IS ERATOSTHENES OUT?

GEORGE LEDIN, JR.
Institute of Chemical Biology, University of San Francisco, San Francisco, Calif.

Two thousand years ago the Alexandrian geographer-astronomer Eratosthenes, a friend of Archimedes, devised a procedure for obtaining a list of primes. His procedure is usually identified as "the Sieve" and basically consists of writing a table of consecutive integers starting from 1 and crossing out all multiples of 2, 3, and so on; all those numbers which remain undeleted are the primes sought. This procedure can be extended to larger tables from 1 to N, but when N is large, the sieve is indeed a cumbersome tool. Nevertheless, Eratosthenes' procedure is the only general way of obtaining primes in an orderly fashion today. Extensive tables have been compiled, but no formula that would yield the n^{th} prime for a given n has been found yet; many a mathematician doubt that such a formula exists. When confronted with the question, "What is the n^{th} prime?" all a mathematician can do is look in a table of primes, and if asked, "Is this number a prime?" the mathematician may not be able to reply at all, for although there are tests for primality, they might not be applicable or may prove insufficient, and if the number given is too large, it might not be listed in the tables. The puzzling aspect of the situation is that, although prime numbers are not randomly distributed along the sequence of integers, their distribution has so far defied all attempts at exact description. Despite the countless efforts, number-theorists are not happy with the idea of settling for the "simple-minded" Eratosthenes' Sieve.

This paper presents two elementary glimpses of modified but simple approaches to the Sieve. The first one is a slight improvement on the original procedure of Eratosthenes, although it is basically the same method, cleverly disguised.

Consider the "Semi-Tribonacci" sequence

T_k: 1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19, 20, 22, 23, $\cdots$

($k = 1, 2, 3, \cdots$)

which obeys the recurrence relation

$$T_{k+3} = T_{k+2} + T_{k+1} - T_k ; \qquad T_1 = 1, \; T_2 = 2 .$$

Notice that all multiples of 3 are absent, since

$$T_{2k} = 3k - 1$$

and

$$T_{2k+1} = 3k + 1 \ .$$

The closed-form formula for these Semi-Tribonacci numbers is:

$$T_k = \tfrac{3}{2}k - (\tfrac{3}{4} + \tfrac{1}{4}(-1)^k) \quad (k = 1, 2, 3, \cdots) \tag{1}$$

Now, if we write the above sequence cancelling all T_k such that

$$T_k \equiv T_n \pmod{2T_n + 3} \quad (T_k > T_n) \tag{2}$$

(i.e., cancel all $T_k \equiv 1 \pmod 5$, $T_k \equiv 2 \pmod 7$, $T_k \equiv 4 \pmod{11}$, etc.) we obtain the "Deleted Semi-Tribonacci Sequence:"

$\overline{T}_k$: 1, 2, 4, 5, 7, 8, 10, 13, 14, 17, 19, 20, 22, $\cdots$, $(k = 1, 2, 3, \cdots)$

And here we can state the following result:

All numbers

$$P_{k+2} = 2\overline{T}_k + 3$$

are prime numbers, and, in fact, all primes (except 2 and 3) are represented in this way. Thus

$$P_{k+2} = 5, \ 7, \ 11, \ 13, \ 17, \ 19, \cdots \ (k = 1, 2, 3, \cdots).$$

The above may seem quite astonishing at first sight. The reader is invited to convince himself that this is, however, true. But, unfortunately, it is only the Sieve covered up. The core of the problem lies in the solution of the following congruences

$$4T_k \equiv (6n - 3 - (-1)^n) \pmod{12n + 6 - 2(-1)^n} \tag{3}$$

which is, to put it mildly, quite a problem by itself.

The second glimpse offers a simpler disguise, but cleverer. Consider the array

$$\begin{array}{cccccccc} 4 & 7 & 10 & 13 & 16 & 19 & \cdots \\ 7 & 12 & 17 & 22 & 27 & 32 & \cdots \\ 10 & 17 & 24 & 31 & 38 & 45 & \cdots \\ 13 & 22 & 31 & 40 & 49 & 58 & \cdots \\ 16 & 27 & 38 & 49 & 60 & 71 & \cdots \\ \cdot & & & & & & \\ \cdot & & & & & & \\ \cdot & \cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdots \\ \cdot & \cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdots \end{array} \tag{4}$$

The array is symmetric about its main diagonal, for as it is readily seen, each k^{th} row and k^{th} column are equal, and the numbers are obtained from arithmetic progressions. The differences are: first line, 3, second line, 5, third line, 7, and so on. We are now prepared to formulate the following statement:

If the number N is a member of the above array, then $2N + 1$ is composite; however, if N is not found in this array, then $2N + 1$ is prime. ($2N + 1$ is prime if and only if N is not a member of the above array.)

The proof is very simple. Designate the n^{th} term of the k^{th} row (or k^{th} term of the n^{th} column) by a_{nk}. Then, since

$$a_{n1} = 4 + 3(n - 1), \quad a_{n2} = 7 + 5(n - 1),$$

etc., in general we have

$$a_{nk} = 1 + 3k + (1 + 2k)(n - 1). \tag{5}$$

or more simply

$$(6) \qquad a_{nk} = k + (2k + 1)n = a_{kn} = n + (2n + 1)k$$

Now suppose N is found in the array. Then $N = k + (2k + 1)n$ and therefore

$$2N + 1 = 2(k + (2k + 1)n) + 1 = 2k + 1 + 4kn + 2n = 2k + 1 + 2n(2k + 1)$$
$$= (2k + 1)(2n + 1)$$

which means that $2N + 1$ is the product of at least two factors (neither of which is unity) and hence, composite. The converse is proven similarly.

The following example may be useful to compare the powerfulness of the array (4) as opposed to the naive Sieve. Let us suppose that we wanted to find out whether 437 was or was not a prime. Using the rudimentary approach of the Sieve, we would test for divisibility of all primes up to

$$[\sqrt{437}] = 20 ,$$

that is, we would see if 437 is divisible by 3, 5, 7, 11, 13, 17, and 19. Instead of proceeding this way, let us apply the reasoning provided to us by the array's approach.

If 437 is not a prime, we can find an N in the array such that

$$2N + 1 = 437 .$$

This would yield $N = 218$. Is 218 a member of the array? If it is, we should be able to find it as some n^{th} element of some k^{th} row. Thus, we should be able to solve for n the equation

$$k + (2k + 1)n = 218$$

(if we fail, this would mean that 218 is not in the array, and that 437 is prime). First, we find a bound on k by solving the quadratic

$$2(k^2 + k) = 218 ,$$

and this yields

$$k^2 + k - 109 = 0$$

or $k = 10$.

Thus,

$$k = 10, \quad n = 208/21 \quad \text{(no good)}$$
$$k = 9, \quad n = 209/19 = 11$$

and we get

$$a_{9,11} = a_{11,9} = 218.$$

Therefore 218 is contained in the array, and 437 is not prime.

In fact, if we had tried it using the Sieve method we would have found out, sooner or later, that $437 = 19 \cdot 23$. For large numbers, the array test is tedius although shorter than Eratosthenes'.

Nowadays, with the advent of superfast computers, much of the sieve work is done electronically at very high speeds. Still, the job of classifying larger numbers as primes is very difficult and can only be simplified by choosing specific patterns within sequences of identifiable properties. That, for example, is the case of the 3,376-digit number $(2^{11,213} - 1)$ which belongs to the "Mersenne" family of primes and is presently considered the largest known prime number. Other, modern, more effective sieves are inevitably based on the Sieve or its principle.

Despite the fact that mathematics has progressed immeasurably and contemporary mathematicians have the benefit of ultra-sophisticated tools and techniques, Eratosthenes' method has survived the severe test of twenty centuries. Indeed, Erathosthenes is still not out.

★ ★ ★ ★ ★