REPRESENTATIONS FOR REAL NUMBERS

C. J. EVERETT[1]

1. **Introduction.** In a recent paper[2] [1] B. H. Bissinger generalized continued fractions by iteration of more general decreasing functions than the $1/x$ of the classical case. We extend here the algorithm by which real numbers are represented as decimals of base p, to general continuous increasing functions on $(0, p)$, including the classical x/p as special case. This sets up a correspondence from real numbers to sequences of integers mod p. Weak sufficient conditions are given that the correspondence be one-one. In the one-one case, algebraic examples are noted. The limit involved in the inscribed polygon problem appears here in a natural way. In the many-one case, the algorithm defines a set L of limit numbers which is perfect and nowhere dense. These sets are closely related to the Cantor perfect set. Finally, the relation between the above theory and the topological transformations F_1 of the unit interval into itself is studied. The latter yield sequences $\{F_p\}$ of our functions, $p=2, 3, \cdots$, and their structure is reflected in the limit sets $L_2, L_3, \cdots$.

2. **The algorithm.** Let $p \geqq 2$ be a fixed integer and $f(t)$ a continuous, strictly increasing function on the interval $0 \leqq t \leqq p$, with $f(0)=0$ and $f(p)=1$ (cf. [4]).

Such a function may be used to associate with every real number $\gamma_0 \geqq 0$, a sequence $\{c_\nu\}$ of integers, with $0 \leqq c_0 < \infty$, $0 \leqq c_\nu \leqq p-1$, $\nu=1, 2, \cdots$, by way of the following algorithm. We write, for $\gamma_0 \geqq 0$,

$$\text{(A)} \quad \begin{aligned} \gamma_0 &= c_0 + f(\gamma_1), & c_0 &\leqq \gamma_0 < c_0 + 1, \\ & & & 0 \leqq c_0 < \infty,\ 0 \leqq \gamma_1 < p, \\ \gamma_1 &= c_1 + f(\gamma_2), & c_1 &\leqq \gamma_1 < c_1 + 1, \\ & & & 0 \leqq c_1 \leqq p-1,\ 0 \leqq \gamma_2 < p, \end{aligned}$$

and so on.

Thus, at each step, c_ν is the greatest integer in γ_ν, and $\gamma_{\nu+1}$ is the uniquely defined real number on the interval $0 \leqq t < p$ such that $f(\gamma_{\nu+1}) = \gamma_\nu - c_\nu$, where $0 \leqq \gamma_\nu - c_\nu < 1$.

Presented to the Society, September 17, 1945; received by the editors July 10, 1945, and, in revised form, April 9, 1946.

[1] The work here reported was supported by the Wisconsin Alumni Research Foundation.

[2] The numbers in brackets refer to the references cited at the end of the paper.

Since $\gamma_1, \gamma_2, \cdots$ are all on $0 \leqq t < p$, it follows that $c_1, c_2, \cdots$ are integers satisfying $0 \leqq c_\nu \leqq p-1$. Hence we have a correspondence

(B) $$\gamma_0 \rightarrow \{c_\nu\}$$

from all reals $\gamma_0 \geqq 0$, to sequences of integers as described.

3. **Termination in $p-1$.** We ask now whether, under the algorithm, sequences may appear terminating in $p-1, p-1, \cdots$. Such is the case if and only if, for the function $f(t)$:

(C) There exists a $\gamma_0 = p-1+f(\gamma_0)$, $p-1 < \gamma_0 < p$.

Obviously such a $\gamma_0 \rightarrow \{p-1, p-1, \cdots\}$ under (A). On the other hand, if a number δ_0 under (A) yields a sequence terminating in $p-1, p-1, \cdots$, this implies that some δ_ν itself yields $p-1, p-1, \cdots$. Suppose then that γ_0 $(=\delta_\nu)$ under (A) gives $\gamma_0 = p-1+f(\gamma_1)$, $\gamma_1 = p-1+f(\gamma_2), \cdots, \gamma_\nu = p-1+f(\gamma_{\nu+1})$, and so on. If (C) is false, it follows from continuity of $f(t)$ that:

(D) $$f(t) > t-(p-1), \quad \text{for all } t \text{ on } p-1 \leqq t < p.$$

We should then have $\gamma_1-(p-1) < f(\gamma_1) = \gamma_0-(p-1)$, $\gamma_2-(p-1) < f(\gamma_2) = \gamma_1-(p-1)$, and so on, with $p > \gamma_0 > \gamma_1 > \gamma_2 > \cdots > p-1$. Hence $t_0 = \lim \gamma_\nu$ exists, with $p-1 \leqq t_0 < p$. But from $\gamma_\nu = p-1+f(\gamma_{\nu+1})$, we have $t_0 = p-1+f(t_0)$, a contradiction.

Indeed, the condition "not C" is equivalent to (D), and (D) may in turn be rephrased as a slope condition

(D′) $$(f(p)-f(t))/(p-t) < 1 \quad \text{on } p-1 \leqq t < p.$$

Moreover, if a γ_0 satisfying (C) exists, then not only γ_0 but every δ_0 on $\gamma_0 < \delta_0 < p$ will yield $\{p-1, p-1, \cdots\}$ under (A). For $\gamma_0 = p-1+f(\gamma_0) < \delta_0 < p$ implies $\delta_0 = p-1+f(\delta_1)$, hence $\gamma_0 < \delta_1 < p$, and so on. Since our final object is to obtain a one-one correspondence (B), we assume from this point on the necessary condition (D′). *The correspondence* (B) *then maps all reals* $\gamma_0 \geqq 0$ *onto non-*$(p-1)$*-terminating sequences.*

4. **Upper and lower limits.** Let $\{c_\nu\}$ be an arbitrary sequence of integers with $0 \leqq c_0 \leqq \infty$; $0 \leqq c_\nu \leqq p-1$, $\nu = 1, 2, \cdots$, not $(p-1)$-terminating. We define $C_\nu^\lambda = c_\lambda + f(c_{\lambda+1} + \cdots + f(c_{\lambda+\nu})$ and $\Gamma_\nu^\lambda = c_\lambda + f(c_{\lambda+1} + \cdots + f(c_{\lambda+\nu}+1)$, where the last parentheses are ν-fold. Then, from monotonicity, one has $c_\lambda \leqq C_\nu^\lambda \leqq C_{\nu+1}^\lambda < \Gamma_{\nu+1}^\lambda \leqq \Gamma_\nu^\lambda \leqq c_\lambda + 1$, so that the limits $C^\lambda = \lim C_\nu^\lambda$, $\Gamma^\lambda = \lim \Gamma_\nu^\lambda$ exist and satisfy

(E) $$c_\lambda \leqq C^\lambda \leqq \Gamma^\lambda \leqq c_\lambda + 1.$$

Since $C_\nu^\lambda = c_\lambda + f(C_{\nu-1}^{\lambda+1})$ we have $C^\lambda = c_\lambda + f(C^{\lambda+1})$ and similarly

$\Gamma^\lambda = c_\lambda + f(\Gamma^{\lambda+1})$. Now since the sequence is not $(p-1)$-terminating, for every λ there is a $c_{\lambda+\mu} \leqq p-2$. Moreover, $\Gamma_{\mu+\nu}^{\lambda} = c_\lambda + f(c_{\lambda+1} + \cdots + f(\Gamma_\nu^{\lambda+\mu})$, and $\Gamma^\lambda = c_\lambda + f(c_{\lambda+1} + \cdots + f(\Gamma^{\lambda+\mu})$. By (E), $\Gamma^{\lambda+\mu} \leqq c_{\lambda+\mu} + 1 \leqq p-1$, so that we have

$$\text{(E}'\text{)} \qquad c_\lambda \leqq C^\lambda \leqq \Gamma^\lambda < c_\lambda + 1,$$

and, as already shown,

$$\text{(F)} \qquad C^\lambda = c_\lambda + f(C^{\lambda+1}), \qquad \Gamma^\lambda = c_\lambda + f(\Gamma^{\lambda+1}).$$

But (E′, F) imply that, under (A), the numbers C^0 and Γ^0 yield the original sequence $\{c_\nu\}$. We call these the lower and upper limit numbers of the sequence.

If $f(t)$ satisfies (D′), *the correspondence* (B) *maps all reals $\gamma_0 \geqq 0$ onto all non-$(p-1)$-terminating sequences. Every such sequence is indeed the map of its limit numbers C^0, Γ^0.*

Now if γ_0 yields $\{c_\nu\}$ under (A), then

$$\text{(G)} \qquad \overset{0}{C_\nu} \leqq \gamma_0 = c_0 + f(c_1 + \cdots + f(c_\nu + f(\gamma_{\nu+1}) < \overset{0}{\Gamma_\nu}, \qquad \text{all } \nu,$$

and hence $C^0 \leqq \gamma_0 \leqq \Gamma^0$.

Also, if γ_0' and γ_0'' yield $\{c_\nu\}$ under (A), and if $\gamma_0' \leqq \gamma_0 \leqq \gamma_0''$, then γ_0 yields $\{c_\nu\}$. For

$$c_0 \leqq \gamma_0' = c_0 + f(\gamma_1') \leqq \gamma_0 \leqq \gamma_0'' = c_0 + f(\gamma_1') < c_0 + 1,$$

hence $\gamma_0 = c_0 + f(\gamma_1)$ and $\gamma_1' \leqq \gamma_1 \leqq \gamma_1''$, and so on.

It follows that γ_0 *yields* $\{c_\nu\}$ *under* (A) *if and only if* $C^0 \leqq \gamma_0 \leqq \Gamma^0$. Thus the correspondence (B) is actually a mapping of disjoint closed sets $[C^0, \Gamma^0]$ on all non-$(p-1)$-terminating sequences. The sequences $\{c_\nu\}$ fall into two classes according as $C^0 < \Gamma^0$ or $C^0 = \Gamma^0$. The correspondence (B) thus splits into two parts:

$$\text{(B}'\text{)} \qquad [C^0, \Gamma^0] \to \{c_\nu\}, \qquad C^0 < \Gamma^0,$$

$$\text{(B}''\text{)} \qquad C^0 = \Gamma^0 \to \{c_\nu\}.$$

In the case (B″) the C_ν^0 and Γ_ν^0 converge to $C^0 = \Gamma^0 = \gamma_0$ with errors thus (see G):

$$\text{(H)} \qquad 0 \leqq \gamma_0 - \overset{0}{C_\nu} < \overset{0}{\Gamma_\nu} - \overset{0}{C_\nu}; \qquad 0 < \overset{0}{\Gamma_\nu} - \gamma_0 \leqq \overset{0}{\Gamma_\nu} - \overset{0}{C_\nu}.$$

We note here two properties of the sequence $\{p-1, p-1, \cdots\}$ of later use. Although this sequence does not appear under the algorithm, nevertheless the lim C_ν^0 exists and is p. For $p-1 < C_\nu^0 < C_{\nu+1}^0 < p$ and $t_0 = \lim C_\nu^0$ satisfies $p-1 < t_0 \leqq p$. But $C_\nu^0 = p-1+f(C_{\nu-1}^1) = p-1+f(C_{\nu-1}^0)$. Hence $t_0 = p-1+f(t_0)$, and by (D′), $t_0 = p$.

Also, $p-1+f(p-1+\cdots+f(p-1+f(p-2)\geqq C_{\nu-1}{}^0$ where the first expression contains ν $(p-1)$'s. Thus the sequence $p-1+f(p-2)$, $p-1+f(p-1+f(p-2), \cdots$ has limit p.

5. **Terminating sequences.** We call a sequence $\{c_\nu\}$ with $c_\nu=0$, $\nu>N$ for some N, *terminating*. There exist numbers $\gamma_0>0$ yielding $\{0, 0, \cdots\}$ under (A) if and only if $f(t)$ has the property:

(I) There exists a $\gamma_0=f(\gamma_0)$, $0<\gamma_0<1$.

Clearly such a γ_0 yields $\{0, 0, \cdots\}$ under (A). Suppose that $\gamma_0>0$ yields $\{0, 0, \cdots\}$ and that (I) is false. By continuity of $f(t)$ we have

$$\text{(J)} \qquad f(t) < t \quad \text{for all } t \text{ on } 0 < t \leqq 1,$$

and $0<\gamma_0=f(\gamma_1)<\gamma_1=f(\gamma_2)<\gamma_2 \cdots$. Hence $0<\gamma_0<\gamma_1<\gamma_2<\cdots<1$, and $t_0=\lim \gamma_\nu$ exists with $0<t_0\leqq 1$. But from $\gamma_\nu=f(\gamma_{\nu+1})$ follows $t_0=f(t_0)$, a contradiction.

Obviously "not I" is equivalent to (J), and (J) may be restated in slope form

$$\text{(J}'\text{)} \qquad f(t) - f(0)/t < 1 \qquad \text{on } 0 < t \leqq 1.$$

If a γ_0 exists satisfying (I) then not only γ_0 but also every δ_0 on $0\leqq\delta_0<\gamma_0$ will yield $\{0, 0, \cdots\}$ under (A). Hence for a one-one correspondence (B), (J′) is necessary, and we assume from this point on that $f(t)$ satisfies (D′) and (J′).

Under these restrictions, the sequence $\{0, 0, \cdots\}$ has $C^0=\Gamma^0=0$, and since in any sequence $\{c_\nu\}$, $C^0=c_0+\cdots+f(C^\lambda)$, $\Gamma^0=c_0+\cdots+f(\Gamma^\lambda)$, it follows that every terminating sequence has $\Gamma^0=C^0$ and falls under (B″).

We remark here that if $\{d_\nu\}$ is a terminating sequence $\{d_1, d_2, \cdots, d_\nu, 0, 0, \cdots\}$, then the associated limit numbers $D^0=\Delta^0=d_0+f(d_1+\cdots+f(d_\nu+f(D^{\nu+1})=d_0+f(d_1+\cdots+f(d_\nu)$, since $D^{\nu+1}=0$.

6. **The many-one case.** Suppose then that $f(t)$ satisfies (D') and (J′) and consider the algorithm (A) only as it applies to numbers γ_0 on the interval $[0, p)=(0\leqq t<p)$. The correspondences (B′, B″) then map the interval $[0, p)$ onto all non-$(p-1)$-terminating sequences $\{c_\nu\}$ with $0\leqq c_\nu\leqq p-1$.

Let L be the set of all limit numbers C^0, Γ^0 on $[0, p)$ (equal or not) of all such sequences, and G the complement of L in $[0, p)$. The points of L are then the numbers $C^0=\Gamma^0$ occurring under (B″), including the limits of all terminating sequences, together with the end points $C^0<\Gamma^0$ of the closed intervals under (B′). The points of G are those of all the open intervals (C^0, Γ^0) in (B′). Since G is a union of (nonoverlapping, indeed, non-abutting) intervals, *G is open, and L is closed.*

We write $[0, p) = L + G$, and $L = L' + L''$, where L' is the set of end points under (B′) and L'' the set of $C^0 = \Gamma^0$ under (B″).

Since *the intervals of G are countable, so is the set L'.* We now show that L is dense in itself. It then follows that *L is perfect, L (hence also L'') has the power of the continuum. Since the limits of terminating sequences are countable, the set of limits of non-terminating sequences for which $C^0 = \Gamma^0$ is of the power of the continuum* (cf. [3]).

Indeed, every point λ of L is a limit point of limit numbers $D^0 = \Delta^0$ of *terminating* sequences $\{d_\nu\}$. First let $\lambda = C^0 = \Gamma^0$ for $\{c_\nu\}$. Then $\lambda = \lim C_\nu^0 = \lim \Gamma_\nu^0$ and $C_\nu^0 < \Gamma_\nu^0$. The numbers C_ν^0 are in L, being limit numbers of terminating sequences. Since the sequence $\{c_\nu\}$ is not $(p-1)$-terminating, a subsequence of $\{\Gamma_\nu^0\}$ has $\Gamma_\nu^0 = c_0 + f(c_1 + \cdots + f(c_\nu + 1)$ with $c_\nu + 1 \leqq p - 1$, and these Γ_ν^0 are thus in L, being limit numbers of terminating sequences $\{c_0, c_1, \cdots, c_\nu + 1, 0, 0, \cdots\}$ in our class. Hence λ is a limit point of points of L.

Second, let $\lambda = C^0 < \Gamma^0$ for $\{c_\nu\}$. Then the sequence $\{c_\nu\}$ is not terminating, and a subsequence of $\{C_\nu^0\}$ is properly increasing to C^0 as a limit point.

Finally, let $C^0 < \Gamma^0 = \lambda$ for $\{c_\nu\}$. Since $\{c_\nu\}$ is not $(p-1)$-terminating, a proper subsequence of $\{\Gamma_\nu^0\}$ with $c_\nu + 1 \leqq p - 1$ is properly decreasing to Γ^0 as a limit point. Hence L is dense in itself.

If $f(t)$ admits a sequence $\{d_\nu\}$ with $D^0 < \Delta^0$, that is, *if the correspondence* (B) *is not one-one, then the set L is nondense on* $[0, p)$. If (a, b) is a subinterval: $0 \leqq a < b < p$, we show that (a, b) contains a subinterval containing no point of L. If (a, b) itself contains no point of L, (a, b) will serve. However if a point λ of L is in (a, b) and if $\lambda = C^0 < \Gamma^0$ or $C^0 < \Gamma^0 = \lambda$ for some $\{c_\nu\}$ then the interval (a, b) intersects (C^0, Γ^0) in an interval containing only points of G. The only case remaining is $\lambda = C^0 = \Gamma^0$ in (a, b), $\lambda = \lim C_\nu^0 = \lim \Gamma_\nu^0$. But $C_\nu^0 < c_0 + f(c_1 + \cdots + f(c_\nu + f(D^0) < c_0 + f(c_1 + \cdots + f(c_\nu + f(\Delta^0) < \Gamma_\nu^0$. The inner numbers define an interval of G, interior to (a, b) for sufficiently large ν.

7. **An example.** Consider for $p = 3$ the function $f(t)$ defined by $f(0) = 0$, $f(4/3) = 1/3$, $f(5/3) = 2/3$, $f(3) = 1$, and elsewhere by the broken line connecting these points. It is clear that $4/3$ and $5/3$ yield $\{1, 1, \cdots\}$ under the algorithm. Moreover, for this sequence, $C^0 = 4/3$ and $\Gamma^0 = 5/3$ as is seen graphically from the sequences $1 + f(1)$, $1 + f(1 + f(1)$, $\cdots$ and $1 + f(2)$, $1 + f(1 + f(2)$, $\cdots$. Imagine that we blacken the intervals $(i + f(C^0), i + f(\Gamma^0))$, $i = 0, 1, 2$. The first of these defines three intervals $(j + f(0 + f(C^0), j + f(0 + f(\Gamma^0)), j = 0, 1, 2$, and the last similarly, all of which we blacken. (Graphically, the

process amounts to projecting the function values above each black interval onto the three 45° lines and thence down to the t-axis.) Repetition of this process yields a set of open intervals of total length $1/3+2/3+3(2/3)(1/4)+\cdots=1/3+2/3(1+3/4+(3/4)^2+\cdots)=3$. It follows that the set of black intervals exhausts the set G, and the complement L is of measure zero, perfect, and nondense on $[0, 3)$. While this is not quite the Cantor "middle-third" set it has precisely the same structure.

8. **Sufficient conditions for one-one correspondence.** Let $c_0<\gamma_0<\delta_0<c_0+1$ and $\gamma_0, \gamma_1, \cdots, \gamma_n$; $\delta_0, \delta_1, \cdots, \delta_n$, be the numbers resulting from the first n steps of the algorithm. We say that the slopes $f(\delta_i)-f(\gamma_i)/\delta_i-\gamma_i$, $i=1, \cdots, n$, are *connected.*

In order that the correspondence (B) *be one-one it is sufficient that*:

(K) *There exists an integer n such that the product of every n connected slopes is less than one.*

Suppose that (B) is not one-one, and let X' be the class of all intervals (C^0, Γ^0) under (B$'$). Then there must be in X' an interval of maximal length. For this interval, write $\Gamma^0-C^0=(f(\Gamma^1)-f(C^1)/\Gamma^1-C^1)\cdots(f(\Gamma^n)-f(C^n)/\Gamma^n-C^n)(\Gamma^n-C^n)$. The interval (C^n, Γ^n) is in X', hence these n connected slopes have product not less than 1, contradicting (K).

Stronger sufficient conditions are:

$$\text{(K}'\text{)} \qquad f(t_2) - f(t_1)/t_2 - t_1 < 1, \qquad 0 \leqq t_1 < t_2 \leqq p.$$

(K$''$) There exists a β such that $0<\beta<1$, $f(t_2)-f(t_1)/t_2-t_1\leqq\beta$ on $0\leqq t_1<t_2\leqq p$.

In case (K$''$) obtains, we note that $\Gamma_\nu^0-C_\nu^0=(f(\Gamma_{\nu-1}^1)-f(C_{\nu-1}^1))/\Gamma_{\nu-1}^1-C_{\nu-1}^1)\cdots(f(\Gamma_1^{\nu-1})-f(C_1^{\nu-1})/\Gamma_1^{\nu-1}-C_1^{\nu-1})\ (f(c_\nu+1)-f(c_\nu)/(c_\nu+1-c_\nu)\leqq\beta^\nu$, so that from (H) the error in the Γ_ν^0 and C_ν^0 approximations to $\gamma_0=C^0=\Gamma^0$ is not greater than β^ν.

Although the slope condition (K$'$) is sufficient for one-one (B), it is far from necessary. We shall construct functions of arbitrarily great slope for which (B) is one-one.

Consider the set of all ratios (note: *not* slopes) $f(b+f(a+1)-f(b+f(a))/f(b+1)-f(b)$ where a, b are arbitrary integers on $0, 1, \cdots, p-1$. Of these there are only a finite number, each less than one, since the numerator is the difference of function values on a proper subinterval of $(b, b+1)$. Let M be the maximum of these ratios, $M<1$.

Now consider the intervals $(b+f(a), b+f(a+1))$ and suppose that the ratio of inner to outer slope of $f(t)$ on each of these intervals is bounded above from $1/M$, that is:

(L) There exists a $k<1/M$ such that

$$(f(t_2)-f(t_1)/t_2-t_1)/(f(b+f(a+1)-f(b+f(a)/f(a+1)-f(a))\leqq k$$

for all t_1, t_2 on

$$b+f(a)\leqq t_1<t_2\leqq b+f(a+1),$$

or equivalently:

(L′) There exists a $k<1/M$ such that whenever

$$b+f(a)\leqq t_1<t_2\leqq b+f(a+1) \quad\text{and}\quad t_2-t_1\leqq\tau\cdot(b+f(a+1)-(b+f(a)),$$

we must have $f(t_2)-f(t_1)\leqq k\tau(f(b+f(a+1)-f(b+f(a))$.

The condition (L′) *is sufficient for one-one* (B).

For, by definition of M, $f(b+f(a+1)-f(b+f(a)\leqq M(f(b+1)-f(b))$ and $c+f(b+f(a+1)-(c+f(b+f(a))\leqq M((c+f(b+1)-(c+f(b))$. Now use (L′) on the interval $(c+f(b), c+f(b+1))$ and we have

$$f(c+f(b+f(a+1)-f(c+f(b+f(a)\leqq kM(f(c+f(b+1)-f(c+f(b))$$
$$\leqq kM^2(f(c+1)-f(c)).$$

By iteration of this process, one obtains

$$f(c_1+\cdots+f(c_\nu+1)-f(c_1+\cdots+f(c_\nu)$$
$$\leqq k^{\nu-2}M^{\nu-1}(f(c_1+1)-f(c_1))\leqq(kM)^{\nu-2}M,$$

which approaches zero since $k<1/M$.

While this discussion is cumbersome, it nevertheless shows that a function $f(t)$ defined arbitrarily (consistent with monotonicity) at $t=0, 1, 2, \cdots, p-1, p$, and then at all $t=b+f(a)$, a, b on $0, \cdots, p-1$, and elsewhere by the broken line connecting these points, must satisfy (L) with $k=1$, since the ratio of inner to outer slope on the straight segments is unity.

Thus the broken line function connecting $f(0)=0$, $f(1)=e>0$, (e arbitrarily small constant), $f(1+e)=1-e$, $f(2)=1$ (for $p=2$) yields a one-one (B). The slope on $(1, 1+e)$ however is $(1-2e)/e$, which may be arbitrarily large.

9. **Algebraic examples of the one-one case.** *Example* 1. Let $f(t)=t^n/p^n$ for an integer $p\geqq 2$ and an integer n on $1\leqq n<p$. One verifies the properties of §2, and condition (K″) with $\beta=n/p$. For $n=1$, our theory reduces to the classical decimals with base p. In the general case let q be an integer not greater than $p-2$, and let C^0 be the limit for sequence $\{q, q, \cdots\}$. Then $C^0=q+f(C^0)$, and the number $\alpha=C^0/p$ satisfies $p\alpha=q+\alpha^n$ or $\alpha^n-p\alpha+q=0$, where $0\leqq q/p\leqq\alpha<(q+1)/p\leqq p-1/p$. Thus the equation $x^n-px+q=0$, $1\leqq n<p$, $0\leqq q$

$\leqq p-2$, has exactly one real root on $[0, 1)$, namely $\alpha=(1/p)\cdot(q+f(q+f(q+\cdots$.

In particular, for $n=2$, $p=3$, $q=1$, $x^2-3x+1=0$, $\alpha=(3-5^{1/2})/2$ is approximately

$$(1/3)C_3^0 = (1/3)\left(1+\frac{1}{9}\left(1+\frac{1}{9}\left(1+\frac{1}{9}\right)^2\right)^2\right),$$

with error not greater than $\beta^3=(2/3)^3$.

Example 2. For $f(t)=(1+t)^{1/n}-1$, $p=2^n-1$, $n>1$, one has slope on $(0, p)$ not greater than $1/n$. We consider $\gamma=1+f(q+f(q+\cdots)$ where $0<q\leqq 2^n-2$. We have $\gamma=1+f(q+\gamma-1)=1+(q+\gamma)^{1/n}-1$ or $\gamma^n-\gamma-q=0$. Thus, the equation $x^n-x-q=0$, $n>1$, $0<q\leqq 2^n-2$, has only one real root γ on $(1, 2]$, namely the number γ above.

For instance, $n=2$, $p=3$, $q=1$, $x^2-x-1=0$, $\gamma=1+f(1+f(1+\cdots$. The successive C_ν^0 are $1+f(1)=2^{1/2}=(1+1^{1/2})^{1/2}$ (from here on radicals are "nested"), $1+f(1+f(1)=(1+(1+1^{1/2})^{1/2})^{1/2}$, and so on. Hence $(1+5^{1/2})/2=(1+(1+(1+\cdots)^{1/2}$.

Recalling the remark at the end of §4, and using $n=2$, $p=3$, $q=2$, $x^2-x-2=0$, $\gamma=2=1+f(2+f(2+\cdots)$, the successive approximations being $1+f(2)=3^{1/2}$, $1+f(2+f(2)=(2+3^{1/2})^{1/2}$, $1+f2+f2+f2=(2+(2+3^{1/2})^{1/2})^{1/2}$, and so on. But using the sequence $p-1+f(p-1+\cdots+f(p-2)$, we have $3=2+f2+f2+\cdots$ with approximations $2+f(1)=1+2^{1/2}$, $2+f2+f1=1+(2+2^{1/2})^{1/2}$, whence $2=(2+(2+(2+\cdots)^{1/2}$ which is the classical limit occurring in the inscribed polygon theory [2].

Finally, for $n=3$, $p=2^n-1=7$, $q=6$, $x^3-x-6=0$, $\gamma=2=1+f(6+f(6+\cdots$, the approximations being $1+f6=7^{1/3}$, $1+f6+f6=[6+7^{1/3}]^{1/3}$, or again using the $(p-2)$-terminating approximations, $2=\{6+[6+(6+\cdots)^{1/3}]^{1/3}\}^{1/3}$.

10. "Spectra" of the topological maps of the unit interval. Let $T=\{F_1(t)\}$ be the class of all continuous increasing functions on $0\leqq t\leqq 1$ with $F_1(0)=0$, $F_1(1)=1$. These are the topological mappings of the unit interval onto itself [5]. If p is any integer not less than 2 and $f(t)$ is of the type in §2:

(M) $f(t)$ continuous increasing on $0\leqq t\leqq p$; $f(0)=0$, $f(p)=1$, then $F(t)=f(pt)$ is in the class T. Thus all our functions may be regarded as magnifications of the functions of T by a factor p in the t-direction. Conversely, if $F_1(t)$ is in T and $p\geqq 2$, then $F_p(t)=F_1(t/p)$ is a function of type (M). Hence for every $F_1(t)$ in T we regard the sequence of functions $\{F_1(t), F_2(t), F_3(t), \cdots\}$ where $F_p(t)=F_1(t/p)$ for $p\geqq 2$. The associated sequence of perfect sets L_p of limit numbers of $F_p(t)$ is a curious sort of "spectrum" for F_1.

For a fixed $F_1(t)$ in T, the correspondence $(t, F_1(t)) \leftrightarrow (nt, F_1(t))$, $0 \leqq t \leqq 1$, is a one-one correspondence of the points on the curves $y = F_1(t)$ and $y = F_n(t)$. This induces a one-one correspondence between the points of the curves F_n and F_{n+1}, namely, $(nt, F_1(t)) \leftrightarrow ((n+1)t, F_1(t))$, $0 \leqq t \leqq 1$. The latter may be used to show that the slopes s_n, s_{n+1} of the chords at corresponding points of F_n and F_{n+1} satisfy $s_n > s_{n+1}$. Hence if some F_n satisfies (K′) so do all succeeding F_n, and thus $L_p = [0, p)$, $p \geqq n$. If F_1 is of bounded slope, there will exist an F_n of slope everywhere less than one. Moreover, one can show that if F_n satisfies (D′) and (J′), so does F_{n+1}. This leads to the question whether $L_p = [0, p)$ implies $L_{p+1} = [0, p+1)$. This is in fact *not* the case.

Example. The broken line function F_1 defined by $F_1(0) = 0$, $F_1(4/9) = 1/3$, $F_1(5/9) = 2/3$, $F_1(1) = 1$ has $L_2 = [0, 2)$, since the product of every two connected slopes of F_2 is less than one (condition (K) with $n = 2$; note that the test (L) fails). But $F_3(t)$ is the function of §7 with L_3 of measure zero. However $L_p = [0, p)$, $p \geqq 4$, since the maximum slope of F_3 is one and all successors therefore have slope less than one.

11. **Unsolved problems.** (1) State simple necessary and sufficient conditions on $f(t)$ such that (B) be one-one. (2) Do there exist functions $f(t)$ which give $C^0 < \Gamma^0$ for non-terminally-periodic sequences $\{c_\nu\}$? (3) Do functions exist with sets L of every measure between 0 and p? (4) The limits of periodic sequences of period k are algebraic numbers of degree n^k at most for the function of Example 1, §9. Characterize algebraically all such limits.

References

1. B. H. Bissinger, *A generalization of continued fractions*, Bull. Amer. Math. Soc. vol. 50 (1944) pp. 868–876.

2. R. Courant and H. Robbins, *What is mathematics?* Oxford University Press, New York, 1941, p. 125.

3. E. W. Hobson, *Theory of functions of a real variable*, vol. 1, 3d ed., Cambridge University Press, 1927, p. 117 ff.

4. O. Perron, *Irrationalzahlen*, Berlin, 1921, p. 90 ff.

5. J. Schreier and S. Ulam, *Eine Bemerkung über die Gruppe der topologischen Abbildungen der Kreislinie auf sich selbst*, Studia Mathematica vol. 5 (1935) pp. 156–159.

University of Wisconsin