THE STRUCTURE OF POPULAR DIFFERENCE SETS

J. WOLF

Abstract. We show that the set of popular differences of a large subset of $\mathbb{Z}_N$ does not always contain the complete difference set of another large set. For this purpose we construct a so-called *niveau set*, which was first introduced by Ruzsa in [Ruz87] and later used in [Ruz91] to show that there exists a large subset of $\mathbb{Z}_N$ whose sumset does not contain any long arithmetic progressions. In this paper we make substantial use of measure concentration results on the multi-dimensional torus and Esseen's Inequality.

1. Introduction

Let G be a finite Abelian group of order N. Suppose that A is a subset of G of cardinality linear in N, and define the set of γ-*popular* differences of A to be

$$D_\gamma(A) := \{x \in G : A * -A(x) \geq \gamma\},$$

where we have written A for the indicator function of the subset A, and the convolution $f * g$ of two functions $f, g : G \to \mathbb{C}$ is defined as $f * g(x) = \mathbb{E}_{y \in G} f(y)\overline{g(x-y)}$. In other words, $D_\gamma(A)$ is the set of elements of G which can be written as a difference of elements of A in at least γN different ways. Because we are considering subsets of G of size linear in N, we shall take γ to be a small constant throughout the paper. Is it true that $D_\gamma(A)$ always contains the complete difference set $A_0 - A_0$ for some large set A_0? Our aim in this paper is to show that this is not always so. More precisely, when $G = \mathbb{F}_2^n$ and $G = \mathbb{Z}_N$ with N a prime, we prove that there exists a set A of linear size such that any set A_0 whose difference set is contained in $D_\gamma(A)$ has density $o(1)$. Here $o(1)$ denotes a quantity tending to 0 as the order N of the group G tends to infinity.

Theorem 1.1. *Let $G = \mathbb{F}_2^n$ or $G = \mathbb{Z}_N$. Then there exists a set $A \subsetneq G$ of size greater than $N/3$ with the property that any set A_0 whose difference set is contained in the set $D_\gamma(A)$ of γ-popular differences of A has density $o(1)$.*

Apart from being an interesting question in its own right, this problem has arisen in the context of counting the number of sum-free subsets of an Abelian group G, notably in

Date: July 2008.

the work of Lev, Łuczak and Schoen [LŁS01] and Green and Ruzsa [GR05]. The first team of authors pursued the following strategy: Suppose every sum-free set A contained a small subset E with large difference set. The small cardinality of E implies that there are relatively few such sets, and from the fact that the difference set is large it follows that there are only few sets A corresponding to a given E, since for a sum-free set A we have $A \subseteq G\backslash(A-A) \subseteq G\backslash(E-E)$. By taking a random subset of A with suitable probability, one can obtain a small set E which has the property that its difference set contains the set $D_\gamma(A)$ of popular differences of A. Therefore the argument we just sketched implies an upper bound on the number of sum-free sets A whenever $D_\gamma(A)$ is large. For those A with few popular differences, the following proposition from [LŁS01] can be used in conjunction with Kneser's Theorem to obtain an upper bound in the remaining case. Its proof consists of a simple averaging argument on the Cayley graph on $\mathbb{Z}_N$ generated by $D_\gamma(A)$.

Proposition 1.2. *Let X be a subset of G, and let γ be a positive constant. Suppose that the set of γ-popular differences $D_\gamma(X)$ satisfies*

$$|D_\gamma(X)| \leq 2|X| - 5\sqrt{\gamma N|X-X|}.$$

Then there exists a subset $X' \subseteq X$ such that

$$|X\backslash X'| \leq \sqrt{\gamma N|X-X|} \quad \textit{and} \quad X'-X' \subseteq D_\gamma(X).$$

Green and Ruzsa [GR05] used this proposition to show that it suffices to remove ϵN elements from a set of size greater than $(1/3+\epsilon)N$ with few (more precisely, up to $\epsilon^3N^2/27$) Schur triples in order to make it sum-free, which allows them to strengthen the result of Lev, Łuczak and Schoen on the number of sum-free subsets of G.

The result we present in this paper shows that the condition on the size of the set of popular differences in Proposition 1.2 cannot be removed, which by the preceding discussion rules out simpler approaches to counting sum-free sets of Abelian groups. Before dealing with the case of the group $G = \mathbb{Z}_N$ with N a prime in Section 3, we first describe a combinatorial approach in the model setting of $G = \mathbb{F}_2^n$.

2. Vector Spaces over Finite Fields

The case where G is a finite-dimensional vector space over the field of two elements is often a good model for what happens in the cyclic groups $\mathbb{Z}_N$, and generally easier to deal with as we have additional geometric structure available. We refer the reader to the excellent survey [Gre05] for a plentiful supply of examples confirming this assertion.

For $x \in \mathbb{F}_2^n$, let $|x|$ denote the number of non-zero coordinates of the vector x. In this section we shall show that in the model setting $\mathbb{F}_2^n$, the set $A \subseteq \mathbb{F}_2^n$ defined by

$$A := \left\{ x \in \mathbb{F}_2^n : |x| \geq \frac{n}{2} + \frac{\sqrt{3n}}{2} \right\}$$

is an example of a set whose popular difference set does not contain the complete difference set of any other large set.

The set A described above is the finite field analogue of a so-called *niveau set*, which was originally introduced by Ruzsa in [Ruz87] and later used in [Ruz91] to show that there exists a subset of $\mathbb{Z}_N$ whose sumset does not contain any long arithmetic progressions. It is a versatile construction that has received a fair amount of attention since. For example, a modified version of such a set can be used to show that Chang's Theorem on the structure of the large Fourier spectrum of a function is tight [Gre03]. We shall discuss the original construction in more detail in Section 3.

First we need to show that the set A thus constructed has the required size, that is, that it contains a positive proportion of all elements of $\mathbb{F}_2^n$. For the remainder of this section, we write $N := 2^n$ for the size of the group.

Lemma 2.1. *The set $A \subseteq \mathbb{F}_2^n$ as defined above has size at least* $(1 - \exp(-1/2))N$.

Proof. By definition, the size of A can be written as

$$|A| = \sum_{j=0}^{\frac{n}{2}+\frac{\sqrt{3n}}{2}} \binom{n}{j},$$

which equals the probability that a random variable X with binomial distribution $B(n, 1/2)$ takes values at most $\sqrt{3n}/2$ above its mean. We use a standard Chernov-type tail estimate, details of which can be found in [JŁR00] or Appendix A of [AS00].

Lemma 2.2. *Suppose X is a random variable with binomial distribution. Then for any $0 \leq \epsilon \leq 1$, we have the estimates*

$$\mathbb{P}(X \leq (1-\epsilon)\mathbb{E}X) \leq \exp(-\epsilon^2 \mathbb{E}X/2)$$

and

$$\mathbb{P}(X \geq (1+\epsilon)\mathbb{E}X) \leq \exp(-\epsilon^2 \mathbb{E}X/3).$$

It follows immediately from the second inequality that the density of A is at least $1 - \exp(-1/2)$, which means that A contains more than a third of all elements of $\mathbb{F}_2^n$. □

Next we show that the set of popular differences $D_\gamma(A)$ is contained in the complement of a *Hamming ball* centred at 1, which is defined as

$$B_t(1) := \{x \in \mathbb{F}_2^n : |x| \geq n-t\}.$$

Note that our finite field niveau set A is in fact itself a Hamming ball of radius $n/2-\sqrt{3n}/2$.

Lemma 2.3. *Let the set $A \subseteq \mathbb{F}_2^n$ and the Hamming ball $B_t(1)$ be defined as above. Then for any real $t \leq 3n/(4\log(\gamma^{-1}))$, we have*

$$D_\gamma(A) \subseteq B_t(1)^C.$$

Proof. We shall show that if $z \in \mathbb{F}_2^n$ is such that $|z| = n-t$, then the number of ways of writing z as a difference (or, equivalently, as a sum since we are performing addition modulo 2) of two elements of A is bounded above by $N\exp(-3n/4t)$. So suppose that z is the sum of two vectors x and y which both lie in A. Without loss of generality, we can assume that the first t coordinates of z are 0s, and the remaining $n-t$ coordinates are 1s. Writing

$$(z_1, z_2, ...z_t, z_{t+1}, ..., z_n) \equiv (x_1, x_2, ...x_t, x_{t+1}, ..., x_n) + (y_1, y_2, ...y_t, y_{t+1}, ..., y_n),$$

we observe (again without loss of generality) that the number of 1s amongst the coordinates $x_{t+1}, ..., x_n$ is bounded above by $(n-t)/2$. But we require that x be an element of A, so that the number of 1s amongst $x_1, ..., x_t$ is at least $n/2+\sqrt{3n}/2-(n-t)/2 = t/2+\sqrt{3n}/2$. Hence the number of possible vectors x, which for fixed z in turn immediately determine y, is bounded above by

$$2\sum_{i=\frac{t}{2}+\frac{\sqrt{3n}}{2}}^{t}\binom{t}{i}\sum_{j=0}^{\frac{1}{2}(n-t)}\binom{n-t}{j}.$$

The first sum can be bounded above by $2^t\exp(-(\sqrt{3n}/t)^2 t/4) = 2^t\exp(-3n/4t)$ by the first inequality of Lemma 2.2, and the second sum clearly equals 2^{n-t-1} by the binomial theorem. The result follows. □

Finally, we need to exploit the geometric information we have just gathered. It is not unreasonable to expect to be able to bound the size of *any* set whose difference set is contained in the complement of a large Hamming ball. A result to this effect was already proved by Kleitman [Kle66] (see also page 202 of [AS00]).

Here we shall use a simple instance of *measure concentration* on the discrete cube, which turns out to adapt well to a more general context. We include the proof for the sake of completeness and in order to motivate our work in Section 3.3. More background on the concentration of measure phenomenon in general compact metric groups will be presented in the introduction to that section.

Lemma 2.4. *Let A_0 be any subset of $\mathbb{F}_2^n$ with the property that $A_0 - A_0 \subseteq B_t(1)^C$. Then the density of A_0 is bounded above by* $\exp(-t^2/4n)$*.*

Proof. For ease of notation let us also define the Hamming ball centred at 0 in the obvious way by setting

$$B_t(0) := \{x \in \mathbb{F}_2^n : |x| < t\}.$$

This is just the usual ball associated with the so-called *Hamming metric* on $\mathbb{F}_2^n$ defined by setting $d(x, y) = |x - y|$. In other words, the distance between x and y equals the number of coordinates in which they differ. It is easy to see that

$$A_0 - A_0 \subseteq B_t(1)^C \quad \Rightarrow \quad A_0 + B_t(1) \cap A_0 = \emptyset,$$

which in turn implies that

$$\overline{A_0} + B_t(0) \cap A_0 = \emptyset,$$

where we have used the bar to denote the set $(1, 1, \ldots, 1) + A_0$ of *antipodal vectors* of A_0. But the set $\overline{A_0} + B_t(0)$ is just the set of elements of $\mathbb{F}_2^n$ at Hamming distance less than t from some element in $\overline{A_0}$. It is this observation which inspires us to use the following classical measure concentration result in the discrete cube, which can be found on page 172 of [McD89] or page 31 of [Led01].

Theorem 2.5. *Let μ denote the uniform measure on $\mathbb{F}_2^n$. Given any subset C of $\mathbb{F}_2^n$, we have the inequality*

$$\mu(C + B_t(0)) \geq 1 - \frac{\exp(-t^2/2n)}{\mu(C)}.$$

It was shown by Harper [Har66] that this inequality is sharp if the set C is a Hamming ball.

Proof of Lemma 2.4, continued: Applying Theorem 2.5 to the set $\overline{A_0}$, we immediately deduce that

$$\mu(\overline{A_0} + B_t(0)) \geq 1 - \frac{\exp(-t^2/2n)}{\mu(A_0)},$$

but the fact that $\overline{A_0} + B_t(0) \cap A_0 = \emptyset$ implies that

$$1 - \frac{\exp(-t^2/2n)}{\mu(A_0)} + \mu(A_0) \leq 1,$$

which after rearranging concludes the proof. □

Combining Lemma 2.3 and Lemma 2.4, we have proved the main result of this section. It asserts that $D_\gamma(A)$ only contains the complete difference set of sets of density $o(1)$.

Theorem 2.6. *There exists a set $A \subseteq \mathbb{F}_2^n$ of size greater than $N/3$ with the property that the set $D_\gamma(A)$ of γ-popular differences does not contain the complete difference set of any set of density greater than*

$$\exp\left(-9n/64 \log^2(\gamma^{-1})\right).$$

3. From the Model Case to $\mathbb{Z}_N$

We now focus our attention on the finite Abelian group $\mathbb{Z}_N$ with N a large prime, whose characters are of the form $x \mapsto e(rx/N) := \exp\left(2\pi i r x/N\right)$. In this more general context, we define a *niveau set* $A \subseteq \mathbb{Z}_N$ as the set

$$A := \left\{ x \in \mathbb{Z}_N : \Re \sum_{i=1}^{k} \gamma_i(x) \geq \epsilon\sqrt{k} \right\},$$

for some judiciously chosen set of characters $\gamma_1, \gamma_2, ..., \gamma_k$. The precise value of the parameters ϵ and k will be determined in the course of the argument, but ϵ should always be thought of as a fixed constant and k as growing roughly like $\log N$ to some small power. As already mentioned in Section 2, this construction was originally introduced by Ruzsa in [Ruz87] and later used in [Ruz91] to give an example of a subset of $\mathbb{Z}_N$ whose sumset does not contain any long arithmetic progressions. We shall follow his analysis of the properties of such a set very closely in Section 3.1, where we show that A contains a positive proportion of all elements of $\mathbb{Z}_N$. In order to be able to give an estimate for the size of A, we need the characters to behave roughly "independently" in the following sense:

Definition. *We say that a set of characters $(\gamma_i(x) = e(r_i x/N))_{i=1}^k$ is* K-independent *if $\sum_{i=1}^k \lambda_i r_i \equiv 0 \mod N$ has no solutions satisfying $\sum_i |\lambda_i| \leq K$. We shall also sometimes refer to the corresponding k-tuple $(r_i)_{i=1}^k \subseteq \mathbb{Z}_N^k$ as K-independent.*

We first of all need to make sure that such a set of characters actually exists, otherwise Definition 3 would be rather pointless.

Lemma 3.1. *The number of k-tuples in $\mathbb{Z}_N$ which are* not *K-independent is bounded above by*

$$(2K + 1)^k N^{k-1}.$$

In other words, there exists a set of k characters with the K-independence property provided that K satisfies the inequality $K < N^{1/k}/4$.

Proof. A very crude but effective counting argument will do the job: Every k-tuple which is not K-independent satisfies by definition an equation in k variables with coefficients between $-K$ and K. There are at most $(2K+1)^k$ such equations. □

From now on we assume that we are dealing with a set of K-independent characters whenever we make reference to the niveau set A. Having set up the basics, we now turn to proving the analogues of Lemmas 2.1, 2.3 and 2.4 in Sections 3.1, 3.2 and 3.3, respectively.

3.1. Estimating the Size of the Niveau Set.

The following lower bound on the cardinality of the niveau set A is proved in [Ruz91]. It is the analogue of Lemma 2.1 in the case $G = \mathbb{Z}_N$.

Proposition 3.2. *Let $\epsilon = 1/4$ and suppose $k \ll \log N/\log\log N$. Then the set A with parameters ϵ and k as defined above has cardinality at least $N/3$.*

For the sake of clarity, self-containedness and because we want to use a very similar argument later on, we give a concise exposition of Ruzsa's proof in this section. We shall proceed in two steps. First, we compare the character sum appearing in the definition of A to a sum of independent random variables distributed uniformly on the unit circle. Second, we approximate this sum of independent random variables by a normal distribution, which allows us to perform explicit computations.

A crucial tool in proving the first step is the following theorem in probability theory, which is known as *Esseen's Inequality*. It dates back to Esseen [Ess45] and independently Berry [Ber41], but see Shiryayev [Shi84] for a general introductory reference.

Theorem 3.3. *Let F_1, F_2 be probability distribution functions with corresponding characteristic functions ϕ_1, ϕ_2. Assume F_1' exists and is pointwise bounded by a constant V. Then*

$$\sup_x |F_1(x) - F_2(x)| \ll \frac{V}{T} + \int_0^T \frac{|\phi_1(t) - \phi_2(t)|}{t} dt.$$

We briefly recall that the *characteristic function* ϕ_X of a random variable X is defined to be $\phi_X(t) := \mathbb{E}\exp(itX)$, and that therefore the probability density function of a random variable is the inverse Fourier transform of its characteristic function. From now on we shall be using the notation $a \ll b$ to indicate that there exists an absolute constant c such that $a \le cb$.

A special case of Theorem 3.3, also known as the *Berry-Esseen Inequality*, will help us complete the second step. It measures the total variation distance between a sum of independent identically distributed random variables and the normal distribution, in other words, it gives us information about the rate of convergence in the Central Limit Theorem. More precisely, let $X_1, X_2, \ldots, X_k$ be independent random variables, each distributed uniformly on the unit circle, and define their sum to be

$$X := \sum_{j=1}^{k} X_j \text{ with real part } \widetilde{X} := \Re X.$$

Let $\sigma := \sqrt{k/2}$ denote the standard deviation of $\widetilde{X}$. The following formulation of the Berry-Esseen Inequality is taken from page 374 of [Shi84].

Theorem 3.4. *Let $\widetilde{X}$ be defined as above, and let Φ denote the standard normal distribution function. Then*

$$\sup_x |F_{\widetilde{X}/\sigma}(x) - \Phi(x)| \ll \frac{\mathbb{E}|\widetilde{X}|^3}{\sigma^4},$$

provided that the third absolute moment $\mathbb{E}|\widetilde{X}|^3$ is finite.

In order to estimate the difference between two characteristic functions effectively using Theorem 3.3, we need to consider the moments of the corresponding random variables. Given a random variable $\widetilde{X}$ as defined above, we can express its l^{th} moment $\widetilde{\mu}_l := \mathbb{E}\widetilde{X}^l$ as

$$\widetilde{\mu}_l = \frac{1}{2^l} \sum_{i=0}^{l} \binom{l}{i} \widetilde{\mu}_{i,l-i} \quad \text{by writing} \quad \widetilde{\mu}_{i,j} := \mathbb{E}X^i \overline{X}^j.$$

We set up analogous expressions for the character sum defining A by writing

$$f(x) := \sum_{j=1}^{k} \gamma_j(x) \text{ with real part } \widetilde{f}(x) := \Re f(x) \text{ and } l^{th} \text{ moment } \widetilde{\nu}_l := \frac{1}{N} \sum_{x=1}^{N} \widetilde{f}(x)^l.$$

The l^{th} moment of $\widetilde{f}$ can likewise be expanded as

$$\widetilde{\nu}_l = \frac{1}{2^l} \sum_{i=0}^{l} \binom{l}{i} \widetilde{\nu}_{i,l-i} \quad \text{upon setting} \quad \widetilde{\nu}_{i,j} := \frac{1}{N} \sum_{x=1}^{N} f(x)^i \overline{f(x)^j}.$$

Let $F_{\widetilde{X}}$, $F_{\widetilde{f}}$ denote the obvious distribution functions, and write $\phi_{\widetilde{X}}$, $\phi_{\widetilde{f}}$ for the corresponding characteristic functions.

We are interested in the distribution of $\widetilde{f}$. More precisely, in order to estimate the size of A we want to count the number of elements $x \in \mathbb{Z}_N$ such that $\widetilde{f}(x) \geq \epsilon\sqrt{k}$. This means that $1 - F_{\widetilde{f}}(\epsilon\sqrt{k})$ is the quantity we are ultimately interested in.
Our first lemma shows that K-independence guarantees that the lower moments of $\widetilde{f}$ and $\widetilde{X}$ are equal.

Lemma 3.5. *With the moments $\widetilde{\mu}_l$ and $\widetilde{\nu}_l$ defined as above and the characters $\gamma_1, \gamma_2, \dots, \gamma_k$ assumed to be K-independent, we have $\widetilde{\nu}_l = \widetilde{\mu}_l$ for all $l = 1, 2, \dots, K$.*

Proof. Under the assumption of K-independence, it is not too difficult to compute the mixed moments explicitly. Indeed, we can rewrite $\widetilde{\nu}_{i,j}$ as

$$\frac{1}{N}\sum_{x=1}^{N}\left(\sum_{m=1}^{k}\gamma_m(x)\right)^i\left(\sum_{n=1}^{k}\overline{\gamma_n(x)}\right)^j = \frac{1}{N}\sum_{\substack{m_1,\dots,m_i\\ n_1,\dots,n_j}}\sum_{x=1}^{N} e((r_{m_1}+\dots+r_{m_i}-r_{n_1}-\dots-r_{n_j})x/N).$$

Whenever $i + j \leq K$, the latter sum equals zero by K-independence unless $m_1, \dots, m_i$ is a permutation of $n_1, \dots, n_j$, in which case it equals N. We compare this with

$$\widetilde{\mu}_{i,j} = \mathbb{E}\left(\sum_{m=1}^{k} X_m\right)^i\left(\sum_{n=1}^{k}\overline{X_n}\right)^j = \sum_{m_1,\dots,m_i=1}^{k}\sum_{n_1,\dots,n_j=1}^{k}\mathbb{E}X_{m_1}\dots X_{m_i}\overline{X}_{n_1}\dots\overline{X}_{n_j}.$$

Again, since X_i is independent of X_j for $i \neq j$, the expectation is non-zero only when $m_1, \dots, m_i$ is a permutation of $n_1, \dots, n_j$, in which case it equals 1. Hence $\widetilde{\nu}_{i,j} = \widetilde{\mu}_{i,j}$ for all $i + j \leq K$, and the result follows as stated. □

In order to usefully estimate the difference between the two characteristic functions we also need to infer a decent bound on the l^{th} moment $\widetilde{\mu}_l$.

Lemma 3.6. *For any even integer $l \leq K$ and $\widetilde{\mu}_l$ defined as above, we have the upper bound*

$$\widetilde{\mu}_l \leq \min\left\{k^l, \frac{l!}{2^l(l/2)!}k^{l/2}\right\}.$$

Proof. The first part of the bound is obvious, and the second follows from the fact that the only non-zero mixed moments $\widetilde{\mu}_{i,l-i}$ are those for which $i = l/2$, when they are of magnitude $k^{l/2}(l/2)!$. □

We are now ready to carry out the first step of the argument, namely showing that $\widetilde{f}$ and $\widetilde{X}$ are close in distribution using Theorem 3.3.

Proposition 3.7. *Under the same assumptions as before, $\widetilde{f}$ and $\widetilde{X}$ are close in distribution in the sense that*

$$\sup_x |F_{\widetilde{X}}(x) - F_{\widetilde{f}}(x)| \ll \min\left\{\frac{1}{\sqrt{K}}, \frac{\sqrt{k}}{K}\right\}.$$

Proof. In order to apply Esseen's Inequality, we first need to verify that $F'_{\widetilde{X}}$ exists and is bounded above by a suitable constant. As we have already mentioned, it is a well-known fact in probability theory that the density function of a random variable is the inverse Fourier transform of its characteristic function, hence

$$F'_{\widetilde{X}}(x) \leq \int_{-\infty}^{\infty} |\phi_{\widetilde{X}}(t)| dt.$$

We thus require the following bounds on the characteristic function $\phi_{\widetilde{X}}$ of $\widetilde{X}$, which we state here without proof. The interested reader is referred to [Ruz91] for details.

Lemma 3.8. *There exist constants $a, b > 0$ and $T_0 > 1$ such that $\phi_{\widetilde{X}}$ satisfies*

$$|\phi_{\widetilde{X}}(t)| \leq \begin{cases} \exp\left(-akt^2\right) & |t| \leq T_0\sigma \\ (b|t|)^{-k/2} & |t| > T_0\sigma \end{cases}.$$

It is immediate to deduce that $F'_{\widetilde{X}}(x)$ is bounded above by a constant times the standard deviation σ. Next we observe that by Taylor's Theorem with remainder we can write

$$\phi_{\widetilde{X}}(t) = \sum_{j=1}^{l-1} \frac{\widetilde{\mu}_j}{j!}(it)^j + \delta\widetilde{\mu}_l \frac{|t|^l}{l!},$$

and similarly

$$\phi_{\widetilde{f}}(t) = \sum_{j=1}^{l-1} \frac{\widetilde{\nu}_j}{j!}(it)^j + \delta\widetilde{\nu}_l \frac{|t|^l}{l!}$$

for some $|\delta| \leq 1$. With the benefit of hindsight, this allows us to justify why we were so keen to compare moments in the first place. K-independence gave us through Lemma 3.5 that all moments $\widetilde{\mu}_j$ and $\widetilde{\nu}_j$ up to order K were equal, and thus

$$|\phi_{\widetilde{X}}(t) - \phi_{\widetilde{f}}(t)| \leq 2\widetilde{\mu}_K \frac{|t|^K}{K!}.$$

It now follows from Theorem 3.3 that for any $T > 1$,

$$\sup_x |F_{\widetilde{X}}(x) - F_{\widetilde{f}}(x)| \ll \frac{\sigma}{T} + \widetilde{\mu}_K \frac{T^K}{K!K}.$$

Using the bound on $\widetilde{\mu}_K$ derived in Lemma 3.6 and setting $T = \sigma(K!/\widetilde{\mu}_K)^{1/(K+1)}$ followed by a short compuation concludes the proof of Proposition 3.7. □

We have thus successfully approximated $\widetilde{f}$ by $\widetilde{X}$. It remains to compare a suitably normalized version of $\widetilde{X}$ to a standard normal random variable. The following proposition states that $\widetilde{X}$ is close to a normal distribution with mean 0 and standard deviation σ.

Proposition 3.9. *Let $\widetilde{X}$ be defined as above, and let Φ denote the standard normal distribution function. Then*

$$\sup_x |F_{\widetilde{X}/\sigma}(x) - \Phi(x)| \ll \frac{1}{\sigma}.$$

Proof. This is a straightforward application of Theorem 3.4. The third absolute moment $\mathbb{E}|\widetilde{X}|^3$ can be bounded by the Cauchy-Schwarz Inequality as

$$\mathbb{E}|\widetilde{X}|^3 \leq (\mathbb{E}|\widetilde{X}|^2)^{1/2}(\mathbb{E}|\widetilde{X}|^4)^{1/2}.$$

Splitting X_j into real and imaginary parts $X_j = R_j + iI_j$, we first observe that $\mathbb{E}R_j^2 = 1/2$ and $\mathbb{E}I_j^2 = 1/2$ as well as $\mathbb{E}R_j^4 = 3/8$. It is not hard to see that X_i and X_j are independent if and only if the pairs (R_i, I_i) and (R_j, I_j) are independent (but see page 273 of [Shi84] for a justification of this claim), which yields

$$\mathbb{E}\widetilde{X}^2 = \mathbb{E}\sum_{j,l=1}^{k} R_j R_l = \sum_{j=1}^{k} \mathbb{E}R_j^2 + \sum_{j \neq l=1}^{k} \mathbb{E}R_j R_l = \frac{k}{2}$$

and

$$\mathbb{E}\widetilde{X}^4 = \mathbb{E}\sum_{j=1}^{k} R_j^4 + \sum_{j,l=1}^{k} \mathbb{E}R_j^2 \mathbb{E}R_l^2 = \left(\frac{k}{2}\right)^2 + \frac{3}{4}k.$$

This implies that $\mathbb{E}|\widetilde{X}|^3 \ll \sigma^3$, and the result follows as claimed from Theorem 3.4. □

We remark that in fact Ruzsa [Ruz91] proves the slightly stronger error term of σ^{-2}, but we shall not need to do so here. Proposition 3.9 completes the second step of the argument, so we are now in a position to estimate the size of the niveau set A.

Proof of Proposition 3.2. Bearing in mind that by definition of the distribution function $F_{\widetilde{X}/\sigma}(x) = F_{\widetilde{X}}(\sigma x)$, we deduce from Propositions 3.7 and 3.9 the existence of two constants c and c' such that

$$F_{\widetilde{f}}(\epsilon\sqrt{k}) \leq F_{\widetilde{X}}(\epsilon\sqrt{k}) + c\min\left\{\frac{1}{\sqrt{K}}, \frac{\sqrt{k}}{K}\right\} \leq \Phi(\sqrt{2}\epsilon) + c\min\left\{\frac{1}{\sqrt{K}}, \frac{\sqrt{k}}{K}\right\} + c'\frac{1}{\sqrt{k}}.$$

It is easy to compute that for $\epsilon \leq 1/4$, the value of the standard normal distribution function Φ at $\sqrt{2}\epsilon$ is bounded above by $2/3$, so that the size of the set A is at least $N/3$. In fact, the density can be made arbitrarily close to $1/2$ by choosing ϵ small enough. We also need to ensure that the error term $\sqrt{k}/K$ tends to 0 as N tends to infinity, and that K satisfies $K \ll N^{1/k}$. We therefore require that k grow at most like a constant times $\log N/\log\log(N)$. This proves Proposition 3.2 for N sufficiently large. □

3.2. **Counting the Number of Representations in $A - A$.** This section is devoted to proving the analogue of Lemma 2.3 for the finite Abelian group $\mathbb{Z}_N$. More precisely, we shall show that the popular difference set $D_\gamma(A)$ is contained in the complement of a ball $B_t(1)$, which in this context will be defined as

$$B_t(1) := \left\{ x \in \mathbb{Z}_N : \sum_{i=1}^{k} |\gamma_i(x) + 1| \leq t \right\} = \left\{ x \in \mathbb{Z}_N : \sum_{i=1}^{k} |\cos(\pi x r_i/N)| \leq \frac{t}{2} \right\}$$

using the same set $\gamma_1, \dots, \gamma_k$ of K-independent characters as the niveau set A. Of course we hope to be able to take the radius t as large as possible.

Proposition 3.10. *Let $\epsilon > 0$ and suppose that $k \ll \log N/\log\log N$. Let the niveau set A with parameters ϵ and k be defined as above, and write $t := \beta k$. Then provided that β and ϵ are bounded above by a suitable function of γ, we have the inclusion*

$$D_\gamma(A) \subseteq B_t(1)^C.$$

Let us first observe, as is done in Ruzsa's original paper [Ruz91], that the complete difference set $A - A$ is contained in the complement of the ball $B_{4\epsilon\sqrt{k}}(1)$. Indeed, for arbitrary $x, y \in A$, we have

$$2\epsilon\sqrt{k} \leq \Re\left[\sum_{i=1}^{k} \gamma_i(x) + \sum_{i=1}^{k} \gamma_i(y) \right],$$

which in turn is bounded above by

$$\left| \sum_{i=1}^{k} \gamma_i\left((x+y)/2\right)\left(\gamma_i\left((x-y)/2\right) + \gamma_i\left(-(x-y)/2\right)\right) \right| \leq \sum_{i=1}^{k} \left|\cos(\pi(x-y)r_i/N)\right|.$$

This proves our claim. It stands to reason that the set of popular differences $D_\gamma(A)$ should be contained in the complement of a much larger ball around 1. However, a trivial adaptation of the method we used in the model setting $\mathbb{F}_2^n$, that is, coordinate-wise counting, falls short of what is required.

Recall that we would like to show that for fixed $z \in B_t(1)$, the number of representations of z as a difference $x - y$ with x and y in A is strictly less than γN. In other words, our aim is to establish that for fixed $z \in B_t(1)$, there are few elements x such that both $x \in A$ and $x - z \in A$. This condition is equivalent to counting the number of elements $x \in \mathbb{Z}_N$ that satisfy both $\Re \sum_{j=1}^{k} \gamma_j(x) > \epsilon\sqrt{k}$ and $\Re \sum_{j=1}^{k} \gamma_j(x-z) > \epsilon\sqrt{k}$, under the assumption that $\sum_{j=1}^{k} |\gamma_j(z) + 1| = \beta k$ with $\beta = t/k$. As before, we write

$$f(x) := \sum_{j=1}^{k} \gamma_j(x) \quad \text{with real part} \quad \widetilde{f}(x) := \Re f(x),$$

but now we also need

$$g(x) = \sum_{j=1}^{k} \gamma_j(x-z) \quad \text{with real part} \quad \widetilde{g}(x) = \Re g(x).$$

Thus we are interested in an upper bound on the probability that both $\widetilde{f}$ and $\widetilde{g}$ are greater than $\epsilon\sqrt{k}$, under the hypothesis that $\sum_{j=1}^{k} |\gamma_j(z) + 1| = \beta k$. It turns out that when the parameter β is small enough, the functions $\widetilde{f}$ and $\widetilde{g}$ are sufficiently negatively correlated for this probability to be less than γ.

In order to prove this, we shall use techniques very similar to the ones we used to establish a lower bound on the size of A in the preceding section. We shall first compare the joint distribution of $(\widetilde{f}, \widetilde{g})$ with the joint distribution of two sums of appropriately defined independent random variables, and then compare their distribution to a suitable bi-variate normal.

It should be obvious at this point that we will need a 2-dimensional analogue of Esseen's Inequality, which can be found in [Sad66] and [Ber45] (with better bounds in the former).

Theorem 3.11. *Let F_1, F_2 be 2-dimensional distribution functions, and let ϕ_1, ϕ_2 be the corresponding characteristic functions. Write $\widetilde{\phi_i}(s,t) = \phi_i(s,t) - \phi_i(s,0)\phi_i(0,t)$ for $i = 1, 2$, and set*

$$\gamma_1 := \sup_{x,y} \frac{\partial F_2(x,y)}{\partial x} \quad , \quad \gamma_2 := \sup_{x,y} \frac{\partial F_2(x,y)}{\partial y}.$$

Then for any $T > 0$, the total variation distance $\sup_{x,y} |F_1(x,y) - F_2(x,y)|$ is bounded above by

$$\frac{2}{(2\pi)^2} \int_{-T}^{T} \int_{-T}^{T} \left| \frac{\widetilde{\phi_1}(s,t) - \widetilde{\phi_2}(s,t)}{st} \right| dsdt$$

plus an additional error term of the form

$$\frac{2}{\pi}\int_{-T}^{T}\left|\frac{\phi_1(s,0)-\phi_2(s,0)}{s}\right|ds+\frac{2}{\pi}\int_{-T}^{T}\left|\frac{\phi_1(0,t)-\phi_2(0,t)}{t}\right|dt+\frac{(6\sqrt{2}+8\sqrt{3})(\gamma_1+\gamma_2)}{T}.$$

As a more or less immediate corollary we have the 2-dimensional Berry-Esseen Inequality, the precise statement of which is taken from [Sad66].

Theorem 3.12. *Let $\widetilde{X}$ and $\widetilde{Z}$ be sums of k independent identically distributed mean-zero random variables $\widetilde{X}_i$, $\widetilde{Z}_i$, respectively. Let Φ_ρ denote the distribution function of a standard bi-variate normal distribution with correlation ρ. Suppose that $\widetilde{X}$ and $\widetilde{Z}$ have correlation ρ, and denote their joint distribution function by $F_{(\widetilde{X},\widetilde{Z})}$. Then*

$$\sup_{x,z}|F_{(\widetilde{X}/\sigma,\widetilde{Z}/\sigma)}(x,z)-\Phi_\rho(x,z)|\ll\frac{\widetilde{\mu}_{3,0}^{abs}+\widetilde{\mu}_{0,3}^{abs}}{\sigma^2(1-\rho^2)^2\min\{\widetilde{\mu}_{2,0}^{3/2},\widetilde{\mu}_{0,2}^{3/2}\}},$$

where we have written

$$\widetilde{\mu}_{i,j}:=\mathbb{E}\widetilde{X}^i\widetilde{Z}^j \quad \textit{and} \quad \widetilde{\mu}_{i,j}^{abs}:=\mathbb{E}|\widetilde{X}^i\widetilde{Z}^j|.$$

Let us put our idea into practice and first compare the joint distribution of $\widetilde{f}$ and $\widetilde{g}$ to the joint distribution of two sums of sequences of independent random variables with correlation ρ. In addition to

$$X:=\sum_{j=1}^{k}X_j \quad \text{with real part} \quad \widetilde{X}:=\Re X,$$

we now also define

$$Z:=\sum_{j=1}^{k}\gamma_j(-z)X_j \quad \text{with real part} \quad \widetilde{Z}:=\Re Z,$$

where the X_i are independently and uniformly distributed on the unit circle as in Section 3.1. We first show that $(\widetilde{f},\widetilde{g})$ and $(\widetilde{X},\widetilde{Z})$ are close in distribution using Theorem 3.11.

Proposition 3.13. *Let $(\widetilde{X},\widetilde{Z})$ and $(\widetilde{f},\widetilde{g})$ be defined as above, and let their joint distribution functions be denoted by $F_{(\widetilde{X},\widetilde{Z})}$ and $F_{(\widetilde{f},\widetilde{g})}$, respectively. Then the total variation distance satisfies*

$$\sup_{x,z}|F_{(\widetilde{X},\widetilde{Z})}(x,z)-F_{(\widetilde{f},\widetilde{g})}(x,z)|\ll\min\left\{\frac{1}{\sqrt{K}},\frac{\sqrt{k}}{K}\right\}.$$

Proof. We need to consider the characteristic functions

$$\phi_{(\widetilde{f},\widetilde{g})}(s,t) = \frac{1}{N}\sum_{x=1}^{N} \exp\left(i(s\widetilde{f}(x) + t\widetilde{g}(x))\right) \quad \text{and} \quad \phi_{(\widetilde{X},\widetilde{Z})}(s,t) = \mathbb{E}\exp\left(i(s\widetilde{X} + t\widetilde{Z})\right).$$

It is easy to check that the partial derivatives of $F_{(\widetilde{X},\widetilde{Z})}$ are bounded above by a constant times the standard deviation σ. Indeed, let $\eta(s,t)$ denote the joint probability density function of $(\widetilde{X},\widetilde{Z})$. By definition, we have

$$\sup_{x,z} \frac{\partial F_{(\widetilde{X},\widetilde{Z})}(x,z)}{\partial x} = \int_{-\infty}^{z} \eta(x,t)dt,$$

which by positivity of the probability density function η is bounded above by

$$\int_{-\infty}^{\infty} \eta(x,t)dt = F'_{\widetilde{X}}(x).$$

The final expression is exactly the same term as in the 1-dimensional case, which we bounded by a constant times σ using Lemma 3.8. An analogous inequality holds for the partial derivative with respect to z.

The second and third term in the bound in Theorem 3.11 are bounded above just as in the 1-dimensional case. It remains to estimate the main error term, and we shall proceed as before by comparing moments. As in the proof of Proposition 3.2, we can write

$$\phi_{(\widetilde{X},\widetilde{Z})}(s,t) = \sum_{j=1}^{l-1} \frac{i^j}{j!}\mathbb{E}(s\widetilde{X} + t\widetilde{Z})^j + \delta\frac{\mathbb{E}|s\widetilde{X} + t\widetilde{Z}|^l}{l!}$$

with $|\delta| \leq 1$, and similarly with $(\widetilde{X},\widetilde{Z})$ replaced by $(\widetilde{f},\widetilde{g})$. Let's have a closer look at $\mathbb{E}(s\widetilde{X} + t\widetilde{Z})^l$, which can be expressed as

$$\sum_{i=1}^{l}\binom{l}{i}s^it^{l-i}\mathbb{E}\widetilde{X}^i\widetilde{Z}^{l-i} = \frac{1}{2^l}\sum_{i=1}^{l}\binom{l}{i}s^it^{l-i}\sum_{c=1}^{i}\sum_{d=1}^{l-i}\binom{i}{c}\binom{l-i}{d}\mathbb{E}X^c\overline{X}^{i-c}Z^d\overline{Z}^{l-i-d}.$$

After defining the mixed moments

$$\xi_{i,j,c,d} := \mathbb{E}X^c\overline{X}^{i-c}Z^d\overline{Z}^{j-d} \quad \text{and} \quad \theta_{i,j,c,d} := \mathbb{E}f(x)^c\overline{f(x)}^{i-c}g(x)^d\overline{g(x)}^{j-d},$$

the expression for the l^{th} moment becomes

$$\mathbb{E}(s\widetilde{X} + t\widetilde{Z})^l = \frac{1}{2^l}\sum_{i=1}^{l}\binom{l}{i}s^it^{l-i}\sum_{c=1}^{i}\sum_{d=1}^{l-i}\binom{i}{c}\binom{l-i}{d}\xi_{i,l-i,c,d}.$$

As in the 1-dimensional case, we need a lemma saying that for independent characters, the low mixed moments $\xi_{i,j,c,d}$ and $\theta_{i,j,c,d}$ are equal.

Lemma 3.14. *For all $1 \leq c \leq i, 1 \leq d \leq j$ and $i + j \leq K$, we have that $\xi_{i,j,c,d} = \theta_{i,j,c,d}$.*

Proof. It is easily checked that under the given conditions both expressions reduce to the number of sequences $(m_1, \dots, m_c, n_1, \dots, n_c)$ and $(m'_1, \dots, m'_c, n'_1, \dots, n'_c)$ that are permutations of each other. □

We also need to prove a bound on $\mathbb{E}|s\widetilde{X} + t\widetilde{Z}|^l$ for even l in the style of Lemma 3.6.

Lemma 3.15. *For any even integer $l \leq K$ and $\widetilde{X}$, $\widetilde{Z}$ defined as above, we have*

$$\mathbb{E}|s\widetilde{X} + t\widetilde{Z}|^l \leq \frac{k^{l/2} l!}{2^l (l/2)!} (|s| + |t|)^l.$$

Proof. This is a straightforward computation just as in the 1-dimensional case. The moment $\xi_{i,j,c,d}$ is easily to be seen non-zero only when $2(c + d) = i + j$, in which case its absolute value is bounded above by $k^{c+d}(c + d)!$. The l^{th} moment is therefore bounded by

$$\frac{1}{2^l} \sum_{i=1}^{l} \binom{l}{i} s^i t^{l-i} \sum_{c=1}^{i} \binom{i}{c} \binom{l-i}{l/2 - c} k^{l/2} (l/2)!.$$

The sum over c in this expression is no greater than

$$\sum_{c=1}^{l/2} \binom{i}{c} \binom{l-i}{l/2 - c} k^{l/2} (l/2)!$$

and by Vandermonde convolution, the sum over the binomial coefficients actually equals $\binom{l}{l/2}$. The statement of the lemma now follows as claimed. □

We have now gathered enough information to estimate the main error term in Theorem 3.11. A not too lengthy computation using Lemmas 3.14 and 3.15 concludes the proof of Proposition 3.13 for the appropriate choice of the parameter T. □

It remains to compare the joint distribution of $(\widetilde{X}, \widetilde{Z})$ to a bi-variate standard normal distribution, and we shall do so using Theorem 3.12 in the following proposition.

Proposition 3.16. *Let $\widetilde{X}$ and $\widetilde{Z}$ be defined as above, and write $F_{\widetilde{X},\widetilde{Z}}$ for their joint distribution function. Let Φ_ρ denote the standard bi-variate normal distribution function with correlation ρ. Then*

$$\sup_{x,z} |F_{(\widetilde{X}/\sigma, \widetilde{Z}/\sigma)}(x, z) - \Phi_{-1+\beta}(x, z)| \ll \frac{1}{\sigma^{1/2}}.$$

Proof. We have already seen in Proposition 3.9 that the third absolute moment of $\widetilde{X}$ is bounded above by σ^3. A similar analysis can be carried out for $\widetilde{Z}$. For instance, writing

$z_j = -zr_j/N$ for $r_1, \ldots, r_k \in \mathbb{Z}_N$ corresponding to the characters $\gamma_1, \ldots, \gamma_k$, we find that

$$\mathbb{E}\widetilde{Z}^2 = \mathbb{E}(\sum_{j=1}^{k} \cos 2\pi z_j R_j - \sin 2\pi z_j I_j)^2 = \sum_{j=1}^{k} (\cos 2\pi z_j)^2 \mathbb{E}R_j^2 + (\sin 2\pi z_j)^2 \mathbb{E}I_j^2 = \frac{k}{2}.$$

Therefore the third absolute moments $\widetilde{\mu}_{3,0}^{abs}$ and $\widetilde{\mu}_{0,3}^{abs}$ are both bounded by σ^3. Finally, we need to check that $\widetilde{X}$ and $\widetilde{Z}$ have the required correlation, so we compute the covariance

$$\mathbb{E}\widetilde{X}\widetilde{Z} = \mathbb{E}\sum_{j=1}^{k} R_j \sum_{l=1}^{k} \cos 2\pi z_l R_l - \sin 2\pi z_l I_l = \sum_{j=1}^{k} \cos 2\pi z_j \mathbb{E}R_j^2 = (-1+\beta)\frac{k}{2}$$

by the condition we imposed on the $(z_j)_{j=1}^k$ by requiring that $z \in B_t(1)$. Thus the correlation, which is always a dimension-less quantity, of the two random variables $\widetilde{X}/\sigma$ and $\widetilde{Z}/\sigma$ with mean 0 and variance 1 is

$$\rho = \frac{\mathbb{E}\widetilde{X}\widetilde{Z}}{\sqrt{\mathbb{E}\widetilde{X}^2\mathbb{E}\widetilde{Z}^2}} = -1+\beta.$$

Proposition 3.16 now follows from Theorem 3.12. □

Last but not least, now that we have successfully approximated the distribution of $(\widetilde{f}, \widetilde{g})$ by a bi-variate normal distribution, we turn to computing the corresponding bi-variate probability.

Lemma 3.17. *Let $\epsilon, \beta, \gamma > 0$ be constants, and let $\widetilde{X}$ and $\widetilde{Z}$ be bivariate standard normal random variables with correlation $-1+\beta$. Then*

$$\mathbb{P}(\widetilde{X} \geq \sqrt{2}\epsilon \wedge \widetilde{Z} \geq \sqrt{2}\epsilon) \leq \gamma,$$

provided that β and ϵ are sufficiently small compared with γ.

Proof. We shall confine ourselves to asserting that the probability in question is less than γ provided that β is sufficiently small. This can be made precise using, for example, simple approximations to the bivariate normal with large correlation coefficient such as those in [AK94]. □

Summarising our work in this section, we have shown that $D_\gamma(A)$ is contained in the complement of a ball $B_t(1)$, where the parameter $\beta = t/k$ can be taken to be a small constant depending on γ, that is, the radius t can be taken to be of order k. This compares favourably with the statement of Lemma 2.3 in the model setting $\mathbb{F}_2^n$, where $n = \log N$ played the rôle of the parameter k.

3.3. **Using Concentration of Measure on the Torus.**

In this final section we prove the $\mathbb{Z}_N$-analogue of Lemma 2.4, that is, we show that for an appropriately chosen parameter t the complement of a ball $B_t(1)$ contains only difference sets of sets of density $o(1)$.

Proposition 3.18. *Let β be a constant and write $t=\beta k$ with $k \ll \sqrt{\log N}$. Let A_0 be any subset of $\mathbb{Z}_N$ with the property that $A_0 - A_0 \subseteq B_t(1)^C$. Then the density of A_0 is bounded above by* $\exp(-\beta^2 k/72)$.

By considering the map

$$\Psi : \mathbb{Z}_N \to \mathbb{T}^k,$$

which takes $x \mapsto (\arg\gamma_1(x), \arg\gamma_2(x), ..., \arg\gamma_k(x))/2\pi$, we move the problem to the k-dimensional torus $\mathbb{T}^k$, where appropriate measure concentration results are known. For an exhaustive survey of all aspects of measure concentration we recommend the book [Led01], and in particular Chapter 4 on concentration in product spaces. It puts into context as well as generalizes the classical probabilistic inequalities by Talagrand, which in turn are based on martingale results by Hoeffding (1963) and Azuma (1967). The precise statement of Theorem 3.19 below can be taken from page 71 of [Led01], or page 173 of [McD89], whose excellent survey article emphasizes applications to combinatorial and discrete structures.

Theorem 3.19. *Let G be a compact metric group with a translation invariant metric d and let*

$$G = G_0 \supseteq G_1 \supseteq ... \supseteq G_n = \{1_G\}$$

be a decreasing sequence of closed subspaces of G. Let $a_i = diam(G_{i-1}/G_i)$, and write $l = (\sum_{i=1}^n a_i^2)^{1/2}$. Let μ be Haar measure on G. Then for any measurable subset E of G, we have

$$\mu(E + B_d(0,t)) \geq 1 - \frac{\exp(-t^2/2l^2)}{\mu(E)}.$$

For the application we have in mind, let $G = \mathbb{T}^k$ be equipped with normalised product measure μ and metric $d(s,t) = \sum_{i=1}^k |\sin \pi(s_i - t_i)|$. It is easily checked that d is indeed a translation invariant metric on G which encapsulates the antipodal concept. Setting $G_i = \mathbb{T}^{k-i}$, the diameter a_i of each quotient G_{i-1}/G_i equals 1, whence $l^2 = k$. Denote by $C_t(1)$ the ball

$$C_t(1) := \left\{ x \in \mathbb{T}^k : \sum_{j=1}^k |\cos(\pi x_i)| \leq \frac{t}{2} \right\}.$$

The reader may care to verify that $C_t(1)$ coincides with a ball in the metric d as defined above of radius $t/2$ about the point $(1/2, 1/2, \ldots, 1/2) \in \mathbb{T}^k$. We thus have the following quantitative statement of measure concentration in $\mathbb{T}^k$ with respect to the metric d.

Corollary 3.20. *Let the metric d be defined as above, and let E be a measurable subset of $\mathbb{T}^k$. We have the bound*

$$\mu(\overline{E} + C_t(1)) \geq 1 - \frac{\exp(-t^2/8k)}{\mu(E)},$$

where the bar indicates translation by $(1/2, 1/2, \ldots, 1/2) \mod 1$.

Recall that in the model setting $\mathbb{F}_2^n$ in Section 2, we used the fact that for any subset $A_0 \subseteq \mathbb{F}_2^n$,

$$A_0 - A_0 \subseteq B_t(1)^C \quad \Rightarrow \quad \overline{A_0} + B_t(0) \cap A_0 = \emptyset.$$

In the group $\mathbb{Z}_N$ it follows from the fact that Ψ is linear and injective that any subset $A_0 \subseteq \mathbb{Z}_N$ with the property that $A_0 - A_0 \subseteq B_t(1)^C$ satisfies

$$\Psi(A_0) - \Psi(A_0) \subseteq \Psi(B_t(1)^C) = \Psi(\mathbb{Z}_N) \setminus \Psi(B_t(1)) = \Psi(\mathbb{Z}_N) \cap C_t(1)^C \subseteq C_t(1)^C,$$

and further that

$$\Psi(A_0) + C_t(1) \cap \Psi(A_0) = \emptyset \quad \Rightarrow \quad (\Psi(A_0) + C_{t/3}(1)) + C_{t/3}(1) \cap (\Psi(A_0) + C_{t/3}(1)) = \emptyset.$$

The set $\Psi(A_0) + C_{t/3}(1)$ is a union of balls in $\mathbb{T}^k$ centred at the image points of A_0 under the map Ψ. Corollary 3.20 now gives us a bound on the measure of this set of the form

$$\mu(\overline{\Psi(A_0)} + C_{t/3}(1)) \leq \exp(-t^2/72k). \tag{1}$$

We are almost done. Because the characters $\gamma_1, \ldots, \gamma_k$ are K-independent, we expect the image of $\mathbb{Z}_N$ under the map Ψ to be roughly uniformly distributed in $\mathbb{T}^k$. As we shall see shortly, this implies that the translates of the ball $C_{t/3}(1)$ generate a set of measure proportional to the density of A_0, so that we will be able to infer a bound on this density from the bound on the measure of $\overline{\Psi(A_0)} + C_{t/3}(1)$. The remainder of this section serves to make these remarks more precise.

We first turn to the equidistribution of $\mathbb{Z}_N$ under the map Ψ. We have already seen in the preceding sections that K-independence of the characters $\gamma_1, \ldots, \gamma_k$ gives us rather precise information about their distribution, and we are about to exploit this fact yet again. Let us define the *discrepancy* of a set of points $y_1, ..., y_N$ in $\mathbb{T}^k$ by

$$\operatorname{disc}(y_1, \ldots, y_N) := \sup_{B^\infty \subseteq \mathbb{T}^k} \left| \frac{|\{i : y_i \in B^\infty\}|}{N} - \mu(B^\infty) \right|,$$

where the supremum is taken over all L^∞-balls $B^\infty \subseteq \mathbb{T}^k$ and μ is, of course, Lebesgue measure as before. We shall be able to give a bound on the discrepancy of the set $\Psi(\mathbb{Z}_N)$ using the following proposition known as the *Erdős-Turán-Koksma Inequality.* It can be viewed as a quantitative version of Kronecker's Equidistribution Theorem and is taken from page 15 of [DT97].

Proposition 3.21. *Let $y_1, ..., y_N$ be points in $\mathbb{T}^k$, and let $K \in \mathbb{N}$. Then the discrepancy $\operatorname{disc}(y_1, \ldots, y_N)$ satisfies the bound*

$$\operatorname{disc}(y_1, \ldots, y_N) \leq \left(\frac{3}{2}\right)^k \left(\frac{2}{K+1} + \sum_{0<\|h\|_\infty \leq K} \frac{1}{r(h)} \left|\frac{1}{N}\sum_{i=1}^N e(h \cdot y_i)\right|\right),$$

where $r(h) = \prod_{i=1}^k \max\{1, |h_i|\}$ for $h = (h_1, ..., h_k) \in \mathbb{Z}^k$.

It should be noted (and is discussed at length in [NP73]) that Proposition 3.21 is very closely related to the Berry-Esseen Inequality. Its proof is again purely Fourier analytic, and we use it here as a black box for pure convenience. As an immediate corollary we have the following result for K-independent characters, once again illustrating the principle that K-independence of characters is the Fourier analytic (and quantitative) analogue of the notion of independence of random variables.

Corollary 3.22. *Given the map Ψ defined as above by a set $\gamma_1, \ldots, \gamma_k$ of K-independent characters, we have the bound*

$$\operatorname{dics}(\Psi(\mathbb{Z}_N)) \ll \left(\frac{3}{2}\right)^k \frac{1}{K}.$$

In other words,

$$|\{x \in \mathbb{Z}_N : \Psi(x) \in B_\eta^\infty\}| = \mu(B_\eta^\infty)N + O((3/2)^k N/K)$$

for all L^∞-balls $B_\eta^\infty \in \mathbb{T}^k$ of side length $\eta \gg K^{-1/k}$.

Recall that in Section 3 we were forced to choose $K \ll N^{1/k}$ in order for a set of K-independent characters of cardinality k to exist. This implies that we are able to resolve down to subcubes of side length $\eta \gg N^{-1/k^2}$. It is this restriction that is chiefly responsible for our bound in Theorem 1.1 in the case $G = \mathbb{Z}_N$.

Finally, we are able to make the transition from a bound on the measure of $\overline{\Psi(A_0)} + C_{t/3}(1)$ to a bound on the density of A_0.

Lemma 3.23. *Let $k \ll \sqrt{\log N}$, and let $\gamma_1, \dots, \gamma_k$ be a set of K-independent characters. Let the map Ψ be defined as above. Then for any set $A_0 \subseteq \mathbb{Z}_N$ we have*

$$|A_0| \leq \mu(\Psi(A_0) + C_{t/3}(1))N.$$

Proof. First note that $C_{t/3}(1)$ always contains the L^∞-ball $B^\infty_{t/3k}$ of side length $t/3k = \beta/3$, which implies

$$\mu(\Psi(A_0) + C_{t/3}(1)) \geq \mu(\Psi(A_0) + B^\infty_{\beta/3}).$$

Now divide $\mathbb{T}^k$ into η^{-k} subcubes of sidelength η satisfying $\eta \gg N^{-1/k^2}$ and $\eta < \beta/3$. This determines the constant required in the growth rate of k. By averaging and Corollary 3.22, at least $|\Psi(A_0)|/\eta^k N$ of these subcubes contain at least one point of $\Psi(A_0)$. Suppose these non-empty subcubes are indexed by the set $I \subseteq [\eta^{-k}]$, so that $|I| \gg |\Psi(A_0)|/\eta^k N$. But by our choice of η the subcubes B_i are smaller than the L^∞-balls $B^\infty_{\beta/3}$. It follows that

$$\mu(\Psi(A_0) + B^\infty_{\beta/3}) \geq \mu\left(\cup_{i\in I} B_i\right) = \sum_{i \in I} \mu(B_i) \gg \frac{|A_0|}{\eta^k N}\eta^k,$$

and therefore we obtain the lemma as stated. □

Lemma 3.23 and Equation (1) combine to conclude the proof of Proposition 3.18. We now bring together Propositions 3.2, 3.10 and 3.18 in order to state the main result of this paper.

Theorem 3.24. *There exists a set $A \subseteq \mathbb{Z}_N$ of size greater than $N/3$ with the property that any set A_0 whose difference set is contained in the set $D_\gamma(A)$ of γ-popular differences of A has density*

$$\exp\left(-c_\gamma \sqrt{\log N}\right),$$

where c_γ is a small constant depending on γ.

4. Remarks

Our analysis in Section 3 only relied on measure concentration in the d-dimensional torus and our ability to pick a set of independent characters. Therefore, it is evident that our methods will yield the statement of Theorem 1.1 in any finite Abelian group.

It would be interesting to establish whether the bounds in Theorem 1.1 could be improved to give a power-type decay as in Theorem 2.6.

Some interesting observations regarding the question whether Theorem 2.6 is best possible were recently made by Sanders [San08].

Acknowledgements. The author would like to thank Ben Green for posing the problem and many helpful discussions. She is also grateful to Geoffrey Grimmett for sharing his insights into Esseen's Inequality, and an anonymous editor for pointing out the reference [Kle66]. The author was supported by NSF grant DMS-0635607 for part of the period during which this work was carried out.

References

[AK94] W. Albers and W.C.M. Kallenberg. A simple approximation to the bivariate normal distribution with large correlation coefficient. *J. Multivariate Anal.*, 49:87–96, 1994.

[AS00] N. Alon and J. Spencer. *The probabilistic method.* Wiley, 2000.

[Ber41] A. C. Berry. The accuracy of the Gaussian approximation to the sum of independent variates. *Trans. Amer. Math. Soc.*, 49:122–136, 1941.

[Ber45] H. Bergstrom. On the central limit theorem in the case $\mathbb{R}^k$, $k > 1$. *Skand. Aktuarietidskr.*, 2(3):106–127, 1945.

[DT97] M. Drmota and R.F. Tichy. *Sequences, Discrepancies and Applications.* Springer, 1997.

[Ess45] Carl-Gustav Esseen. Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian law. *Acta Math.*, 77:1–125, 1945.

[GR05] B.J. Green and I. Ruzsa. Sum-free sets in abelian groups. *Israel J. Math.*, 147:157–189, 2005.

[Gre03] B.J. Green. Some constructions in the inverse spectral theory of cyclic groups. *Combin. Probab. Comput.*, 12(2):127–138, 2003.

[Gre05] B.J. Green. Finite field models in additive combinatorics. In *Surveys in combinatorics 2005*, volume 327 of *London Math. Soc. Lecture Note Ser.*, pages 1–27. Cambridge Univ. Press, Cambridge, 2005.

[Har66] L.H. Harper. Optimal numberings and isoperimetric problems on graphs. *J. Combinatorial Theory*, 1:385–393, 1966.

[JŁR00] S. Janson, T. Łuczak, and A. Rucinski. *Random graphs.* Wiley-Interscience Series in Discrete Mathematics and Optimization, 2000.

[Kle66] D. Kleitman. On a combinatorial conjecture by Erdös. *J. Combinatorial Theory*, 1:209–214, 1966.

[Led01] M. Ledoux. *The concentration of measure phenomenon.* AMS Mathematical Surveys and Monographs, 2001.

[LŁS01] V. Lev, T. Łuczak, and T. Schoen. Sum-free sets in abelian groups. *Israel J. Math.*, 125:347–367, 2001.

[McD89] C. McDiarmid. On the method of bounded differences. In *Surveys in combinatorics, 1989 (Norwich, 1989)*, volume 141 of *London Math. Soc. Lecture Note Ser.*, pages 148–188. Cambridge Univ. Press, Cambridge, 1989.

[NP73] H. Niederreiter and W. Philipp. Berry-Esseen bounds and a theorem of Erdös and Turán on uniform distribution mod 1. *Duke Math. Journal*, 40(3):633–649, 1973.

[Ruz87] I.Z. Ruzsa. Essential components. *Proc. London Math. Soc. (3)*, 54(1):38–56, 1987.

[Ruz91] I.Z. Ruzsa. Arithmetic progressions in sumsets. *Acta Arith.*, 2:191–202, 1991.

[Sad66] S.M. Sadikova. Two-dimensional analogues of an inequality of Esseen with applications to the Central Limit Theorem. *Theory of Probability and Its Applications*, 11:325–335, 1966.

[San08] T. Sanders. Popular difference sets. Available at http://arxiv.org/abs/0807.5106, 2008.

[Shi84] A.N. Shiryayev. *Probability*. Springer, 1984.

Department of Pure Mathematics and Mathematical Statistics, Wilberforce Road, Cambridge CB3 0WB, U.K.

Institute for Advanced Study, School of Mathematics, Einstein Drive, Princeton NJ 08540, U.S.A.
E-mail address: `julia.wolf@cantab.net`