Linear equations with the Euler totient function

by

Florian Luca (Morelia) and Pantelimon Stănică (Monterey, CA)

1. Introduction. In [1], Bager called an integer $n \geq 3$ a *Phibonacci* number if $\phi(n) = \phi(n-1) + \phi(n-2)$, where ϕ is the Euler function. He asked if there are any composite Phibonacci numbers. A quick computer search reveals that $n = 1037 = 17 \cdot 61$ is a composite Phibonacci number. It is still open whether there are any even Phibonacci numbers: if they exist, they should be greater than 10^{1600} (see [1]).

In this paper, we prove a general result concerning linear relations among the values of the Euler function at nearby integers. As a byproduct of this result, it will follow that for most positive integers n the sums of the Euler functions at integers close to n are distinct. In particular, the set of Phibonacci numbers is of asymptotic density zero. We also look at the subset of Phibonacci numbers which are primes, and we show that the sum of the reciprocals of the members of this set is finite. Hence, either this set is finite, or infinite but the series of reciprocals of its members is convergent. Similar results hold with the Euler function ϕ replaced by the sum of divisors function σ.

Problems of a similar nature were considered previously, most notably in the series of papers [5]–[8], for the sets of positive integers n such that $\phi(n) = \phi(n+1)$, or $\sigma(n) = \sigma(n+1)$, or $\omega(n) = \omega(n+1)$, or $\Omega(n) = \Omega(n+1)$ or $\tau(n) = \tau(n+1)$, where σ, ω, Ω and τ are the sum of divisors, the number of prime divisors, the number of prime power (> 1) divisors and the total number of divisors functions of n, respectively. All such sets of positive integers were shown to have asymptotic density zero. Our paper is inspired by the papers in the above series.

Throughout this paper, we use the Vinogradov symbols $\gg$ and $\ll$ and the Landau symbols O and o with their usual meanings. The constants implied by such symbols are absolute. We write x for a large positive real number, and p and q for prime numbers. If $\mathcal{A}$ is a set of positive integers,

we write $\mathcal{A}(x) = \mathcal{A} \cap [1, x]$. We write $\log x$ for the maximum between the natural logarithm of x, denoted $\ln x$, and 1. Thus, all logarithms which will appear are ≥ 1. We use $c_1, c_2, \ldots$ for positive computable constants which are absolute.

2. The results. Let $t \geq 1$ be a positive integer. Let $\mathbf{a} = (a_0, \ldots, a_t)$ be a vector with integer components, not all zero. Write $\mathcal{A}_{\mathbf{a}}$ for the set of all positive integers n such that

$$(1) \qquad \sum_{i=0}^{t} a_i \phi(n+i) = 0.$$

Here and in what follows, if $\mathcal{A}$ is a subset of the positive integers and b is a positive integer, then $\mathcal{A} + b = \{a + b : a \in \mathcal{A}\}$. Put $H(\mathbf{a}) = \max\{|a_i| : i = 0, \ldots, t\}$. Given a positive real number x we put $y = \exp(\log x / \log \log x)$. Note that $\mathcal{A}_{(1,-1)}$ coincides with the set of positive integers n such that $\phi(n) = \phi(n+1)$, and that the set of Phibonacci numbers is $\mathcal{A}_{(1,1,-1)} + 2$.

THEOREM 2.1. *Let $C(t, \mathbf{a}) = t^3 \log H(\mathbf{a})$. Then the estimate*

$$\#\mathcal{A}_{\mathbf{a}}(x) \ll C(t, \mathbf{a}) \frac{x \log \log \log x}{\sqrt{\log \log x}}$$

holds uniformly in $\mathbf{a}$ and $1 \leq t < y$.

Note that the above estimate is nontrivial only when

$$t < c_0(\mathbf{a}) \frac{(\log \log x)^{1/6}}{(\log \log \log x)^{1/3}},$$

where $c_0(\mathbf{a})$ is some constant depending on $\mathbf{a}$.

COROLLARY 2.2. *Let $\varepsilon > 0$ be fixed. Then, for a set of positive integers n of asymptotic density one, the numbers*

$$\sum_{i \in \mathcal{I}} \phi(n+i) \quad \text{for } \mathcal{I} \subset \{0, 1, \ldots, t(n)\},$$

where $t(n) = \lfloor (1/(4 \ln 2) - \varepsilon) \log \log \log n \rfloor$, are all distinct.

Now let

$$\mathcal{P}_{\mathbf{a}} = \{n \in \mathcal{A}_{\mathbf{a}} : n + i \text{ is prime for some } i = 0, \ldots, t \text{ with } a_i \neq 0\}.$$

We have the following result.

THEOREM 2.3. *Let $D(t, \mathbf{a}) = tH(\mathbf{a}) + t^2 \log\log|t|$. Then the estimate*

$$\mathcal{P}_{\mathbf{a}}(x) \ll D(t, \mathbf{a}) \frac{x}{(\log x)^{5/4}}$$

holds uniformly in $\mathbf{a}$ with $H(\mathbf{a}) \leq y$, and $1 \leq t < y$.

Let $\mathcal{P}$ be the set of prime Phibonacci numbers. Note that $\mathcal{P}$ is contained in $\mathcal{P}_{(1,1,-1)} + 2$. By partial summation, Theorem 2.3 implies that

$$\sum_{p \in \mathcal{P}} \frac{1}{p} < \infty.$$

Also, while Theorem 2.1 shows that the set of Phibonacci numbers is of asymptotic density zero, the upper bound on the counting function of this set is not strong enough to allow us to decide whether the sum of the reciprocals of all the Phibonacci numbers is convergent. We would like to propose this as a conjecture.

CONJECTURE 2.4. *Show that $\sum_{n \, Phibonacci} 1/n < \infty$.*

3. The proofs. For a positive integer n we write $P(n)$ and $p(n)$ for the largest and smallest prime factor of n, respectively. Recall that a positive integer m is called *powerfull* if $p^2 \,|\, m$ whenever $p \,|\, m$.

For the proof of Theorem 2.1, we need the following lemma.

LEMMA 3.1. *Assume that x is a positive real number. Put $z = \log\log x$ and $w = \lfloor c_1 \log\log\log x \rfloor$, where $c_1 = 3/(2\ln 2)$. Let $\mathcal{A}(x)$ be the set of positive integers $n \leq x$ with the following properties:*

(i) *If $d \,|\, n$ is powerfull, then $d < z$.*
(ii) *If $d \,|\, n$ and $d > x^{1/3}$, then $P(d) > y$.*
(iii) *If $p \,|\, n$ is a prime, then $p - 1$ is not divisible by 2^w.*

Then $\mathcal{A}(x)$ contains all positive integers $n \leq x$ with $O(x/\sqrt{\log\log x})$ exceptions.

Proof. We start with an upper bound on the set

$$\mathcal{A}_1(x) = \{n \leq x : n \text{ fails condition (i)}\}.$$

For each $n \in \mathcal{A}_1(x)$, there exists a powerfull $d > z$ dividing n. For a fixed value of d, the number of such $n \leq x$ is $\leq x/d$. Summing up over all d, we get

$$(2) \qquad \#\mathcal{A}_1(x) \leq x \sum_{\substack{d > z \\ d \, \text{powerfull}}} \frac{1}{d} \ll \frac{x}{\sqrt{z}} = \frac{x}{\sqrt{\log\log x}},$$

where the last estimate follows by partial summation from the well-known estimate

$$\{m \leq s : m \text{ powerfull}\} \ll s^{1/2}.$$

(see, for example, Theorem 14.4 in [10]). Now let

$$\mathcal{A}_2(x) = \{n \le x : n \text{ fails condition (ii)}\}.$$

Put

$$\mathcal{B}(s) = \{n \le s : P(n) \le y\}.$$

It is well-known (see, for example, Section III.5.4 in Tenenbaum's book [11]) that

(3) $\#\mathcal{B}(s) = \Psi(s, y) = x \exp(-(1 + o(1))u \log u)$,

where $u = \log s/\log y$, uniformly in $s \in [x^{1/3}, x]$.

   To estimate $\mathcal{A}_2(x)$, let $d > x^{1/3}$ be such that $P(d) < y$. The number of $n \le x$ which are multiples of d is $\le x/d$. Thus,

$$\#\mathcal{A}_2(x) \le x \sum_{\substack{x^{1/3} < d \le x \\ P(d) < y}} \frac{1}{d} = x \int_{x^{1/3}}^{x} \frac{d(\Psi(s, y))}{s}.$$

By partial integration and using estimate (3), we find that with

$$v = \log(x^{1/3})/\log y = (1/3) \log \log x,$$

we have the estimate

(4) $\#\mathcal{A}_2(x) \le x \log x \exp(-(1 + o(1))v \log v) = o\left(\dfrac{x}{\sqrt{\log \log x}}\right).$

Finally, let

$$\mathcal{A}_3(x) = \{n \le x : p \mid n \text{ for some } p \text{ with } 2^w \mid p - 1\}.$$

If we fix such a prime p, then the number of $n \le x$ which are multiples of p is $\le x/p$. Thus,

(5) $\#\mathcal{A}_3(x) \le x \sum_{\substack{p \le x \\ p \equiv 1 \,(\mathrm{mod}\, 2^w)}} \frac{1}{p} \ll \dfrac{x \log \log x}{2^w} \ll \dfrac{x}{\sqrt{\log \log x}},$

where we have used the known fact that

(6) $\sum_{\substack{p \le s \\ p \equiv 1 \,(\mathrm{mod}\, d)}} \frac{1}{p} \ll \dfrac{\log \log s}{\phi(d)}$

uniformly for $1 \le d \le s$ (see the bound (3.1) in [4] or Lemma 1 in [2]). The conclusion of the lemma follows from estimates (2), (4) and (5). ∎

   Proof of Theorem 2.1. We assume that $1 \le t \le n - t$. For a nonzero integer n we write $\nu(n)$ for the order at which 2 divides n.

   If $n = \prod_{p^\alpha \| n} p^\alpha$, we also put

$$\nu'(n) = \nu(n) + \sum_{p^\alpha \| n} \alpha \nu(p - 1).$$

It is clear that $\nu'(n)$ is strongly additive.

We look only at those positive integers n such that

$$n \in \mathcal{A}_0(x) = \bigcap_{i=0}^{t}(\mathcal{A}(x) - i),$$

where $\mathcal{A}(x)$ is the set appearing in Lemma 3.1. In particular, $n + i \in \mathcal{A}(x)$ for all $i = 0, 1, \ldots, t$. By Lemma 3.1, the number of excluded integers up to x does not exceed

$$O\left(\frac{tx}{\sqrt{\log\log x}}\right)$$

uniformly for $1 \leq t \leq x$. Hence, it suffices to prove the estimate stated by the theorem only for those positive integers n in $\mathcal{A}_0(x)$.

For a positive integer n we write $m(n)$ for the largest powerfull divisor of n. Then

$$\begin{align}
(7) \qquad \nu'(n+i) &= \nu(\phi(n+i)) + 1 + \sum_{\substack{p^\alpha \| n+i \\ \alpha > 1, p > 2}} (\alpha - 1)\nu(p-1) \\
&= \nu(\phi(n+i)) + O(\log(m(n+i))) \\
&= \nu(\phi(n+i)) + O(\log z) \\
&= \nu(\phi(n+i)) + O(\log\log\log x).
\end{align}$$

We recall the following obvious fact.

LEMMA 3.2. *Let $(b_i)_{i=0}^{t}$ be a finite sequence of nonzero integers. If*

$$\sum_{i=0}^{t} b_i = 0,$$

then there exists $i < j$ such that $\nu(b_i) = \nu(b_j)$.

Assume now that $n \in \mathcal{A}_{\mathbf{a}}(x) \cap \mathcal{A}_0(x)$. By Lemma 3.2, there exist $i < j$ such that $a_i a_j \neq 0$ and $\nu(a_i\phi(n+i)) = \nu(a_j\phi(n+j))$. We fix such i and j. We then get

$$|\nu(\phi(n+i)) - \nu(\phi(n+j))| = |\nu(a_i) - \nu(a_j)| \leq A,$$

where $A = \lfloor \log H(\mathbf{a})/\ln 2 \rfloor$. Together with estimate (7), we arrive at the conclusion that

$$(8) \qquad |\nu'(n+i) - \nu'(n+j)| = O(A + \log\log\log x) \ll A\log\log\log x.$$

We now fix $j \in \{1, \ldots, t\}$ and put

$$\mathcal{A}_{0,j}(x) = \{n \in \mathcal{A}_0(x) : |\nu'(n) - \nu'(n+j)| \leq c_2 A\log\log\log x\},$$

where c_2 is the constant implied by inequality (8). Note that if n satisfies (8), then $n + i \in \mathcal{A}_{0,j-i}(x)$. Thus, it follows that in order to prove the estimate

claimed by Theorem 2.1, it suffices to show that

$$
(9) \qquad \#\mathcal{A}_{0,j}(x) \ll At \, \frac{\log\log\log x}{\sqrt{\log\log x}}
$$

uniformly in $1 \le t \le x$. There are two possible ways to proceed in order to prove (9). One possibility is to use the Barban–Vinogradov theorem on the joint distribution of the strongly additive functions $\nu'(n)$ and $\nu'(n+j)$ [3, Theorem 20.1, p. 262]. In fact, if t is fixed, then (9) follows, even with an extra saving of $1/\log\log\log\log x$, directly from the statement of the above theorem in [3] and Lemma 3.1. Unfortunately, for the purpose of that theorem the shift j must be fixed, whereas we want our result to be uniform in it as well. The second method, which is the one we choose to follow, was used in [6] to deal with the set of n such that $\omega(n) = \omega(n+1)$.

Given $n \in \mathcal{A}_{0,j}(x)$, define the integers a, b, κ, ℓ by

$$
\begin{aligned}
n + j &= a\kappa, & P(a) &\le p(\kappa), & a &\le x^{1/3}, & ap(\kappa) &> x^{1/3}; \\
n &= b\ell, & P(b) &\le p(\ell), & b &\le x^{1/3}, & bp(\ell) &> x^{1/3}.
\end{aligned}
$$

We shall assume that $p(\kappa) \le p(\ell)$, the case $p(\kappa) > p(\ell)$ being similar. For $1 \le s \le x^{1/3}$, let $\mathcal{N}(s)$ be the number of $n \in \mathcal{A}_{0,j}(x)$ with $s \le p(\kappa) < s^3$. Note that if we set $s_u = x^{3^{-u}}$, we have

$$
(10) \qquad \#\mathcal{A}_{0,j}(x) \le \sum_{u=1}^{\infty} \mathcal{N}(s_u).
$$

Note also that since $n + j \in \mathcal{A}_0(x)$, it follows that $ap(\kappa) > x^{1/3}$ is a divisor of $n+j$, and by property (i) of the set $\mathcal{A}(x)$, it follows that $p(\kappa) = P(ap(\kappa)) > y$. Thus, if n is counted by $\mathcal{N}(s_u)$, then $s_u^3 > y$. Hence, $x^{3^{-u+1}} > y$, which leads to $3^{u-1} < \log\log x$, therefore $u < 1 + \log\log\log x/\log 3$. In particular, the sum appearing on the right hand side of estimate (10) is finite.

We now turn our attention to estimating $\mathcal{N}(s)$. If n is counted by $\mathcal{N}(s)$, then the numbers a, b, κ and ℓ defined above satisfy

$$
(11) \qquad
\begin{aligned}
a\kappa - b\ell &= j, & a\kappa &< x, & |\nu'(a\kappa) - \nu'(b\ell)| &\ll A\log\log\log x, \\
s &\le p(\kappa) < p(\ell), & a &\le x^{1/3}, & b &\le x^{1/3}, \\
P(a) &\le p(\kappa) \le s^3, & a &> x^{1/3}/p(\kappa) > x^{1/3}/s^3.
\end{aligned}
$$

Since all of the primes in κ or ℓ are at least s, we have

$$
\omega(\kappa) \le \frac{\log x}{\log s} \quad \text{and} \quad \omega(\ell) \le \frac{\log x}{\log s}.
$$

Since both n and $n + j$ satisfy property (ii) of Lemma 3.1, it follows that

$$
\nu'(\kappa) \le w\omega(\kappa) \ll \frac{\log x \log\log\log x}{\log s}
$$

and a similar upper bound holds for $\nu'(\ell)$. Thus,

$$|\nu'(a) - \nu'(b)| \leq |\nu'(a\kappa) - \nu'(b\ell)| + |\nu'(\kappa) - \nu'(\ell)| \leq c_3 A \frac{\log x \log \log \log x}{\log s},$$

where $c_3 > 0$ is an absolute constant. We now fix the integers a, b appearing in (11) and count the number of pairs κ, ℓ there can be. Let $d = \gcd(a, j)$. Since $j \leq t$, $\min\{p(\kappa), p(\ell)\} \geq y > t$, it follows that ℓ is coprime to d, therefore $d = \gcd(b, j)$. Let $a = da_0$, $b = db_0$, $j = dj_0$, and let κ_0, ℓ_0 denote the unique integers which satisfy

$$a_0 \kappa_0 - b_0 \ell_0 = j_0, \quad 0 \leq \kappa_0 < b_0, \quad 0 \leq \ell_0 \leq a_0.$$

Hence, if κ, ℓ satisfy (11), we have some integer m with

$$(12) \quad \begin{aligned} &\kappa = b_0 m + \kappa_0, \quad \ell = a_0 m + \ell_0, \\ &0 \leq m \leq \frac{x}{ab_0} = \frac{xd}{ab}, \quad p((b_0 m + \kappa_0)(a_0 m + \ell_0)) \geq s. \end{aligned}$$

Thus, it suffices to count the number of m satisfying (12). This is easily done as on page 3 in [6], by using either Brun's method or Selberg's sieve (see, e.g., [9, Theorem 3.1, p. 101]). Noting that $a \leq x^{1/3}$ and $b_0 \leq b \leq x^{1/3}$, we find that the number of such m is

$$\leq \frac{c_4 x}{\phi(a)\phi(b_0)(\log s)^2} \leq \frac{c_4 x \phi(d)}{\phi(a)\phi(b)(\log s)^2},$$

where $c_4 > 0$ is an absolute constant. Summing up over all possible values of $d \mid j$ and using the fact that

$$\sum_{d \mid j} \phi(d) = j,$$

we see that the number of such integers m when a and b are fixed is

$$\leq \frac{c_4 x j}{\phi(a)\phi(b)(\log s)^2} \leq \frac{c_4 x t}{\phi(a)\phi(b)(\log s)^2}.$$

Thus, from (11) and (12), we have

$$(13) \quad \mathcal{N}(s) \leq \sum_{\substack{a > x^{1/3}/s^3 \\ P(a) < s^3}} \sum_{\substack{b \leq x^{1/3} \\ |\nu'(a) - \nu'(b)| \leq \frac{Ac_3 \log x \log \log \log x}{(\log s)^2}}} \frac{c_4 x t}{\phi(a)\phi(b)(\log s)^2}$$

$$= \frac{c_4 x t}{(\log s)^2} \sum_{\substack{a > x^{1/3}/s^3 \\ P(a) < s^3}} \frac{1}{\phi(a)} \sum_{|t - \nu'(a)| \leq \frac{Ac_3 \log x \log \log \log x}{(\log s)^2}} \sum_{\substack{b \leq x^{1/3} \\ \nu'(b) = t}} \frac{1}{\phi(b)}.$$

We will need the following lemma.

LEMMA 3.3. *Let t be any fixed positive integer. Let*

$$\mathcal{B}_t(x) = \{n \leq x : \nu'(n) = t\}.$$

Then

$$\sum_{n\in\mathcal{B}_t(x)} \frac{1}{\phi(b)} \leq \frac{c_5\log x}{\sqrt{\log\log x}}.$$

Assume for the moment that Lemma 3.3 is proved. Then Lemma 3.3 together with (13) show that

$$(14) \qquad \mathcal{N}(s) \leq \frac{c_6 xt(\log x)^2\log\log\log x}{\sqrt{\log\log x}(\log s)^2} \sum_{\substack{a>x^{1/3}/s^3 \\ P(a)<s^3}} \frac{1}{\phi(a)}.$$

The lemma on p. 3 of [6] shows that

$$\sum_{\substack{a>x^{1/3}/s^3 \\ P(a)<s^3}} \frac{1}{\phi(a)} \leq c_7\exp\left(-\frac{\log(x^{1/3}/s^3)}{6\log s}\right).$$

Recalling that $s_u = x^{3^{-u}}$, this gives

$$\mathcal{N}(s_u) \leq \frac{c_8 xt\log\log\log x}{\sqrt{\log\log x}} 3^{2u}\exp\left(-\frac{1}{6}\left(3^{u-1}-3\right)\right),$$

and now summing over u, using the fact that the series

$$\sum_{u\geq 0} 3^{2u}\exp\left(-\frac{1}{6}\left(3^{u-1}-3\right)\right)$$

converges, as well as estimate (10), we get the desired estimate (9).

Proof of Lemma 3.3. The function ν' is strongly additive. Furthermore,

$$(15) \qquad E(x) = \sum_{p^\alpha\leq x} \frac{\nu'(p^\alpha)}{p^\alpha} = \sum_{2<p\leq x} \frac{\nu'(p)}{p} + O\left(\sum_{p\geq 2} \frac{\log p}{p^2}\right)$$

$$= \sum_{a\geq 1} a \sum_{\substack{1\leq p\leq x \\ p\equiv 1+2^a\,(\mathrm{mod}\,2^{a+1})}} \frac{1}{p} + O(1) \ll \log\log x\cdot\left(\sum_{a\geq 1} \frac{a}{2^a}\right)$$

$$\ll \log\log x,$$

where we have again used estimate (6) for all $d = 2^a \leq x$.

By Halász's theorem (see Theorem 21.5, p. 303 in [3]),

$$\#\mathcal{B}_t(x) \ll \frac{x}{\sqrt{E(x)}} \ll \frac{x}{\sqrt{\log\log x}}.$$

In particular, by partial summation, we get

$$(16) \qquad \sum_{n \in \mathcal{B}_t(x)} \frac{1}{n} \leq \sum_{n \leq y} \frac{1}{n} + \int_y^x \frac{d(\#\mathcal{B}_t(s))}{s}$$

$$\leq \log y + \frac{\#\mathcal{B}_t(s)}{s}\Big|_{s=y}^{s=x} + \int_y^x \frac{\#\mathcal{B}_t(s)}{s^2}\, ds$$

$$\ll \frac{\log x}{\log \log x} + O(1) + \int_y^x \frac{ds}{s\sqrt{\log \log s}}$$

$$\ll \frac{\log x}{\log \log x} + O(1) + \frac{1}{\sqrt{\log \log x}} \int_y^x \frac{ds}{s}$$

$$\ll \frac{\log x}{\sqrt{\log \log x}},$$

uniformly in t. We now use the known fact that

$$\frac{1}{\phi(n)} \ll \frac{\sigma(n)}{n^2} = \frac{1}{n}\sum_{d|n}\frac{1}{d},$$

to get

$$\sum_{n \in \mathcal{B}_t(x)} \frac{1}{\phi(n)} \ll \sum_{n \in \mathcal{B}_t(x)} \frac{1}{n}\sum_{d|n}\frac{1}{d},$$

which, by making the substitution $n = md$, changing the order of summation, and using the fact that ν' is strongly additive, becomes

$$\sum_{n \in \mathcal{B}_t(x)} \frac{1}{n} \leq \sum_{1 \leq d \leq x} \frac{1}{d^2} \sum_{m \in \mathcal{B}_{t-\nu'(d)}} \frac{1}{m}.$$

From estimate (16), we immediately get

$$\sum_{n \in \mathcal{B}(x)} \frac{1}{\phi(n)} \ll \sum_{1 \leq d \leq x} \frac{\log(x/d)}{d^2 \sqrt{\log \log(x/d)}}$$

$$\leq \frac{\log x}{\sqrt{\log \log x}} \sum_{d \geq 1} \frac{1}{d^2} \ll \frac{\log x}{\sqrt{\log \log x}},$$

which completes the proof of Lemma 3.3 and of Theorem 2.1. ∎

Proof of Corollary 2.2. Let $\varepsilon > 0$ be given and put

$$t(x) = \lfloor (1/(4\ln 2) - \varepsilon)\log\log\log x \rfloor.$$

Given distinct subsets $\mathcal{I}$ and $\mathcal{J}$ of $\{0, 1, \ldots, t(x)\}$, let $\mathcal{A}_{\mathcal{I},\mathcal{J}}(x)$ be the set of $n \leq x$ such that

$$\sum_{i \in \mathcal{I}} \phi(n + i) = \sum_{j \in \mathcal{J}} \phi(n + j).$$

Clearly, we may assume that $\mathcal{I}$ and $\mathcal{J}$ are disjoint and then the above relation is a recurrence of type (1) with $t = t(x)$ and $a_i \in \{0, \pm 1\}$. Since $t(x) < y$ for large x, Theorem 2.1 yields

$$\#\mathcal{A}_{\mathcal{I},\mathcal{J}}(x) \ll \frac{x(\log\log\log x)^4}{\sqrt{\log\log x}}.$$

Since the number of pairs $\mathcal{I}$, $\mathcal{J}$ of distinct subsets of $\{0, \ldots, t(x)\}$ does not exceed $2^{2t(x)+2}$, we get

$$\#\Big(\bigcup_{\substack{\mathcal{I},\mathcal{J} \subset \{0,\ldots,t(x)\} \\ \mathcal{I} \neq \mathcal{J}}} \mathcal{A}_{\mathcal{I},\mathcal{J}}(x) \Big) \ll \frac{x(\log\log\log x)^4 2^{2t(x)}}{\sqrt{\log\log x}}$$

$$= \frac{x(\log\log\log x)^4}{(\log\log x)^{2\varepsilon\ln 2}} = o(x).$$

The conclusion of the corollary follows from the above estimate by observing that if $n \leq x$ is not in the union of $\mathcal{A}_{\mathcal{I},\mathcal{J}}(x)$ over all pairs of distinct subsets $\mathcal{I}$, $\mathcal{J}$ of $\{0, \ldots, t(x)\}$, then all sums $\sum_{i \in \mathcal{I}} \phi(n+i)$ for $\mathcal{I} \in \{0, 1 \ldots, t(n)\}$ are distinct. ∎

Proof of Theorem 2.3. We put $\mathcal{B}(x) = \{n \leq x : P(n) \leq y\}$. By estimate (3), it follows easily that $\#\mathcal{B}(x) = o(x/(\log x)^2)$. We write $\mathcal{P}_1(x)$ for the set of $n \leq x - t$ such that $n + i \notin \mathcal{B}(x)$ for any $i = 0, \ldots, t$. Clearly, the number of positive integers $n \leq x$ such that either $n > x - t$ or n is not in $\mathcal{P}_1(x)$ is

$$(17) \qquad x - \#\mathcal{P}_1(x) \leq t + (t+1)\#\mathcal{B}(x) \ll \frac{tx}{(\log x)^2}.$$

We now let $\alpha \in (0, 1)$ be a constant to be determined later, let $i \neq j$ with $a_i \neq 0$ be in $\{0, \ldots, t\}$, and put

$$\mathcal{P}_{2,i,j}(x) = \{n \in \mathcal{P}_1(x) : n + i \text{ is prime and } \nu(\phi(n+j)) < \alpha\log\log x\},$$

and

$$\mathcal{P}_{3,i}(x) = \{n \in \mathcal{P}_1(x) : n + i \text{ is prime and}$$
$$\nu(\phi(n+j)) > \alpha\log\log x \text{ for all } j \neq i\}.$$

It is clear that the union of $\mathcal{P}_{2,i,j}(x)$ for all $i \neq j$ with $a_i \neq 0$ together with the union of $\mathcal{P}_{3,i}(x)$ for all i with $a_i \neq 0$ make up $\mathcal{P}_{\mathbf{a}}(x) \cap \mathcal{P}_1(x)$.

We now estimate $\mathcal{P}_{2,i,j}(x)$. If $n \in \mathcal{P}_{2,i,j}(x)$, then $P = n + i = (n+j) + (i-j)$ is a prime and $n + j = Qm$, where $Q \geq P(m) > y > |i-j|$ and $\omega(m) \leq \nu(\phi(n+j)) + 1 \leq K = \lfloor\alpha\log\log x\rfloor + 1$. This leads to the equation

$$P - (i-j) = Qm,$$

where P and Q are primes, and $Q \leq x/m$. Fixing m and noting that $i - j$ is coprime to p (because $P = n + i = P(n+i) > y > t > |i-j|$), it follows,

by Brun's sieve, that the number of solutions (P, Q) of the above equation once i, j and m are fixed is

$$\leq \frac{c_1 x}{\phi(m)(\log(x/m))^2} \cdot \frac{|i - j|}{\phi(|i - j|)}.$$

Since $|i - j| \leq t$, we have $|i - j|/\phi(|i - j|) \ll \log \log t$. Furthermore, since $x/m > Q > y$, we get $\log(x/m) > \log y = \log x/\log \log x$. Thus, the number of solutions does not exceed

$$\frac{c_2 (\log \log x)^2 \log \log t}{\phi(m)(\log x)^2}.$$

Summing up over all possible values of $m \leq x$ with $\omega(m) \leq K$, we get

$$(18) \qquad \#\mathcal{P}_{2,i,j}(x) \leq \frac{c_2 (\log \log x)^2 (\log \log t)}{(\log x)^2} \sum_{k=0}^{K} \sum_{\substack{m \leq x \\ \omega(\phi(m))=k}} \frac{1}{\phi(m)}.$$

It is easy to see from the multinomial formula that

$$\sum_{\substack{m \leq x \\ \omega(m)=k}} \frac{1}{\phi(m)} \leq \frac{1}{k!} \left(\sum_{p^\alpha \leq x} \frac{1}{\phi(p^\alpha)} \right)^k$$

$$\leq \frac{1}{k!} \left(\sum_{p \leq x} \frac{1}{p-1} + \sum_{\alpha \geq 2} \sum_{p \geq 2} \frac{1}{p^{\alpha-1}(p-1)} \right)$$

$$= \frac{1}{k!} (\log \log x + O(1))^k.$$

Using the fact that $k! > (k/e)^k$, we get

$$\sum_{\substack{m \leq x \\ \omega(m)=k}} \frac{1}{\phi(m)} \leq \left(\frac{e \log \log x + O(1)}{k} \right)^k.$$

The function $k \mapsto ((e \log \log x + O(1))/k)^k$ is increasing for $k \leq K \leq \alpha \log \log x + 2$ for large x, because $\alpha < 1$. Hence,

$$\sum_{\substack{m \leq x \\ \omega(m)=k}} \frac{1}{\phi(m)} \leq \left(\frac{e \log \log x + O(1)}{k} \right)^k$$

$$\leq \left(\frac{e \log \log x + O(1)}{\alpha \log \log x + 2} \right)^{\alpha \log \log x + 2} \ll (\log x)^\beta,$$

where $\beta = \alpha \ln(e/\alpha)$. Inserting the above estimate into (18), we get

$$(19) \qquad \#\mathcal{P}_{2,i,j}(x) \leq \frac{c_3 (\log \log x)^3 \log \log t}{(\log x)^{2-\beta}},$$

and summing up over all possible values for i and j, we get

$$(20) \qquad \sum_{\substack{0 \le i \le t \\ a_i \ne 0}} \sum_{\substack{0 \le j \le t \\ j \ne i}} \#\mathcal{P}_{2,i,j}(x) \ll \frac{c_3 t^2 (\log\log x)^3 \log\log t}{(\log x)^{2-\beta}}.$$

We now estimate $\#\mathcal{P}_{3,i}(x)$. Let $n \in \mathcal{P}_{3,i}(x)$. Reducing equation (1) modulo 2^K, we get $a_i(p-1) \equiv 0 \ (\mathrm{mod}\ 2^K)$, where $p = n + i$. Since $a_i \ne 0$, it follows that if we write $\alpha_i = \nu(a_i)$, then $p - 1 \equiv 0 \ (\mathrm{mod}\ 2^{K-\alpha_i})$. The number of such primes $p \le x$ is

$$\le \frac{x}{\phi(2^{K-\alpha_i}) \log(x/2^{K-\alpha_i})} \ll \frac{x 2^{\alpha_i}}{2^K \log x} \le \frac{x H(\mathbf{a})}{(\log x)^\gamma},$$

where $\gamma = 1 + \alpha \ln 2$. Summing up over all possible i, we get

$$(21) \qquad \sum_{\substack{0 \le i \le t \\ a_i \ne 0}} \#\mathcal{P}_{3,i}(x) \ll \frac{t H(\mathbf{a})}{(\log x)^\gamma}.$$

Choosing α such that $\beta = \gamma$, we deduce from (17), (20) and (21) that

$$\#\mathcal{P}_{\mathbf{a}}(x) \le D(t, \mathbf{a}) \frac{x(\log\log x)^3}{(\log x)^\beta}.$$

The equation $\beta = \gamma$ leads to $2 - \alpha \ln(e/\alpha) = 1 + \alpha \ln 2$, whose solution in the interval $(0,1)$ is $\alpha = 0.373365\dots$, leading to $\beta = \gamma > 1.25887\dots > 5/4$, which completes the proof of the theorem. ∎

References

[1]   A. Bager, *Proposed Problem AMME2833*, Amer. Math. Monthly 87 (1980), 404; a solution, ibid. 88 (1981), 622.
[2]   N. L. Bassily, I. Kátai and M. Wijsmuller, *On the prime power divisors of the iterates of the Euler-ϕ function*, Publ. Math. Debrecen 55 (1999), 17–32.
[3]   P. D. T. A. Elliott, *Probabilistic Number Theory. II*, Springer, New York, 1980.
[4]   P. Erdős, A. Granville, C. Pomerance and C. Spiro, *On the normal behavior of the iterates of some arithmetic function*, in: Analytic Number Theory, Birkhäuser, Boston, 1990, 165–204.
[5]   P. Erdős, C. Pomerance and A. Sárközy, *On locally repeated values of certain arithmetic functions. II*, Acta Math. Hungar. 49 (1987), 251–259.
[6]   —, —, —, *On locally repeated values of certain arithmetic functions. III*, Proc. Amer. Math. Soc. 101 (1987), 1–7.
[7]   —, —, —, *On locally repeated values of certain arithmetic functions. IV*, Ramanujan J. 1 (1997), 227–241.
[8]   P. Erdős, A. Sárközy and C. Pomerance, *On locally repeated values of certain arithmetic functions. I*, J. Number Theory 21 (1985), 319–332.
[9]   H. Halberstam and H.-E. Richert, *Sieve Methods*, Academic Press, London, 1974.
[10] A. Ivić, *The Riemann Zeta-Function, Theory and Applications*, Dover Publ., Mineola, NY, 2003.

[11] G. Tenenbaum, *Introduction to Analytic and Probabilistic Number Theory*, Cambridge Univ. Press, 1995.

Instituto de Matemáticas
Universidad Nacional Autónoma de México
C.P. 58089, Morelia, Michoacán, México
E-mail: fluca@matmor.unam.mx

Applied Mathematics Department
Naval Postgraduate School
Monterey, CA 93943, U.S.A.
E-mail: pstanica@nps.edu