Geometric and Functional Analysis
Vol. 2, No. 1 (1992)
1016-443X/92/01001-28$1.50+0.20/0

UNIFORM DILATIONS

N. ALON AND Y. PERES

Abstract

Every sufficiently large finite set X in $[0,1)$ has a dilation $nX \bmod 1$ with small maximal gap and even small discrepancy. We establish a sharp quantitative version of this principle, which puts into a broader perspective some classical results on the distribution of power residues. The proof is based on a second-moment argument which reduces the problem to an estimate on the number of edges in a certain graph . Cycles in this graph correspond to solutions of a simple Diophantine equation: The growth asymptotics of these solutions, which can be determined from properties of lattices in Euclidean space, yield the required estimate.

1. Introduction

Let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ denote the one dimensional torus, i.e., the set of real numbers modulo 1. A subset X of T is called *ε-dense* if it intersects every interval of length ε in $\mathbb{T}$. A *dilation* of X is a subset $nX = \{nx : x \in X\}$ where n is an integer (and each number nx is reduced, of course, modulo 1). S. Glasner proved in [G] that for every infinite subset X of $\mathbb{T}$ and for each $\varepsilon > 0$ there exists an ε-dense dilation nX of X. Motivated by this result, Berend and Peres [BP] defined $k(\varepsilon)$ as the minimal integer k such that for any set $X \subset \mathbb{T}$ of cardinality at least k, some dilation nX is ε-dense. They proved that

$$\Omega\left(\frac{1}{\varepsilon^2}\right) \leq k(\varepsilon) \leq O(1/\varepsilon)^{O(1/\varepsilon)} \tag{1}$$

N.A. – Research supported in part by a U.S.A.-Israel BSF grant.
Y.P. – Partially supported by a Weizmann Postdoctoral Fellowship.
AMS Classification - Primary: 11K38, Secondary: 11K06, 11J13.
Keywords: Discrepancy, density mod 1, dilation, second moment method.

and raised the problem of determining the correct order of magnitude of $k(\varepsilon)$. Here we prove the following result, which determines this order of magnitude almost precisely.

THEOREM 1.1. *For any $\gamma > 0$ there exists a positive $\varepsilon_0 = \varepsilon_0(\gamma)$ such that for $\varepsilon < \varepsilon_0$, every set $X \subset \mathbb{T}$ of cardinality at least $\frac{1}{\varepsilon^{2+\gamma}}$ has an ε-dense dilation nX,*

$$\left(\textit{in other words} \quad \Omega\left(\frac{1}{\varepsilon^2}\right) \leq k(\varepsilon) \leq \frac{1}{\varepsilon^{2+\gamma}}\right) .$$

Our first proof of Theorem 1.1, which is based on a second-moment probabilistic argument, actually yields a stronger result. For a finite sequence $X = \{x_1, \ldots, x_k\}$ in $\mathbb{T}$ and any subinterval of $\mathbb{T}$, write

$$\mathrm{disc}(X, I) = \left|\frac{1}{k}\sum_{j=1}^{k} 1_I(x_j) - \text{length } (I)\right| .$$

Note that here X is a sequence rather than a set, i.e., it may contain repeated elements. Recall from [KN] that the *discrepancy* of X is defined by

$$\mathrm{disc}(X) = \sup_I \mathrm{disc}(X, I)$$

where the supremum is over all intervals in $\mathbb{T}$. Observe that if $\mathrm{disc}(X) < \varepsilon$ then X is ε-dense. Thus the following result extends theorem 1.1. (Extending the theorem in this direction was originally suggested by D. Berend [unpublished], who obtained less precise bounds by different methods prior to our work).

THEOREM 1.2. *For any $\gamma > 0$ there exists an integer $k_0 = k_0(\gamma)$, such that for every $k > k_0$ and every sequence X consisting of k distinct points in $\mathbb{T}$, there is a dilation nX satisfying*

$$\mathrm{disc}(nX) \leq k^{\gamma - 1/2} .$$

This estimate is, of course, sharp up to the γ error-term, by the lower bound in (1).

The proof of theorem 1.1 goes through particularly smoothly when the set X to be dilated consists of rationals with the same prime denominator (see the next section). In the general case, the proof hinges on the following proposition.

PROPOSITION 1.3. *Let $\{x_1, \ldots, x_k\}$ be an arbitrary set of k distinct numbers in the unit interval $[0,1)$. Denote by h_m the number of pairs (i,j) with $1 \le i < j \le k$, such that $m(x_i - x_j)$ is an integer. For each $\alpha > 0$, if k is sufficiently large then for any k points $\{x_1, \ldots, x_k\}$ in $[0,1)$ and for every $m \ge 1$, the partial sum $H_m = \sum_{\ell=1}^{m} h_\ell$ satisfies*

$$H_m \le (km)^{1+\alpha} .$$

The trivial upper bound for H_m defined above is km^2. Simple examples show that the bound above is sharp, up to the α error-term in the exponent.

Actually, in section 5 a slightly tighter upper bound is established (corollary 5.2) and this allows an improvement of the estimate in theorem 1.1 (see proposition 6.1 in section 6). The rest of the paper is organized as follows. In the next section we consider the case in which the set X to be dilated consists of rationals with the same prime denominator p. This case puts into a broader context some classical results and conjectures concerning the distribution of quadratic residues modulo p. A dilation of a set modulo p is its image under a linear function; in section 3 we show that a smaller maximal gap $(\mathrm{mod}\, p)$ may be obtained by considering images of a given set X under polynomials of higher degree. Section 4 contains the proofs of theorems 1.1 and 1.2. In section 5 we establish proposition 1.3, using a simple combinatorial argument which reduces it to an estimate on the growth of solutions of a certain Diophantine equation. In section 6 we present an alternative proof of theorem 1.1 which uses some classical harmonic analysis.

While this approach does not yield the bounds on discrepancy obtained via the probabilistic approach, it seems better suited for obtaining extensions of theorem 1.1 in which restrictions are placed on the multiplier n.

For instance, we show that under the hypotheses of theorem 1.1, there actually exists a *prime* n for which nX is ε-dense (this answers a question suggested by R.L. Graham [private communication]).

If the hypotheses of that theorem are strengthened and we assume the cardinality of X is at least $\left(\frac{1}{\varepsilon}\right)^{4+\gamma}$, then we show that an ε-dense dilation of the form n^2X (for some integer n) exists.

The final section 7 contains a few additional applications of the techniques.

2. Dilating Subsets of $\mathbf{Z}_p$

For p prime and $\frac{1}{p} \leq \varepsilon < 1$, let $k(\varepsilon, p)$ be the minimal integer k such that for any set X of k distinct rationals in $\mathbf{T}$ with the same denominator p, there exists an ε-dense dilation nX. Obviously in this definition we may assume that $\varepsilon p \in \mathbf{Z}$ And reformulate it as follows: $k(\varepsilon, p)$ is the minimal k, such that for every subset X of cardinality k of the cyclic group $\mathbf{Z}_p$, there is $a \in \mathbf{Z}_p$ such that aX intersects every interval of εp consecutive elements of $\mathbf{Z}_p$. In [BP] it is shown, using a simple probabilistic argument, that

$$\sup\{k(\varepsilon, p) \mid p > \frac{1}{\varepsilon},\ p \text{ prime}\} \geq \frac{c}{\varepsilon} \log \frac{1}{\varepsilon} .$$

As observed by the first author of the present paper (cf. [BP]), this lower bound may be slightly improved to $\frac{c}{\varepsilon} \log(1/\varepsilon) \log\log\log(1/\varepsilon)$ by considering the set

$$X = \{j^2/p (\mathrm{mod}\, 1) : j \in \mathbf{Z}\} \tag{2}$$

where $p \equiv 3 \bmod 4$, and invoking the recent result of Graham and Ringrose [GR] which asserts that for infinitely many primes p, the smallest quadratic nonresidue modulo p exceeds $c \log p \log\log\log p$ (this result holds for primes $p \equiv 3 \bmod 4$ as well). Note that the quadratic-residue construction shows that *if* the inequality $k(\varepsilon, p) \leq (1/\varepsilon)^{4/3}$ holds for large primes p, it would imply that the length of the maximum gap between consecutive quadratic residues modulo a prime p congruent to $3 \bmod 4$ does not exceed $p^{1/4}$, and would improve the best known upper estimate of $O(p^{1/4} \log p)$, due to Burgess [Bu2], for this quantity. Similarly, if $k(\varepsilon, p) \leq O\left((1/\varepsilon)^{1+\gamma}\right)$ for every $\gamma > 0$, this would imply the Vinogradov conjecture concerning the smallest quadratic nonresidue modulo such primes. Unfortunately, we are unable to obtain such an upper bound for $k(\varepsilon, p)$, but we present a very simple proof of the following.

PROPOSITION 2.1. *For every prime p and every $\varepsilon > 0$ for which εp is an integer*

$$k(\varepsilon, p) \leq 4/\varepsilon^2 .$$

Proof: We actually prove the following

Reformulation. Let p be prime and let $X = \{x_1, \ldots, x_k\} \subset \mathbf{Z}_p$ be a set of cardinality k. If δp is an integer and $k\delta^2 \geq 1$, then there is an element $a \in \mathbf{Z}_p$ such that the set $aX = \{ax_1, \ldots, ax_k\}$ intersects each interval of length $2\delta p - 1$ in $\mathbf{Z}_p$.

(Note that the proposition follows by applying this reformulation with the largest δ for which $2\delta p - 1 \leq \varepsilon p$).

Let a and b be two random elements of $\mathbf{Z}_p$, chosen independently and uniformly according to a uniform distribution on $\mathbf{Z}_p$. Fix an interval I of length δp in $\mathbf{Z}_p$, and define random variables $Y_1^I, \ldots, Y_k^I$ by letting $Y_j^I = 1$ if $ax_j + b \in I$, and $Y_j^I = 0$ otherwise. Denote $Y^I = \sum_{j=1}^k Y_j$. Obviously the expectation of Y^I is

$$E(Y^I) = \sum_{j=1}^{k} E(Y_j^I) = k\delta \ .$$

The crucial (and simple) fact is that the random variables $ax_1+b, \ldots, ax_k+b$ are pairwise independent and therefore the same holds for the variables $Y_1^I, \ldots, Y_k^I$. Consequently, the variance of Y^I is precisely

$$\mathrm{Var}(Y^I) = \sum_{j=1}^{k} \mathrm{Var}(Y_j^I) = k\delta(1-\delta) \ .$$

By Chebyshev's inequality we conclude that

$$\mathbb{P}\mathrm{rob}\ [Y^I = 0] \leq \frac{\mathrm{Var}(Y^I)}{E(Y^I)^2} = \frac{1-\delta}{k\delta} \ .$$

Let $\mathcal{F}$ be a family of $\lceil 1/\delta \rceil$ intervals (of length δp each) in $\mathbf{Z}_p$ that cover $\mathbf{Z}_p$.

By the above inequality and our hypothesis,

$$\mathbb{P}\mathrm{rob}[\exists I \in \mathcal{F} : Y^I = 0] \leq \lceil 1/\delta \rceil \frac{1-\delta}{k\delta} < \frac{1}{k\delta^2} \leq 1 \ .$$

Hence there exists a choice of a and b such that for each $I \in \mathcal{F}$, we have $Y^I > 0$. Fix such a, b. This implies that every interval of $2\delta p - 1$ elements of $\mathbf{Z}_p$, contains an element $ax_j + b$ for some $1 \leq j \leq k$. Therefore every such interval contains at least one element ax_j, completing the proof. □

The probabilistic argument above is easily extended to cover simultaneous dilation of several sets. This yields the following generalization of proposition 2.1, whose proof is analogous.

PROPOSITION 2.2. *Let p be a prime, and let $X_1, \ldots, X_r$ be subsets of the cyclic group $\mathbf{Z}_p$ with $|X_i| = k_i$. If $\delta p \in \mathbf{Z}$ and*

$$\sum_{i=1}^{r} \frac{1}{k_i} \leq \delta^2 ,$$

then there is an element $a \in \mathbf{Z}_p$ such that each of the sets aX_i intersects each interval of length $2\delta p - 1$ in $\mathbf{Z}_p$.

If G is a subgroup of index r of the multiplicative group $\mathbf{Z}_p^*$, then the last proposition implies that every interval of length $3r\sqrt{p}$ in $\mathbf{Z}_p$ intersects all the cosets of G (since a dilation $\bmod p$ merely interchanges the cosets). Actually, for *cosets* better bounds are available using more sophisticated methods (cf. [BU1]) but nothing near the truth has been established.

The method described above can be used to obtain dilations with small discrepancy.

THEOREM 2.3. *For every set X of k distinct rationals in $\mathbf{T}$ with the same prime denominator p, there exists a dilation nX satisfying* $\mathrm{disc}(nX) = O\left(k^{-1/2}(\log k)^{3/2}\right)$. *(Note that here thinking of nX as a sequence or a set makes no difference, since no repeated elements will occur in any dilation nX for which $n \not\equiv 0 \bmod p$).*

This estimate is sharp, up to the logarithmic factor. This is because by applying it to the quadratic residue construction (2), it implies that for every prime p, the number of quadratic residues in any interval of length ℓ does not deviate from $\ell/2$ by more than $O(\sqrt{p}\,(\log p)^{3/2}))$. However, an easy argument of Schur (cf. [D]) shows that there are intervals in which this deviation is at least $\Omega(\sqrt{p})$. Of course the Polya-Vinogradov inequality (cf. [D]) implies that the number of quadratic residues mod p in any interval of length ℓ does not deviate from $\ell/2$ by more than $O(\sqrt{p}\log p)$. This estimate is better than our estimate above, but the latter holds for dilations of arbitrary subsets of cardinality $p/2$ of $\mathbf{Z}_p$, and not only when a group structure is present.

For a subset $V = \{v_1, \ldots, v_k\}$ of $\mathbf{Z}_p$ and an interval I of cardinality $|I|$ in $\mathbf{Z}_p$, define

$$\mathrm{disc}(V, I) = \left| \frac{1}{k} \sum_{j=1}^{k} 1_I(v_j) - \frac{|I|}{p} \right|$$

and $\mathrm{disc}(V) = \max_I \mathrm{disc}(V, I)$ taken over all intervals I in $\mathbf{Z}_p$. The following proposition implies theorem 2.3.

PROPOSITION 2.4. *Let p be a prime and let $X = \{x_1, \ldots, x_k\} \subset \mathbb{Z}_p$ be a set of cardinality k. If $m \in \mathbb{Z}$ and $k \geq m2^{2m+1}$ then there exists $a \in \mathbb{Z}_p$ such that* $\mathrm{disc}(aX) \leq (m+2)2^{1-m}$.

Proof: Let s be the smallest integer such that $2^s \geq p$. For each $1 \leq i \leq m$, partition the interval of integers $\{0, 1, 2, \ldots, 2^s - 1\}$ into disjoint intervals of length 2^{s-i} each, and denote by Φ_i the family of 2^i intervals so obtained, after all numbers are reduced modulo p. Define $\Phi = \bigcup_{i=1}^{m} \Phi_i$. Let a, b be random elements of $\mathbb{Z}_p$, chosen independently according to a uniform distribution on $\mathbb{Z}_p$. Consider the random set $aX+b = \{ax_1+b, \ldots, ax_k+b\}$ in $\mathbb{Z}_p$. Our objective is to show that with positive probability,

$$\forall I \in \Phi \qquad \mathrm{disc}(aX + b, I) < 2^{-m} \ . \tag{3}$$

To do so, fix an interval $I \in \Phi_i$. For each $j \in \{1, 2, \ldots, k\}$ let $Y_j^I = 1$ if $ax_j + b \in I$ and $Y_j^I = 0$ otherwise.

Define $Y^I = \sum_{j=1}^{k} Y_j^I$.

Obviously $E(Y^I) = k\frac{|I|}{p} = k\frac{2^{s-i}}{p}$ and by pairwise independence $\mathrm{Var}(Y^I) = k\frac{|I|}{p}\left(1 - \frac{|I|}{p}\right) < k\frac{2^{s-i}}{p}$.

Therefore, by Chebyshev's inequality:

$$\mathbb{P}\mathrm{rob}\left[\mathrm{disc}(aX + b,\ I) \geq 2^{-m}\right] = \mathbb{P}\mathrm{rob}\left[|Y - E(Y)| \geq k2^{-m}\right] \leq$$
$$\leq \frac{\mathrm{Var}(Y^I)}{2^{-2m}k^2} < \frac{2^{s-i}}{p2^{-2m}k} \leq \frac{2^{2m+1-i}}{k}$$

where the last step uses the fact that $2^s \leq 2p$.

Since there are 2^i intervals in Φ_i, we find that

$$\mathbb{P}\mathrm{rob}[\exists I \in \Phi : \mathrm{disc}[aX + b, I) \geq 2^{-m}] <$$
$$< \sum_{i=1}^{m} 2^i \frac{2^{2m+1-i}}{k} = \frac{m2^{2m+1}}{k} \leq 1$$

as asserted in (3).

It follows that there are $a, b \in \mathbb{Z}_p$ for which (3) holds.

We claim that for these a and b, (which we now fix)

$$\mathrm{disc}(aX + b) \leq (m+2)2^{1-m} \ . \tag{4}$$

To see this it clearly suffices to verify that $\mathrm{disc}(aX+b, J) \leq (m+2)2^{-m}$ for every *initial* interval J, i.e., an interval of the form $\{0, 1, \ldots, \ell\}$. Now for each initial interval J there are two initial intervals J_1 and J_2 such that $J_1 \subset J \subset J_2$ and the right endpoints of J_1 and J_2 are $(t-1)2^{s-m}$ and $t2^{s-m}$, respectively, for some $0 \leq t < 2^m$. Since J_1 and J_2 are both disjoint unions of at most m intervals in Φ, it follows from (3) that

$$\mathrm{disc}(aX+b, J_\nu) < m2^{-m}$$

for $\nu = 1, 2$. This together with the fact that $\frac{1}{p}|J \backslash J_1|$ and $\frac{1}{p}|J_2 \backslash J|$ are at most $\frac{1}{p}2^{s-m} \leq 2^{1-m}$ each, implies that

$$\mathrm{disc}(aX+b, J) \leq (m+2)2^{-m}$$

establishing the claim (4). Since discrepancy is not changed by a cyclic shift,

$$\mathrm{disc}(aX) \leq (m+2)2^{1-m}$$

completing the proof. □

Remark: The last proposition can be obviously extended to the case of several sets, in the same way that proposition 2.2 generalizes proposition 2.1.

3. Dispersing Subsets of $\mathbb{Z}_p$ by Applying Polynomials

The pairwise independence of the random variables $ax_j + b$ employed in the proofs of propositions 2.1 and 2.4, has been used in [ABI] for derandomizing several parallel algorithms. Next, using similar ideas, we prove an extension of theorem 2.1 which shows that the maximal gap between consecutive elements of a set in $\mathbf{Z}_p$ can be further decreased if we consider its images under polynomials of higher degree. Somewhat surprisingly, there is no discrepancy analogue; this is explained at the end of the section.

THEOREM 3.1. *Let p be prime and let $X = \{x_1, \ldots, x_k\} \subset \mathbf{Z}_p$. If $d \geq 2$ is even, then there is a polynomial f of degree less than d over $\mathbf{Z}_p$, such that the image $f(X)$ intersects every interval of length*

$$c(d)pk^{-d/(d+2)} \quad \text{in} \quad \mathbf{Z}_p \ ,$$

(here $c(d) > 0$ depends only on d).

Proof: Let $a_0, \ldots, a_{d-1}$ be d random elements of $\mathbf{Z}_p$, chosen independently and uniformly in $\mathbf{Z}_p$. Construct the random polynomial $f(t) = \sum_{i=0}^{d-1} a_i t^i$ and note that for any d distinct elements $\{u_1, \ldots, u_d\}$ in $\mathbf{Z}_p$, the d random variables $f(u_1), \ldots, f(u_d)$ are independent (this is just the invertibility of a van-der-Monde matrix).

Take $\delta > 1/k$ to be specified later, and fix an interval I of length δp. For $1 \leq j \leq k$, define $W_{I,j} = 1 - \delta$ if $f(x_j) \in I$ and $W_{I,j} = -\delta$ otherwise and let $W_I = \sum_{j=1}^{k} W_{I,j}$. Clearly $E(W_I) = 0$; when computing the d'th moment of W_I using the multinomial expansion and d-wise independence, the dominant contribution comes from products of $d/2$ second moments, since $k\delta > 1$. Thus $E(W_I^d) \leq C_0(d)[k \operatorname{Var}(W_{I,1})]^{d/2} \leq C_0(d)(k\delta)^{d/2}$. Now it follows that

$$\mathbf{P}\text{rob}[W_I = -k\delta] \leq E(W_I^d)/(k\delta)^d \leq C_0(d)(k\delta)^{-d/2} .$$

Letting $\mathcal{F}$ be a collection of $2/\delta$ intervals of length δp whose union covers $\mathbf{Z}_p$, we get

$$\mathbf{P}\text{rob}[\exists I \in \mathcal{F} : W_I = -k\delta] \leq C_1(d) k^{-d/2} \delta^{-1-d/2} .$$

If the right-hand side is less than 1, i.e., if $\delta > C_2(d) k^{-d/(d+2)}$, then for some choice of the polynomial f, the set $f(X)$ intersects every interval $I \in \mathcal{F}$. This completes the proof. □

Remark: There is no discrepancy analogue of the last proposition , in the sense that the exponent $-1/2$ which occurs in theorem 2.3 cannot be improved upon by taking polynomial images $f(X)$ instead of dilations. More precisely, let $u_0, \ldots, u_{p-1}$ be independent random variables taking values $0, 1$ with equal probabilities and consider the random subset

$$X = \{i \mid 0 \leq i < p, \quad u_i = 1\} \quad \text{of} \quad \mathbf{Z}_p .$$

Then there exists $A > 0$ such that for every $d \geq 2$, with probability tending to 1 as $p \to \infty$;

(i) All polynomials f of degree less than d over $\mathbf{Z}_p$ satisfy

$$\operatorname{disc}(f(X)) > [Adp \log p]^{-1/2} .$$

and

(ii) $|X| > p/3$.

We sketch one way to check (i), since (ii) is clear. It is a well known fact that the probability that a simple random walk on the integers stays in an interval of size $L < \sqrt{p}$ for p consecutive steps decays like $\exp[-cp/L^2]$, for some constant $c > 0$, as $p \to \infty$. Inserting $L = \left[\frac{p}{Ad \log p}\right]^{1/2}$ here, yields a probability smaller than p^{-d} for large A. The relevance of this to discrepancy becomes apparent by considering initial intervals. The p^d random sets obtained by applying polynomials of degree less than d to X, are not all distributed like X, but sufficiently close so that the argument applies.

4. Uniform Dilations in $\mathbb{T}$: Proofs

In this section we prove theorems 1.1 and 1.2. Basically, we proceed as in section 2, but the pairwise independence used there is no longer available. Instead, we need to estimate covariances and it is here that proposition 1.3 is useful (that proposition is established in the next section).

Proof of Theorem 1.1: Given a set of k points $X = \{x_1, \ldots, x_k\}$ in $\mathbb{T}$, our objective is to obtain a sufficiently dense dilation of X. To this end, we choose a large integer A (whose size depends on the actual numbers x_i), pick an integer a, chosen randomly in the interval $\{1, 2, \ldots, A\}$ according to a uniform distribution, and pick a real b chosen uniformly in the interval $[0, 1)$.

Define $z_i = ax_i + b \pmod 1$ and $aX + b = \{z_1, \ldots, z_k\}$. The main step in the proof is the estimate of the expectation and variance of the number of elements of $aX + b$ that lie in a fixed interval. Fix an interval I of length $\varepsilon \leq \frac{1}{2}$ in $\mathbb{T}$. Let $Y_j = Y_j^I$ be the indicator random variable which is 1 if $z_j \in I$ and 0 otherwise. Define $Y = Y^I = \sum_{j=1}^k Y_j$. Obviously $E(Y) = k\varepsilon$, but the computation of the variance is more complicated. Moreover, the variance depends on the actual choice of A, and we shall concentrate on estimating its value when $A \to \infty$. Clearly

$$\mathrm{Var}(Y) = \sum_{j=1}^{k} \mathrm{Var}(Y_j) + 2 \sum_{1 \leq i < j \leq k} \mathrm{Cov}(Y_i, Y_j) \leq k\varepsilon + 2 \sum_{1 \leq i < j \leq k} \mathrm{Cov}(Y_i, Y_j) \tag{5}$$

where

$$\mathrm{Cov}(Y_i, Y_j) = E(Y_i Y_j) - E(Y_i)E(Y_j) = \mathbf{Prob}[z_i \in I \quad \text{and} \quad z_j \in I] - \varepsilon^2 \, .$$

The quantity $\mathbb{P}\text{rob}[z_i \in I \text{ and } z_j \in I]$ can be estimated in terms of the difference $x_i - x_j$ (and A). Indeed, by conditioning on the value of a we see that

$$\begin{aligned}\mathbb{P}\text{rob}[z_i \in I, z_j \in I] &= \frac{1}{A}\sum_{n=1}^{A} \mathbb{P}\text{rob}[z_i \in I,\ z_j \in I \mid a = n] \\ &= \frac{1}{A}\sum_{n=1}^{A} \psi_\varepsilon(nx_i - nx_j)\end{aligned} \tag{6}$$

where $\psi_\varepsilon(t) = \max\{\varepsilon - |t|, 0\}$ for $|t| \le 1/2$ and ψ_ε is continued with period 1 to $\mathbb{R}$. Now we consider three possible cases.

Case 1. $x_i - x_j$ is irrational. In this case Weyl's equidistribution lemma (cf. [KN]) and (6) give

$$\lim_{A\to\infty} \mathbb{P}\text{rob}[z_i \in I,\ z_j \in I] = \int_{-\varepsilon}^{\varepsilon} \psi_\varepsilon(t)dt = \varepsilon^2$$

so that $\lim_{A\to\infty} \text{Cov}(Y_i, Y_j) = 0$.

Case 2. $x_i - x_j = c/d$, where c/d is a reduced fraction and $1/d \ge \varepsilon$. In this case ψ_ε vanishes at all multiples of $x_i - x_j$ which are not integers, so that

$$\lim_{A\to\infty} \text{Cov}(Y_i, Y_j) = \frac{1}{d}\psi_\varepsilon(0) - \varepsilon^2 < \frac{\varepsilon}{d}\,.$$

Case 3. $x_i - x_j = c/d$, where c/d is a reduced fraction and $1/d < \varepsilon$. Here

$$\lim_{A\to\infty} E(Y_iY_j) = \frac{\varepsilon}{d} + \frac{2}{d}\sum_{n=1}^{\lfloor \varepsilon d\rfloor}\left(\varepsilon - \frac{n}{d}\right) = \frac{\varepsilon}{d} + \frac{1}{d}\left(2\varepsilon - \frac{\lfloor \varepsilon d\rfloor + 1}{d}\right)\lfloor \varepsilon d\rfloor\,.$$

Put $\lfloor \varepsilon d\rfloor = \varepsilon d - t$, where $0 \le t < 1$. Then

$$\begin{aligned}\lim_{A\to\infty} E(Y_iY_j) &= \frac{\varepsilon}{d} + \frac{1}{d}\left(2\varepsilon - \frac{\varepsilon d + 1 - t}{d}\right)(\varepsilon d - t) \\ &= \varepsilon^2 + \frac{t(1-t)}{d^2} \le \varepsilon^2 + \frac{1}{4d^2}\,.\end{aligned}$$

Hence, in this case

$$\lim_{A\to\infty} \text{Cov}(Y_i, Y_j) \le \frac{1}{4d^2}\,.$$

As in proposition 1.3, denote by h_d the number of pairs (i,j) where $1 \le i < j \le k$ such that $d(x_i - x_j)$ is an integer. Combining the results in cases 1,2 and 3 with equation (5) we obtain

$$\lim_{A\to\infty} \mathrm{Var}(Y) \le k\varepsilon + 2\sum_{d\le 1/\varepsilon} h_d \frac{\varepsilon}{d} + \frac{1}{2}\sum_{d>1/\varepsilon} \frac{h_d}{d^2}\ .$$

Define $H_d = \sum_{\nu=1}^{d} h_\nu$ (in particular, $H_1 = 0$) and let $d' = \lfloor 1/\varepsilon \rfloor$.

Using summation by parts we get

$$\begin{aligned}\lim_{A\to\infty} \mathrm{Var}(Y) &\le k\varepsilon + 2\sum_{d\le 1/\varepsilon} H_d\varepsilon\left(\frac{1}{d} - \frac{1}{d+1}\right) + 2H_{d'}\frac{\varepsilon}{d'} \\ &\quad + \frac{1}{2}\sum_{d>1/\varepsilon} H_d\left(\frac{1}{d^2} - \frac{1}{(d+1)^2}\right) \\ &\le k\varepsilon + 2\sum_{d\le\frac{1}{\varepsilon}} H_d\frac{\varepsilon}{d^2} + 2H_{d'}\frac{\varepsilon}{d'} + \sum_{d>1/\varepsilon} H_d\frac{1}{d^3}\ . \qquad (7)\end{aligned}$$

Substituting the estimates given in proposition 1.3 into the inequality (7) we obtain

COROLLARY 4.1. *For every $\alpha > 0$ there exists $k_0 = k_0(\alpha)$ such that if $k > k_0$ then the following holds. Given a set $X = \{x_1, \ldots, x_k\}$ of k elements in $\mathbf{T}$ and a large integer A, construct the random set $aX + b \pmod 1$ where a is a random integer in $\{1, 2, \ldots, A\}$ and b is a random real in $\mathbf{T}$, independent of a. Then for a fixed interval $I \subset \mathbf{T}$ of length ε, the random variable $Y = Y^I$ giving the cardinality of $(aX + b) \cap I$ satisfies $E(Y^I) = k\varepsilon$ and $\lim_{A\to\infty} \mathrm{Var}(Y^I) \le k^{1+\alpha}\varepsilon^{1-\alpha}$.*

To complete the proof of theorem 1.1, consider a set of $\lceil 1/\varepsilon \rceil$ intervals of measure ε each, whose union covers $\mathbf{T}$. By corollary 4.1 and Chebyshev's inequality, given $\alpha > 0$, if ε is sufficiently small, A is sufficiently large and

$$\lceil 1/\varepsilon \rceil \frac{k^{1+\alpha}\varepsilon^{1-\alpha}}{k^2\varepsilon^2} < 1$$

then with positive probability $aX + b$ (and hence also aX) intersects each interval of the family and hence each interval of length 2ε in $\mathbf{T}$.

The extra factor of 2 is unimportant, so this concludes the proof of the upper bound on $k(\varepsilon)$ in theorem 1.1. The lower bound was noted in [BP], where it was observed that if X is the Farey sequence

$$X_m = \{j/\ell \mid 1 \le j \le \ell \le m \ , \ gcd(j,\ell) = 1\}$$

then no dilation of X_m intersects the interval $\left(0, \frac{1}{m}\right)$ and $|X_m| \ge \Omega(m^2)$ (in fact, as is well known, $|X_m| = (1+o(1))\frac{3}{\pi^2}m^2$ as $m \to \infty$). □

Proof of Theorem 1.2: We argue as in the proof of proposition 2.4. Let $m \ge 1$. For each $i \in \{1,2,\ldots,m\}$, denote by $\mathcal{F}_i$ the family of 2^i disjoint intervals

$$\mathcal{F}_i = \{[(\nu-1)2^{-i},\ \nu 2^{-i}) : 1 \le \nu \le 2^i\}$$

and define $\mathcal{F} = \bigcup_{i=1}^m \mathcal{F}_i$. By combining Chebychev's inequality with corollary 4.1, we obtain the following. For every $\alpha > 0$, if k and A are sufficiently large and

$$\sum_{i=1}^{m} 2^i \frac{k^{1+\alpha}(2^{-i})^{1-\alpha}}{(k2^{-m})^2} < 1 \ , \tag{8}$$

then with positive probability

$$\forall I \in \mathcal{F} \qquad \mathrm{disc}(aX+b, I) \le 2^{-m} \ , \tag{9}$$

(here $a \in \{1,2,\ldots,A\}$ and $b \in \mathbb{T}$ are chosen randomly as in corollary 4.1). The reasoning used in the proof of Proposition 2.4 shows that if an integer a and an element b of $\mathbb{T}$ satisfy (9), then

$$\mathrm{disc}(aX) = \mathrm{disc}(aX+b) \le 2(m+1)2^{-m} \ . \tag{10}$$

Now the left-hand-side of (8) is at most

$$c(\alpha)k^{\alpha-1}2^{(2+\alpha)m} \ . \tag{11}$$

Therefore, by choosing α sufficiently small, and then using the largest m for which (11) is less than 1, (10) implies the assertion of the theorem. □

Remark: Theorem 1.2 can be improved in (at least) two directions. These improvements may be combined, but we describe them separately.

First, the theorem extends to simultaneous dilations of several sequences, in the spirit of proposition 2.2.

Second, by its very nature, the probabilistic method, when exhibiting a dilation with small discrepancy, necessarily exhibits many such dilations. To state this in a sharp form, recall that the upper Banach density $BD^*(S)$ of a set of integers S is defined by

$$BD^*(S) = \lim_{A\to\infty} \frac{1}{A} \sup_{L\in\mathbb{Z}} |S \cap \{L+1,\ L+2, \ldots, L+A\}| \ .$$

COROLLARY 4.2. *Let $\gamma > \delta > 0$. If k is sufficiently large, then for every sequence $X = \{x_1, \ldots, x_k\}$ of k distinct points in $\mathbb{T}$:*

$$BD^*\{n \in \mathbb{Z} \mid \mathrm{disc}(nX) > k^{\gamma-1/2}\} \leq k^{-\delta} \ . \tag{12}$$

Proof: The statement corresponding to (12) but involving ordinary upper density is essentially contained in the previous proof. Indeed, if $k^{\gamma-1/2} > 2(m+1)2^{-m}$ then

$$\frac{1}{A}\left|\{n : 1 \leq n \leq A,\ \mathrm{disc}(nX) \geq k^{\gamma-1/2}\}\right|$$

is bounded above by the quantity in (11) for large k. To obtain (12) with upper Banach density, repeat the argument leading to the proof of theorem 1.3, with the integer a chosen randomly in intervals of the form $[L+1,\ L+A]$ rather than $[1, A]$. □

5. Counting Differences with Small Denominators

Proposition 1.3, which we prove in this section, may be naturally viewed as an upper bound on the number of edges in a certain graph Γ. To obtain such a bound, we introduce an auxiliary parameter r and count the number of closed walks of length r in Γ. For this purpose we first require an estimate on the number of solutions of a certain Diophantine equation.

LEMMA 5.1. *Let $r, m > 1$ be integers. Denote by $\Lambda_r(m)$ the number of solutions of the equation*

$$\frac{a_1}{b_1} + \frac{a_2}{b_2} + \cdots + \frac{a_r}{b_r} = 0 \tag{13}$$

in which a_i, b_i are integers satisfying $0 < |a_i| < b_i \leq m$. Then

$$\Lambda_r(m) \leq (3m)^r (r \log m)^{2^{r+1}} \ .$$

Proof: For any two positive integers s, m denote by $Q_r(s,m)$ the number of integer solutions of (13) satisfying $|a_i| \leq s$ and $0 < b_i \leq m$. We first obtain an upper bound for $Q_r(s,m)$ assuming s is large and then derive the required bound for $\Lambda_r(m)$. Let $b_1, \ldots, b_r$ be r fixed positive integers which do not exceed m, and let $L = lcm(b_1, \ldots, b_r)$ be their least common multiple. Define

$$v = (L/b_1, \ldots, L/b_r)$$

and denote by $v^{\perp}$ the $(r-1)$-dimensional lattice consisting of all vectors in $\mathbb{Z}^r$ which are orthogonal to v. The number of solutions of (13) with these fixed denominators b_i and with integers $a_1, \ldots, a_r$ where $|a_i| \leq s$ is precisely the number of points in the lattice $v^{\perp}$ which lie inside the cube $[-s,s]^r$. By lemma 1 in [S] and the corollary that follows, the determinant of this lattice is equal to the Euclidean norm of v, i.e., to

$$\left(\sum_{i=1}^{r} L^2/b_i^2 \right)^{1/2} \geq L/m \ .$$

(Note that the greatest common divisor of the coordinates of v is 1, and hence the 1-dimensional lattice consisting of all integral multiples of v is primitive, as required in [S]).

For large s we can bound the number of points of $v^{\perp}$ inside the cube by estimating volumes. The $(r-1)$ volume of the intersection of the hyperplane spanned by $v^{\perp}$ with the cube $[-s,s]^r$ is

$$U \leq \sqrt{2}(2s)^{r-1} \ ,$$

by the result of Ball [Ba] (actually, any trivial estimate of the form $c(r)s^{r-1}$ suffices here). Therefore the number of points in $v^{\perp} \cap [-s,s]^r$ is at most

$$(1+o(1))Um/L \leq 2(2s)^{r-1}m/L$$

as $s \to \infty$. Summing over all possible choices of $1 \leq b_i \leq m$ we get

$$Q_r(s,m) \leq 2^r s^{r-1} m \sum_{1 \leq b_1, \ldots, b_r \leq m} \ell cm(b_1, \ldots, b_r)^{-1} \ . \tag{14}$$

Denoting by $d(L) = \sum_{h|L} 1$ the number of divisors of L, (14) implies that

$$Q_r(s,m) \leq 2^r s^{r-1} m \sum_{L=1}^{m^r} \frac{d(L)^r}{L} \ , \tag{15}$$

since every $L \in [1, m^r]$ is the l.c.m. of at most $d(L)^r$ vectors $(b_1, \ldots, b_r)$. An elementary inductive argument (given in detail in section 6.5 of [H]) shows that for $M > 4$,

$$\sum_{L=1}^{M} \frac{d(L)^r}{L} < (\log M)^{2^r} .$$

Substituting this into (15) gives that for all $m > 2$ and all $s > s_0(m)$,

$$Q_r(s, m) \leq 2^r s^{r-1} m (r \log m)^{2^r} . \tag{16}$$

provided $s_0(m)$ is sufficiently large. Next, we claim that for all s, m

$$\Lambda_r(m) Q_r(s, 1) \leq Q_r(3ms + m, m) . \tag{17}$$

To see this, map every pair of solutions: $\{(a_i, b_i)\}_{i=1}^r$ contributing to $\Lambda_r(m)$ and $\{(c_i, 1)\}_{i=1}^r$ contributing to $Q_r(s, 1)$, to the solution $\{(a_i + 3cb_i, b_i)\}_{i=1}^r$ of (13) in which

$$|a_i + 3c_i b_i| \leq m + 3ms .$$

This mapping is obviously one to one, implying (17). Finally, since the $(r-1)$-dimensional sublattice of $\mathbf{Z}^r$ consisting of vectors orthogonal to $(1, 1, \ldots, 1)$ has determinant $\sqrt{r}$, and any hyperplane through the center of the cube $[-s, s]^r$ intersects that cube in a set of $(r-1)$-volume of at least $(2s)^r$ (cf. [Ba]) we have

$$Q_r(s, 1) \geq \frac{1 + o(1)}{\sqrt{r}} (2s)^{r-1}$$

(actually, the elementary lower bound $Q_r(s, 1) \geq \left(\frac{s}{r}\right)^{r-1}$ would suffice). In conjunction with (16) and (17) this yields, for large s,

$$\Lambda_r(m) \leq \frac{2^r (3ms + m)^{r-1} m}{(2s)^{r-1} r^{-1/2}} (r \log m)^{2^r} \leq (3m)^r (r \log m)^{2^{r+1}}$$

as claimed. □

Proof of Proposition 1.3: Let $\{x_1, \ldots, x_k\}$ be an arbitrary set of k distinct numbers in $[0, 1)$. Construct a labeled graph Γ on the vertex set $\{x_1, \ldots, x_k\}$ as follows. For each i, j such that $1 \leq i < j \leq k$, put an edge labeled by (a, b) between x_i and x_j iff b is an integer, $1 \leq b \leq m$, and $b(x_i - x_j) = a$ is an integer. (Note that we allow multiple edges in Γ).

The number of edges in Γ is precisely the number H_m defined in the statement of the proposition. Define $D = \frac{1}{k} H_m$. Clearly the average degree in Γ is $2D$. Let Γ' be the graph obtained from Γ by repeatedly deleting from Γ vertices of degree less than D, if there are any such vertices. Since this process increases the average degree, it must terminate in a nonempty graph in which every degree is at least D. Let $r \geq 2$ be an even integer. Choose arbitrarily a fixed vertex w_0 of Γ' and consider walks of length $r/2$ in Γ' which start at w_0: A walk is determined by a sequence of (not necessarily distinct) adjacent *edges* in Γ'.

Any two such walks which end at the same vertex of Γ', may be combined to form a closed walk of length r in Γ', starting and ending at w_0. If the edges along this closed walk are labeled successively by

$$(a_1, b_1), (a_2, b_2), \ldots, (a_r, b_r)$$

then necessarily $\frac{a_1}{b_1} + \cdots + \frac{a_r}{b_r} = 0$. Thus we have a one to one mapping from closed walks of length r (starting at w_0) to solutions of (13) which satisfy $0 < |a_i| < b_i \leq m$ (these are the solutions counted by $\Lambda_r(m)$). For each $1 \leq i \leq k$, denote by T_i the number of walks of length $r/2$ in Γ' which start at w_0 and end at x_i (if $x_i \notin \Gamma'$ we set $T_i = 0$). Clearly $\sum_{i=1}^{k} T_i \geq D^{r/2}$. By the considerations above, the number of closed walks of length r in Γ' which start at w_0 is at least

$$\sum_{i=1}^{k} T_i^2 \geq \frac{1}{k} \left(\sum_{i=1}^{k} T_i \right)^2 \geq \frac{1}{k} D^r \ .$$

Consequently

$$\Lambda_r(m) \geq \frac{1}{k} D^r \ .$$

Using lemma 5.1, this implies

$$D \leq 3k^{1/r} m (r \log m)^{\frac{2r+1}{r}}$$

and therefore

$$H_m = kD \leq 3k^{1+1/r} m (r \log m)^{\frac{2r+1}{r}} \ . \tag{18}$$

Taking $r > 1/\alpha$ completes the proof. □

Remarks: (i) As observed by J. Bourgain [private communication] it is possible to prove proposition 1.3 using harmonic analysis, along the lines of [Bo].

(ii) By allowing r to vary with m, the estimate in proposition 1.3 may be improved. Specifically, for $m > e^{10}$ let r be the least even integer greater than $\log\log m$. Substituting this value of r into (18) yields

COROLLARY 5.2. *With the hypothesis and notation of proposition 1.3, the conclusion can be sharpened to*

$$H_m \leq 3(km)^{1+(\log\log m)^{-1}}$$

provided $m > e^{10}$.

Finally we note that it is possible to get (slightly) better estimates for the sum in (14), but this does not yield a noticeable improvement of the bound in the preceding Corollary.

6. The Harmonic Analysis Approach and Restricted Multipliers

We start by reproving Theorem 1.1 in the following slightly sharper form, and then establish some extensions.

PROPOSITION 6.1. *Using the notation of the introduction, the inequality*

$$k(\varepsilon) \leq \left(\frac{1}{\varepsilon}\right)^{2+\frac{3}{\log\log(1/\varepsilon)}}$$

holds, provided $\varepsilon > 0$ *is sufficiently small.*

We note that this may also be established by the probabilistic method of section 4; the key point is the application of corollary 5.2 rather than proposition 1.3. Throughout this section, we use the notation

$$e_m(t) = e^{2\pi imt} .$$

Let us start by recalling a classical fact concerning "bump" functions.

LEMMA 6.2. *There exist, for* $0 < \varepsilon < 1$, *nonnegative functions* $g_\varepsilon : \mathbb{R} \to \mathbb{R}$ *of period 1 such that*

(i) $g_\varepsilon(t) = 0$ *for* $\varepsilon \leq |t| \leq 1/2$.

(ii) The Fourier coefficients $\widehat{g}_\varepsilon(m) = \int_0^1 g_\varepsilon(t)e_m(t)dt$ *satisfy* $\widehat{g}_\varepsilon(0) = 1$ *and*

$$\forall m \in \mathbf{Z} \quad |\widehat{g}_\varepsilon(m)| \leq C \exp\left(-\sqrt{\varepsilon|m|}\right) ,$$

where C *is an absolute constant. (Actually, a slower rate of decay would suffice to establish theorem 1.1.)*

Proof: For the convenience of the reader, we reproduce the well-known argument. By the easy direction of the Denjoy-Carleman theorem (cf. [K], chapter V) there exists a nonnegative function $g : \mathbf{R} \to \mathbf{R}$ which vanishes off $(-\frac{1}{2}, \frac{1}{2})$ and satisfies $\int_{\mathbf{R}} g(t)dt = 1$ and

$$\left| \int_{\mathbf{R}} g(t)e^{2\pi ist}dt \right| \leq C \exp\left(-|s|^{1/2}\right)$$

for some $C > 0$ and all real s. Explicitly, one may take g to be the infinite convolution $\star_{\ell=10}^{\infty}\varphi_\ell$ where $\varphi_\ell(t) = \frac{\ell^{3/2}}{2}$ if $|t| < \ell^{-3/2}$ and $\varphi_\ell(t) = 0$ otherwise. Then for each $0 < \varepsilon < 1$, define $g_\varepsilon(t) = \frac{1}{\varepsilon}g(t/\varepsilon)$ for $|t| < 1/2$ and continue g_ε with period 1. Since $\widehat{g}_\varepsilon(m) = \frac{1}{\varepsilon}\int_{\mathbf{R}} g(t/\varepsilon)e^{2\pi imt}dt = \int_{\mathbf{R}} g(u)e^{2\pi im\varepsilon u}du$, we are done. Observe that for this construction,

$$\int_{-\varepsilon}^{\varepsilon} g_\varepsilon(t)^2 dt = \frac{1}{\varepsilon}\int_{\mathbf{R}} g(t)^2 dt . \tag{19}$$

□

Proof of Proposition 6.1: Let $\varepsilon > 0$, and suppose that $X = \{x_1, \ldots, x_k\}$ is a set of k points in $\mathbf{T}$ such that for every integer n, the dilation $nx(\mathrm{mod}\,1)$ is *not* ε-dense. We shall derive from this assumption an upper bound on k, implying the proposition. By our hypothesis, for each integer n there is some $\lambda_n \in \mathbf{T}$ such that nX and the interval $(\lambda_n - \varepsilon/2,\ \lambda_n + \varepsilon/2)$, both taken mod 1, are disjoint. Employing the function g_ε from the previous lemma, we have for every $N > 1$,

$$0 = \frac{1}{N}\sum_{n=1}^{N}\sum_{j=1}^{k} g_\varepsilon(nx_j + \lambda_n) = \frac{1}{N}\sum_{n=1}^{N}\sum_{j=1}^{k}\sum_{m=-\infty}^{\infty} \widehat{g}_\varepsilon(m)e_m(nx_j + \lambda_n) .$$

Since $\widehat{g}_\varepsilon(0) = 1$, separating this coefficient gives

$$k \leq \left| \frac{2}{N}\sum_{n=1}^{N}\sum_{j=1}^{k}\sum_{m=1}^{M} \widehat{g}_\varepsilon(m)e_m(nx_j + \lambda_n) \right| + 2k \sum_{m=M+1}^{\infty} \left|\widehat{g}_\varepsilon(m)\right| , \tag{20}$$

for any choice of $M > 1$.

The rightmost sum is easy to estimate:

$$\sum_{m=M+1}^{\infty} \left|\widehat{g}_\varepsilon(m)\right| \le C \sum_{m=M+1}^{\infty} \exp\left(-\sqrt{m\varepsilon}\right) \le \frac{C}{\varepsilon} \int_{M\varepsilon}^{\infty} e^{-\sqrt{t}} dt$$
$$= \frac{C}{\varepsilon} \int_{\sqrt{M\varepsilon}}^{\infty} 2ue^{-u} du = \frac{2C}{\varepsilon} \left(\sqrt{M\varepsilon} + 1\right) e^{-\sqrt{M\varepsilon}} .$$

Setting

$$M = \lfloor \frac{4}{\varepsilon} \log^2 \left(\frac{1}{\varepsilon}\right) \rfloor \tag{21}$$

we see that for $\varepsilon < \varepsilon_0$, certainly $\sum_{m=M+1}^{\infty} |\widehat{g}_\varepsilon(m)| < 1/4$. Substituting this into (20) gives

$$\frac{k}{4} \le \left| \frac{1}{N} \sum_{n=1}^{N} \sum_{m=1}^{M} \sum_{j=1}^{k} \widehat{g}_\varepsilon(m) e_m(nx_j + \lambda_n) \right| .$$

Now we square both sides and use Cauchy-Schwartz twice:

$$\frac{k^2}{16} \le \frac{1}{N} \sum_{n=1}^{N} \left| \sum_{m=1}^{M} \widehat{g}_\varepsilon(m) \sum_{j=1}^{k} e_m(nx_j + \lambda_n) \right|^2 \le$$
$$\le \frac{1}{N} \sum_{n=1}^{N} \left(\sum_{m=1}^{M} \left|\widehat{g}_\varepsilon(m)\right|^2 \right) \left(\sum_{m=1}^{M} \left| \sum_{j=1}^{k} e_m(nx_j + \lambda_n) \right|^2 \right) . \tag{22}$$

By Bessel's inequality and (19), there is an absolute constant c_1 such that

$$\sum_{m=1}^{M} \left|\widehat{g}_\varepsilon(m)\right|^2 \le \frac{c_1}{\varepsilon} .$$

Denoting $c_2 = 16c_1$, (22) implies that

$$k^2 \le \frac{c_2}{N\varepsilon} \sum_{n=1}^{N} \sum_{m=1}^{M} \left| \sum_{j=1}^{k} e_m(nx_j + \lambda_n) \right|^2 =$$
$$= \frac{c_2}{\varepsilon} \sum_{m=1}^{M} \sum_{j=1}^{k} \sum_{\ell=1}^{k} \frac{1}{N} \sum_{n=1}^{N} e_m(nx_j - nx_\ell) . \tag{23}$$

Next we let $N \to \infty$, and observe that for fixed m we have a nonzero contribution from a pair x_j, x_ℓ iff $m(x_j - x_\ell)$ is an integer. Recalling the definition of the quantities h_m and H_m (which depend on $x_1, \ldots, x_k$) from proposition 1.3, we get

$$k^2 \leq \frac{c_2}{\varepsilon} \sum_{m=1}^{M} [k + 2h_m] = \frac{c_2}{\varepsilon}(kM + 2H_M)$$

Using proposition 1.3, this immediately yields theorem 1.1. Utilizing corollary 5.2 instead, we conclude that for some $c_3 > 0$, if $\varepsilon < \varepsilon_0$ then our initial assumption that no dilation of X is ε-dense, forces the inequality

$$k^2 \leq \frac{c_3}{\varepsilon}(kM)^{1+(\log\log M)^{-1}}$$

to hold. Recalling our choice of M from (21), we see that necessarily

$$k < \left(\frac{1}{\varepsilon}\right)^{2+3(\log\log 1/\varepsilon)^{-1}}$$

provided $\varepsilon > 0$ is sufficiently small.

This completes the proof. □

To motivate our final theorem, we mention some results from [BP]. A set of integers S is called a *Glasner set* if for any infinite set $X \subset \mathbb{T}$ there exists an ε-dense dilation nX with $n \in S$. If there is a finite k such that the same holds for any $X \subset \mathbb{T}$ of cardinality k, then the smallest k with this property is denoted by $k_S(\varepsilon)$ (otherwise, we define $k_S(\varepsilon) = \infty$). In [BP] it was shown that any set $S \subset \mathbb{Z}$ of positive upper Banach density is a Glasner set and $k_S(\varepsilon) \leq O(\frac{1}{\varepsilon})^{O(\frac{1}{\varepsilon})}$.

In fact, corollary 4.2 from section 4 guarantees that if S has positive upper Banach density then for every $\beta > 0$ we have $k_S(\varepsilon) \leq O\left((\frac{1}{\varepsilon})^{2+\beta}\right)$. This is also easy to verify by the method of this section. Thinner sets defined arithmetically are more interesting. In [BP] it was proved that the image $f(\mathbb{Z})$ of a nonconstant polynomial with integer coefficients is a Glasner set, but the techniques used there did not yield any bound on the corresponding $k_S(\varepsilon)$. The method used to establish proposition 6.1 is particularly well-suited for obtaining such a bound and enables us to prove:

THEOREM 6.3.

(i) *For any $\delta > 0$, if $0 < \varepsilon < \varepsilon_1(\delta)$ then every set X in $\mathbb{T}$ of cardinality*

$$k > \frac{1}{\varepsilon^{2+\delta}} ,$$

has an ε-dense dilation pX with p prime.

(ii) *Let f be a polynomial of degree $L \geq 1$ with integer coefficients and let $\delta > 0$. If $0 < \varepsilon < \varepsilon_2(f, \delta)$, then any set X in $\mathbb{T}$ of cardinality*

$$k > \left(\frac{1}{\varepsilon}\right)^{2L+\delta} ,$$

has an ε-dense dilation of the form $f(n)X$, for some $n \in \mathbb{Z}$. [Actually, $\varepsilon_2(f, \delta)$ depends on f only through its degree L and the greatest common divisor of its coefficients].

Proof: (i) As in proposition 6.1., we assume that $X = \{x_1, \ldots, x_k\}$ is a set of cardinality k in $\mathbb{T}$ such that pX is *not* ε-dense for all primes p, and derive an upper bound on k. If $\{p_n\}$ is the sequence of primes in ascending order, this assumption implies that $0 = \frac{1}{N}\sum_{n=1}^{N}\sum_{j=1}^{k} g_\varepsilon(p_n x_j + \lambda_n)$ for suitable $\{\lambda_n\}$ and any $N \geq 1$. Proceeding exactly as in proposition 6.1., we arrive at the analogue of (23):

$$k^2 \leq \frac{c_2}{\varepsilon}\sum_{m=1}^{M}\sum_{j=1}^{k}\sum_{\ell=1}^{k}\frac{1}{N}\sum_{n=1}^{N} e_m(p_n(x_j - x_\ell)) \tag{24}$$

with the same value of M (given in (21)) and the same constant c_2. For any irrational number α, the sequence $\{p_n\alpha\}$ is equidistributed modulo 1 (cf. [V], chapter 11). Also, it is well known that for any two relatively prime integers a and b, the sequence $\{p_n a \bmod b\}_{n\geq 1}$ is asymptotically equidistributed among the $\varphi(b)$ residue classes which are relatively prime to b (cf. [D]), and hence

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\exp\left(2\pi i p_n\frac{a}{b}\right) = \frac{1}{\varphi(b)}\sum_{a'}\exp\left(2\pi\frac{a'}{b}\right) \tag{25}$$

where the right-hand sum is over all $1 \leq a' \leq b$ such that $gcd(a', b) = 1$. But the sum of all primitive roots of unity of order b (i.e., the sum of the roots

of the b-th cyclotomic polynomial) is precisely $\mu(b)$ where μ is the Möbius function, by the usual formula for the cyclotomic polynomials. Thus the right-hand side of (25) equals $\frac{\mu(b)}{\varphi(b)}$. If $x_j - x_\ell = \frac{a}{b}$ is a reduced fraction, then the denominator of $m(x_j - x_\ell)$ as a reduced fraction is $b' = \frac{b}{gcd(m,b)}$ and therefore

$$\left| \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e_m(p_n(x_j - x_\ell)) \right| = \left| \frac{\mu(b')}{\varphi(b')} \right| \leq \frac{O(\log\log b')}{b'}$$
$$= \frac{2gcd(m,b)}{b} O(\log\log b) \ .$$

Denoting by $d(b)$ the number of divisors of b, we have

$$O(\log\log b) \sum_{m=1}^{M} gcd(m,b) \leq O(\log\log b) \sum_{\substack{r|b \\ r\leq M}} \frac{M}{r} \cdot r$$
$$\leq O(\log\log b) M d(b) \leq C_\gamma M b^\gamma \ ,$$

where C_γ is a constant depending only on $\gamma > 0$ (cf. [H], chapter 6, theorem 5.2 for the rightmost inequality). Consequently for fixed x_j, x_ℓ as above,

$$\sum_{m=1}^{M} \left| \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e_m(p_n(x_j - x_\ell)) \right| \leq 2C_\gamma M b^{\gamma-1} \ . \tag{26}$$

Let $\widetilde{h}_b$ be the number of pairs (j,ℓ) with $1 \leq j < \ell \leq k$ such that $x_j - x_\ell$ as a *reduced* fraction has denominator b. Also, let $\widetilde{H}_b = \sum_{i=1}^{b} \widetilde{h}_i$. As $N \to \infty$, we can bound the right hand side of (24) by the sum over all $1 \leq j, \ell \leq k$ of the left-hand side in (26) and infer that for some constant C'_γ

$$k^2 \leq C'_\gamma \frac{M}{\varepsilon} \left[k + \sum_{b=2}^{\infty} \widetilde{h}_b b^{\gamma-1} \right] = C'_\gamma \frac{M}{\varepsilon} \left[k + \sum_{b=2}^{\infty} \widetilde{H}_b \left(b^{\gamma-1} - (b+1)^{\gamma-1} \right) \right] \tag{27}$$

(using summation by parts). Next, for $b \geq k$ we use the trivial inequality $\widetilde{H}_b \leq k^2$ and for $b < k$ the inequality $\widetilde{H}_b \leq H_b \leq (kb)^{1+\gamma}$ provided by

proposition 1.3 (assuming k is sufficiently large). Thus

$$\sum_{b=2}^{\infty} \widetilde{H}_b \left(b^{\gamma-1} - (b+1)^{\gamma-1}\right) \leq \sum_{b=2}^{k} H_b b^{\gamma-2} + k^2 k^{\gamma-1} \leq$$

$$\leq k^{1+\gamma} \left(\sum_{b=2}^{k} b^{2\gamma-1} + 1 \right) \leq \frac{1}{\gamma} k^{1+3\gamma} .$$

Inserting this into (27), we find that

$$k^2 \leq C''_\gamma \frac{M}{\varepsilon} k^{1+3\gamma} .$$

Recalling the value of M from (21), we see that our assumption that pX is never ε-dense forces the inequality

$$k^{1-3\gamma} < \left(\frac{1}{\varepsilon}\right)^{2+\gamma}$$

to hold, provided $\varepsilon > 0$ is sufficiently small.

This implies the assertion of the theorem.

(ii) We first suppose the coefficients of f have g.c.d. 1. Starting from the assumption that all dilations of $X = \{x_1, \ldots, x_k\}$ of the form $f(n)X$ are not ε-dense, one arrives as above at the inequality

$$k^2 \leq \frac{C_2}{\varepsilon} \sum_{m=1}^{M} \sum_{j=1}^{k} \sum_{\ell=1}^{k} \frac{1}{N} \sum_{n=1}^{N} e_m \left(f(n)(x_j - x_\ell)\right) . \tag{28}$$

If $x_j - x_\ell$ is irrational then

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e_m \left(f(n)(x_j - x_\ell)\right) = 0$$

by Weyl's equidistribution theorem (cf. [KN]). If $x_j - x_\ell$ is rational, $m \geq 1$ and

$$m(x_j - x_\ell) = \frac{a}{b} \tag{29}$$

is a reduced fraction we invoke Hua's estimate

$$\left| \sum_{n=0}^{b-1} \exp\left(2\pi i f(n) \frac{a}{b}\right) \right| \leq C_{\delta,L} b^{1-1/L+\delta} \tag{30}$$

valid for arbitrary $\delta > 0$ (cf. [H], chapter 7, theorem 10.1). Here the assumption on the coefficients of f is needed. Note that even though Weil's inequality gives a much better bound, $(L-1)b^{1/2}$ when b is prime, in general (30) is sharp except possibly for the δ-error term. This may be seen by considering the case in which $b = p^L$ is a prime power and $f(t) = t^L$. From (29),(30) we immediately infer

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e_m\left(f(n)(x_j - x_\ell)\right) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \exp\left(2\pi i f(n)\frac{a}{b}\right) \leq$$
$$\leq C_{\delta,L} b^{-1/L+\delta} .$$

Inserting this into (28) and repeating the argument in part (i) we get

$$k^2 \leq C'_{\delta,L} \frac{1}{\varepsilon^{2+\gamma}} k^{2-1/L+3\delta} ,$$

which implies the assertion of (ii) under our assumption on the coefficients of f. In the general case, write $f = qf_1$ where $q \in \mathbb{Z}$ and the coefficients of f_1 have no common divisor. If a set $X \subset \mathbb{T}$ of cardinality k has the property that $f(n)X$ is never ε-dense for any n, then the set $qX(\mathrm{mod}\,1)$, which has cardinality at least $\lceil k/q \rceil$, has the same property with respect to f_1. This completes the proof. □

It would be interesting to determine the best possible power of $1/\varepsilon$ required in Theorem 6.3, part (ii), and in particular to decide if it is bigger than $2 + \delta$ (for some fixed $\delta > 0$) for $f(t) = t^2$.

7. Concluding Remarks

We conclude with two applications of the foregoing.

COROLLARY 7.1. *Let a and b be relatively prime integers. The set of rationals in $[0, 1)$ with a terminating expansion in base a, such that in their base b expansion one of the digits $0, 1, 2, \ldots, b-1$ does not appear, is a finite set.*

Proof: We give the details assuming a is prime. The general case can be deduced from this via the Chinese Remainder Theorem. Consider the subgroup generated by b in the multiplicative group modulo a^m. The index of this subgroup is monotone nondecreasing in m, but is eventually constant

(this follows from Hensel's lemma, for instance). Denote by r the ultimate value of this index. Also, fix $0 < \varepsilon < \frac{1}{b}$. By Corollary 4.2 there exists k such that for any set X of at least k points in $\mathbf{T}$, the integers n such that $nX \bmod 1$ fails to be ε-dense constitute a set of upper density less than $\frac{1}{r}(1-\frac{1}{a})$.

Fix such a choice of k. Next we show that any rational x of the form $x = \frac{c}{a^m}$, where c is not divisible by a, has all digits $0, 1, 2, \ldots, b-1$ appearing in its base b representation, provided $\frac{1}{r}(a-1)a^{m-1} \geq k$. This will obviously complete the proof.

The left hand side of the last inequality is precisely the cardinality of the set

$$X = \{b^j x \bmod 1 : j \geq 1\}.$$

By our choice of k, there exists an ε-dense dilation nX with n congruent to some power of b modulo a^m. For such n, however, $nX \equiv X \bmod 1$. We conclude that for every $d \in \{0, 1, 2, \ldots, b-1\}$ there is some point of X in the interval $[\frac{d}{b}, \frac{d+1}{b})$, as claimed. □

Remark: It is possible to extract from the proof above bounds on the cardinality of the set of rationals whose finiteness is established there. Corollary 7.1 was motivated by the note [W] in which C. R. Wall shows that there are precisely 16 terminating decimals in the ternary Cantor set and exhibits them explicitly. We mention that establishing finiteness of the set of integers $m \geq 1$ for which some digit is missing in the decimal expansion of 5^{-m} is an open problem, first raised in [F].

Finally, let us show how similar ideas yield a very simple proof, communicated to us by H. Furstenberg, of the following result of K. Mahler.

COROLLARY 7.2. *[M] For every integer $b > 1$ and every irrational α, there exists an integer n such that in the base b expansion of $n\alpha$ each of the digits $0, 1, 2, \ldots, b-1$ occurs infinitely often.*

Proof: [H. Furstenberg] Denote by X the set of limit points of the sequence $\{b^j \alpha \bmod 1\}_{j \geq 1}$. A moments reflection shows that X is infinite. By Glasner's lemma there exists an ε-dense dilation nX for, say, $\varepsilon = \frac{1}{2b}$. For each $d \in \{0, 1, 2, \ldots, b-1\}$ there exists some $x \in X$ such that $\frac{d}{b} < nx < \frac{d+1}{b}$. This means that d appears infinitely often in the base b expansion of $n\alpha$. □

Remark: Mahler's original proof (which relies on the geometry of numbers) is longer but yields an explicit bound on the multiplier n which depends only on the base b and not on α. The proof presented here, besides its simplicity, also allows restricting the multiplier n to lie in a prescribed Glasner set (see section 6 or [BP] for the definition of these sets). Thus, for every irrational α, the set of integers n for which each possible digit appears infinitely often in the base b expansion of $n\alpha$, has density 1 and contains squares, primes, etc.

Acknowledgements.

Our foremost thanks go to Lennart Carleson for indicating the Harmonic Analysis approach and to Ze'ev Rudnick for several illuminating conversations. We are also grateful to Dani Berend, Daniel Kleitman and Peter Sarnak for their helpful suggestions.

References

[ABI] N. ALON, L. BABAI, A. ITAI, A fast and simple randomized parallel algorithm for the maximal independent set problem, J. Algorithms 7 (1986), 567-583.

[Ba] K. BALL, Cube Slicing in R^n, Proc. Amer. Math Soc. 97 (1986), 465-473.

[BP] D. BEREND, Y. PERES, Asymptotically dense dilations of sets on the circle, J. London Math. Soc., to appear.

[Bo] J. BOURGAIN, On $\Lambda(p)$-subsets of squares, Israel Journal of Math. 67 (1989), 291-311.

[Bu1] D.A. BURGESS, On character sums and primitive roots, Proc. London Math. Soc. 3,12 (1962), 179-192.

[Bu2] D.A. BURGESS, A note on the distribution of residues and non-residues, J. London Math. Soc. 38 (1963), 253-256.

[D] H. DAVENPORT, Multiplicative Number Theory, Second Edition, Springer Verlag, New York, 1980, chapters 22 and 23.

[F] H. FURSTENBERG, Intersections of Cantor sets and transversality of semigroups, in: Problems in Analysis, a symposium in honor of S. Bochner (R.C. Gunning, ed.), Princeton University Press (1970).

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

[GR] S.W. GRAHAM, C.J. RINGROSE, Lower bounds for least quadratic nonresidues, in: "Analytic Number Theory: Proceedings of a Conference in honor of P.T. Bateman", B.C. Berndt et. al. eds., Birkhauser, Boston, 1990.

[H] L.K. HUA, Introduction to Number Theory, Springer-Verlag, Berlin, 1982.

[K] Y. KATZNELSON, An Introduction to Harmonic Analysis, second edition, Dover, New York, 1976.

[KN] L. KUIPERS, H. NIEDRREITER, Uniform Distribution of Sequences, John Wiley and Sons, 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] W.M. SCHMIDT, Asymptotic formulae for point lattices of bounded determinant and subspaces of bounded height, Duke Math. J. 35 (1968), 327-339.

[V] I.M. VINOGRADOV, The Method of Trigonometrical sums in the Theory of Numbers, Interscience, London, 1954.

[W] C.R. WALL, Terminating decimals in the Cantor ternary set, The Fibonacci Quarterly, May 1990, 98-101.

N. Alon
School of Mathematical Sciences
Raymond and Beverly Sackler
Faculty of Exact Sciences
Tel Aviv University
Tel Aviv, Israel 69978

Y. Peres
Dept. of Mathematics
Stanford University
Present Address:
Mathematics Department
Yale University
New Haven, Connecticut 06520
USA

Submitted: October 1991