CHAPTER III

PRIME NUMBERS

§ 1. The primes. Factorization of a natural number m into primes. Any number >1 which has no natural divisors except itself and 1 is called a *prime number*, or simply a *prime*. *A necessary and sufficient condition for a natural number $m>1$ to be a prime is that m should not be the product of two natural numbers less than m.* In fact, if m is a prime, then it cannot be the product $a\cdot b$ of two natural numbers less than m, since, if it could, the numbers a and b would be greater than 1 and therefore the number m would have a divisor greater than 1 and less than m, which would contradict the assumption that m is a prime. This proves the necessity of the condition. On the other hand, if the number m is not a prime, then it has a divisor a such that $1<a<m$ and hence $m=a\cdot b$, where b must be a natural number less than m, since $a>1$. Thus the number m is the product of two natural numbers, each of them less than m. Thus the sufficiency of the condition is proved.

Thus the definition itself provides a method by means of which one can decide whether a given natural number $n>1$ is a prime or not. In fact, it suffices to divide the number n by the numbers $2,3,\ldots,n-1$ successively and see whether any of these numbers divides the number n; if none of them does, then (and only then) the number n is a prime.

A natural number which is neither 1 nor a prime is said to be *composite*. Such a number is representable as the product of two positive integers each less than the number in question. Consequently if n is a composite number, $n=ab$, where a and b are natural numbers each less than n; it follows that each of the numbers a, b is greater than 1. Interchanging, if necessary, the rôles of a and b, we may assume that $a\leqslant b$, whence $a^2\leqslant ab=n$, and consequently $a\leqslant\sqrt{n}$. Hence we have

THEOREM 1. *If a natural number n is composite, then it has a divisor a such that $1<a\leqslant\sqrt{n}$.*

It follows that in order to decide whether a natural number $n>1$ is a prime it suffices to divide it by numbers greater than 1 and not greater than $\sqrt{n}$, successively.

We now prove

THEOREM 2. *Every natural number >1 has at least one prime divisor.*

Proof. Let n be a natural number >1. Obviously the number n has some divisors greater than 1, since the number n itself is such a divisor. Denote by p the least of them. If p were not a prime, then we would have $p=ab$, where a,b would be natural numbers greater than 1 and less than p. Thus the number a would be a divisor of n at the same time greater than 1 and less than p, contrary to the definition of p. Therefore p is a prime and this completes the proof of theorem 2.

As an immediate consequence of theorems 1 and 2 we have

COROLLARY 1. *Every composite number n has at least one prime divisor $\leqslant\sqrt{n}$.*

COROLLARY 2. *Every natural number >1 is the product of a finite number of prime factors.* (Clearly, trivial products of one factor are not excluded).

Proof. Suppose to the contrary that corollary 2 is untrue. Then there exists a least natural number $n>1$ which is not the product of prime numbers. In virtue of theorem 2 number n has a prime divisor p and $n=pn_1$, where n_1 is a natural number. We cannot have $n_1=1$; for, in that case we would have $n=p$ and the corollary 2 would be true.

Therefore $n_1>1$ and $n=pn_1>n_1$. Hence $n_1<n$, and from the definition of number n we infer that n_1 is the product of prime numbers. Then, however, the number $n=pn_1$ is also the product of prime numbers, contrary to the definition of number n. Thus the assumption that corollary 2 is untrue results in a contradiction. Corollary 2 is thus proved.

A question arises whether there exists a method which would enable us to represent a given natural number as a product of prime numbers. We show that, although the calculations involved may be very long, such a method does exist. It is sufficient to prove that for a given natural number one can either find the required factorization for the number n itself or reduce the problem to finding such a factorization of a number less than n.

If n is a natural number >1, then, dividing it by $2,3,\ldots,n$ successively, we find its least divisor which, as we know, is a prime p. We then have $n=pn_1$, where n_1 is a natural number. If $n_1=1$, then $n=p$ and the desired representation is completed. If $n_1>1$, then in order to find the representation of n it suffices to find the representation for the number n_1, less than n. Continuing, we proceed similarly with n_1 in place of n. It is clear that after a finite number of steps less than n we ultimately obtain the representation of number n as a product

$$n=pp'p''\ldots p^{(k-1)}$$

of prime factors. If in this product some identical factors occur, then replacing them by the powers of suitable prime numbers we can rewrite the representation in the form

$$n = q_1^{a_1} q_2^{a_2} \dots q_s^{a_s}, \tag{1}$$

where $q_1, q_2, \dots, q_n$ are all different prime numbers, i.e. for instance, $q_1 < q_2 < \dots < q_s$ and a_i $(i = 1, 2, \dots, s)$ are natural numbers. Such representation of a natural number n is called the *factorization of n into prime numbers*.

In factorization (1) of number n the numbers $q_1, q_2, \dots, q_s$ are all the prime divisors of the number n. In fact, if the number n were divisible by a prime number q different from the numbers $q_1, q_2, \dots, q_s$, then, for $i = 1, 2, \dots, s$, we would have $(q, q_i) = 1$, since the prime number q has only two divisors, q and 1, and $q \neq q_i$. Therefore any two different prime numbers are relatively prime. We would also have $(q, q_i^{a_i}) = 1$ for $i = 1, 2, \dots, s$, whence, in virtue of (1) and theorem 6^a of Chapter I, $(q, n) = 1$, contrary to the assumption that n is divisible by q.

We see that the numbers q_i $(i = 1, 2, \dots, s)$, as well as the number of them, are uniquely determined by number n (as the prime divisors of n). Moreover, also the exponents $a_1, a_2, \dots, a_s$ are uniquely determined by n. In particular, the number a_1 can be defined as the greatest natural number for which $q_1^{a_1} \mid n$, since in the case $q_1^{a_1+1} \mid n$ we would have $q_1 \mid q_2^{a_2} \dots q_s^{a_s} \mid n$, which is impossible. Therefore, since we have assumed that $q_1, q_2, \dots, q_s$ is an increasing sequence, factorization (1) is unique.

This leads us to the following

THEOREM 3. *Any natural number can be represented in one and only one way as a product of primes. (Clearly enough two factorizations are regarded as being identical if they differ in the order of the factors).*

As has been proved above, for every natural number $n > 1$ we are able to find the factorization into primes effectively provided we are not daunted by long calculations, which may possibly occur.

In some cases these are too long to be carried out even with the aid of the newest technical equipment. For instance this happens in the case of the number $2^{101}-1$, which has 31 digits. (We know that this number is composite.) We do not know any of its prime divisors, although we do know that the least of them has at least 8 digits. We do not know any of the prime divisors of the number $F_{19} = 2^{2^{19}}+1$, either. It is not known whether this number is a prime or not. We know a prime divisor of the number F_{1945}, namely $5 \cdot 2^{1947}+1$, though we do not know any other of its prime divisors, which, as we know, do exist.

An example of a number which can easily be proved to be composite but none of whose prime divisors are known is the number F_{17}^2.

THEOREM 4. *If a natural number n is greater than 2, then between n and $n!$ there is at least one prime number.*

Proof. Since $n > 2$, the number $N = n!-1$ is greater than 1, whence, in virtue of theorem 2, it has a prime divisor, p. Number p cannot be less than or equal to n, since, if it could, it would divide 1, which is impossible. Consequently $p > n$. On the other hand, $p \leqslant N$, p as a divisor of N. Thus we conclude that $n < p \leqslant n!-1 < n!$, which completes the proof.

It follows that for each natural number n there exists a prime number greater than n; therefore there are infinitely many prime numbers. In particular, there exist prime numbers having at least ten thousand digits, but we do not know any one of them. The greatest prime number which is known so far is the number $2^{11213}-1$; it has 3376 digits. The proof that it is a prime number was carried out in 1963.

EXERCISES. **1.** Given a prime each of whose digits (in the decimal expansion) equals 1, prove that the number of the digits must be prime. (The converse implication is not true).

Proof. Let n be such a number having s digits in the decimal expansion, each equal to 1; suppose that s is a composite number, i.e. $s = ab$, where a, b are natural numbers, each greater than 1.

We then have $n = \frac{10^s-1}{9} = \frac{10^{ab}-1}{9}$. But $10^a-1 \mid 10^{ab}-1$, whence $\frac{10^a-1}{9} \Big| n$. $\frac{10^a-1}{9}$ is a natural number > 1, since $a > 1$. Since $b > 1$, we have $\frac{10^a-1}{9} < \frac{10^{ab}-1}{9} = n$.

From this we conclude that number n has a divisor $\frac{10^a-1}{9}$, less than n and greater than 1, which is impossible. This completes the proof.

To see that the converse implication does not hold we note, for example, that $111 = 3 \cdot 37$ and $11111 = 41 \cdot 271$. We do not know whether the sequence 11, 111, 1111, ... contains infinitely many terms which are prime numbers. M. Kraïtchik [2] (Chapter III) has proved that number $(10^{23}-1)/9$ is a prime, and D. H. Lehmer [1] has proved that number $(10^{37}-1)/9$ is composite.

2. Prove that there exist infinitely many natural numbers which are not of the form a^2+p, where a is an integer and p a prime.

Proof. Such are for instance the numbers $(3n+2)^2$, where $n = 1, 2, \dots$ Suppose, to the contrary, that for a natural number n we have $(3n+2)^2 = a^2+p$, where a is an integer and p a prime. Then, plainly, a cannot equal 0; consequently, we may assume that a is a natural number. Then $3n+2 > a$, so $3n+2-a > 0$. But $p = (3n+2-a)(3n+2+a)$, whence $3n+2-a = 1$ and $3n+2+a = p$, which implies that $p = 6n+3 = 3(2n+1)$, which is impossible.

Remark. It can be proved that for every natural number k there are infinitely many k-powers of natural numbers which are not of the form a^k+p, where a is an integer and p a prime. (cf. Clement [2]).

As verified by Euler, each odd natural number n, with $1 < n < 2500$, is of the form $n = 2a^2+p$, where a is an integer and p a prime. This is not true for n equal to 5777 and 5993, cf. Dickson [8], Vol. I, p. 424. I do not know whether there exist

infinitely many odd natural numbers that are not of the form $2a^2+p$, where a is an integer and p a prime.

3. Prove that every number of the form 8^n+1 is composite.

Proof. For each natural number n we have $2^n+1 \mid 2^{3n}+1 = 8^n+1$ and, clearly, $1 < 2^n+1 < 8^n+1$. This proves that the number 8^n+1 is composite.

Remark. We do not know whether there are infinitely many prime numbers of the form 10^n+1 ($n = 1, 2, \ldots$), or whether every number of the form 12^n+1 is composite ($n > 1$).

§ 2. The Eratosthenes sieve. Tables of prime numbers. It is an immediate consequence of corollary 1 of § 1 that, if a natural number $n > 1$ is not divisible by any prime number $\leqslant \sqrt{n}$, then n is a prime number.

It follows that in order to obtain all the prime numbers which occur in the sequence $2, 3, 4, \ldots, m$, where m is a given natural number, it suffices to remove all the multiples kp of the prime numbers $p \leqslant \sqrt{m}$ with $k > 1$ from the sequence. Thus, in particular, to obtain all the primes occurring in the sequence $2, 3, \ldots, 100$ it is sufficient to remove from the sequence all the numbers greater than $2, 3, 5$ and 7 and divisible by at least one of these numbers.

An easy method of finding consecutive prime numbers was given by a Greek mathematician Eratosthenes. We consider the sequence $2, 3, 4, \ldots$ Then, since 2 is the first prime number p_1, we remove from the sequence all the numbers greater than p_1 and divisible by 2. The first of the remaining numbers is $3 = p_2$. We now remove all the numbers greater than p_2 and divisible by p_2. The first of the remaining numbers is $5 = p_3$. Suppose that after the nth step we have found the nth prime number p_n. We remove from the sequence all the numbers greater than p_n and divisible by p_n. The least number which has not yet been removed is the $n+1$-th prime number.

If the sequence of the natural numbers from 2 onwards is replaced by the sequence of natural numbers $2, 3, \ldots, N$, the above procedure terminates after the kth step, where p_k is the greatest prime number $\leqslant \sqrt{N}$.

Thus we obtain $p_1 = 2$, $p_2 = 3$, $p_3 = 5$, $p_4 = 7$, $p_5 = 11$, $p_6 = 13$, $p_7 = 17$, $p_8 = 19$, $p_9 = 23$, $p_{10} = 29$, $p_{25} = 97$, $p_{100} = 541$, $p_{200} = 1223$, $p_{1000} = 7917$, $p_{1229} = 9973$, $p_{1230} = 10007$. It has recently been computed that $p_{6000000} = 104395301$ (cf. Editorial Note [1]). D. Blanuša [1] has found the following simple geometric interpretation of the Eratosthenes sieve. In the Cartesian system of coordinates the set A of points $\left(0, \frac{1}{m}\right)$, $m = 1, 2, \ldots$, and the set B of points $(n+1, 0)$, $n = 1, 2, \ldots$, are considered. Each point of the set A is connected with each point of the set B by a straight line. Then the set of the abscissae of the intersections of the straight lines with the straight line $y = 1$ is precisely the set of composite numbers.

In fact, the equation of the line joining points $\left(0, \frac{1}{m}\right)$ and $(n+1, 0)$ is $x/(n+1)+my = 1$. This line intersects the line $y = -1$ at the point whose abscissa is $x = (m+1)(n+1)$. But, since m and n are natural numbers, x is a composite number. Conversely, if x is a composite number, then $x = (m+1)(n+1)$, where m, n are natural numbers, and consequently it is the abscissa of the intersection of the line joining the point $\left(0, \frac{1}{m}\right)$ and the point $(n+1, 0)$ with the line $y = -1$.

There exist printed tables of the prime numbers less than eleven millions. Cf. D. N. Lehmer [1]. In that table for each natural number not greater than 10170000 the least prime divisor greater than 2, 3, 5, 7 is given. Cf. also Kulik, Poletti, Porter [1], where the primes of the eleventh million are listed.

Jacob Philip Kulik, a mathematician of Polish origin (born in 1793 in Lwów, died in 1863 in Prague), prepared a manuscript (to the writing of which he devoted 20 years of his life) under the title *Magnus Canon Divisorum pro omnibus numeris par* 2, 3, 5 *non divisilibus et numerorum primorum interjacentium ad Millies centum, millia accuratius ad* 100330201 *usque. Authore Jacobo Philippo Kulik Galiciano Leopoliensis Universitate Pragensi Matheseos sublimioris Prof. publ. ac ord.* At present the manuscript is owned by the Vienna Academy of Sciences. This manuscript was used when the table for prime numbers less than twelve millions were being prepared. (Some mistakes in it were then corrected.)

An article about J. P. Kulik and his work together with his portrait has recently been published by I. Ya. Depman [1]. For the history of tables of prime numbers, see ibid. pp. 594-601.

In 1959 C. L. Baker and F. J. Gruenberger made microcards containing all the prime numbers less than 104395301, cf. Baker and Gruenberger [1].

§ 3. The differences between consecutive prime numbers. As in the preceding section let p_n denote the nth prime number and let $d_n = p_{n+1} - p_n$ for $n = 1, 2, \ldots$ The first hundred of the terms of the infinite sequence $d_1, d_2, \ldots$ are the following:

1,	2,	2,	4,	2,	4,	2,	4,	6,	2
6,	4,	2,	4,	6,	6,	2,	6,	4,	2
6,	4,	6,	8,	4,	2,	4,	2,	4,	14
4,	6,	2,	10,	2,	6,	6,	4,	6,	6
2,	10,	2,	4,	2,	12,	12,	4,	2,	4
6,	2,	10,	6,	6,	6,	2,	6,	4,	2
10,	14,	4,	2,	4,	14,	6,	10,	2,	4
6,	8,	6,	6,	4,	6,	8,	4,	8,	10
2,	10,	2,	6,	4,	6,	8,	4,	2,	4
12,	8,	4,	8,	4,	6,	12,	2,	18,	6

Number 2 is the only even number which is a prime (since even numbers greater than 2 are divisible by 2, they are composite). Thus numbers p_n for $n>1$ are odd and, consequently, the numbers $d_n = p_{n+1} - p_n$ are even.

Looking at the table presented above (p. 115), one can raise the question whether for each natural number k there exists at least one number n for which $d_n = 2k$? We do not know the answer to this question.

We present here the table of the least natural numbers n for which $d_n = 2k$ with $2k \leqslant 30$ together with the prime numbers p_n, p_{n+1} such that $p_{n+1} - p_n = 2k$.

$2k$	n	p_n	p_{n+1}	$2k$	n	p_n	p_{n+1}	$2k$	n	p_n	p_{n+1}
2	2	3	5	12	46	199	211	22	189	1129	1151
4	4	7	11	14	30	113	127	24	263	1669	1693
6	9	23	29	16	282	1831	1847	26	367	2477	2503
8	24	89	97	18	99	523	541	28	429	2971	2999
10	34	139	149	20	154	887	907	30	590	4297	4327

(Cf. D. H. Lehmer [10].)

It has been found that the least consecutive prime numbers whose difference is 100 are the numbers 396733 and 396833. The table of the numbers d_n with $n < 600$ has been given by P. Erdös and A. Rényi [1] (1). The table of d_n with $n \leqslant 1233$ has been given by M. Colombo [1].

The table of the least numbers p_n for which $p_{n+1} - p_n = 2k$ with $2k \leqslant 156$ has been given by D. H. Lehmer [10].

Over a hundred years ago the conjecture was raised that for every even number $2k$ there exist infinitely many natural numbers n such that $d_n = 2k$ (de Polignac [1]). For $k = 2$ this conjecture is equivalent to the conjecture that there exist infinitely many pairs of twin primes, i.e. pairs of consecutive odd numbers n each of which is a prime. The first ten such pairs are 3 and 5, 5 and 7, 11 and 13, 17 and 19, 29 and 31, 41 and 43, 59 and 61, 71 and 73, 101 and 103, 107 and 109. H. Tietze has given a table of twin primes less than 300000 presenting the greater number of each pair. They are 2994 in number (Cf. Tietze [1] and also Früchtl [1]. See also Selmer and Nesheim [1], where the numbers n are

(1) There are some mistakes in the table:

instead of $d_{256} = 12$ it should be $d_{256} = 2$,

instead of $d_{314} = 6$ it should be $d_{314} = 4$,

instead of $d_{429} = 18$ it should be $d_{429} = 28$,

instead of $d_{462} = 18$ it should be $d_{462} = 28$.

It should be also $d_{579} = 2$.

given for which $6n+1$ and $6n-1$ are both prime and less than 200000. Compare also Sexton [1] and [2].) D. H. and E. Lehmer [1] have found that there are 152892 pairs of twin primes less than 30000000. The greatest of the known pairs of twin primes is the pair $10^{12}+9649$ and $10^{12}+9651$.

It can be proved that the problem whether there exist infinitely many pairs of twin primes is equivalent to the question whether there exist infinitely many natural numbers n for which n^2-1 has exactly 4 natural divisors.

We note here that in order to obtain from the sequence of consecutive integers $1, 2, \ldots, n$ the prime numbers p for which also $p+2$ is prime one has to remove for each composite number k the number $k-2$ provided all the composite numbers have already been removed (for instance by means of the Erathostenes sieve) from this sequence (cf. Golomb [1]).

W. A. Golubew has asked whether for a natural number n there is at least one pair of twin primes between n^3 and $(n+1)^3$.

It has been proved that the series of the reciprocals of the prime numbers of the pairs of twin primes is finite or convergent (Brun [1]) (1).

The sum of the series

$$(\tfrac{1}{3}+\tfrac{1}{5})+(\tfrac{1}{5}+\tfrac{1}{7})+(\tfrac{1}{11}+\tfrac{1}{13})+(\tfrac{1}{17}+\tfrac{1}{19})+(\tfrac{1}{29}+\tfrac{1}{31})+\ldots$$

has been calculated with an accuracy to three decimal places by E. S. Selmer [1]. In § 14 we shall see that the series of the reciprocals of all the prime numbers is divergent.

Another question to which the answer is not known is whether there exist infinitely many primes p for which p, $p+2$, $p+6$ and $p+8$ are all prime numbers. A quadruple of the primes of this type is called simply a *quadruplet*. The first six consecutive quadruplets are obtained for $p = 5, 11, 101, 191, 821, 1481$. G. H. Hardy and J. E. Littlewood [1] found that there are 165 quadruplets less than 1000000. C. R. Sexton [3] settled the number of those contained between 1000000 and 2000000, which is 295. It has recently been found by W. A. Golubew [1], [2], [3], [4] that there are 897 quadruplets less than ten millions.

It is easy to prove that for a given quadruplet such that the least of the primes it contains is greater than 5 any two numbers entering into it differ only in their least digits, which are 1, 3, 7 and 9, respectively. Clearly, each quadruplet forms two pairs of twin numbers.

However, there are pairs of twin numbers not separated by a prime number which do not form a quadruplet. Such are the pairs 179,

(1) An "elementary" proof of the theorem of Brun is to be found in a book of E. Landau [2].

181 and 191, 193, for instance. The latter form a quadruplet with the pair 197, 199. The pairs of twin numbers 419, 421 and 431, 433 are not separated by any prime number; neither of them forms a quadruplet with any other pair of prime numbers. The pairs of twin numbers 809, 811, 821, 823 and 1019 1021, 1031, 1033 have the same property.

It seems a natural question to ask whether there exists an arbitrarily large number of consecutive pairs of twin numbers not separated by prime numbers. We know a number of triplets of such pairs. Such are for instance 179, 181; 191, 193; 197, 199 or 809, 811; 821, 823; 827, 829 or 3359, 3361; 3371, 3373; 3389, 3391 or 4217, 4219; 4229, 4231; 4241, 4243 or 6761, 6763; 6779, 6781; 6791, 6793. We also know an example of four such pairs: 9419, 9421; 9431, 9433; 9437, 9439; 9461, 9463.

It can be proved that if $p \neq 5$ and the numbers p, $p+2$, $p+6$ and $p+8$ are prime, then, dividing p by 210, we obtain 11, 101, or 191 as the remainder.

Turning back to the numbers d_n we note that it is easy to prove that they can be arbitrarily large. In fact, let m denote an arbitrary natural number greater than 1. Let p_n be the greatest prime number $\leqslant m!+1$. The numbers $m!+k$ are composite for $k = 2, 3, \ldots, m$ (since $k \mid m!+k$ for $k = 2, 3, \ldots, m$). Therefore $p_{n+1} \geqslant m!+m+1$ and consequently $d_n = p_{n+1}-p_n \geqslant m$.

On the other hand, we cannot prove that the numbers d_n $(n = 1, 2, \ldots)$ tend to infinity. There are natural numbers n such that $d_n = d_{n+1}$. For instance, $n = 2$, 15, 36, 39, 46. There are also natural numbers n for which $d_n = d_{n+1} = d_{n+2}$: for instance $n = 54$, 464, 682, 709, 821, 829. However we do not know whether for each natural number k there exists a natural number n such that $d_n = d_{n+1} = d_{n+2} = \ldots = d_{n+k}$.

P. Erdös and P. Turán [2] have proved that there exist infinitely many natural numbers n such that $d_n < d_{n+1}$ and also infinitely many numbers n for which $d_n > d_{n+1}$.

It has been proved that for every two natural numbers m and k there exists a natural number n such that each of the numbers $d_n, d_{n+1}, \ldots$ $\ldots, d_{n+k}$ is greater than m. In other words, there exist arbitrarily many consecutive prime numbers such that the differences of the successive ones are arbitrarily large (Erdös [8]). The differences of consecutive prime numbers were the subject of extensive investigations by G. Ricci (cf. Ricci [1], [2]).

§ 4. Goldbach's conjecture. Under this name the conjecture that every even number greater than 2 is the sum of two prime numbers is known. The conjecture has been verified directly for the even numbers up to 100000 (cf. Pipping [3], [4]).

In 1959 Y. Wang [1] proved that each sufficiently large even number is the sum of two natural numbers of which one has at most two prime factors and the other at most three, this greatly improving the first result of this kind obtained by Brun [2] in 1920. Recently, Wang [2] has proved that every sufficiently large even number is the sum of a prime and a natural number which has at most four prime factors.

It follows from Goldbach's conjecture that every odd integer has infinitely many representations of the form $p+q-r$, where p, q, r are prime numbers. This result, not easy to prove, is due to J. G. van der Corput [2]. He also proved that almost every even numbers is a sum of two odd prime numbers. This means that for each positive number ε for every sufficiently large natural number N the number of even natural numbers $< N$ which fail to be sums of two primes is less than εN (van der Corput [1]).

According to A. Desboves [1] every natural number $\leqslant 10000$ of the form $4k+2$ is the sum of two primes, each being of the form $4k+1$. This of course could be true only if number 1 were regarded as a prime. Thus, in particular, $2 = 1+1$, $6 = 1+5$, $14 = 1+13$, $38 = 1+37$, $62 = 1+61$.

Another problem closely connected with the conjecture of Goldbach is whether for a given even natural number n the number $G(n)$ of all possible decompositions of n into the sum of two prime numbers increases to infinity together with the number n. N. Pipping [1], [2] has calculated the function $G(n)$ for even natural numbers n less than 5000 and some others. We have $G(4) = G(6) = 1$, $G(8) = 2$, $G(10) = 3$, $G(12) = 2$, $G(14) = 3$, $G(16) = G(18) = G(20) = 4$, $G(22) = 5$, $G(24) = 6$. Further, we have $G(158) = 9$ and the tables suggest that $G(2n) \geqslant 10$ for $2n \geqslant 158$. Similarly $G(188) = 10$ and it seems plausible that $G(2n) > 10$ for $2n > 188$. The least even number $2n$ for which $G(2n) \geqslant 100$ is 840; actually we have $G(840) = 102$. The greatest number $2n$ for which $G(2n) < 100$ is probably the number $2n = 4574$.

It follows from the conjecture of Goldbach that each odd number greater than 7 is the sum of three odd primes. In fact, if n is an odd natural number > 7, then $n-3$ is an even number > 4. Consequently, in view of Goldbach's conjecture, it is the sum of two primes, each of them odd of course. Thus every odd natural number greater than 7 is the sum of three odd primes.

We do not know whether every odd number > 7 is the sum of three odd primes though the difficulty in solving this question is only of a technical nature, since I. Vinogradov proved in 1937 that for odd natural numbers greater than a certain effectively computable constant a the answer is positive. Later K. G. Borozdkin [1] proved that $a \leqslant \exp(\exp 16{,}038) < 3^{3^{15}}$. In view of this result it suffices to answer the

problem for odd numbers n with $7 < n \leqslant a$, which for a given natural number is a matter of simple but perhaps tedious computations.

The situation is quite different as regards the question whether every even number is a difference of two prime numbers. Here no method of solution is known, even as tedious as that of the previous problem.

A. Schinzel [13] has proved that Goldbach's conjecture implies that every odd number >17 is the sum of three different primes. It follows from the results of Vinogradov that each sufficiently large odd number is such a sum. The conjecture that every even number >6 is the sum of two different prime numbers can also be proved to be equivalent to the conjecture that every natural number >17 is the sum of three different prime numbers (Sierpiński [22]).

In 1930 L. Schnirelman [1] proved elementarily that there exists a number s such that every natural number >1 is representable as the sum of at most s primes. Yin Wen Lin [1] has proved by refining Schnirelman's method that every natural number, from some point onwards, is the sum of 18 at most primes. From the theorem of Vinogradov (quoted above) we see that every sufficiently large natural number is representable as the sum of at most four primes; Vinogradov's proof, however, is not elementary.

It can easily be proved that there exist infinitely many natural numbers which cannot be represented as the sums of less than three primes (compare exercise 2 below).

It has also been conjectured that every odd number >5 is the sum of a prime number and a number of the form $2p$, where p is a prime (Dickson [8], vol. I, p. 424).

EXERCISES. **1.** Prove that every natural number > 11 is the sum of two composite numbers.

Proof. Let n be a natural number greater than 11. If n is even, i.e. $n = 2k$, then $k \geqslant 6$ and $n-6 = 2(k-3)$, which, in view of the fact that $k \geqslant 6$, shows that $n-6$ is a composite number. If n is odd, i.e. $n = 2k+1$, then $k \geqslant 6$ and so $n-9 = 2(k-4)$ is a composite number.

2. Prove that there exist infinitely many natural odd numbers which cannot be represented as the sum of less than three primes.

Proof. Such are, for instance, the numbers $(14k+3)^2$, where $k = 1, 2, \ldots$ In fact, the numbers themselves are not primes. They cannot be represented as the sum of two primes either; for, if they could, then, since they are odd, one of the primes would be equal to 2, which would give $(14k+3)^2 = 2+p$, where p would be a prime. Hence $p = 7(28k^2+12k+1)$, which is impossible.

Remark. It can be proved elementarily that there exist infinitely many odd numbers which are sums of three different primes but are not sums of less than three different primes (cf. Sierpiński [30]).

3. Prove that the conjecture of Goldbach is equivalent to the conjecture that every even number > 4 is the sum of three prime numbers.

Proof. It follows from Goldbach's conjecture that for a natural number $n > 1$ we have $2n = p+q$, where p and q are prime numbers. Hence $2(n+1) = 2+p+q$, that is, every arbitrarily chosen even number > 4 can be represented as the sum of three primes. On the other hand, if every even number > 4 is the sum of three primes, i.e., if for $n > 2$ we have $2n = p+q+r$, where p, q, r are primes, then at least one of the numbers p, q, r must be even, and consequently equal to 2. Suppose that, for instance, $r = 2$. Then $2(n-1) = p+q$ for $n-1 > 1$, which implies the conjecture of Goldbach.

4. Prove that none of the equations $x^2+y^2 = z^2$, $x^2+y^2+z^2 = t^2$, $x^2+y^2+z^2+t^2 = u^2$ is solvable in prime numbers.

Proof. For the first of the equations the result follows from the fact, proved in Chapter II, § 3, that for any solution of the equation in natural numbers at least one of the numbers must be divisible by 4.

Now suppose that there are primes x, y, z, t for which the equation $x^2+y^2+z^2=t^2$ is satisfied. As was proved in Chapter II, § 10, at least two of the numbers x, y, z must be even; since they are primes, each of them is equal to 2. Thus $t^2-z^2 = 8$. But since z, t are primes and obviously odd ones, the equality $(t-z)(t+z) = 8$ implies that $t-z \geqslant 2$ and consequently $t+z \leqslant 4$, which is impossible, since t, z are odd primes.

Finally suppose that there exist primes x, y, z, t, u satisfying the equation $x^2+y^2+z^2+t^2 = u^2$. Clearly, the number u must be greater than 2, and thus it is odd. Therefore at least one of the numbers x, y, z, t must be odd. If precisely one of them were odd, say t, then we would have $x = y = z = 2$, whence $12+t^2 = u^2$ and consequently $(t-u)(t+u) = 12$, whence, since $t-u \geqslant 2$, $t+u \leqslant 6$. But this is impossible since u, t are different odd primes. In the other case, i. e. if three of the primes x, y, z, t were odd, and the fourth of them were even, then $u^2 = x^2+y^2+z^2+t^2$ would be of the form $4k+3$, which is impossible.

5. Find all the solutions of the equation $x^2+y^2+z^2+t^2+u^2 = v^2$ in primes x, y, z, t, u, v with $x \leqslant y \leqslant z \leqslant t \leqslant u \leqslant v$.

Solution. There is precisely one such solution, namely $2^2+2^2+2^2+2^2+3^2 = 5^2$, for it is easy to prove that only one of the numbers x, y, z, t, u, can be odd. So we have $4 \cdot 2^2+u^2 = v^2$, whence $(v-u)(v+u) = 16$, $v-u \geqslant 2$, $v+w \leqslant 8$, so $u = 3$, $v = 5$.

§ 5. Arithmetical progressions whose terms are prime numbers.

Arithmetical progressions consisting of ten different prime numbers are known, for instance $199+210k$, for $k = 0, 1, 2, \ldots, 9$.

V. Seredinsky has found that the numbers $60060k+4943$ ($k = 0, 1, 2, \ldots, 12$) form an arithmetical progression consisting of 13 different prime numbers. We do not know, however, whether there exists an arithmetical progression consisting of a hundred different prime numbers. We shall prove that if such a progression existed then the difference of its terms would have more than thirty digits.

To this end we prove the following theorem.

THEOREM 5. *If n and r are natural numbers, $n > 1$ and if n terms of the arithmetical progression m, $m+r, \ldots, m+(n-1)r$ are odd prime numbers, then the difference r is divisible by every prime number less than n* (cf. Dickson [8], vol. I, p. 425).

Proof. Suppose that m, $n>1$ and r are given natural numbers and that each of the numbers $m, m+r, \ldots, m+(n-1)r$ is an odd prime number. We must have $m<n$, since otherwise the composite number $m+mr=m(1+r)$ would be a term of the arithmetical progression. Let p denote a prime number less than n and let $r_0, r_1, \ldots, r_{p-1}$ be the remainders obtained by dividing the numbers $m, m+r, \ldots, m+(p-1)r$ by p, respectively. The latter are clearly less than p and moreover they are all different from zero, since otherwise one of the prime numbers being greater than $m \geqslant n > p$ would be divisible by the prime p, which is impossible.

Therefore the remainders can take only the values $1, 2, \ldots, p-1$, which are $p-1$ in number. From this we infer that for some two integers k and l such that $0 \leqslant k < l \leqslant p-1$ we have $r_k = r_l$. Consequently, $p \mid (m+lr)-(m+kr)$ and hence $p \mid (l-k)r$. But $0 < l-k \leqslant p-1 < p$, and therefore $p \mid r$. Since p was an arbitrary prime number less than n, the theorem follows.

From Theorem 5 we derive the following

COROLLARY. *If there exists an increasing arithmetical progression consisting of $n>2$ prime numbers, then the difference of this sequence is divisible by the product P_n of all the prime numbers less than n, and consequently it is $\geqslant P_n$.*

In particular, the difference of an arithmetical progression consisting of three different prime numbers must be $\geqslant P_3 = 2$. There exists precisely one arithmetical progression consisting of prime numbers whose difference is 2, namely 3, 5, 7.

It is known that there exist infinitely many arithmetical progressions consisting of three prime numbers each. The proof of this fact, however, is difficult (cf. van der Corput [2] and Chowla [2]).

The problem of the existence of infinitely many such arithmetical progressions is, clearly, equivalent to the question whether the equation $p+r=2q$ has infinitely many solutions in prime numbers p, q, r, with $p \neq r$. It follows from the conjecture H (cf. § 8) that for every natural number n and every prime number $p \geqslant n$ there exist infinitely many increasing arithmetical progressions, each consisting of n terms which are prime numbers, the first term being p.

Here are now some examples of arithmetical progressions consisting of three prime numbers whose first terms are equal to 3: 3, 7, 11; 3, 11, 19; 3, 13, 23; 3, 17, 31; 3, 23, 43; 3, 31, 59; 3, 37, 71; 3, 41, 79; 3, 43, 83.

The difference of an arithmetical progression consisting of four prime numbers must be $\geqslant P_4 = 6$. There are known many arithmetical progressions consisting of four prime numbers each and having the difference

equal to 6, e.g. 5, 11, 17, 23; 11, 17, 23, 29; 41, 47, 53, 59; 61, 67, 73, 79. It follows from the conjecture H that there are infinitely many such progressions, consisting, in addition, of consecutive prime numbers. In particular, such are the progressions 251, 257, 263, 269; 1741, 1747, 1753, 1759.

The difference of an arithmetical progression consisting of five different prime numbers must also be greater than or equal to 6. There exists precisely one arithmetical progression consisting of five different prime numbers whose difference is equal to 6. This is 5, 11, 17, 23, 29. To see that indeed there is precisely one such progression, we note that among five numbers forming an arithmetical progression whose difference is 6 one term must be divisible by 5. Similarly, we easily prove that there exists precisely one arithmetical progression consisting of five prime numbers whose difference is 12 — this is the progression 5, 17, 29, 41, 49 — and that there is no progression with the difference 18 or 24. However, it follows from the conjecture H that there exist infinitely many arithmetical progressions consisting of six prime numbers each and having the difference equal to 30. E.g. 7, 37, 67, 97, 127, 157; 541, 571, 601, 631, 661, 691.

It follows from the above corollary that in every arithmetical progression consisting of seven prime numbers the difference must be divisible by 30. It is easy to prove that there is no arithmetical progression consisting of seven primes whose difference is less than 150. However, there is precisely one arithmetical progression whose difference is 150; namely 7, 157, 307, 457, 607, 757, 907. The reason for this is that in every arithmetical progression consisting of seven natural numbers at least one of them must be divisible by 7.

In virtue of the corollary the difference of an arithmetical progression consisting of ten different prime numbers must be $\geqslant P_{10} = 210$. A progression whose difference is equal to 210 is formed by the numbers $199+210k$, where $k = 0, 1, 2, \ldots, 9$. It follows from the conjecture H that there are infinitely many such progressions.

In virtue of the corollary the difference of an arithmetical progression consisting of a hundred prime numbers would have to be divisible by the product of all prime numbers less than a hundred, and thus it would have more than thirty digits (in the scale of ten). We are not able to find, at least for the time being, any such arithmetical progressions. We do not know any proof of the existence of such an arithmetical progression either.

§ 6. Primes in a given arithmetical progression. Here is a problem on primes in arithmetical progressions of different type than those considered in § 5: for what natural numbers a and b does the arithmetical progression $ak+b$, $k = 1, 2, \ldots$, contain infinitely many prime numbers?

It is clear that, if $(a, b) = d > 1$, then there is no prime in the arithmetical progression $ak+b$, $k = 1, 2, \ldots$, because, for any k, $ak+b = d(ka/b+b/d)$ is a composite number (a/d, b/d are natural numbers). Therefore a necessary condition for the existence of infinitely many primes in an arithmetical progression $ak+b$ is that $(a, b) = 1$.

In the year 1837 Lejeune Dirichlet proved that this condition is also sufficient. The proof given by Lejeune Dirichlet is not elementary. Later the proof was simplified. The simplest proof of this theorem (though still very complicated) makes up chapter VIII (p. 73-78) of the book by E. Trost [3].

We shall prove in the sequel several particular cases of this theorem: in Chapter V with $a = 4$, $b = 1, 3$ (theorems 7 and 7a), in Chapter VI with $b = 1$, a being arbitrary (theorem 11a), in Chapter IX with $a = 8$. $b = 3, 5, 7$ (theorems 1, 2, 3) and with $a = 5$, $b = 4$ (theorem 4).

The following two theorems are equivalent:

T. *If a and b are natural numbers such that $(a, b) = 1$, then there exist infinitely many primes of the form $ak+b$, where k is a natural number.*

T_1. *If a and b are natural numbers such that $(a, b) = 1$, then there exists at least one prime number p of the form $ak+b$ where k is a natural number* (1).

Proof. Trivially, T implies T_1. It is sufficient to prove the converse, that is, that T_1 implies T. We may suppose that $a > 1$ because for $a = 1$ the theorem follows from the fact that theorem T holds. Let a, b be two given natural numbers such that $(a, b) = 1$. Then, of course, $(a^m, b) = 1$. Hence, by theorem T_1, there exists a prime p such that $p = a^m k+b$, for a natural number k. But, since $a > 1$, $a^m \geqslant 2^m > m$. Hence $p > m$. Thus we have proved that for any natural number m there exists a prime of the form $ak+b$ which is greater than m. This shows that there exist infinitely many primes of this form.

It will be proved later (Chapter V, theorem 9) that every prime of the form $4t+1$ is a sum of two perfect squares. Using this result we prove the following corollary of theorem T:

Corollary. *For every natural number n there exists a prime p such that $p = a^2+b^2$, where a, b are natural numbers each greater than n.*

Proof. Let n be a natural number. According to T, there exists a prime $q > n$ which is of the form $4t-1$. Then, clearly

$$\left(4(1^2+q)^2(2^2+q)^2\ldots(n^2+q)^2, q\right) = 1.$$

(1) The proof of the equivalence of theorems T and T_1 was given by me in the year 1950 (cf. Sierpiński [12], p. 526). Six years later the problem of the equivalence of theorems T and T_1 was formulated in The Amer. Math. Monthly as E 1218 (1956), p. 342; and solved ibid. by D. Zeitlin (1957, p. 46), cf. V. S. Hanly [1].

Hence, by theorem T, we infer that there exists a natural number k such that the number

$$p = 4(1^2+q)^2(2^2+q)^2\ldots(n^2+q)^2k-q$$

is a prime, necessarily of the form $4t+1$.

Thus the existence of the numbers a, b such that $p = a^2+b^2$, where $a < b$, is proved.

Suppose $a \leqslant n$. Then

$$b^2 = p-a^2 = 4(1^2+q)^2(2^2+q)^2\ldots(n^2+q)^2k-(a^2+q)$$

$$= (a^2+q)\left(4(1^2+q)^2\ldots\left((a-1)^2+q\right)^2\left((a+1)^2+q\right)\ldots(n^2+q)^2k-1\right),$$

where both the factors on the right-hand side of the equality are relatively prime. Consequently they must be squares, but this is impossible because the second of the factors is of the form $4t-1$. Thus we come to the conclusion that $b > a > n$, and this completes the proof of the corollary.

We note here that, according to a theorem of E. Hecke [1], for any two real numbers $c > d \geqslant 0$ there exists a prime p such that $p = a^2+b^2$ where a, b are natural numbers and $c > \dfrac{a}{b} > d$ (cf. Kubilyus [1]).

§ 7. Trinomial of Euler x^2+x+41. It is easy to prove that there is no polynomial $f(x) = a_0x^m+a_1x^{m-1}+\ldots+a_{m-1}x+a_m$ with integral coefficients and $a_0 > 0$ for which the numbers $f(x)$ would be prime for all integral values of x. In fact, as is well known, for sufficiently large x, say for $x > x_0$, the function $f(x)$ is increasing. If for some $x_1 > x_0$, $f(x_1) = p$ is a prime number, then, as can easily be verified, $p \mid f(x_1+p)$, which, in virtue of $f(x_1+p) > f(x_1) = p$, implies that $f(x_1+p)$ is a composite number.

It has also been proved that there is no rational function whose all values would be prime numbers for all integral values of the argument (Buck [1]).

However, there are polynomials of degree two with integral coefficients taking prime values for long sequence of consecutive natural numbers. For example such is the polynomial of Euler $f(x) = x^2+x+41$, whose values are prime numbers for $x = 0, 1, \ldots, 39$. To see this we note that $f(x+1) = f(x)+2(x+1)$. From this we easily infer that for $x = 0, 1, 2, \ldots$ the values $f(x)$ are the partial sums of the series $41+2\cdot1+2\cdot2+2\cdot3+\ldots$ Thus we obtain the values 41, 43, 47, 53, 61, 71, 83, ..., 1601. As can be checked in the tables of prime numbers, each of these numbers is a prime. Since $f(-x) = f(x-1)$, also the numbers $f(-x)$ are prime for $x = 1, 2, \ldots, 40$. Thus for $x = -40, -39, \ldots,$

$-1, 0, 1, \ldots, 39$ the function $f(x)$ takes the values which are all (not necessarily different) prime numbers. The function $f(x)$ has another interesting property: for integral values of x there is no divisor d with $1 < d < 41$ dividing $f(x)$.

In fact, suppose that for an integer x we have $d \mid f(x)$, where $1 < d < 41$. Let r be the remainder obtained by dividing x by d. Then $x = kd+r$, where k is an integer and $0 \leqslant r < d$. But since $f(kd+r) = kd(kd+2r+1)+f(r)$, the relation $d \mid f(x)$ implies $d \mid f(r)$; however this leads to a contradiction. In fact, in virtue of $0 \leqslant r < d < 41$, we must have $0 \leqslant r \leqslant 39$; therefore, as we know, $f(r)$ is a prime number $\geqslant 41$, and so it cannot have a divisor d such that $1 < d < 41$. Thus for an integer x the number $f(x)$ has no divisor d such that $1 < d \leqslant 41$.

This property is particularly relevant to finding whether for a given natural number $x \geqslant 40$ the number $f(x)$ is a prime. For $x = 40$ we have $f(40) = 40 \cdot 41+41 = 41^2$, so the number $f(x)$ is composite. The number $f(41) = 41 \cdot 42+41 = 41 \cdot 43$ is also composite. If $x > 40$ and, if the number $f(x)$ isc omposite, then, by $(x+1)^2 = x^2+2x+1$ and $x^2+x+1 = f(x)$, we obtain $f(x) < (x+1)^2$. Therefore the number $f(x)$ has a prime divisor $p < x+1$ and, in virtue of what we proved above, $41 \leqslant p < x$ (since dividing $f(x)$ by x we obtain the remainder 41). Thus, in particular, the number $f(42) = 42 \cdot 43+41$ is prime; for, plainly, it is not divisible by 41, the only prime number p for which $41 \leqslant p < 42$.

According to E. Trost ([3], p. 41), for x running up to 11000 the function $f(x)$ takes 4506 different values that are prime numbers.

We do not know whether in the sequence $f(x)$ $(x = 1, 2, \ldots)$ there are infinitely many prime numbers. (The answer in the affirmative follows from conjecture H, cf. § 8.)

It follows from the properties of the trinomial $f(x)$ that the trinomial $g(x) = f(x-40) = x^2-79x+1601$ takes values that are (not necessarily different) prime numbers for $x = 0, 1, 2, \ldots, 79$. (We have $g(t) = g(79-t)$ for all t.)

It has been proved by D. H. Lehmer [3] that if there exists a natural number A greater than 41 such that the trinomial x^2+x+A takes values which are prime numbers for all $x = 0, 1, 2, \ldots, A-2$, then A must be greater than $125 \cdot 10^7$ and H. Heilbronn and E. Linfoot [1] have proved that, if A exists, it is unique.

For $x = 0, 1, \ldots, 28$ the values taken by $6x^2+6x+31$ are all different prime numbers of the form $6k+1$; they are contained between 31 and 4909 with the limits included (C. Coxe, cf. van der Pol and Speziali [1]). The values of the binomial $2x^2+29$ are prime numbers for $-28 \leqslant x \leqslant 28$.

It can easily be proved that there exist polynomials of degree n taking prime values for $x = 0, 1, \ldots, n$; however we do not know any

polynomial of degree two or higher in variable x about which we could prove that it takes prime values for infinitely many values of x. In particular, we do not know whether the binomial x^2+1 has this property. W. A. Golubew [5] has presented a list of all natural numbers $x \leqslant 20000$ for which the numbers x^2+1 are prime. D. Shanks [1] has found that there are 11223 numbers $x \leqslant 180000$ with this property. P. Kuhn [1] has proved that there exist infinitely many numbers x^2+1 composed of at most 3 primes (cf. Wang [3]) and B. M. Bredihin [1] has proved that there exist infinitely many primes of the form x^2+y^2+1.

If a polynomial $f(x)$ with integral coefficients takes prime number values for infinitely many x's, then, plainly, the coefficient a_0 at the highest power of variable x must be positive, since for sufficiently large values of x the polynomial has the same sign as a_0. Furthermore, the polynomial $f(x)$ cannot be the product of two polynomials with integral coefficients, since otherwise for sufficiently large values of x the number $f(x)$ would be composite. Therefore the polynomial $f(x)$ is irreducible. However, these conditions are not yet sufficient for $f(x)$ to take prime number values which are even for at least one value of x. In fact, the polynomial x^2+x+4 is irreducible (it has no real root) and for all integers x the numbers x^2+x+4 are composite — they are even natural numbers greater than 4, since, as we know, the number $x^2+x = (x+1)x$ is even and non-negative.

In 1857 W. Bouniakowsky [2] formulated the following conjecture:

If $f(x)$ is an irreducible polynomial with integral coefficients and if N denotes the greatest common divisor of the numbers $f(x)$, x running over all integers, then the polynomial $f(x)/N$ takes prime number values for infinitely many x's (cf. Dickson [8], vol. I, p. 333).

For instance, consider the polynomial $f(x) = x^2+x+4$. Since $f(0)=4$, $f(1) = 6$ and, as we already know, $f(x)$ is an even integer for integer x, then for x running over all integers the greatest common divisor of the numbers $f(x)$ is 2. Consequently it follows from the conjecture of Bouniakowsky, that for infinitely many integers x the number $x(x+1)/2+2$ is prime.

§ 8. The conjecture H. Let s denote a natural number and let $f_1(x)$, $f_2(x), \ldots, f_s(x)$ be polynomials whose coefficients are integers. Suppose that there exist infinitely many natural numbers x for which each of the numbers $f_1(x,), f_2(x), \ldots, f_s(x)$ is a prime. As we learned in § 7, the polynomials $f_i(x)$, $i = 1, 2, \ldots, s$, must be irreducible and the leading coefficient of each of them must be positive. Accordingly, for sufficiently large values of x all the numbers $f_i(x)$, $i = 1, 2, \ldots, s$, can be arbitrarily large. As can easily be verified, this implies that there is no natural number $d > 1$ which divides the number $P(x) = f_1(x)f_2(x)\ldots f_s(x)$ for any

natural value of x. In fact, if such a number could exist, it would be the divisor of the product of s arbitrarily large prime numbers, which is impossible.

We have thus proved that if s is a natural number and $f_1(x), f_2(x), \ldots, f_s(x)$ are polynomials whose coefficients are integers and if for infinitely many natural numbers x the numbers $f_1(x), f_2(x), \ldots, f_s(x)$ are prime, then the polynomials must satisfy the following condition:

CONDITION S. *Each of the polynomials $f_i(x)$ $(i = 1, 2, \ldots, s)$ is irreducible, its leading coefficient is positive and there is no natural number $d > 1$ that is a divisor of each of the numbers $P(x) = f_1(x)f_2(x)\ldots f_s(x)$, x being an integer.*

In 1958 A. Schinzel formulated the following conjecture:

CONJECTURE H. *If s is a natural number and if $f_1(x), f_2(x), \ldots, f_s(x)$ are polynomials with integral coefficients satisfying condition* S, *then there exist infinitely many natural values of x for which each of the numbers $f_1(x), f_2(x), \ldots, f_s(x)$ is prime* (cf. Schinzel et Sierpiński [3], p. 188).

For the case of linear polynomials f_i an equivalent conjecture was formulated earlier by L. E. Dickson [1].

We present here some of the corollaries which follow from conjecture H.

Let n be a given natural number and let $f_1(x) = x^{2^n}+1$, $f_2(x) = x^{2^n}+3$, $f_3(x) = x^{2^n}+7$, $f_4(x) = x^{2^n}+9$. For $P(x) = f_1(x)f_2(x)f_3(x)f_4(x)$ we have $P(0) = 1\cdot3\cdot7\cdot9$ and $P(1) = 2\cdot4\cdot8\cdot10$. Consequently, $\big(P(0), P(1)\big) = 1$. Therefore condition S is satisfied and conjecture H gives the following corollary:

For every natural number n there exist infinitely many natural numbers x for which each of the numbers $x^{2^n}+1$, $x^{2^n}+3$, $x^{2^n}+7$, $x^{2^n}+9$ is a prime (Sierpiński [33]).

This implies that there exist infinitely many quadruplets of prime numbers (cf. § 3), and that there are infinitely many prime numbers of the form x^2+1 as well as of the form x^4+1. W. A. Golubew has calculated that there are only four natural numbers x less than ten millions for which each of the numbers x^2+1, x^2+3, x^2+7, x^2+9 is a prime. These are $x = 10$, 1420, 2080, 2600.

Now let k denote an arbitrary integer and let $f_1(x) = x, f_2(x) = x+2k$. For $p(x) = f_1(x)f_2(x)$ we have $P(1) = 2k+1$, $P(2) = 4(k+1)$. Since clearly $\big(2k+1, 4(k+1)\big) = 1$, the polynomials satisfy condition S. Consequently, according to conjecture H, there exist infinitely many natural numbers x for which the numbers $p = x$ and $q = x+2k$ are both prime numbers. Hence $2k = p-q$, which proves that the number $2k$ admits infinitely many representations as the difference of two prime numbers. This means that the conjecture H implies that every even number has

infinitely many representations as the difference of two prime numbers. It can also be deduced from conjecture H that every even number has infinitely many representations as the difference of two consecutive prime numbers (cf. Schinzel and Sierpiński [3], p. 190).

It follows from conjecture H that if a and b are natural numbers such that $(a, b) = \big(a, b(b+2)\big) = 1$, then there exist infinitely many prime numbers p of the form $ak+b$, where k is a natural number, such that $p+2$ is a prime number. In fact, let $f_1(x) = ax+b$, $f_2(x) = ax+b+2$. For $P(x) = f_1(x)f_2(x)$ we have $P(0) = b(b+2)$, $P(1) = (a+b)\times(a+b+2)$ and $P(1)+P(-1) = 2a^2+2b(b+2)$. If there exists a prime number q such that $q \mid P(x)$ for all integers x, then, if b is odd, $P(0)$, and consequently q, are odd; and if b is even, then, in view of $(a, b) = 1$, a is odd; thus both $a+b$ and $a+b+2$ are odd and, consequently, $P(1)$ is odd, which implies that also q is odd. Therefore, in any case, q is odd. Since we have assumed that $q \mid P(0)$, i.e. $q \mid b(b+2)$ and $q \mid P(1)+P(-1)$, we have $q \mid 2a^2$ and consequently, since q is odd, $q \mid a$. But this is impossible since $\big(a, b(b+2)\big) = 1$. Thus we see that condition S is satisfied. Therefore it follows from conjecture H that there exist infinitely many natural numbers x for which the numbers $f_1(x) = ax+b$ and $f_2(x) = ax+b+2$ are prime. The corollary is thus proved.

It is easy to see that the condition $\big(a, b(b+2)\big) = 1$ is also necessary for the existence of infinitely many prime numbers p of the form $ak+b$ for which also the number $p+2$ is a prime.

Let k be an arbitrary integer and let $f_1(x) = x$, $f_2(x) = 2k+1+2x$. For $P(x) = f_1(x)f_2(x)$ we have $P(1) = 2k+3$, $P(-1) = -(2k-1)$. Since $(2k-1, 2k+3) = 1$ for every integer k, we see that the polynomials satisfy condition S. Then, according to conjecture H, there exist infinitely many natural numbers x for which the numbers $q = x$ and $p = 2k+1+2x$ are both prime.

Hence $2k+1 = p-2q$. Thus conjecture H implies that every odd integer (>0, $=0$ or <0) has infinitely many representations as the difference of a prime number and the double of a prime number.

G. de Rocquigny [1] has asked whether every integer divisible by 6 is the difference of two primes of the form $6k+1$. The positive answer to this is a corollary of conjecture H. In fact, for $f_1(x) = 6x+1$ and $f_2(x) = 6x+6k+1$, $P(x) = f_1(x)f_2(x)$ we have $P(0) = 6k+1$, $P(-k) = -(6k-1)$ and, as is known, $(6k-1, 6k+1) = 1$ for all integers k.

It follows from conjecture H that there exist arbitrarily long arithmetical progressions whose terms are consecutive prime numbers (cf. Schinzel and Sierpiński [3], p. 191).

There are many other corollaries which can be derived from conjecture H, e.g. the conjecture of Bouniakowsky (cf. Schinzel and Sierpiński [3] and Schinzel [15]).

EXERCISE. Prove that conjecture H implies the following assertion. Given two relatively prime integers a and b such that one of them is even and $a > 0$. Then there exist infinitely many prime numbers p such that $ap+b$ is a prime.

Proof. Let $f_1(x) = ax+b$, $f_2(x) = x$. For $P(x) = f_1(x)f_2(x)$ we have $P(1) = a+b$, $P(-1) = a-b$, and since one of the numbers a, b is even, the other, in virtue of $(a, b) = 1$, is odd and so from $(a, b) = 1$, it follows that $(a+b, a-b) = 1$. Therefore $(P(1), P(-1)) = 1$ and this shows that condition S is satisfied. Consequently, from conjecture H we conclude that there exist infinitely many x for which both $f_2(x) = x$ and $f_1(x) = ax+b$ are prime numbers, and this is what was to be proved.

§ 9. The function $\pi(x)$. For any real number x we denote by $\pi(x)$ the number of primes not greater than x. We then have $\pi(1) = 0$, $\pi(2) = 1$, $\pi(3) = \pi(4) = 2$, $\pi(5) = \pi(6) = 3$, $\pi(7) = \pi(8) = \pi(9) = \pi(10) = 4$, $\pi(100) = 25$, $\pi(1000) = 168$, $\pi(10000) = 1229$, $\pi(10^5) = 9592$, $\pi(10^6) = 78498$, $\pi(10^7) = 664579$, $\pi(10^8) = 5761455$. In 1958 D. H. Lehmer [11] calculated that $\pi(10^9)$ equal 50847534 (this was a correction of the result of Bertelsen obtained in 1893) and $\pi(10^{10}) = 455052512$ (cf. Locher-Ernst [1]).

Obviously we have $\pi(p_n) = n$ for $n = 1, 2, \ldots$ P. Erdös has found (cf. Trost [3], pp. 52-53) quite an elementary proof of the inequality

$$\pi(n) \geqslant \frac{\log n}{2\log 2} \quad \text{for} \quad n = 1, 2, \ldots \tag{2}$$

As we proved in Chapter I, § 14, every natural number has a unique representation in the form $k^2 l$, where k and l are natural numbers and, moreover, the number l is square-free. For each of the n numbers $1, 2, \ldots, n$, we have $k^2 l \leqslant n$; so, *a fortiori*, $k^2 \leqslant n$. Therefore $k \leqslant \sqrt{n}$. Consequently the number k can take at most $\sqrt{n}$ different values. The numbers l, being square-free and less than n, can be represented as products of different primes each not greater than n, i.e. as products of primes belonging to the sequence $p_1, p_2, \ldots, p_{\pi(n)}$. The number of such products (including number 1) is $2^{\pi(n)}$. Consequently the numbers l can assume at most $2^{\pi(n)}$ different values. Therefore the number of the products lk^2 (where l is square-free) each being not greater than n, is at most $\sqrt{n}2^{\pi(n)}$. Since every natural number $\leqslant n$ is representable as such a product, we have $n \leqslant \sqrt{n}2^{\pi(n)}$. Hence $\sqrt{n} \leqslant 2^{\pi(n)}$ and, further, taking the logarithm of both sides of the last inequality, we obtain $\frac{1}{2}\log n \leqslant \pi(n)\log 2$, which proves formula (2).

Later on (in § 14) we shall prove stronger inequalities for the function $\pi(n)$. The main interest in equality (2), however, is aroused by the simplicity of its proof.

Let k denote an arbitrary natural number and let $n = p_k$. By formula (2), in view of $\pi(p_k) = k$, we have $k \geqslant \log p_k / 2\log 2$. Therefore

$p_k \leqslant 2^{2k}$ for $k = 1, 2, \ldots$, which, in virtue of the fact that 2^{2k} is a composite number for every $k = 1, 2, \ldots$, proves the inequality

$$p_k < 2^{2k} \quad \text{for} \quad k = 1, 2, \ldots \tag{3}$$

EXERCISES. **1.** Prove that for natural numbers n the inequality

$$\frac{\pi(n-1)}{n-1} < \frac{\pi(n)}{n} \tag{4}$$

holds if and only if n is a prime. For n being composite numbers we have

$$\frac{\pi(n-1)}{n-1} > \frac{\pi(n)}{n}. \tag{5}$$

Proof. If n is a composite number, then $\pi(n) = \pi(n-1)$ and inequality (5) follows.

If n is a prime number, then $\pi(n) = \pi(n-1)+1$, whence

$$\frac{\pi(n)}{n} - \frac{\pi(n-1)}{n-1} = \frac{1}{n}\left(1 - \frac{\pi(n-1)}{n-1}\right). \tag{6}$$

But since $\pi(k) < k$ for $k = 1, 2, \ldots$, (6) implies (4).

2. Given a natural number m, find all the solutions of the equation $\pi(n) = m$ in natural numbers n.

Solution. These are the natural numbers n for which $p_m \leqslant n < p_{m+1}$. Thus for a given natural number m there are $p_{m+1} - p_m$ solutions.

§ 10. Proof of Bertrand's postulate (Theorem of Tchebycheff). For a given real number x we denote by $[x]$ the greatest integer $\leqslant x$. Thus, in particular, we have $[\frac{3}{4}] = 0$, $[-\frac{3}{4}] = -1$, $[\sqrt{2}] = 1$, $[\pi] = 3$. It follows from the definition that for all real numbers x we have $x-1 < [x] \leqslant x$. The equality $[x] = x$ holds if and only if x is an integer. If k is an integer, then for the x's that are real numbers we have $[x+k] = [x]+k$. For any real numbers x, y we have, of course, $[x]+[y] \leqslant [x+y]$. E.g.

$$0 = [\tfrac{1}{2}]+[\tfrac{2}{3}] < [\tfrac{1}{2}+\tfrac{2}{3}] = 1 \quad \text{but} \quad [\tfrac{1}{3}]+[\tfrac{1}{2}] = [\tfrac{1}{3}+\tfrac{1}{2}] = 0.$$

THEOREM 6. *The exponent of a prime p in the factorization into prime numbers of $n!$, where n is a natural number, is*

$$\alpha = \left[\frac{n}{p}\right] + \left[\frac{n}{p^2}\right] + \left[\frac{n}{p^3}\right] + \ldots \tag{7}$$

Proof. Let n, k be two given natural numbers and p a prime number $\leqslant n$. The numbers of the sequence $1, 2, \ldots, n$ which are divisible by p^k are of the form lp^k, where l is a natural number such that $lp^k \leqslant n$, that is $l \leqslant n/p^k$. The number of l's is, of course, $[n/p^k]$. On the other hand, it is clear that the exponent α of the prime p in factorization into prime numbers of the number $n!$ is obtained by adding to the number of the

terms of the sequence $1, 2, \ldots, n$ which are divisible by p the number of the terms divisible by p^2 and then the number of the terms divisible by p^3 and so on. This gives formula (7).

As a simple application of theorem 6 we calculate the number of zeros at the end of the number 100!.

According to formula (7) (for $n = 100$ and $p = 2$) the exponent of the number 2 in the factorization into prime numbers of the number 100! is

$$\left[\frac{100}{2}\right]+\left[\frac{100}{2^2}\right]+\left[\frac{100}{2^3}\right]+\ldots = 50+25+12+6+3+1 = 97.$$

The exponent of number 5 is

$$\left[\frac{100}{5}\right]+\left[\frac{100}{5^2}\right] = 20+4 = 24.$$

Hence it follows that number 100! has 24 zeros at the end in its decimal expansion.

LEMMA 1. *For natural numbers $n > 1$ we have*

$$\binom{2n}{n} > \frac{4^n}{2\sqrt{n}}. \tag{8}$$

Proof. Inequality (8) holds for $n = 2$ because $\binom{4}{2} = 6 > \frac{4^2}{2\sqrt{2}}$. Suppose that inequality (8) holds for a given natural number n. We then have

$$\binom{2n+2}{n+1} = 2\frac{2n+1}{n+1}\binom{2n}{n} > \frac{2(2n+1)4^n}{(n+1)2\sqrt{n}} = \frac{2(2n+1)4^n}{\sqrt{4n}(n+1)\sqrt{n+1}} > \frac{4^{n+1}}{2\sqrt{n+1}},$$

and so, since $(2n+1)^2 > 4n(n+1)$, we infer that $2n+1 > \sqrt{4n(n+1)}$. From this the proof of inequality (8) for $n > 1$ follows by induction.

LEMMA 2. *The product P_n of the prime numbers $\leqslant n$, where n is a natural number, is not greater than 4^n.*

Proof. The lemma is of course true for $n = 1$ and $n = 2$. Let n denote a natural number > 2. We suppose that the lemma holds for the natural numbers $< n$. If n is an even number > 2, then $P_n = P_{n-1}$. Hence the lemma holds for the number n. If, however, $n = 2k+1$, where k is a natural number, then each prime number p such that $k+2 \leqslant p \leqslant 2k+1$ is a divisor of the number

$$\binom{2k+1}{k} = \frac{(2k+1)2k(2k-1)\ldots(k+2)}{1\cdot 2\ldots k}. \tag{9}$$

In view of the fact that

$$(1+1)^{2k+1} > \binom{2k+1}{k} + \binom{2k+1}{k+1} = 2\binom{2k+1}{k},$$

we have

$$\binom{2k+1}{k} < 4^k.$$

Consequently, the product of all the (different) prime numbers such that $k+2 \leqslant p \leqslant 2k+1$ is a divisor of number (9) not greater than 4^k. But since, by the assumption that the lemma is valid for numbers less then n, the product of the prime numbers $\leqslant k+1$ is less than 4^{k+1}, we have $P_n = P_{2k+1} < 4^k \cdot 4^{k+1} = 4^{2k+1} = 4^n$. Hence $P_n < 4^n$. Thus, by induction, the lemma follows.

LEMMA 3. *If p is a prime divisor of number $\binom{2n}{n}$ with $p \geqslant \sqrt{2n}$, then the exponent of p in the factorization into primes of number $\binom{2n}{n}$ is equal to 1.*

By theorem 6 the exponent of the prime p in the factorization into primes of number $(2n)!$ is $\left[\frac{2n}{p}\right]+\left[\frac{2n}{p^2}\right]+\left[\frac{2n}{p^3}\right]+\ldots$ and in the factorization of the number $n!$ the exponent on a prime p is $\left[\frac{n}{p}\right]+\left[\frac{n}{p^2}\right]+\left[\frac{n}{p^3}\right]+\ldots$

In virtue of

$$\binom{2n}{n} = \frac{(2n)!}{(n!)^2}$$

the exponent of the prime p in the factorization into prime numbers of the number $\binom{2n}{n}$ is

$$\alpha = \sum_{k=1}^{\infty}\left[\frac{2n}{p^k}\right] - 2\sum_{k=1}^{\infty}\left[\frac{n}{p^k}\right] = \sum_{k=1}^{\infty}\left(\left[\frac{2n}{p^k}\right] - 2\left[\frac{n}{p^k}\right]\right).$$

If $p \geqslant \sqrt{2n}$, then $p = \sqrt{2n}$ only in the case where $n = 2$. Therefore for $n \neq 2$ we have $p > \sqrt{2n}$, whence $\alpha = \left[\frac{2n}{p}\right] - 2\left[\frac{n}{p}\right] < 2$. Consequently, $\alpha < 2$, that is $\alpha \leqslant 1$ (since α is an integer). This proves lemma 3 for $n \neq 2$. For $n = 2$, however, we verify it directly; we have $\binom{4}{2} = 2\cdot 3$.

LEMMA 4. *Each divisor of number $\binom{2n}{n}$ which is of the form p^r, p being a prime and r a natural number, is not greater than $2n$. We have*

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

Proof. For a prime p such that $p^r \mid \binom{2n}{n}$, the exponent of p in the factorization of $\binom{2n}{n}$ into primes is

$$\alpha = \sum_{k=1}^{\infty} \left(\left[\frac{2n}{p^k} \right] - 2 \left[\frac{n}{p^k} \right] \right) \geqslant r.$$

If p^r were $> 2n$, then we would have $\left[\frac{2n}{p^k}\right] - 2\left[\frac{n}{p^k}\right] = 0$ for $k \geqslant r$; consequently

$$\alpha = \sum_{k=1}^{r-1} \left(\left[\frac{2n}{p^k} \right] - 2 \left[\frac{n}{p^k} \right] \right).$$

But since, for all real x, $[2x] - 2[x] \leqslant 1$, the last equality would imply that $\alpha \leqslant r-1$, which contradicts the fact that $\alpha \geqslant r$. Therefore $p^r \leqslant 2n$. To prove the second part of the lemma we note that since in the factorization of number $\binom{2n}{n}$ only the primes $\leqslant 2n$ can occur, we have $\binom{2n}{n} \leqslant (2n)^{\pi(2n)}$. The lemma is thus proved.

Lemma 5. *If n is a natural number > 2, then none of the primes p for which $\frac{2}{3}n < p \leqslant n$ can be a divisor of number $\binom{2n}{n}$.*

Proof. If $\frac{2}{3}n < p \leqslant n$, then $\frac{2n}{p} < 3$ and $\frac{n}{p} \geqslant 1$. Therefore $\left[\frac{2n}{p}\right] \leqslant 2$, $\left[\frac{n}{p}\right] \geqslant 1$, which gives $\left[\frac{2n}{p}\right] - 2\left[\frac{n}{p}\right] = 0$ [1]. For $k > 1$ we then have $p^k > \frac{4}{9}n^2$ and consequently $\frac{2n}{p^2} < \frac{9}{2n} < 1$ for $n > 4$. Therefore $\left[\frac{2n}{p^k}\right] - 2\left[\frac{n}{p^k}\right] = 0$ for all $k > 1$ and $n > 4$. Hence we conclude that for $n > 4$ the exponent of the prime p in the factorization into primes of number $\binom{2n}{n}$ is zero, which means that $\binom{2n}{n}$ is not divisible by p. This proves the lemma for $n > 4$. To prove it for the remaining cases, that is for $n = 3$ and $n = 4$ we check that the inequalities $\frac{2}{3}n < p \leqslant n$ imply $p = 3$ and that 3 is not a divisor either of $\binom{6}{3} = 20$ or of $\binom{8}{4} = 70$. Lemma 5 is thus proved.

[1] In fact for real numbers x we have $[2x] \leqslant 2x$, $2[x] > 2x - 1$, whence $[2x] - 2[x] < -1$, and consequently, since the left-hand side is an integer, we have $[2x] - 2[x] = 0$.

Lemma 6. *The exponent of a prime number p such that $n < p < 2n$ in the factorization into primes of the number $\binom{2n}{n}$ is equal to 1.*

Proof. For $n < p < 2n$ we have $1 < \frac{2n}{p} < 2$, $\frac{n}{p} < 1$. Therefore $\left[\frac{2n}{p}\right] = 1$, $\left[\frac{n}{p}\right] = 0$. For $k \geqslant 2$ we have $\frac{2n}{p^k} \leqslant \frac{2n}{p^2} < \frac{2}{n}$. Therefore, for $n > 1$, $\frac{2n}{p^k} < 1$ and, consequently, $\left[\frac{2n}{p^k}\right] = 0$, whence of course $\left[\frac{n}{p^k}\right] = 0$. Hence the exponent α of the prime p in the factorization of number $\binom{2n}{n}$ into primes is equal to 1. Clearly, for $n = 1$ there is nothing to prove, since $n < p < 2n$ cannot hold for $n = 1$. The lemma is thus proved.

Lemma 7. *For natural numbers $n \geqslant 14$ we have $\pi(n) \leqslant \frac{1}{2}n - 1$.*

As can easily be verified, we have $\pi(14) = 6 = \frac{14}{2} - 1$. Consequently lemma 7 is true for $n = 14$. Suppose that n is a natural number not less than 15. In the sequence $1, 2, \ldots, n$ the even numbers $4, 6, 8, \ldots, 2\left[\frac{n}{2}\right]$ are composite. Their number is clearly $\left[\frac{n}{2}\right] - 1$. Moreover, in the sequence $1, 2, \ldots, n$ for $n \geqslant 15$, there are other composite numbers which are odd, namely 1, 9, 15. Thus

$$\pi(n) \leqslant n - \left(\left[\frac{n}{2} \right] - 1 + 3 \right) = n - \left[\frac{n}{2} \right] - 2 < \frac{n}{2} - 1$$

(because $\left[\frac{n}{2}\right] > \frac{n}{2} - 1$). Thus $\pi(n) < \frac{n}{2} - 1$ for $n \geqslant 15$, and this completes the proof of the lemma.

Lemma 8. *Let R_n denote the product of the primes p such that $n < p \leqslant 2n$. In the case when there are no such primes, let $R_n = 1$. Then*

$$R_n > \frac{4^{n/3}}{2\sqrt{n}(2n)^{\sqrt{n}/2}} \tag{10}$$

holds for all $n \geqslant 98$.

Proof. It follows immediately from the definition of R_n that $R_n \mid \binom{2n}{n}$. Consequently $\binom{2n}{n} = Q_n R_n$, where Q_n is a natural number. Hence, by lemma 6, we infer that none of the numbers p with $n < p \leqslant 2n$ appears in the factorization into primes of the number Q_n. It follows that each of the primes p which does appear in this factorization must be $\leqslant n$, hence, by lemma 5, it must be $\leqslant \frac{2}{3}n$.

The product of all the different primes p such that $p \mid Q_n$ is, then, not greater than the product of the primes of which none is greater than $\frac{2}{3}n$. Therefore, by lemma 3 and the relation $Q_n \mid \binom{2n}{n}$, the exponent of a prime number p in the factorization of the number Q_n into primes can be greater than 1 only in the case where $p < \sqrt{2n}$. The number of such primes is in virtue of lemma 7 (with $[\sqrt{2n}]$ in place of n — this substitution is justified because, since $n \geqslant 98$, we have $\sqrt{2n} \geqslant 14$) less than $\sqrt{2n}/2$. By lemma 4 the product of the powers of the primes appearing in the factorization into primes of the number $\binom{2n}{n}$ is $< (2n)^{\sqrt{2n}/2}$. We obtain of course the same inequality for the product of the powers of the primes appearing in the factorization into primes of number Q_n. Hence it follows that $Q_n < 4^{2n/3}(2n)^{\sqrt{2n}/2}$. But since $\binom{2n}{n} = Q_n R_n$, in virtue of Lemma 1 we obtain $Q_n R_n > 4^n/2\sqrt{n}$ and thus formula (10) follows.

LEMMA 9. *For natural numbers $k \geqslant 8$ we have $2^k > 18(k+1)$.*

Proof. We have $2^8 = 256 > 18 \cdot 9$. If $2^k > 18(k+1)$, then $2^{k+1} = 2^k + 2^k > 18k + 18 + 18k + 18 > 18k + 36 = 18(k+2)$. Thus, by induction, the lemma follows.

LEMMA 10. *For real numbers $x \geqslant 8$ we have $2^x > 18x$.*

Proof. For a real number $x \geqslant 8$ we have $[x] \geqslant 8$. Hence, by lemma 9, $2^x \geqslant 2^{[x]} > 18([x]+1) > 18x$, whence $2^x > 18x$, as required.

LEMMA 11. *For natural numbers $k \geqslant 6$ we have $2^k > 6(k+1)$.*

Proof. In view of lemma 9 it is sufficient to prove lemma 11 for $k = 6$ and $k = 7$. To do this we check that $2^6 = 64 > 6 \cdot 7$, and $2^7 = 128 > 6 \cdot 8$.

LEMMA 12. *For real numbers $x \geqslant 6$ we have $2^x > 6x$.*

The proof is analogous to that of lemma 10.

LEMMA 13. *If n is a natural number $\geqslant 648$, then $R_n > 2n$.*

Proof. In view of lemma 8 it is sufficient to prove that if $n \geqslant 648$, then $4^{n/3} > 4n\sqrt{n}(2n)^{\sqrt{n}/2}$. To do this we note that, if $n \geqslant 648$, then $\sqrt{2n}/6 > 6$ and, by lemma 12, $2^{\sqrt{2n}/6} > \sqrt{2n}$, whence, raising each side to the power $\sqrt{2n}$, we obtain $2^{n/3} > (2n)^{\sqrt{n}/2}$. But, since, in virtue of $n \geqslant 648$, we have $2n/9 > 8$, by the use of lemma 10 we obtain $2^{2n/9} > 4n$, whence $2^{n/3} > 4n\sqrt{4n} > 4n\sqrt{n}$. This, for $n \geqslant 648$, gives $4^{n/3} > 4n\sqrt{n}(2n)^{\sqrt{n}/2}$. The lemma is thus proved.

LEMMA 14. *If $n \geqslant 648$, then between n and $2n$ there are at least two different prime numbers.*

Proof. It follows from the definition of R_n that if there were at most one prime number between n and $2n$, then we have $R_n \leqslant 2n$, which, for $n \geqslant 648$, is impossible because of lemma 13.

THEOREM 7. *If n is a natural number > 5, then between n and $2n$ there are at least two different prime numbers.*

Proof. For $n = 6$ the theorem is clearly true, since between 6 and 12 there are two primes, 7 and 11. Thus, in virtue of lemma 14, the theorem is to be proved for natural numbers n such that $7 \leqslant n < 648$. In order to do this it is not necessary to verify the theorem for each of the natural numbers $7, 8, \ldots, a = 647$ directly. It is sufficient to define a sequence of prime numbers $q_0, q_1, \ldots, q_m$ such that $q_0 = 7$, $q_k < 2q_{k-2}$ for $k = 2, 3, \ldots, m$ and $q_{m-1} > a$. Let n denote an arbitrary natural number such that $7 \leqslant n \leqslant a$. The first term of the sequence $q_0, q_1, \ldots, q_m$ is $\leqslant n$ and the last but one term is $> a \geqslant n$. Thus there exists a greatest index k with $k < m-1$ such that $q_k \leqslant n$. We have $k+2 \leqslant m$, $n < q_{k+1}$ and thus, in virtue of the relation $q_{k+2} < 2q_k \leqslant 2n$, between n and $2n$ there are at least two prime numbers q_{k+1} and q_{k+2}.

By the use of the tables of prime numbers we can easily check that the sequence defined above is the sequence 7, 11, 13, 19, 23, 37, 43, 73, 83, 139, 163, 277, 317, 547, 631, 653, 1259.

As an immediate corollary to theorem 7 we derive

THEOREM 8 (Tchebycheff). *If n is a natural number > 3, then between n and $2n-2$ there is at least one prime number.*

Proof. For $n = 4$ and $n = 5$ the theorem is true, since between 4 and 6 is the prime 5, and between 5 and 8 is the prime 7. If $n > 5$, then, in virtue of theorem 7, between n and $2n$ there are at least two prime numbers. If the greater of them is $q = 2n-1$, then the other must be $< 2n-2$, since $2n-2$, for $n > 5$, is a composite number. We then have $n < p < 2n-2$. If $q < 2n-1$, then, since $p < q$, we obtain also $n < p < 2n-2$.

Theorem 8 was conjectured by P. Bertrand in 1845 and first proved by P. Tchebycheff in 1850. The proof given above is a modification, due to L. Kalmár, of the proof by P. Erdös [1].

COROLLARY 1. *If n is a natural number > 1, then between n and $2n$ there is at least one prime number.*

Proof. In virtue of theorem 8 the corollary is true for natural numbers > 3. To verify it for $n = 2$ and $n = 3$ we check that between the numbers 2 and 4 is the prime 3 and between the numbers 3 and 6 is the prime 5.

In 1892 J. J. Sylvester [1] proved the following generalization of corollary 1:

If $n > k$, then in the sequence $n, n+1, n+2, \ldots, n+k-1$ there exists at least one number which has a prime divisor $> k$. From this corollary 1 is obtained for $n = k+1$. This generalization was proved also by I. Schur [2] in 1924. A shorter and more elementary proof of it was given by P. Erdös [2] in 1934 (cf. Erdös [15]).

COROLLARY 2. *For natural numbers $k > 1$ we have $p_k < 2^k$.*

Proof. We have $p_2 = 3 < 2^2$. If, for a natural number k, $p_k < 2^k$, then, using corollary 1, we see that between 2^k and 2^{k+1} there is at least one prime number, which is of course greater than p_k. Thus we must have $p_{k+1} < 2^{k+1}$ and, by induction, the corollary follows.

We note that corollary 2 is stronger than inequality (3) of § 9; its proof, however, is much more difficult.

COROLLARY 3. *In the factorization into primes of number $n!$ with $n > 1$ there is at least one prime factor whose exponent is* 1.

Proof. For $n = 2$ the corollary is trivially true. If $n = 2k > 1$, where k is a natural number > 1, then, by corollary 1, there exists a prime number p such that $k < p < 2k$, whence $p < n < 2p$ and consequently p is a divisor of only one of the factors of the product $1, 2, \ldots, n$. On the other hand, if $n = 2k+1$, where k is a natural number, then there exists a prime number p such that $k < p < 2k < n$, whence $2k < 2p$ and therefore $2k+1 < 2p$, i. e. $p < n < 2p$, which proves corollary 3 analogously to the previous case.

As an immediate consequence of corollary 3 we have

COROLLARY 4. *For natural numbers $n > 1$ number $n!$ is not a k-th power with $k > 1$ being a natural number.*

Now, from theorem 7 we derive

THEOREM 9. *For natural numbers $k > 3$ we have $p_{k+2} < 2p_k$.*

Proof. Let k denote a natural number > 3. We then have $p_k > p_3 = 5$. In virtue of theorem 7, between p_k and $2p_k$ there are at least two different prime numbers, but, since the least two prime numbers greater than p_k are p_{k+1} and p_{k+2}, we must have $p_{k+2} < 2p_k$ and this is what was to be proved.

We note that, conversely, theorem 9 immediately implies theorem 7. In fact, suppose that theorem 9 is true. Then if n denotes an arbitrary natural number > 6, i.e. $n \geqslant 7$, we have $p_4 = 7 \leqslant n$. Let p_k be the greatest prime number such that $p_k \leqslant n$. We then have $k > 3$ and $p_{k+1} > n$. Therefore, by theorem 9, $p_{k+2} < 2p_k \leqslant 2n$. Thus we see that between n and $2n$ there are at least two prime numbers, p_{k+1} and p_{k+2}. Thus all that remains is to verify theorem 7 for $n = 6$.

We have thus proved that theorems 7 and 9 are equivalent in the sense that one can easily be deduced from the other.

COROLLARY 1. *We have $p_{k+1} < 2p_k$ for $k = 1, 2, \ldots$*

Proof. For $k = 4, 5, \ldots$ corollary 1 follows immediately from theorem 9. We verify corollary 1 for $k = 1, 2, 3$; $p_2 = 3 < 4 = 2p_1$, $p_3 = 5 < 6 = 2p_2$, $p_4 = 7 < 10 = 2p_3$.

COROLLARY 2. *For natural numbers $k > 1$ we have $p_{k+2} < p_k + p_{k+1}$.*

Proof. For $k > 3$ the relation follows immediately from theorem 9; for, $p_{k+2} < 2p_k < p_k + p_{k+1}$ (since $p_k < p_{k+1}$). We verify that it is also true for $k = 2$ and $k = 3$. In fact, $p_4 = 7 < 3+5 = p_2+p_3$ and $p_5 = 11 < 5+7 = p_3+p_4$.

EXERCISES. **1.** Find the natural numbers n such that n is the sum of all the primes less than n.

Solution. It is clear that the least possible natural number of this kind is $5 = 2+3$. Suppose, further, that $n > 5$ and that n is the sum of all the prime numbers less than n. If p_k is the greatest prime number less than n, then, since $n > 5$, we have $p_k \geqslant 5$. Consequently $k > 2$ and $p_1+p_2+\ldots+p_k = n \leqslant p_{k+1}$. Since $k > 2$, corollary 2 of theorem 9 gives $p_{k+1} < p_{k-1}+p_k$ and consequently $p_1+p_2+\ldots+p_k < p_{k-1}+p_k$, which is clearly impossible. Thus we conclude that only number 5 satisfies the condition of the exercise.

2. Prove that if $n > 1$ and k are natural numbers, then the number $\frac{1}{n} + \frac{1}{n+1} + \ldots + \frac{1}{n+k}$ cannot be an integer.

Proof. If the number in question were an integer we would have $\frac{1}{n} + \frac{1}{n+1} + \ldots + \frac{1}{n+k} \geqslant 1$, whence, since $\frac{1}{n} + \frac{1}{n+1} + \ldots + \frac{1}{n+k} < \frac{k+1}{n}$, we would obtain $k+1 > n$, and consequently $k \geqslant n$. Let p denote the greatest prime number $\leqslant n+k$. We have $2p > n+k$; for, in view of corollary 1 of theorem 8, between p and $2p$ there is a prime q, and for $2p \leqslant n+k$, we would have $p < q < n+k$, contrary to the definition of p. Since $k \geqslant n$, we have $n+k \geqslant 2n$, and, by corollary 1, there is a prime r between n and $2n$. Hence $r < 2n \leqslant n+k$ and the definition of p implies that $r \leqslant p$. But, since $n < r$, we have $n < p \leqslant n+k < 2p$. It follows that among the summands of the sum $\frac{1}{n} + \frac{1}{n+1} + \ldots + \frac{1}{n+k}$ there is only one whose denominator is divisible by the prime p. From this we easily infer that the sum in question cannot be an integer. In fact, reducing the fraction to the same denominator $n(n+1)\ldots(n+k)$, we see that all the numerators but one are divisible by the prime p, this being also a divisor of the denominator. Thus we have proved that none of the partial sums of the harmonic series $\frac{1}{1}+\frac{1}{2}+\frac{1}{3}+\ldots$ can be an integer provided we do not take into account the trivial case where the sum consists only of the first term.

3. Prove that corollary 1 of theorem 8 is equivalent to the following assertion T:

T. *Every finite sequence of consecutive natural numbers which contains at least one prime number contains also at least one number prime to each of the remaining terms of the sequence* (cf. Zahlen [1]).

Proof. Let

(i) $$k, k+1, \ldots, l$$

be a sequence of consecutive natural numbers and p the greatest of the primes contained in this sequence. If $2p$ were $\leqslant l$, then, according to corollary 1 of theorem 8, there would exist a prime number q such that $p < q < 2p \leqslant l$, contrary to the definition of p as the greatest prime number of the sequence (i). Accordingly, we have $l < 2p$. Hence, as can easily be seen, the number p is prime to each of the numbers $1, 2, \ldots, l$ different from p, and consequently, it is of course prime to each term of (i) different from p. We have thus proved that corollary 1 to theorem 8 implies theorem T.

Now we suppose that theorem T holds. Let $n > 1$ be a natural number. According to theorem T, in the sequence

(ii) $$2,\ 3,\ \ldots,\ 2n,$$

containing the prime 2, there exists at least one number p which is prime to each of the remaining terms of the sequence. First we note that p must be a prime number. In fact, if $p = ab$, where a and b are natural numbers each > 1, then the number $a < p$ belongs to sequence (ii) and is not relatively prime to p. Further, if p were $< n$, then $2p < 2n$ and the number $2p \neq p$ would belong to (ii) and $2p$ would not be relatively prime to p. Thus we have $p > n$. But, since $p < 2n$, p belongs to sequence (ii). Moreover $p \neq 2n$ because $n > 1$ and p is a prime. From this we conclude that $n < p < 2n$. We have thus proved that theorem T implies corollary 1 of theorem 8, which, together with the first part of the proof, shows that theorem T and corollary 1 of theorem 8 are equivalent in the sense that one can easily be deduced from the other.

4. Using corollary 1 of theorem 8 prove that for natural numbers k and $n > 2^k$ the least k numbers > 1 divisible by none of the numbers $2, 3, \ldots, n$ are primes.

Proof. If $n > 2^k$, then $n^2 > 2^k n$ and, since, in virtue of corollary 1 of theorem 8, between any two consecutive terms of the sequence $n, 2n, 2^2 n, \ldots, 2^k n$ there is at least one prime number, between n and n^2 there are at least k different prime numbers. Then of course between n and n^2 there exist at least k numbers not divisible by any of the numbers $2, 3, \ldots, n$. Each of these k numbers is a prime, since, if l is such a number and $l = ab$, where a, b are natural numbers > 1, $a \leqslant b$, then we cannot have $a \leqslant n$ (since l is not divisible by any of the numbers $2, 3, \ldots, n$); thus we must have $b \geqslant a > n$, whence $l = ab > n^2$, which is impossible. From this the theorem follows at once.

§ 11. Theorem of H. F. Scherk.

THEOREM 10 (H. F. Scherk). *For every natural number n and a suitable choice of the signs $+$ or $-$ we have*

(11) $$p_{2n} = 1 \pm p_1 \pm p_2 \pm \ldots \pm p_{2n-2} + p_{2n-1}$$

and

(12) $$p_{2n+1} = 1 \pm p_1 \pm p_2 \pm \ldots \pm p_{2n-1} + 2p_{2n}.$$

These formulae were found by H. F. Scherk [1] in 1830, proof of H. F. Scherk's formulae was published by S. S. Pillai [1] in 1928. The proof that will be presented here was published by me in 1952 (Sierpiński [14]). A similar proof was published by R. Teuffel [1] in 1955.

Proof. We say that an infinite sequence $q_1, q_2, \ldots$ has property P if it is an increasing sequence of natural numbers, odd except the first term, such that

(13) $$q_1 = 2,\ q_2 = 3,\ q_3 = 5,\ q_4 = 7,\ q_5 = 11,\ q_6 = 13,\ q_7 = 17$$

and

(14) $$q_{n+1} < 2q_n \quad \text{for} \quad n = 1, 2, \ldots$$

In particular, in view of corollary 1 of theorem 9, the sequence $q_n = p_n$ (for $n = 1, 2, \ldots$) has property P. Accordingly, to prove the theorem of Scherk it is sufficient to prove that for a suitable choice of the signs formulae (11) and (12) are valid for any sequence which has property P.

LEMMA. *If $q_1, q_2, \ldots$ is an infinite sequence having property* P, *then for $n \geqslant 3$ every odd natural number $\leqslant q_{2n+1}$, is of the form $\pm q_1 \pm q_2 \pm \ldots \pm q_{2n-1} + q_{2n}$ provided the signs $+$ or $-$ are suitably chosen.*

Proof of the lemma. It follows from (13) that the lemma is true for $n = 3$, since

$$1 = -q_1+q_2+q_3-q_4-q_5+q_6, \qquad 11 = q_1-q_2-q_3-q_4+q_5+q_6,$$
$$3 = q_1-q_2-q_3+q_4-q_5+q_6, \qquad 13 = q_1-q_2+q_3+q_4-q_5+q_6,$$
$$5 = q_1+q_2+q_3-q_4-q_5+q_6, \qquad 15 = -q_1+q_2+q_3+q_4-q_5+q_6,$$
$$7 = -q_1-q_2-q_3-q_4+q_5+q_6, \qquad 17 = q_1+q_2-q_3-q_4+q_5+q_6,$$
$$9 = q_1+q_2-q_3+q_4-q_5+q_6,$$

We note, that for $n = 2$ the lemma is not true because for no choice of the signs $+$ or $-$ would give us $5 = \pm 2 \pm 3 \pm 5 + 7$.

Now suppose that the lemma is true for a natural number $n \geqslant 3$ and let $2k-1$ be an odd natural number $\leqslant q_{2n+3}$. In view of (14) we have $q_{2n+3} < 2q_{2n+2}$ and consequently $-q_{2n+2} < 2k-1-q_{2n+2} < q_{2n+2}$. Therefore for a suitable choice of the signs $+$ or $-$ we have $0 \leqslant \pm(2k-1-q_{2n+2}) < q_{2n+2}$. In virtue of (14), we have $q_{2n+2} < 2q_{2n+1}$ and consequently

$$-q_{2n+1} \leqslant \pm(2k-1-q_{2n+2}) - q_{2n+1} < q_{2n+1},$$

and, moreover, for a suitable choice of the signs $+$ or $-$ we have

(15) $$0 \leqslant \pm\{\pm(2k-1-q_{2n+2}) - q_{2n+1}\} \leqslant q_{2n+1}.$$

Each of the numbers q_{2n+1} and q_{2n+2} is odd, and so the number in the middle of inequalities (15) is an odd natural number $\leqslant q_{2n+1}$. Conse-

quently, by the use of the inductive assumption, we conclude that for a suitable choice of the signs $+$ or $-$ we have

$$\pm\{\pm(2k-1-q_{2n+2})-q_{2n+1}\} = \pm q_1 \pm q_2 \pm \dots \pm q_{2n-1} + q_{2n}.$$

Hence, if the signs $+$ or $-$ are suitably chosen, we have

$$2k-1 = \pm q_1 \pm q_2 \pm \dots \pm q_{2n} \pm q_{2n+1} + q_{2n+2},$$

which proves the lemma for $n+1$ and at the same time, by induction, for all natural numbers $n \geqslant 3$.

COROLLARY. *For a suitable choice of the signs $+$ or $-$ we have*

$$q_{2n+1} = \pm q_1 \pm q_2 \pm \dots \pm q_{2n-1} + q_{2n}. \tag{16}$$

Proof of the corollary. Since q_{2n+1} is an odd natural number, then for $n \geqslant 3$ formula (16) follows immediately from the lemma. For $n = 1$ and $n = 2$ a straightforward computation shows that, in virtue of (13), $q_3 = q_1 + q_2$ and $q_5 = q_1 - q_2 + q_3 + q_4$.

Now, we are going to prove formulae (11) and (12).

Proof of formula (12). For $n \geqslant 3$ the number $q_{2n+1} - q_{2n} - 1$ is, by (14), an odd natural number $< q_{2n+1}$. Therefore, applying the lemma, we see that for a suitable choice of the signs $+$ or $-$ we have $q_{2n+1} - q_{2n} - 1 = \pm q_1 \pm q_2 \pm \dots \pm q_{2n-1} + q_{2n}$, whence (with $q_i = p_i$, $i = 1, 2, \dots$) formula (12) follows. For $n = 1$ and $n = 2$ a straightforward computation shows that $q_3 = 1 - q_1 + 2q_2$, $q_5 = 1 - q_1 + q_2 - q_3 + 2q_4$. Formula (12) is thus proved for all natural numbers n.

Proof of formula (11). In virtue of (14) we have $q_{2n+2} < 2q_{2n+1}$ and we see that $q_{2n+2} - q_{2n+1} - 1$ is an odd natural number > 0 and $< q_{2n+1}$. Now, applying the lemma, we see that for $n \geqslant 3$ and a suitable choice of the signs $+$ or $-$ we have

$$q_{2n+2} - q_{2n+1} - 1 = \pm q_1 \pm q_2 \pm \dots \pm q_{2n-1} + q_{2n},$$

whence

$$q_{2n+2} = 1 \pm q_1 \pm q_2 \pm \dots \pm q_{2n-1} + q_{2n} + q_{2n+1}. \tag{17}$$

Moreover, by (13), we see that

$$q_2 = 1 + q_1, \quad q_4 = 1 - q_1 + q_2 + q_3, \quad q_6 = 1 + q_1 - q_2 - q_3 + q_4 + q_5,$$

which proves formula (17) for $n = 0, 1$ and 2. Consequently formula (17) is valid for $n = 0, 1, 2, \dots$ and therefore (for $q_i = p_i$, $i = 1, 2, \dots$) formula (11) holds for $n = 1, 2, 3, \dots$ The theorem of Scherk is thus proved.

§ 12. Theorem of H. E. Richert.

LEMMA 1. *If $m_1, m_2, \dots$ is an infinite increasing sequence of natural numbers such that for a certain natural number k the inequality*

$$m_{i+1} \leqslant 2m_i \quad \text{for} \quad i > k \tag{18}$$

holds, and if there exist an integer $a \geqslant 0$ and natural numbers r and $s_{r-1} \geqslant m_{k+r}$ such that each of the numbers

$$a+1, \quad a+2, \quad \dots, \quad a+s_{r-1} \tag{19}$$

is the sum of different numbers of the sequence $m_1, m_2, \dots, m_{k+r-1}$, then for $s_r = s_{r-1} + m_{k+r}$ each of the numbers

$$a+1, \quad a+2, \quad \dots, \quad a+s_r \tag{20}$$

is the sum of different numbers of the sequence $m_1, m_2, \dots, m_{k+r}$ and, moreover, $s_r \geqslant m_{k+r+1}$.

Proof of lemma 1. Suppose that the conditions of the lemma are satisfied. Let n denote a natural number of sequence (20). If $n \leqslant a + s_{r-1}$, then there is nothing to prove, since, by assumption, n is the sum of different terms of the sequence $m_1, m_2, \dots, m_{k+r-1}$. Suppose then that $n > a + s_{r-1}$. Since $s_{r-1} \geqslant m_{k+r}$, we have $n \geqslant a + 1 + m_{k+r}$. Consequently $n - m_{k+r} \geqslant a+1$. Moreover, since n is a term of sequence (20), we have $n \leqslant a + s_r = a + s_{r-1} + m_{k+r}$. So $n - m_{k+r} \leqslant a + s_{r-1}$. Therefore the number $n - m_{k+r}$ is a term of sequence (19) and, consequently, it is the sum of different numbers of the sequence $m_1, m_2, \dots, m_{k+r-1}$. It follows that n is the sum of different numbers of the sequence $m_1, m_2, \dots, m_{k+r}$. Further, in virtue of (18), we have $m_{k+r+1} \leqslant 2m_{k+r}$, so $s_r = s_{r-1} + m_{k+r} \geqslant 2m_{k+1} \geqslant m_{k+r+1}$. The lemma is thus proved.

LEMMA 2. *If $m_1, m_2, \dots$ is an infinite sequence of natural numbers such that formula (18) holds for a natural number k and if there exist an integer $a \geqslant 0$ and a natural number $s_0 \geqslant m_{k+1}$ such that each of the numbers*

$$a+1, \quad a+2, \quad \dots, \quad a+s_0 \tag{21}$$

is the sum of different terms of the sequence $m_1, m_2, \dots, m_k$, then every natural number $> a$ is the sum of different terms of the sequence $m_1, m_2, \dots$

Proof of lemma 2. Suppose that the conditions of the lemma are satisfied. Applying lemma 1 with $r = 1, 2, \dots, l$ succesively, l being a natural number, we conclude that each of the numbers

$$a+1, \quad a+2, \quad \dots, \quad a+s_l \tag{22}$$

is the sum of different terms of the sequence $m_1, m_2, \dots, m_{k+l}$. But since $s_r > s_{r-1}$, $r = 1, 2, \dots, l$, we see that for every natural number n there exist a natural number l such that $n \leqslant a + s_l$. Consequently, every natural number $n > a$ is one of the numbers of the sequence (22), pro-

vided the number l is suitably chosen, accordingly, it is the sum of different terms of the infinite sequence $m_1, m_2, \ldots$ The lemma is thus proved.

Now, let $m_i = p_i$ with $i = 1, 2, \ldots$ In virtue of corollary 1 of theorem 9, the conditions of lemma 2 are satisfied for $a = 6$, $s_0 = 13$, $k = 5$; this is because $13 = p_6$ and each of the numbers $7, 8, \ldots, 19$ is the sum of different prime numbers $\leqslant p_5$.

In fact, $7 = 2+5$, $8 = 3+5$, $9 = 2+7$, $10 = 3+7$, $11 = 11$, $12 = 5+7$, $13 = 2+11$, $14 = 3+11$, $15 = 2+5+7$, $16 = 5+11$, $17 = 2+3+5+7$, $18 = 8+11$, $19 = 3+5+11$. Of course we do not exclude the trivial sums consisting of one term only: number 11 is not the sum of two or more different primes. As a corollary to lemma 2 we obtain

THEOREM 11. *Every natural number >6 is a sum of different prime numbers* (Richert [1], [2]).

Now suppose that $m_i = p_{i+1}$. The conditions of lemma 2 are satisfied for $a = 9$, $s_0 = 19$, $k = 6$, since $19 = p_8 = m_7$ so $s_0 = m_{6+1}$ and, moreover, each of the numbers $10, 11, \ldots, 28$ is the sum of different odd prime numbers $\leqslant m_6 = 19$. In fact, we have $10 = 3+7$, $11 = 11$, $12 = 5+7$, 13, $14 = 3+11$, $15 = 3+5+7$, $16 = 5+11$, $17 = 17$, $18 = 5+13$, $19 = 3+5+11$, $20 = 7+13$, $21 = 3+5+13$, $22 = 5+17$, $23 = 3+7+13$, $24 = 11+13$, $25 = 5+7+13$, $26 = 3+5+7+11$, $28 = 3+5+7+13$. Thus we obtain

THEOREM 12. *Every natural number $\geqslant 10$ is a sum of different odd prime numbers.*

If we admit also number 2 as a summand of the sums, we get

THEOREM 13. *Every natural number $\geqslant 12$ is a sum of two or more different prime numbers.*

As can easily be seen, number 11 is not a sum of two or more different prime numbers. Number 17 is not a sum of two or three different prime numbers (but $17 = 2+3+5+7$). One can also prove, using elementary methods only, that there exist infinitely many odd numbers which are not the sums of less than three prime numbers.

Here are four theorems due to A. Mąkowski [6]:

Every natural number >55 is a sum of different prime numbers of the form $4k-1$.

Every natural number >121 is a sum of different prime numbers of the form $4k+1$.

Every natural number >161 is a sum of different prime numbers of the form $6k-1$.

Every natural number >205 is a sum of different prime numbers of the form $6k+1$.

The lower bounds given in the theorems are sharp, i.e. cannot be replaced by still lower ones.

§ 13. A conjecture on prime numbers. Several years ago I formulated the following conjecture P.

CONJECTURE P. *If the numbers $1, 2, 3, \ldots, n^2$ with $n > 1$ are arranged in n rows each containing n numbers:*

$$
\begin{array}{rrrrr}
1, & 2, & 3, & \ldots, & n \\
n+1, & n+2, & n+3, & \ldots, & 2n \\
2n+1, & 2n+2, & 2n+3, & \ldots, & 3n \\
\ldots & \ldots & \ldots & \ldots & \ldots \\
(n-1)n+1, & (n-1)n+2, & \ldots, & & n^2
\end{array} \tag{23}
$$

then each row contains at least one prime number (Schinzel et Sierpiński [3]).

The first row of table (23) contains of course ($n > 1$) number 2. The assertion that for $n > 1$ the second row contains at least one prime number is another formulation of corollary 1 to theorem 8. It easily follows from the inequality of J. B. Rosser and L. Schoenfeld (cf. § 15) that for $n > e^k$ each of the first k rows contains a prime.

As can be verified on the basis of the tables of A. E. Western [1] and D. H. Lehmer [10] conjecture P is true for $1 < n \leqslant 4500$. Since the last two rows of table (23) consist of numbers $(n-1)^2, (n-1)^2+1, \ldots, n^2$, conjecture P implies that between two consecutive squares of natural numbers there are at least two prime numbers. Further, since in every interval whose end-points are cubes of two consecutive natural numbers there are two squares of two consecutive natural numbers, we see that conjecture P implies that between the cubes of any two consecutive natural numbers there are at least two prime numbers. The last statement has not been proved yet, but it follows from the results of A. E. Ingham of 1937 that the number of primes between n^3 and $(n+1)^3$ tends to infinity with n.

As an immediate consequence of conjecture P we obtain the assertion that between any two triangular numbers there is at least one prime number. Namely, if we arrange natural numbers in rows in such a manner that in the nth row we put n consecutive natural numbers, i.e. if we form the table

$$
\begin{array}{r}
1 \\
2, \ 3 \\
4, \ 5, \ 6 \\
7, \ 8, \ 9, \ 10 \\
11, \ 12, \ 13, \ 14, \ 15 \\
\ldots\ldots\ldots\ldots
\end{array}
$$

then each but the first of its rows contains a prime number. We do not know whether the above statement is true.

In 1932 R. Haussner [1] formulated a conjecture that, *for a natural number k, between two consecutive multiplies of the prime number p_k both less than p_{k+1}^2 there is at least one prime number*. This conjecture was verified by Haussner for prime numbers $p_k < 100$. Conjecture P for a prime n is an immediate consequence of the conjecture of Haussner. As has been noticed by L. Skula, conjecture P implies that for every natural number $n > 1$ also $(n+1)$-th row and $(n+2)$-th row of table (23) contain at least one prime number each.

In fact, it follows from conjecture P for the number $n+1$ that among the numbers $n^2-1, n^2, \ldots, n(n+1)$ there is at least one prime number and, since for $n > 2$ the first two terms of the sequence are composite numbers, at least one prime number is to be found among the numbers $n^2+1, n^2+2, \ldots, (n+1)n$. (This of course is also true for $n = 2$.) It follows from conjecture P for the number $n+1$ that among the numbers $n^2+n+1, n^2+n+2, \ldots, (n+1)^2$ there is at least one prime number; thus, clearly, there is at least one prime number among the numbers $n^2+n-1, n^2+n, \ldots, n^2+2n$ (since the number $(n+1)^2$ is composite).

With reference to table (23) is should be mentioned that A. Schinzel has formulated a conjecture that, *if n is a natural number >1 and k a natural number less than n and relatively prime to n, then in the k-th column of table* (23) *there is at least one prime number*. In other words, if k and n are natural numbers relatively prime and $k < n$, then among the numbers

$$k,\ k+n,\ k+2n,\ \ldots,\ k+(n-1)n$$

there is at least one prime number. A. Gorzelewski has verified this for the natural numbers $n \leqslant 100$ (cf. Schinzel [15]).

It is necessary to note here that Yu. V. Linnik proved in 1947 the existence of a constant C such that if $(k, n) = 1$ and $1 \leqslant k < n$ the least prime number in the arithmetical progression $k, k+n, k+2n, \ldots$ is less than n^C. Pan-Cheng-Tun [1] has calculated that $C \leqslant 10000$ (cf. also H. Fluch [1]).

As observed by A. Schinzel [15], a conjecture somewhat stronger than conjecture P can be formulated. Namely, one can conjecture that, if x is a real number $\geqslant 117$, then between x and $x+\sqrt{x}$ there is at least one prime number. This conjecture, P_1, follows from A. E. Western and D. H. Lehmer's tables for $117 \leqslant x \leqslant 20.3 \cdot 10^6$. It was Legendre who formulated the conjecture that for sufficiently large numbers x there is at least one prime number between x and $x+\sqrt{x}$.

We now show how conjecture P for $n \geqslant 117$ is derived from conjecture P_1. Let n denote a natural number $\geqslant 117$ and let k be a natural number less than n. We have $kn \geqslant 117$ and so, by conjecture P_1, there exists a prime number p such that $kn < p < kn+\sqrt{kn}$. But, since $k < n$, we have $\sqrt{kn} < n$; consequently there exists at least one prime number among the terms of the sequence $kn+1, kn+2, \ldots, (k+1)n$. Since this is valid for every natural number $k < n$, we see that (for $n \geqslant 117$) in each row of table (23) from the second onwards there is at least one prime number. (In the first row, however, for $n > 1$ at least the prime 2 occurs.) Thus we see that conjecture P for $n \geqslant 117$ follows from conjecture P_1. For $n < 117$ conjecture P has been proved by a straightforward verification.

As observed by A. Schinzel [15] a still stronger conjecture than P_1 can be formulated, namely that for each real number $x \geqslant 8$ between x and $x+(\log x)^2$ at least one prime number occurs. If we set $x = p_n$ with $n > 4$, then we obtain the inequality $p_{n+1}-p_n < (\log p_n)^2$ for all $n > 4$. It was H. Cramér [1] who conjectured that $\overline{\lim}(p_{n+1}-p_n)/(\log p_n)^2 = 1$.

There is another conjecture about the difference of two consecutive prime numbers, namely the following conjecture of N. L. Gilbreath formulated in 1958. We form a table of natural numbers in this manner: in the first row we write the differences of consecutive prime numbers (i.e. the numbers $p_{n+1}-p_n$, $n = 1, 2, \ldots$), in the second row we write the modules of the differences of the consecutive numbers of the first row. In each of the following rows we write the modules of the differences of the consecutive terms of the preceding row. The conjecture of Gilbreath is that the first term of each row is equal to 1.

Here is an example of the first 10 rows obtained in this way:

1, 2, 2, 4, 2, 4, 2, 4, 6, 2
1, 0, 2, 2, 2, 2, 2, 2, 4
1, 2, 0, 0, 0, 0, 0, 2
1, 2, 0, 0, 0, 0, 2
1, 2, 0, 0, 0, 2
1, 2, 0, 0, 2
1, 2, 0, 2
1, 2, 2
1, 0
1

The conjecture of Gilbreath has been verified for the first 63418 rows with the aid of the electronic computer SWAC. In general it has not been proved as yet (cf. Killgrove and Ralston [1]).

§ 14. Inequalities for the function $\pi(x)$. Now we are going to deduce some corollaries from lemma 9 of § 10. Since R_n denotes the product of prime numbers p such that $n < p \leqslant 2n$ and the number of such primes

is $\pi(2n)-\pi(n)$ (and by corollary 1 of theorem 8, § 10, for every natural number n at least one such prime p exists) and, moreover, each of those primes is less than $2n$, then $R_n \leqslant (2n)^{\pi(2n)-\pi(n)}$. It follows from formula (10) of § 10 that for natural numbers $n \geqslant 98$ we have

$$(2n)^{\pi(2n)-\pi(n)} > \frac{4^{n/3}}{2\sqrt{n}(2n)^{\sqrt{n}/2}};$$

taking the logarithm of each side of the last inequality, we conclude that for $n \geqslant 98$

$$\pi(2n)-\pi(n) > \frac{n}{3\log 2n}\left(\log 4 - \frac{3\log 4n}{2n} - \frac{3\log 2n}{2n}\right) \tag{24}$$

holds. But, as we know,

$$\lim_{x\to\infty} \frac{\log x}{x} = 0;$$

therefore

$$\lim_{n\to\infty}\left(\pi(2n)-\pi(n)\right) = +\infty.$$

It follows that for every natural number k there exists a natural number m_k such that for $n \geqslant m_k$ there are at least k prime numbers between n and $2n$.

Further, since $\log x/x$ is, for $x > e$, a decreasing function of x, we have for $n \geqslant 2500$

$$\frac{3\log 4n}{2n} + \frac{3\log 2n}{2n} = 6\left(\frac{\log 4n}{4n} + \frac{\log\sqrt{2n}}{\sqrt{2n}}\right)$$

$$\leqslant 6\left(\frac{\log 4\cdot 2500}{4\cdot 2500} + \frac{\log\sqrt{2\cdot 2500}}{\sqrt{2\cdot 2500}}\right) < 0{,}37;$$

hence

$$\log 4 - \frac{3\log 4n}{2n} - \frac{3\log 2n}{\sqrt{2n}} > 1{,}38 - 0{,}37 > 1. \tag{25}$$

In virtue of (24), formula (25) gives the inequality of Finsler,

$$\pi(2n)-\pi(n) > \frac{n}{3\log 2n}, \tag{26}$$

which holds not only for $n \geqslant 2500$ but, as can easily be verified, for all natural numbers $n > 1$.

It is even easier to obtain the second inequality of Finsler. We note that for natural numbers n we have $\binom{2n}{n} < 4^n$ (this follows immediately

from the binomial formula applied to $(1+1)^{2n} > \binom{2n}{n}$). In view of the relation $R_n \mid \binom{2n}{n}$ we see that $R_n < 4^n$ and from the definition of the number R_n we infer that $R_n \geqslant n^{\pi(2n)-\pi(n)}$. Consequently, $n^{\pi(2n)-\pi(n)} < 4^n$ and hence

$$\pi(2n)-\pi(n) < \frac{n\log 4}{\log n} < \frac{7n}{5\log n}$$

since, as can easily be verified, $\log 4 < \frac{7}{5}$. From this, using (26), we get (cf. Finsler [1] and Trost [3], Satz 32)

$$\frac{n}{3\log 2n} < \pi(2n)-\pi(n) < \frac{7n}{5\log n} \quad \text{for} \quad n > 1. \tag{27}$$

It follows from (27) that

$$\pi(2n) > \frac{n}{3\log 2n} \quad \text{for} \quad n > 1$$

and, since for $n \geqslant 4$ we have $n > n/2 \geqslant [n/2] > n/2-1 > n/4$, and since $\log(2[n/2]) \leqslant \log n$, we see that

$$\pi(n) \geqslant \pi\left(2\left[\frac{n}{2}\right]\right) > \frac{[n/2]}{3\log^2[n/2]} > \frac{n}{12\log n} \quad \text{for} \quad n \geqslant 4,$$

$$\pi(n) > \frac{n}{12\log n} \quad \text{for} \quad n > 1; \tag{28}$$

for, as can easily be verified, the inequality holds for $n = 2$ and $n = 3$ as well.

We are going to prove that

$$\pi(2^k) < \frac{2^{k+1}}{k\log 2} \quad \text{holds for natural numbers } k. \tag{29}$$

As can easily be seen, formula (29) holds for natural numbers $k \leqslant 6$ because $\log 2 < 1$. Now suppose it is valid for a natural number $k \geqslant 6$. In virtue of (27) (with 2^k in place of n) and formula (29) we have

$$\pi(2^{k+1}) < \pi(2^k) + \frac{7\cdot 2^k}{5k\log 2} < \frac{2^{k+1}}{k\log 2}\left(1+\frac{7}{10}\right).$$

But, since for $k \geqslant 6$ we have $(k+1)(1+\frac{7}{10}) < 2k$,

$$\pi(2^{k+1}) < \frac{2^{k+2}}{(k+1)\log 2}$$

and thus by induction inequality (29) follows.

Now let n denote a natural number >1. There exists a natural number k such that $2^k \leqslant n < 2^{k+1}$, whence $(k+1)\log 2 > \log n$. Hence, by (29), we have

$$\pi(n) \leqslant \pi(2^{k+1}) < \frac{2^{k+2}}{(k+1)\log 2} < \frac{4n}{\log n}.$$

From this we see that

(30) $$\pi(n) < \frac{4n}{\log n} \quad \text{for the natural numbers } n > 1.$$

By replacing n by p_n in (28) and (30) and by the fact that $\pi(p_n) = n$ we obtain

$$\frac{p_n}{12\log p_n} < n < \frac{4p_n}{\log p_n};$$

consequently, since $p_n > n$ (for $n = 1, 2, \ldots$), we infer that

$$p_n > \frac{n}{4}\log p_n > \frac{n\log n}{4} \quad \text{and} \quad p_n < 12n\log p_n,$$

whence $\log p_n < \log 12 + \log n + \log\log p_n$. But, in virtue of corollary 2 to theorem 8 of § 10, we see that $p_n < 2^n$, whence $\log p_n < n\log 2$ and $\log\log p_n < \log n + \log\log 2$. Since $\log 2 < 1$, for $n \geqslant 12$ we have $n > 12\log 2$ and hence $\log n > \log 12 + \log\log 2$. Therefore, for $n \geqslant 12$, we have $\log p_n < 2\log n + \log 12 + \log\log 2 < 3\log n$. Consequently, $p_n < 36n\log n$ for all $n \geqslant 12$ and, as is easy to verify, also for $2 \leqslant n < 12$.

Thus we arrive at the final conclusion that

(31) $$\frac{n\log n}{4} < p_n < 36n\log n \quad \text{for} \quad n > 1.$$

From formula (28) we derive the following corollary:

For every natural number s there exists a natural number which can be represented as the sum of two prime numbers in more than s different ways.

Proof. Suppose that for a natural number s there is no natural number which can be represented as the sum of two prime numbers in more than s ways. Let n denote a natural number >1. Let us consider all the pairs (p, q) where p, q are prime numbers, neither of them greater than n. The number of such pairs is clearly $[\pi(n)]^2$. We divide the set of the pairs (p, q) into classes by saying that (p, q) belongs to the kth class if $p+q = k$. Since $p \leqslant n$ and $q \leqslant n$, we have $k \leqslant 2n$. By assumption, for a given $k \leqslant 2n$ in the kth class there are at most s different pairs. Since the number of all the classes is less than $2n$, the number of the pairs (p, q) is less than $2ns$.

Consequently, $[\pi(n)]^2 < 2ns$ and since, by formulae (28), $[\pi(n)]^2 > n^2/12^2(\log n)^2$, we have $2\cdot 12^2 s(\log n)^2 > n$. But, as is known, $e^x > x^3/3!$ for all $x \geqslant 0$, whence, for $x = \log n$, we have $6n > (\log n)^3$. Therefore $12^2 s(\log n)^2 > (\log n)^3$ for $n > 1$, whence $\log n < 12^2 s$ for all $n > 1$, which for sufficiently large n is not true. Consequently the assumption that for a natural number s there is no natural number which can be represented as the sum of two prime numbers in more than s ways leads to a contradiction. The corollary is thus proved. The conjecture has been formulated that the number of all possible decompositions into the sum of two primes of an even natural number increases with n to infinity.

Remark. Numbers which can be represented as sums of two primes in more than one way must be even provided we do not regard two representations as being different if they differ only in the order of the factors. In fact, if an odd number n is the sum of two primes, then of course one of them must be even, i.e. equal to 2, and consequently the other is $n-2$ and we see that the representation of n as the sum of two primes is unique apart from the order of the factors.

By a slight modification of the proof of corollary 1 one can prove that for every natural number s there exists a natural number which can be represented as the sum of three squares of prime numbers in more than s different ways. P. Erdös [4] has proved that for each natural number s there exists a natural number which is representable as the sum (resp. as the difference) of the squares of two primes in more than s different ways.

It follows immediatelv from (30) that

$$\lim_{n\to\infty} \frac{\pi(n)}{n} = 0.$$

In consequence of relation (31), for natural numbers $n > 1$ we have $\log n + \log\log n - \log 4 < \log p_n < \log n + \log\log n + \log 36$. Hence immediately

(32) $$\lim_{n\to\infty} \frac{\log p_n}{\log n} = 1.$$

Now we are going to derive a corollary from inequality (31). In virtue of (31) we have

$$\frac{1}{p_k} > \frac{1}{36\,k\log k} \quad \text{for} \quad k = 2, 3, \ldots,$$

whence for natural numbers $n > 2$ we deduce that

$$\sum_{k=2}^{n} \frac{1}{p_k} > \frac{1}{36}\sum_{k=2}^{n} \frac{1}{k\log k}.$$

But, as we know, $\log(1+x) < x$ for $0 < x < 1$, whence, for $k = 2, 3, \ldots$,

$$\log(k+1) - \log k = \log\left(1+\frac{1}{k}\right) < \frac{1}{k},$$

consequently,

$$\frac{\log(k+1)}{\log k} < 1 + \frac{1}{k\log k}$$

and

$$\log\log(k+1) - \log\log k = \log\frac{\log(k+1)}{\log k} < \log\left(1+\frac{1}{k\log k}\right) < \frac{1}{k\log k}.$$

Thus we have

$$\frac{1}{k\log k} > \log\log(k+1) - \log\log k \quad \text{for} \quad k = 2, 3, \ldots, n.$$

Hence (for natural $n > 2$) we have

$$\sum_{k=2}^{n} \frac{1}{k\log k} > \log\log(n+1) - \log\log 2 > \log\log(n+1)$$

(since $\log\log 2 < 0$).

We then have

$$\sum_{k=2}^{\infty} \frac{1}{p_k} > \frac{1}{36}\log\log(n+1).$$

From this we see that the series of the reciprocals of the consecutive prime numbers, i.e. the series

$$\tfrac{1}{2}+\tfrac{1}{3}+\tfrac{1}{5}+\tfrac{1}{7}+\tfrac{1}{11}+\tfrac{1}{13}+\ldots,$$

is divergent.

§ 15. The prime number theorem and its consequences. It follows from formulae (28) and (30) of § 14 that there exist positive numbers (e.g. $a = \frac{1}{12}$, $b = 4$) such that

$$a < \pi(n) : \frac{n}{\log n} < b$$

for natural numbers $n > 1$.

In 1896 J. Hadamard and Ch. de la Vallée Poussin proved that

$$\lim_{x\to\infty}\left(\pi(x) : \frac{x}{\log x}\right) = 1. \tag{33}$$

Nowadays owing to the new methods created by A. Selberg [1] and P. Erdös [10], this formula, known under the name of the *prime number theorem*, can be proved "elementarily", though the proof is very

complicated. We will not present it here (1). If $\pi(n) : \frac{n}{\log n} = h(n)$, then e.g. $h(10^3) = 1.159$, $h(10^4) = 1.132$, $h(10^5) = 1.104$, $h(10^6) = 1.084$, $h(10^7) = 1.071$, $h(10^8) = 1.061$, $h(10^9) = 1.053$, $h(10^{10}) = 1.048$.

A better approximation for the function $\pi(x)$ is obtained by the function

$$\int_0^x \frac{dt}{\log t}.$$

J. E. Littlewood has proved that the difference $\pi(x) - \int_0^x \frac{dt}{\log t}$ takes infinitely many positive values and infinitely many negative values for x running over all natural numbers.

Proofs of the theorem of Littlewood and of the other theorems mentioned in this chapter, which require analytical methods, can be found in a book of K. Prachar [1].

In formula (33) setting $x = p_n$, by $\pi(p_n) = n$ we obtain

$$\lim_{n\to\infty} \frac{n\log p_n}{p_n} = 1,$$

whence, by (32), we get

$$\lim_{n\to\infty} \frac{p_n}{n\log n} = 1, \tag{34}$$

and consequently we see that an approximate value for p_n is the number $n\log n$, provided n is sufficiently large.

It follows immediately from (34) that

$$\lim_{n\to\infty} \frac{p_{n+1}}{p_n} = 1.$$

J. B. Rosser [1] has proved that for all natural numbers n the inequality $p_n > n\log n$ holds.

More information about $\pi(n)$ than that can be derived from formula (33) is given by the theorem of J. B. Rosser and L. Schoenfeld [1] stating that

$$\frac{n}{\log n - \frac{1}{2}} < \pi(n) < \frac{n}{\log n - \frac{3}{2}} \tag{35}$$

for every natural number $n \geqslant 67$.

(1) Cf. e.g. Trost [3], Chapter VII: *Elementarer Beweis des Primzahlsatzes*, pp. 66-73; see also Leveque [1], vol. II, p. 229-263, chapter 7: *The prime number theorem*.

Clearly formula (33) follows at once from (35).

But even from inequality (35) we are unable to derive certain simple properties of the function $\pi(n)$. For example such is the case with the theorem of E. Landau (cf. Landau [3], vol. I, pp. 215-216) stating that $\pi(2n) < 2\pi(n)$ holds for sufficiently large numbers n, which means that there are more prime numbers in the interval $0 < x \leqslant n$ than there are in the interval $n < x \leqslant 2n$, provided n is large enough. In this connection we may ask whether for natural numbers $x > 1$ and $y > 1$ the inequality

$$\pi(x+y) \leqslant \pi(x)+\pi(y) \tag{36}$$

holds. This, clearly, would imply that the inequality $\pi(2n) \leqslant 2\pi(n)$ is valid for any natural n. Inequality (36) has been proved by A. Schinzel [15] for $\min(x, y) \leqslant 146$ and has been verified by S. L. Segal [1] for $x+y \leqslant 100000$. With reference to the function $\pi(x)$ we note that the function assigning to a pair of natural numbers k and x the number of positive integers $\leqslant x$ having precisely k prime divisors, resp. k natural divisors, has also been investigated and the formulas describing its asymptotic behaviour have been found (cf. Sathe [1], Selberg [2], resp. Leveque [1]).

Now let a and b be two real numbers such that $0 < a < b$. Since, as can easily be seen, $\lim\limits_{x\to\infty} \dfrac{\log ax}{\log bx} = 1$, by (33), we have

$$\lim_{x\to\infty} \frac{\pi(bx)}{\pi(ax)} = \frac{b}{a}.$$

Consequently, since $0 < a < b$, $\pi(bx) > \pi(ax)$, provided n is large enough. This proves the following assertion:

If a and b are two positive real numbers and $a < b$, then for sufficiently large real numbers x there is at least one prime number between ax and bx.

In particular, if $a = 1$ and $b = 1+\varepsilon$ where ε is an arbitrary positive real number, it follows that there is at least one prime number between n and $n(1+\varepsilon)$ provided n is large enough.

Now let $c_1, c_2, \ldots, c_m$ be an arbitrary finite sequence consisting of digits. Let a be the number whose digits are $c_1, c_2, \ldots, c_m$. Applying the corollary just derived from formula (33) we see that $\pi(an) < \pi\big((a+1)n\big)$ holds for sufficiently large numbers n. Consequently, there exists a natural number s such that $\pi(a\cdot 10^s) < \pi\big((a+1)\cdot 10^s\big)$. Therefore there exists a prime number p such that $a\cdot 10^s < p < (a+1)\cdot 10^s$.

Thus the first m digits of number p are identical with the corresponding digits of number a. This means that the first m digits of number p are $c_1, c_2, \ldots, c_m$. Thus, as another consequence of formula (33), we obtain the following corollary:

For an arbitrary finite sequence $c_1, c_2, \ldots, c_m$ of digits there exists a prime number whose first m digits are $c_1, c_2, \ldots, c_m$ [1].

Let x denote a real number > 0. For sufficiently large natural numbers n we have $nx > 2$; so $\pi(nx) \geqslant 1$. It follows from (34) that

$$\lim_{n\to\infty} \frac{p_{\pi(nx)}}{\pi(nx)\log\pi(nx)} = 1. \tag{37}$$

But, in virtue of (33), we have

$$\lim_{n\to\infty} \frac{\pi(nx)\log nx}{nx} = 1, \tag{38}$$

whence $\lim\limits_{n\to\infty} \big(\log\pi(nx)+\log\log nx-\log nx\big) = 0$, which proves that

$$\lim_{n\to\infty} \frac{\log\pi(nx)}{\log nx} = 1. \tag{39}$$

From formulae (37), (38) and (39) we infer that

$$\lim_{n\to\infty} \frac{p_{\pi(nx)}}{n} = 1.$$

We have thus proved that formula (33) implies the fact, observed by H. Steinhaus, that for every real number $x > 0$ there exists an infinite sequence of prime numbers $q_1, q_2, \ldots$ such that

$$\lim_{n\to\infty} \frac{q_n}{n} = x.$$

Finally, let a and b be two arbitrary real numbers such that $a < b$. It follows from the above corollary to formula (33) that, if q is a sufficiently large prime number, then there exists a prime number p such that $aq < p < bq$, whence $a < p/q < b$.

This proves that the set of the quotients p/q, p and q being prime numbers, is dense in the set of positive real numbers.

---

[1] Cf. Sierpiński [10] and Trost [3], p. 42 (theorem 20), see also Sierpiński [24]. There a stronger theorem is proved.