THE TEACHING OF MATHEMATICS
2000, Vol. III, 2, pp. 63–82

SELECTED CHAPTERS FROM ALGEBRA

I. R. Shafarevich

Abstract. This paper is the fourth part of the publication "Selected chapters from algebra", the first three having been published in previous issues of the Teaching of Mathematics, Vol. I (1998), 1–22, Vol. II, 1 (1999), 1–30, Vol. II, 2 (1999), 65–80, and Vol. III, 1 (2000), 15–40.

AMS Subject Classification: 00 A 35

Key words and phrases: Primes, function $\pi(n)$, Chebyshev inequality, asymptotical law of distribution of primes.

CHAPTER IV. PRIMES

1. Infinity of the number of primes

In this chapter we return to the question which we have already dealt with in Chapter I. It was proved there that each natural number can be uniquely represented as a product of primes. Therefore, when the multiplication is concerned, the primes are the simplest elements and all the natural numbers can be obtained by multiplying primes, similarly to the fact that they can be obtained by the operation of addition starting from the number 1. From this point of view, the interest for the set of all primes can be easily understood. There are four primes among the first ten natural numbers: 2, 3, 5, 7. Further primes can be found by dividing each of the consequent numbers by previously found primes, in order to decide whether it is a prime itself. In this way we find the following 25 primes among the first one hundred natural numbers:

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97.$$

How far does this sequence continue? The question arose already in the antique times. The answer was given by Euclid:

Theorem 1. *The number of primes is infinite.*

We give several proofs of this theorem.

First proof—the one contained in Euclid's "Elements". Suppose we have found n primes: $p_1, p_2, \dots, p_n$. Consider the number $N = p_1 p_2 \cdots p_n + 1$. As we saw

This paper is an English translation of: И. Р. Шафаревич, *Избраные главы алгебры. Глава IV. Простые числа*, Математическое образование, N 1(4), янв.–мар. 1998, Москва, стр. 2–21. In the opinion of the editors, the paper merits wider circulation and we are thankful to the author for his kind permission to let us make this version.

in §2 of Chapter I, each number has at least one prime divisor. In particular, N has a prime divisor. But none of the numbers p_1, p_2, ..., p_n can divide N. To see this, let p_i be a divisor of N. Then $N - p_1 \cdots p_n$ must be divisible by p_i, but since $N - p_1 \cdots p_n = 1$, this is impossible. It follows that this prime divisor must be different from each p_i, $i = 1, \dots, n$, which means that after each n primes there must be at least one additional prime. This proves the theorem.

Second proof. According to the theorem of the section "Set algebra" of Chapter III, the number of numbers which are smaller than the given number N and relatively prime with it, is given by the formula

$$N\left(1-\frac{1}{p_1}\right)\left(1-\frac{1}{p_2}\right)\cdots\left(1-\frac{1}{p_n}\right), \tag{1}$$

where p_1, ..., p_n are all prime divisors of N. We shall prove the theorem by contradiction. Suppose that the number of primes is finite and that p_1, ..., p_n are all of them. Set $N = p_1 \cdots p_n$. Substituting in formula (1) we obtain for each factor $p_i(1-\frac{1}{p_i})$ the expression $p_i - 1$, and for the whole product (1) the expression $(p_1-1)(p_2-1)\cdots(p_n-1)$. As we know that there exist primes greater than 2 (for example, 3), the number obtained is *greater* than 1. Hence, there exists a number a, smaller than N, relatively prime with N and different from 1. But a has at least one prime divisor which must be contained among the numbers p_1, ..., p_n, and so a cannot be relatively prime with N. We obtained a contradiction which proves the theorem.

The infinite sequence of primes is, on the other hand, very sparsely distributed among natural numbers. For example, there are arbitrary big "gaps" in this sequence, i.e., one can find (successively further away) any given number of consecutive numbers which are not prime. For example, n numbers $(n+1)!+2$, $(n+1)!+3$, ..., $(n+1)!+n+1$ are obviously not primes—the first one is divisible by 2, the second by 3, the last by $n+1$.

For some time mathematicians have searched for a formula expressing primes. For example, Euler found an interesting polynomial $x^2 + x + 41$, which, for 40 values of x—from 0 to 39—obtains prime values. However, it is obvious that for $x = 40$ its value is a nonprime number 41^2. It is not hard to conclude that there cannot exist a polynomial $f(x)$ which takes prime values for all natural values $x = 0, 1, 2, \dots$ (not even speaking about the possibility that its values are *all* of the primes). We shall show this on an example of a polynomial of second degree $ax^2 + bx + c$ with integer coefficients a, b, c. Suppose that for $x = 0$ the polynomial has a prime value c. Then for each $x = kc$ its value $ak^2c^2 + bkc + c$ is divisible by c. This value can be equal to c for at most one additional value of k (besides $k = 0$), which can be easily checked. Moreover, there does not exist a polynomial $f(x) = ax^2 + bx + c$ having prime values for each integer x, *starting from some limit.* Indeed, suppose that the values of the polynomial $f(x)$ are prime for each $x \geqslant m$. Set $x = y + m$, $f(y+m) = g(y)$; then all the values of the polynomial $g(y)$ are prime for all integers $y \geqslant 0$, by the assumption, and its coefficients are also integers, since $g(y) = a(y+m)^2 + b(y+m) + c$. The same reasoning also applies to a polynomial of an arbitrary degree n: $f(x) = a_0 + a_1x + \cdots + a_nx^n$. If all of its values for integer

$x \geqslant 0$ are prime, it means that $f(0) = a_0 = p$ is prime, too. Then for each integer k the values $f(kp) = p + a_1kp + \cdots + a_n(kp)^n$ are divisible by p. They can be equal to p only if $p + a_1kp + \cdots + a_n(kp)^n = p$, i.e., $a_1 + a_2kp + \cdots + a_n(kp)^{n-1} = 0$, and the last equation in k is of the degree $n-1$, and according to Theorem 3 of Chapter II it has at most $n-1$ roots. For all other values of k the number $f(kp)$ is divisible by p and different from p, i.e., it is not a prime.

If we suppose that the values of the polynomial $f(x)$ are prime only for integer values of $x \geqslant m$, for a certain number m, then we can set $x = y + m$ and $f(y+m) = g(y)$. The polynomial $g(y) = a_0 + a_1(y+m) + \cdots + a_n(y+m)^n$ is obtained by expanding all the parentheses by the binomial formula and reducing similar terms. Therefore its coefficients are again integers, but it obtains prime values for *all* integers $y \geqslant 0$, which again is a contradiction.

It can be also proved that for an arbitrary number k no polynomial in k variables with integer coefficients exists such that all of its values for all natural values of its variables are primes. Nevertheless, it appears that there is a polynomial of degree 25 with 26 variables, having the following property: if we select those values of that polynomial which are obtained for nonnegative integer values of its variables and which are positive themselves, then the set of such values coincides with the set of primes. Since 26 is equal to the number of letters of the Latin alphabet, it is possible to denote the variables by the letters: a, b, $\dots$, x, y, z. Then the polynomial is of the form:

$$\begin{gathered}
F(a,b,c,d,e,f,g,h,i,j,k,l,m,n,o,p,q,r,s,t,u,v,w,x,y,z) = \\
= (k+2)\{1 - [wz + h + j - q]^2 - [(gk + 2g + k + 1)(h + j) + hz]^2 - \\
-[2n + p + q + z - e]^2 - [16(k+1)^3(k+2)(n+1)^2 + 1 - f^2]^2 - \\
-[e^3(t+2)(a+1)^2 + 1 - o^2]^2 - [(a^2-1)y^2 + 1 - x^2]^2 - \\
-[16r^2y^4(a^2-1) + 1 - u^2]^2 - [(a + u^2(u^2 - a^2) - 1)(n + 4dy)^2 + 1 - (x + cu)^2]^2 - \\
-[n + l + v - y]^2 - [(a-1)l^2 + 1 - m^2]^2 - [ai + k + 1 - l - i]^2 - \\
-[p + l(a - n - 1) + b(2an + 2a - n^2 - 2n - 2) - m]^2 - \\
-[q + y(a - p - 1) + s(2ap + 2a - p^2 - 2p - 2) - x]^2 - \\
-[z + pl(a - p) + t(2ap - p^2 - 1) - pm]^2\}.
\end{gathered}$$

This polynomial has been written here just to impress the reader. Its number of variables is too big. It can be proved that it takes also negative values $-m$, where m is not prime. Hence, it does not give us information about the sequence of primes either.

Long trials convinced the majority of mathematicians that there is no easy formula describing the sequence of primes. There exist “explicit formulae” describing the primes, but they use objects which are even less known than the primes themselves. That is why mathematicians concentrated on the characteristics of the sequence of primes “in total” and not “in parts”. We will deal with this kind of questions in the next sections.

Problems

1. Prove that there are infinitely many primes of the form $3s+2$.

2. The same for the primes of the form $4s+3$.

3. Prove that each two numbers $2^{2^n}+1$ and $2^{2^m}+1$ are relatively prime. Deduce once more the infinity of the number of primes. [*Hint.* Assuming that p is a common divisor of two such numbers, find the remainders of division of 2^{2^m} and 2^{2^n} by p.]

4. Let $f(x)$ be a polynomial with integer coefficients. Prove that there exist infinitely many distinct prime divisors of its values $f(1)$, $f(2)$, $\dots$. (If you do not succeed immediately, solve the problem for the polynomials of the first and of the second degree.)

5. Denote by p_n the n-th prime in the natural order. Prove that $p_{n+1} < p_n^n+1$.

6. Using the notation of Problem 5, prove that $p_n < 2^{2^n}$. Deduce the similar inequality $p_n \leqslant 2^{2^n}+1$ from the result of Problem 3.

7. Using the notation of Problem 5, prove that $p_{n+1} < p_1p_2\cdots p_n$.

2. Euler's proof of the infinity of the number of primes

We shall give another proof of the infinity of the number of primes, which is due to Euler, and which clarifies some general properties of this sequence.

Let us start with the "prehistory", that is, with some simple facts which had been known before Euler started dealing with questions about primes. The question is about how big the following sums can be:

$$1, \quad 1+\frac{1}{2}, \quad 1+\frac{1}{2}+\frac{1}{3}, \quad \dots, \quad 1+\frac{1}{2}+\dots+\frac{1}{n}, \quad \dots$$

Using notation from section 3 of Chapter II, these are the sums $(Sa)_n$, where a is the sequence of the inverses of natural numbers $1, \frac{1}{2}, \frac{1}{3}, \dots$. Since we denoted the sums of the m-th powers of natural numbers from 1 to $n-1$ by $S_m(n)$ (cf. formula (28) of Chapter II), it is natural to denote our sums by $S_{-1}(n)$.

We have come to a concept which we shall often deal with later, so we consider it now in more detail. It refers generally to properties of an *infinite* sequence of positive numbers $s_1, s_2, \dots, s_n, \dots$ (in our case it appeared as the sequence of sums of another sequence, but for the moment that is of no importance). One type of such sequences is called *bounded.* This means that there exists a number C (the same for the whole sequence), such that $s_n < C$ for all $n = 1, 2, 3, \dots$. If the sequence does not have this property, it is called *unbounded.* This means that *no* number C can possess this property, i.e., for each number C there exists an index n such that $s_n \geqslant C$. Finally, it may happen that for each number C there exists an index n such that $s_m \geqslant C$ for all $m = n, n+1, \dots$. In other words, for n sufficiently large, the numbers s_n become arbitrary large. In that case the sequence is called *unboundedly increasing.* For example, the sequence 1, 2, 1, 3, 1, 4, ... , where 1 stands on odd places, and natural numbers stand on even places in succession, is unbounded, but not unboundedly increasing, since one can find the number 1 arbitrarily far in it.

If a sequence $a = a_1, a_2, \ldots, a_n, \ldots$ of positive numbers is given, and $s = Sa$, then $s_{n+1} > s_n$ (since $s_{n+1} = s_n + a_{n+1}$, $a_{n+1} > 0$), and, generally, $s_m > s_n$ for $m > n$. Therefore, such a sequence will be unboundedly increasing if it is unbounded. For example, if all $a_i = 1$, then $s_n = n$ and the sequence $s_1, s_2, \ldots$ is unbounded. But in other cases it may be bounded. An example can be visualised on Fig. 1, where we first divide the segment between 0 and 1 in half and set $a_1 = \frac{1}{2}$, then divide again the segment between 0 and $\frac{1}{2}$ in half and set $a_2 = \frac{1}{4}$, etc. In this way, $a_n = \frac{1}{2^n}$. The result of adding such numbers is represented on Fig. 1 and it is obvious that the sums S_n stay inside our initial segment, i.e., $S_n < 1$.

Fig. 1

It is easy to check the last assertion by calculation. If $a_n = \frac{1}{2^n}$, then

$$(Sa)_n = \frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2^n} = \frac{1}{2}\left(1 + \frac{1}{2} + \cdots + \frac{1}{2^{n-1}}\right),$$

and by formula (12) of Chapter I

$$(Sa)_n = \frac{1}{2}\,\frac{\frac{1}{2^n} - 1}{\frac{1}{2} - 1} = 1 - \frac{1}{2^n},$$

so that $(Sa)_n < 1$ for each n.

We shall show now that in the case of the sequence 1, $\frac{1}{2}$, $\frac{1}{3}$, ... the *first* case appears: although the terms of the sequence decrease, they do not decrease fast enough, and their sums (i.e., $S_{-1}(n)$) increase unboundedly.

Lemma 1. *The sum $S_{-1}(n)$ is, for n sufficiently large, greater than an arbitrary given number.*

Let the number k be given. We assert that for some n (and so also for all greater integers) $S_{-1}(n) > k$. Take n such that $n - 1 = 2^m$ for some m. Divide the sum

$$S_{-1}(n) = 1 + \left(\frac{1}{2}\right) + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \cdots + \left(\frac{1}{2^{m-1}+1} + \cdots + \frac{1}{2^m}\right)$$

in parts as it is shown: in groups contained between two consecutive powers of two. Each parenthesis has the form $\frac{1}{2^{k-1}+1} + \cdots + \frac{1}{2^k}$, and the number of parentheses is equal to m. In each parenthesis we replace each summand by the smallest one entering that parenthesis, that is by the last one. Since the number of summands in such a parenthesis is equal to $2^k - 2^{k-1} = 2^{k-1}$, we obtain that the k-th parenthesis is greater than $\frac{2^{k-1}}{2^k} = \frac{1}{2}$. As a result, we obtain that $S_{-1}(n) > 1 + \frac{m}{2}$. This inequality is valid for each n if $n - 1 = 2^m$. It remains to put $1 + \frac{m}{2} = k$, i.e., $m = 2k - 1$ and $n = 2^{2k-1} + 1$. Then $S_{-1}(n) > k$.

Now we come to Euler's proof. His idea is connected with the method of computing the sums of powers of the divisors of a natural number, which was described in section 3 of Chapter I (cf. formula (13) in Chapter I). Denote the sum of k-th powers of all divisors (including 1 and n) of a natural number n by $\sigma_k(n)$. According to formula (13) of Chapter I, for the number n having canonical factorisation $n = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$,

$$(2) \qquad \sigma_k(n) = \frac{p_1^{k(\alpha_1+1)} - 1}{p_1^k - 1} \frac{p_2^{k(\alpha_2+1)} - 1}{p_2^k - 1} \cdots \frac{p_r^{k(\alpha_r+1)} - 1}{p_r^k - 1}.$$

Formula (2) had been known since the antique times, but it was implicitly assumed that the number k in it was positive. Finally, Euler got interested in it and he posed the question—what would happen if k was integer, but negative? The answer is, of course, that there is no difference, the derivation of formula (2) is completely formal and the same for negative as well as for positive values of k. In particular, it is valid for $k = -1$. The sum of (-1)-st powers (i.e., the inverses) of the divisors of a given number n will be denoted, as before, by $\sigma_{-1}(n)$. Formula (2) gives

$$\sigma_{-1}(n) = \frac{1 - \dfrac{1}{p_1^{\alpha_1+1}}}{1 - \dfrac{1}{p_1}} \cdot \ldots \cdot \frac{1 - \dfrac{1}{p_r^{\alpha_r+1}}}{1 - \dfrac{1}{p_r}}$$

(we interchanged the order of summands in numerators and denominators in each of the fraction). From here (since all the expressions in numerators are less than 1),

$$(3) \qquad \sigma_{-1}(n) < \frac{1}{\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right) \cdots \left(1 - \frac{1}{p_r}\right)}.$$

Let us now replace n in this formula by $n!$ ($p_1, \ldots, p_r$ are now prime divisors of $n!$). The numbers $1, 2, \ldots, n$ are all contained among the divisors of $n!$. Therefore, the sum $\sigma_{-1}(n!)$ definitely contains summands $1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}$, whose sum is equal to $S_{-1}(n+1)$. According to Lemma 1, already the sum $S_{-1}(n+1)$ is greater than any given number k for n sufficiently large. Since other summands in the sum $\sigma_{-1}(n!)$ are positive, the same conclusion is valid for it. If the number of primes were finite and $p_1, \ldots, p_r$ were the whole list of them, we would obtain that

$$\frac{1}{\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right) \cdots \left(1 - \frac{1}{p_r}\right)} > k,$$

where k is an arbitrary number. This is, of course, a contradiction.

The value of the above proof is not that the assumption of finiteness of the number of primes has led to a contradiction, but that it, when the infinity of that number has already been proved, gives some quantitative characteristics of the sequence of primes. Namely, reformulating the result obtained, we can now say that if $p_1, p_2, \dots, p_n, \dots$ is the infinite sequence of primes, then the expression $\dfrac{1}{\left(1-\frac{1}{p_1}\right)\left(1-\frac{1}{p_2}\right)\cdots\left(1-\frac{1}{p_n}\right)}$ becomes greater than any arbitrary number for n sufficiently large. This is, of course, equivalent to the fact that the denominator of the last fraction becomes *smaller* than an arbitrary positive number for n sufficiently large. We have proved

Theorem 2. *If $p_1, p_2, \dots, p_n, \dots$ is the sequence of all primes, then the product $\left(1-\dfrac{1}{p_1}\right)\left(1-\dfrac{1}{p_2}\right)\cdots\left(1-\dfrac{1}{p_n}\right)$, for n sufficiently large, becomes smaller than any given positive number.*

This is a first approximation to our goal. Let us try now to give a more useful form of the characteristic obtained.

Theorem 3. *If $p_1, p_2, \dots, p_n, \dots$ is the sequence of all primes, then the sequence of sums $\dfrac{1}{p_1}+\dfrac{1}{p_2}+\dots+\dfrac{1}{p_n}$ increases unboundedly.*

Derivation of Theorem 3 from Theorem 2 is purely formal: it does not use the fact that $p_1, p_2, \dots, p_n, \dots$ is the sequence of *primes*—it could be an arbitrary sequence of natural numbers which satisfies the conditions of Theorem 2.

Lemma 2. *For each natural number $n > 1$ the inequality*

$$1-\frac{1}{n} \geqslant \frac{1}{4^{1/n}} \tag{4}$$

is valid.

Since both sides of inequality (4) are positive, rasing them to the power of n, we obtain an *equivalent* inequality

$$\left(1-\frac{1}{n}\right)^n \geqslant \frac{1}{4}, \tag{5}$$

which we are going to prove. Expanding the left-hand side by the binomial formula we obtain

$$\left(1-\frac{1}{n}\right)^n = 1-n\,\frac{1}{n}+\frac{n(n-1)}{2}\,\frac{1}{n^2}-\frac{n(n-1)(n-2)}{3!}\,\frac{1}{n^3}+\dots+(-1)^n\frac{1}{n^n}. \tag{6}$$

Absolute values of the summands on the right-hand side of formula (6) form the sequence $C_n^k\frac{1}{n^k}$. We examined such a sequence in connection with the Bernoulli scheme in the section “Language of probability” in Chapter III (formula (35)). More precisely, if in that formula we put $p = \dfrac{1}{n+1}$, $q = 1-\dfrac{1}{n+1} = \dfrac{n}{n+1}$, then we obtain that $p+q = 1$, $p^kq^{n-k} = (n+1)^{-n}n^{n-k}$ and the numbers obtained differ from the ones examined in formula (6) just by the common factor $\left(\frac{n}{n+1}\right)^n$. The expression $(n+1)p-1$ is in our case equal to zero. In the section "Language of probability" of Chapter III we proved that if $k>(n+1)p-1$ (in our case $k>0$), then the $(k+1)$-st term is smaller than the k-th one. This means that the numbers of the sequence $C_n^k\frac{1}{n^k}$, $k=1,2,\dots,n$ decrease monotonously. (We referred here to Chapter III just to stress the connection between different problems that we are dealing with. It would, of course, be easy to write down the ratio of the $(k+1)$-st term of the sequence to the k-th one and conclude that it is less than 1). We can see that in formula (6), the first two terms on the right-hand side cancel. The next two terms (after cancellation which can be done easily) give $\frac{1}{3}-\frac{1}{3n^2}$. This number is not less than $\frac{1}{4}$ for $n\geqslant 2$ (check it yourself!). The rest of the terms can be grouped in pairs, where in each pair the first term is positive and the next negative, but, as we have seen, by absolute value less than the first one. Therefore each pair gives a positive contribution to the sum (6). If n is odd, then the number of summands on the right-hand side of formula (6) is even (it is equal to $n+1$) and the sum is partitioned into $\frac{n+1}{2}$ pairs. If n is even, then after grouping into pairs, there remains the summand $\frac{1}{n^n}$. In such a way, in any case the right-hand side consists of a summand which is not less than $\frac{1}{4}$, and some additional positive summands. This proves inequality (5), and so the lemma itself.

Theorem 3 is now evident. For each p_i we have, according to the Lemma:

$$1-\frac{1}{p_i}\geqslant\frac{1}{4^{1/p_i}}.$$

Multiplying these inequalities for $i=1,\dots,n$ we obtain

$$\left(1-\frac{1}{p_1}\right)\left(1-\frac{1}{p_2}\right)\cdots\left(1-\frac{1}{p_n}\right)\geqslant\frac{1}{4^{\left(\frac{1}{p_1}+\frac{1}{p_2}+\cdots+\frac{1}{p_n}\right)}}.$$

If the sums $\frac{1}{p_1}+\frac{1}{p_2}+\cdots+\frac{1}{p_n}$ were for each n less than a certain value k, it would follow that

$$\left(1-\frac{1}{p_1}\right)\left(1-\frac{1}{p_2}\right)\cdots\left(1-\frac{1}{p_n}\right)\geqslant\frac{1}{4^k}.$$

This contradicts Theorem 2.

We run here into a problem of a new kind. If N is a subset of a finite set M, then we can tell how much "smaller" N is than M, comparing the number of their elements, e.g., computing the ratio $n(N)/n(M)$. But now we have two infinite sets: the set of all natural numbers and the set of all primes contained in it. How can we compare them? Theorem 3 offers one way of comparing, not very easy at first sight. It can be applied to each sequence of natural numbers a: a_1, a_2, $\dots$, a_n, $\dots$. According to Lemma 1, for the sequence of all natural numbers, the sums of their inverses (i.e., the sums $S_{-1}(n)$) increase unboundedly. We can think of the sequence a to be "tightly" distributed among natural numbers if it has the same property, i.e., if the sums $\frac{1}{a_1}$, $\frac{1}{a_1}+\frac{1}{a_2}$, $\dots$, $\frac{1}{a_1}+\frac{1}{a_2}+\cdots+\frac{1}{a_n}$, $\dots$ unboundedly increase. This means that in the sequence a enough natural numbers remained so that the sums of their inverses are not too much less than the sums $S_{-1}(n)$ of the inverses of all natural numbers. If, on the other hand, the sums of inverses of the sequence a remain bounded, we can think of it as "loosely" distributed in the natural row. Theorem 3 states that the sequence of primes is "tight". The most "loose" case is the case of a sequence a having only a finite number of terms.

But there are intermediate cases. For instance, the sequence of squares: 1, 4, 9, ... , n^2, ... It is natural to denote the corresponding sums $1+\frac{1}{4}+\frac{1}{9}+\dots+\frac{1}{n^2}$ by $S_{-2}(n)$. We shall prove that they are bounded by a number not depending on n. We use the same idea as in the proof of Lemma 1. Let m be such that $2^m \geqslant n$. Then $S_{-2}(n) \leqslant S_{-2}(2^m)$. We divide the sum $S_{-2}(2^m) = 1+\frac{1}{2^2}+\frac{1}{3^2}+\dots+\frac{1}{2^{2m}}$ into parts:

$$(1)+\left(\frac{1}{2^2}\right)+\left(\frac{1}{3^2}+\frac{1}{4^2}\right)+\dots+\left(\frac{1}{(2^{m-1}+1)^2}+\dots+\frac{1}{2^{2m}}\right).$$

Each part $\dfrac{1}{(2^{k-1}+1)^2}+\dots+\dfrac{1}{2^{2k}}$ again contains 2^{k-1} terms and the first term is the greatest. Therefore this part cannot be greater than $2^{k-1}\dfrac{1}{(2^{k-1}+1)^2} < 2^{k-1}\dfrac{1}{(2^{k-1})^2} = \dfrac{1}{2^{k-1}}$. Therefore, $S_{-2}(2^m) \leqslant 1+1+\dfrac{1}{2}+\dfrac{1}{2^2}+\dots+\dfrac{1}{2^{m-1}} = 1+\dfrac{1-\frac{1}{2^m}}{1-\frac{1}{2}} \leqslant 1+\dfrac{1}{1-\frac{1}{2}} = 3$. So, none of the sums $S_{-2}(n)$ is greater than 3.

In such a way, Theorem 3 shows that, for example, the primes are distributed more "tightly" in the natural row than the squares.

Problems

1. Prove that for each given integer $k \geqslant 2$ and for all natural n, the sums $S_{-k}(n) = \frac{1}{1^k}+\frac{1}{2^k}+\dots+\frac{1}{n^k}$ are bounded.

2. Let the sequence a be an arithmetic progression: $a_0 = p$, $a_1 = p+q$, $a_2 = p+2q$, ... , $a_n = p+nq$ for some natural p and q. Prove that the sums $\frac{1}{a_0}$, $\frac{1}{a_0}+\frac{1}{a_1}$, ... , $\frac{1}{a_0}+\frac{1}{a_1}+\dots+\frac{1}{a_n}$, ... become unboundedly large for n sufficiently large.

3. Let the sequence a be a geometric progression: $a_0 = c$, $a_1 = cq$, $a_2 = cq^2$, ... , $a_n = cq^n$, ... , where c and q are natural numbers. Is it "tight" or "loose" in the natural row?

4. Let $p_1, \dots, p_n, \dots$ be the sequence of all primes. Prove that the expressions $\dfrac{1}{\left(1-\frac{1}{p_1^2}\right)\left(1-\frac{1}{p_2^2}\right)\cdots\left(1-\frac{1}{p_n^2}\right)}$ are bounded for each n.

3. The function $\pi(n)$

In this section we will try once more to estimate how much the sequence of primes differs from the sequence of all natural numbers. We will replace the more elaborate method of comparing "tight" and "loose" sequences from the previous section by a more naive one, which can be understood more easily. Namely, we will try to answer the naive question—"which portion of the sequence of natural numbers is covered by the primes"— by finding how many primes there are smaller than 10, how many smaller than 100, how many smaller than 1000, etc. For each natural number n, denote by $\pi(n)$ the number of primes not greater than n, so that $\pi(1) = 0$, $\pi(2) = 1$, $\pi(4) = 2, \ldots$. What can be said about the ratio $\frac{\pi(n)}{n}$ when n increases?

First of all, consider what can be learned from tables. Each assertion or question concerning natural numbers can be checked for all natural numbers not exceeding a certain limit N. This fact plays a role in the number theory, which investigates properties of natural numbers, similar to that of the possibility of experimenting in theoretical physics. in particular, one can compute the values $\pi(n)$ for $n = 10^k$, $k = 1, 2, \ldots, 10$. The following table is obtained.

n	$\pi(n)$	$\dfrac{n}{\pi(n)}$
10	4	2.5
100	25	4.0
1000	168	6.0
10000	1229	8.1
100000	9592	10.4
1000000	78498	12.7
10000000	664579	15.0
100000000	5761455	17.4
1000000000	50847534	19.7
10000000000	455059512	22.0

Table 1.

We see that the ratio $\frac{n}{\pi(n)}$ is constantly increasing, which means that $\frac{\pi(n)}{n}$ is decreasing all the time. In other words, the portion of primes among the first n numbers becomes close to zero when n increases. According to the tables, it could be said that "the primes constitute a zero portion among all natural numbers". That was the way Euler formulated this fact, although his reasoning did not contain a full proof. We will now give the precise formulation and then the proof.

Theorem 4. *The ratio $\frac{\pi(n)}{n}$ becomes smaller than any given positive number for n sufficiently large.*

In order to prove the theorem we have to estimate somehow the function $\pi(n)$. For actual calculation of its values we start with the prime 2, then we cancel all the numbers which are multiples of 2 and not exceeding n. Then we take the first remaining number—this will be 3—and repeat the process. We continue till we have exhausted all the numbers not exceeding n. The numbers which are not cancelled (2, 3, etc.) are all primes not exceeding n. This method was used already in the antique times; it is called "the sieve of Eratosthenes".

We will apply the same process in our reasoning. Suppose we have already found the first r primes: $p_1, p_2, \dots, p_r$. Then the remaining primes, not exceeding n, are contained among "noncancelled" numbers, not exceeding n, i.e., among those numbers $m \leqslant n$ which are not divisible by any of the numbers $p_1, p_2, \dots, p_r$. But the number of numbers not exceeding n and not divisible by any of the primes $p_1, p_2, \dots, p_r$ was explored in Chapter III—it is given by the formula in the section Set algebra of Chapter III. As we showed there, the expression in the formula can be replaced with the easier one $n\left(1-\frac{1}{p_1}\right)\cdots\left(1-\frac{1}{p_r}\right)$, where the error is less than 2^r (formula (28) of Chapter III). Hence, the number s of numbers $m \leqslant n$ not divisible by any of the primes $p_1, p_2, \dots, p_r$ satisfies the inequality

$$s \leqslant n\left(1-\frac{1}{p_1}\right)\cdots\left(1-\frac{1}{p_r}\right)+2^r. \tag{7}$$

All $\pi(n)$ primes not exceeding n are contained either among r primes $p_1, p_2, \dots, p_r$, or among s numbers accounted for by inequality (7). In such a way, $\pi(n) \leqslant s + r$, and

$$\pi(n) \leqslant n\left(1-\frac{1}{p_1}\right)\cdots\left(1-\frac{1}{p_r}\right)+2^r+r. \tag{8}$$

Inequality (8) is remarkable because it contains the product $\left(1-\frac{1}{p_1}\right)\cdots\left(1-\frac{1}{p_r}\right)$ which can be estimated using Theorem 2.

Now we can pass to the proof of Theorem 4. Let an arbitrary small positive number ε be given. We have to find a number N, depending on ε, such that $\frac{\pi(n)}{n} < \varepsilon$ is valid for each $n > N$. In the inequality (8) we replace r by a greater number 2^r (cf. Problem 6 in section 2 of Chapter I), in order to obtain a simpler inequality

$$\pi(n) \leqslant n\left(1-\frac{1}{p_1}\right)\cdots\left(1-\frac{1}{p_r}\right)+2^{r+1}. \tag{9}$$

In the inequality (9) there are two summands and we shall choose N so that for each $n \geqslant N$ each of the summands will not exceed $\varepsilon n/2$. Then from the inequality (9) we will conclude that $\pi(n) < \varepsilon n$, and so $\frac{\pi(n)}{n} < \varepsilon$. Recall that till now the number r in our reasoning was arbitrary. We choose it so that the first summand does not exceed $\varepsilon n/2$, and then we choose N such that the second summand does not exceed $\varepsilon n/2$. The first choice is possible according to Theorem 2. It states that for r sufficiently large, the product $\left(1-\frac{1}{p_1}\right)\cdots\left(1-\frac{1}{p_r}\right)$ is less than any arbitrary given positive number. We can take $\varepsilon/2$ to be such a number. Then the first summand in the inequality (9) does not exceed $\varepsilon n/2$. The second summand can be dealt with even more easily. Now r has already been chosen. Choose N such that $2^{r+1} < eN/2$. For this it is enough to choose $N > \dfrac{2^{r+2}}{\varepsilon}$. Then $2^{r+1} < \dfrac{\varepsilon N}{2} \leqslant \dfrac{\varepsilon n}{2}$ for each $n \geqslant N$. Theorem 4 is proved.

Note that if we choose an arithmetic progression $am+b$ even with a very big difference a, i.e., being very "sparse", then the number of terms of this progression not exceeding n is the same as the number of integers m satisfying $am \leqslant n-b$, i.e., $\left[\frac{n-b}{a}\right]$. We saw in section 3 of Chapter III that $\left[\frac{n-b}{a}\right]$ differs from $\frac{n-b}{a}$ by not more than 1. Hence, the number of terms of the progression not exceeding n is not less than $\frac{n-b}{a}-1$. Its quotient with n is not less than $\frac{1}{n}(\frac{n-b}{a}-1)=\frac{1}{a}-\frac{1}{n}\frac{b}{a}-\frac{1}{n}$. When n increases, this number approaches $\frac{1}{a}$ and is not becoming arbitrarily small. Thus, Theorem 4 would become false if we replaced the sequence of primes in it by an arbitrary arithmetic progression. This shows that primes are distributed more sparsely than any arithmetic progression.

Problems

1. Let p_n denote the n-th prime. Prove that for an arbitrarily large positive number C the inequality $p_n > Cn$ is valid for n sufficiently large. [*Hint.* Use the fact that $\pi(p_n)=n$.]

2. Consider natural numbers having the property that, when written in decimal form, they do not contain a certain digit (e.g., 0). Let $q_1, q_2, \dots, q_n, \dots$ be those numbers, written in ascending order, and let $\pi_1(n)$ denotes the number of such numbers not exceeding n. Prove that the ratio $\frac{\pi_1(n)}{n}$ becomes smaller than any arbitrary given positive number, for n sufficiently large. Prove that the sums $\frac{1}{q_1}, \frac{1}{q_1}+\frac{1}{q_2}, \dots, \frac{1}{q_1}+\frac{1}{q_2}+\dots+\frac{1}{q_n}, \dots$ are bounded. [*Hint.* Do not try to copy the proof of Theorem 4. Split the sum to parts, where in each part denominators are contained between 10^k and 10^{k+1}. Then find the number of numbers q_i in such intervals. The answer depends on the digit which is excluded: $r=0$ or $r\neq 0$.]

APPENDIX

Inequalities of Chebyshev for $\pi(n)$

We have put this text in the Appendix mainly for formal reasons, because we have to use logarithms, the knowledge of which is not assumed in the rest of the text. Recall that the *logarithm with basis a of the number x* is a number y such that

$$a^y = x.$$

This is written as $y=\log_a x$. In the sequel it will always be assumed that $a>1$ and that x is a positive number. Basic properties of logarithms, following directly from the definition, are:

$$\log_a(xy)=\log_a x+\log_a y, \quad \log_a c^n = n\log_a c, \quad \log_a a = 1.$$

$\log_a x > 0$ if and only if $x>1$. Logarithm is a monotonous function, i.e., $\log_a x \leqslant \log_a y$ if and only if $x \leqslant y$.

In this text, if the basis of a logarithm is not indicated, it is supposed that it is equal to 2; $\log x$ means $\log_2 x$.

The second reason for putting this text in the Appendix is the following. In the rest of the book, the logic of reasoning was clear, namely, why do we follow a certain road (at least I hope it was so). Here we encounter the case, not rare in mathematical investigations, when it looks as if some new consideration has "jumped in from nowhere", when even the author cannot explain how he came to the conclusion. About such situations Euler used to say: "Sometimes it seems to me that my pencil is smarter than myself". Of course, these are results of long trials and unknown work of psyche.

We are going now to continue our investigations concerning the ratio $\frac{\pi(n)}{n}$ when n is increasing unboundedly. Take another look at Table 1, showing the values of $\pi(n)$ for $n = 10^k$, $k = 1, 2, \dots, 10$. The last column of the table contains the values of the ratio $\frac{n}{\pi(n)}$ for some values of n. We see that when we pass from $n = 10^k$ to $n = 10^{k+1}$, that is when we go down by one row, the value of $\frac{n}{\pi(n)}$ increases always approximately by the same value. Namely, the first number is equal to 2.5; the second differs from it for 1.5, and the further differences are: 2; 2.1; 2.3; 2.3; 2.3; 2.4; 2.3; 2.3. We see that all of these numbers are close to the one: 2.3. Not trying at the moment to explain the meaning of this particular value, let us suppose that also beyond the range of our table the numbers $\frac{n}{\pi(n)}$, when passing from $n = 10^k$ to $n = 10^{k+1}$, increase by amounts which are closer and closer to a certain constant α. This would mean that $\frac{n}{\pi(n)}$ for $n = 10^k$ would be very close to αk. But, if $n = 10^k$, then by the definition $k = \log_{10} n$. It is natural to assume that also for other values of n the ratio $\frac{n}{\pi(n)}$ is very close to $\alpha \log_{10} n$. Thus, $\pi(n)$ is very close to $c\,\dfrac{n}{\log_{10} n}$, where $c = \alpha^{-1}$.

A lot of mathematicians where attracted by the secret of distribution of primes and tried to solve it using tables. In particular, Gauss got interested in this question almost as a child. His interest in mathematics started, it seems, from the child's interest in numbers and forming tables. In general, a lot of great mathematicians showed virtuosity in calculations and were capable of doing immense ones, often by heart (Euler was even fighting insomnia in that way!). When Gauss was only 14, he constructed a table of primes (in fact, a smaller one than our Table 1) and came to the conclusion we have formulated. Later on this conclusion was considered by many mathematicians. But the first result in that direction was proved only half a century later, in 1850, by Chebyshev.

Chebyshev proved the following assertion.

Theorem. *There exist constants c and C such that for all $n > 1$*

$$c\,\frac{n}{\log n} \leqslant \pi(n) \leqslant C\,\frac{n}{\log n}. \tag{10}$$

Before we proceed with the proof, we give some remarks concerning the formulation of the theorem. What is the basis of the logarithms we are using? The answer is: arbitrary. From the definition of logarithms it immediately follows that $\log_b x = \log_b a \log_a x$: it is enough to substitute a by $b^{\log_b a}$ in the relation

$a^{\log_a x} = x$, to obtain $b^{\log_b a \log_a x} = x$, which shows that $\log_b x = \log_b a \log_a x$. Hence, if the inequality (10) is proved for $\log_a n$, then it is true also for $\log_b n$, with the substitution of c and C by $\frac{c}{\log_b a}$ and $\frac{C}{\log_b a}$.

Inequalities (10) express the idea inspired by tables that $\pi(n)$ is "close" to $c\,\frac{n}{\log n}$ for some c. The question why in our hypothetical reasoning there appeared one constant c, and in the theorem there appear two of them—c and C—and whether it is possible to use only one constant, will be discussed after the proof of the theorem.

The key to the proof of Chebyshev's theorem are properties of binomial coefficients C_n^k: mostly the fact that they are integers and some properties about their divisibility by primes. We shall recall these properties before we pass to the proof.

First of all, there is a proposition proved in section 3 of Chapter II which says that the sum of all binomial coefficients C_n^k for $k = 0, 1, \ldots, n$ is equal to 2^n. Since the sum of positive summands is greater than any of them, we deduce that

$$C_n^k \leqslant 2^n. \tag{11}$$

We shall particularly need large binomial coefficients. We saw in Chapter II that for even $n = 2m$ the coefficient C_{2m}^m is greater than all the others, and for odd $n = 2m+1$ there exist two equal coefficients C_{2m+1}^m and C_{2m+1}^{m+1} which are greater than the others. We draw our attention to them, particularly to

$$C_{2n}^n = \frac{2n(2n-1)\cdots(n+1)}{1\cdot 2\cdot\ldots\cdot n}. \tag{12}$$

If we group the factors of the numerator with the factors of the denominator taken in reverse order, we obtain

$$C_{2n}^n = \frac{2n}{n}\cdot\frac{2n-1}{n-1}\cdot\ldots\cdot\frac{n+1}{1}.$$

Obviously, no factor in the last formula is less than 2, so

$$C_{2n}^n \geqslant 2^n. \tag{13}$$

Consider now properties of divisibility of binomial coefficients by primes. Factors in the numerator in the expression (12) are obviously divisible by all primes greater than n and not exceeding $2n$. These primes cannot divide any factor in the denominator, and so they do not cancel and they are divisors of C_{2n}^n. The number of primes between n and $2n$ is equal to $\pi(2n) - \pi(n)$ and all of them are greater than n, hence

$$C_{2n}^n \geqslant n^{\pi(2n)-\pi(n)}. \tag{14}$$

An analogous assertion is valid for the "middle" coefficients $C_{2n+1}^n = C_{2n+1}^{n+1}$ with an odd lower index. If we write them as

$$C_{2n+1}^n = \frac{(2n+1)\cdots(n+2)}{1\cdot 2\cdot\ldots\cdot n},$$

we see that $\pi(2n+1)-\pi(n+1)$ of primes, greater than $n+1$ and not exceeding $2n+1$, enters the numerator and cannot be cancelled with the denominator. Since they are greater than $n+1$, we have

$$C_{2n+1}^{n} > (n+1)^{\pi(2n+1)-\pi(n+1)}. \tag{15}$$

The inequalities (14) and (15) already reveal important connections between binomial coefficients and prime numbers.

Finally, we state the last of the properties of binomial coefficients which we need for the proof; although it is quite easy, it is not as obvious as the previous ones.

Lemma. *For an arbitrary binomial coefficient C_n^k, any power of a prime dividing it does not exceed n.*

We draw the attention to the fact that we are not speaking about the *exponent* of the power but about the *power itself.* In other words, we assert that if p^r divides C_n^k, where p is a prime, then $p^r \leqslant n$. For example, $C_9^2 = 9 \cdot 4$ is divisible by 9 and by 4, and both of these numbers do not exceed 9.

Write the binomial coefficient in the form

$$C_n^k = \frac{n(n-1)\dots(n-k+1)}{1\cdot 2\cdot \ldots \cdot k}. \tag{16}$$

The prime p we are dealing with has to divide the numerator of this fraction. Denote by m the factor in numerator which contains the maximal power of p (or one of those having such a property), and by p^r this maximal power. Obviously, $n \geqslant m \geqslant n-k+1$. Set $n-m=a$, $m-(n-k+1)=b$, then $a+b=k-1$ and C_n^k can be written in the form

$$C_n^k = \frac{(m+a)(m+a-1)\cdots(m+1)m(m-1)\cdots(m-b)}{k!}. \tag{17}$$

The factor m is now the most important for us and we write down the product in the numerator as having a factors to the left and b factors to the right of it. Rearrange the denominator analogously: $k! = (1\cdot 2\cdot \ldots \cdot a)(a+1)\cdots(a+b)(a+b+1)$. Since $(a+1)(a+2)\cdots(a+b)$ is divisible by $b!$, this product (denominator) has the form $a!\,b!\,l$, where l is an integer. Now we can rewrite C_n^k in the following form

$$C_n^k = \frac{m+a}{a}\cdot\frac{m+a-1}{a-1}\cdot\ldots\cdot\frac{m+1}{1}\cdot\frac{m-1}{1}\cdot\ldots\cdot\frac{m-b}{b}\cdot\frac{m}{l}, \tag{18}$$

where we transferred the factor $\frac{m}{l}$ to the end.

Note that in each of the factors $\frac{m+i}{i}$ or $\frac{m-j}{j}$ ($i=1,\dots,a$, $j=1,\dots,b$) the power of p entering the numerator completely cancels with the denominator, and hence after cancellation only the denominator could be divisible by p (though it can also be relatively prime with p). Really, consider, for instance, fractions $\frac{m+i}{i}$ (factors of the type $\frac{m-j}{j}$ can be treated in the same way). Let i be divisible exactly by p^s, i.e., $i = p^s u$, where u is relatively prime with p. If $s<r$, then $m+i$ is

also divisible exactly by p^s: setting $m = p^r v$ (recall that m is divisible by p^r), we obtain that $m + i = p^s(u + p^{r-s}v)$. If $s \geqslant r$, then for the same reasons $m + i$ is divisible by p^r and taking into account the way m was chosen (it is divisible by the greatest power of p of all the numbers between $n - k + 1$ and n and this power is p^r), we conclude that the number $m + i$ cannot be divisible by a greater power of p than the r-th. Thus, p^r cancels and in the numerator there remains a number not divisible by p. As a result we see that among all the factors in the expression (18), only the last one can contain p as a factor. But the power of p which divides m is p^r, and this means the product (18) cannot be divisible by a greater power of p than p^r. Since p^r divides m, and $m \leqslant n$, it is $p^r \leqslant n$. The Lemma is proved.

Let us see what it says about the canonical factorisation $C_n^k = p_1^{\alpha_1} \cdots p_m^{\alpha_m}$. First of all, the primes $p_1, \ldots, p_m$ can appear just from the numerator of the expression (16), therefore all $p_i \leqslant n$ and so $m \leqslant \pi(n)$. According to the Lemma, $p_i^{\alpha_i} \leqslant n$ for $i = 1, \ldots, m$. As a result we obtain that

$$C_n^k \leqslant n^{\pi(n)}. \tag{19}$$

Now we can proceed with the proof of the Chebyshev's theorem itself, i.e., with the proof of the inequalities (10). Note that it is enough to prove these inequalities for all values of n starting with a certain fixed limit n_0. For all $n < n_0$ these inequalities can then be obtained by decreasing the constant c and increasing the constant C. If we wanted to obtain values of these constants explicitly and in the most economic way, then we could check, using tables of primes, that the inequalities (10) are valid for values $n \leqslant n_0$ (in our arguments n_0 will not be a large number).

We start with juxtaposition of inequalities (13) and (19) for the binomial coefficient C_{2n}^n. We obtain that $2^n \leqslant C_{2n}^n \leqslant (2n)^{\pi(2n)}$ and hence

$$2^n \leqslant (2n)^{\pi(2n)}. \tag{20}$$

Taking logarithms with basis 2 of both sides (recall that we will write $\log_2 x = \log x$) and using monotonicity of the logarithm, we obtain that $n \leqslant \pi(2n) \log 2n$ and so

$$\pi(2n) \geqslant \frac{n}{\log 2n} = \frac{1}{2} \frac{2n}{\log 2n},$$

i.e., the left of the two inequalities (10) with the constant $c = \frac{1}{2}$. But for the time being it is proved only for even values of n. For odd values of the form $2n + 1$ we use the monotonicity of the logarithm and of the function $\pi(n)$. It follows that

$$\pi(2n+1)\log(2n+1) \geqslant \pi(2n)\log 2n.$$

Substituting here the obtained inequality for $\pi(2n)$, we see that

$$\pi(2n+1) \geqslant \frac{n \log 2n}{\log(2n)\log(2n+1)} = \frac{n}{\log(2n+1)}.$$

Since it is always $n \geqslant \frac{1}{3}(2n+1)$, it follows that

$$\pi(2n+1) \geqslant \frac{1}{3}\frac{2n+1}{\log(2n+1)}.$$

Thus, the left inequality (10) is proved for odd n with the constant $c = \frac{1}{3}$. So, the left inequality (10) is valid for all n and $c = \frac{1}{3}$.

We proceed to the proof of the right inequality in (10). We shall prove it by induction on n. Let, first of all, n be even. We will write $2n$ instead of it. Taking the inequality (11) for the coefficient C_{2n}^{n} (i.e., substitute in C_n^k n by $2n$ and k by n) together with the inequality (14), as a consequence we obtain

$$n^{\pi(2n)-\pi(n)} \leqslant 2^{2n}$$

and, passing to logarithms,

$$\pi(2n) - \pi(n) \leqslant \frac{2n}{\log n}, \qquad \pi(2n) \leqslant \pi(n) + \frac{2n}{\log n}. \tag{21}$$

In accordance with the inductive hypothesis, suppose that our inequality has been proved: $\pi(n) \leqslant C\frac{n}{\log n}$ with a constant C whose value we shall make more precise later. Substituting in the formula (21), we obtain:

$$\pi(2n) \leqslant C\,\frac{n}{\log n} + \frac{2n}{\log n} = \frac{(C+2)n}{\log n}.$$

We would like to prove the inequality $\pi(2n) \leqslant \frac{C\cdot 2n}{\log 2n}$ and for that we have to choose the constant C in such a way that the inequality

$$\frac{(C+2)n}{\log n} \leqslant \frac{2Cn}{\log 2n} \tag{22}$$

is valid for all n, starting from some limit.

This is just a simple school exercise. Cancel in the inequality both sides by n, remark that $\log 2n = \log 2 + \log n = \log n + 1$ and denote $\log n$ by x. Then the inequality (22) takes the form

$$\frac{C+2}{x} \leqslant \frac{2C}{x+1}.$$

Multiplying both sides by $x(x+1)$ (as $x > 0$) and transforming, we write it in the form $(C-2)x \geqslant C+2$. Obviously, C has to be chosen so that $C - 2 > 0$. Setting, e.g., $C = 3$, we obtain that it is valid for $C = 3$ and all $x \geqslant 5$. Since x denotes $\log n$, this means that the necessary inequality would be valid if $n \geqslant 2^5 = 32$, $2n \geqslant 64$.

It remains to consider the case of odd values of the form $2n+1$. Compare the inequality (11) (substituting in it n by $2n+1$ and k by n) with the inequality (15). We obtain the inequality

$$2^{2n+1} \geqslant (n+1)^{\pi(2n+1)-\pi(n+1)}$$

and, taking logarithms, the inequality

$$2n+1 \geqslant (\pi(2n+1) - \pi(n+1))\log(n+1).$$

From here, using the inductive hypothesis about $\pi(n+1)$, we obtain, as before

$$\pi(2n+1) \leqslant C\,\frac{n+1}{\log(n+1)} + \frac{2n+1}{\log(n+1)}.$$

The inequality that we need: $\pi(2n+1) \leqslant C\frac{2n+1}{\log(2n+1)}$ will be proved if we can check that

$$(23) \qquad C\,\frac{n+1}{\log(n+1)} + \frac{2n+1}{\log(n+1)} \leqslant C\,\frac{2n+1}{\log(2n+1)}$$

for a suitable choice of the constant C and for all n starting from some limit. This is again an exercise of purely school type, though a bit harder than the previous one. In order to compare various terms in the inequality more easily, replace on the left-hand side $2n+1$ by a greater value $2(n+1)$:

$$(24) \qquad C\,\frac{n+1}{\log(n+1)} + \frac{2n+1}{\log(n+1)} \leqslant \frac{(C+2)(n+1)}{\log(n+1)}.$$

In order to transform the right-hand side, note that $2n+1 \geqslant \frac{3}{2}(n+1)$ for $n \geqslant 1$, $\log(2n+1) \leqslant \log(2n+2) = \log(n+1)+1$. Hence,

$$(25) \qquad \frac{2n+1}{\log(2n+1)} \geqslant \frac{(3/2)(n+1)}{\log(n+1)+1}.$$

Comparing inequalities (24) and (25) we see that the inequality (23) will be proved if we prove that

$$\frac{(C+2)(n+1)}{\log(n+1)} \leqslant \frac{(3/2)C(n+1)}{\log(n+1)+1}.$$

Cancelling both sides by $n+1$ and putting $\log(n+1) = x$, we arrive at the inequality

$$\frac{C+2}{x} \leqslant \frac{(3/2)C}{x+1},$$

which can be solved completely in the same way as in the previous case. It is enough to multiply both sides by $x(x+1)$ and reduce similar terms. We obtain the inequality $(C+2)x + C + 2 \leqslant \frac{3}{2}Cx$, i.e., $(\frac{1}{2}C - 2)x \geqslant C+2$. Setting $C = 6$, we see that the inequality is valid for $x \geqslant 8$, i.e., for $n+1 \geqslant 2^8$, $2n+1 \geqslant 511$. Thus, the right inequality (10) is proved with the constant $C = 6$ and for all values of n starting with 511. The Theorem is proved.

Note that Theorem 4 appears as an easy consequence of the Theorem just proved. Really, since $\pi(n) < C\frac{n}{\log n}$, we have $\frac{\pi(n)}{n} \leqslant \frac{C}{\log n}$. And as a logarithm changes monotonously and increases unboundedly ($\log 2^k = k$), $\frac{\pi(n)}{n}$ becomes less than any arbitrary positive number. But, the proof of the theorem of Chebyshev was based on completely different considerations than the proof of Theorem 4.

At the end, we return once more to assertions which can be made by considering Table I. Starting from it, we came to the claim that $\frac{n}{\pi(n)}$ is close to $\log_{10} n$ with a certain value of the constant C: the first decimal figures of the number C^{-1} are

2.3. Hence it can be concluded that $\pi(n)$ is close to $C^{-1}\frac{n}{\log_{10} n}$. This expression can be given a simpler form $\frac{n}{\log_e n}$, if we use a new basis of logarithms e such that $C\log_{10} n = \log_e n$. But, as it was said earlier, it is always $\log_b x = \log_b a \cdot \log_a x$, and so our relation will be fulfilled if $C = \log_e 10$. Substituting the value $x = b$ into the relation $\log_b x = \log_b a \cdot \log_a x$, we obtain that $\log_b a \cdot \log_a b = 1$ and the relation $C = \log_e 10$ which we are interested in can be rewritten as $C^{-1} = \log_{10} e$.

14-year-old Gauss turned his attention to these relations and tried to guess which number e could be, so that $\log_{10} e$ is close to $(2.3)^{-1}$. Such a number at that time was well known, thanks to the fact that the logarithm with such a basis has a lot of useful properties. This number is commonly denoted by e. The logarithm with the basis e is called *natural* and is denoted by ln: $\log_e x = \ln x$. Here, to the end of this page, we have to consider that the reader is familiar with the concept of the natural logarithm.

In such a way, a natural assertion which can be deduced from tables is that $\pi(n)$ becomes close to $\frac{n}{\ln n}$ when n increases unboundedly. The theorem of Chebyshev which has been just proved states (if we use natural logarithms) that there exist two such constants c and C, that $c\frac{n}{\ln n} < \pi(n) < C\frac{n}{\ln n}$, starting from some n. The hypothetical refinement deduced from tables asserts that the inequalities $c\frac{n}{\ln n} < \pi(n) < C\frac{n}{\ln n}$ are valid for n large enough *whichever constants $c < 1$ and $C > 1$ we take.* This assertion is called *the asymptotical law of distribution of primes.* It was stated by Gauss and some other mathematicians at the end of XVIII and the beginning of XIX century. After the proof of the inequalities of Chebyshev in 1850 it seemed that all that was needed was a better choice and approaching of the constants c and C. However, the asymptotical law of distribution of primes was proved just half a century later, at the end of XIX century, using completely new ideas, proposed by Riemann.

Problems

1. Prove that $p_n > an\log n$ for a certain constant $a > 0$ [*Hint.* Use the fact that $\pi(p_n) = n$.]

2. Prove that $\log n < \sqrt{n}$, starting from some limit (find it). [*Hint.* Reduce the problem to proving the inequality $2^x > x^2$ for real x, starting from some limit. Let $n \leqslant x \leqslant n+1$, where n is an integer. Reduce to proving the inequality $2^n \geqslant (n+1)^2$ and use the induction.]

3. Prove that $p_n < Cn^2$ for some constant C. [*Hint.* Apply the inequality of the previous problem and use the fact that $n = \pi(p_n)$.]

4. Prove that $p_n < An\log n$ for some constant A.

5. Prove that that the largest exponent a for which p^a divides $n!$ is equal to $\left[\frac{n}{p}\right] + \left[\frac{n}{p^2}\right] + \dots + \left[\frac{n}{p^k}\right]$. Here $\left[\frac{r}{s}\right]$ is the incomplete quotient of dividing r by s, the sum extends to all k for which $p^k \leqslant n$, p denotes an arbitrary prime and n an arbitrary natural number.

6. Using the result of Problem 5 give a new proof of the Lemma in the Appendix.

7. Prove that if $p_1, \ldots, p_r$ are all the primes between m and $2m+1$, then their product does not exceed 2^{2m}.

8. Determine the constants c and C such that the inequality (10) is valid for all n.

9. Try to find as large as possible a constant c and as small as possible a constant C, for which the inequality (10) is valid for all n, starting from some limit.

I. R. Shafarevich,
Russian Academy of Sciences,
Moscow, Russia