Proceedings of the Edinburgh Mathematical Society (1979), 22, 145–160 ©

EXCEPTIONAL SETS IN UNIFORM DISTRIBUTION

by R. C. BAKER
(Received 26th November 1977)

1. Introduction

Let B be a measurable set of real numbers in $(0, 1)$ of Lebesgue measure $|B|$ and let $x_1, \ldots, x_n$ be real. Then

$$Z(B; x_1, \ldots, x_n)$$

denotes the number of j $(1 \leqq j \leqq n)$ for which the fractional part $\{x_j\} \in B$. The *discrepancy* of $x_1, \ldots, x_n$ is

$$D(x_1, \ldots, x_n) = n^{-1} \sup_I |Z(I; x_1, \ldots, x_n) - n|I||$$

where the supremum is taken over all intervals I in $[0, 1]$.

Let $g_1(x)$, $g_2(x), \ldots$ be a sequence of differentiable functions on the finite interval $[\alpha, \beta]$. Throughout the paper we assume that $g_1'(x)$ and $g_k'(x) - g_j'(x)$ are positive and monotonic non-decreasing in $[\alpha, \beta]$ whenever $k > j \geqq 1$. We also assume that for some $p > 0$, $C > 0$,

$$g_k'(\beta) \leqq Ck^p \qquad (k \geqq 1), \tag{1}$$

and that there are numbers $c > 0$ and a, $0 \leqq a < 1$, such that

$$g_k'(x) - g_j'(x) \geqq c \tag{2}$$

whenever $j \geqq 1$ and $k \geqq j + Cj^a$. Evidently $p \geqq 1 - a$. We write

$$F(B, n, x) = Z(B; g_1(x), \ldots, g_n(x)) - n|B|$$

for $n \geqq 1$, $\alpha \leqq x \leqq \beta$, and

$$D(m, n, x) = D(g_{m+1}(x), \ldots, g_{m+n}(x))$$

for $m \geqq 0$, $n \geqq 1$. In this paper we are interested in the *exceptional sets*

$$E_q = \{x \in [\alpha, \beta] : \limsup_{n\to\infty} n^q D(0, n, x) > 0\}$$

and

$$E(B) = \{x \in [\alpha, \beta] : \limsup_{n\to\infty} n^{-1} |F(B, n, x)| > 0\}.$$

To make sure that $|E(B)| = 0$ we consider only open sets B with 'thin tail', that is

$$B = I_1 \cup I_2 \cup \ldots \cup I_n \cup \ldots \tag{3}$$

where I_1, I_2,.. are the distinct component intervals of B arranged in order of decreasing length, and

$$b(B) = \liminf_{m\to\infty} \frac{\log |I_m|^{-1}}{\log m} > 1.$$

The Hausdorff dimension of a real set A is written dim A.

Theorem 1. *We have*

$$\dim E_q \leqq 1-(1-a-2q)/(p+2q) \quad (0<q<\tfrac{1}{2}(1-a)).$$

This improves my previous upper bound (2, 4), which was

$$\min\{1-(1-a-2q)/(p+2q+\tfrac{1}{2}(1-a)),\ 1-(1-a-3q)/(p+2q)\}.$$

Theorem 2. *Let f denote the polynomial*

$$f(y)=(by-1)(py+1-a-p)-p(3-y)(1-y).$$

For $b>1$ let $t=\max(b^{-1}, 1-(1-a)/p)$. Then, since $f(t)<0$, $f(1)>0$, and $f'(y)>0$ $(t \leqq y \leqq 1)$, f has a unique zero γ in $(t, 1)$. We have

$$\dim E(B) \leqq \gamma \qquad \textit{whenever} \qquad b(B) \geqq b.$$

Theorem 2 improves Theorem 5.1 of (3). Let $g_j(x)=a_jx$ where $a_1, a_2,$.. is a strictly increasing sequence of positive integers. Then $a=0$ and

$$f(y)=(by-1)(py+1-p)-p(3-y)(1-y).$$

In this particular case, it is shown in (3) that

$$\dim E(B) \leqq \delta \qquad \text{whenever} \qquad b(B) \geqq b,$$

where δ is the unique zero in $(t, 1)$ of the smaller function

$$F(y)=(by-1)(py+1-p)-p(5-y)(1-y), \quad \text{so that } \delta>\gamma.$$

It seems highly unlikely that Theorem 2 is best possible, but Theorem 1 might be. Some examples in Section 4 yield bounds beyond which the theorems cannot be improved.

Theorem 3. *Let ψ be a function on the positive integers such that*

$$Kk^{-\gamma} \leqq \psi(k) \leqq 1 \qquad (k=1,2\ldots)$$

for some $K>0$ and γ, $0<\gamma<1$. Write $\Psi(n)=\Sigma_{k=1}^{n}\psi(k)$.
Let $a_1, a_2,$.. be strictly increasing positive integers with

$$a_k \leqq Ck^p \qquad (k \geqq 1)$$

for some $C>0$ and $p \geqq 1$. Let $\alpha_1, \alpha_2,$.. be real numbers. Write $N(n, x)$ for the number of solutions $k \leqq n$ of

$$\{a_kx-\alpha_k\}<\psi(k).$$

Then

$$N(n, x) \sim \Psi(n) \qquad \text{as} \qquad n \to \infty$$

except for a set of x in [0, 1] *of Hausdorff dimension at most*

$$1-(1-\gamma)/(p+2\gamma).$$

Theorem 3 refines a result of LeVeque (Theorem 3 of (9)) that $N(n, x) \sim \Psi(n)$ for almost all x. I have little idea how far the upper bound obtained could be sharpened.

2. Some lemmas

In this section we collect together some preliminary results. Lemma 1 is similar to a result on p. 106 of (7). For a real set A, $A(\text{mod } 1)$ denotes the set of fractional parts $\{x\}$ $(x \in A)$.

Lemma 1. *Let* $x_1, y_1, \ldots, x_n, y_n$ *be real, then*

$$|D(x_1, \ldots, x_n) - D(y_1, \ldots, y_n)| \leqq 2 \max_i |x_i - y_i|.$$

Proof. Write $d = \max_i |x_i - y_i|$. Let I be a subinterval of [0, 1] with endpoints a, b $(a < b)$ and write $J = [a-d, b+d]$ (mod 1). Then, if K denotes the complement of J in [0, 1],

$$Z(J; y_1, \ldots, y_n) + Z(K; y_1, \ldots, y_n) = n = n(|J| + |K|),$$

or

$$Z(J; y_1, \ldots, y_n) - n\,|J| = -(Z(K; y_1, \ldots, y_n) - n|K|).$$

Either J or K is an interval, so

$$Z(J; y_1, \ldots, y_n) - n|J| \leqq nD(y_1, \ldots, y_n).$$

Now it is clear that

$$\begin{aligned} Z(I; x_1, \ldots, x_n) - n|I| &\leqq Z(J; y_1, \ldots, y_n) - n\,|I| \\ &\leqq Z(J; y_1, \ldots, y_n) - n|J| + n(|J| - |I|) \\ &\leqq nD(y_1, \ldots, y_n) + 2\,nd. \end{aligned}$$

A similar argument shows that

$$Z(I; x_1, \ldots, x_n) - n|I| \geqq -nD(y_1, \ldots, y_n) - 2\,nd.$$

Therefore

$$nD(x_1, \ldots, x_n) \leqq nD(y_1, \ldots, y_n) + 2\,nd.$$

Reversing the roles of x's and y's, the lemma follows.

Lemma 2. *Let F be a non-negative function on* $[\alpha, \beta]$. *Suppose*

$$|F(x) - F(y)| \leqq U|y - x| \qquad (\alpha \leqq x \leqq \beta)$$

and

$$\int_\alpha^\beta F^2(x)\,dx \leqq V.$$

Let

$$E = \{x \in [\alpha, \beta] : F(x) \geqq d > 0\}.$$

There is a covering of E with intervals $J_1, \ldots, J_h$ *such that for* $0 < \sigma \leqq 1$,

$$\sum_{j=1}^{h} |J_j|^\sigma < C_1(1 + UVd^{-3})^{1-\sigma}(Vd^{-2})^\sigma,$$

where C_1 *is a numerical constant.*

Proof. This is a slight variant of Lemma 1 of (4).

Lemma 3. *Let g be a function on* $[\alpha, \beta]$ *whose derivative is monotonic non-decreasing with*

$$0 < G \leqq g'(x) \leqq H \quad (\alpha \leqq x \leqq \beta). \tag{4}$$

Let I be an interval in $[0, 1]$ *and let*

$$F = \{x \in [\alpha, \beta] : \{g(x)\} \in I\}.$$

Then F comprises intervals $J_1, \ldots, J_m$ *with*

$$\sum_{j=1}^{m} |J_j|^\sigma < C_2|I|^\sigma(H^{1-\sigma} + G^{-\sigma})$$

for $0 < \sigma \leqq 1$, *where* C_2 *depends only on* α, β.

Proof. We have

$$F = [u_1, v_1] \cup [u_2, v_2] \cup \ldots \cup [u_m, v_m]$$

where $m \geqq 0$, $\alpha \leqq u_1 \leqq v_1 < u_2 < v_2 < \ldots < u_m \leqq v_m \leqq \beta$,

$$g(v_j) - g(u_j) = |I| \quad (1 < j < m), \tag{5}$$

$$g(v_j) - g(v_{j-1}) = 1 \quad (1 < j < m), \tag{6}$$

$$\max(g(v_1) - g(u_1), g(v_m) - g(u_m)) \leqq |I|. \tag{7}$$

Suppose for a moment that $m > 2$. As g is a convex function,

$$\frac{g(v_j) - g(v_{j-1})}{v_j - v_{j-1}} \leqq \frac{g(v_j) - g(u_j)}{v_j - u_j} \quad (1 < j < m),$$

or in view of (5), (6),

$$v_j - u_j \leqq (v_j - v_{j-1})|I| \quad (1 < j < m).$$

https://doi.org/10.1017/S0013091500016266 Published online by Cambridge University Press

Thus

$$\sum_{1<j<m} (v_j - u_j)^\sigma \leqq |I|^\sigma \sum_{1<j<m} (v_j - v_{j-1})^\sigma$$
$$\leqq |I|^\sigma (m-2)^{1-\sigma} \Big(\sum_{1<j<m} (v_j - v_{j-1}) \Big)^\sigma$$

using Hölder's inequality. But

$$\sum_{1<j<m} (v_j - v_{j-1}) \leqq \beta - \alpha,$$
$$m - 2 = g(v_{m-1}) - g(v_1) \leqq (\beta - \alpha) H$$

in view of (4). Thus (even if $m \leqq 2$)

$$\sum_{1<j<m} (v_j - u_j)^\sigma \leqq |I|^\sigma (\beta - \alpha) H^{1-\sigma}.$$

Moreover,

$$(v_1 - u_1)^\sigma + (v_m - u_m)^\sigma \leqq 2(G^{-1}|I|)^\sigma$$

from (4), (7). This proves the lemma.

Lemma 4. *Let B be an open set in $[0, 1]$. Suppose there are measurable sets $G_1, G_2, \ldots,$ such that*

(i) *B is the union of G_m and m intervals $J_{m1}, \ldots, J_{mm}$,*

(ii) $c = \liminf_{m \to \infty} \dfrac{\log |G_m|^{-1}}{\log m} > 0.$

Then $b(B) \geqq c + 1$.

Proof. We have, for $\epsilon > 0$,

$$|G_m| < m^{-c+\epsilon}$$

for sufficiently large m. In the notation of (3), let $I_{j(1)}, \ldots, I_{j(m)}$ be the component intervals containing $J_{m1}, \ldots, J_{mm}$ respectively. Then for large m,

$$m|I_{2m}| \leqq \sum_{k>m} |I_k| \leqq |B| - |I_{j(1)} \cup \ldots \cup I_{j(m)}|$$
$$\leqq |G_m| < m^{-c+\epsilon},$$

so that

$$|I_{2m}|^{-1} > m^{c+1-\epsilon}.$$

Obviously $b(B) \geqq c + 1$.

Lemma 5. *For $m \geqq 0$, $n \geqq 1$, we have*

$$\int_\alpha^\beta (nD(m, n, x))^2 \, dx \leqq C_3 n (m+n)^a \log^2(n+1),$$

where C_3 is independent of m and n.

https://doi.org/10.1017/S0013091500016266 Published online by Cambridge University Press

Proof. This is established on p. 424 of (4).

We introduce the notation

$$n_k = [\exp(k^{1/2})] \quad (k \geqq 1).$$

The significant properties of this integer sequence are that $n_{k+1}/n_k \to 1$ as $k \to \infty$ and

$$\sum_{k \geqq 1} n_k^{-\epsilon} < \infty$$

for every $\epsilon > 0$.

Lemma 6. *We have*

$$D(0, n_k, x) < n_k^{-q}$$

for sufficiently large k, except for a set of x in $[\alpha, \beta]$ *of Hausdorff dimension at most*

$$1 - (1 - a - 2q)/(p + q) \quad (0 < q < \tfrac{1}{2}(1 - a)).$$

Proof. It suffices to show that whenever

$$1 > \sigma > 1 - (1 - a - 2q)/(p + q),$$

the set $A(n)$ of x in $[\alpha, \beta]$ for which

$$D(0, n, x) \geqq n^{-q}$$

can be covered by intervals $J(n, 1)$, $J(n, 2), \ldots$ with

$$\sum_{j \geqq 1} |J(n, j)|^{\sigma} \leqq C_4 n^{-\epsilon} \quad (n \geqq 1) \tag{8}$$

where $\epsilon > 0$ and C_4 are independent of $n \geqq 1$. For then, given $K \geqq 1$, the set A of x belonging to infinitely many $A(n_k)$ can be covered by the family of intervals

$$J(n_k, j) \; (j \geqq 1, k \geqq K).$$

We have

$$\sum_{k \geqq K} \sum_{j \geqq 1} |J(n_k, j)|^{\sigma} \leqq C_4 \sum_{k \geqq K} n_k^{-\epsilon} \to 0$$

as $K \to \infty$, yielding $\dim A \leqq \sigma$, and indeed $\dim A \leqq 1 - (1 - a - 2q)/(p + q)$.

To get these coverings, we apply Lemma 2 with $F(x) = n\,D(0, n, x)$, $d = n^{1-q}$, so that we may take

$$V = C_3 n^{1+a} \log^2(n + 1)$$

in view of Lemma 5, and

$$U = 2n \max_{j \leqq n} g_j'(\beta) \leqq 2Cn^{p+1}$$

in view of Lemma 1. Thus $A(n)$ may be covered by intervals $J(n, 1)$, $J(n, 2), \ldots$ with

https://doi.org/10.1017/S0013091500016266 Published online by Cambridge University Press

$$\sum_{j \geqq 1} |J(n, j)|^\sigma \leqq C_5 \log^2(n+1)(n^{2+a+p-3(1-q)})^{1-\sigma}(n^{1+a-2(1-q)})^\sigma$$

where C_5 is independent of n. We obtain (8) on noting that

$$p + 3q - (1-a) < \sigma(p+q).$$

This proves Lemma 6.

Lemma 7. *Let $h(m, n, x)$ $(m \geqq 0, n \geqq 1)$ be functions satisfying the following conditions on $[\alpha, \beta]$:*

$$|h(m, n, x) - h(m, n, y)| \leqq d(m, n)|y - x|,$$

$$d(m, n) \sup |h(m, n, x)| \leqq C_6 k^\sigma n^{\mu-\sigma} \quad (0 \leqq m \leqq k, 1 \leqq n \leqq k), \tag{9}$$

$$\int_\alpha^\beta h^2(m, n, x)dx \leqq C_7 k^\rho n^{\nu-\rho} \quad (0 \leqq m \leqq k, 1 \leqq n \leqq k).$$

Here C_6, C_7, σ, μ, ν, ρ are independent of k, n, with $\mu \geqq \sigma + 1$, $\nu \geqq \rho + 1$. Suppose further that

$$nD(m, n, x) \leqq h(m, n, x) \quad (m \geqq 0, n \geqq 1, \alpha \leqq x \leqq \beta).$$

Then if $0 < \lambda < \min(\frac{1}{2}\mu, \frac{1}{4}(\nu + \mu))$ we have

$$D(0, n, x) < n^{\lambda-1}$$

for sufficiently large n, except for a set of x of Hausdorff dimension at most

$$(\mu + \nu - 4\lambda)/(\mu - 2\lambda).$$

Proof. This is a slight variant of Theorem 4 of (4).

Lemma 8. *Suppose that*

$$\limsup_{n \to \infty} n^{-1}|F(B, n, x)| > 0.$$

Then

$$\limsup_{k \to \infty} n_k^{-1}|F(B, n_k, x)| > 0.$$

Proof. Let $n \geqq n_1$, then $n_k \leqq n \leqq n_{k+1}$ for some $k \geqq 1$. We clearly have

$$F(B, n_k, x) + n_k|B| \leqq F(B, n, x) + n|B| \leqq F(B, n_{k+1}, x) + n_{k+1}|B|$$

so that

$$n_k^{-1}F(B, n_k, x) - (n_{k+1} - n_k)n_k^{-1} \leqq n^{-1}F(B, n, x) \leqq n_k^{-1}F(B, n_{k+1}, x) + (n_{k+1} - n_k)n_k^{-1}.$$

If $n_k^{-1}F(B, n_k, x) \to 0$ as $k \to \infty$ then, in view of $n_{k+1}/n_k \to 1$ as $k \to \infty$, we evidently have $n^{-1}F(B, n, x) \to 0$ as $n \to \infty$. This proves the lemma.

Lemma 9. *Let Q be a Borel set in $[\alpha, \beta]$ having Hausdorff dimension greater than σ, then there is a positive measure μ supported on Q such that*

$$\mu([x, y]) \leqq (y-x)^{\sigma} (\alpha \leqq x < y \leqq \beta). \tag{10}$$

Proof. By Theorems 47 and 48 of (9), Q has a compact subset of positive measure with respect to the function t^{σ}. The existence of μ now follows from Theorem 3 of Chapter II of (6).

3. Proofs of Theorems 1 and 2

The new idea in the proof of Theorem 1 is to use the smoothness of $D(m, n, x)$ (Lemma 1) rather than smoothness of a trigonometric sum that majorizes $D(0, n, x)$ (as in (2, 4)).

Proof of Theorem 1. We apply Lemma 7, taking $h(m, n, x) = nD(m, n, x)$. In view of Lemma 5, we may take

$$\rho = a, \quad \nu = 1 + a + \epsilon,$$

for any $\epsilon > 0$. In view of Lemma 1, we may take

$$d(m, n) = 2n \max_{m+1 \leqq j \leqq m+n} g_j'(\beta) \leqq 2n(m+n)^p,$$

and thus (9) holds with

$$\sigma = p, \quad \mu = 2 + p.$$

Write $\lambda = 1 - q$, where $0 < q < \frac{1}{2}(1-a)$. The condition

$$\lambda < \min(\tfrac{1}{2}\mu, \tfrac{1}{4}(\mu + \nu)) = \min(\tfrac{1}{2}(2+p), \tfrac{1}{4}(3 + p + a + \epsilon))$$

is satisfied because $p \geqq 1 - a$. Thus

$$D(0, n, x) < n^{-q}$$

for sufficiently large n, except for a set of Hausdorff dimension at most

$$\frac{\mu + \nu - 4\lambda}{\mu - 2\lambda} = \frac{p + 4q - (1-a) + \epsilon}{p + 2q}.$$

Theorem 1 follows immediately.

In Theorem 2, the improvement of the result of (3) is obtained by the device of splitting $E(B)$ into two subsets, so that integrals

$$\int_{\alpha}^{\beta} n^2 D(m, n, x)^2 d\mu(x)$$

are no longer needed.

Proof of Theorem 2. Suppose that

$$\eta = \dim E(B) > \gamma.$$

Then $f(\eta) > 0$, and we can find a positive d that satisfies

$$\frac{p\eta + 1 - a - p}{3 - \eta} > d > \frac{p(1-\eta)}{b\eta - 1}. \tag{11}$$

We write, in the notation of (3),

$$S(n) = \bigcup_{j \leqq n^d} I_j, \; T(n) = \bigcup_{j > n^d} I_j.$$

Thus

$$F(B, n, x) = F(S(n), n, x) + F(T(n), n, x). \tag{12}$$

Now in view of Lemma 8,

$$E(B) = \{x \in [\alpha, \beta] : \limsup_{k \to \infty} n_k^{-1} |F(B, n_k, x)| > 0\}.$$

It follows from (12) that

$$E(B) \subset P \cup Q \tag{13}$$

where

$$P = \{x \in [\alpha, \beta] : \limsup_{k \to \infty} n_k^{-1} |F(S(n_k), n_k, x)| > 0\}$$

and

$$Q = \{x \in [\alpha, \beta] : \limsup_{k \to \infty} n_k^{-1} |F(T(n_k), n_k, x)| > 0\}.$$

We can readily estimate dim P. We have

$$F(S(n), n, x) = \sum_{j \leqq n^d} (Z(I_j; g_1(x), \ldots, g_n(x)) - n|I_j|)$$

so that

$$n^{-1} F(S(n), n, x) \leqq n^d D(0, n, x).$$

It now follows from Lemma 6 that

$$\dim P \leqq 1 - (1 - a - 2d)/(p + d).$$

But, from (11),

$$\dim E(B) = \eta > 1 - (1 - a - 2d)/(p + d) \geqq \dim P. \tag{14}$$

Combining (13) and (14), it is clear that

$$\dim Q \geqq \eta. \tag{15}$$

Now select a number c, $\eta^{-1} < c < b$, and a number σ, $c^{-1} < \sigma < \eta$, such that

$$d > \frac{p(1-\sigma)}{c\sigma - 1}. \tag{16}$$

Since Q is a Borel set having dimension greater than σ, there is a positive measure μ

https://doi.org/10.1017/S0013091500016266 Published online by Cambridge University Press

satisfying (10) supported on Q. We have

$$\int_\alpha^\beta n^{-1}|F(T(n), n, x)|d\mu(x) \leqq \int_\alpha^\beta n^{-1}Z(T(n); g_1(x), \dots, g_n(x))d\mu(x) + \int_\alpha^\beta |T(n)|d\mu(x). \tag{17}$$

Now for large n,

$$|T(n)| = \sum_{j>n^d} |I_j| < \sum_{j>n^d} j^{-c} < n^{-\epsilon} \tag{18}$$

for some $\epsilon > 0$. We need a similar estimate for

$$\int_\alpha^\beta n^{-1}Z(T(n);\ g_1(x), \dots, g_n(x))d\mu(x).$$

Write $E(k, j)$ for the set of x in $[\alpha, \beta]$ such that

$$\{g_k(x)\} \in I_j.$$

Then

$$\int_\alpha^\beta Z(T(n);\ g_1(x), \dots, g_n(x))d\mu(x) = \sum_{k=1}^{n} \sum_{j>n^d} \mu(E(k, j)).$$

We can estimate $\mu(E(k, j))$ by combining (10) with Lemma 3 and (1). We have

$$\mu(E(k, j)) < C_2|I_j|^\sigma(Ck^{p(1-\sigma)} + g_1'(\alpha)^{-\sigma})$$
$$< j^{-c\sigma}n^{p(l-\sigma)} \qquad (j > n^d, 1 \leqq k \leqq n)$$

if n is sufficiently large. Thus

$$\int_\alpha^\beta n^{-1}Z(T(n);\ g_1(x), \dots, g_n(x))d\mu(x) < n^{p(1-\sigma)} \sum_{j>n^d} j^{-c\sigma}$$
$$< 2n^{p(1-\sigma)-d(c\sigma-1)}. \tag{19}$$

The last exponent of n is negative because of (16). Combining (17), (18) and (19), we certainly have

$$\sum_{k=1}^{\infty} \int_\alpha^\beta n_k^{-1}|F(T(n_k), n_k, x)|d\mu(x) < \infty.$$

But then the series

$$\sum_{k=1}^{\infty} n_k^{-1}|F(T(n_k), n_k, x)|$$

converges for almost all x with respect to $d\mu$. Since the series diverges at every point of Q, the support of μ, we have a contradiction. This proves that

$$\dim E(B) \leqq \gamma.$$

4. Examples

(i) Let $0 < p \leqq 1$, $a = 1 - p$. By taking $g_j(x) = [j^p]x$, we show that the bound $4q/(p + 2q)$ of Theorem 1 cannot be reduced below $2q/(p + q)$.

If x is real, write $w(x)$ for the supremum of all η for which

$$\liminf_{q\to\infty} q^{\eta}\|qx\| = 0.$$

Here $\|\cdot\|$ denotes distance from the nearest integer. We write

$$X(\eta) = \{x \in [0, 1] : w(x) \geqq \eta\}.$$

Then for $\eta > 1$,

$$\dim X(\eta) = 2/(\eta + 1).$$

(This was proved by V. Jarnik and A. Besicovitch; (1) is the best reference).

Write $a_j = [j^p]$. Let $\epsilon > 0$. The discrepancy of $a_1x, a_2x, \ldots, a_nx$ satisfies

$$D(0, n, x) > n^{-\epsilon - p/\eta} \tag{20}$$

for infinitely many n, whenever $x \in X(\eta)$. To see this, we follow the argument of Theorem 3.3 of (7), Chapter 2. Suppose $\epsilon < \eta/2$. There are infinitely many positive integers s and corresponding integers t such that

$$|x - t/s| < s^{-1-\eta+\epsilon}.$$

Write $n = [s^{(\eta-2\epsilon)/p}]$. Then for $1 \leqq j \leqq n$,

$$a_jx = k_j/s + \theta_j$$

where k_j is an integer and

$$|\theta_j| < s^{-1-\epsilon}.$$

The interval $I = (s^{-1-\epsilon}, s^{-1} - s^{-1-\epsilon})$ thus contains none of the points $\{a_1x\}, \ldots, \{a_nx\}$, and therefore

$$D(0, n, x) \geqq |I| > \tfrac{1}{2}s^{-1} > \tfrac{1}{4}n^{-p/(\eta-2\epsilon)}$$

for sufficiently large s.

It follows that

$$X(\epsilon + p/q) \subset E_q \qquad (0 < q < p)$$

for any $\epsilon > 0$, and therefore

$$\dim E_q \geqq 2q/(p + q) \qquad (0 < q < p = 1 - a).$$

In case $p = 1$, we can be more precise. The discrepancy of $x, 2x, \ldots, nx$ satisfies

$$D(0, n, x) < n^{\epsilon - 1/\eta}$$

for sufficiently large n, unless $x \in X(\eta)$. This is Theorem 3.2 of (7), Chapter 2. In other words,

$$E_q \subset X(q^{-1} - \epsilon)$$

for $0 < q < 1$, $0 < \epsilon < q^{-1}$. We easily deduce that

$$\dim E_q = 2q/(1 + q) \qquad (0 < q < 1).$$

(ii) Let $p = 2$, $a = 0$. By taking $g_j(x) = j^2x$, we show that the bound $(1+4q)/(2+2q)$ of Theorem 1 cannot be reduced below $2q$.

If x is irrational and $q_1, q_2, \ldots$ are the denominators of the continued fraction of x, write

$$q_{k+1} = q_k^{\mu_k}.$$

Note that $\limsup_{k\to\infty} \mu_k = w(x)$ from the elementary theory of continued fractions. Let $k(1), k(2), \ldots$ be the indices for which $q_k \not\equiv 2 \pmod 4$ and let

$$\theta(x) = \limsup_{j\to\infty} \mu_{k(j)}$$

The following result is easily deduced from Satz XIII of (5) using Koksma's inequality (7, p. 143). If $\theta(x) = \theta > 1$, the discrepancy of $1^2x, 2^2x, \ldots, n^2x$ satisfies

$$D(0, n, x) > n^{-\epsilon - 1/(\theta+1)} \tag{21}$$

for infinitely many n.

Now the techniques of (1) may easily be adapted to show that for $\eta > 1$, the set $Y(\eta) = \{x \in [0, 1] : \theta(x) \geqq \eta\}$ has dimension $2/(\eta+1)$. Since (21) implies

$$Y(q^{-1} - 1 + \epsilon) \subset E_q$$

for $0 < q < \frac{1}{2}$, $\epsilon > 0$, we have

$$\dim E_q \geqq 2q \qquad (0 < q < \tfrac{1}{2}).$$

(iii) Let $b > 1$. Let $a_1 < a_2 < \ldots$ be any integers with $a_{n+1}/a_n \to 1$ as $n \to \infty$, and let $g_j(x) = a_jx$. We shall show that there is an open set G in $(0, 1)$ with

$$b(G) \geqq b, \quad \dim E(G) \geqq b^{-1}.$$

With more calculation, our construction works for $a_n = [n^p]$ $(p > 0)$. Thus the upper bound γ of Theorem 2 could never be reduced below b^{-1}.

To construct G we use the Cantor set $C(\rho)$, where ρ is defined by

$$\log 2/\log \rho^{-1} = b^{-1},$$

so that $0 < \rho < \frac{1}{2}$.

If J is the union of m disjoint closed intervals $[\alpha_i, \beta_i]$, write J^ρ for the union of

$$[\alpha_i, \alpha_i + (\beta_i - \alpha_i)\rho], [\beta_i - (\beta_i - \alpha_i)\rho, \beta_i] \quad (1 \leqq i \leqq m).$$

Thus J^ρ is the union of $2m$ disjoint closed intervals.

Define $J(0), J(1), \ldots$ by induction as follows: $J(0) = [0, 1]$, $J(m) = J(m-1)^\rho$ $(m > 0)$. We readily see that $J(m)$ is the union of 2^m disjoint closed intervals of length ρ^m. It is shown in (6), Chapter III that

$$C(\rho) = \bigcap_{m=1}^{\infty} J(m)$$

has Hausdorff dimension $\log 2/\log \rho^{-1}$. We write $C'(\rho)$ for the set of irrational numbers in $C(\rho)$ that are not endpoints of any interval of $J(m)$ $(m \geqq 1)$.

Our open set G is

$$G = \bigcup_{r=1}^{\infty} K_r,$$

where K_r is the interior of the set $a_r J(a_r^2)$ (mod 1). The number of intervals comprising K_r is at most $2^{a_r^2+1}$. Thus the set $\bigcup_{r=1}^{k} K_r$ comprises h_k intervals, where

$$h_k \leqq \sum_{r=1}^{k} 2^{a_r^2+1} \leqq k 2^{a_k^2+1}.$$

Let m be a positive integer, $m \geqq h_1$. Then for some $k = k(m)$,

$$h_k \leqq m \leqq h_{k+1}.$$

We can express G as the union of m intervals $J_{m1}, \ldots, J_{mm}$ with the set

$$G_m = \bigcup_{r>k} K_r.$$

Moreover, for large m,

$$\begin{aligned} |G_m| &\leqq \sum_{r>k} |K_r| \leqq \sum_{r>k} a_r |J(a_r^2)| \\ &\leqq \sum_{r>k} a_r (2\rho)^{a_r^2} < a_k (2\rho)^{a_k^2} \end{aligned}$$

in view of

$$a_{j+1}(2\rho)^{a_{j+1}^2} < \tfrac{1}{2} a_j (2\rho)^{a_j^2} \qquad \text{for large } j.$$

Now

$$\begin{aligned} \frac{\log |G_m|^{-1}}{\log m} &\geqq \frac{-a_k^2 \log 2\rho - \log a_k}{\log h_{k+1}} \\ &\geqq \frac{a_k^2(\log \rho^{-1} - \log 2) - \log a_k}{(1 + a_{k+1}^2)\log 2 + \log(k+1)} \end{aligned}$$

so that

$$\liminf_{m\to\infty} \frac{\log |G_m|^{-1}}{\log m} \geqq \frac{\log \rho^{-1}}{\log 2} - 1 = b - 1.$$

It follows from Lemma 4 that $b(G) \geqq b$.

We now observe that if $x \in C'(\rho)$, then $\{a_r x\} \in K_r$ for $r \geqq 1$. Hence $\{a_r x\} \in G$ for $r \geqq 1$. Obviously

$$C'(\rho) \subset E(G),$$

and it follows that $\dim E(G) \geqq b^{-1}$.

5. Proof of Theorem 3

We use a lemma of a rather different nature from those in Section 2. Let $d(m)$ denote the number of divisors of a positive integer m and (s, t) the greatest common

https://doi.org/10.1017/S0013091500016266 Published online by Cambridge University Press

divisor of positive integers s and t. If I is an interval of the real line write E_I for the union of all intervals $I+u$ (u integer) and $X(I,x)$ for the indicator function of E_I.

Lemma 10. *For any intervals $J_1, \ldots, J_n$ of length $\leqq 1$,*

$$\int_0^1 \left\{\sum_{k=1}^n (X(J_k, a_k x) - |J_k|)\right\}^2 dx \leqq 2 \sum_{k=1}^n |J_k| d(a_k).$$

Proof. It is shown on p. 217 of (8) that

$$\int_0^1 \left\{\sum_{k=1}^n (X(J_k, a_k x) - |J_k|)\right\}^2 dx \leqq 2 \sum_{k=1}^n |J_k| a_k^{-1} \sum_{j=1}^k (a_j, a_k),$$

and on p. 219 of the same paper that

$$\sum_{j=1}^k (a_j, a_k) \leqq a_k d(a_k).$$

Lemma 10 follows on combining these two inequalities.

We introduce some further notations. Let $\rho(y, A)$ denote the distance from the real number y to the set A. If I is an interval with endpoints a, b ($a < b$), and $\delta > 0$, we write I_δ for the interval $[a - \delta|I|, b + \delta|I|]$. Define

$$Y(I, \delta, x) = \max\{0, 1 - (\delta|I|)^{-1}\rho(x, E_I)\}$$

and

$$Z(I, \delta, x) = X(I_\delta, x).$$

It is clear that for any real x,

$$X(I, x) \leqq Y(I, \delta, x) \leqq Z(I, \delta, x). \tag{22}$$

Proof of Theorem 3. There are intervals $I_1, I_2, \ldots$, with $|I_j| = \psi(j)$ such that

$$N(n, x) = \sum_{j=1}^n X(I_j, a_j x).$$

Let $\epsilon > 0$. We shall show that

$$\limsup_{k\to\infty} \Psi(n_k)^{-1} \sum_{j=1}^{n_k} Y(I_j, n_k^{-\epsilon}, a_j x) \leqq 1 \tag{23}$$

except for a set W of x having dimension at most

$$(p + 3\gamma + 7\epsilon - 1)/(p + 2\gamma).$$

It follows from (22) and (23) that

$$\limsup_{k\to\infty} \Psi(n_k)^{-1} N(n_k, x) \leqq 1 \tag{24}$$

outside W. Taking ϵ arbitrarily close to 1 we find that (24) holds outside a set of dimension at most $1 - (1-\gamma)/(p + 2\gamma)$. A similar argument applies to $\liminf_{k\to\infty} \Psi(n_k)^{-1} N(n_k, x)$. We can now complete the proof by arguing as in Lemma 8. Thus it suffices to consider (23).

Write

$$M(n,x)=\max\left\{0,\sum_{j=1}^{n}\left(Y(I_j,n^{-\epsilon},a_jx)-\int_0^1 Z(I_j,n^{-\epsilon},t)dt\right)\right\}$$

and

$$P(n,x)=\sum_{j=1}^{n}\left(Z(I_j,n^{-\epsilon},a_jx)-\int_0^1 Z(I_j,n^{-\epsilon},t)dt\right).$$

Then in view of (22), whenever $M(n,x)\neq 0$ we have

$$0<M(n,x)\leqq P(n,x),$$

hence

$$\int_0^1 M(n,x)^2dx\leqq\int_0^1 P(n,x)^2dx.$$

We now apply Lemma 10, together with upper bounds for $d(m)$ and a_k, to get

$$\int_0^1 P(n,x)^2dx\leqq 2\Psi(n)(1+2n^{-\epsilon})\max_{j\leqq n} d(a_j)<\Psi(n)n^{\epsilon}$$

for sufficiently large n. We also observe that for any I, $\delta>0$, and real x, y,

$$Y(I,\delta,x)-Y(I,\delta,y)=\sum_{j=1}^{r}\int_{V_j}\pm(\delta|I|)^{-1}dt$$

where $V_1,\dots,V_r$ are intervals of total length $\leqq|y-x|$. Consequently if x, y are real,

$$\begin{aligned}|M(n,x)-M(n,y)|&\leqq\left|\sum_{j=1}^{n}\{Y(I_j,n^{-\epsilon},a_jx)-Y(I_j,n^{-\epsilon},a_jy)\}\right|\\&\leqq n^{\epsilon}\sum_{j=1}^{n}|I_j|^{-1}a_j|y-x|\\&\leqq CK^{-1}n^{p+\gamma+1+\epsilon}|y-x|.\end{aligned}$$

We now apply Lemma 2 with $[\alpha,\beta]=[0,1]$, $F(x)=M(n,x)$, $U=CK^{-1}n^{p+\gamma+1+\epsilon}$, $V=\Psi(n)n^{\epsilon}$ and $d=\Psi(n)n^{-\epsilon}$. For large n we have a covering of

$$\{x\in[0,1]:\quad M(n,x)\geqq\Psi(n)n^{-\epsilon}\}$$

by intervals $J_{n1},J_{n2},\dots$ such that for $0<\sigma<1$,

$$\sum_{j\geqq 1}|J_{nj}|^{\sigma}<C_8(n^{p+\gamma+1+5\epsilon}\Psi^{-2}(n))^{1-\sigma}(n^{3\epsilon}\Psi^{-1}(n))^{\sigma}$$

where C_8 is independent of n. Since

$$\Psi(n)>n^{1-\gamma-\epsilon}$$

for large n, we have

$$\sum_{j\geqq 1}|J_{nj}|^{\sigma}<C_8n^{p+3\gamma-1+7\epsilon-\sigma(p+2\gamma)}.$$

If $\sigma > (p + 3\gamma + 7\epsilon - 1)/(p + 2\gamma)$, the exponent of n is negative. Arguing as in the proof of Lemma 6 it follows that

$$M(n_k, x) < \Psi(n_k)n_k^{-\epsilon} \quad (k \geqq k_0(x)) \tag{25}$$

except for a set of x of dimension at most $(p + 3\gamma + 7\epsilon - 1)/(p + 2\gamma)$. Since (25) implies (23), this completes the proof of Theorem 3.

REFERENCES

(1) A. BAKER and W. M. SCHMIDT, Diophantine approximation and Hausdorff dimension, *Proc. London Math. Soc.* (3) **21** (1970), 1–11.

(2) R. C. BAKER, Slowly growing sequences and discrepancy modulo one, *Acta Arith.* **23** (1973), 279–293.

(3) R. C. BAKER, Khinchin's conjecture and Marstrand's theorem, *Mathematika* **21** (1974), 248–260.

(4) R. C. BAKER, Dyadic methods in the measure theory of numbers, *Trans. Amer. Math. Soc.* **221** (1976), 419–432.

(5) H. BEHNKE, Zur Theorie der diophantischen Approximationen, *Abh. Math. Sem. Univ. Hamburg* **3** (1924), 261–318.

(6) J-P. KAHANE and R. SALEM, *Ensembles parfaits et series trigonometriques* (Hermann, Paris, 1963).

(7) L. KUIPERS and H. NIEDERREITER, *Uniform distribution of sequences* (Wiley, New York, 1974).

(8) W. J. LEVEQUE, On the frequency of small fractional parts in certain real sequences III, *J. für reine und angew. Math.* **202** (1959), 215–220.

(9) C. A. ROGERS, *Hausdorff measures* (Cambridge University Press, 1970).

ROYAL HOLLOWAY COLLEGE
EGHAM
SURREY