62. A Remark on a Theorem of Copeland-Erdös

By Iekata SHIOKAWA
Tokyo University of Education

(Comm. by Kôsaku YOSIDA, M. J. A., April 18, 1974)

Let $g \geq 2$ be a fixed integer. An infinite sequence $a_1a_2\cdots$ of non-negative integers not greater than $g-1$ is said to be normal to base g, if for every positive integer l and every sequence $B=b_1b_2\cdots b_l$ of digits $0, 1, \cdots, g-1$, of length l we have

$$\lim_{n\to\infty} \frac{1}{n} N_n(B) = g^{-l},$$

where $N_n(B)$ is the number of indices i, $1 \leq i \leq n$, for which $a_i a_{i+1}\cdots a_{i+l-1} = b_1b_2\cdots b_l$. Any positive integer n can be expressed uniquely in the form

$$n = \sum_{i=1}^{k} a_i g^{k-i}$$

where each $a_i = a_i(n)$ is one of $0, 1, \cdots, g-1$, and $k=k(n)$ is the integer such that $g^{k-1} \leq n < g^k$, and we shall denote the sequence $a_1a_2\cdots a_{k(n)}$ by $B(n)$. An increasing sequence $\{m_1, m_2, \cdots\}$ of positive integers is said to be normal to base g, if the sequence of digits $B(m_1)B(m_2)\cdots$ is normal to base g. In 1946 Copeland-Erdös [1] proved that *any increasing sequence* $\{m_1, m_2, \cdots\}$ *of positive integers such that for every* $\theta < 1$ *the number of* m_j*'s up to* x *exceeds* x^θ *provided* x *is sufficiently large, is normal to any base.* This theorem implies the normality (to any base) of the sequence of prime numbers, and this is the only known proof of this fact. In this paper we shall make a remark that the theorem of Copeland-Erdös is, in some sense, the best possible. Indeed we shall prove the following

Theorem. *For any fixed integer* $g \geq 2$ *and any fixed positive number* $\theta < 1$ *we can construct a non-normal (to base* g*), increasing sequence of positive integers such that*

$$x^\theta < \sum_{m_j \leq x} 1 < g^2 x^\theta$$

for all sufficiently large x.

To prove the theorem we need the following lemma.

Lemma. *Let* b *be any one of* $0, 1, \cdots, g-1$, *and let* $\varepsilon < 1/3$ *be any fixed positive number. Denote by* $T(b; k, \varepsilon)$ *the number of sequences* $B = b_1b_2\cdots b_k$ *of* 0*'s,* 1*'s,* $\cdots$*,* $g-1$*'s of length* k *such that* $N(b, B) > (g^{-1}+\varepsilon)k$, *where* $N(b, B)$ *be the number of* b*'s contained in the sequence* B. *Then we have*

$$T(b; k, \varepsilon) > g^k \exp(-16g\varepsilon^2 k)$$