ACTA ARITHMETICA
Online First version

The set of values of any finite iteration of Euler's φ function contains long arithmetic progressions

by

R. Balasubramanian (Chennai),
Jean-Marc Deshouillers (Bordeaux) and Sanoli Gun (Chennai)

To Henryk Iwaniec, on his 75th birthday,
with admiration and friendship

Abstract. Assuming the validity of Dickson's conjecture, we show that the set of values of iterated Euler's totient φ function $\varphi \circ \cdots \circ \varphi$ (n times) contains arbitrarily long arithmetic progressions with an explicitly given common difference D_a depending only on a. This extends a previous result (case $a = 1$) of Deshouillers, Eyyunni and Gun. In particular, this implies that this set has upper Banach density at least $1/D_a > 0$.

1. Introduction. In an earlier article [2], the second and third authors along with Eyyunni investigated the existence of long arithmetic progressions among the set of values of the φ function over the natural numbers. In this article, we study the values of iterated Euler's totient φ function, defined by

$$\varphi^{(0)} = \mathrm{Id}_{\mathbb{N}} \quad \text{and} \quad \forall a \geq 1\colon\ \varphi^{(a)} = \varphi \circ \varphi^{(a-1)}$$

at natural numbers. As in the previous article, we study this question under the assumption of Dickson's conjecture [3], which is a predecessor of the Hardy–Littlewood prime k-tuples conjecture and also of Schinzel's Hypothesis H. Let us recall its statement.

Conjecture 1 (Dickson's conjecture). *Let s be a positive integer and $F_1, \ldots, F_s$ be linear polynomials with integral coefficients and positive leading coefficients such that their product has no fixed prime divisor. Then there exist infinitely many natural numbers n such that $F_1(n), \ldots, F_s(n)$ are all primes.*

2020 *Mathematics Subject Classification*: Primary 11B83; Secondary 11B05, 11N32, 11N64.
Key words and phrases: iterated values of Euler's totient φ function, Dickson's conjecture, long arithmetic progression, Banach density.
Received 1 June 2023; revised 14 June 2023.
Published online 24 January 2024.

Remark 1.1. The only case where Dickson's conjecture is known to be true is for $s = 1$, thanks to Dirichlet.

Our main result is the following.

Theorem 1. *Suppose that Dickson's conjecture is true and let $a \geq 2$ be a positive integer. There exists a positive integer D_a such that for any positive integer H there exist positive integers $M, m_1, \ldots, m_H$ such that for all h in $[1, H]$,*

$$\varphi^{(a)}(m_h) = D_a h + M. \tag{1}$$

Moreover, we can take

$$D_a = 2^{2a} P_a Q_a, \tag{2}$$

where Q_a is the product of distinct primes dividing $2^i - 1$ for $1 \leq i \leq a$, and P_a is the product of primes between 5 and $2a+1$ which are coprime to Q_a.

Remark 1.2. Theorem 1 implies that the set $\varphi^{(a)}(\mathbb{N})$ has a positive upper Banach density, provided that Dickson's conjecture holds true.

Remark 1.3. It would be interesting to prove unconditionally that the set $\varphi^{(a)}(\mathbb{N})$ contains an arbitrarily long arithmetic progression with some fixed common difference, or even that $\varphi^{(a)}(\mathbb{N})$ has a positive upper Banach density, even for $a = 1$.

2. Some intermediate results. From now on, the letters p and q, with or without index or subscript, will denote prime numbers, a an integer larger than 1, and H an integer larger than 2^a.

In this section, we will prove a few lemmas leading to the proof of the theorem.

We start by defining hyper Sophie Germain primes and fixed prime divisors of a polynomial as they will play an important role in our proof.

Definition 1. Let v be a positive integer. A prime number p is said to be a *v-hyper Sophie Germain prime* if all the numbers

$$\frac{p}{2} - \frac{1}{2}, p, 2p + (2-1), \ldots, 2^{v-1}p + (2^{v-1} - 1)$$

are prime numbers.

Remark 2.1. With the standard definition, we can say that all the numbers $p/2 - 1/2, p, 2p + (2-1), \ldots, 2^{v-2}p + (2^{v-2} - 1)$ are Sophie Germain primes ([1]); the sequence $p/2 - 1/2, p, 2p + (2-1), \ldots, 2^{v-1}p + (2^{v-1} - 1)$ is called a *Cunningham chain of first type with length v*, after [1].

([1]) Sophie Germain investigated those primes p such that $2p + 1$ is prime in the early 19th century in her study of Fermat's problem.

Definition 2. Let $F(t) \in \mathbb{Z}[t]$ be a polynomial with integer coefficients. A prime number p is called a *fixed prime divisor* of F if p divides $F(t)$ for all integers t.

2.1. Existence of infinitely many a-hyper Sophie Germain primes under Dickson's conjecture

Lemma 2. *Suppose that Dickson's conjecture is true. Let $a \geq 2$ and c be positive integers and b be an integer such that $b, 2b+(2-1), \dots, 2^a b+(2^a-1)$ are coprime to c. The arithmetic progression with difference c and first term b contains infinitely many a-hyper Sophie Germain primes.*

Proof. Consider the polynomial G defined by

$$G(t) = (ct+b)(2ct+2b+2-1)\cdots(2^a ct+2^a b+2^a-1). \tag{3}$$

We claim that G has no fixed prime divisor.

If a prime number p divides c, it cannot be a fixed divisor of G as otherwise $G(0) \equiv 0 \pmod{p}$, i.e.

$$b(2b+2-1)\cdots(2^a b+2^a-1) \equiv 0 \pmod{p}$$

implies that $\gcd(c, 2^i b+2^i-1) > 1$ for some $1 \leq i \leq a-1$, a contradiction to the hypothesis.

If a prime number p does not divide c, then we choose an integer t_0 such that

$$ct_0+b+1 \equiv 0 \pmod{p}$$

and hence, for this choice of t_0, we have

$$G(t_0) \equiv (-1)^{a+1} \not\equiv 0 \pmod{p}.$$

Thus G has no fixed prime divisor. Hence, by Dickson's conjecture, there exist infinitely many n such that

$$cn+b \text{ and } 2^i(cn+b)+(2^i-1)\ \forall 1 \leq i \leq a$$

are prime numbers. ∎

Lemma 3. *Suppose that Dickson's conjecture is true. Let $a \geq 2$ and c_1 be positive integers and let d be an integer such that $d, 2d+2-1, \dots, 2^a d+2^a-1$ are coprime to c_1; also let $\ell_1, \dots, \ell_g$ be distinct prime numbers which are coprime to c_1. Choose b such that*

$$b \equiv \begin{cases} d \pmod{c_1}, \\ -1 \pmod{\ell_i} \text{ for all integers } i \in [1, g]. \end{cases} \tag{4}$$

The arithmetic progression with difference $c = c_1\ell_1\cdots\ell_g$ and first term b contains infinitely many a-hyper Sophie Germain primes.

Proof. Consider again the polynomial G defined by (3). From the choice of b and the given assumptions, we see that $b, 2b+2-1, \dots, 2^a b+2^a-1$ are coprime to c_1 as well as to $\ell_1, \dots, \ell_g$. Hence,

$$b, 2b+2-1, \dots, 2^a b + 2^a - 1$$

are coprime to c. Thus Lemma 3 follows from Lemma 2. ■

2.2. Construction of a suitable family of primes. From now on, we assume that $a \geq 2$ and that H is an integer larger than $2^a - 1$. Before proceeding further, let us fix some notation. Set

$$T_a = \{p : \exists i \in [1, a-1] : p \mid 2^i - 1\}, \quad Q_a = \prod_{p \in T_a} p, \quad P_a = \prod_{\substack{p \leq 2a+1 \\ \gcd(p, Q_a)=1}} p. \tag{5}$$

The proof of Theorem 1 will make use of a family of $(a+1)$-hyper Sophie Germain primes satisfying some congruences and size properties, the existence of which is asserted by the following proposition.

Proposition 4. *Let*

$$\Pi_H = \prod_{\substack{2a+3 \leq p \leq aH \\ \gcd(p, Q_a)=1}} p.$$

Let u be a positive integer such that

$$u \equiv \begin{cases} 0 \pmod{2^{2a} P_a}, \\ 2^{a+1} \pmod{p} \ \text{if } p \mid Q_a. \end{cases} \tag{6}$$

Assuming Dickson's conjecture, one can find H many $(a+1)$-hyper Sophie Germain primes $p_1, \dots, p_H$ such that

$$\begin{aligned} & p_1 > 2^{2a} P_a Q_a \Pi_H + 1, \\ & \forall 1 \leq h \leq H-1 : \ p_{h+1} > 2^{2a} P_a Q_a \frac{p_h - 1}{2} p_h \prod_{i=1}^{a} (2^i p_h + 2^i + 1). \end{aligned} \tag{7}$$

Further, for all $h, k \in [1, H]$ with $h \neq k$ and for all p in $[2a+3, aH]$ with $\gcd(p, Q_a) = 1$, these primes satisfy the following relations, for $1 \leq i \leq a$:

$$\gcd\left(2^{2a+i-1} P_a Q_a h + 2^{i-1} u + (2^i - 1)\frac{p_h - 1}{2^{-a+1}}, p\right) = 1, \tag{8}$$

$$\gcd\left(\frac{2^{2a} P_a Q_a (k-h)}{2^{-i+1}} + (2^i - 1)\frac{p_k - 1}{2^{-a+1}}, \frac{p_h - 1}{2}\right) = 1, \tag{9}$$

$$\forall p \mid Q_a : p_h \equiv -1 \pmod{p}. \tag{10}$$

Proof. Set $R_a = 2^{2a} P_a Q_a$. We prove by induction that for any h between 1 and H, we can find $(a+1)$-hyper Sophie Germain primes $p_1, \dots, p_h$ such that

$$\text{(11)} \qquad \begin{aligned} & p_1 > R_a \Pi_H + 1, \\ & \forall \ell \in [1, h-1]\colon\ p_{\ell+1} > R_a \frac{p_\ell - 1}{2} p_\ell \prod_{i=1}^{a} (2^i p_\ell + 2^i + 1). \end{aligned}$$

Further, for all ℓ in $[1, h]$ and all primes p in $[2a+3, aH]$ with $\gcd(p, Q_a) = 1$, we have for $1 \le i \le a$:

$$\text{(12)} \qquad \gcd\left(\frac{R_a \ell + u}{2^{-i+1}} + (2^i - 1)\frac{p_\ell - 1}{2^{-a+1}}, p\right) = 1,$$

$$\text{(13)} \qquad \forall k < \ell \le h:\ \gcd\left(\frac{R_a}{2^{-i+1}}(\ell - k) + (2^i - 1)\frac{p_\ell - 1}{2^{-a+1}}, \frac{p_k - 1}{2}\right) = 1,$$

$$\text{(14)} \qquad \forall k < \ell \le h:\ \gcd\left(\frac{R_a}{2^{-i+1}}(k - \ell) + (2^i - 1)\frac{p_k - 1}{2^{-a+1}}, \frac{p_\ell - 1}{2}\right) = 1.$$

The construction of p_1 proceeds as follows.

For $h = 1$, conditions (13) and (14) are empty. Condition (11) will be satisfied as soon as we know that there are infinitely many $(a+1)$-hyper Germain primes satisfying (12).

For p in $[2a+3, aH]$, we can always find a residue class $r_1(p)$ modulo p such that none of the classes

$$\frac{r_1(p) - 1}{2}, r_1(p), \ldots, 2^a r_1(p) + (2^a - 1), 2^{i-1}(R_a + u) + 2^{a-2}(2^i - 1)(r_1(p) - 1)$$

for $1 \le i \le a$ is equivalent to 0 modulo p. This is possible as we have to avoid at most $2a + 2$ residue classes modulo p.

For $p \mid Q_a$ we can choose $r_1(p) \equiv -1 \pmod{p}$.

Having found suitable residue classes $r_1(p)$ for any prime p in $[2a+3, aH]$, the Chinese remainder theorem permits us to find a positive integer $s(1)$ such that, for each prime p in $[2a+3, aH]$ with $\gcd(p, Q_a) = 1$, none of the numbers

$$\begin{aligned} &\frac{s(1) - 1}{2}, s(1), \ldots, 2^a s(1) + (2^a - 1), (R_a + u) + 2^{a-1}(s(1) - 1), \ldots, \\ &\qquad\qquad 2^{a-1}(R_a + u) + 2^{a-1}(2^a - 1)(s(1) - 1) \end{aligned}$$

is congruent to 0 modulo p. Further, for any p dividing Q_a, we have $s(1) \equiv -1 \pmod{p}$, and hence all the numbers

$$\frac{s(1) - 1}{2}, s(1), \ldots, 2^a s(1) + (2^a - 1)$$

are congruent to -1 modulo p. Thus, by Lemma 3 the arithmetic progression with difference $Q_a \Pi_H$ and first term $(s(1) - 1)/2$ contains infinitely many $(a+1)$-hyper Sophie Germain primes satisfying (12), and thus we can find such a prime satisfying also (11).

We now apply induction to complete the proof of Proposition 4.

Assume that for some h between 1 and $H-1$, we have constructed a family of h many $(a+1)$-hyper Sophie Germain primes satisfying (10) and (12)–(14). Now we would like to construct p_{h+1}. It is enough to show that there exist infinitely many $(a+1)$-hyper Sophie Germain primes p_ℓ satisfying (10) and (12)–(14), where ℓ and h are replaced by $h+1$. Our new relation (14) is trivially satisfied as soon as p_{h+1} is large enough. For each $\ell < h+1$, one can choose an integer $r_{h+1}(\ell)$ such that for all primes p in $[2a+3, aH]$ with $\gcd(p, Q_a) = 1$ we have

$$\gcd\bigl(2^{i-1}(R_a(h+1)+u) + 2^{a-1}(2^i-1)(r_{h+1}(\ell)-1), p\bigr) = 1 \quad \text{for } 1 \le i \le a.$$

Further, $r_{h+1}(\ell)$ satisfies the relation

$$\gcd\left(\frac{R_a(h+1-\ell)}{2^{-i+1}} + (2^i-1)\frac{r_{h+1}(\ell)-1}{2^{-a+1}}, \frac{p_\ell - 1}{2}\right) = 1 \quad \text{for } 1 \le i \le a.$$

It is possible to find such $r_{h+1}(\ell)$ as we need to avoid at most $2a+2$ residue classes modulo $(p_\ell - 1)/2$. Arguing as we did previously, we can find a positive integer $s(h+1)$ such that all the numbers

$$\frac{s(h+1)-1}{2}, s(h+1), \ldots, 2^a s(h+1) + (2^a - 1),$$
$$2^{i-1}(R_a + u) + 2^{a-2}(2^i-1)(s(h+1)-1)$$

for $1 \le i \le a$ are coprime to Π_H and $(s(h+1)-1)/2$ satisfies (10). By the Chinese remainder theorem and Dickson's conjecture, there exist infinitely many $(a+1)$-hyper Sophie Germain primes which satisfy (10), (12) and (13), and we can choose one of them which is sufficiently large to also satisfy (11) and (14); we call such a prime p_{h+1}. This completes the induction. ∎

3. Proof of Theorem 1. We notice that, without loss of generality, it is enough to prove Theorem 1 with $H \ge 2^a$, which we assume from now on, thus being in a position to apply Proposition 4.

3.1. Construction of an auxiliary polynomial F. We consider the set $\{p_1, \ldots, p_H\}$ introduced in Proposition 4, and for h in $[1, H]$ we define the integer n_h by

$$n_h = (p_h - 1)2^{a-1}. \tag{15}$$

We notice that, thanks to (7), the numbers $n_h/2^a$ as h varies from 1 to H are pairwise coprime. We recall Definition 5 and further let

$$A = 2^{2a} Q_a \Pi_H \prod_{h=1}^{H} n_h^2.$$

We select a positive integer u satisfying (6) and a positive integer B satisfying

$$B \equiv \begin{cases} 0 \pmod{2^{2a} Q_a \Pi_H}, \\ -(u + 2^{2a} Q_a h) \pmod{(n_h/2^a)^2} \text{ for all integers } h \text{ in } [1, H]. \end{cases} \tag{16}$$

For h in $[1, H]$ and i in $[1, a]$, we define the polynomials $F_{h,i}$ by

$$F_{h,i}(t) = \frac{At + B + u + 2^{2a} P_a Q_a h}{2^{-i+1} n_h} + (2^i - 1) \tag{17}$$

and we let

$$F = \prod_{h=1}^{H} \prod_{i=1}^{a} F_{h,i}.$$

Note that each $F_{h,i}$ is a linear polynomial with integer coefficients and positive leading coefficient.

PROPOSITION 5. *The polynomial F has no fixed prime divisor.*

Proof. If p does not divide A, the congruence $F(t) \equiv 0 \pmod p$ has at most aH solutions in $\mathbb{Z}/p\mathbb{Z}$. Now if p is larger than aH, then p is not a fixed divisor of F.

If p divides A, then either p is in $[2, aH]$ or p divides Q_a or $p = (p_h - 1)/2$ for some $1 \le h \le H$. In this case, $F(t) \equiv 0 \pmod p$ is equivalent to

$$\prod_{h=1}^{H} \prod_{i=1}^{a-1} \left(\frac{B + u + 2^{2a} P_a Q_a h}{2^{-i+1} n_h} + (2^i - 1) \right) \equiv 0 \pmod p. \tag{18}$$

Note that for any h in $[1, H]$, we have

$$B + u + 2^{2a} P_a Q_a h \equiv 0 \pmod{n_h^2}.$$

Hence, 2 is not a fixed divisor of F. In addition, if we apply (7), then we can also conclude that $(p_h - 1)/2$ does not divide

$$\prod_{i=1}^{a-1} \left(\frac{B + u + 2^{2a} P_a Q_a h}{2^{-i+1} n_h} + (2^i - 1) \right).$$

If $p = (p_h - 1)/2$ divides

$$\prod_{i=1}^{a-1} \left(\frac{B + u + 2^{2a} P_a Q_a k}{2^{-i+1} n_k} + (2^i - 1) \right)$$

for some $k \neq h$, then

$$2^{2a+i-1} P_a Q_a (k - h) + (2^i - 1) n_k \equiv 0 \pmod p$$

for some $1 \le i \le a$, a contradiction to (9). If any prime p in $[2a + 3, aH]$ and coprime to Q_a satisfies (18), then

$$2^{2a+i-1} P_a Q_a h + 2^{i-1} u + n_h (2^i - 1) \equiv 0 \pmod p$$

for some $1 \le i \le a$ and $1 \le h \le H$, a contradiction to (8). Hence, the only possible fixed prime divisors of F satisfy $p \in [3, 2a+1]$ or $p \mid Q_a$. Now if $p \mid Q_a$, then (18) implies that

$$\prod_{h=1}^{H}\prod_{i=1}^{p-2}(2^{i-1}u + (2^i - 1)n_h) \equiv 0 \pmod{p}.$$

Hence, for some $1 \le i_0 \le p-2$ and $1 \le h \le H$, we have

$$2^{i_0-a-1}u + (2^{i_0} - 1)\frac{n_h}{2^a} \equiv 0 \pmod{p}.$$

Since $u \equiv 2^{a+1} \pmod{p}$ and, by construction, $n_h/2^a \equiv -1 \pmod{p}$, we have

$$2^{i_0-a-1}u + (2^{i_0} - 1)\frac{n_h}{2^a} \equiv 1 \not\equiv 0 \pmod{p}.$$

Since $a \ge 2$ and $2^2 - 1 = 3$, we see that 3 divides Q_a. Hence, we need to consider only primes in $[5, 2a+1]$ and coprime to Q_a. Finally, if p is in $[5, 2a+1]$ and coprime to Q_a, then

$$2^{p+i-1}u + (2^{p+i} - 1)n_h \equiv 2^i u + (2^{i+1} - 1)n_h \pmod{p}$$

for $0 \le i \le a-p$, and hence (18) can be written as

$$\prod_{h=1}^{H}\prod_{i=1}^{p-2}\left(2^{i-1}u + (2^i - 1)n_h\right)^{d_i} \equiv 0 \pmod{p} \tag{19}$$

for some positive integers $d_i \ge 1$. Recalling (6), we deduce from (19) that

$$(2^i - 1)\frac{n_h}{2^a} \equiv 0 \pmod{p}$$

for some $1 \le i \le p-2$ and $1 \le h \le H$. This is a contradiction as p is coprime to $n_h/2^a$ and Q_a. Thus the polynomial F has no fixed prime divisor. ∎

3.2. End of the proof of Theorem 1. Since the polynomials $F_{h,i}$ have positive leading coefficients and F has no fixed prime divisor, Dickson's conjecture implies that we can find a positive integer t_0 such that for each h in $[1, H]$ the value of $F_{h,i}$ for $1 \le i \le a$ at t_0 is a prime number strictly larger than p_H.

Let us write $q_h = F_{h,1}(t_0)$ and $M = At_0 + B + u$. Then for each $2 \le i \le a$, $F_{h,i}(t_0) = 2^i q_h + (2^i - 1)$ is a prime number and

$$q_h = \frac{M + 2^{2a}P_aQ_ah}{n_h} + 1. \tag{20}$$

For any $1 \le h \le H$, using (20), we can write

$$\begin{aligned}\varphi^{(a)}&((2^{a-1}p_h + 2^{a-1} - 1)(2^{a-1}q_h + 2^{a-1} - 1))\\ &= \varphi^{(a-1)}(2^2(2^{a-2}p_h + 2^{a-2} - 1)(2^{a-2}q_h + 2^{a-2} - 1))\\ &\vdots\\ &= \varphi(2^a p_h q_h)\\ &= 2^{2a-1}(p_h - 1)(q_h - 1).\end{aligned}$$

The last equality follows from the fact that, by construction, q_h is a prime larger than p_h and thus coprime to it. Using (15) and (20), we can write, for any h in $[1, H]$,

$$\varphi^{(a)}((2^{a-1}p_h + 2^{a-1} - 1)(2^{a-1}q_h + 2^{a-1} - 1)) = M + 2^{2a}P_aQ_ah.$$

This completes the proof of Theorem 1.

Acknowledgements. The authors are grateful to Suhita Hazra, Rashi Lunia, Sunil L Naik and Purusottam Rath for some stimulating discussions.

This work benefitted from the support of the SPARC project 445. The second author is grateful to the Institute of Mathematical Sciences, Chennai, and the Harish-Chandra Research Institute, Prayagraj, for providing excellent working conditions. All authors would also like to thank the Indo-French Program in Mathematics (IFPM). The third author would also like to thank DAE number theory plan project for partial financial support.

References

[1] A. Cunningham, *On hyper-even numbers and on Fermat's numbers*, Proc. London Math. Soc. (2) 5 (1907), 237–274.

[2] J.-M. Deshouillers, P. Eyyunni and S. Gun, *On the local structure of the set of values of Euler's φ function*, Acta Arith. 199 (2021), 103–109.

[3] L. E. Dickson, *A new extension of Dirichlet's theorem on prime numbers*, Messenger Math. 33 (1904), 155–161.

R. Balasubramanian, Sanoli Gun
The Institute of Mathematical Sciences
HBNI, C.I.T. Campus, Taramani
Chennai 600113, Tamil Nadu, India
E-mail: balu@imsc.res.in
sanoli@imsc.res.in

Jean-Marc Deshouillers
Institut de Mathématiques de Bordeaux
Université de Bordeaux, CNRS, Bordeaux INP
33400 Talence, France
E-mail: jean-marc.deshouillers@math.u-bordeaux.fr