CHAPTER VII

REPRESENTATION OF NUMBERS BY DECIMALS IN A GIVEN SCALE

§ 1. Representation of natural numbers by decimals in a given scale. Let g be a given natural number >1. We say that a *natural number N is expressed as a decimal in the scale of g* if

(1) $$N = c_m g^m + c_{m-1} g^{m-1} + \ldots + c_1 g + c_0,$$

where m is an integer $\geqslant 0$ and c_n $(n = 0, 1, 2, \ldots, m)$ are integers with the property

(2) $$0 \leqslant c_n \leqslant g-1 \quad \text{for} \quad n = 0, 1, \ldots, m \text{ and } c_m \neq 0.$$

If each number of the sequence

(3) $$0,\ 1,\ 2,\ \ldots,\ g-1$$

is denoted by a special symbol, the symbols are called the *digits* and formula (1) can be rewritten in the form

$$N = (\gamma_m \gamma_{m-1} \ldots \gamma_1 \gamma_0)_g,$$

where γ_n is the digit which denotes the number c_n.

If $g \leqslant 10$, the digits 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 are taken as the symbols to denote the numbers of (3). For example,

$N = (10010)_2$ means $N = 1\cdot 2^4 + 0\cdot 2^3 + 0\cdot 2^2 + 1\cdot 2 + 0 = 18,$

$N = (5603)_7$ means $N = 5\cdot 7^3 + 6\cdot 7^2 + 0\cdot 7 + 3 = 2012.$

THEOREM 1. *Any natural number may be uniquely expressed as a decimal in the scale of g (g being a natural number >1), i.e. it can be rewritten in form (1), where the numbers c_n $(n = 0, 1, \ldots, m)$ are integers which satisfy inequalities (2).*

Proof. Suppose that a natural number N can be represented in form (1), where c_n $(n = 0, 1, \ldots, m)$ are integers satisfying conditions (2).

Let n denote one of the numbers $0, 1, 2, \ldots, m-1$. In virtue of (1) we have

(4) $$\frac{N}{g^n} = c_m g^{m-n} + c_{m-1} g^{m-n-1} + \ldots + c_n + \frac{c_{n-1}}{g} + \frac{c_{n-2}}{g^2} + \ldots + \frac{c_0}{g^n}.$$

But in view of (2),

$$0 \leqslant \frac{c_{n-1}}{g} + \frac{c_{n-2}}{g^2} + \ldots + \frac{c_0}{g^n} \leqslant \frac{g-1}{g} + \frac{g-1}{g^2} + \ldots + \frac{g-1}{g^n} = 1 - \frac{1}{g^n}.$$

Hence, by (4), we infer that

$$\left[\frac{N}{g^n}\right] = c_m g^{m-n} + c_{m-1} g^{m-n-1} + c_{n+1} g + c_n$$

and similarly

$$\left[\frac{N}{g^{n+1}}\right] = c_m g^{m-n-1} + c_{m-1} g^{m-n-2} + \ldots + c_{n+1}.$$

These formulae show that

(5) $$c_n = \left[\frac{N}{g^n}\right] - g\left[\frac{N}{g^{n+1}}\right] \quad \text{for any} \quad n = 0, 1, \ldots, m.$$

In virtue of (1) and (2), we also have

$$g^m \leqslant N \leqslant (g-1)(g^m + g^{m-1} + \ldots + g + 1) = g^{m+1} - 1 < g^{m+1},$$

whence $m \log g \leqslant \log N < (m+1)\log g$ and therefore

$$m \leqslant \frac{\log N}{\log g} < m+1,$$

which proves

(6) $$m = \left[\frac{\log N}{\log g}\right].$$

Formulae (6) and (5) show that if N is represented as (1) and conditions (2) are satisfied, then the numbers m and c_n $(n = 0, 1, \ldots, m)$ are uniquely defined by number N. This proves that for a given natural number N (with a fixed natural number $g > 1$) there is at most one representation (1) such that conditions (1) are satisfied.

Therefore in order to prove the theorem it is sufficient to show that for any natural number N and a natural number $g > 1$ there is at least one representation (1) (conditions (2) being satisfied).

Let N_1 and c_0 be the quotient and the remainder yielded by the division of N by g. We then have $N = c_0 + gN_1$. Replacing N by N_1 we find the quotient N_2 and the remainder c_1 from the division of N_1 by g. Continuing, we proceed similarly with N_2 in place of N_1 and so on.

It is clear that the quotients consecutively obtained, when positive, decrease because $N_{n+1} \leqslant N_n/g$. Since they are non-negative integers, for some $k \geqslant 1$ we must ultimately obtain $N_k = 0$. Let m denote the greatest index for which $N_m \neq 0$. We have the following sequence of equalities:

$$N = c_0 + gN_1, \quad N_1 = c_1 + gN_2, \quad \ldots, \quad N_{m-1} = c_{m-1} + gN_m, \quad N_m = c_m.$$

Hence we easily obtain the desired representation of N, namely $N = c_0 + c_1 g + c_2 g^2 + \ldots + c_m g^m$, where $c_m \neq 0$ because $N_m \neq 0$, and the numbers c_n $(n = 0, 1, \ldots, m)$, being remainders obtained from the division by g, satisfy condition (2).

Thus we have proved theorem 1 and, at the same time, we have found an algorithm for finding the representation of N as a decimal in the scale of g. The algorithm is the following: we divide N by g and denote the remainder by c_0 and the quotient by N_1; then we divide N_1 by g and denote the remainder by c_1 and the quotient by N_2. We proceed in this way until we obtain the quotient $N_{m+1} = 0$. This, as we have just seen, leads to a representation of N in form (1).

Since in the scale of $g = 2$ there are only two digits, 0 and 1, from theorem 1 we deduce the following

COROLLARY. *Any natural number may be uniquely expressed as the sum of different powers (the exponents being non-negative integers) of number* 2.

For example: $100 = 2^6 + 2^5 + 2^2$, $29 = 2^4 + 2^3 + 2^2 + 2^0$, $M_n = 2^n - 1 = 2^{n-1} + 2^{n-2} + \ldots + 2 + 2^0$.

EXERCISES. 1. Find the decimals in the scale of 2 of the first twelve prime numbers.

Answer: 10, 11, 101, 111, 1011, 1101, 10001, 10011, 10111, 11101, 11111, 100101.

2. Prove that for every natural number m there exists a prime whose representation as a decimal in the scale of 2 is such that the last digit is 1 and the preceding m digits are equal to zero.

Proof. By theorem 11 of Chapter VI, for a natural number m there exists a prime p of the form $2^{m+1}k + 1$, where k is a natural number. In the representation of this number as a decimals in the scale of 2, m of the last $m+1$ digits are 0 and one, at the very end, is equal to unity.

Remark. It is known that there are prime numbers whose digits in the scale of 2 are all 1. There are 23 known numbers of this kind; the greatest of them has 11213 digits (each equal ot 1) in the scale of 2. We do not know whether there exist infinitely many primes of this kind. (Clearly, they coincide with the primes of the form $2^n - 1$.) There are known primes whose decimals in the scale of 2 consist of digits all equal to zero with the exception of the first and the last digits. For example: 11, 101, 10001, 100000001 and 10000000000000001. These are all the known primes of this form, we do not know whether there exist any other such primes. They are the primes of Fermat of the form $2^{2^n} + 1$.

3. Prove that for any natural number $s > 1$ there exist at least two primes which, presented as decimals in the scale of 2, have precisely s digits.

Proof. For $s = 2$ and $s = 3$ the result follows from exercise 1. If $s \geqslant 4$, then $2^{s-1} > 5$ and, by theorem 7 of Chapter III, it follows that between 2^{s-1} and 2^s there are at least two primes. On the other hand, if n is a natural number with the property $2^{s-1} \leqslant n < 2^s$, then it has, of course, s digits in the scale of 2.

4. Prove that the last digit of the representation as a decimal in the scale of 12 of any arbitrary square is a square.

Proof. If the last digit of a natural number is $0, 1, \ldots, 11$, then the last digit (in the scale of 12) of the square of them is 0, 1, 4, 9, 4, 1, 0, 4, 9, 4, 1, respectively.

Remark. It has been proved that other scales with this property (proved above for the scale of 12) are only the numbers 2, 3, 4, 5, 8, 16. Cf. Müller [1].

5. Prove that there exist infinitely many natural numbers n that are not divisible by 10 and such that number n', obtained from n by reversing the order of the digits in the representation of n as a decimal in the scale of 10, is a divisor of n and $n:n' > 1$.

Proof. As is easy to verify, the following numbers have the desired property:

$$9899\ldots9901 = 9 \cdot 1099\ldots9989$$

and

$$8799\ldots9912 = 4 \cdot 2199\ldots9978$$

where the number of 9's in the middle is arbitrary but equal on either side of the equality.

It can be proved that the least natural number > 9 with this property is the number $8712 = 4 \cdot 2178$ and that the numbers written above exhaust the class of the numbers of this property. Cf. Subba Rao [1]. The problem whether such numbers exist had been formulated by D. R. Kaprekar.

6. Prove that any natural number may be uniquely expressed in the form

$$(*) \qquad n = a_1 \cdot 1! + a_2 \cdot 2! + \ldots + a_m \cdot m!,$$

where m is a natural number, $a_m \neq 0$ and a_j $(j = 1, 2, \ldots, m)$ are integers such that $0 \leqslant a_j \leqslant j$ for $j = 1, 2, \ldots, m$.

Proof. Suppose that a natural number n admits two representations in the form (*). We then have

$$a_1 \cdot 1! + a_2 \cdot 2! + \ldots + a_m \cdot m! = a_1' \cdot 1! + a_2' \cdot 2! + \ldots + a_m' \cdot m!.$$

Let k denote the greatest natural number such that $a_k \neq a_k'$, i.e. $a_k' > a_k$, say. Therefore $a_k' - a_k \geqslant 1$, whence

$$k! \leqslant a_k' \cdot k! - a_k \cdot k! = a_1 \cdot 1! + \ldots + a_{k-1}(k-1)! - a_1' \cdot 1! - \ldots - a_{k-1}'(k-1)!$$
$$\leqslant 1 \cdot 1! + 2 \cdot 2! + \ldots + (k-1)(k-1)! = k! - 1 < k!,$$

which is impossible.

Now let s denote a natural number. Consider all the expansions of the form (*) with $m \leqslant s$ and $0 \leqslant a_j \leqslant j$ for $j = 1, 2, \ldots, m$. As is easy to calculate, the number of them is equal to $(1+1)(2+1)\ldots(s+1) = (s+1)!$. Therefore the number of the expansions excluding those which give $n = 0$ is $(s+1)!-1$. In virtue of what we have proved above, different expansions of the form (*) give different n's. On the other hand, any expansion of the form (*) with $m \leqslant s$ produces a natural number $\leqslant 1\cdot 1! + 2\cdot 2! + \ldots + m\cdot m! = (m+1)!-1 \leqslant (s+1)!-1$. Hence, trivially, any natural number $\leqslant (s+1)!-1$ can be obtained as an n for a suitable expansion of the form (*) with $m \leqslant s$.

7. For fixed natural numbers g and s let $f(n)$ denote the sum of the sth powers of the digits in the scale of g of the natural number n. Prove that for any natural number n the infinite sequence

$$n, f(n), ff(n), fff(n), \ldots \tag{i}$$

is periodic.

Proof. Clearly, in order to show that sequence (i) is periodic it is sufficient to prove that there is a number which occurs as different terms of (i).

In other words, it is sufficient to prove that not all terms of (i) are different. Let n denote a natural number and let $n = a_0 + a_1 g + \ldots + a_{k-1} g^{k-1}$ be the representation of n as a decimal in the scale of g. We have $f(n) = a_0^s + a_1^s + \ldots a_{k-1}^s \leqslant k(g-1)^s < kg^s$. But, as we know, g^k/k increases to infinity with k; so for k large enough we have $g^k/k > g^{s+1}$. Therefore $kg^s < g^{k-1} \leqslant n$. From this we easily infer that for sufficiently large n, say for $n > m$, we have $f(n) < n$. This shows that after any term of the sequence that is greater than m there occurs a term less than the term in question. Consequently, for none of the terms all the terms that follows it are greater than m (for this would produce a decreasing infinite sequence of natural numbers). Thus we have proved that the sequence contains infinitely many terms that are not greater than m and this shows that the sequence must contain different terms that are equal, and this is what was to be proved.

Remark. For $g = 10$ and $s = 2$, Porges [1] has proved that the period of sequence (i) consists of either one term equal to 1 or the following eight terms: 4, 16, 37, 58, 89, 145, 42, 20. For example, if $n = 3$ we have the sequence 3, 9, 81, 65, 61, 37, 58, …, 16, 37, …; if $n = 5$, we have the sequence 5, 25, 29, 85, 89, 145, …, 58, 89, …; if $n = 7$, we have the sequence 7, 49, 97, 130, 10, 1, 1, 1, … A generalization of the results of Porges has been obtained by B. M. Stewart [1]. The case where $g = 10$ and $s = 3$ has been considered by K. Iséki [1]. He has proved that there are 9 possible periods of the sequence of the form (i). These are: one term periods, the term being any of the numbers 1, 153, 370, 371, 407; period consisting of two numbers, either of 136 and 244 or of 919 and 1459; finally, periods consisting of three numbers, either of 55, 250, 133 or of 160, 217, 252 (see also Iséki [2]).

K. Chikawa, K. Iséki and T. Kusakabe [1] proved that in the case where $g = 10$, $s = 4$ there are six possible periods of sequence (i). These are: periods consisting of one number, which can be any of the numbers 1, 1634, 8208, 9474; a period consisting of the numbers 2178, 6514; a period consisting of seven numbers 13139, 6725, 4338, 4514, 1138, 4179, 9219 (see also Chikawa, Iséki, Kusakabe and Shibamura [1]).

8. Prove that the period of sequence (i) of exercise 7 may begin arbitrarily far.

Proof. This follows immediately from the fact that for every natural number n there exists a natural number $m > n$ such that $f(m) = n$. In fact, for any natural number s the sum of the s-th powers of the digits (in the scale of g) of the number $m = \frac{g^n-1}{g-1}$ is n and, moreover, if $n > 1$, we have $m > n$; if $n = 1$, then we put $m = g$.

9. Find the tables of addition and multiplication of decimals in the scale of 7.

Answer:

+	1	2	3	4	5	6
1	2	3	4	5	6	10
2	3	4	5	6	10	11
3	4	5	6	10	11	12
4	5	6	10	11	12	13
5	6	10	11	12	13	14
6	10	11	12	13	14	15

×	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	4	6	11	13	15
3	3	6	12	15	21	24
4	4	11	15	22	26	33
5	5	13	21	26	34	42
6	6	15	24	33	42	51

§ 2. Representations of numbers by decimals in negative scales.

THEOREM 2. *If g is an integer < -1, then any integer N may be uniquely expressed as a decimal of form (1), where c_n $(n = 0, 1, \ldots, m)$ are integers such that*

$$0 \leqslant c_n < |g| \quad \textit{for} \quad n = 0, 1, \ldots, m \tag{7}$$

and $c_m \neq 0$.

The theorem is due to Andrzej Wakulicz and Z. Pawlak, [1], who have found it as an aid to computation with the use of electronic computers.

Proof. Let g be an integer < -1 and $x = N$ an arbitrary integer. Denote by c_0 the remainder left when x is divided by $|g_0|$. We have $0 \leqslant c_0 < |g|$ and $x = c_0 + gx_1$, where x_1 is an integer. Hence $gx_1 = x - c_0$ and so $|gx_1| \leqslant |x| + c_0 \leqslant |x| + |g| - 1$, whence $|x_1| \leqslant (|x| + |g| - 1)/|g|$. If $(|x| + |g| - 1)/|g| \geqslant |x|$, then $|x| + |g| - 1 \geqslant |g|\,|x|$, i.e. $|g| - 1 \geqslant (|g| - 1)|x|$, whence, by $|g| > 1$, we see that $|x| \leqslant 1$, so $x = 0, 1$ or -1. If $x = 0$ or $x = 1$, then $x = c_0$. If $x = -1$, then $x = |g| - 1 + g = c_0 + g$, where $c_0 = |g| - 1$. Therefore it remains to consider the case where $(|x| + |g| - 1)/|g| < |x|$. We have $|x_1| < |x|$ and we may apply the procedure which we have just applied to x, to x_1. Continuing, we proceed in this way until, after a finite number of steps, we obtain a representation of N in form (1), where c_n $(n = 0, 1, \ldots, m)$ are integers satisfying conditions (7).

In order to prove that the representation of N in form (1), conditions (7) being satisfied, is unique, it is sufficient to note that N divided by $|g|$ leaves the remainder c_0, $(N - c_0)/g$ divided by $|g|$ leaves the remainder c_1 and so on. Hence it follows that the numbers $c_0, c_1, c_2, \ldots$ are uniquely defined by number N; so the representation of N in form (1) is unique. Theorem 2 is thus proved.

Examples: $-1 = (11)_{-2}$, $10 = (11110)_{-2}$, $-10 = (1010)_{-2}$, $16 = (10000)_{-2}$, $-16 = (110000)_{-2}$, $25 = (1101001)_{-2}$, $-25 = (111011)_{-2}$, $100 = (110100100)_{-2} = (10201)_{-3}$.

§ 3. Infinite fractions in a given scale. Let g denote a natural number >1 and x a real number. Let $x_1 = x-[x]$. We have $0 \leqslant x_1 < 1$. Further, let $x_2 = gx_1-[gx_1]$, then again $0 \leqslant x_2 < 1$. Continuing, we define x_3 as $gx_2-[gx_2]$ and so on. Thus we obtain an infinite sequence x_n ($n = 1, 2, \ldots$) defined by the conditions

(8) $$x_1 = x-[x], \quad x_{n+1} = gx_n-[gx_n] \quad \text{for} \quad n = 1, 2, \ldots$$

These formulae imply

(9) $$0 \leqslant x_n < 1 \quad \text{for} \quad n = 1, 2, \ldots$$

Let

(10) $$c_n = [gx_n] \quad \text{for} \quad n = 1, 2, \ldots$$

In virtue of (9) we have $0 \leqslant gx_n < g$; therefore, by (10), $0 \leqslant c_n < g$, and, since numbers (10) are integers, we have

(11) $$0 \leqslant c_n \leqslant g-1 \quad \text{for} \quad n = 1, 2, \ldots$$

Formulae (8) and (11) give

$$x = [x]+x_1, \quad x_1 = \frac{c_1+x_2}{g}, \quad x_2 = \frac{c_2+x_3}{g}, \quad \ldots, \quad x_n = \frac{c_n+x_{n+1}}{g}.$$

Hence, for $n = 1, 2, \ldots$,

(12) $$x = [x]+\frac{c_1}{g^2}+\frac{c_2}{g^2}+\ldots+\frac{c_n}{g^n}+\frac{x_{n+1}}{g^n}.$$

Since, by (9), $0 \leqslant \frac{x_{n+1}}{g^n} < \frac{1}{g^n}$ and in virtue of $g \geqslant 2$, g^n increases to infinity with n, we see that $\lim_{n\to\infty} \frac{x_{n+1}}{g^n} = 0$. Therefore, by (12), we obtain the following expansion of number x into an infinite series:

(13) $$x = [x]+\frac{c_1}{g}+\frac{c_2}{g^2}+\frac{c_3}{g^3}+\ldots$$

where, by (11), numbers c_n are digits in the scale of g.

Thus we have proved that every real number x has a representation (at least one) in form (13) for any given natural scale $g > 1$, where numbers c_n are digits in the scale of g.

Suppose that a real number x is represented in form (13) (where c_n are integers satisfying conditions (11)). For any $n = 1, 2, \ldots$ we set

(14) $$r_n = [x]+\frac{c_1}{g}+\frac{c_2}{g^2}+\ldots+\frac{c_n}{g^n}.$$

We have

$$x-r_n = \frac{c_{n+1}}{g^{n+1}}+\frac{c_{n+2}}{g^{n+2}}+\ldots,$$

whence, by (11),

$$0 \leqslant x-r_n \leqslant \frac{g-1}{g^{n+1}}+\frac{g-1}{g^{n+2}}+\ldots = \frac{1}{g^n},$$

the equality $x-r_n = 1/g^n$ being possible only in the case where $c_{n+1} = c_n = \ldots = g-1$, i.e. where all the digits of the representation are equal to $g-1$ from a certain n onwards. Then $x = r_n+1/g^n$, and so, by (14), x is the quotient of an integer by a power of number g. If m is the least natural number such that $c_n = g-1$ for $n \geqslant m$, then in the case of $m = 1$, by (13), we would have $x = [x]+1$, which is impossible. If, however, $m > 1$, then $c_{m-1} \neq g-1$, therefore, by (11), $c_{m-1} < g-1$, that is, $c_{m-1} \leqslant g-2$, which shows that number $c'_{m-1} = c_{m-1}+1$ is also a digit in the scale of g; consequently number x has a representation

$$x = [x]+\frac{c_1}{g}+\frac{c_2}{g^2}+\ldots+\frac{c_{m-2}}{g^{m-2}}+\frac{c'_{m-1}}{g^{m-1}}+\frac{0}{g^m}+\frac{0}{g^{m+1}}+\ldots,$$

which is different from (13).

It is easy to prove that, conversely, if x is the quotient of an integer by a power of number g, then x has two different representations in form (13), where c_n are integers satisfying conditions (11). In one of them all c_n's except a finite number are equal to zero, in the other from a certain n onwards all c_n's are equal to $g-1$.

If a real number x is not the quotient of an integer by a power of number g, then

$$0 \leqslant x-r_n < \frac{1}{g^n} \quad \text{for} \quad n = 1, 2, \ldots,$$

whence $0 \leqslant g^n x-g^n r_n < 1$. Hence, since by (14) number $g^n r_n$ is an integer, we see that $g^n r_n = [g^n x]$, this being also true for $n = 0$ provided r_0 is defined as $[x]$. We then have

(15) $$g^n r_n = [g^n x] \quad \text{and} \quad g^{n-1} r_{n-1} = [g^{n-1} x] \quad \text{for} \quad n = 1, 2, \ldots$$

But, in view of (14), $r_n - r_{n-1} = \frac{c_n}{g^n}$ for any $n = 1, 2, \ldots$, whence $c_n = g^n r_n - g g^{n-1} r_{n-1}$ which, by (15), implies

(16) $$c_n = [g^n x] - g[g^{n-1}x], \quad n = 1, 2, \ldots$$

This shows that any real number x which is not the quotient of an integer by a power of g has precisely one representation as series (13), where c_n are integers satisfying conditions (11). This representation is denoted by

(17) $$x = [x] + (0, c_1 c_2 c_3 \ldots)_g.$$

Formula (16), which gives the nth digit, is simple; however, it is not easy in general to compute the value of its right-hand side. For example, for $g = 10$ formula (16) gives for the 1000th digit of the decimal of $\sqrt{2}$ the value $c_{1000} = [10^{1000}\sqrt{2}] - 10[10^{999}\sqrt{2}]$, which is not easy to calculate.

We have just proved that in order to obtain the representation of a real number as a decimal (17) we may apply the following algorithm: $x_1 = x - [x]$, $c_1 = [gx_1]$, $x_2 = gx_1 - c_1$, $c_2 = [gx_2]$, $x_3 = gx_2 - c_2, \ldots, x_n = gx_{n-1} - c_{n-1}$, $c_n = [gx_n], \ldots$

We have also proved that representation (13) is finite (i.e. all its digits are zero from a certain n onwards) if and only if x is the quotient of an integer by a power of number g. It is easy to prove that this condition is equivalent to saying that x is a rational number equal to an irreducible fraction whose denominator is a product of primes each of which is a divisor of g. The necessity of this condition is evident. On the other hand, if $x = l/m$, where l is an integer and m a natural number such that any prime divisor of m is a divisor of g, then, if $g = q_1^{a_1} q_2^{a_2} \ldots q_s^{a_s}$ denotes the factorization of g into primes, $m = q_1^{\lambda_1} q_2^{\lambda_2} \ldots q_s^{\lambda_s}$, where $\lambda_1, \lambda_2, \ldots, \lambda_s$ are non-negative integers. Let k be a natural number such that $ka_i \geqslant \lambda_i$ for any $i = 1, 2, \ldots, s$. Then $m \mid g^k$, so $g^k = hm$, where h is a natural number. Hence $x = l/m = hl/g^k$, which gives the sufficiency of the condition.

Thus we see that if a real number x is not a rational number which is an irreducible fraction with a denominator such that any prime divisor of it divides g, then number x has precisely one representation in form (13), where c_n $(n = 1, 2, \ldots)$ are digits in the scale of g. Moreover, the representation is infinite and has infinitely many digits different from $g-1$. The representation is to be obtained by the use of the algorithm presented above.

The algorithm for representing a real number x as a decimal may also be applied in the case where g is a real number > 1. Then formulae (8), (9), (10) and (12) are still valid. However, the only proposition about c_n's $(n = 1, 2, \ldots)$ which remains true is that they satisfy the inequalities $0 \leqslant c_n < g$ and that they are integers. For example, for $g = \sqrt{2}$, $x = \sqrt{2}$ the representation given by the algorithm is

$$\sqrt{2} = 1 + \frac{1}{(\sqrt{2})^3} + \frac{1}{(\sqrt{2})^9} + \frac{1}{(\sqrt{2})^{12}} + \frac{1}{(\sqrt{2})^{21}} + \ldots$$

However, there is also another representation of $\sqrt{2}$ in the form (13). This is

$$\sqrt{2} = \frac{1}{\sqrt{2}} + \frac{1}{(\sqrt{2})^3} + \frac{1}{(\sqrt{2})^5} + \frac{1}{(\sqrt{2})^7} + \ldots$$

For $g = \sqrt{2}$ and $x = (2\sqrt{2}+1)/4$ we have two representations in the form (13):

$$\frac{2\sqrt{2}+1}{4} = \frac{1}{\sqrt{2}} + \frac{1}{(\sqrt{2})^6} + \frac{1}{(\sqrt{2})^8} + \ldots = \frac{1}{\sqrt{2}} + \frac{1}{(\sqrt{2})^4} + \ldots,$$

the latter being given by the algorithm. We also have

$$\frac{2}{1} + \frac{\sqrt{2}}{4} = \frac{1}{(\sqrt{2})^4} + \frac{1}{(\sqrt{2})^5} + \frac{1}{(\sqrt{2})^6} + \ldots = \frac{1}{(\sqrt{2})^2} + \frac{1}{(\sqrt{2})^5} + \frac{1}{(\sqrt{2})^7} + \ldots$$

where the second representation is given by the algorithm. See also Gelfond [1].

§ 4. Representations of rational numbers by decimals. Now let x be a rational number which is equal to an irreducible fraction $1/m$ and suppose that the representation of x as a decimal is of the form (13), where c_n $(n = 1, 2, \ldots)$ are digits in the scale of g where g is an integer > 1. Let x_n $(n = 1, 2, \ldots)$ be numbers defined by formulae (8). Then, as we know, formulae (9) and (10) hold. In virtue of (8) we have $mx_1 = l - [x]$. Consequently mx_1 is a natural number and, since, by (8), we have $mx_{n+1} = gmx_n - m[gx_n]$ for any $n = 1, 2, \ldots$, then, by induction, we infer that all the numbers mx_n are integers and, moreover, by (9), that they satisfy the inequalities $0 \leqslant mx_n < m$ for $n = 1, 2, \ldots$ If for some n we have $x_n = 0$, then, by (8), $x_j = 0$ for all $j \geqslant n$. Hence, by (10), $c_j = 0$ for $j \geqslant n$ and representation (13) for x is finite. Further, suppose that $x_n \neq 0$ for all $n = 1, 2, \ldots$ We then have $0 < mx_n < m$ for $n = 1, 2, \ldots$ and so the numbers $mx_1, mx_2, \ldots, mx_m$ can take only $m-1$ different values $1, 2, \ldots, m-1$. It follows that there exist natural numbers h and s such that $h+s \leqslant m$ and $mx_h = mx_{h+s}$, which, by (8), proves that $x_n = x_{n+s}$ for $n > h$ and therefore, by (10), $c_n = c_{n+s}$ for $n \geqslant h$. This proves that the infinite sequence of digits in (17) is periodic. We have thus proved the following theorem:

THEOREM 3. *The representation of a rational number in form* (13), *where g is a natural number greater than* 1, *is periodic. The number of digits in the period, as well as the number, not less than* 0, *of digits preceding the period, is less than the denominator of the rational number in question.*

Consider an arbitrary infinite sequence $c_1, c_2, \ldots$, where c_n $(n = 1, 2, \ldots)$ are digits in the scale of g. Then the c_n's satisfy condition (11), whence it follows that infinite series (13) is convergent and its sum x is a real number. It follows from theorem 3 that, if the sequence $c_1, c_2, \ldots$ is not periodic, then x is an irrational number. In order to prove the converse it is sufficient to show that if a sequence of digits $c_1, c_2, \ldots$ is periodic, then number (17) is rational.

Suppose then that the sequence $c_1, c_2, \ldots$ is periodic. This means that for some natural numbers s and h the equality $c_{n+s} = c_n$ holds whenever $n \geqslant h$. We then have

$$\begin{aligned}
&\frac{c_1}{g} + \frac{c_2}{g^2} + \ldots \\
&= \frac{c_1}{g} + \frac{c_2}{g^2} + \frac{c_{h-1}}{g^{h-1}} + \frac{c_h}{g^h} + \frac{c_{h+1}}{g^{h+1}} + \ldots + \\
&\qquad + \frac{c_{h+s-1}}{g^{h+s-1}} + \frac{c_h}{g^{h+s}} + \frac{c_{h+1}}{g^{h+s+1}} + \ldots + \frac{c_{h+s-1}}{g^{h+2s-1}} + \ldots \\
&= \frac{c_1}{g} + \frac{c_2}{g^2} + \ldots + \frac{c_{h-1}}{g^{h-1}} + \left(\frac{c_h}{g^h} + \frac{c_{h+1}}{g^{h+1}} + \ldots + \frac{c_{h+s-1}}{g^{h+s-1}}\right)\left(1 + \frac{1}{g^s} + \frac{1}{g^{2s}} + \ldots\right) \\
&= \frac{c_1}{g} + \frac{c_2}{g^2} + \ldots + \frac{c_{h-1}}{g^{h-1}} + \left(\frac{c_h}{g^h} + \frac{c_{h+1}}{g^{h+1}} + \ldots + \frac{c_{h+s-1}}{g^{h+s-1}}\right)\frac{g^s}{g^s - 1} \\
&= \frac{c_1 g^{h+s-2} + c_2 g^{h+s-3} + \ldots + c_{h-1} g^s + c_h g^{s-1} + \ldots + c_{h+s-1}}{g^{h-1}(g^s - 1)} - \\
&\qquad - \frac{c_1 g^{h-1} + c_2 g^{h-2} + \ldots + c_{h-1}}{g^{h-1}(g^s - 1)} \\
&= \frac{(c_1 c_2 \ldots c_{h+s-1})_g}{g^{h-1}(g^s - 1)} - \frac{(c_1 c_2 \ldots c_{h-1})_g}{g^{h-1}(g^s - 1)}.
\end{aligned}$$

Thus we see that the sum of the series is a rational number

$$\frac{(c_1 c_2 \ldots c_{h+s-1})_g - (c_1 c_2 \ldots c_{h-1})_g}{g^{h-1}(g^s - 1)}.$$

This, however, in the form as is written, is not necessarily an irreducible fraction.

This formula may also serve as a *rule for reducing periodic fractions* in a given scale of $g > 1$.

We have thus proved the following

THEOREM 3ª. *For a given scale of $g > 1$, where g is a natural number, the numbers which admit representations in form* (13) *such that the sequence of the digits is periodic are precisely rational numbers* (finite representations are understood to be periodic, the period consisting of one number being equal either to 0 or to $g-1$).

As an immediate consequence of theorem 3ª we note that if a number x has a non-periodic representation as a decimal in a scale of g, then x is irrational. On the basis of this fact it is easy to prove that a number a whose decimal (in the scale of 10) is obtained by writing 0 for the integer and the consecutive natural numbers to the right of the decimal point, i.e. number

$$a = 0.12345678910111121314\ldots,$$

is an irrational number. In fact, if the decimal of a were recurring, then, since all numbers 10^n $(n = 1, 2, \ldots)$ occur in it, arbitrarily long sequences consisting of 0's would appear; consequently, the period would necessarily consist of number 0 only. But this is impossible since infinitely many 1's occur in the decimal.

EXERCISES. 1. Write the number $\frac{1}{99^2}$ as a decimal.

Answer:

$$\frac{1}{99^2} = 0.\dot{0}\dot{0}\dot{0}\dot{1}\dot{0}\dot{2}\dot{0}\dot{3}\ldots\dot{0}\dot{8}\dot{0}\dot{9}\dot{1}\dot{0}\dot{1}\dot{1}\dot{1}\dot{2}\ldots\dot{9}\dot{6}\dot{9}\dot{7}\dot{9}\dot{9}.$$

(the dots above the digits indicate that the digits form the period). The period (which starts exactly at the decimal point) is obtained by writing down all the natural numbers from 0 to 99 excluding 98 written as decimals. As proved by J. W. L. Glaisher [1] a more general formula holds

$$\frac{1}{(g-1)^2} = (0.\dot{0}\dot{1}\dot{2}\dot{3}\ldots\dot{\overline{g-3}}\,\dot{\overline{g-1}})_g$$

whenever g is a natural number > 2.

2. Using the representation of number e as a decimal $e = 2.718281828\ldots$ write number e as a decimal in the scale of 2 up to the 24th decimal place.

Answer:

$$e = (10.101101111110000101010001\ldots)_2.$$

This representation has been given by G. Peano [1]. He writes a point and an exclamation mark in place of 0 and 1, respectively; therefore this equality has the form

$$e = (!.,!.!!.!!!!!!....!.!.!...!)_2.$$

3. Using the representation of number π as a decimal $\pi = 3.14159265\ldots$ write the number π as a decimal in the scale of 2 up to the 24th decimal place.

Answer:

$$\pi = (11.001001000011111101101010\ldots)_2$$

(cf. G. Peano [1], p. 177).

4. Prove that in any infinite decimal fraction there are arbitrarily long sequences of digits that appear infinitely many times.

Proof. Let $0.c_1c_2c_3\ldots$ denote an infinite decimal fraction and m a natural number. Consider all the sequences that consist of m digits which appear in the sequence $c_1c_2\ldots$, i.e. all the sequences

(18) $$c_{km+1}, c_{km+2}, \ldots, c_{km+m} \quad \text{where} \quad k = 0, 1, \ldots$$

We divide the set of sequences into classes by saying that two sequences belong to the same class if and only if the terms of one are equal to the corresponding terms of the other. Clearly, the number of classes of sequences consisting of m terms is not greater than 10^m. Consequently it is a finite number. But, on the other hand, there are infinitely many sequences of form (18); so at least one of the classes contains infinitely many of them.

Remark. As a special case of the theorem just proved, we note that in any infinite decimal fraction at least one digit appears infinitely many times. (If, moreover, the number is irrational, there are at least two digits that appear infinitely many times each.) However, for numbers $\sqrt{2}$ and π we are unable to establish which two of the digits have this property. As was noticed by L. E. J. Brouwer, we do not know whether the sequence 0123456789 appears in the representation of number π as a decimal.

The decimals of e and π up to the 2053th decimal place are to be found on page 14 of a paper of G. W. Reitwiesner [1].

The number π is given up to the 100000th decimal place in D. Shanks and J. W. Wrench Jr. [1].

5. Prove that the number $(cc\ldots c)_{10}$, whose digits in the scale of 10 are all equal to c, with $c = 2$, $c = 5$ or $c = 6$, is not of the form m^n, where m and n are natural numbers > 1.

Proof. Numbers 2, 5 and 6 are not divisible by any square of a natural number > 1. Therefore none of them can be of the form m^n, where m and n are natural numbers > 1. Numbers whose last two digits are 22, 55 or 66 are not divisible by the numbers 4, 25 and 4 respectively, which would be the case if they were of the form m^n, where m and n are natural numbers > 1. A number > 4 whose digits (in the scale of 10) are all equal to 4 is divisible by 4 but not divisible by 8. Consequently it cannot be an nth power of a natural number m with $n \geqslant 3$. If $44\ldots4 = m^2$, then the number $11\ldots1$ would be a square; but this is impossible since the last two digits of a square of a natural number cannot be 11.

Remark. R. Obláth [1] showed that, if any of the numbers $33\ldots3$, $77\ldots7$, $88\ldots8$, $99\ldots9$ is greater than 10, then it cannot be of the form m^n, where m, n are natural numbers > 1. It is still an open question whether the number $11\ldots1$ can be of that form.

6. Write the number $\frac{1}{10}$ as a decimal in the scale of 2 and in the scale of 3.

Answer:

$$\tfrac{1}{10} = (0.0\dot{0}01\dot{1})_2 = (0.\dot{0}02\dot{2})_3.$$

7. Write the number $\frac{1}{61}$ as a decimal in the scale of 10.

Answer:

$$\tfrac{1}{61} = (0.\dot{0}1639344262295081967213114754098360655737704918032786885245\dot{9})_{10}.$$

Remark. It can be proved that the period of the decimal of number 1/97 consists of 96 digits and that of number 1/1913 consists of 1912 digits. We do not know whether there exist infinitely many natural numbers $n > 2$ such that the decimal of number $1/n$ has the period consisting of $n-1$ digits. To this class belong the numbers $n = 313$, 1021, 1873, 2137, 3221, 3313. It can be proved that primes for which 10 is a primitive root have this property.

§ 5. Normal numbers and absolutely normal numbers.

Let g be a natural number > 1; we write a real number x: $x = [x] + (0.c_1c_2c_3\ldots)_g$ as a decimal in the scale of g. For any digit c (in the scale of g) and every natural number n we denote by $l(c, n)$ the number of those digits of the sequence $c_1, c_2, \ldots, c_n$ which are equal to c. If

$$\lim_{n\to\infty} \frac{l(c, n)}{n} = \frac{1}{g}$$

for each of the g possible values of c, then number x is called *normal* in the scale of g. For example number

$$\frac{1234567890}{9999999999}$$

is normal in the scale of 10; number $\frac{1}{10}$ is normal in the scale of 2 but it is not normal in the scale of 3. If x is a normal number in the scale of 10, then $x/2$ is not necessarily a normal number. For example, $x = 0.\dot{1}35798204\dot{6}$ is a normal number and $x/2 = 0.\dot{0}67899102\dot{3}$ is not.

A number which is normal in any scale is called *absolutely normal*. The existence of absolutely normal numbers was proved by E. Borel [1]. His proof is based on the measure theory and, being purely existentional, it does not provide any method for constructing such a number. The first effective example of an absolutely normal number was given by me in the year 1916 (Sierpiński [5], see also H. Lebesgue [1]). As was proved by Borel, almost all (in the sense of the measure theory) real numbers are absolutely normal. However, as regards most of the commonly used numbers, we either know them not to be normal or we are unable to decide whether they are normal or not. For example, we do not know whether the numbers $\sqrt{2}$, π, e are normal in the scale of 10. Therefore, though according to the theorem of Borel almost all numbers are absolutely normal, it was by no means easy to construct an example of an absolutely normal number. Examples of such numbers are indeed fairly complicated.

D. G. Champernowne [1] proved in 1933 that the number a (which we proved in § 4 to be irrational) is normal in the scale of 10. He formulated the conjecture that the number whose decimal is obtained by writing 0 for the integers and the consecutive prime numbers (instead of consecutive natural numbers) to the right of the decimal point, i.e. number $0.2357111317\ldots$, is normal in the scale of 10. The conjecture, and a more general theorem have been proved by A. H. Copeland and P. Erdös [1]. Other interesting properties of normality have been investigated by W. M. Schmidt [1].

§ 6. Decimals in the varying scale. Let $g_1, g_2, \ldots$ be an infinite sequence of natural numbers >1, x a real number. We define infinite sequences $c_1, c_2, \ldots$ and $x_1, x_2, \ldots$ as follows:

$$c_0 = [x],\ x_1 = x-c_0,\quad c_1 = [g_1x_1],\ x_2 = g_1x_1-c_1,\ c_2 = [g_2x_2],\ \ldots,\ c_n = [g_nx_n],\ x_{n+1} = g_nx_n-c_n,\quad n = 1, 2, \ldots \tag{19}$$

It is clear that $0 \leqslant x_n < 1$ and $0 \leqslant c_n \leqslant g_n-1$ hold for any $n = 1, 2, \ldots$

Comparing formulae (19) and the algorithm of § 3, we see that the digit c_1 has been defined as if it were the corresponding digit in the scale of g_1, c_2 as if it were the corresponding digit in the scale of g_2 and so on. Moreover, formulae (19) give

$$x = c_0+\frac{c_1}{g_1}+\frac{c_2}{g_1g_2}+\frac{c_3}{g_1g_2g_3}+\ldots+\frac{c_n}{g_1g_2\ldots g_n}+\frac{x_{n+1}}{g_1g_2\ldots g_n}. \tag{20}$$

Since for $n = 1, 2, \ldots$ we have $g_n \geqslant 2$ and $0 \leqslant x_{n-1} < 1$, the last summand in (20) is non-negative and less than $1/2^n$, and consequently it tends to zero as n increases to infinity. This gives the following expansion of number x in an infinite series:

$$x = c_0+\frac{c_1}{g_1}+\frac{c_2}{g_1g_2}+\frac{c_3}{g_1g_2g_3}+\ldots \tag{21}$$

If $g_1 = g_2 = \ldots = g$, this coincides with the ordinary representation of x as a decimal in the scale of g.

Now we put $g_n = n+1$, $n = 1, 2, \ldots$ Then (21) assumes the form

$$x = c_0+\frac{c_1}{2!}+\frac{c_2}{3!}+\frac{c_3}{4!}+\ldots, \tag{22}$$

where c_0, c_n $(n = 1, 2, \ldots)$ are integers and

$$0 \leqslant c_n < n \quad (n = 1, 2, \ldots). \tag{23}$$

It is easy to prove that if x is a rational, algorithm (19) leads to a finite representation in form (22), where c_n $(n = 1, 2, \ldots)$ satisfy

inequalities (23). However, any rational admits also another infinite representation in form (22). This follows from the following identity:

$$c_0+\frac{c_1}{2!}+\frac{c_2}{3!}+\ldots+\frac{c_{n-1}}{n!}+\frac{c_n}{(n+1)!}$$
$$= c_0+\frac{c_1}{2!}+\ldots+\frac{c_{n-1}}{n!}+\frac{c_n-1}{(n+1)!}+\frac{n+1}{(n+2)!}+\frac{n+2}{(n+3)!}+\frac{n+3}{(n+4)!}+\ldots$$

As regards representations of type (21) see E. Strauss [1] and G. Cantor [1]; representations of the type (22) have been investigated by C. Stéphanos [1] and G. Faber [1].

Let us mention some other expansions of real numbers into infinite series.

Let x denote a positive real number. Denote by k_1 the least natural number satisfying the inequality $k_1x > 1$. We set $k_1x = 1+x_1$ and have $x_1 > 0$. We proceed similarly with x_1 in place of x, i.e. we find the least natural number k_2 such that $k_2x_1 > 1$ and we put $k_2x_1 = 1+x_2$ and so on. The expansion of x into an infinite series thus obtained is as follows

$$x = \frac{1}{k_1}+\frac{1}{k_1k_2}+\frac{1}{k_1k_2k_3}+\ldots,$$

where k_n $(n = 1, 2, \ldots)$ are natural numbers and $k_{n+1} \geqslant k_n$ for $n = 1, 2, \ldots$

It can be proved that each positive real number has precisely one representation in this form and that a sufficient and necessary condition for x to be an irrational number is that $\lim_{n\to\infty} k_n = +\infty$ (Sierpiński [3]).

The expansion thus obtained for number e is as follows:

$$e = \frac{1}{1}+\frac{1}{1\cdot 1}+\frac{1}{1\cdot 1\cdot 2}+\frac{1}{1\cdot 1\cdot 2\cdot 3}+\ldots$$

Let a be a natural number >2. Using the identity

$$\frac{a-\sqrt{a^2-4}}{2} = \frac{1}{a}+\frac{a^2-2-\sqrt{(a^2-2)^2-4}}{2a}$$

one easily proves that for $a_1 = a$, $a_{n+1} = a_n^2-2$ $(n = 1, 2, \ldots)$

$$\frac{a-\sqrt{a^2-4}}{2} = \frac{1}{a_1}+\frac{1}{a_1a_2}+\frac{1}{a_1a_2a_3}+\ldots \tag{24}$$

This series converges rapidly because, as is easily proved by induction, $a_n > 2^{2^{n-1}}$, $n = 1, 2, \ldots$

In particular, for $a=3$, we obtain $a_1=3$, $a_2=7$, $a_3=47$, $a_4=2207$, $a_5=4870847$ and so on. Hence

$$\frac{3-\sqrt{5}}{2}=\frac{1}{3}+\frac{1}{3\cdot 7}+\frac{1}{3\cdot 7\cdot 47}+\frac{1}{3\cdot 7\cdot 47\cdot 2207}+\dots$$

This expansion is to be found under the name of *Pell's series* in a book by E. Lucas [2], p. 331.

If a is even, $a=2b$, $b>1$, from (24) we derive the following expansion:

$$b-\sqrt{b^2-1}=\frac{1}{2b_1}+\frac{1}{2b_1 2b_2}+\frac{1}{2b_1 2b_2 2b_3}+\dots$$

where $b_1=b$ and $b_{n+1}=2b_n^2-1$ for $n=1,2,\dots$

It is worth noticing that the following expansion into an infinite product is valid:

$$\sqrt{\frac{b+1}{b-1}}=\left(1+\frac{1}{b_1}\right)\left(1+\frac{1}{b_2}\right)\left(1+\frac{1}{b_3}\right)\dots$$

Some particular cases of this expansion (for $b=2$, $b=3$ and some others) were given by G. Cantor [2] in 1879.

Now let x_0 denote an irrational number such that $0<x_0<1$. Let a_1 be the greatest natural number such that $x_0<\frac{1}{a_1}$. Let $x_1=\frac{1}{a_1}-x_0$. We then have $0<x_1<1$. We proceed similarly with x_1 in place of x_0 and obtain the greatest natural number a_2 such that $x_1<\frac{1}{a_2}$. We put $x_2=\frac{1}{a_2}-x_1$ and so on. Thus we obtain an infinite sequence of natural numbers $a_1, a_2\dots$ and an infinite sequence of irrational numbers $x_1, x_2, \dots$ such that $0<x_n<1$ for $n=0,1,2,\dots$ and $x_n=\frac{1}{a_n}-x_{n-1}$ for $n=1,2,\dots$ Moreover, $\frac{1}{a_n+1}<x_{n-1}<\frac{1}{a_n}$ for $n=1,2,\dots$ Hence $-x_{n-1}<-\frac{1}{a_n-1}$ and so

$$\frac{1}{a_{n+1}+1}<x_n<\frac{1}{a_n}-x_{n-1}<\frac{1}{a_n}-\frac{1}{a_n+1}=\frac{1}{a_n(a_n+1)}.$$

It follows that $a_{n+1}+1>a_n(a_n+1)$ and so $a_{n+1}\geqslant a_n(a_n+1)$ for $n=1,2,\dots$ From this, by induction, we easily infer that $a_{n+2}>2^{2^n}$ for

$n=1,2,\dots$ Numbers a_n increase rapidly to infinity with n. It follows from the definition of numbers a_n and x_n $(n=1,2,\dots)$ that

$$x_0=\frac{1}{a_1}-\frac{1}{a_2}+\frac{1}{a_3}-\dots+\frac{(-1)^{n-1}}{a_n}+(-1)^n x_n. \tag{25}$$

Since $0<x_n<\frac{1}{a_{n+1}}$, in view of the fact that $\lim\limits_{n\to\infty} a_{n+1}=+\infty$, we have $\lim\limits_{n\to\infty} x_n=0$. Therefore formula (25) gives us an expansion of the irrational number x_0 into an infinite rapidly convergent series

$$x_0=\frac{1}{a_1}-\frac{1}{a_2}+\frac{1}{a_3}-\frac{1}{a_4}+\dots \tag{26}$$

where a_n $(n=1,2,\dots)$ are natural numbers satisfying the inequalities

$$a_{n+1}\geqslant a_n(a_n+1) \quad \text{for} \quad n=1,2,\dots \tag{27}$$

We have thus proved that any irrational number x_0, $0<x_0<1$, may be expressed in form (26).

It can be proved that every irrational number between 0 and 1 has precisely one representation of this form and that a real number x_0 which can be expressed in form (26), where a_n $(n=1,2,\dots)$ are natural numbers satisfying conditions (27), is an irrational number (Sierpiński [4]).