ACTA ARITHMETICA
LXVI.4 (1994)

On a result of Mahler on the decimal expansions of $(n\alpha)$

by

Daniel Berend (Beer-Sheva) and
Michael D. Boshernitzan (Houston, Tex.)

1. Introduction and the main results. It is well known that, given any irrational α, the sequence $(n\alpha)_{n=1}^{\infty}$ is dense modulo 1. (It is also uniformly distributed modulo 1, but this is of no consequence here.) In particular, given any digits $a_1, a_2, \ldots, a_k$, there exists a positive integer m for which the decimal expansion of $m\alpha$ contains this block of digits. It was proved by Mahler [M] that, moreover, there necessarily exists an m for which the decimal expansion of $m\alpha$ contains the given block infinitely often. Mahler also established an upper bound for the minimal value M of the number m with that property; $M = M(k)$ depends only on the number k of digits, but not on α:

$$M(k) < 10^{2k+1} .$$

Mahler's original proof is based on the geometry of numbers.

In this paper we give a shorter proof of Mahler's result (see Section 2), which at the same time yields a better upper bound:

$$M(k) < 2 \cdot 10^{k+1} .$$

This result is best possible up to a constant factor. In fact, we show that

$$M(k) \geq 8 \cdot (10^k - 1) .$$

(Actually, the factor 8 can be replaced by any real number less than 10 for sufficiently large k—see Example 3.1.)

Of course, there is nothing special about the base 10. Mahler's theorem refers equally to any base $g \geq 2$, and the upper bound for $M(g,k)$ he obtains in this general case is:

$$M(g,k) < g^{2k+1} .$$

(Note that even the finiteness of $M(g,k)$ is not obvious.)

Research of the second author supported in part by NSF Grant No. DMS-9003450.

Our first result improves upon this bound. A *g-block of length k* is a sequence of length k with entries in $\{0,1,\ldots,g-1\}$.

THEOREM 1.1. *Let α be an irrational, $g \ge 2$ an integer and B a g-block of length k. Then there exists a positive integer $m < 2g^{k+1}$ such that the g-ary expansion of $m\alpha$ contains the block B infinitely often.*

The theorem is proved in Section 2.

Remark. As is well known, the g-ary expansion of almost every α (in the sense of the Lebesgue measure) contains every g-block infinitely often (and even in the "right" frequency). The theorem thus relates mainly to numbers α which are "badly behaved" in base g.

It is easy to see (Proposition 3.1) that

$$M(g,k) \ge g^k - 1\,.$$

Thus the gap between our upper and lower bounds is just by a factor of $2g$, which is constant (for fixed g). In Section 3 we shall discuss improvements upon this lower bound. Our lower bounds depend on the arithmetic nature of g (i.e., its factorization into a product of primes), and may hint that there is no simple formula for $M(g,k)$.

The density modulo 1 of the sequence $(n\alpha)$ is but a special case of a result which asserts that, given any polynomial P with real coefficients, at least one of which (besides the constant term) is irrational, the sequence $P(n)$ is dense modulo 1. (More well-known is Weyl's even stronger result by which this sequence is uniformly distributed modulo 1 [W].) It turns out that Mahler's result is true in this more general setting as well.

THEOREM 1.2. *Let $g \ge 2$ be an integer and $P \in \mathbb{R}[x]$ a polynomial with at least one irrational coefficient besides the constant term. Then for each finite g-block there exists a positive integer m such that B appears infinitely often in the g-ary expansion of $P(m)$.*

Remark. It can be shown (although this does not follow from the considerations of this paper) that there exists an effective upper bound, in terms of g, the length of B and the degree of P, on the least m satisfying the conclusion of the theorem.

EXAMPLE 1.1. There are numerous sequences in which one can find, given any g-block, an element whose g-ary expansion contains the block infinitely often. Such are the sequences $(\ln n)$ (consider numbers n of the form 2^m and use Mahler's result), $(\ln\ln n)$ (take n's of the form 2^{2^m} and use Theorem 1.2 for linear polynomials) and (n^θ) for θ positive rational non-integer (if $\theta = p/q$ take n's of the form $2m^q$ and use Theorem 1.2). On the other hand, we do not know whether the sequences $(\ln\ln\ln n)$ and (n^θ) with irrational θ share this property. More generally, we note that the question of infinite repetitions of blocks is usually harder than the question of density mod 1 which holds for the above sequences. For a large class of regularly growing sequences (defined by certain formulae or recurrences), the questions of density and of uniform distribution modulo 1 can be resolved by means of simple tests [B], but we doubt that such criteria can be formulated for infinite repetition problems.

We note in conclusion that our approach was influenced by an idea due to Furstenberg who, employing a certain result of Glasner [G], provided a very short proof of the finiteness of $M(g,k)$ (see [AP, Cor. 7.2]).

2. The improved upper bound. Let $\mathbb{T}=\mathbb{R}/\mathbb{Z}$ be the circle group. A set $E\subseteq\mathbb{T}$ is *ε-dense* (or, alternatively, *forms an ε-net*) in $\mathbb{T}$ if every interval of length ε meets E.

It is obvious that $M(g,k)$ cannot be divisible by g, so that proving a weak inequality in Theorem 1.1 is equivalent to proving the strict inequality.

Proof of Theorem 1.1. Denote by E the set of all limit points in $\mathbb{T}$ of the sequence $\{g^n\alpha : n\ge 0\}$. We distinguish between several (somewhat overlapping) cases:

Case I: E contains a rational point $r=p/q$ (where $(p,q)=1$) with $g^k<q\le g^{k+1}$. In this case the set $\{0,r,2r,\ldots,(q-1)r\}$ forms a $\frac{1}{g^k}$-net in $\mathbb{T}$. Hence the g-ary expansion of some mr, $1\le m<q$, starts with $0.B$, and not all of the following digits are 0, neither are they all $g-1$. It follows that the g-ary expansion of $m\alpha$ contains the block B infinitely often. (Note that in this case we could have replaced the upper bound $2g^{k+1}$ by g^{k+1}.)

Case II: $0\in E$. Replacing α by $-\alpha$ if necessary, we may assume that the g-ary expansion of α contains arbitrarily large blocks consisting of 0. Take a sequence (n_j) of positive integers such that the g-ary expansion of $g^{n_j}\alpha$ starts with the block 0^j, but that of $g^{n_j-1}\alpha$ does not start with 0. For each fixed positive integer d, consider the sequence $g^{n_j-d}\alpha$ (which is well-defined for sufficiently large j). Replacing (n_j) by a subsequence thereof, we may assume that each of these sequences converges in $\mathbb{T}$, say $g^{n_j-d}\alpha\to r_d=p_d/q_d$ (where $(p_d,q_d)=1$). Obviously, for each d we have $q_d<q_{d+1}\le gq_d$. Hence $g^k<q_d\le g^{k+1}$ for an appropriate choice of d. Since $r_d\in E$ for each d, this yields a reduction to the preceding case.

Case III: E contains a rational point $r=p/q$ (where $(p,q)=1$) with $q\le g^{k+1}$. Carrying out the construction of Case II, with α replaced by $q\alpha$, we find rationals s_d in qE with finite g-ary expansion and corresponding rationals $r_d=p_d/q_d$ in E such that $g^{n_j-d}\alpha\to r_d$ and $qs_d=r_d$ for each d. As in the preceding case we have $q_d<q_{d+1}\le gq_d$, and since $q_0\le g^{k+1}$ there exists an r_d whose denominator is in the range $(g^k,g^{k+1}]$, bringing us again back to Case I.

Case IV: E contains no rational point $r=p/q$ with $q\le 2g^k$. We first claim that a point $\beta \in E$ and a rational $r = p/q$ can be found such that:

(a) $|\beta - p/q| < 1/(2g^{k+1}q)$.
(b) $2g^k < q$.
(c) $q \le 2g^{k+1}$.
(d) $(p,q) = 1$.

In fact, starting with any $\beta_0 \in E$, we can find a rational $r_0 = p_0/q_0$ such that conditions (a), (c) and (d) are satisfied with β and r replaced by β_0 and r_0, respectively. Choose inductively points $\beta_i \in E$, $i = 0,1,2,\ldots$ with $g\beta_{i+1} = \beta_i$ for each i. Next choose rationals $r_i = p_i/q_i$, $i = 1,2,\ldots$ (in reduced form) with $gr_{i+1} = r_i$ and $|\beta_i - r_i| = g^{-i}|\beta_0 - r_0|$ for each i. Clearly, $q_i \le q_{i+1} \le gq_i$ for each i. If $q_i = q \le 2g^k$ for all sufficiently large i, then some rational $r = p/q$ appears infinitely often in the sequence (r_i), in which case $r \in E$, contradicting our assumption. Consequently, $q_i \to \infty$, whence for a suitable i the rational r_i satisfies conditions (a)–(d).

Now the set $\{r, 2r, \ldots, qr\}$ forms a $\frac{1}{2g^k}$-net in $\mathbb{T}$. Since $|m\beta - mr| < 1/(2g^k)$ for $1 \le m \le q$, the set $\{\beta, 2\beta, \ldots, q\beta\}$ forms a $\frac{1}{g^k}$-net in $\mathbb{T}$. We conclude as in Case I.

This completes the proof.

3. Lower bounds. In this section we shall discuss the question of lower bounds on $M(g,k)$. A simple observation is

PROPOSITION 3.1. *$M(g,k) \ge g^k - 1$ for every $g \ge 2$, $k \ge 1$.*

In fact, considering a number of the form $\alpha = \sum_{j=1}^{\infty} g^{-n_j}$, where $n_{j+1} - n_j \to \infty$, and the block B consists of k consecutive $(g-1)$'s, we easily see that the g-ary expansion of $m\alpha$ will not contain B infinitely often for any $m < g^k - 1$.

The bound provided by Proposition 3.1 may be equal exactly to $M(g,k)$. This is the case, for example, for $g = 2$, $k = 1$ and for $g = 3$, $k = 1$. It is, however, usually possible to improve on this lower bound, as we first see for composite g.

THEOREM 3.1. *Let a be a proper divisor of g. Then*

$$M(g,k) \ge a(g^k - 1), \qquad k \ge 1 .$$

Taking $a = 1$ we obtain Proposition 3.1. Of course, the best result is obtained in general by selecting a as the maximal proper divisor of g. Thus Theorem 3.1 improves Proposition 3.1 unless g is a prime.

The proof of Theorem 3.1 is almost the same as that of Proposition 3.1, except that we choose the "bad number" this time as

$$\alpha = \frac{g}{a}\sum_{j=1}^{\infty} g^{-n_j}\,.$$

The least multiple of g/a containing the block consisting of k consecutive $(g-1)$'s is the number

$$g^{k+1} - g = \frac{g}{a}\cdot a(g^k - 1)\,.$$

Consequently, the least positive integer m for which $m\alpha$ contains the block B infinitely often is $a(g^k-1)$, which proves the theorem.

Even more can be said if g is not a prime power.

THEOREM 3.2. *If g is not a prime power, then for every $\varepsilon > 0$ there exists a positive integer $K = K(\varepsilon)$ such that*

$$M(g,k) \ge (1-\varepsilon)g^{k+1}, \qquad k \ge K\,.$$

Proof. Let p be a prime divisor of g. Since g is not a prime power, $\log p/\log g$ is irrational. Therefore one can find positive integers l and r such that $g^l < p^r < (1+\varepsilon)g^l$. We first claim that the g-ary expansion of no positive multiple of p^r contains the block B, consisting of $r-l$ consecutive $(g-1)$'s, within its r lowest digits. In fact, if this were possible, then by multiplying this multiple of p^r by an appropriate power of g, we would get a number of the form mp^r whose block of r lowest digits starts with the block B. Since g^r is divisible by p^r, we can find such a number with exactly r digits. But then for this number mp^r we have

$$g^r - g^l \le mp^r < g^r\,.$$

As all three numbers involved in the inequality are multiples of p^r, this is inconsistent with the fact that $p^r > g^l$. Thus the minimal mp^r containing a block consisting of $k \ge r-l$ consecutive $(g-1)$'s is at least

$$\sum_{i=l+1}^{l+k} (g-1)g^i = g^{l+1}(g^k-1)\,.$$

Now set

$$\alpha = \left(\frac{p}{g}\right)^r \sum_{j=1}^{\infty} g^{-n_j}\,,$$

where $n_{j+1} - n_j \to \infty$. The foregoing discussion implies that the smallest m for which $m\alpha$ contains the block consisting of k $(\ge r-l)$ consecutive $(g-1)$'s infinitely often is at least

$$\frac{g^{l+1}(g^k-1)}{p^r} > \left(1-\frac{\varepsilon}{2}\right)(g^{k+1}-g) > (1-\varepsilon)g^{k+1}$$

for sufficiently large k. This completes the proof.

Example 3.1. For $g = 10$, taking $p = 2,\ l = 0,\ r = 1$ in the proof we see that $M(10,k) \ge 5 \cdot (10^k - 1)$ for $k \ge 1$. Taking $p = 5,\ l = 2,\ r = 3$, we obtain a better result, namely $M(10,k) \ge 8 \cdot (10^k - 1)$ for $k \ge 1$. With $p = 2,\ l = 3,\ r = 10$ we get $M(10,k) \ge 9.765 \cdot (10^k - 1)$ for $k \ge 7$.

We do not know whether it is true in general that $M(g,k) < g^{k+1}$. However, except for the cases $g = 2$ and $g = 3$, mentioned earlier, we never have $M(g,1) = g-1$. In view of Theorem 3.1 we have to prove this assertion only for prime g. The following theorem includes this case.

Theorem 3.3. *Let $g \ge 5$ be an odd integer. Then*

$$M(g,1) \ge \tfrac{3}{2}(g-1).$$

Proof. Take

$$\alpha = \tfrac{1}{2} + \sum_{j=1}^{\infty} g^{-n_j},$$

where $n_{j+1} - n_j \to \infty$. One easily writes down the g-ary expansion of α and of multiples $m\alpha$. It is easily checked that if $g \equiv 1 \pmod 4$ then the digit $(g-3)/2$ appears at most finitely many times in the expansion of $m\alpha$ for every $m < \frac{3}{2}(g-1)$. The same is true for the digit $g-2$ if $g \equiv 3 \pmod 4$. This proves the theorem.

4. The polynomial Mahler theorem. In this section we prove Theorem 1.2.

Define (for the purposes of this section) the *complexity* of a polynomial $P(x) = a_0 + a_1x + \ldots + a_dx^d$ with real coefficients as the least common denominator of the numbers $a_1, \ldots, a_d$ if they are all rational and as ∞ otherwise.

Lemma 4.1. *Given $\varepsilon > 0$ and a positive integer d, there exists a positive integer M such that for every polynomial P of degree d with complexity at least M, the set $\{P(n) : n \in \mathbb{N}\}$ is ε-dense modulo* 1.

Proof. If P is of infinite complexity, then the sequence $(P(n))_{n=1}^{\infty}$ is uniformly distributed modulo 1, and in particular dense modulo 1. Suppose therefore that the coefficients $a_1, \ldots, a_d$ of P are all rational, and let Q be the complexity of P. Let $x_n = \{P(n)\}$ be the fractional part of $P(n)$, $n = 1, \ldots, Q$. We ought to show that the set $\{x_n : 1 \le n \le Q\}$ is ε-dense in $[0,1]$ if Q is large enough. We shall prove, moreover, that even the discrepancy $D_Q = D_Q(x_1, \ldots, x_Q)$ must be small as Q becomes large. Indeed, according to LeVeque's Inequality (see, for example, [KN, Ch. 1, Th. 2.4]) we have

$$D_Q \le \left(\frac{6}{\pi^2} \sum_{h=1}^{\infty} \frac{1}{h^2} \left| \frac{1}{Q} S(h,Q,P) \right|^2 \right)^{1/3},$$

where

$$S(h,Q,P)=\sum_{n=1}^{Q}\exp(2\pi ihP(n))\,.$$

Setting $h'=h/(h,Q)$ and $Q'=Q/(h,Q)$, and employing some well-known estimates of exponential sums involving a rational polynomial [C], [S] (which are, up to a multiplicative constant, best possible; Hua's original estimates [H, Ch. 7, Th. 10.1] would suit our purposes as well), we obtain

$$|S(h,Q,P)|=(h,Q)\cdot|S(h',Q',P)|\le hC_1(d)Q'^{1-1/d},$$

where $C_1(d)$ depends only on d. Thus

$$D_Q\le\left(C_2(d)\sum_{h=1}^{\infty}\frac{1}{h^2}\left(\frac{Q}{h}\right)^{-2/d}\right)^{1/3}\le C_3(d)Q^{-1/(2d)}\,.$$

Consequently, if Q is sufficiently large, then $D_Q<\varepsilon$, which completes the proof.

Proof of Theorem 1.2. Let B be a g-block of an arbitrary length k. Write:

$$P(x)=a_0+a_1x+\ldots+a_dx^d\,.$$

Let a_l $(1\le l\le d)$ be an irrational coefficient of P. Let $\varepsilon=1/g^{k+2}$. Take M as in Lemma 4.1. One easily verifies that the set of limit points modulo 1 of the sequence $(g^na_l)_{n=1}^{\infty}$ is infinite, whence there exists a sequence (n_j) such that $g^{n_j}a_l\to b_l\pmod 1$ where b_l is either irrational or is a rational number with denominator at least M. Replacing (n_j) by a subsequence thereof, we may assume that each of the subsequences $(g^{n_j}a_i)$, $1\le i\le d$, converges modulo 1, say $g^{n_j}a_i\to b_i\pmod 1$. Consider the polynomial

$$P_0(x)=b_0+b_1x+\ldots+b_dx^d\,.$$

By Lemma 4.1 we can find a positive integer m such that the g-ary expansion of the number $P_0(m)$ modulo 1 starts with the block $B01$. But then the number $P(m)$ contains the block B infinitely often in its expansion. This proves the theorem.

References

[AP] N. Alon and Y. Peres, *Uniform dilations*, J. Geom. Funct. Anal. 2 (1992), 1–28.

[B] M. Boshernitzan, *Uniform distribution and Hardy fields*, J. Analyse Math., to appear.

[C] J. R. Chen, *On Professor Hua's estimate of exponential sums*, Sci. Sinica 20 (1977), 711–719.

[G] S. Glasner, *Almost periodic sets and measures on the torus*, Israel J. Math. 32 (1979), 161–172.

[H] L. K. Hua, *Introduction to Number Theory*, Springer, New York, 1982.
[KN] L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, Wiley, New York, 1974.
[M] K. Mahler, *Arithmetical properties of the digits of the multiples of an irrational number*, Bull. Austral. Math. Soc. 8 (1973), 191–203.
[S] S. B. Stečkin, *An estimate of a complete rational trigonometric sum*, Analytic Number Theory, Mathematical Analysis and their Applications, Trudy Mat. Inst. Steklov. 143 (1977), 188–207, 211; English transl.: Proc. Steklov Inst. Math. (1980), 201–220.
[W] H. Weyl, *Über die Gleichverteilung von Zahlen mod. Eins*, Math. Ann. 77 (1916), 313–352.

DEPARTMENT OF MATHEMATICS
AND COMPUTER SCIENCE
BEN-GURION UNIVERSITY
BEER-SHEVA 84105
ISRAEL

DEPARTMENT OF MATHEMATICS
RICE UNIVERSITY
HOUSTON, TEXAS 77251
U.S.A.

Received on 10.8.1992 (2292)