

An elementary determination of prime numbers

by Jörg W. Müller

Bureau International des Poids et Mesures, F-92312 Sèvres Cedex

Abstract

By the systematic use of residues, the location of prime numbers, for a given interval beyond an initial value N , can be easily determined. The only primes which have to be known in advance are those below \sqrt{N} .

1. Introduction

Prime numbers are usually determined from the sequence of integers by some selection process. The best known example is the sieve of Eratosthenes where, for a given upper limit N , all multiples of the known primes $p_1 = 2$, $p_2 = 3$, $p_3 = 5$, $p_4 = 7$, up to $p_n \leq \sqrt{N}$, are eliminated. The remaining integers are the prime numbers sought. This procedure is equivalent to identifying the consecutive integers $2, 3, 4, \dots, N$ which are *not* divisible by any of the primes $p \leq \sqrt{N}$. For large values of N , the method becomes rather cumbersome.

It would be desirable, therefore, to have available a recurrence relation for prime numbers which, for example, would allow us, on the basis of the known primes p_1, p_2, \dots, p_n , to calculate the next prime number p_{n+1} .

Unfortunately, no such recurrence is known today, and apparently only few people seem still to believe in its very existence. In fact, the required numerical results can always be obtained by some "brute force" technique, although no real insight can then be expected. Considering the many unsuccessful attempts that have been made, no excessive hope should be invested in such an enquiry, especially as it may have no practical value. In spite of all this, a fresh approach may be useful.

2. Basis of the approach

What follows is essentially an elementary exercise in congruences. As is well known, residues play a major role in all questions of divisibility, and therefore also for prime numbers. We shall try to take advantage of this situation.

If, for two given natural numbers N and m , we write

$$N = r \pmod{m}, \tag{1}$$

this means that

$$N = k m + r ,$$

with k an integer; m is called the module and r the residue. Two integers N_1 and N_2 with the same residue r (for a given m) are said to be congruent, or equal modulo m .

Usually residues are restricted to the range

$$0 \leq r \leq m-1 , \quad (2)$$

but this is arbitrary and can be changed. It is equally possible to use instead of r , for example, residues of the form

$$r+m, \quad r+2m \quad \text{or} \quad r-m = -(m-r)$$

since they are all equal modulo m .

For what follows it is helpful to introduce the quantity

$$v = m - r , \quad (3)$$

which is a kind of "negative residue" and therefore called "nessi" for short (in allusion to the Loch Ness monster). While r indicates by how much N exceeds a certain multiple of the module m , v describes by how much it fails to reach the next one. Their sum is therefore equal to m , or more generally $0 \pmod{m}$. The relation between residue and nessi is shown graphically in Fig. 1.

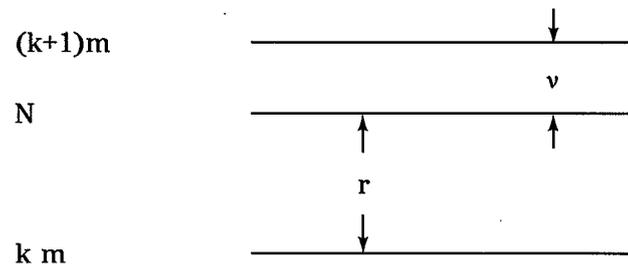


Fig. 1: Display of the relation between residue r and nessi v ($k = \text{integer}$).

Both residues and nesses are always positive integers. For a simple numerical illustration (with $m = 7$) consider the number 11:

$$11 = 4 \pmod{7} = -3 \pmod{7} ,$$

where $r = 4$ and $v = 7 - 4 = 3$.

We note that, since the integer N in (1) is an arbitrarily chosen initial value (i.e. a prime or a composed number), there may be zeros among its nesses and residues. They are only absent if N is a prime number.

From now onwards our modules will always be prime numbers. For $m = p_i$ relation (1) can be written with (3) as

$$N = r_i \pmod{p_i} = -v_i \pmod{p_i} . \quad (4)$$

3. Evaluation of primes

Let us now try to locate the next prime number P beyond N , for which we put

$$P = N + d . \quad (5)$$

The quantity sought is the distance d .

Since residues are additive, we obtain from (4) and (5)

$$P = d - v_i \pmod{p_i} . \quad (6)$$

However, the necessary condition for P to be prime is that *none* of its residues vanish, i.e. that

$$v_i \neq d , \quad \text{for all relevant primes } p_i . \quad (7)$$

In consequence, the prime number in (5) is determined by the smallest value of d which does *not* appear among the nesses of N . This solves our problem.

In the listing of the nessi values v_i (for an illustration see below) we must obviously also include the values

$$v_i + k p_i , \quad k = 1, 2, 3, \dots , \quad (8)$$

as they are all congruent modulo p_i . Limits on these quantities are considered later.

If we go far enough in listing the nessi values v_i , we may find additional (higher) values (d') which are also missing. They correspond to the further prime numbers

$$P' = N + d' ,$$

which are thus obtained as a welcome by-product.

4. Practical applications

To illustrate the simplicity of the method described above, we suggest two numerical applications.

a) Evaluation of some prime numbers beyond 100

For simplicity, $N = 100$ is chosen. The prime numbers p_i to be considered are below \sqrt{N} , thus $p_1 = 2$, $p_2 = 3$, $p_3 = 5$ and $p_4 = 7$.

Since $100 = 34 \cdot 3 - \underline{2} = 15 \cdot 7 - \underline{5}$, we find for the initial nessi values

$$v_1 = 0, v_2 = 2, v_3 = 0 \text{ and } v_4 = 5.$$

Repeated addition of p_i then gives the nessi values listed in Table 1.

Table 1: Table of the nessi values v_i , for $N = 100$ and $v_i \leq 15$.

$i =$	1	2	3	4
$p_i =$	2	3	5	7
$v_i =$	0	2	0	5
	2	5	5	12
	4	8	10	
	6	11	15	
	8	14		
	10			
	12			
	14			

By searching in Table 1 for missing integers we find

$$d = 1, 3, 7, 9, 13.$$

Since $P = N + d$, we readily obtain the first primes beyond 100, namely

$$P = 101, 103, 107, 109, 113.$$

This illustrates both the simplicity and the efficiency of the approach.

b) *Primes beyond 12 553*

The list of primes given in Riesel [1] ends at $N = p_{1500} = 12\,553$. Let us determine the subsequent ones, say in a range of 50.

Since $\sqrt{p_{1500}} \cong 122$, the largest prime to be considered in our analysis is $p_{29} = 109$. This should still be manageable "by hand".

The list of the nessi values can now be shortened. Since the distances d we are looking for are even, only even values of v_i will be given (after the first line). We then arrive at Table 2.

Table 2: List of (even) nessi values v_i , for $N = p_{1500}$, up to $v_i^{(last)} \geq 50$.

$p_i =$	3	5	7	11	13	17	19	23	29	31	37	41
$v_i =$	2	2	5	9	5	10	6	5	4	2	27	34
	8	12	12	20	18	44	44	28	62	64	64	116
	14	22	26	42	44	78	82	74				
	20	32	40	64	70							
	26	42	54									
	32	52										
	38											
	44											
	50											
				$p_i =$	43	47	53	59	61	67	71	73
				$v_i =$	3	43	8	14	13	43	14	3
					46	90	114	132	74	110	156	76
					132							
				$p_i =$	79	83	89	97	101	103	107	109
				$v_i =$	8	63	85	57	72	13	73	91
					166	146	174	154	274	116	180	200

Inspection of Table 2 reveals that

$$d = 16, 24, 30, 36 \text{ and } 48$$

are the only even integers absent up to 50. This yields five prime numbers following $p_{1500} = 12\,553$ within an interval of 50, namely

$$P = 12\,569, 12\,577, 12\,583, 12\,589 \text{ and } 12\,601.$$

5. Final remarks

The two examples just discussed show that the method suggested for evaluating prime numbers works well, at least for relatively small numbers. However, the procedure can readily be programmed and should then be applicable to much larger primes. Only the last nessi values need be kept for later use : all the others are held in a common memory which may be inspected for absent values.

As we can always choose an odd number N as starting point, only even values of v_i are candidates for d , and the long column for $p_i = 2$ may be omitted. From the first even value of v_i onwards, the others are obtained by the repeated addition of $2p_i$.

We leave open the question of whether the procedure described may actually be considered a recurrence formula; this is a problem of semantics. In any case, it leads to equivalent results.

The number of primes that have to be known in advance increases with N , but their values are limited to $p_i \leq \sqrt{N}$. This is reminiscent of the similar condition which holds for formulae that permit the exact evaluation of $\Pi(N)$, the number of primes up to N , based on an approach first suggested by Legendre (for details, see [1]).

It may be of interest to obtain an estimate for the fraction $1/R$ of prime numbers actually needed in the calculation. By using the well-known approximation

$$\Pi(N) \cong \frac{N}{\ln N}, \quad (9)$$

it is easy to calculate

$$R = \frac{\Pi(N)}{\Pi(\sqrt{N})} \cong \sqrt{N} \frac{\ln \sqrt{N}}{\ln N} = \frac{1}{2} \sqrt{N}. \quad (10)$$

The reduction rapidly becomes quite important. The numerical values compare favourably with those applicable in the two examples given above.

When choosing an upper limit D for the range of explored values d , we must ensure that no new prime p_i will be required. If the last prime accounted for is p_f , the limit D must be chosen below $(p_{f+1})^2 - N$. For the two examples considered before, this leads to the upper limits

$$D_a = 11^2 - 100 - 1 = 20$$

and

$$D_b = 113^2 - 12\,553 - 1 = 215,$$

which have both been respected. Without this condition $(p_{f+1})^2$ would be taken as prime. If the limit is not fixed in advance, the inclusion e.g. of p_{f+1} and p_{f+2} among the primes p_i provides a simple protection.

After listing a number of nessi values v_i with N as starting point, one may wish to update them with respect to a higher reference value N' which, for example, can be the last prime number (with the highest value d) determined previously. This is easily done. If we start, for a given prime p_i , with the last nessi value - which can always be arranged to be even and not smaller than d , as is the case in Table 2 -, then the initial values v'_i for the new round are given by

$$v'_i = v_i^{(\text{last})} - d.$$

To them we now add multiples of $2p_i$, at least up to a limit D' . This allows us to work with smaller nessi numbers for which we determine (as before) the integers $d' \leq D'$ which do *not* appear in the list, in order to obtain the subsequent primes $P' = N' + d'$.

It will be realized that the suggested method is equivalent to the sieve of Eratosthenes. All we have done is introduce a grid that starts at N and is based on residues which can readily be updated. Once the nessi values are known for N , none of the usual divisions is needed; all are replaced by simple additions.

This short note is probably the last, in a rather long series of BIPM reports issued in the past 30 years, to appear while I am still an active staff member of the BIPM.

In looking back I am deeply impressed by the decisive support and the friendly encouragement I have always received during this period from so many colleagues, present or past. To thank them all most cordially, I take the opportunity to dedicate this report to each of them. I realize only too well that without their unfailing help, in whichever form it reached me, nothing could have been achieved.

Finally, I would like to encourage researchers to devote their effort not only to solving the practical problems posed by their daily work, but also on occasion to giving some thought to those which, considered by many as too difficult or irrelevant, are often left aside. In basic questions, even small advances are welcome, as they may prove to be necessary steps on the way to bigger ones.

Reference

- [1] H. Riesel: "Prime Numbers and Computer Methods for Factorization" (Birkhäuser, Basel, 1994²)

(April 1996)