COUNTING INTEGERS WITH A SMOOTH TOTIENT

W. D. BANKS, J. B. FRIEDLANDER, C. POMERANCE,
AND I. E. SHPARLINSKI

Abstract. In an earlier paper we considered the distribution of integers n for which Euler's totient function at n has all small prime factors. Here we obtain an improvement that is likely to be best possible.

1. Introduction

Our paper [1] considers various multiplicative problems related to Euler's function φ. One of these problems concerns the distribution of integers n for which $\varphi(n)$ is y-smooth (or y-friable), meaning that all prime factors of $\varphi(n)$ are at most y. Let $\Phi(x, y)$ denote the number of $n \le x$ such that $\varphi(n)$ is y-smooth. Theorem 3.1 in [1] asserts that the following bound holds:

> *For any fixed $\varepsilon > 0$, numbers x, y with $y \ge (\log\log x)^{1+\varepsilon}$, and $u = \log x/\log y \to \infty$, we have the bound $\Phi(x, y) \le x/\exp((1 + o(1))u \log\log u)$.*

In this note we establish a stronger bound. Merging Propositions 2.3 and 3.2 below we prove the following result.

Theorem 1.1. *For any fixed $\varepsilon > 0$, numbers x, y with $y \ge (\log\log x)^{1+\varepsilon}$, and $u = \log x/\log y \to \infty$, we have*

$$\Phi(x, y) \le x \exp\big(-u(\log\log u + \log\log\log u + o(1))\big).$$

One might wonder about a matching lower bound for $\Phi(x, y)$, but this is very difficult to achieve since it depends on the distribution of primes p with $p-1$ being y-smooth. Let $\psi(x, y)$ denote the number of y-smooth integers at most x, and let $\psi_\pi(x, y)$ denote the number of primes $p \le x$ such that $p-1$ is y-smooth. It has been conjectured (see [15] and the discussion therein) that in a wide range one has $\psi_\pi(x, y)/\pi(x) \sim$

Date: June 4, 2019.
2010 *Mathematics Subject Classification.* 11N25, 11N37.
Key words and phrases. Euler function, smooth numbers.

$\psi(x,y)/x$. Assuming a weak form of this conjecture, Lamzouri [9] has shown that there is a continuous monotonic function $\sigma(u)$ such that

$$\sigma(u) = \exp\bigl(-u(\log\log u + \log\log\log u + o(1))\bigr) \qquad (u \to \infty),$$

and such that $\Phi(x, x^{1/u}) \sim \sigma(u)x$ as $x \to \infty$ with u bounded. The function σ is explicitly identified as the solution to the integral equation

$$u\sigma(u) = \int_0^u \sigma(u-t)\rho(t)\,dt,$$

where ρ is the Dickman–de Bruijn function.

In light of Lamzouri's theorem, it seems likely that we have equality in Theorem 1.1.

Our proof of Theorem 1.1 is given as two results: Proposition 2.3 for the case when $y \le x^{1/\log\log x}$ and Proposition 3.2 for the case when $y \ge \exp(\sqrt{\log x \log\log x}\,)$. Note that the ranges of Propositions 2.3 and 3.2 have a significant overlap. In the first range we use a variant of Rankin's trick. In the second range we use a variant of the Hildebrand approach [7] for estimating $\psi(x,y)$.

Our proof is adaptable to multiplicative functions similar in structure to Euler's φ-function. For example, in [14] a version of our theorem is used for the distribution of squarefree $n \le x$ with $\sigma(n)$ being y-smooth, where σ is the sum-of-divisors function.

The original purpose of this paper was to correct an error in the proof of [1, Theorem 3.1], kindly pointed out to us by Paul Kinlaw. In fact, our treatment there of the sum $\sum_{p\le y} p^{-c}$ is flawed for larger values of y. Being able now to establish a likely best-possible result was an unexpected bonus.

In a recent paper, Pollack [10] shows (as a special case) that for any fixed number $\alpha > 1$,

$$\Phi(x, (\log x)^{1/\alpha}) \le x^{1-(\alpha+o(1))\log\log\log x/\log\log x}$$

as $x \to \infty$. A slightly stronger inequality follows from our Theorem 1.1, though in Pollack's result the inequality applies to sets more general than the $(\log x)^{1/\alpha}$-smooth integers.

Our paper [1] also considered the distribution of integers n for which $\varphi(n)$ is a square and the distribution of squares in the image of φ. These results have attracted interest and since then have been improved and extended in various ways; see [5, 6, 11, 12].

In what follows, $P(n)$ denotes the largest prime factor of an integer $n > 1$, and $P(1) = 1$. The letter p always denotes a prime number; the letter n always denotes a positive integer. As usual in the subject, we write $\log_k x$ for the kth iterate of the natural logarithm, assuming that the argument is large enough for the expression to make sense.

We use the notations $U = O(V)$ and $U \ll V$ in their standard meaning that $|U| \le cV$ for some constant c, which throughout this paper may depend on the real positive parameters ε, δ, η. We also use the notations $U \sim V$ and $U = o(V)$ to indicate that $U/V \to 1$ and $U/V \to 0$, respectively, when certain (explicitly indicated) parameters tend to infinity.

2. Small y

2.1. **Dickman–de Bruijn function.** As above, we denote by ρ the Dickman–de Bruijn function; we refer the reader to [8] for an exact definition and properties. For the first range it is useful to have the following two estimates involving this function.

Lemma 2.1. *Let $\eta > 0$ be arbitrarily small but fixed. For $A \ge 2$ we have*

$$\sum_{n\ge 1} A^n \rho(n) \ll \exp\left(\frac{(1+\eta)A}{\log A}\right).$$

Proof. It is sufficient to prove the result for large numbers A. Since $\rho(n) \le 1$, the sum up to $A/(\log A)^2$ is $\ll \exp(A/\log A)$, hence we need only consider integers $n > A/(\log A)^2$. We have for $t > 1$,

(2.1)
$$\rho(t) = \exp\left(-t\left(\log t + \log_2 t - 1 + \frac{\log_2 t - 1}{\log t} + O\left(\frac{(\log_2 t)^2}{(\log t)^2}\right)\right)\right);$$

see for example de Bruijn [3, (1.5)]. Consequently, if $n > A/(\log A)^2$ and A is large enough, then

$$A^n\rho(n) < \exp(n(\log A - \log n - \log_2 n + 1)).$$

In the case $n > A$, this implies that

$$A^n\rho(n) < \exp(-n\log_2 n + n) < \exp(-n),$$

and so the contribution to the sum when $n > A$ is $O(1)$. Now assume that $A/(\log A)^2 < n \le A$. Let $f(t) = t(\log A - \log t - \log_2 t + 1)$. For any $\theta \ge 1/\log A$ one sees that

$$\begin{aligned} f\left(\frac{\theta A}{\log A}\right) &= \frac{\theta A}{\log A}\left(-\log\theta + \log_2 A - \log_2\left(\frac{\theta A}{\log A}\right) + 1\right) \\ &= -\frac{\theta A}{\log A}(\log\theta + C_{A,\theta}), \end{aligned}$$

where

$$C_{A,\theta} = \log\left(\frac{\log A + \log\theta - \log_2 A}{\log A}\right) - 1.$$

Hence, when A is large enough depending on the choice of η, we have

$$f\left(\frac{\theta A}{\log A}\right) \le -\frac{\theta A}{\log A}(\log\theta - (1+\eta/2)) \qquad (\theta > 1/\log A).$$

Since this last expression reaches a maximum when $\theta = e^{\eta/2}$, we have $f(t) \le e^{\eta/2}A/\log A < (1+3\eta/4)A/\log A$ for all $t > A/(\log A)^2$, and so

$$\sum_{A/(\log A)^2 < n \le A} A^n\rho(n) < A\exp\left(\frac{(1+3\eta/4)A}{\log A}\right) \ll \exp\left(\frac{(1+\eta)A}{\log A}\right),$$

which completes the proof of the lemma. □

To prove the main results of this paper, we need information about the distribution of primes p with $p-1$ suitably smooth. The following statement, which is [15, Theorem 1] (see also [1, Equation (2.3)]), suffices for our purposes.

Lemma 2.2. *For* $\exp(\sqrt{\log t \log_2 t}\,) \le y \le t$ *and with* $u_t = \log t/\log y$ *we have*

$$\psi_\pi(t,y) = \sum_{\substack{p\le t\\ P(p-1)\le y}} 1 \ll u_t\rho\left(u_t\right)\frac{t}{\log t} = \rho(u_t)\frac{t}{\log y}.$$

It is useful to observe that the range in Lemma 2.2 includes the range

$$y \le t \le y^{\log y/2\log_2 y}.$$

2.2. **Bound on $\Phi(x,y)$ for $(\log_2 x)^{1+\varepsilon} \le y \le x^{1/\log_2 x}$.** We give a proof of the following result.

Proposition 2.3. *Fix* $\varepsilon > 0$. *For* $(\log_2 x)^{1+\varepsilon} \le y \le x^{1/\log_2 x}$, *and* $u = \log x/\log y \to \infty$, *we have*

$$\Phi(x,y) \le x\exp(-u(\log_2 u + \log_3 u + o(1))).$$

Proof. We may assume that u is large and shall need to do so at various points in the proof. We may also assume that $\varepsilon < 1$. Let $\delta > 0$ be arbitrarily small but fixed. We prove that

$$\Phi(x,y) \le x\exp(-u(\log_2 u + \log_3 u - \delta + o(1))) \qquad (u\to\infty),$$

which is sufficient for the desired result.

Put

$$c = 1 - (\log_2 u + \log_3 u - \delta)/\log y,$$

so that $c < 1$ for u sufficiently large. Also, $u < \log x$ implies that

$$1 - c = \frac{\log_2 u + \log_3 u - \delta}{\log y} < \frac{\log_3 x + \log_4 x}{(1+\varepsilon)\log_3 x} < 1 - \frac{\varepsilon}{2},$$

for u sufficiently large, so we may assume that $1 > c > \varepsilon/2$. We have

$$\Phi(x,y) \le x^c \sum_{\substack{n \le x \\ P(\varphi(n)) \le y}} \frac{1}{n^c} \le x^c \prod_{\substack{p \le x \\ P(p-1) \le y}} \left(1 - \frac{1}{p^c}\right)^{-1}. \tag{2.2}$$

Note that $x^c = x\exp(-u(\log_2 u + \log_3 u - \delta))$, so via (2.2) it suffices to prove that

$$-\sum_{\substack{p \le x \\ P(p-1) \le y}} \log\left(1 - \frac{1}{p^c}\right) = o(u), \tag{2.3}$$

as $u \to \infty$. Note that, using $c > \varepsilon/2$,

$$-\sum_{\substack{p \le x \\ P(p-1) \le y}} \log\left(1 - \frac{1}{p^c}\right) = \sum_{\substack{p \le x \\ P(p-1) \le y}} \sum_{k \ge 1} \frac{1}{kp^{ck}} \ll \sum_{\substack{p \le x \\ P(p-1) \le y}} \frac{1}{p^c}.$$

To establish (2.3) and hence the desired result, it is sufficient to show that, as $u \to \infty$,

$$\sum_{\substack{p \le x \\ P(p-1) \le y}} \frac{1}{p^c} = o(u). \tag{2.4}$$

Put

$$z = \frac{\log y}{2\log_2 y}, \tag{2.5}$$

and consider primes $p \le x$ with $P(p-1) \le y$ in two ranges:

(1) $p \le y^z$,
(2) $p > y^z$.

Note that the second range contains primes only in the case that $y^z \le x$.

To estimate the first range for p, we have

$$\sum_{\substack{p \le y^z \\ P(p-1) \le y}} \frac{1}{p^c} \le \sum_{1 \le k < z+1} \sum_{\substack{y^{k-1} < p \le y^k \\ P(p-1) \le y}} \frac{1}{p^c}.$$

For the inner sum we use Lemma 2.2 together with partial summation and the fact that $y^{1-c} = e^{-\delta} \log u \log_2 u$ getting that

$$\sum_{\substack{y^{k-1} < p \le y^k \\ P(p-1) \le y}} \frac{1}{p^c} \ll \rho(k) \frac{y^{k(1-c)}}{\log y} + \int_{y^{k-1}}^{y^k} \rho(k-1) \frac{1}{t^c \log y} \, dt$$

$$\ll \rho(k-1) \frac{y^{k(1-c)}}{(1-c)\log y} \ll \rho(k-1) \left(e^{-\delta} \log u \log_2 u \right)^k .$$

We use Lemma 2.1 with $A = e^{-\delta} \log u \log_2 u$ and $\eta = \delta$, finding that

$$\sum_{\substack{p \le y^z \\ P(p-1) \le y}} \frac{1}{p^c} \ll A \exp\left(\frac{(1+\delta)A}{\log A} \right).$$

Since $(1+\delta)A/\log A \sim (1+\delta)e^{-\delta} \log u$ as $u \to \infty$, and $(1+\delta)e^{-\delta} < 1$, this shows that the sum in (2.4) is $O(u^{1-\delta'})$ for some $\delta' > 0$ depending on the choice of δ. Thus we have (2.4) for primes in the first range.

Now we turn to the second range. As mentioned earlier, we may assume that $y^z \le x$. By de Bruijn [2, (1.6)] we have

$$\psi(t, y) \le t / e^{u_t \log u_t} \qquad (y^z < t \le x), \tag{2.6}$$

where u_t is as in Lemma 2.2, for u sufficiently large. Ignoring that p is prime we have the bound

$$\sum_{\substack{y^z < p \le x \\ P(p-1) \le y}} \frac{1}{p^c} \le \sum_{\substack{y^z - 1 < n \le x \\ P(n) \le y}} \frac{1}{n^c} \le 1 + \sum_{z+1 \le k \le u} \sum_{\substack{y^{k-1} < n \le y^k \\ P(n) \le y}} \frac{1}{n^c}. \tag{2.7}$$

Next, we put

$$y_0 = \exp\left((\log_2 x)^2 \right)$$

and consider separately the cases $y \ge y_0$ and $y < y_0$. In the case that $y \ge y_0$, using (2.6) the inner sum on the right side of (2.7) satisfies

$$\sum_{\substack{y^{k-1} < n \le y^k \\ P(n) \le y}} \frac{1}{n^c} \le \frac{\psi(y^k, y)}{y^{kc}} + \int_{y^{k-1}}^{y^k} \frac{c\,\psi(t, y)}{t^{c+1}} \, dt$$

$$\le k^{-k} y^{k(1-c)} + (k-1)^{-(k-1)} \int_{y^{k-1}}^{y^k} t^{-c} \, dt$$

$$\ll k^{-(k-1)} \frac{y^{k(1-c)}}{1-c}$$

$$\le k \log y \cdot \exp(-k(\log k - \log_2 u - \log_3 u + \delta)).$$

Since $y \ge y_0$, $k \ge z$, with z given by (2.5), and $u < \log x$, we have

$$\begin{aligned}\log k - \log_2 u - \log_3 u &\ge \log z - \log_2 u - \log_3 u\\ &\ge \log_2 y - \log_3 y - \log 2 - \log_2 u - \log_3 u\\ &\ge \frac{7}{8}\log_2 y - \log_2 u - \log_3 u\\ &\ge \frac{7}{4}\log_3 x - \log_2 u - \log_3 u > \frac{1}{2}\log_3 x\end{aligned}$$

provided that u is large. Hence,

$$\sum_{\substack{y^{k-1}<n\le y^k\\ P(n)\le y}} \frac{1}{n^c} \ll \exp(-k)\log y$$

and so the sum in (2.7) is $O(\exp(-z)\log y) = O(1)$.

It remains to handle the second range when $y < y_0$. In this case, we use an Euler product for a second time, getting that

$$\begin{aligned}\sum_{\substack{n\le x\\ P(n)\le y}} n^{-c} < \prod_{p\le y}\left(1-p^{-c}\right)^{-1} \ll \exp\left(\sum_{p\le y} p^{-c}\right)\\ = \exp\left(\mathrm{li}(y^{1-c})(1+O(1/\log y)) + O(|\log(1-c)|)\right),\end{aligned}$$

where we have used [13, Equation (2.4)] in the last step. Now

$$\mathrm{li}(y^{1-c}) = (1+o(1))\frac{y^{1-c}}{(1-c)\log y} = \frac{1+o(1)}{e^{\delta}}\log u,$$

as $u \to \infty$, and

$$|\log(1-c)| < \log_2 y < 2\log_3 x \ll \log_2 u.$$

Therefore

$$\sum_{\substack{n\le x\\ P(n)\le y}} n^{-c} \le u^{e^{-\delta/2}}$$

for u sufficiently large. This completes the proof. □

3. Large y

3.1. **A version of the Hildebrand identity.** We begin this section by proving an analog of the Hildebrand identity which is adapted to our function $\Phi(x, y)$. Note that it is given as an inequality, but it would not be hard to account for the excess on the higher side.

Lemma 3.1. *For $x \ge y \ge 2$ we have*

$$\Phi(x,y) \le \frac{1}{\log x}\int_1^x \frac{\Phi(t,y)}{t}\,dt + \frac{1}{\log x}\sum_{\substack{d\le x\\ P(\varphi(d))\le y}} \Phi\left(\frac{x}{d},y\right)\Lambda(d).$$

Proof. By partial summation, we have

$$\sum_{\substack{n\le x\\ P(\varphi(n))\le y}} \log n = \Phi(x,y)\log x - \int_1^x \frac{\Phi(t,y)}{t}\,dt. \tag{3.1}$$

On the other hand, we have

$$\begin{aligned}\sum_{\substack{n\le x\\ P(\varphi(n))\le y}} \log n &= \sum_{\substack{n\le x\\ P(\varphi(n))\le y}} \sum_{d\,|\,n} \Lambda(d) = \sum_{\substack{d\le x\\ P(\varphi(d))\le y}} \sum_{\substack{m\le x/d\\ P(\varphi(md))\le y}} \Lambda(d)\\ &\le \sum_{\substack{d\le x\\ P(\varphi(d))\le y}} \Phi\left(\frac{x}{d},y\right)\Lambda(d).\end{aligned}$$

Substituting (3.1) on the left side and solving the resulting inequality for $\Phi(x,y)$ gives the result. □

3.2. Bound on $\Phi(x,y)$ for $y \ge \exp(\sqrt{\log x \log_2 x}\,)$.

Proposition 3.2. *For $y \ge \exp(\sqrt{\log x \log_2 x}\,)$, and $u = \log x/\log y \to \infty$, we have*

$$\Phi(x,y) \le x\exp(-u(\log_2 u + \log_3 u + o(1))).$$

Proof. Let $\delta > 0$ be arbitrarily small but fixed, and put

$$g(u) = \exp(-u(\log_2 u + \log_3 u - \delta)).$$

It suffices to show that $\Phi(x,y) \ll xg(u)$ for x, y in the given range.

For any given $u \ge 3$, which without loss of generality we may assume, let Γ_u be the supremum of $\Phi(x,y)/(xg(u))$ for all x, y with $y = x^{1/u}$, so that trivially $\Gamma_u \le 1/g(u)$. Further, let

$$\gamma_u = \sup\{\Gamma_v : 3 \le v \le u\}.$$

Our goal is to show that γ_u is bounded. Towards this end, we may assume that $u \ge u_0 \ge 3$, where u_0 is a suitably large constant, depending on the choice of δ. Since γ_u is nondecreasing as a function of u, we may assume that

$$\gamma_u \ge 1 \qquad (u \ge u_0), \tag{3.2}$$

for otherwise γ_u is clearly bounded. We further assume that u_0 is large enough so that

$$\frac{1}{\log v} + \frac{1}{\log v \log_2 v} \le \delta \qquad (v \ge u_0). \tag{3.3}$$

Let N be such that

$$u_0 \le N \le \exp(\sqrt{\log x / \log_2 x}\,) - 1.$$

We claim that for u_0 large enough

$$\sup_{N<u\le N+1} \Gamma_u \le \gamma_N. \tag{3.4}$$

By induction, this implies that $\gamma_u \le \gamma_{u_0}$ for all $u \ge u_0$, and therefore

$$\Phi(x,y) \le \gamma_{u_0} x g(u)$$

for all $u \ge u_0$, and the result would follow.

One other observation is that $g(u) \sim e^{-\delta} g(u+1) \log u \log_2 u$ as $u \to \infty$, so that with u_0 large and $u_0 \le N < u \le N+1$, we have

$$g(N) \le g(u) \log u \log_2 u \quad \text{and} \quad g(N-1) \le g(u)(\log u \log_2 u)^2. \tag{3.5}$$

To establish (3.4) we first consider the term

$$\mathfrak{T}_1 = \frac{1}{\log x} \int_1^x \frac{\Phi(t,y)}{t}\, dt$$

in Lemma 3.1. We split the range of integration as follows:

$$\int_1^x = \int_1^{y^{u_0}} + \int_{y^{u_0}}^{y^N} + \int_{y^N}^x .$$

We have trivially that

$$\int_1^{y^{u_0}} \frac{\Phi(t,y)}{t}\, dt < y^{u_0}. \tag{3.6}$$

We show that for u_0 sufficiently large, we have

$$y^{u_0} \le x g(u)/g(u_0). \tag{3.7}$$

Since $y^{u_0} = x^{u_0/u}$, (3.7) is equivalent to showing that for

$$D(u) = \left(1 - \frac{u_0}{u}\right) \log x - \log g(u_0) - u(\log_2 u + \log_3 u - \delta),$$

we have

$$D(u) \ge 0. \tag{3.8}$$

Note that the hypothesis $y \ge \exp(\sqrt{\log x \log_2 x})$ implies that $\log x > u^2(\log_2 u + \log_3 u)$. By considering $D'(u)$ and using (3.3), we see that

$D(u)$ is increasing for $u \ge u_0$ and u_0 sufficiently large. Since $D(u_0) = 0$, this proves (3.8), which establishes (3.7), and so via (3.6) we have

$$\int_1^{y^{u_0}} \frac{\Phi(t,y)}{t}\,dt \le xg(u)/g(u_0). \tag{3.9}$$

Also,

$$\int_{y^{u_0}}^{y^N} \frac{\Phi(t,y)}{t}\,dt \le \gamma_N I,$$

where

$$I = \int_{y^{u_0}}^{y^N} g(\log t/\log y)\,dt = \int_{u_0}^{N} g(v) y^v \log y\,dv = \int_{u_0}^{N} g(v)\,d(y^v).$$

Thus, I is equal to

$$y^v g(v)\Big|_{u_0}^{N} + \int_{u_0}^{N} \left(\log_2 v + \log_3 v - \delta + \frac{1}{\log v} + \frac{1}{\log v \log_2 v}\right) g(v) y^v\,dv$$
$$< y^N g(N) + \frac{\log_2 N + \log_3 N}{\log y} I,$$

where we have used (3.3) in the last step. Assuming u_0 is sufficiently large (and thus so are x and y), we see that

$$\int_{y^{u_0}}^{y^N} \frac{\Phi(t,y)}{t}\,dt < 2\gamma_N y^N g(N). \tag{3.10}$$

Finally,

$$\int_{y^N}^{x} \frac{\Phi(t,y)}{t}\,dt \le \int_{y^N}^{x} \frac{\Phi(t,t^{1/N})}{t}\,dt \le \gamma_N g(N)(x - y^N). \tag{3.11}$$

Thus, using (3.9), (3.10), and (3.11), we have

$$\begin{aligned} \mathfrak{T}_1 &\le \frac{xg(u)}{g(u_0)\log x} + \frac{2\gamma_N x}{\log x} g(N) \\ &\le \frac{2\gamma_N \log u \log_2 u + 1/g(u_0)}{\log x} xg(u), \end{aligned} \tag{3.12}$$

assuming that u_0 is sufficiently large, where we used (3.5) for the last step.

Next, we consider the second term

$$\mathfrak{T}_2 = \frac{1}{\log x} \sum_{\substack{d \le x \\ P(\varphi(d)) \le y}} \Phi\left(\frac{x}{d}, y\right) \Lambda(d)$$

in Lemma 3.1, and begin by estimating the contribution from terms $d \le y$. For such d we have $y^{N-1} \le x/d$ (since $y^N \le x$), which implies that $y \le (x/d)^{1/(N-1)}$. Hence, this part of $\mathfrak{T}_2$ is at most

$$\frac{1}{\log x}\sum_{d\le y}\Phi\left(\frac{x}{d},y\right)\Lambda(d) \le \frac{1}{\log x}\sum_{d\le y}\Phi\left(\frac{x}{d},\left(\frac{x}{d}\right)^{1/(N-1)}\right)\Lambda(d)$$
$$\le \frac{\gamma_N x}{\log x} g(N-1)\sum_{d\le y}\frac{\Lambda(d)}{d}.$$

Hence, by the Mertens formula

$$\begin{aligned}\frac{1}{\log x}\sum_{d\le y}\Phi\left(\frac{x}{d},y\right)\Lambda(d) &\le \frac{2\gamma_N x\log y}{\log x} g(N-1)\\ &\le \frac{2\gamma_N x(\log u\log_2 u)^2}{u} g(u),\end{aligned}\tag{3.13}$$

assuming that u_0 is sufficiently large and using (3.5).

Next, we consider the contribution from terms $d = p^a > y$ for which $p \le y$ (and thus the positive integer a is at least two), finding from the trivial bound $\Phi(x/p^a, y) \le x/p^a$ that

$$\frac{1}{\log x}\sum_{\substack{p\le y\\ p^a>y}}\Phi\left(\frac{x}{p^a},y\right)\log p \le \frac{x}{\log x}\sum_{\substack{p\le y\\ p^a>y}}\frac{\log p}{p^a} \ll \frac{x}{\sqrt{y}\log x}.\tag{3.14}$$

The remaining terms are of the form $d = p^a$ with $p > y$, and since $P(\varphi(d)) \le y$ we conclude that $a = 1$, i.e., $d = p$. Therefore, we need to estimate

$$\frac{1}{\log x}\sum_{\substack{y<p\le x\\ P(p-1)\le y}}\Phi\left(\frac{x}{p},y\right)\log p = \frac{1}{\log x}\sum_{1\le k<u} S_k,\tag{3.15}$$

where

$$S_k = \sum_{\substack{y^k<p\le\min\{x,y^{k+1}\}\\ P(p-1)\le y}}\Phi\left(\frac{x}{p},y\right)\log p.$$

We also denote

$$T_k = \sum_{\substack{y^k<p\le\min\{x,y^{k+1}\}\\ P(p-1)\le y}}\frac{\log p}{p}.$$

For integers $k \le u/2$ we use the bound

$$S_k \le \gamma_N x g(u-k-1)T_k \le \gamma_N x\log u\log_2 u\cdot g(u-k)T_k,$$

whereas for larger integers $k > u/2$, the trivial bound $\Phi(x/p, y) \le x/p$ and (3.2) together imply that

$$S_k \le \gamma_N x T_k;$$

consequently, using (3.15),

$$\frac{1}{\log x} \sum_{\substack{y<p\le x\\ P(p-1)\le y}} \Phi\left(\frac{x}{p}, y\right) \log p \le \frac{\gamma_N x \log u \log_2 u}{\log x} \sum_{1\le k\le u/2} g(u-k) T_k + \frac{\gamma_N x}{\log x} \sum_{u/2<k<u} T_k. \tag{3.16}$$

Next, define

$$h(k) = \exp(-k(\log k + \log_2(k+1) - 1))$$

and note that from (2.1) we have

$$k\rho(k) \ll h(k). \tag{3.17}$$

By partial summation, using Lemma 2.2 together with (3.17), we see that there is an absolute constant c_0 such that for $1 \le k < u$ we have

$$T_k = \sum_{\substack{y^k<p\le \min\{x,y^{k+1}\}\\ P(p-1)\le y}} \frac{\log p}{p} \le c_0 h(k) \log y.$$

Using this bound in (3.16) along with the simple bound

$$h(k) \le \frac{g(u)}{u} \qquad (k > u/2)$$

leads to

$$\frac{1}{\log x} \sum_{\substack{y<p\le x\\ P(p-1)\le y}} \Phi\left(\frac{x}{p}, y\right) \log p \le \frac{c_0 \gamma_N x \log u \log_2 u}{u} \sum_{1\le k\le u/2} g(u-k) h(k) + \frac{c_0 \gamma_N x}{u} g(u). \tag{3.18}$$

To bound the sum in (3.18), we start with the estimate

$$\log g(u-k) = -(u-k)(\log_2 u + \log_3 u - \delta) + O\left(\frac{k}{\log u}\right), \tag{3.19}$$

which holds uniformly for $1 \le k \le u/2$. Using (3.19) and assuming that u_0 is sufficiently large depending on δ, we derive that

$$g(u-k)h(k) \le g(u) e^{B_u(k)} \qquad (1 \le k \le u/2), \tag{3.20}$$

where

$$B_u(k) = k(\log_2 u + \log_3 u - \log k - \log_2(k+1) + 1 - \delta/2).$$

Note that

$$\begin{aligned}\frac{dB_u(k)}{dk} &= \log_2 u + \log_3 u - \delta/2 \\ &\qquad - \log k - \log_2(k+1) - \frac{k}{(k+1)\log(k+1)} \\ &= \log\left(e^{-\delta/2}\frac{\log u}{k}\right) + \log\frac{\log_2 u}{\log(k+1)} - \frac{k}{(k+1)\log(k+1)}.\end{aligned}$$

Therefore, the function B_u reaches its maximum for some $k = k_0$ with

$$k_0 = e^{-\delta/2}\log u + O\left(\frac{\log u}{\log_2 u}\right)$$

and, since for a constant $C > 0$ the derivative is bounded independently of u for any k in the interval

$$k \in \left[e^{-\delta/2}\log u - C\frac{\log u}{\log_2 u}, e^{-\delta/2}\log u + C\frac{\log u}{\log_2 u}\right],$$

we obtain

$$\begin{aligned}\max_{1\le k\le u/2} B_u(k) &= B_u\left(e^{-\delta/2}\log u\right) + O\left(\frac{\log u}{\log_2 u}\right) \\ &= e^{-\delta/2}\log u + O\left(\frac{\log u}{\log_2 u}\right).\end{aligned}$$

This implies via (3.20) that

$$\max_{1\le k\le u/2} g(u-k)h(k) \le g(u)u^{1-\delta/3} \qquad (1 \le k \le u/2), \tag{3.21}$$

if u_0 is sufficiently large. Moreover, for any fixed constant $c > 1$, it is easy to see that B_u is decreasing for $k \ge c\log u$ if u_0 is sufficiently large depending on δ and c, and after a simple estimate we have

$$\max_{c\log u\le k\le u/2} B_u(k) \le (c - c\log c)\log u.$$

In particular, with $c = 3$ (and noting that $3 - 3\log 3 = -0.295\cdots$), this implies via (3.20) that

$$\max_{3\log u\le k\le u/2} g(u-k)h(k) \le g(u)u^{-1/4}. \tag{3.22}$$

Splitting the range of the summation in (3.18) according to whether $k \le 3\log u$ or $k > 3\log u$, and using (3.21) and (3.22), respectively, we have

$$\begin{aligned}\sum_{1\le k\le u/2} g(u-k)h(k) &\le \sum_{1\le k\le 3\log u} g(u)u^{1-\delta/3} + \sum_{3\log u<k\le u/2} g(u)u^{-1/4}\\ &\le 3g(u)u^{1-\delta/3}\log u + g(u)u^{3/4}\\ &\le g(u)u^{1-\delta/4}\end{aligned}$$

if δ is small enough and u_0 sufficiently large. Inserting this bound into (3.18), it follows that

$$\begin{aligned}\frac{1}{\log x}\sum_{\substack{y<p\le x\\ P(p-1)\le y}} \Phi\left(\frac{x}{p},y\right)\log p & \\ \le \frac{c_0\gamma_N x\log u\log_2 u}{u}g(u)u^{1-\delta/4} &+ \frac{c_0\gamma_N x}{u}g(u)\\ \le c_0\gamma_N u^{-\delta/5}g(u)x, &\end{aligned} \tag{3.23}$$

again assuming that u_0 is large.

Combining the bounds (3.13), (3.14) and (3.23) we obtain

$$\begin{aligned}\mathfrak{T}_2 \le \frac{2\gamma_N x(\log u\log_2 u)^2}{u}g(u) + c_0\gamma_N u^{-\delta/5}g(u)x &\\ + O\left(\frac{x}{\sqrt{y}\log x}\right). &\end{aligned} \tag{3.24}$$

We deduce from Lemma 3.1 and the bounds (3.12) and (3.24), that for u_0 large,

$$\Phi(x,y) \le \gamma_N g(u)x.$$

This establishes our claim (3.4), and the proposition is proved. □

4. Comments

The bound of Proposition 2.3, taken at the lower range with $y = (\log_2 x)^{1+\varepsilon}$, and thus with

$$u = \frac{\log x}{(1+\varepsilon)\log_3 x},$$

implies that

$$\begin{aligned}\Phi\left(x,(\log_2 x)^{1+\varepsilon}\right) &\le x\exp\left(-\frac{\log x}{1+\varepsilon} + O\left(\frac{\log x\log_4 x}{\log_3 x}\right)\right)\\ &= x^{\varepsilon/(1+\varepsilon)+o(1)},\end{aligned}$$

Hence $\Phi\left(x,\log_2 x\right) = x^{o(1)}$. Although we do not have any lower bounds for this range that are much better than the trivial bound $\Phi(x,y) \ge$

$\psi(x, y)$, this does suggest the existence of a phase transition near the point $y = \log_2 x$. Using the same heuristic as in Erdős [4], one should have quite small values of y with $\Phi(x, y) = x^{1-o(1)}$. In particular this should hold for any y of the shape $(\log x)^\varepsilon$, with $\varepsilon > 0$ fixed. It is interesting to recall that for the classical function $\psi(x, y)$ there is a well-known phase transition near the point $y = \log x$; see [3].

Acknowledgement

The authors are very grateful to Paul Kinlaw for pointing out a problem in our previous paper [1].

This work was partially supported by NSERC (Canada) Discovery Grant A5123 (for J.F.) and the Australian Research Council Grant DP170100786 (for I.S.).

References

1. W. D. Banks, J. B. Friedlander, C. Pomerance, and I. E. Shparlinski, *Multiplicative structure of values of the Euler function*, in High Primes and Misdemeanours: Lectures in honour of the sixtieth birthday of Hugh Cowie Williams, A. van der Poorten, ed., Fields Inst. Comm. **41** (2004), pp. 29–47. (pp. 1, 2, 4, and 15)
2. N. G. de Bruijn, *On the number of positive integers $\leq x$ and free of prime factors $> y$*, Nederl. Acad. Wetensch., Proc. Ser. A **54** (1951), 50–60. (p. 6)
3. ———, *On the number of positive integers $\leq x$ and free of prime factors $> y$, II*, Indag. Math. **28** (1966), 239–247. (pp. 3 and 15)
4. P. Erdős, *On the normal number of prime factors of $p - 1$ and some related problems concerning Euler's φ-function*, Q. J. Math., Oxford Ser. **6** (1935), 205–213. (p. 15)
5. T. Freiberg, *Products of shifted primes simultaneously taking perfect power values*, J. Aust. Math. Soc. **92** (2012) ,145–154. (p. 2)
6. T. Freiberg and C. Pomerance, *A note on square totients*, Int. J. Number Theory **11** (2015), 2265–2276. (p. 2)
7. A. J. Hildebrand, *On the number of positive integers $\leq x$ and free of prime factors $> y$*, J. Number Theory **22** (1986), 289–307. (p. 2)
8. A. J. Hildebrand and G. Tenenbaum, *Integers without large prime factors*, J. Théor. Nombres Bordeaux **5** (1993), 411–484. (p. 3)
9. Y. Lamzouri, *Smooth values of the iterates of the Euler φ-function*, Canadian J. Math. **59** (2007), 127–147. (p. 2)
10. P. Pollack, *Popular subsets for Euler's φ-function*, Preprint, 2018. (p. 2)
11. ———, *How often is Euler's totient a perfect power?*, Preprint, 2018. (p. 2)
12. P. Pollack and C. Pomerance, *Square values of Euler's function*, Bull. London Math. Soc. **46** (2014), 403–414. (p. 2)
13. C. Pomerance, *Two methods in elementary analytic number theory*, in Number theory and applications, R. A. Mollin, ed., Kluwer Academic Publishers, Dordrecht, 1989, pp. 135–161. (p. 7)

14. ______, *On amicable numbers*, in Analytic number theory (in honor of Helmut Maier's 60th birthday), M. Rassias and C. Pomerance, eds., Springer, Cham, Switzerland, 2015, pp. 321–327. (p. 2)
15. C. Pomerance and I. E. Shparlinski, *Smooth orders and cryptographic applications*, Proc. ANTS-V, Sydney, Australia, Springer Lecture Notes in Computer Science **2369**, (2002), pp. 338–348. (pp. 1 and 4)

Department of Mathematics, University of Missouri, Columbia, MO 65211, USA
E-mail address: bankswd@math.missouri.edu

Department of Mathematics, University of Toronto, Toronto, Ontario M5S 2E4, Canada
E-mail address: frdlndr@math.utoronto.ca

Mathematics Department, Dartmouth College, Hanover, NH 03755, USA
E-mail address: carl.pomerance@dartmouth.edu

Department of Pure Mathematics, University of New South Wales, Sydney, NSW 2052, Australia
E-mail address: igor.shparlinski@unsw.edu.au