1}^{R(M)} \ell_r^2 < C_0 M^{-2} \log M. \tag{15}$$

A comparison of (14) and (15) reveals that

$$\theta(M) < C_0 \frac{N^2}{M^2} \log M.$$

For $N^{3/5} \leq M \leq N$ we have $\log M \leq \log N$ hence

$$\sum_{N^{3/5} \leq M \leq N} \theta(M) < C_0 N^2 \log N \int_{N^{3/5}-1}^{N} \frac{dM}{M^2}$$

so

$$S_4 < 2C_0 N^{7/5} \log N. \tag{16}$$

Next we estimate S_3. By (5)

$$S_3 = \sum_{M_1 < M < N^{3/5}} \theta(M) \leq \sum_{M_1 < M < N^{3/5}} \sum_{j=1}^{v} \sigma_M \left(\frac{jM^2}{N} \right). \tag{17}$$

For $M < N^{3/5}$ we see from (6) that $v + 1$ is at least $N^{2/5}$, which in turn exceeds 10^4 for N sufficiently large. Then, by Lemma 3,

$$\begin{aligned}
\sum_{M_1 < M < N^{3/5}} \sum_{10^4 < j \leq v} \sigma_M \left(\frac{jM^2}{N} \right) &< \sum_{M_1 < M < N^{3/5}} \frac{N^2}{M^2} \sum_{10^4 < j < \infty} \left(\frac{1}{j} \right)^2 \\
&< 10^{-4} N^2 \sum_{M_1 < M < N^{3/5}} \frac{1}{M^2} \\
&< 10^{-4} N^{3/2},
\end{aligned} \tag{18}$$

for N sufficiently large. Accordingly by (17) and (18)

$$S_3 < 10^{-4} N^{3/2} + \sum_{M_1 < M < N^{3/5}} \sum_{j=1}^{10^4} \sigma_M \left(\frac{jM^2}{N} \right). \tag{19}$$

Let $\varepsilon > 0$. For N sufficiently large in terms of ε

$$R(M) < \left(\frac{3}{\pi^2} + \varepsilon \right) M^2$$

hence

$$\sigma_M \left(\frac{jM^2}{N} \right) = R(M) \delta_M \left(\frac{jM^2}{N} \right) < \left(\frac{3}{\pi^2} + \varepsilon \right) M^2 \delta_M \left(\frac{jM^2}{N} \right)$$

and so

$$\sigma_M \left(\frac{jM^2}{N} \right) < \left(\frac{3}{\pi^2} + \varepsilon \right) \frac{N}{j} \left(\frac{jM^2}{N} \delta_M \left(\frac{jM^2}{N} \right) \right). \tag{20}$$

It follows from Lemma 2 and (3) that for $j \leq 10^4$ and $M \leq N^{3/5}$

$$\frac{jM^2}{N} \delta_M \left(\frac{jM^2}{N} \right) = f \left(\frac{jM^2}{N} \right) + O \left(\frac{\log N}{N} \right).$$

Thus, by (4), for N sufficiently large in terms of ε

$$\frac{jM^2}{N}\delta_M\left(\frac{jM^2}{N}\right) < (1+\varepsilon) f\left(\frac{jM^2}{N}\right). \tag{21}$$

For each integer j with $1 \le j \le 10^4$ we find from (20) and (21) that

$$\sum_{M_1<M<N^{3/5}} \sigma_M\left(\frac{jM^2}{N}\right) < \left(\frac{3}{\pi^2}+\varepsilon\right)(1+\varepsilon)\frac{N}{j}\sum_{M_1<M<N^{3/5}} f\left(\frac{jM^2}{N}\right). \tag{22}$$

The function f is continuous and it is increasing on $(1, 4)$ and decreasing on $(4, \infty)$. Accordingly, with $\Delta = 1/\log N$, we have

$$\sum_{M_1<M<N^{3/5}} f\left(\frac{jM^2}{N}\right)$$

$$< \left(\sum_{1\le k<(N^{3/5}-M_1)/[\Delta\sqrt{N}]} f\left(\frac{j(M_1+k[\Delta\sqrt{N}])^2}{N}\right)[\Delta\sqrt{N}]\right) + O\left(\frac{\sqrt{N}}{\log N}\right)$$

which is, for N sufficiently large,

$$< \left(\sum_{1\le k<N^{1/5}} f\left(\frac{j(\lambda\sqrt{N}+O(1)+k(\Delta\sqrt{N}+O(1)))^2}{N}\right)(\Delta\sqrt{N}+O(1))\right) + O\left(\frac{\sqrt{N}}{\log N}\right).$$

Therefore, for N sufficiently large in terms of ε,

$$\sum_{M_1<M<N^{3/5}} f\left(\frac{jM^2}{N}\right) < (1+\varepsilon)N^{1/2}\sum_{1\le k<N^{1/5}} f\left(j(\lambda+k\Delta)^2 + O\left(k^2N^{-1/2}\right)\right)\cdot\Delta$$

$$< (1+\varepsilon)^2N^{1/2}\int_\lambda^\infty f(jt^2)dt. \tag{23}$$

Thus, by (22) and (23),

$$\sum_{j=1}^{10^4}\sum_{M_1<M<N^{3/5}} \sigma_M\left(\frac{jM^2}{N}\right)$$
$$< \left(\frac{3}{\pi^2}+\varepsilon\right)(1+\varepsilon)^3N^{3/2}\sum_{j=1}^{10^4}\frac{1}{j}\int_\lambda^\infty f(jt^2)dt. \tag{24}$$

Evaluating with MAPLE we find that

$$\sum_{j=1}^{10^4} \frac{1}{j} \int_{\lambda}^{\infty} f(jt^2)dt < 2.8640. \tag{25}$$

Therefore, by (24) and (25), for N sufficiently large,

$$\sum_{j=1}^{10^4} \sum_{M_1 < M < N^{3/5}} \sigma_M \left(\frac{jM^2}{N}\right) < 0.8706\, N^{3/2}. \tag{26}$$

By (19) and (26)

$$S_3 < 0.8707\, N^{3/2} \tag{27}$$

for N sufficiently large. Further, by (13), (16) and (27),

$$S_2 < 1.8313\, N^{3/2}$$

for N sufficiently large. Our result now follows from (8) and (10).

5 The Lower Bound in Theorem 1

The value of the smallest q_i in S_2 exceeds $\sqrt{N}$ and so

$$S_2 \geq [\sqrt{N}]u_2 + \theta([\sqrt{N}]) + \theta([\sqrt{N}] + 1) + \cdots + \theta(N)$$

hence, by (7) and (9),

$$S_2 \geq \left(1 - \frac{3}{\pi^2}\right) N^{3/2} + O(N \log N) + \theta([\sqrt{N}]) + \cdots + \theta(N). \tag{28}$$

Certainly

$$\theta([\sqrt{N}]) + \cdots + \theta(N) \geq \sum_{N^{1/2} < M < N^{3/5}} \theta(M)$$

and for M with $M < N^{3/5}$ we see from (6) that $v + 1$ is at least $N^{2/5}$. Therefore, by (5), for N sufficiently large

$$\sum_{N^{1/2} < M < N^{3/5}} \theta(M) > \sum_{N^{1/2} < M < N^{3/5}} \sum_{j=2}^{10^4} \sigma_M \left(\frac{jM^2}{N}\right)$$

and so, by (28),

$$S_2 > \left(1 - \frac{3}{\pi^2}\right) N^{3/2} + O(N \log N) + \sum_{j=2}^{10^4} \sum_{N^{1/2} < M < N^{3/5}} \sigma_M\left(\frac{jM^2}{N}\right). \tag{29}$$

We shall now estimate the double sum in (29). Let $\varepsilon > 0$. For N sufficiently large in terms of ε

$$R(M) > \left(\frac{3}{\pi^2} - \varepsilon\right) M^2$$

hence

$$\sigma_M\left(\frac{jM^2}{N}\right) = R(M)\delta_M\left(\frac{jM^2}{N}\right) > \left(\frac{3}{\pi^2} - \varepsilon\right) M^2 \delta_M\left(\frac{jM^2}{N}\right)$$

and so

$$\sigma_M\left(\frac{jM^2}{N}\right) > \left(\frac{3}{\pi^2} - \varepsilon\right) \frac{N}{j}\left(\frac{jM^2}{N}\delta_M\left(\frac{jM^2}{N}\right)\right). \tag{30}$$

It follows from Lemma 2 and (3) that for $j \le 10^4$ and $M \le N^{3/5}$

$$\frac{jM^2}{N}\delta_M\left(\frac{jM^2}{N}\right) = f\left(\frac{jM^2}{N}\right) + O\left(\frac{\log N}{N}\right).$$

Thus, by (4), for N sufficiently large in terms of ε

$$\frac{jM^2}{N}\delta_M\left(\frac{jM^2}{N}\right) > (1-\varepsilon) f\left(\frac{jM^2}{N}\right). \tag{31}$$

For each integer j with $2 \le j \le 10^4$ we find from (30) and (31) that

$$\begin{aligned} &\sum_{N^{1/2} < M < N^{3/5}} \sigma_M\left(\frac{jM^2}{N}\right) \\ &\qquad > \left(\frac{3}{\pi^2} - \varepsilon\right)(1-\varepsilon)\frac{N}{j} \sum_{N^{1/2} < M < N^{3/5}} f\left(\frac{jM^2}{N}\right). \end{aligned} \tag{32}$$

The function f is continuous and it is increasing on $(1, 4)$ and decreasing on $(4, \infty)$. Accordingly, with $\Delta = 1/\log N$, we have

$$\sum_{N^{1/2}<M<N^{3/5}} f\left(\frac{jM^2}{N}\right)$$
$$\geq \left(\sum_{1\leq k<(N^{3/5}-N^{1/2})/[\Delta\sqrt{N}]} f\left(\frac{j([\sqrt{N}]+k[\Delta\sqrt{N}])^2}{N}\right)[\Delta\sqrt{N}]\right) + O\left(\frac{\sqrt{N}}{\log N}\right)$$

which is, for N sufficiently large,

$$\geq \left(\sum_{1\leq k<N^{1/10}} f\left(\frac{j(\sqrt{N}+O(1)+k(\Delta\sqrt{N}+O(1)))^2}{N}\right)(\Delta\sqrt{N}+O(1))\right) + O\left(\frac{\sqrt{N}}{\log N}\right).$$

Therefore, for N sufficiently large in terms of ε,

$$\sum_{N^{1/2}<M<N^{3/5}} f\left(\frac{jM^2}{N}\right) > (1-\varepsilon)N^{1/2} \sum_{1\leq k<N^{1/10}} f(j(1+k\Delta)^2 + O(k^2N^{-1/2}))\cdot\Delta$$
$$> (1-\varepsilon)^2 N^{1/2}\int_1^\infty f(jt^2)dt. \tag{33}$$

Thus, by (32) and (33),

$$\sum_{j=2}^{10^4}{}^{\cdot} \sum_{N^{1/2}<M<N^{3/5}} \sigma_m\left(\frac{jM^2}{N}\right)$$
$$> \left(\frac{3}{\pi^2}-\varepsilon\right)(1-\varepsilon)^3 N^{3/2}\sum_{j=2}^{10^4}\frac{1}{j}\int_1^\infty f(jt^2)dt. \tag{34}$$

Evaluating with MAPLE we find that

$$\sum_{j=2}^{10^4}\frac{1}{j}\int_1^\infty f(jt^2)dt > 1.5098. \tag{35}$$

Therefore by (34) and (35), for N sufficiently large

$$\sum_{j=2}^{10^4}\sum_{N^{1/2}<M<N^{3/5}} \sigma_M\left(\frac{jM^2}{N}\right) > 0.4589\, N^{3/2}. \tag{36}$$

By (8), (10), (29) and (36) we see that

$$S(N) > \left(1 - \frac{1}{\pi^2} + 0.458\right) N^{3/2} > 1.35\, N^{3/2}$$

for N sufficiently large and the result now follows.

References

1. C. Cobeli, A. Zaharescu, The Haros-Farey sequence at two hundred years, a survey, *Acta Universitatis Apulensis* **5** (2003), 1–38.
2. R.R. Hall, A note on the Farey series, *J. London Math. Soc.* **2** (1970), 139–148.
3. G.H. Hardy, E.M. Wright, *An introduction to the theory of numbers*, 5th ed., Oxford, 1979.
4. D. Kruyswijk, H.G. Meijer, On small denominators and Farey sequences, *Indagtiones Math.* **39** (1977), 332–337.
5. C.L. Stewart, On a sum associated with the Farey series, Report ZW88, *Mathematisch Centrum* (1976), 1–11.