Isolating the prime numbers

Adrian M. Stokes

February 29, 2024

Abstract

Prime numbers greater than 3 belong to the number sequences $6n \pm 1$ where $n \geq 1$. These sequences also include the composites that are not divisible by 2 and/or 3 and therefore their factors must also be of the form $6n \pm 1$. This allows all of the $6n \pm 1$ composites to be equivalently written in the form of factors $(6n_1 \pm 1)(6n_2 \pm 1)$, where n_1 and $n_2 \geq 1$, creating three sub-sequences that exclude prime numbers. Finding and isolating the prime numbers can be achieved by selecting a number range and creating a set of $6n \pm 1$ numbers for that range before subtracting the subsets $(6n_1 \pm 1)(6n_2 \pm 1)$ to isolate and identify all the primes in the set.

1 The sequences $6n \pm 1$

All prime numbers greater than 3 belong to the sequences $6n \pm 1$ where $n \geq 1$ since all other natural numbers greater than 3 are composites divisible by 2 and/or 3. For a given value of n, $6n \pm 1$ produces a number pair, which for $n = 1$, 2 and 3 results in twin primes i.e. 5 and 7, 11 and 13, and 17 and 19 respectively. For the purpose of isolating primes greater than 3 from the composites that belong to the same sequences, however, the problem is that when $n \geq 4$ prime/composite pairs are also generated and furthermore when $n \geq 20$ composite pairs are observed too [1]. Nevertheless, $6n \pm 1$ is a promising place to start since separating the primes from the composites in these sequences would enable all primes greater than 3 to be readily isolated and identified. A useful observation at this juncture is that the factors for all $6n \pm 1$ composites must also belong to the $6n \pm 1$ sequences otherwise these composites would be divisible by 2 and/or 3 and therefore not conform to $6n \pm 1$. For example, the first three composites in these sequences are $5 \times 5 = 25$, $5 \times 7 = 35$ and $7 \times 7 = 49$. We will return to this train of thought shortly but first we will take a diversion that will help to explain how a solution for isolating primes emerged.

2 Squares and primes

In the course of my investigation, I had the idea that primes may occur at relatable intervals between successive square numbers. This was prompted by the thought that composite numbers can be represented geometrically by squares and rectangles with primes forming intermediate, irregular shapes inbetween. To test this, I created a spreadsheet with each row commencing with a successive square number so that the first row started with $1 \times 1 = 1$, the next with $2 \times 2 = 4$, then $3 \times 3 = 9$ and so on. The rows formed sequences of successive numbers so each row was a number line offset relative to the previous row by virtue of commencing with a larger square number. So the first row started 1, 2, 3,..., x, the second 4, 5, 6, ..., x, the third 9, 10, 11, ..., x and the rows continued to an upper limit of 243^2 to ensure an adequate sample size. Next I set about identifying all the primes in the resulting table by using an online database of primes and shading all of the cells containing prime numbers using two colours, one for $6n-1$ and another for $6n+1$ primes.

Interestingly, prime rich vertical and diagonal lines emerged across the table much the same or at least remarkably similar to those of a Ulam spiral [2]. However, the two colours of shading also revealed that these lines sometimes included only $6n-1$ or $6n+1$ primes whilst others contained a mix of the two. The unbroken diagonal sequence of primes in Table 1 below was especially interesting because it was a particularly good example of a pattern whereby a $6n+1$ prime was followed by two $6n-1$ primes (shown in bold) with this pattern repeated throughout the sequence (note the sequence commences with a single $6n-1$ prime before entering the pattern described). It was this pattern that led me to revisit the factors of the $6n \pm 1$ composites.

Table 1

41	43	**47**	**53**	61	**71**	**83**	97	**113**	**131**
151	**173**	**197**	223	**251**	**281**	313	**347**	**383**	421
461	**503**	547	**593**	**641**	691	**743**	**797**	853	**911**
971	1033	**1097**	**1163**	1231	**1301**	**1373**	1447	**1523**	**1601**

3 The factors of the $6n \pm 1$ composites

Let's look at first three composites of the sequences and in particular their factors. First we note that 25 belongs to the sequence $6n+1$ where $n = 4$. Its factors are 5×5, which can also be written as $(6n_1 - 1)(6n_2 - 1)$ where both values of $n = 1$. Now let's consider 49. This also belongs to the sequence $6n+1$ where $n = 8$. This time, though, its factors are 7×7, which can equivalently be written as $(6n_1 + 1)(6n_2 + 1)$ where again both values of $n = 1$. We therefore have two different examples of $6n+1$ composites, one that has two $6n-1$ factors whilst the other has two $6n+1$ factors. In both cases it is the multiplication of like signs that gives a $6n+1$ composite. Turning to 35, its factors are 5 and 7. So in this example we have the multiplication of factors with unlike signs giving rise to a $6n-1$ composite where $n = 6$.

In summary, we have two types of $6n+1$ composites and only one type of $6n-1$ composite as defined by the signs for 1 within the factors $(6n_1 \pm 1)(6n_2 \pm 1)$. This is the opposite situation to that in Table 1 where, excluding the first value, we have one $6n+1$ prime for every two $6n-1$ primes. It is as if the different types of composites in the form of their factors is producing the distribution we see in Table 1 by a subtraction process.

We are now closing in on a solution to isolating the primes from the composites. Key to this is the ability to express each and every $6n \pm 1$ composite in an equivalent factors form $(6n_1 \pm 1)(6n_2 \pm 1)$. On the other hand, primes can only be expressed in the form $6n \pm 1$ since their only factors are themselves and 1. This means that for a selected number range we can calculate subsets of composites using the factors form knowing primes are excluded. We can then subtract these subsets from the $6n \pm 1$ sets for the same number range to leave only primes. Armed with this conclusion, we can now write the formulae for isolating prime numbers greater than 3.

Let's start by letting $6n-1$ be set A and $6n+1$ set B. From set A we will subtract the subset of composites $\{(6n_1-1)(6n_2+1)\}$, A_c, to leave the $6n-1$ primes, A_p. So:

$A_p = A - A_c$

For set B, we need to deduct two subsets, one for $\{(6n_1-1)(6n_2-1)\}$ composites, which we'll call B_c, and one for $\{(6n_1+1)(6n_2+1)\}$ composites, which we'll call B_k. This gives us the following formula:

$B_p = B - B_c - B_k$

To test these findings, I created a model using a spreadsheet workbook to find and isolate all of the primes greater than 3 in the sequence of natural numbers up to 50,000. The model correctly identified and isolated 5,131 primes (excluding 2 and 3) for this range in agreement with an online database of prime numbers. A sample of 1,500 numbers from the sequence were checked methodically to ensure that both primes and composites were correctly identified. No errors or omissions were found.

4 Additional thoughts

Although the findings described in the "Squares and primes" section above help to explain how the thought process evolved, ultimately they have no bearing on the conclusion reached. However, it is interesting to observe the emergence of lines with significant concentrations of primes and how those lines differ in their proporions of $6n-1$ and $6n+1$ primes.

Of perhaps greater interest is the possibility that two sets of an equal number elements (i.e. $6n-1$ and $6n+1$) contain differing proportions of composites and primes owing to the inequality in composite subsets between the two sets. A deeper analysis of this possibility would be needed but a brief review suggests that there are offsetting factors that mean the primes across both sets are very nearly distributed $50:50$ but not quite. In fact, in the small number of different number sequences I have considered, there always seems to be a slight excess of set A primes relative to set B primes. It would be easy to assume that the differences are so slight that probability could offer an explanation but given the findings presented here it is perhaps a line of enquiry worth exploring further. If it is found that there is indeed some bias in the distribution of primes in the context of the two $6n \pm 1$ sequences it would be interesting to see if such effects had wider implications beyond the distribution of prime numbers.

References

[1] N. Lord. Prime numbers and the sequences 6n ± 1. *The Mathematical Gazette*, 98(541):126–128., 2014.

[2] M.L. Stein, S.M. Ulam, and M.B. Wells. A visual display of some properties of the distribution of primes. *The American Mathematical Monthly*, 71(5):516–520., 1964.