COLLOQUIUM MATHEMATICUM

VOL. 123 2011 NO. 1

ON THE DISTRIBUTION OF THE PARTIAL SUM OF EULER'S TOTIENT FUNCTION IN RESIDUE CLASSES

BY

YOUNESS LAMZOURI, M. TIP PHAOVIBUL and ALEXANDRU ZAHARESCU (Urbana, IL)

Abstract. We investigate the distribution of $\Phi(n) = 1+\sum_{i=1}^{n} \varphi(i)$ (which counts the number of Farey fractions of order n) in residue classes. While numerical computations suggest that $\Phi(n)$ is equidistributed modulo q if q is odd, and is equidistributed modulo the odd residue classes modulo q when q is even, we prove that the set of integers n such that $\Phi(n)$ lies in these residue classes has a positive lower density when $q = 3, 4$. We also provide a simple proof, based on the Selberg–Delange method, of a result of T. Dence and C. Pomerance on the distribution of $\varphi(n)$ modulo 3.

1. Introduction. Let φ denote Euler's totient function, which counts the number of positive integers less than n that are coprime to n. Define

$$\Phi(n) := 1 + \sum_{i=1}^{n} \varphi(i).$$

Then $\Phi(n)$ is the number of Farey fractions of order n, which also corresponds to the number of lattice points (x, y) with $0 \leq x \leq y \leq n$ that are visible from the origin. C. Pomerance gave an outline in [8] (see exercise 20 page 145) of the proof that there are infinitely many values of $\Phi(n)$ in every residue class modulo 3. His idea is to exploit the fact that the Dirichlet series $L(s) := \sum_{n=1}^{\infty} \chi_3(\varphi(n))/n^s$ has a pole at $s = 1$, where χ_3 is the unique non-principal character modulo 3. This gives a motivation to study the distribution of $\Phi(n)$ modulo 3, and more generally one can ask for an asymptotic formula for the number of positive integers $n \leq x$ such that $\Phi(n) \equiv k \bmod q$. In this paper, we investigate this question in the cases $q = 3$ and $q = 4$. Note that $\Phi(n)$ is odd for all $n \geq 2$. We define

$$\mathcal{A}_k := \{n \geq 1 : \Phi(n) \equiv k \bmod 3\}, \quad \mathcal{B}_j := \{n \geq 1 : \Phi(n) \equiv j \bmod 4\}.$$

Moreover, if $A \subseteq \mathbb{N}$, we denote by $|A(x)|$ the number of positive integers $j \leq x$ with $j \in A$.

2010 *Mathematics Subject Classification*: Primary 11N69; Secondary 11N64.
Key words and phrases: Euler's totient function, distribution in residue classes.

DOI: 10.4064/cm123-1-8

Table 1

x	$\lvert\mathcal{A}_0(x)\rvert$	$\lvert\mathcal{A}_1(x)\rvert$	$\lvert\mathcal{A}_2(x)\rvert$
10^3	298	337	365
10^4	3400	3290	3310
10^5	33260	33283	33457
10^6	332701	333142	334157
10^7	3333156	3334029	3332815
10^8	33332106	33325232	33342662

Table 2

x	$\lvert\mathcal{B}_1(x)\rvert$	$\lvert\mathcal{B}_3(x)\rvert$
10^3	475	524
10^4	5162	4837
10^5	49703	50296
10^6	497269	502730
10^7	4988042	5011957
10^8	49990438	50009561

Tables 1 and 2 suggest that $\Phi(n)$ is equidistributed among the three residue classes modulo 3, and among the classes 1 and 3 modulo 4. One can also remark that the convergence seems to be very fast, and that the error term tends to be approximately of the size of the square root of the main term. To further investigate the distribution of $\Phi(n)$ in residue classes, we have performed numerical computations for all moduli $3 \le q \le 100$ (for x up to 10^7) and noticed that a similar phenomenon occurs. Indeed these numerical investigations may suggest that for all $q \ge 3$, $\Phi(n)$ is equidistributed among the residue classes in $C(q)$ where

$$C(q) := \begin{cases} \{a \bmod q\} & \text{if } q \text{ is odd,} \\ \{a \bmod q : (a,2)=1\} & \text{if } q \text{ is even.} \end{cases}$$

We define

$$E_q(x) := \max_{j\in C(q)} \left| \frac{\{n \le x : \Phi(n) \equiv j \bmod q\}}{x} - \frac{1}{|C(q)|} \right|.$$

Table 3 below contains values of $E_q(x)$ for $3 \le q \le 10$, and x up to 10^8.

Table 3

x	$E_3(x)$	$E_4(x)$	$E_5(x)$	$E_6(x)$
10^2	.033333	.090000	.060000	.033333
10^3	.035000	.025000	.031000	.035333
10^4	.006667	.016300	.004600	.006667
10^5	.001237	.002970	.002180	.001237
10^6	.000823	.002731	.000601	.000632
10^7	.000007	.001195	.000169	.000069
10^8	.000093	.000096	.000061	.000093

x	$E_7(x)$	$E_8(x)$	$E_9(x)$	$E_{10}(x)$
10^2	.057142	.090000	.058889	.060000
10^3	.015857	.034000	.019888	.031000
10^4	.004842	.011500	.006911	.004600
10^5	.001277	.004260	.002689	.002180
10^6	.000670	.001887	.000843	.000601
10^7	.000220	.000931	.000145	.000170
10^8	.000084	.000181	.000084	.000061

In 1909, E. Landau [5] proved that the number of integers $n \le x$ having all prime divisors in r residue classes modulo q (with $r < \phi(q)$) is asymptotic to

$$C(r,q)\frac{x}{\log^{1-r/\phi(q)} x} \quad \text{as } x \to \infty, \tag{1.1}$$

where $C(r,q)$ is a positive constant. Since the condition $q \nmid \varphi(n)$ implies that n has no prime divisors in the residue class $1 \bmod q$, it follows from Landau's result that $\varphi(n)$ is divisible by q for almost all integers n. This shows that $\Phi(n)$ stays constant modulo q for a large proportion of the time, then it changes precisely at those integers n such that $q \nmid \varphi(n)$. If $(q,6)=1$, W. Narkiewicz [7] showed that $\varphi(n)$ is equidistributed among the residue classes relatively prime to q. However, as we shall see later, the distribution of $\varphi(n)$ in residue classes modulo 3 and 4 has a different behavior. Indeed we shall prove that $\varphi(n)$ has more values that are congruent to $1 \bmod 3$ than to $2 \bmod 3$, and $4 \nmid \varphi(n)$ implies that $\varphi(n) \equiv 2 \bmod 4$ for all $n \geq 3$. Exploiting these irregularities and using sieve theory, we show that the sets $\mathcal{A}_k$ and $\mathcal{B}_j$ have positive lower densities for $k=0,1,2$ and $j=1,3$. We should also note that our idea would not work in general, since such irregularities do not exist when the modulus q is coprime to 6, by the result of Narkiewicz.

Theorem 1. *For $j=1,3$ we have*

$$\liminf_{x\to\infty} \frac{|\mathcal{B}_j(x)|}{x} \geq \delta_1,$$

where $\delta_1 = 9/896 > 1/100$.

Remark 1. The poor value of δ_1 is not only due to the use of sieve theory, but also to the difficulty of understanding the gaps between consecutive primes. Indeed, for $n \geq 5$, Lemma 2.1 below shows that $\Phi(n) \not\equiv \Phi(n-1) \bmod 4$ if and only if $n = p^k$ or $n = 2p^k$, where p is a prime congruent to 3 modulo 4.

In the case where $q=3$ we prove

Theorem 2. *For $k=0,1,2$ we have*

$$\liminf_{x\to\infty} \frac{|\mathcal{A}_k(x)|}{x} \geq \delta_2$$

for some $\delta_2 > 0$.

Remark 2. An explicit computation allows one to take

$$\delta_2 \approx 0.0003159363.$$

The proof of Theorem 2 relies on understanding the distribution of $\varphi(n)$ in the residue classes 1 and 2 modulo 3. For $i=1,2$ let

$$N_i(x) := \{n \leq x : \varphi(n) \equiv i \bmod 3\}.$$

In [2], T. Dence and C. Pomerance proved an asymptotic formula for $|N_i(x)|$ using a combinatorial argument along with Landau's result (1.1) and Wirsing's theorem (see [10]) on mean values of multiplicative functions. Using a direct approach based on the Selberg–Delange method we provide a simpler proof of the result from [2]. Moreover, we can also exhibit lower order terms in the asymptotics of $|N_i(x)|$. We have

THEOREM 3. *Let x be large and $K \geq 2$ be a positive integer. Then there exist explicit constants λ_j, β_j for $j = 1, \ldots, K-1$ such that*

$$|N_1(x)| = \lambda \frac{x}{\sqrt{\log x}} + \sum_{j=1}^{K-1} \lambda_j \frac{x}{\log^{j+1/2} x} + O_K\left(\frac{x}{\log^{K+1/2} x}\right),$$

$$|N_2(x)| = \beta \frac{x}{\sqrt{\log x}} + \sum_{j=1}^{K-1} \beta_j \frac{x}{\log^{j+1/2} x} + O_K\left(\frac{x}{\log^{K+1/2} x}\right),$$

where

$$\lambda = \frac{2^{3/2}}{3^{3/4}\pi} \prod_{p \equiv 2 \bmod 3} \left(1 - \frac{1}{p^2}\right)^{-1/2} \left(1 + \frac{1}{2} \prod_{p \equiv 2 \bmod 3} \left(1 - \frac{2}{p(p+1)}\right)\right),$$

and β is given by the same expression as for λ, except that the factor $1 + \frac{1}{2}\prod_{p\equiv 2 \bmod 3}\left(1 - \frac{2}{p(p+1)}\right)$ is replaced by $1 - \frac{1}{2}\prod_{p\equiv 2 \bmod 3}\left(1 - \frac{2}{p(p+1)}\right)$. Moreover $\lambda \approx 0.6109136202$ and $\beta \approx 0.3284176245$.

The asymptotic expansion in Theorem 3 can also be obtained along the lines provided by J. Kaczorowski in [4].

2. Preliminary lemmas. First we characterize the values of n for which $\Phi(n) \not\equiv \Phi(n-1) \bmod q$, when $q = 3, 4$.

LEMMA 2.1. *Let $n \geq 5$ be a positive integer. Then $\Phi(n) \not\equiv \Phi(n-1) \bmod 4$ if and only if $n = p^k$ or $n = 2p^k$, where p is a prime congruent to 3 modulo 4 and k is a positive integer. Moreover $\Phi(n) \not\equiv \Phi(n-1) \bmod 3$ if and only if $n = 3^a m$ where $a = 0$ or 1 and m is divisible only by primes that are congruent to 2 modulo 3.*

Proof. Write $n = 2^{\alpha_0} p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_l^{\alpha_l}$ where the p_i are odd primes and the α_i are non-negative integers. Then $\varphi(n) = 2^{\alpha_0 - 1} \prod_{i=1}^{l} (p_i - 1) p_i^{\alpha_i - 1}$. Hence, if $\alpha_0 \geq 2$, $p_i \equiv 1 \bmod 4$ for some $1 \leq i \leq l$, or $l \geq 2$ then $\varphi(n) \equiv 0 \bmod 4$. Conversely if $n = p^k$ or $n = 2p^k$ where p is a prime congruent to 3 modulo 4 then $\varphi(n) = p^{k-1}(p-1) \equiv 2 \bmod 4$.

Similarly writing $n = 3^a m$ with $3 \nmid m$ one can see that $3 \mid \varphi(n)$ if and only if $a \geq 2$ or m is divisible by a prime $p \equiv 1 \bmod 3$. ■

In order to prove Theorems 1 and 2 we shall need the following applications of Selberg's upper bound sieve.

LEMMA 2.2. *Let a be a fixed non-negative integer and d be a positive integer such that d is even if $a = 0$, and d is odd if $a \geq 1$. Then as $x \to \infty$*

we have, uniformly in d,

$$|\{p \leq x : p \equiv 3 \bmod 4,\ 2^a p + d \text{ is prime}\}| \leq (4 + o(1)) \prod_{p>2} \left(1 - \frac{1}{(p-1)^2}\right) f_1(d) \frac{x}{\log^2 x},$$

where $f_1(d)$ is the multiplicative function satisfying $f_1(p^k) = \frac{p-1}{p-2}$ if $p > 2$ and $f_1(2^k) = 1$, for all positive integers k.

Proof. This is a corollary of Theorem 3.12 of [3]. ∎

LEMMA 2.3. *Let x be large and $1 \leq d \leq \log^2 x$ be a positive integer. Then as $x \to \infty$ we have*

$$|\{n \leq x : p \mid n(n+d) \Rightarrow p \not\equiv 1 \bmod 3\}| \leq (2 + o(1)) \prod_{p} \left(1 - \frac{w(p)}{p}\right)\left(1 - \frac{1}{p}\right)^{-1} f_2(d) \frac{x}{\log x},$$

where $w(p) = 2$ if $p \equiv 1 \bmod 3$ and $w(p) = 0$ otherwise. Moreover $f_2(d)$ is the multiplicative function defined by $f_2(p^k) = \frac{p-1}{p-2}$ if $p \equiv 1 \bmod 3$ and $f_2(p^k) = 1$ otherwise, for any positive integer k.

Proof. This follows from Theorem 5.1 of [3] by taking $\kappa = 1$ and $L = 2 \log\log x$ there. ∎

Lastly we prove estimates for mean values of the multiplicative functions f_1 and f_2 that arise in the sieve bounds of Lemmas 2.2 and 2.3.

LEMMA 2.4. *Let $\mathcal{P}$ be a set of odd prime numbers, and d be a positive integer divisible only by primes $p \notin \mathcal{P}$. Let f be the multiplicative function defined by*

$$f(p^k) = \begin{cases} \dfrac{p-1}{p-2} & \textit{if } p \in \mathcal{P}, \\ 1 & \textit{if } p \notin \mathcal{P}, \end{cases}$$

for any positive integer k. Then for any $\epsilon > 0$ we have

$$\sum_{\substack{n \leq x \\ d \mid n}} f(n) = \frac{x}{d} \prod_{p \in \mathcal{P}} \left(1 - \frac{1}{(p-1)^2}\right)^{-1} + O_{\epsilon,d}(x^{\epsilon}).$$

Proof. First note that

$$f(n) = \prod_{\substack{p \mid n \\ p \in \mathcal{P}}} \frac{p-1}{p-2},$$

where the product equals 1 if n is divisible only by primes $p \notin \mathcal{P}$. If $d \mid n$

then writing $n = dm$ we deduce that $f(n) = f(m)$. Hence we get

$$\sum_{\substack{n\le x\\ d|n}} f(n) = \sum_{m\le x/d} f(m). \tag{2.1}$$

Let h be the multiplicative function defined by $h = f * \mu$, where $*$ is the Dirichlet convolution and μ is the Möbius function. Then one can check that $h(p) = f(p)-1 = \frac{1}{p-2}$ for $p \in \mathcal{P}$ and $h(p) = 0$ otherwise. Moreover for prime p and $k \ge 2$ we have $h(p^k) = f(p^k) - f(p^{k-1}) = 0$. Let $y = x/d$. Then we obtain

$$\begin{aligned}\sum_{m\le y} f(m) &= \sum_{m\le y}\sum_{r|m} h(r) = \sum_{r\le y} h(r)\left[\frac{y}{r}\right] \\ &= y\sum_{r\le y}\frac{h(r)}{r} + O\biggl(\sum_{r\le y} h(r)\biggr).\end{aligned} \tag{2.2}$$

The error term on the RHS of the above estimate is

$$\ll \prod_{p\le y}(1+h(p)) \le \prod_{p\le y}\left(1+\frac{1}{p-2}\right) \ll \log y. \tag{2.3}$$

Moreover, for any $\epsilon > 0$, the series

$$\sum_{r=1}^{\infty}\frac{h(r)}{r^{\epsilon}} = \prod_{p\in\mathcal{P}}\left(1+\frac{1}{p^{\epsilon}(p-2)}\right)$$

is absolutely convergent. This shows that

$$\sum_{r>y}\frac{h(r)}{r} \le \frac{1}{y^{1-\epsilon}}\sum_{r=1}^{\infty}\frac{h(r)}{r^{\epsilon}} \ll_{\epsilon} y^{-1+\epsilon},$$

which implies

$$\sum_{r\le y}\frac{h(r)}{r} = \prod_{p\in\mathcal{P}}\left(1-\frac{1}{(p-1)^2}\right)^{-1} + O_{\epsilon}(y^{-1+\epsilon}).$$

Thus, the result follows upon combining this estimate with (2.1)–(2.3). ■

3. Proof of Theorem 1. Let $M(x)$ be the set of positive integers $5 \le n \le x$ such that $4 \nmid \varphi(n)$. Then write $M(x) = \{b_1, \ldots, b_m\}$ with $5 \le b_1 < \cdots < b_m \le x$ where $m = |M(x)|$, and set $b_0 = 5$ and $b_{m+1} = [x]$. Using Lemma 2.1 along with the prime number theorem for arithmetic progressions we obtain

$$\begin{aligned}|M(x)| &= (\pi(x;4,3) + \pi(x/2;4,3)) + O(\sqrt{x}) \\ &= \frac{3x}{4\log x} + O\biggl(\frac{x}{\log^2 x}\biggr).\end{aligned} \tag{3.1}$$

Put $r = [|M(x)|/2]$ and let $L \leq \log x$ be a positive real number to be chosen later. Furthermore, define $T_d(x) = |\{0 \leq i \leq r : b_{2i+1} - b_{2i} = d\}|$ for all positive integers $d \geq 1$. Hence, we infer from (3.1) that

$$|\mathcal{B}_3(x)| \geq \sum_{i=0}^{r}(b_{2i+1} - b_{2i}) = \sum_{d\geq 1} dT_d(x) \geq L\Big(r - \sum_{d\leq L} T_d(x)\Big) \geq L\bigg(\frac{3x}{8\log x} - \sum_{d\leq L} T_d(x)\bigg) + O\bigg(\frac{Lx}{\log^2 x}\bigg). \tag{3.2}$$

What remains is to obtain a good upper bound for $\sum_{d<L} T_d(x)$. Let us define $K(x) = \{1 \leq a \leq x : a = p \text{ or } a = 2p \text{ where } p \equiv 3 \bmod 4\}$ and let $K_d(x) = |\{(k_1, k_2) \in K^2(x) : k_1 < k_2 \text{ and } k_2 - k_1 = d\}|$. Then

$$\sum_{d\leq L} T_d(x) \leq \sum_{d\leq L} K_d(x) + O(L\sqrt{x}). \tag{3.3}$$

This reduces to finding an upper bound for $\sum_{d\leq L} K_d(x)$. Note that $K_d(x) = 0$ when $d \equiv 2 \bmod 4$. This leaves us with the following cases:

Case 1: $d \equiv 0 \bmod 4$. There are two possible ways for this to occur, namely when $k_1 = p$ and $k_2 = q$ or $k_1 = 2p$ and $k_2 = 2q$, where p and q are primes congruent to 3 modulo 4. Therefore, Lemma 2.2 gives

$$K_d(x) \leq (6 + o(1)) \prod_{p>2}\bigg(1 - \frac{1}{(p-1)^2}\bigg) f_1(d) \frac{x}{\log^2 x}. \tag{3.4}$$

Case 2: $d \equiv 1 \bmod 2$. This can occur when $k_1 = p$ and $k_2 = 2q$ or $k_1 = 2p$ and $k_2 = q$. In this case we deduce from Lemma 2.2 that

$$K_d(x) \leq (4 + o(1)) \prod_{p>2}\bigg(1 - \frac{1}{(p-1)^2}\bigg) f_1(d) \frac{x}{\log^2 x}. \tag{3.5}$$

Hence, using (3.4) and (3.5), and appealing to Lemma 2.4 with $\mathcal{P}$ being the set of primes $p > 2$, we get

$$\begin{aligned}\sum_{d\leq L} K_d(x) &= \sum_{\substack{d\leq L\\ d\equiv 0 \bmod 4}} K_d(x) + \sum_{\substack{d\leq L\\ d\equiv 1 \bmod 2}} K_d(x)\\ &\leq \bigg(6 \sum_{\substack{d\leq L\\ d\equiv 0 \bmod 4}} f_1(d) + 4 \sum_{\substack{d\leq L\\ d\equiv 1 \bmod 2}} f_1(d)\bigg) \prod_{p>2}\bigg(1 - \frac{1}{(p-1)^2}\bigg)\frac{x}{\log^2 x}(1 + o(1))\\ &= \bigg(\frac{7}{2} + o(1)\bigg) L \frac{x}{\log^2 x}.\end{aligned}$$

Thus, by combining the last estimate with equations (3.2) and (3.3) we obtain

$$|\mathcal{B}_3(x)| \geq \frac{3}{8}L\frac{x}{\log x} - \left(\frac{7}{2}+o(1)\right)L^2\frac{x}{\log^2 x} + O\left(\frac{Lx}{\log^2 x}\right).$$

We choose $L = \frac{3}{56}\log x$ to finally deduce

$$|\mathcal{B}_3(x)| \geq (\delta_1 + o(1))x,$$

where $\delta_1 = \frac{9}{896} \approx 0.0100446429$. The corresponding lower bound for $|\mathcal{B}_1(x)|$ can be obtained along the same lines. ∎

4. Proof of Theorem 2. Let $N(x)$ be the set of positive integers $n \leq x$ such that $3 \nmid \varphi(n)$, and write $N(x) = \{a_1, \dots, a_k\}$ with $1 \leq a_1 < \dots < a_k \leq x$, and $k = |N(x)|$. Put $a_0 = 1$ and $a_{k+1} = [x]$. Since the number of positive integers $1 \leq i \leq k$ such that $a_i \in N_2(x)$ or $a_{i+1} \in N_2(x)$ is at most $2N_2(x)$, using Theorem 3 we deduce that

$$\begin{aligned}(4.1)\quad &|\{1 \leq i \leq k : (a_i, a_{i+1}) \in N_1^2(x)\}| \\ &\geq N(x) - 2N_2(x) = N_1(x) - N_2(x) \geq (\delta + o(1))\frac{x}{\sqrt{\log x}},\end{aligned}$$

where $\delta = \lambda - \beta \approx 0.2824959957$. Let $L \leq \log^2 x$ be a positive real number to be chosen later, and suppose that for some positive integer $1 \leq i \leq k$ we have $(a_i, a_{i+1}) \in N_1^2(x)$ (that is, $\varphi(a_i) \equiv \varphi(a_{i+1}) \equiv 1 \bmod 3$) and $\min(a_i - a_{i-1}, a_{i+1} - a_i, a_{i+2} - a_{i+1}) \geq L$. Then for $j = 0, 1, 2$, there are at least $[L]$ integers $n \in [a_{i-1}, a_{i+2}]$ such that $\Phi(n) \equiv j \bmod 3$. Let $R(L)$ be the set of such integers i, and define

$$S_d(x) = |\{(b_1, b_2) \in N^2(x) : b_1 < b_2 \text{ and } b_2 - b_1 = d\}|$$

for all positive integers $d \geq 1$. Since the number of integers $1 \leq i \leq k$ such that $\min(a_i - a_{i-1}, a_{i+1} - a_i, a_{i+2} - a_{i+1}) < L$ is bounded by $3\sum_{d<L} S_d(x)$, we infer from (4.1) that

$$|R(L)| \geq (\delta + o(1))\frac{x}{\sqrt{\log x}} - 3\sum_{d\leq L} S_d(x). \tag{4.2}$$

On the other hand there are at least $[|R(L)|/3]$ positive integers $i \in R(L)$ such that the intervals $[a_{i-1}, a_{i+2}]$ are disjoint. Hence, for $j = 0, 1, 2$ we have

$$|\mathcal{A}_j(x)| \geq \left(\frac{\delta}{3} + o(1)\right)\frac{Lx}{\sqrt{\log x}} - L\sum_{d\leq L} S_d(x) + O\left(\frac{x}{\sqrt{\log x}}\right). \tag{4.3}$$

In order to obtain an upper bound for $\sum_{d\leq L} S_d(x)$, we use sieve theory. Indeed by Lemma 2.1 we know that all $n \in N(x)$ are not divisible by any prime $p \equiv 1 \bmod 3$. Therefore, Lemma 2.3 gives

$$S_d(x) \leq (2C_0 + o(1)) \prod_{p\equiv 1 \bmod 3}\left(1 - \frac{2}{p}\right)\left(1 - \frac{1}{p}\right)^{-2} f_2(d)\frac{x}{\log x},$$

where

$$C_0 = \lim_{y\to\infty} \prod_{\substack{p\equiv 1 \bmod 3\\ p\le y}} \Big(1-\frac{1}{p}\Big)^2 \prod_{p\le y}\Big(1-\frac{1}{p}\Big)^{-1}$$
$$= \frac{3}{2}\lim_{y\to\infty} \prod_{\substack{p\equiv 1 \bmod 3\\ p\le y}} \Big(1-\frac{1}{p}\Big) \prod_{\substack{p\equiv 2 \bmod 3\\ p\le y}}\Big(1-\frac{1}{p}\Big)^{-1} \approx 3.5082655141,$$

using a computation of A. Languasco and A. Zaccagnini [6]. Furthermore, using Lemma 2.4 with $\mathcal{P}$ being the set of primes $p\equiv 1 \bmod 3$, we get

$$\sum_{d\le L} S_d(x) \le (2C_0+o(1))L\frac{x}{\log x}.$$

Combining this with (4.3) we obtain

$$|\mathcal{A}_j(x)| \ge (1+o(1))\bigg(\frac{\delta}{3}\frac{Lx}{\sqrt{\log x}} - 2C_0\frac{L^2x}{\log x}\bigg) + O\bigg(\frac{x}{\sqrt{\log x}}\bigg)$$

for $j=0,1,2$. Thus, choosing $L=\alpha_0\sqrt{\log x}$ with $\alpha_0=\delta/(12C_0)$ we deduce that

$$|\mathcal{A}_j(x)| \ge (\delta_2+o(1))x,$$

where $\delta_2=\delta^2/(72C_0)\approx 0.0003159363$, completing the proof. ∎

5. The distribution of Euler's function modulo 3: Proof of Theorem 3. For $i=1,2$ let $M_i(x)$ be the set of positive integers $n\le x$ such that $3\nmid\varphi(n)$ and $\varphi(n)\equiv i \bmod 3$. Then one can easily check that $n\in N_1(x)$ if and only if $n\in M_1(x)$ or $n=3d$ with $d\in M_2(x/3)$. This implies

$$|N_1(x)| = |M_1(x)| + |M_2(x/3)|, \tag{5.1}$$

and similarly we get

$$|N_2(x)| = |M_2(x)| + |M_1(x/3)|. \tag{5.2}$$

Hence it suffices to estimate $|M_1(x)|$ and $|M_2(x)|$.

Let n be a positive integer such that $3\nmid n$ and $3\nmid\varphi(n)$, and write $n=\prod_{j=1}^k p_j^{a_j}$. Then $\varphi(n)=\prod_{j=1}^k p_j^{a_j-1}(p_j-1)$ and therefore $p_j\equiv 2 \bmod 3$ for all $1\le j\le k$. Moreover one has

$$\varphi(n)\equiv(-1)^{\sum_{j=1}^k a_j-k}\equiv(-1)^{\Omega(n)-\omega(n)} \bmod 3, \tag{5.3}$$

where $\Omega(n)$ (respectively $\omega(n)$) is the number of distinct prime factors of n counted with (respectively without) multiplicity. Let f and g be the arithmetic functions defined by

$$f(n)=\begin{cases}\dfrac{1+(-1)^{\Omega(n)-\omega(n)}}{2} & \text{if } \ p\,|\,n \text{ implies } p\equiv 2 \bmod 3,\\ 0 & \text{otherwise,}\end{cases}$$

$$g(n)=\begin{cases}\dfrac{1-(-1)^{\Omega(n)-\omega(n)}}{2} & \text{if } \ p\,|\,n \text{ implies } p\equiv 2 \bmod 3,\\ 0 & \text{otherwise.}\end{cases}$$

Then we deduce from (5.3) that

$$|M_1(x)|=\sum_{n\le x} f(n) \quad \text{and} \quad |M_2(x)|=\sum_{n\le x} g(n). \tag{5.4}$$

The Dirichlet series of f and g are defined by

$$L_f(s):=\sum_{n=1}^{\infty}\frac{f(n)}{n^s} \quad \text{and} \quad L_g(s):=\sum_{n=1}^{\infty}\frac{g(n)}{n^s},$$

respectively, and are absolutely convergent for $\mathrm{Re}(s)>1$. Our idea is to express $L_f(s)$ (respectively $L_g(s)$) as a power of the Riemann zeta function $\zeta(s)$ times a function $\mathcal{H}_1(s)$ (respectively $\mathcal{H}_2(s)$) which is analytic in the half plane $\mathrm{Re}(s)\ge 1$, and then use the Selberg–Delange method to estimate $\sum_{n\le x} f(n)$ (respectively $\sum_{n\le x} g(n)$). There are two characters modulo 3, the principal character χ_0 and the real character χ_3 defined by $\chi_3(n)=\left(\frac{n}{3}\right)$. We prove

Proposition 5.1. *Let $s\in\mathbb{C}$ with $\mathrm{Re}\,(s)>1$. Then*

$$L_f(s)=\zeta(s)^{1/2}\mathcal{H}_1(s) \quad \textit{and} \quad L_g(s)=\zeta(s)^{1/2}\mathcal{H}_2(s),$$

where

$$\mathcal{H}_1(s):=\frac{\left(1-\dfrac{1}{3^s}\right)^{1/2}\left(1+\displaystyle\prod_{p\equiv 2 \bmod 3}\left(1-\dfrac{2}{p^s(p^s+1)}\right)^{-1}\right)}{2\left(L(s,\chi_3)\displaystyle\prod_{p\equiv 2 \bmod 3}\left(1-\dfrac{1}{p^{2s}}\right)\right)^{1/2}},$$

$$\mathcal{H}_2(s):=\frac{\left(1-\dfrac{1}{3^s}\right)^{1/2}\left(1-\displaystyle\prod_{p\equiv 2 \bmod 3}\left(1-\dfrac{2}{p^s(p^s+1)}\right)^{-1}\right)}{2\left(L(s,\chi_3)\displaystyle\prod_{p\equiv 2 \bmod 3}\left(1-\dfrac{1}{p^{2s}}\right)\right)^{1/2}}.$$

Moreover both $\mathcal{H}_1(s)$ and $\mathcal{H}_2(s)$ can be analytically continued in a region $\mathrm{Re}(s)\ge 1-c_0/(1+\log(|\mathrm{Im}(s)|+2))$ *for some constant $c_0>0$.*

Proof. We shall only prove the statement for $L_f(s)$, since the argument for $L_g(s)$ is similar. Since the function $(-1)^{\Omega(n)-\omega(n)}$ is multiplicative, we get

$$L_f(s) = \frac{1}{2}\prod_{p\equiv 2 \bmod 3}\left(1-\frac{1}{p^s}\right)^{-1} + \frac{1}{2}\prod_{p\equiv 2 \bmod 3}\left(1+\sum_{a=1}^{\infty}\frac{(-1)^{a-1}}{p^{as}}\right) \tag{5.5}$$
$$= \frac{1}{2}\prod_{p\equiv 2 \bmod 3}\left(1-\frac{1}{p^s}\right)^{-1} + \frac{1}{2}\prod_{p\equiv 2 \bmod 3}\left(1+\frac{1}{p^s+1}\right)$$
$$= \frac{1}{2}\prod_{p\equiv 2 \bmod 3}\left(1-\frac{1}{p^s}\right)^{-1}\left(1+\prod_{p\equiv 2 \bmod 3}\left(1-\frac{2}{p^s(p^s+1)}\right)^{-1}\right)$$

for $\mathrm{Re}(s) > 1$. On the other hand we have

$$\sum_{k\geq 1}\sum_{p\equiv 2 \bmod 3}\frac{1}{kp^{ks}} = \frac{1}{2}\sum_{k\geq 1}\sum_{p}\frac{\chi_0(p)}{kp^{ks}} - \frac{1}{2}\sum_{k\geq 1}\sum_{p}\frac{\chi_3(p)}{kp^{ks}} \tag{5.6}$$
$$= \frac{1}{2}\log L(s,\chi_0) - \frac{1}{2}\log L(s,\chi_3) + \frac{1}{2}\sum_{k\geq 1}\sum_{p}\frac{\chi_3(p)^k-\chi_3(p)}{kp^{ks}}.$$

Now if $p \equiv 1 \bmod 3$ or k is odd then $\chi_3(p)^k - \chi_3(p) = 0$. This yields

$$\sum_{k\geq 1}\sum_{p}\frac{\chi_3(p)^k-\chi_3(p)}{kp^{ks}} = \sum_{m\geq 1}\sum_{p\equiv 2 \bmod 3}\frac{1}{mp^{2ms}} = -\sum_{p\equiv 2 \bmod 3}\log\left(1-\frac{1}{p^{2s}}\right).$$

Combining this with (5.6) we obtain

$$\prod_{p\equiv 2 \bmod 3}\left(1-\frac{1}{p^s}\right)^{-1} \tag{5.7}$$
$$= L(s,\chi_0)^{1/2}L(s,\chi_3)^{-1/2}\prod_{p\equiv 2 \bmod 3}\left(1-\frac{1}{p^{2s}}\right)^{-1/2}$$
$$= \left(1-\frac{1}{3^s}\right)^{1/2}\zeta(s)^{1/2}L(s,\chi_3)^{-1/2}\prod_{p\equiv 2 \bmod 3}\left(1-\frac{1}{p^{2s}}\right)^{-1/2}.$$

The result then follows from (5.5), along with the fact that $L(s,\chi_3)$ is entire and does not vanish in a region $\mathrm{Re}(s) \geq 1 - c_0/(1+\log(|\mathrm{Im}(s)|+2))$, for some constant $c_0 > 0$. ■

Proof of Theorem 3. Using the Selberg–Delange method (more precisely Theorem 3 in Chapter II.5 of [9]) we infer from Proposition 5.1 that

$$\sum_{n\leq x} f(n) = \alpha\frac{x}{\sqrt{\log x}} + \sum_{k=1}^{K-1}\frac{\alpha_k x}{\log^{k+1/2} x} + O_K\left(\frac{x}{\log^{K+1/2} x}\right), \tag{5.8}$$

where

$$\alpha = \frac{\mathcal{H}_1(1)}{\Gamma(1/2)} = \frac{3^{1/4}}{\sqrt{2}\,\pi} \prod_{p\equiv 2 \bmod 3} \left(1-\frac{1}{p^2}\right)^{-1/2} \left(1+\prod_{p\equiv 2 \bmod 3}\left(1-\frac{2}{p(p+1)}\right)\right),$$

since $\Gamma(1/2)=\sqrt{\pi}$ and $L(1,\chi_3)=\pi/3^{3/2}$, which follows from the Dirichlet class number formula (see Chapter 6 of [1]). Moreover the constants α_k are defined by

$$\alpha_k := \frac{1}{\Gamma(1/2-k)} \sum_{l+j=k} \frac{1}{l!}\mathcal{H}_1^{(l)}(1)s_j,$$

and s_j are the coefficients of the Laurent series of $s^{-1}((s-1)\zeta(s))^{1/2}$ around the point $s=1$.

Analogously to (5.8) we obtain a similar asymptotics for $\sum_{n\le x} g(n)$ with different constants α' and α'_k where α' is given by the same expression as for α except that the factor $1+\prod_{p\equiv 2 \bmod 3}\big(1-\frac{2}{p(p+1)}\big)$ is replaced by $1-\prod_{p\equiv 2 \bmod 3}\big(1-\frac{2}{p(p+1)}\big)$. Finally the result follows upon combining these asymptotic formulas with (5.1), (5.2) and (5.4). ∎

One can also arrive at the conclusion of Theorem 3 by applying the same method to the Dirichlet series

$$\sum_{n=1}^{\infty} \frac{\chi_0(\varphi(n))}{n^s} \quad \text{and} \quad \sum_{n=1}^{\infty} \frac{\chi_3(\varphi(n))}{n^s},$$

upon noting that

$$\sum_{\substack{n\le x \\ \varphi(n)\equiv i \bmod 3}} 1 = \frac{1}{2}\Big(\sum_{n\le x}\chi_0(\varphi(n)) + \epsilon_i \sum_{n\le x}\chi_3(\varphi(n))\Big),$$

with $\epsilon_1 = 1$ and $\epsilon_2=-1$.

Acknowledgments. The authors are grateful to the referee for many useful comments and suggestions.

The first author is supported by a postdoctoral fellowship from the Natural Sciences and Engineering Research Council of Canada. The second author was supported from the NSF grant DMS-0838434 "EMSW21MCTP: Research Experience for Graduate Students". Research of the third author is supported by the NSF grant DMS-0901621.

REFERENCES

[1] H. Davenport, *Multiplicative Number Theory*, 3rd ed., Grad. Texts in Math. 74, Springer, New York, 2000.

[2] T. Dence and C. Pomerance, *Euler's function in residue classes*, in: Paul Erdős (1913–1996), Ramanujan J. 2 (1998), 7–20.

[3] H. Halberstam and H.-E. Richert, *Sieve Methods*, London Math. Soc. Monogr. 4, Academic Press, London, 1974.
[4] J. Kaczorowski, *Some remarks on factorization in algebraic number fields*, Acta Arith. 43 (1983), 53–68.
[5] E. Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*, Teubner, Leipzig, 1909; 2nd ed., Chelsea, New York, 1953.
[6] A. Languasco and A. Zaccagnini, *On the constant in the Mertens product for arithmetic progressions. II. Numerical values*, Math. Comp. 78 (2009), 315–326.
[7] W. Narkiewicz, *On distribution of values of multiplicative functions in residue classes*, Acta Arith. 12 (1966/1967), 269–279.
[8] P. Pollack, *Not Always Buried Deep: A Second Course in Elementary Number Theory*, Amer. Math. Soc., Providence, RI, 2009.
[9] G. Tenenbaum, *Introduction to Analytic and Probabilistic Number Theory*, Cambridge Stud. Adv. Math. 46, Cambridge Univ. Press, Cambridge, 1995.
[10] E. Wirsing, *Über die Zahlen, deren Primteiler einer gegebenen Menge angehören*, Arch. Math. 7 (1956), 263–272.

Youness Lamzouri, M. Tip Phaovibul, Alexandru Zaharescu
Department of Mathematics
University of Illinois at Urbana-Champaign
1409 W. Green Street
Urbana, IL 61801, U.S.A.
E-mail: lamzouri@math.uiuc.edu
phaovib1@illinois.edu
Zaharescu@math.uiuc.edu

Received 12 November 2010;
revised 15 January 2011 (5444)