Musing on the primes

There are Infinitely many Primes – I

But how many proofs of this?

Numbers have been a subject of fascination from the most ancient times, and people keep coming up with families of numbers: integers, rational numbers, numbers, real numbers, complex numbers, prime numbers, Fermat numbers, Bernoulli numbers, ... Mathematics teacher D R Kaprekar (1905–1985) found many new families, giving them curious names like Dattatreya numbers, Demlo numbers, monkey numbers, and so on. India's great mathematician S Ramanujan who made a large number of discoveries in number theory found a new family of numbers which he called 'highly composite numbers'. Back in the Greek era, Pythagoras, steeped in mysticism, referred to numbers as sacred, lucky, evil and so on. (Sacred numbers are difficult to find these days. But 13 continues to be unlucky!) For the rest of this article, when we use the word 'number' we mean natural number or positive integer, i.e., one of the numbers 1,2,3,4,5, ...

feature

V G Tikekar

Prime numbers

On the vast canvas of numbers there is one special category, the *prime numbers* (or just 'primes'), which have been a source of interest to mathematicians since ancient times. Not only do they display very beautiful and surprising properties, they also find unexpected application in fields like coding and cryptography.

Early in our encounter with numbers, we discover that there are infinitely many of them, meaning that their supply can never be exhausted. For, no matter how large the number that we name, we can produce a larger one by adding 1 to it.

Keywords: *Numbers, prime, composites, infinite, factorial, coprime, Euclid, contradiction, Pólya, Fermat number*

A number n exceeding 1 is said to be *prime* if it has no divisors among the set of natural numbers, other than 1 and itself. If n does have divisors other than 1 and n it is called *composite*. Note that the number 1 does not get classified by these two definitions. We call 1 a *unit*; it is neither prime nor composite. So the primes are these numbers:

$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, \ldots,$

and the composites are:

$4, 6, 8, 9, 10, 12, 14, 15, 16, 18, 20, 21, 22, 24, \ldots.$

After 2, every even number is composite, so we cannot find stretches of consecutive numbers which are all prime. But we do seem to find long stretches of consecutive numbers which are all composite. For example, 14, 15, 16 is a stretch of three such numbers, and 24, 25, 26, 27, 28 is a stretch of five composites. Here is a stretch of seven composites: $90, 91, 92, 93, 94, 95, 96$. Can we find even longer stretches of composites? Can we, say, find a billion consecutive numbers, all composite? The surprising answer is: Yes!

Here is a simple argument showing why. Let n be any number, $n > 1$. Consider the following $n - 1$ consecutive numbers defined using the factorial function (recall that $n!$ is the product $1 \times 2 \times 3 \times \cdots \times (n-1) \times n$):

$$n!+2,\ n!+3,\ n!+4,\ n!+5,\ \ldots,\ n!+(n-1),\ n!+n.$$

These $n - 1$ numbers are all composite; for, $n!+2$ is a multiple of 2; $n!+3$ is a multiple of 3; ...; $n!+n$ is a multiple of n. (In general, $n!+k$ is a multiple of k if k lies between 2 and n.) So simply by choosing n to be extremely large, we can construct very long stretches of consecutive composite numbers. (Example: Take $n = 6$; we get the following stretch of five consecutive composites: 722, 723, 724, 725, 726.) This establishes the claim.

How many primes are there?

There are obviously infinitely many composites (indeed, every even number after 2 is composite), but we cannot be so sure about the primes. For one thing, they start to thin out! For example, there are 168 primes between 1 and 1000; 135 primes between 1000 and 2000; 127 primes between 2000 and 3000; ...and 98 primes between 20000 and 21000. The number is clearly coming down, so we may wonder whether a point will come, far down the number line, when they vanish altogether.

This question was posed by the ancient Greeks, and answered: they proved that there is no `last prime'; in short, there are infinitely many primes. The oldest proof known of this remarkable claim is found in the great text written by Euclid, *The Elements*. Since then many more proofs have been found by famous mathematicians.

Euclid's proof

It is curious that Euclid's beautiful proof is found in a book that is generally considered to be a text in geometry! But in fact there are several topics in this book which would nowadays be regarded as part of number theory.

Euclid's proof is based on the principle of 'proof by contradiction' . It starts by supposing that what we wish to prove is *false*, then examines what follows from this supposition—in the hope of finding something contrary. If such a contradiction is found, it shows that what was assumed at the start necessarily has to be false. In other words, the statement we wish to prove must be true. This strange-sounding strategy for proof is a corner stone for the development of modern mathematics. Let's see how Euclid carries out this strategy.

Let the primes be $p_1, p_2, p_3, \ldots$; here $p_1 = 2$ is the first prime, $p_2 = 3$ is the second prime, $p_3 = 5$ is the third prime, and so on. Suppose there is a 'last prime' p_n. (This is precisely the supposition we hope to demolish.) We now construct the following number X by adding 1 to the product of all these primes:

$$X = p_1 p_2 p_3 \cdots p_n + 1.$$

It should be clear that X leaves remainder 1 when divided by p_1. In fact it leaves remainder 1 when divided by each of the primes $p_1, p_2, p_3, \ldots, p_n$. This means, in particular, that: *X is not divisible by any of the primes $p_1, p_2, p_3, \ldots, p_n$.*

What kind of number is X? It is either prime or composite. If it is the former then we have a new

prime number (X itself), different from $p_1, p_2, p_3, \ldots, p_n$. If X is not prime then it has a prime divisor q different from $p_1, p_2, p_3, \ldots, p_n$. (It cannot be any of these since X is not divisible by any of these primes.) *Whichever possibility happens, we obtain a prime number different from* $p_1, p_2, p_3, \ldots, p_n$. *So* $\{p_1, p_2, p_3, \ldots, p_n\}$ *cannot be the complete set of primes.* Hence there cannot be a last prime number p_n, and the number of primes is infinite.

We can try out this argument with some actual numbers to see how it works.

- Imagine we thought that 3 is the last prime number (!); then the set of primes is $\{2, 3\}$, and $X = (2 \times 3) + 1 = 7$. It happens that 7 is prime, so we have found a new prime, contrary to our supposition that 3 is the last prime.
- Imagine we thought that 5 is the last prime number; then the set of primes is $\{2, 3, 5\}$, and $X = (2 \times 3 \times 5) + 1 = 31$. It happens that 31 is prime, so we have found a new prime, contrary to our supposition that 5 is the last prime.
- Similarly, if we imagined 13 to be the last prime number, so that the set of primes is $\{2, 3, 5, 7, 11, 13\}$, then $X = (2 \times 3 \times 5 \times 7 \times 11 \times 13) + 1 = 30031$. It happens that 30031 is composite, and its prime factorization is $30031 = 59 \times 509$. So we have found two new primes (59 and 509), contrary to our supposition that 13 is the last prime.

Notice how carefully Euclid has framed the argument. He has never claimed that X is prime, only that a new prime will be found by this means whether X is prime or composite. The proof is indeed a classic.

Variants of Euclid's proof

There are other proofs of the infinitude of primes that closely resemble Euclid's proof but are not the same (though they are clearly modelled on Euclid's proof). We sketch a few here.

(1) Instead of using the number

$$X = p_1 p_2 p_3 \cdots p_n + 1$$

we could as well work with the number $Y = p_1 p_2 p_3 \cdots p_n - 1$. We need $n > 1$, to avoid triviality. The rest of the proof is the same as earlier.

- If $n = 2$ we get $Y = (2 \times 3) - 1 = 5$ which is prime.
- If $n = 4$ we get $Y = (2 \times 3 \times 5 \times 7) - 1 = 209 = 11 \times 19$ which yields two new primes, 11 and 19.

The same reasoning works in all cases; we see that there must be a prime number other than $p_1, p_2, p_3, \ldots, p_n$.

(2) We could also use the factorial function. If K is the supposed largest prime we could work with the number defined by $Z = K! + 1$. Once again the same reasoning works and yields new primes.

(3) The following proof is due to the German mathematician Ernst Kummer (1810--1893), and it is a genuine proof by contradiction. Suppose that p_n is the last prime and that $\{p_1, p_2, p_3, \ldots, p_n\}$ is the complete set of primes. As before we construct the number $X = p_1 p_2 p_3 \cdots p_n - 1$. This number must have a prime divisor, and the divisor must be one of the primes $p_1, p_2, p_3, \ldots, p_n$, because we have supposed that these are *all* the primes that exist. Suppose that the prime divisor of X thus defined is p_k.

Now, clearly, p_k is a divisor of $X + 1$ too (since $X + 1$ is the product of the primes $p_1, p_2, \ldots, p_n$). But if p_k divides X as well as $X + 1$, then p_k must divide the difference between $X + 1$ and X, which is 1. This however is absurd: no prime number can be a divisor of 1. So we have found the desired contradiction, and the conclusion follows that there are infinitely many primes.

A presentation not based on 'proof by contradiction'

There is even a way of presenting Euclid's proof in which we do not emphasize the contradictory aspect; it would not be called a proof by contradiction. We phrase it in a positive manner by claiming that: *Given any finite set S of primes, it is possible to find a prime number that is not in S.*

The idea is exactly the same: we construct a number N which is 1 more than the product of all the numbers in S. Then N is either a prime number, or it has a prime divisor q. Either way we obtain a new prime (N or q) which does not lie in S.

It is fairly obvious that all these are 'children' of Euclid's proof.

Pólya's proof

In contrast, here is a proof which is genuinely different. It is due to the great mathematician educator George Pólya (Christian Goldbach had had exactly the same idea), and it uses the Fermat numbers F_n defined by:

$$F_n = 2^{2^n} + 1.$$

For example we have: $F_0 = 2^1 + 1 = 3$; $F_1 = 2^2 + 1 = 5$; and following these:

$$F_2 = 2^4 + 1 = 17,$$
$$F_3 = 2^8 + 1 = 257,$$
$$F_4 = 2^{16} + 1 = 65537.$$

The five numbers listed are all primes, but that should not fool us, for the very next Fermat number is not prime:

$$F_5 = 2^{32} + 1 = 4294967297 = 641 \times 6700417.$$

Pólya observed that these numbers have the following very nice property:

> *The Fermat numbers are mutually coprime:* $\gcd(F_m, F_n) = 1$ *for all* $m \neq n$.

This follows from the fact that the Fermat numbers are all odd (which is obvious), and they obey the following identity for all $n \geq 1$:

$$F_n - 2 = F_0 \times F_1 \times F_2 \times \cdots \times F_{n-1}.$$

For example, take $n = 3$; we have $255 = 3 \times 5 \times 17$. We shall leave the proof of the identity as an exercise, as also the proof of coprimeness of the Fermat numbers.

Taking the claim as proved for now, we show how it implies that there are infinitely many primes. Each Fermat number has associated with it its own set of prime divisors. These sets must all be disjoint (this is what 'coprime' implies). So for each number n we have a non-empty set of primes corresponding to n. Taking the union of these sets, we see that there must be infinitely many primes. *Note that this is not a proof by contradiction.*

Exercises

(1) Prove the relation $F_{n-2} = F_0 \times F_1 \times F_2 \times \cdots \times F_{n-1}$ for the Fermat numbers. Hint: Use the principle of induction.

(2) Show how the above relation, together with the fact that the Fermat numbers are odd, implies that these numbers are mutually coprime. Hint: Suppose some prime p divides both F_m and F_n where $m > n$. Using the above identity show that p must divide 2. But this is absurd, since p must be odd.

References

[1] Association of Math Teachers of India (AMTI), *The Wonder World of Kaprekar Numbers*

[2] Robert Kanigel, *The Man Who Knew Infinity: A Life of the genius Ramanujan*, Maxwell Macmillan International, 1991, p. 232

[3] Paul Hoffman, *Archimedes' Revenge*, Ballantine Books, 1988, p. 7

PROF. V.G. TIKEKAR retired as the Chairman of the Department of Mathematics, Indian Institute of Science, Bangalore, in 1994. He has been actively engaged in the field of mathematics research and education and has taught, served on textbook writing committees, lectured and published numerous articles and papers on the same. Prof. Tikekar may be contacted on vgtikekar@gmail.com.