# On the solutions to $\phi(n) = \phi(n+k)$

*S.W. Graham*,* *Jeffrey J. Holt and Carl Pomerance*†

*For Andrzej Schinzel on his sixtieth birthday*

**Abstract.** We study the number and nature of solutions of the equation $\phi(n) = \phi(n+k)$, where $\phi$ denotes Euler's phi-function. We exhibit some families of solutions when $k$ is even, and we conjecture an asymptotic formula for the number of solutions in this case. We show that our conjecture follows from a quantitative form of the prime $k$-tuples conjecture. We also show that the prime $k$-tuples conjecture implies that there are arbitrarily long arithmetic progressions of equal $\phi$-values.

1991 Mathematics Subject Classification: Primary: 11N25; Secondary: 11Y99.

Keywords: Euler's phi-function, prime $k$-tuples conjecture, Hypothesis H.

## 1. Introduction

Our objective in this paper is to study the number and nature of the solutions to the equation

$$\phi(n) = \phi(n+k) \tag{1}$$

for a fixed value of $k$. Here $\phi(n)$ is Euler's $\phi$-function which counts the number of positive integers less than or equal to $n$ that are relatively prime to $n$. As we will be considering the number of solutions to (1), it is convenient to define the function

$$P(k;x) = \big|\{n \le x : \phi(n) = \phi(n+k)\}\big|. \tag{2}$$

In 1972, M. Lal and P. Gillard [11] used an IBM 1620, Model 1, to determine all solutions to (1) for each $1 \le k \le 30$ in the range $1 \le n \le 10^5$. They produced a table of values for $P(k;x)$ for each $k$ in the stated range and $x$ taken in increments of $10^4$. Other authors have extended the searches of Lal and Gillard in the case $k = 1$. The most extensive computations currently are due to R. Baillie [1], who found 306 solutions of $\phi(n) = \phi(n+1)$ up to $10^8$.

*Supported in part by an NSA grant.
†Supported in part by an NSF grant.

Using 18 Sun Sparc 5 workstations, we have extended Lal and Gillard's computations of $P(k;x)$. Our computations were performed in three stages. For the first stage, we used *Mathematica* software to compute all solutions to (1) for $1 \le k \le 30$ and $n \le 10^8$. In addition to counting the number of solutions to (1) for each $k$, we also saved the solutions for further analysis. The built-in *Mathematica* function `EulerPhi` was used to compute the $\phi$-values in this stage. In the case of $k = 1$, our computations agree with those of Baillie.

For the second stage, we extended our computations of solutions to (1) to $1 \le k \le 100$ and $n \le 10^{10}$. We used the C++ programming language to implement simple sieving and scanning procedures to compute the necessary $\phi$-values and then look for solutions to (1). The values of $P(k;x)$ computed in this stage were compared to those computed using *Mathematica* in the first stage to verify the consistency of the two programs. A summary of these computations is contained in Table 1.

The final computing stage was motivated by the scarcity of solutions to (1) found in the first two stages when $k \equiv 3 \bmod 6$. We used the same sieving procedure as in the second stage to compute $\phi$-values, but altered the scanning procedure to search only for solutions to (1) corresponding to $1 \le k \le 100$ and $k \equiv 3 \bmod 6$. We checked all values of $n$ satisfying $10^{10} \le n \le 10^{11}$, and found two solutions to (1) in this range. These solutions are discussed in more detail in the next section.

L. Moser [13] noted that if $p$ and $2p-1$ are both odd primes and $n = 2(2p-1)$, than $\phi(n) = \phi(n+2)$. More generally, A. Schinzel [17] observed that if $p$ and $2p-1$ are primes that do not divide the even number $k$, and

$$n = (2p-1)k, \tag{3}$$

then $n$ is a solution of (1). There is a conjecture due to Dickson [4] known as the prime $k$-tuples conjecture; a special case of this conjecture is that there are infinitely many primes $p$ with $2p-1$ prime. Therefore, Dickson's conjecture combined with Schinzel's observation implies that (1) has infinitely many solutions when $k$ is even. In this paper, we generalize Schinzel's observation to obtain more solutions to (1). By appealing to a quantitative form of Dickson's conjecture, we conditionally prove an asymptotic formula for the number of solutions to (1) when $k$ is even.

Our computations imply that solutions of (1) are very sparse when $k$ is odd, especially when $k \equiv 3 \bmod 6$. Despite this, we believe it is likely that there are infinitely many solutions to (1) for each $k$, and that our numerical evidence is simply not extensive enough to overwhelmingly suggest that this is the case. It is interesting to note that little is known unconditionally about the number of solutions to (1). In 1956, W. Sierpiński [21] showed that for each $k$ there is at least one value of $n$ such that (1) is satisfied. (The proof is easy: Let $p$ be the smallest prime not dividing $k$, and then set $n = (p-1)k$.) In 1958, Schinzel [17] showed that there are at least two solutions to (1) for all $k \le 8 \times 10^{47}$, and in the following year Schinzel and A. Wakulicz [19] extended this result to all $k \le 2 \times 10^{58}$.

We also consider arithmetic progressions of equal $\phi$ values. For example, we show that Dickson's conjecture implies that for any $q$, there is some $k$ and infinitely many $n$ such that

$$\phi(n) = \phi(n+k) = \dots = \phi(n+qk).$$

In addition, we have also determined the number of such progressions for $1 \le k \le 100$ and $n \le 10^{10}$. A summary of these results is given in Section 5.

## 2. Discussion of Table 1

Table 1 gives values of $P(k;x)$ for $1 \le k \le 100$ and $x = 10^8, 10^9$, and $10^{10}$. It is immediately evident from the data in Table 1 that the solutions to (1) are much more common for $k$ even than for $k$ odd. When $k$ is odd, the case $k \equiv 3 \bmod 6$ is particularly striking. Up to $n = 245$ there are a few solutions for various values of $k \equiv 3 \bmod 6$, but then they appear to die out. As we mentioned in the introduction, we extended our search up to $n \le 10^{11}$ for these $k$. We found only three such values of $n$ with $245 < n \le 10^{11}$. The three values are given in Table 2; they correspond to $k = 27, 81$, and 81 respectively.

The following observations may help to explain why solutions are so rare when $k \equiv 3 \bmod 6$. If $\phi(n) = \phi(n')$, where $n$ and $n'$ are both large and close together, then $\phi(n)/n$ is approximately equal to $\phi(n')/n'$. The fraction $\phi(m)/m$ depends solely on the prime factors of $m$, and is principally determined by the small prime factors of $m$. When $n' = n + k$, where $k \equiv 3 \bmod 6$, then one of $n$ and $n'$ is divisible by 2 (say $n$ for the sake of this discussion), while the other is not, and either both are divisible by 3 or neither are. Thus the smallest prime 2 pushes $\phi(n)/n$ and $\phi(n')/n'$ apart, and the next smallest prime 3 can not help narrow the difference. This requires $n'$ (which is odd) to be divisible by a large number of small primes greater than 3, and $n$ (which is even) to be free of small prime factors greater than 3. Moreover, since $n'$ is divisible by a large number of primes, $\phi(n')$ must be divisible by a large power of 2, which in turn forces $n$ to be divisible by either a large power of 2 or odd primes $p$ such that $p-1$ is divisible by a large power of 2, or some combination of both. (These properties are evident in our three large solutions above.) All of these constraints seem to conspire to push additional solutions to (1) for $k \equiv 3 \bmod 6$ to very high levels. Note that the three large solutions all correspond to values of $k$ that are powers of 3. In the discussion above, this form of $k$ presents the least difficulty, so perhaps it is not surprising that these solutions appear first. In general, the more small prime factors $k$ has, the more difficult it will be to overcome the constraints described above. In the case of odd $k \not\equiv 3 \bmod 6$, the effects of these constraints are not as pronounced. They are evident, though. For example, examine the number of solutions when $k$ is 5 mod 10.

By contrast, the situation for $k$ even is much clearer. As mentioned earlier, there are many more solutions to (1) in these cases, and most of these solutions can be explained in a simple manner. Upon careful study of our computed solutions to (1) for even values of $k$, a generalization of Schinzel's observation (3) emerges. Once the proper form of the generalization is found, the proof follows easily.

| $x \backslash k$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $10^8$ | 306 | 125986 | 2 | 69131 | 95 | 356157 | 336 | 38157 | 2 | 36579 |
| $10^9$ | 651 | 957214 | 2 | 517863 | 202 | 2679538 | 737 | 281341 | 2 | 267252 |
| $10^{10}$ | 1267 | 7558421 | 2 | 4047331 | 433 | 20969365 | 1608 | 2171817 | 2 | 2048937 |

| $x \backslash k$ | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|
| $10^8$ | 293 | 196539 | 306 | 66036 | 4 | 21532 | 341 | 139506 | 335 | 52606 |
| $10^9$ | 655 | 1455738 | 649 | 492095 | 4 | 154490 | 691 | 1020710 | 718 | 385863 |
| $10^{10}$ | 1367 | 11257702 | 1354 | 3818031 | 4 | 1171252 | 1460 | 7834367 | 1424 | 2958324 |

| $x \backslash k$ | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
|---|---|---|---|---|---|---|---|---|---|---|
| $10^8$ | 2 | 16771 | 378 | 152936 | 109 | 14521 | 3 | 36869 | 361 | 246694 |
| $10^9$ | 2 | 117755 | 784 | 1118648 | 233 | 101821 | 3 | 268658 | 743 | 1801600 |
| $10^{10}$ | 2 | 883419 | 1614 | 8580206 | 494 | 761625 | 4 | 2054653 | 1540 | 13781327 |

| $x \backslash k$ | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 |
|---|---|---|---|---|---|---|---|---|---|---|
| $10^8$ | 347 | 12416 | 4 | 11768 | 11 | 78597 | 321 | 10748 | 5 | 29761 |
| $10^9$ | 726 | 85540 | 4 | 81766 | 17 | 560491 | 715 | 73725 | 5 | 211968 |
| $10^{10}$ | 1507 | 634597 | 4 | 607615 | 29 | 4229528 | 1530 | 545468 | 5 | 1598241 |

| $x \backslash k$ | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 |
|---|---|---|---|---|---|---|---|---|---|---|
| $10^8$ | 336 | 142433 | 364 | 9900 | 6 | 9356 | 397 | 86062 | 371 | 10169 |
| $10^9$ | 733 | 1027337 | 777 | 65586 | 6 | 62695 | 813 | 613668 | 798 | 67937 |
| $10^{10}$ | 1513 | 7797557 | 1598 | 480878 | 6 | 461469 | 1673 | 4631577 | 1746 | 495624 |

| $x \backslash k$ | 51 | 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 60 |
|---|---|---|---|---|---|---|---|---|---|---|
| $10^8$ | 7 | 8634 | 367 | 56315 | 70 | 21054 | 4 | 7808 | 373 | 155252 |
| $10^9$ | 7 | 57053 | 775 | 395121 | 130 | 147809 | 4 | 51634 | 757 | 1103503 |
| $10^{10}$ | 7 | 415358 | 1630 | 2953076 | 278 | 1111035 | 4 | 376503 | 1577 | 8317199 |

| $x \backslash k$ | 61 | 62 | 63 | 64 | 65 | 66 | 67 | 68 | 69 | 70 |
|---|---|---|---|---|---|---|---|---|---|---|
| $10^8$ | 344 | 17177 | 4 | 7462 | 60 | 71373 | 348 | 7063 | 6 | 26705 |
| $10^9$ | 705 | 121821 | 4 | 47955 | 143 | 502432 | 714 | 45861 | 6 | 184487 |
| $10^{10}$ | 1486 | 921142 | 4 | 346381 | 320 | 3767424 | 1457 | 331696 | 6 | 1372323 |

| $x \backslash k$ | 71 | 72 | 73 | 74 | 75 | 76 | 77 | 78 | 79 | 80 |
|---|---|---|---|---|---|---|---|---|---|---|
| $10^8$ | 356 | 62338 | 315 | 6443 | 6 | 6517 | 371 | 65363 | 362 | 17402 |
| $10^9$ | 741 | 434134 | 665 | 42124 | 6 | 41606 | 831 | 460092 | 738 | 117874 |
| $10^{10}$ | 1501 | 3237951 | 1338 | 304249 | 6 | 298381 | 1787 | 3448231 | 1500 | 868497 |

| $x \backslash k$ | 81 | 82 | 83 | 84 | 85 | 86 | 87 | 88 | 89 | 90 |
|---|---|---|---|---|---|---|---|---|---|---|
| $10^8$ | 4 | 6016 | 395 | 85835 | 89 | 5869 | 5 | 6176 | 366 | 119562 |
| $10^9$ | 4 | 38611 | 798 | 598785 | 193 | 37285 | 5 | 37612 | 758 | 838515 |
| $10^{10}$ | 4 | 278338 | 1614 | 4463837 | 463 | 267235 | 5 | 264553 | 1504 | 6275296 |

| $x \backslash k$ | 91 | 92 | 93 | 94 | 95 | 96 | 97 | 98 | 99 | 100 |
|---|---|---|---|---|---|---|---|---|---|---|
| $10^8$ | 340 | 5744 | 5 | 5536 | 84 | 49553 | 374 | 13392 | 6 | 14485 |
| $10^9$ | 738 | 35682 | 5 | 34725 | 162 | 339930 | 756 | 91555 | 6 | 97414 |
| $10^{10}$ | 1659 | 253155 | 5 | 247588 | 363 | 2513443 | 1575 | 677947 | 6 | 713868 |

Table 1: Summary of values for $P(k;x)$

$$\begin{aligned}\phi(4135966808) &= \phi(4135966835) = 2052864000 = 2^{11}3^6 5^3 11^1,\\ 4135966808 &= 2^3 241^1 331^1 6481^1,\\ 4135966835 &= 5^1 7^1 11^1 13^1 19^1 23^1 31^1 61^1;\end{aligned}$$

$$\begin{aligned}\phi(12407900424) &= \phi(12407900505) = 4105728000 = 2^{12}3^6 5^3 11,\\ 12407900424 &= 2^3 3^1 241^1 331^1 6481^1,\\ 12407900505 &= 3^1 5^1 7^1 11^1 13^1 19^1 23^1 31^1 61^1;\end{aligned}$$

$$\begin{aligned}\phi(15720219515) &= \phi(15720219596) = 7834337280 = 2^{15}3^3 5^1 7^1 11^1 23^1,\\ 15720219515 &= 5^1 7^1 11^1 13^1 17^1 23^1 29^1 277^1,\\ 15720219596 &= 2^2 577^1 691^1 9857^1.\end{aligned}$$

Table 2: Exceptional solutions to $\phi(n) = \phi(n+k)$.

**Theorem 1.** *Suppose that $j$ and $j+k$ have the same prime factors (so that $k$ is even), and let $g = (j, j+k)$. Suppose that for a positive integer $r$,*

$$\frac{j}{g}r + 1 \quad \textit{and} \quad \frac{j+k}{g}r + 1 \tag{4}$$

*are both primes that do not divide $j$. If*

$$n = j\Big(\frac{j+k}{g}r + 1\Big), \tag{5}$$

*then $\phi(n) = \phi(n+k)$.*

*Proof.* We have

$$\phi(n) = \phi(j)\frac{j+k}{g}r = \phi(j)(j+k)\frac{r}{g}$$

and

$$\phi(n+k) = \phi\Big((j+k)\Big(\frac{j}{g}r + 1\Big)\Big) = j\phi(j+k)\frac{r}{g}.$$

As $j$ and $j+k$ have the same prime factors, it follows (see [10], Theorem 62) that

$$\phi(j)(j+k) = j\phi(j+k),$$

which completes the proof. □

In order to determine how many solutions $n$ of (1) are of the form (5) in Theorem 1, we first need to find values of $j$ for which $j$ and $j+k$ have the same prime factors. For $k$ even, $2 \le k \le 30$, Table 3 contains all values of $j$ which satisfy the hypotheses of Theorem 1. Further analysis of Table 3 is given in the next section.

| $k$ | $j$ |
|---|---|
| 2 | 2 |
| 4 | 4 |
| 6 | 2, 3, 6, 12, 18, 48 |
| 8 | 8 |
| 10 | 10, 40 |
| 12 | 4, 6, 12, 24, 36, 96 |
| 14 | 2, 14, 98 |
| 16 | 16 |
| 18 | 6, 9, 18, 36, 54, 144 |
| 20 | 5, 20, 80 |
| 22 | 22 |
| 24 | 3, 8, 12, 24, 48, 72, 192 |
| 26 | 26 |
| 28 | 4, 28, 196 |
| 30 | 2, 6, 10, 15, 18, 20, 24, 30, 45, 50, 60, 90, 120, 150, 162, 240, 270, 375, 450, 720, 1250, 2400 |

Table 3: Values of $j$ for which $j$ and $j+k$ have the same prime factors.

## 3. When do $j$ and $j+k$ have the same prime factors?

Let $P$ be a finite set of primes, and let $1 = n_1 < n_2 < \ldots$ be the integers composed of primes in $P$. A result of A. Thue [22] (see also G. Pólya [14] and R. Tijdeman [23]) states that $\lim_{i\to\infty}(n_{i+1} - n_i) = \infty$. Therefore, for a given $k$, there are only finitely many values of $j$ such that $j$ and $j+k$ have the same prime factors. Tijdeman's result is effective, so that in principle one could determine all the desired pairs via an exhaustive search. Unfortunately, the bounds are so large that such a search is not feasible. However, it turns out that for each even $k \le 30$, elementary techniques suffice to completely determine all possible values of $j$.

To illustrate the techniques used, we focus on the case $k = 6$. As $k = 2 \cdot 3$, it follows that there are three possibilities for the prime factors of $j$ and $j+k$: $j = 2^a$, $j + k = 2^c$; $j = 3^b$, $j + k = 3^d$; $j = 2^a 3^b$, $j + k = 2^c 3^d$; in each case, $a, b, c, d \ge 1$. In the first case, we have

$$2^a + 6 = 2^c. \tag{6}$$

Setting $a = 1$ and $c = 3$ yields a solution to (6) which corresponds to $j = 2$. Reducing mod 4 shows that there are no other solutions to (6). The case $j = 3^b$, $j + k = 3^d$ is similar and yields $j = 3$.

For the last case we obtain the equation $2^a 3^b + 6 = 2^c 3^d$. Dividing through by 6, we have

$$2^{a-1}3^{b-1} + 1 = 2^{c-1}3^{d-1}. \tag{7}$$

Reducing mod 2 implies that either $a = 1$ or $c = 1$. If $c = 1$, then (7) reduces to

$$2^{a-1}3^{b-1} + 1 = 3^{d-1}. \tag{8}$$

There are no solutions with $d = 1$, so we must have $d \geq 2$. Reducing mod 3 tells us that $b = 1$. Upon writing $x = a - 1$ and $y = d - 1$, we see that (8) simplifies to

$$2^x + 1 = 3^y. \tag{9}$$

Now (9) has solutions with $(x, y) = (1,1)$ and $(3,2)$. These solutions correspond to $j = 12$ and $j = 48$ respectively. Any other solution must have $x \geq 4$, which we now assume. Reducing mod 16 gives $y \equiv 0 \bmod 4$; this in turn gives $3^y \equiv 1 \bmod 5$. However, we then obtain the congruence $2^x \equiv 0 \bmod 5$, which clearly has no solutions.

The case $a = 1$ can be done in a similar fashion; it yields a solution corresponding to $j = 18$ and no others. Thus the values of $j$ given in Table 3 form a complete list of the values of $j$ such that $j$ and $j + 6$ have the same prime factors.

We have shown that the other entries in Table 3 are complete. The proofs use only elementary congruence arguments similar to the one above; however, there are many tedious cases, and we shall not give the details here. Some of the cases can be simplified by appealing to results of W.J. LeVeque [12], J.W.S. Cassels [3], and R. Scott [20].

## 4. Asymptotics

From Theorem 1, we see that for a fixed even number $k$, to prove that $\phi(n) = \phi(n+k)$ for infinitely many $n$, it suffices to show that there are infinitely many integers $r$ such that $r + 1$ and $2r + 1$ are both prime. Unfortunately, this is a very difficult open problem. It is, however, a special case of Dickson's prime $k$-tuples conjecture [4] and Schinzel's Hypothesis H [18]. Instead of stating these well-known conjectures in all their generality, we give only the following simple special case, which is all we shall need.

**Conjecture 1.** *Let $a_1, a_2, \ldots, a_g$ be distinct positive integers. Then there are infinitely many integers $r$ such that $a_1 r + 1, a_2 r + 1, \ldots, a_g r + 1$ are all prime.*

G.H. Hardy and J.E. Littlewood [9] formulated a more quantitative form of at least part of the prime $k$-tuples conjecture, and P.T. Bateman and R.A. Horn [2] generalized this by formulating Hypothesis H*, which is a quantitative form of Hypothesis H. Again we only give the following special case.

Let $C_2$ be the "twin-prime constant" given by

$$C_2 = \prod_{p>2} \left(1 - (p-1)^{-2}\right) \doteq 0.660161815847\,.$$

**Conjecture 2.** *Suppose that $a$ and $b$ are relatively prime natural numbers with $b < a$. Then, as $x \to \infty$,*

$$\sum_{\substack{r \leq x \\ ar+1 \text{ prime} \\ br+1 \text{ prime}}} 1 \sim 2C_2 \prod_{\substack{p \mid ab(a-b) \\ p>2}} \left(\frac{p-1}{p-2}\right) \int_2^x \frac{1}{\log(at)\log(bt)}\, dt. \tag{10}$$

We note that the integral on the right side of (10) can be replaced by $x/\log^2 x$ without changing the asymptotics. Using *Mathematica* and the information in Table 3, we determined the number of solutions to (1) for even $k \le 30$ and $n \le 10^8$ that are of the form (5) given in Theorem 1. Table 4 provides a summary of the number of such solutions as well as the proportion of the total number of solutions that these special solutions represent. Based on the evidence, we conjecture that these special solutions have density 1 among the set of all solutions to (1). To support this conjecture, we prove that it follows from Conjecture 2. But first we give the following unconditional result.

**Theorem 2.** *Let $P_1(k;x)$ be the number of solutions $n \le x$ to $\phi(n) = \phi(n+k)$ that are not in the form* (5) *given in Theorem* 1. *In particular, when $k$ is odd, $P_1(k;x) = P(k;x)$. Then for every $k$, there is some $x_0(k)$ such that if $x \ge x_0(k)$, then $P_1(k;x) < x/\exp(\log^{1/3} x)$.*

*Proof.* P. Erdős, C. Pomerance, and A. Sárközy [6] proved Theorem 2 in the case $k = 1$. We claim that a small modification of their proof yields the general case. We indicate the changes in their argument needed to do this, and we refer the reader to [6] and [16] for the rest of the details.

As in Theorem 2 of [6], let $l = \exp(\log^{1/3} x)$, $L = \exp\bigl(\frac{1}{8}(\log x)^{1/3}\log\log x\bigr)$, and let $P(n)$ denote the largest prime factor of $n$. We assume that

(i) $P(n) \ge L^2$ and $P(n+k) \ge L^2$,

(ii) if $r^a$ divides $n$ or $n+k$ and $a \ge 2$, then $r^a \le l^3$.

As shown in [6], the number of $n \le x$ not satisfying (i) and (ii) is $o(x/l)$. From these conditions, we see that there are primes $p, p' \ge L^2$ and integers $m, m'$ such that $n = mp, n+k = m'p'$ and $(m,p) = (m',p') = 1$. From this and the assumption that $\phi(n) = \phi(n+k)$, we see that

$$\begin{aligned} p'\bigl(\phi(m)m' - \phi(m')m\bigr) &= \phi(m)(mp+k) - p'\phi(m')m \\ &= m\phi(n) + m\phi(m) + k\phi(m) - \phi(n+k)m - \phi(m')m \\ &= m\phi(m) - m\phi(m') + k\phi(m). \end{aligned}$$

Now we separate the $n$'s into two classes. Class A consists of those $n$ for which $\phi(m)/m \ne \phi(m')/m'$, and class B consists of those with $\phi(m)/m = \phi(m')/m'$. The same proof as in [6], Theorem 2, shows that class A contributes $o(x/l)$. To complete the proof, we will show that all of the $n$'s in class B are of the form given in Theorem 1.

Now assume that $n$ is in class B. From the equations $\phi(m)(p-1) = \phi(m')(p'-1)$ and $\phi(m)/m = \phi(m')/m'$, we have

$$p' - 1 = \frac{\phi(m)}{\phi(m')}(p-1) = \frac{m}{m'}(p-1).$$

Therefore

$$mp + k = m'p' = m'(p'-1) + m' = m(p-1) + m'. \tag{11}$$

We deduce that $k = m' - m$; in other words, $\phi(m)/m = \phi(m+k)/(m+k)$. Therefore, $m$ and $m+k$ have the same prime factors. Now let $g = (m, m+k)$.

| $k$ | A | B | C |
|---|---|---|---|
| 2 | 125986 | 125199 | 0.993753 |
| 4 | 69131 | 68295 | 0.987907 |
| 6 | 356157 | 353616 | 0.992866 |
| 8 | 38157 | 37229 | 0.975679 |
| 10 | 36579 | 35479 | 0.969928 |
| 12 | 196539 | 193729 | 0.985703 |
| 14 | 66036 | 65110 | 0.985977 |
| 16 | 21532 | 20505 | 0.952304 |
| 18 | 139506 | 136715 | 0.979994 |
| 20 | 52606 | 51305 | 0.975269 |
| 22 | 16771 | 15646 | 0.932920 |
| 24 | 152936 | 149779 | 0.979357 |
| 26 | 14521 | 13531 | 0.931823 |
| 28 | 36869 | 35820 | 0.971548 |
| 30 | 246694 | 242021 | 0.981058 |

Table 4: A $= P(k; 10^8)$.
B = Number of solutions of the form (5) in Theorem 1.
C = The ratio B/A.

Then

$$\frac{m+k}{g}(p'-1) = \frac{m}{g}(p-1) \quad \text{and} \quad \left(\frac{m+k}{g}, \frac{m}{g}\right) = 1. \tag{12}$$

We deduce that $m/g$ divides $p'-1$, so there is some $r$ such that

$$p' = \frac{m}{g}r + 1 \quad \text{and} \quad p = \frac{m+k}{g}r + 1. \tag{13}$$

All of this together shows that $n$ is of the form given in Theorem 1, and this completes the proof. □

**Corollary 1.** *If $k > 0$ is even, let $c(k) = \sum^* \frac{g}{j(j+k)} \prod^* \frac{p-1}{p-2}$, where $\sum^*$ runs over all $j$ such that $j$ and $j+k$ have the same prime factors, $\prod^*$ runs over all primes $p > 2$ such that $p | jk(j+k)/g^3$, and $g = (j, j+k)$. Then $0 < c(k) < \infty$ and if Conjecture 2 is true, then, as $x \to \infty$,*

$$P(k; x) \sim 2C_2 c(k) \frac{x}{\log^2 x}. \tag{14}$$

*Proof.* First, we note that if we fix an even number $k$, there is at least one number $j$ such that $j$ and $j+k$ have the same prime factors, namely $j = k$. Further, as noted at the beginning of Section 3, it follows from [22] that there are only finitely many such integers $j$. Thus, $0 < c(k) < \infty$. Now assume that Conjecture 2 is true. For each $j$ satisfying the hypotheses of Theorem 1, the formula (10), with $a = (j+k)/g$ and $b = j/g$, gives a conditional estimate for the number of pairs of the form (4) which are both primes. Summing over each $j$ yields the expression

$$2C_2 \sum^{*} \prod^{*} \left(\frac{p-1}{p-2}\right) \int_2^{\frac{g}{j(j+k)}x} \frac{1}{\log\left(\frac{j}{g}t\right) \log\left(\frac{j+k}{g}t\right)}\, dt, \tag{15}$$

which is asymptotically equal to the right-hand side of (14). Theorem 2 asserts that the additional solutions not in the form of Theorem 1 are negligible, so the result follows. □

We remark that for numerical purposes it can be advantageous to replace the right side of (14) with (15). In Table 5, we give the number of solutions predicted by the formula given in (15) up to $x = 10^{10}$ for each even $k \le 30$. We also give the ratio of the predicted number of solutions to the computed number of solutions.

We note that there are other families of solutions to (1) that have the same flavor as Theorem 1. For instance, suppose $b$ is even and $k = bc$. If

$$p,\ p+b,\ (c+1)p-c, \text{ and } (c+1)(p+b)-c$$

are all prime, and $n = (p+b)((c+1)p-c)$, then $\phi(n) = \phi(n+k)$. In this case, we can use sieve methods to show that the number of such solutions for $n \le x$ is $O(x^{1/2}\log^{-4} x)$. Thus, solutions of this type will contribute to the growth of $P(k;x)$, but if we assume that Conjecture 2 holds, such solutions will not occur frequently enough to alter the formula in (14).

One might ask if a typical solution to the equation $\phi(a) = \phi(b)$ with $a < b$ is of the form given in Theorem 1 (with $n = a$ and $k = b - a$). Especially in light of the above results it is tempting to conjecture this is the case. To specify things, let $P_0(k;x)$ be the number of solutions $n \le x$ of (1) in the form considered in Theorem 1. Thus, $P(k;x) = P_0(k;x) + P_1(k;x)$, and Corollary 1 asserts that if Conjecture 2 is true, then $P(k;x) \sim P_0(k;x)$ as $x \to \infty$ for each fixed even number $k$. Is it true that

$$\sum_{k\le x} P(k;x) \sim \sum_{k \le x} P_0(k;x)?$$

Or perhaps

$$\sum_{k\le x,\ k \text{ even}} P(k;x) \sim \sum_{k \le x} P_0(k;x)?$$

In fact the answer to both questions is a resounding "No!"

We first note that from the argument in [15] there is a positive constant $\alpha$ such that

$$\sum_{k\le x} P(k;x) \ge \sum_{k\le x,\ k \text{ even}} P(k;x) > x^{1+\alpha}, \tag{16}$$

for all sufficiently large values of $x$. (As reported in [15], it had been previously shown in unpublished correspondence of Davenport and Heilbronn that $\frac{1}{x}\sum_{k\le x} P(k;x)$ tends to infinity as $x$ does.) From [7] we may take $\alpha$ in (16) as any number less than $1 - e^{-1/2}$. It follows from an old conjecture of Erdős that

| $k$ | A | B | C |
|---|---|---|---|
| 2 | 7558421 | 7556879 | 0.999796 |
| 4 | 4047331 | 4043534 | 0.999062 |
| 6 | 20969365 | 20957509 | 0.999435 |
| 8 | 2171817 | 2168840 | 0.998629 |
| 10 | 2048937 | 2043622 | 0.997406 |
| 12 | 11257702 | 11244877 | 0.998861 |
| 14 | 3818031 | 3813803 | 0.998893 |
| 16 | 1171252 | 1166329 | 0.995797 |
| 18 | 7834367 | 7821921 | 0.998411 |
| 20 | 2958324 | 2951746 | 0.997776 |
| 22 | 883419 | 877908 | 0.993761 |
| 24 | 8580206 | 8567413 | 0.998509 |
| 26 | 761625 | 756547 | 0.993333 |
| 28 | 2054653 | 2049723 | 0.997600 |
| 30 | 13781327 | 13762152 | 0.998609 |

Table 5: A = $P(k; 10^{10})$.
B = Number of solutions predicted by Corollary 1.
C = The ratio B/A.

one may take any $\alpha < 1$ in (16). From [15] we have that both sums in (16) are $\leq x^2/\exp((1+o(1))\log x \log\log\log x/\log\log x)$, and from a conjecture there, one can deduce that the first sum in (16) is equal to this expression.

So what can we say about $\sum_{k\leq x} P_0(k;x)$? The following result shows that it is asymptotic to a constant times $x$, so it is much smaller than the sums in (16).

**Theorem 3.** *For a natural number $m$ let $\gamma(m)$ denote the largest squarefree divisor of $m$. Let*

$$c = \sum_{\substack{1\leq a<b \\ (a,b)=1}} \sum_{\substack{r \\ ar+1 \text{ prime} \\ br+1 \text{ prime} \\ (b,ar+1)=1}} \frac{br^2}{(ar+1)(br+1)^2\gamma(a)\gamma(b)}.$$

*Then $c < \infty$ and $\sum_{k\leq x} P_0(k;x) \sim cx$ as $x \to \infty$.*

*Proof.* First we show $c < \infty$. Note that the summand in the definition of $c$ is less than $1/abr\gamma(a)\gamma(b)$. From Brun's method (for example, see Theorem 2.3 in [8], where we choose $g=2,\ a_1=a,\ a_2=b,\ b_1=b_2=1,\ \delta=1$), uniformly in $a, b, y \geq 1$,

$$\sum_{\substack{y\leq r<2y \\ ar+1 \text{ prime} \\ br+1 \text{ prime}}} \frac{1}{r} \ll \frac{b-a}{\phi(b-a)} \cdot \frac{a}{\phi(a)} \cdot \frac{b}{\phi(b)} \cdot \frac{1}{\log^2 2y} \ll \frac{a^{1/2}b^{1/2}}{\log^2 2y},$$

where we have used [10], Theorem 328, for the last inequality. We apply this with $y = 2^j$, $j = 0, 1, \ldots$, and so get that

$$\sum_{\substack{r \\ ar+1 \text{ prime} \\ br+1 \text{ prime}}} \frac{1}{r} \ll a^{1/2} b^{1/2}. \tag{17}$$

To show $c < \infty$ it suffices to show that

$$\sum_{a=1}^{\infty} \frac{1}{a^{1/2}\gamma(a)} < \infty, \tag{18}$$

since by (17),

$$c \ll \sum_{a,b} \frac{1}{a^{1/2} b^{1/2} \gamma(a)\gamma(b)} < \left( \sum_{a=1}^{\infty} \frac{1}{a^{1/2}\gamma(a)} \right)^2.$$

To see (18), note that $1/a^{1/2}\gamma(a)$ is a multiplicative function whose value at the prime power $p^j$ is $1/p^{1+j/2}$. Thus the sum in (18) is equal to

$$\prod_{p \text{ prime}} \left(1 + \frac{1}{p^{3/2}} + \frac{1}{p^2} + \ldots\right) = \prod_{p \text{ prime}} \left(1 + \frac{1}{(p^{1/2} - 1)p}\right) < \infty.$$

Now we complete the proof of the theorem. A solution to $\phi(n) = \phi(n + k)$ of the type considered in Theorem 1 corresponds to a quadruple $g, r, a, b$ where $a < b$, $(a, b) = 1$, $ar + 1$ and $br + 1$ are prime, $\gamma(a)\gamma(b) | g$, and $ar + 1$, $br + 1$ do not divide $g$. The correspondence is that $n = ga(br + 1)$ and $k = g(b - a)$. Thus, $\sum_{k \le x} P_0(k; x)$ is the number of such 4-tuples $g, r, a, b$ with $ga(br + 1) \le x$.

For $a, b, r$ given, we count the number of corresponding $g$'s. This is the number of $g$'s with $g \le x/a(br + 1)$, $g \equiv 0 \bmod \gamma(a)\gamma(b)$ and $(g, (ar + 1)(br + 1)) = 1$. For this count to be nonzero, it is necessary that $(ab, (ar + 1)(br + 1)) = 1$. (Note that $(ab, br + 1) = 1$ all the time and $(a, ar + 1) = 1$ all the time, so the only condition is that $(b, ar + 1) = 1$.) So the number of $g$'s is the number of $h$'s with $h \le x/a(br + 1)\gamma(a)\gamma(b)$ and $(h, (ar + 1)(br + 1)) = 1$. The number of $h$'s is $\le x/abr\gamma(a)\gamma(b)$, and as we have seen above, the sum of this expression over legal choices for $a, b, r$ converges. So we may ignore, say, those values of $a, b, r$ with $abr \ge x^{1/10}$. For any choice of $a, b, r$ the number of $h$'s is equal to

$$\frac{x}{a(br + 1)\gamma(a)\gamma(b)} \cdot \frac{ar}{ar + 1} \cdot \frac{br}{br + 1},$$

with an error at most 2 in absolute value. Since we need only consider those $a, b, r$ with $abr < x^{1/10}$, the errors are negligible and the theorem is proved. □

One might wonder how it can be that $\sum_{k \le x} P_0(k; x)$ can be so much smaller than $\sum_{k \le x} P(k; x)$. Part of the mystery might be explained by the expression $c(k)$ in Corollary 2, which decays rapidly as $k$ grows.

| $q \setminus k$ | 6 | 12 | 18 | 24 | 30 | 36 | 42 | 48 |
|---|---|---|---|---|---|---|---|---|
| 2 | 329118 | 183712 | 130686 | 102936 | 245103 | 73784 | 121860 | 58188 |
| 3 | 8461 | 4852 | 3520 | 2807 | 5518 | 2018 | 1883 | 1617 |
| 4 | 0 | 0 | 0 | 0 | 450 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 16 | 0 | 0 | 0 |

| $q \setminus k$ | 54 | 60 | 66 | 72 | 78 | 84 | 90 | 96 |
|---|---|---|---|---|---|---|---|---|
| 2 | 52824 | 167599 | 44813 | 41783 | 39172 | 69009 | 99269 | 33073 |
| 3 | 1489 | 3206 | 1265 | 1184 | 1107 | 1136 | 2387 | 974 |
| 4 | 0 | 60 | 0 | 0 | 0 | 0 | 204 | 0 |
| 5 | 0 | 10 | 0 | 0 | 0 | 0 | 7 | 0 |

Table 6: Number of solutions to (19) for $n \le 10^{10}$.

## 5. Arithmetic progressions

In the course of our investigation, we also determined all solutions for $n \le 10^{10}$ to the equation

$$\phi(n) = \phi(n+k) = \phi(n+2k) = \ldots = \phi(n+qk) \tag{19}$$

for $1 \le k \le 100$. In the case $k = 1$, there is the well-known progression $\phi(5186) = \phi(5187) = \phi(5188)$; we found no other such progressions with common difference 1. Erdős [5] has conjectured that (19) is solvable for $k = 1$ and any arbitrary $q$. Note that a solution with $q > 2$ immediately implies that $\phi(n) = \phi(n+3)$, and so $n > 10^{11}$. On the other hand, we know of no reason why such solutions should not exist. In general, for values of $k$ that are not multiples of 6, we found only a few progressions of length 3, and none longer. Specifically, there is exactly one progression of length 3 when $k$ is in the set

$$\{1, 2, 4, 5, 8, 11, 14, 23, 25, 26, 28, 29, 31, 37, 38, 41, 46, 47, 52, 53, 55, 56, \\ 58, 59, 62, 67, 71, 73, 74, 76, 79, 80, 85, 86, 89, 92, 94, 97, 98\}.$$

There are exactly two progressions of length 3 when $k$ is in the set

$$\{16, 17, 22, 32, 34, 43, 44, 61, 82, 83, 88\}.$$

Finally, there are exactly three progressions of length 3 when $k = 64$ or $k = 68$. We found no more than three progressions of length 3 for any values of $k \le 100$ that is not a multiple of 6.

By contrast, we found many solutions to (19) when $k$ is a multiple of 6. Table 6 contains a summary of the number of such solutions up to $q = 5$, the largest value of $q$ for which a progression was found among the numbers up to $10^{10}$. The first progression of length 6 that we found has $k = 30$; it is

$$\phi(583200) = \phi(583230) = \phi(583260) = \phi(583290) = \phi(583320) = \phi(583350),$$

with common value 155520.

Theorem 1 may be generalized to a result on arithmetic progressions.

**Theorem 4.** *Suppose that $j, j+k, \ldots, j+qk$ all have the same prime factors. Define $B = j(j+k)\cdots(j+qk)$. For $i = 0, \ldots, q$, define*

$$b_i = \frac{B}{j+ik},$$
$$g = \gcd(b_0, b_1, \ldots, b_q),$$
$$a_i = \frac{b_i}{g} = \frac{B}{(j+ik)g}.$$

*Suppose that for some positive integer $r$,*

$$a_0 r + 1, a_1 r + 1, \ldots, a_q r + 1$$

*are all primes that do not divide $j$. If*

$$n = j(a_0 r + 1) = \frac{Br}{g} + j,$$

*then*

$$\phi(n) = \phi(n+k) = \ldots = \phi(n+qk).$$

The proof is a straightforward extension of the proof of Theorem 1; we leave it as an amusing exercise for the reader.

This theorem gives an explanation for why we found a preponderance of arithmetic progressions when $k$ is a multiple of 6. For if $k$ is a multiple of 6, then the hypotheses of Theorem 4 are satisfied with $q = 2$ and $a_0 = 6, a_1 = 3, a_2 = 2$. If $6r+1, 3r+1$, and $2r+1$ are all prime and if $n = k(6r+1)$, then $\phi(n) = \phi(n+k) = \phi(n+2k)$, and Conjecture 1 predicts that there are infinitely many such $n$. In fact, Conjecture 1 gives infinitely many solutions to (19) for any $q$.

**Corollary 2.** *Assume that Conjecture 1 is true. Then for any positive integer $q$, there exists a positive integer $k$ and infinitely many positive integers $n$ such that*

$$\phi(n) = \phi(n+k) = \ldots = \phi(n+qk).$$

*Proof.* Let $j$ be the product of all primes $p \leq q+1$, and take $k = j$. Then $\{j, j+k, \ldots, j+qk\} = \{j, 2j, \ldots, (q+1)j\}$. Since all prime divisors of $1, 2, \ldots, q+1$ divide $j$, we see that $j, j+k, \ldots, j+qk$ all have the same prime factors. In this case, $a_i = L/(i+1)$, where $L = \mathrm{LCM}[1, 2, \ldots, q+1]$. It remains to note that Conjecture 1 implies that there are infinitely many integers $r$ such that $a_0 r+1, a_1 r+1, \ldots, a_q r+1$ are all prime, so that Theorem 4 completes the proof. □

We have used the construction given in the last proof to search for long arithmetic progressions of equal phi-values. We took the above construction with $q = 9$, so that $j = k = \prod_{p \leq 10} p = 210$ and

$$a_i = \frac{2520}{i+1}$$

for $i = 0, \ldots, 9$. We searched for values of $r < 10^9$ such that $a_i r + 1$ is prime for $i = 0, \ldots, 9$. To speed up the search, we sifted out all values of $r$ for which

$a_i r + 1$ has a prime divisor $< 200$ for some $i$. The only solution we found was $r = 950077810$. Consequently, taking

$$n = 210(2520r + 1) = 502781177052210 \quad \text{and} \quad k = 210$$

gives an arithmetic progression of length 10 with equal phi-values. In other words,

$$\phi(502781177052210 + 210i) = 114921411897600$$

for $0 \le i \le 9$.

All of the data described here, as well as the programs used, are available upon request from the second author.

**Acknowledgements.** We thank R. Tijdeman for directing us to references [14], [22], and [23]. We thank A. Granville for a helpful conversation. We would also like to thank K. Ford for helpful conversations and for sharing unpublished work with us.

## References

[1] Baillie, R., Table of $\phi(n) = \phi(n+1)$. Math. Comp. 29 (1975), 329–330.

[2] Bateman, P.T., Horn, R.A., A heuristic asymptotic formula concerning the distribution of prime numbers. Math. Comp. 16 (1962), 363–367.

[3] Cassels, J.W.S., On the equation $a^x - b^y = 1$. Amer. J. Math. 75 (1953), 159–162.

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

[5] Erdős, P., Some remarks on Euler's $\phi$-function and some related problems. Bull. Amer. Math. Soc. 51 (1945), 540–545.

[6] Erdős, P., Pomerance, C., Sárközy, A., On locally repeated values of certain arithmetic functions, II. Acta Math. Hungar. 49 (1987), 251–259.

[7] Friedlander, J.B., Shifted primes without large prime factors. In: Number theory and applications (ed. by R.A. Mollin; NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. 265), 393–401. Kluwer Acad. Publ., Dordrecht 1989.

[8] Halberstam, H., Richert, H.-E., Sieve methods. Academic Press, London 1974.

[9] Hardy, G.H., Littlewood, J.E., Some problems in "Partitio Numerorum," III: On the expression of a number as a sum of primes. Acta Math. 44 (1923), 1–70.

[10] Hardy, G.H., Wright, E.M., An introduction to the theory of numbers. The Clarendon Press, Oxford Univ. Press, New York 1979.

[11] Lal, M., Gillard, P., On the equation $\phi(n) = \phi(n+k)$. Math. Comp. 26 (1972), 579–583.

[12] LeVeque, W.J., On the equation $a^x - b^y = 1$. Amer. J. Math. 74 (1952), 325–331.

[13] Moser, L., Some equations involving Euler's totient function. Amer. Math. Monthly 56 (1949), 22–23.

[14] Pólya, G., Zur arithmetischen Untersuchung der Polynome. Math. Z. 1 (1918), 143–148.

[15] Pomerance, C., Popular values of Euler's function. Mathematatika 27 (1980), 84–89.

[16] — On the distribution of amicable numbers, II. J. Reine Angew. Math. 325 (1981), 183–188.

[17] Schinzel, A., Sur l'équation $\phi(x+k) = \phi(x)$. Acta Arith. 4 (1958), 181–184.

[18] Schinzel, A., Sierpiński, W., Sur certaines hypothèses concernant les nombres premiers. Acta Arith. 4 (1958), 185–208; erratum 5 (1959), 259.

[19] Schinzel, A., Wakulicz, A., Sur l'équation $\phi(x+k) = \phi(x)$, II. Acta Arith. 5 (1959), 425–426.

[20] Scott, R., On the equations $p^x - b^y = c$ and $a^x + b^y = c^z$. J. Number Theory 44 (1993), 153–165.

[21] Sierpiński, W., Sur une propriété de la fonction $\phi(n)$. Publ. Math. Debrecen 4 (1956), 184–185.

[22] Thue, A., Bemerkungen über gewisse Näherungsbrüche algebraischer Zahlen. Skrift. Vidensk. Selsk. Christ. (1908), Nr. 3.

[23] Tijdeman, R., On integers with many small prime factors. Compositio Math. 26 (1973), 319–330.