BOOLEAN FUNCTIONS WITH SMALL SPECTRAL NORM, REVISITED

TOM SANDERS

1. Introduction

The purpose of this note is to present the argument from [San16] for groups of the form $\mathbb{F}_2^n$. Throughout G will denote such a group; the whole point of our arguments is that the results will not depend on n, but we shall touch on this later in the introduction.

Before motivating the problem we need a couple of definitions. Given $f : G \to \mathbb{R}$ we define its *Fourier transform* and *spectral norm* respectively by

$$\widehat{f}(r) := \mathbb{E}_{x\in G} f(x)(-1)^{r^t x} \text{ for all } r \in G, \text{ and } \|f\|_A := \sum_r |\widehat{f}(r)|.$$

It has been known since [KM93, Theorem 4.12] that if a *Boolean function* (meaning a function taking only the values 0 and 1) has small spectral norm then (an approximation to) it can be easily learnt [KM93, p1338]. (See also [Man94].) In view of this it is natural to ask how rich the class of Boolean functions with small spectral norm is, and this was part of what motivated the paper [GS08a].

There are some obvious members: if $V \leqslant G$ then it is easy to check that $\|1_V\|_A = 1$ and so if $V_1, \dots, V_L \leqslant G$ then[1] $f := \sum_{i=1}^L \pm 1_{V_i}$ is certainly integer-valued and it also has $\|f\|_A \leqslant L$. Our aim is to prove the following sort of converse.

Theorem 1.1. *Suppose that f is integer-valued with $\|f\|_A \leqslant M$. Then there are subspaces $V_1, \dots, V_L \leqslant G$ such that*

$$f = \sum_V \pm 1_V \text{ and } L \leqslant \exp(M^{3+o(1)}).$$

This improves on the bound $L \leqslant \exp(\exp(O(M^4)))$ in [GS08a, Theorem 1.3]. On the other hand if f is a sum of k maps of the form $x \mapsto (-1)^{r_i^t x}$ with the r_is independent then $\|f\|_A = O(\sqrt{k})$ so that we certainly need $L = \Omega(M^2)$ above.

The above result includes a structure theorem for Boolean functions (with small spectral norm) as they are functions taking particular integer values, but does not address the question of which linear combinations of indicator functions of subspaces lead to Boolean functions.

Various sub-classes of the Boolean functions with small spectral norm have been studied: for example the symmetric functions in [AFH12, Theorem 1.1]; the low-degree functions in

[1] This notation means that there are signs $\varepsilon_i \in \{-1, 1\}$ such that $f = \sum_{i=1}^L \varepsilon_i 1_{V_i}$.

[TWXZ13, Lemma 4]; and functions with small Fourier support in [STV14, Theorem 1.3]; and the functions in small[2] ambient group [STV14, Theorem 1.2].

This last result is particularly worth mentioning as the dependence on the size of the ambient group is very mild.

Theorem 1.2 ([STV14, Theorem 1.2]). *Suppose that f is Boolean with $\|f\|_A \leqslant M$. Then there are subspaces $V_1, \ldots, V_L \leqslant G$ such that*

$$f = \sum_V \pm 1_V \text{ and } L \leqslant 2^{M^2} n^{2M}.$$

This result is stronger than Theorem 1.1 unless $|G| \geqslant \exp(\exp(M^{2+o(1)}))$.

Weaker structural statements have also been proved about Boolean functions with small spectral norm including [Gro97, Theorem 4], [STV14, Theorems 1.1 &1.4] and [TWXZ13, Lemmas 4 & 8].

2. Outline and conditional proof of main theorem

The overall structure is not wildly different to that in [GS08a]. We use an induction over almost integer-valued functions, where we say that $f : G \to \mathbb{R}$ is *ϵ-almost integer-valued* if there is a function $f_{\mathbb{Z}}$ such that $\|f - f_{\mathbb{Z}}\|_\infty \leqslant \epsilon$. If, as will always be the base, $\epsilon < \frac{1}{2}$ then $f_{\mathbb{Z}}$ is uniquely determined.

Given a finite non-empty set S in G we write μ_S for the uniform probability measure on S. Given a function f on G we then write

$$f * \mu_S(x) := \mathbb{E}_{x+S} f \text{ for all } x \in G.$$

In particular, if $V \leqslant G$ then $1_A * \mu_V(x)$ is the relative density of A on $x + V$.

The idea is to keep splitting f up into pieces with smaller spectral norm, though they may also be less close to being integer-valued. We do this in two parts: given f we find a subspace that it correlates with. This is done first by passing to a set with small doubling using (Proposition 2.1 below), and then using a version of Freiman's theorem (Proposition 2.2 below). This result is discussed more in §5 and may be of some separate interest.

Proposition 2.1. *There is an absolute $C > 0$ such that the following holds. Suppose that f is ϵ-almost integer-valued; and $\|f\|_A \leqslant M$ with $\epsilon \leqslant \exp(-CM)$. Then there is a set $A \subset \operatorname{supp} f_{\mathbb{Z}}$ such that $|A + A| \leqslant \exp(O(M \log M))|A|$ and $|A| \geqslant \exp(-O(M \log M))|\operatorname{supp} f_{\mathbb{Z}}|$.*

Proposition 2.2. *Suppose that $A \subset G$ has $|A + A| \leqslant K|A|$. Then there is some $V \leqslant G$ with $|V| \geqslant \exp(-\log^{3+o(1)} K)|A|$ and $|A \cap V| \geqslant \exp(-\log^{1+o(1)} K)|V|$.*

The key point is that the subspace we find is quite large (in particular larger than that in [San12, Theorem A.1]), while the relative density of A on that subspace is also quite (in particular larger than that in [San13, Theorem 1.4]).

For the second part, given this subspace we try to make f behave continuously on it to restore the property of being almost integer-valued. This is the purpose of Proposition 2.3 and is roughly an analogue to [GS08a, §3].

[2]It is perhaps better to say not enormous rather than small here.

Proposition 2.3. *Suppose that $V \leqslant G$; $\|f\|_A \leqslant M$; and $\epsilon \in (0,1]$ and $p \geqslant 2$ are parameters. Then there is a subspace $U \leqslant V$ with* $\operatorname{codim}_V U = O(p\epsilon^{-2}\log \epsilon^{-1})$ *and*

$$\|f - f * \mu_U\|_{L_p(\mu_W)} \leqslant \epsilon M \text{ for all } W \in G/U.$$

With these in hand we produce our key iteration lemma:

Lemma 2.4. *There is an absolute constant $C > 0$ such that the following holds. Suppose that f is ϵ-almost integer-valued; $\|f\|_A \leqslant M$; $\eta > 0$ is a parameter and $\epsilon \leqslant \exp(-CM)$. Then there if some $V \leqslant G$ such that $f * \mu_V$ is $(\epsilon + \eta)$-almost integer-valued, $(f * \mu_V)_{\mathbb{Z}} \not\equiv 0$ and $|V| \geqslant \exp(-O(M^{2+o(1)}\max\{\log \eta^{-1}, M\}))|\operatorname{supp} f_{\mathbb{Z}}|$.*

Proof. Apply Proposition 2.1 (possible provided $\epsilon \leqslant \exp(-CM)$). This gives us a set $A \subset \operatorname{supp} f_{\mathbb{Z}}$ which we can put into Proposition 2.2 to get $U \leqslant G$ such that

$$|U| \geqslant \exp(-M^{3+o(1)})|\operatorname{supp} f_{\mathbb{Z}}| \text{ and } |A \cap U| \geqslant \exp(-M^{1+o(1)})|U|.$$

Apply Proposition 2.3 with a parameter p to be optimised and we get $V \leqslant G$ with $|V| \geqslant \exp(-O((p+M)M^{2+o(1)}))|\operatorname{supp} f_{\mathbb{Z}}|$ such that

$$\|f - f * \mu_V\|_{L_p(\mu_W)} \leqslant 2^{-4} \text{ for all } W \in G/V.$$

We may assume that $\epsilon \leqslant 2^{-4}$ and hence by the triangle inequality that

$$\|f_{\mathbb{Z}} - f * \mu_V\|_{L_p(\mu_W)} \leqslant \frac{1}{8} \text{ for all } W \in G/V,$$

and so $f * \mu_V$ (which is constant on W) is certainly $\frac{1}{8}$-almost integer-valued, but this can be bootstrapped. Suppose $W \in G/V$. Then

$$\mu_W(\{x \in W : (f * \mu_V)_{\mathbb{Z}}(x) \neq f_{\mathbb{Z}}(x)\})\, 2^{-p} \leqslant \|(f * \mu_V)_{\mathbb{Z}} - f_{\mathbb{Z}}\|^p_{L_p(\mu_W)} \leqslant \left(\epsilon + \frac{1}{8}\right)^p \leqslant 4^{-p}.$$

Writing $W = z + V$ for some $z \in G$ it follows that

$$\begin{aligned} |f * \mu_V(z) - (f * \mu_V)_{\mathbb{Z}}(z)| &\leqslant |(f - f_{\mathbb{Z}}) * \mu_V(z)| + |(f_{\mathbb{Z}} - (f * \mu_V)_{\mathbb{Z}}) * \mu_V(z)| \\ &\leqslant \epsilon + O\left(M \mu_W(\{x \in W : (f * \mu_V)_{\mathbb{Z}}(x) \neq f_{\mathbb{Z}}(x)\})\right). \end{aligned}$$

We conclude that $f * \mu_V$ is $(\epsilon + O(M2^{-p}))$-almost integer-valued from our earlier estimate for the measure. Finally, if $(f * \mu_V)_{\mathbb{Z}} \equiv 0$ then we have

$$\mu_W(\{x \in W : 0 \neq f_{\mathbb{Z}}(x)\}) \leqslant 2^{-p} \text{ for all } W \in G/V,$$

but by averaging there is some $W \in G/V$ such that $\mu_W(\operatorname{supp} f_{\mathbb{Z}}) \geqslant \mu_W(A) \geqslant \mu_U(A \cap U)$ and so we take $p = O(\max\{\log \mu_U(A \cap U)^{-1}, \log M\eta^{-1}\})$ such that $f * \mu_V$ is $(\epsilon + \eta)$-almost integer-valued, and $2^{-p} < \mu_U(A \cap U)$. The lemma follows. □

The result we shall prove (from which Theorem 1.1 follows immediately) is then the following.

Theorem 2.5. *There is an absolute constant $C>0$ such that the following holds. Suppose that f is ϵ-almost integer-valued; $\|f\|_A \leqslant M$; and $\epsilon \leqslant \exp(-CM)$. Then there are subspaces $V_1,\dots,V_L \leqslant G$ such that*

$$f_{\mathbb{Z}} = \sum_V \pm 1_V \text{ and } L \leqslant \exp(M^{3+o(1)}).$$

Proof. Let $C>0$ be the absolute constant in the statement of Lemma 2.4. Let $\epsilon_i := 2^i\epsilon + 4^{i-2M-4}\exp(-CM)$. We shall define functions f_i such that

$$f_i \text{ is } \epsilon_i\text{-almost integer-valued, } \|f_{i+1}\|_A \leqslant \|f_i\|_A - \frac{1}{2},$$

and so $(f_i - f_{i+1})_{\mathbb{Z}}$ can be written as a ± 1 sum of at most $\exp(M^{3+o(1)})$ cosets of a subspace V_i. We set $f_0 := f$ which is certainly ϵ_0-almost integer-valued. At stage $i \leqslant 2M+1$ apply Lemma 2.4 with $\eta = 4^{-2M-3}\exp(-CM)$ which is possible provided $\epsilon \leqslant \exp(-C'M)$. We get $V_{i+1} \leqslant G$ with

$$|V_{i+1}| \geqslant \exp(-M^{3+o(1)})|\operatorname{supp} f_{\mathbb{Z}}| \text{ and } f_i * \mu_{V_{i+1}} \text{ is } (\epsilon_i + \eta) - \text{almost integer-valued.}$$

Put $f_{i+1} := f_i - f_i * \mu_{V_{i+1}}$. Then f_{i+1} is $2\epsilon_i + \eta \leqslant \epsilon_{i+1}$ almost integer-valued. Moreover, since

$$|\operatorname{supp}(f * \mu_{V_{i+1}})_{\mathbb{Z}}| \leqslant 2|\operatorname{supp} f_{\mathbb{Z}}|$$

and $(f * \mu_{V_{i+1}})_{\mathbb{Z}}$ is invariant on cosets of V_{i+1} it follows from the lower bound on $|V_{i+1}|$ that $(f * \mu_{V_{i+1}})_{\mathbb{Z}}$ takes non-zero integer values on at most $\exp(M^{3+o(1)})$ translates of V_{i+1}. Moreover, the value of $(f * \mu_{V_{i+1}})_{\mathbb{Z}}$ on each of these is an integer between $-(M+1)$ and $(M+1)$. It follows that $(f_i - f_{i+1})_{\mathbb{Z}} = (f * \mu_{V_{i+1}})_{\mathbb{Z}}$ can be written as a ± 1 sum of at most $\exp(M^{3+o(1)})$ cosets of V_{i+1}.

Finally, since $(f * \mu_{V_{i+1}})_{\mathbb{Z}}$ is not identically 0 it follows that $\|f * \mu_{V_{i+1}}\|_A \geqslant 1 - \epsilon_{i+1} \geqslant \frac{1}{2}$ and hence $\|f_{i+1}\|_A \leqslant \|f_i\|_A - \frac{1}{2}$. In view of this the iteration terminates in $2M$ steps and unpacking what that means we have the result. □

The $3+o(1)$ in Theorem 2.5 arises at two different points. The first is in the application of Freiman's theorem. While we do not know how to improve that result, in this case there is a lot more structural information available to us in the proof of Proposition 2.1 and better bounds can be achieved in this setting (at the expense of the wider applicability of the result). The second is in Proposition 2.3 where, at least for fixed p (*e.g.* $p=2$) it is unclear how to improve the dependencies given the example of f being a sum of maps of the form $x \mapsto (-1)^{r_i^t x}$ for the r_is independent.

One of the key purposes of this note is to help with the understanding of [San16]. Inevitably that paper is rather more complicated but we have followed the overall structure of that work closely here. Roughly speaking Lemma 2.4 corresponds to [San16, Lemma 10.2], Proposition 2.3 to [San16, Proposition 7.1], and Proposition 2.2 to [San16, Proposition 8.1]; the least similar is Proposition 2.1 which corresponds to a combination of [San16, Lemma 9.1], [San16, Proposition 9.2] (and the Balog-Szemerédi-Gowers lemma).

3. Quantitative continuity

In this section we shall prove the following

Proposition (Proposition 2.3). *Suppose that $V \leqslant G$; $\|f\|_A \leqslant M$; and $\epsilon \in (0,1]$ and $p \geqslant 2$ are parameters. Then there is a subspace $U \leqslant V$ with $\operatorname{codim}_V U = O(p\epsilon^{-2}\log\epsilon^{-1})$ and*

$$\|f - f * \mu_U\|_{L_p(\mu_W)} \leqslant \epsilon M \text{ for all } W \in G/U.$$

To prove this we shall need the following corollary of work of Croot, Łaba and Sisask [CŁS11]. One could also proceed using Chang's Lemma [TV06, Lemma 4.35]. We have chosen this approach because it is closer to the argument in [San16, §7], where the appropriate localisation of Chang's Lemma would be more involved.

Lemma 3.1. *Suppose that $V \leqslant G$, $f \in A$ and $\epsilon \in (0,1]$, $p \geqslant 2$ are parameters. Then there is a subspace $U \leqslant V$ with $\operatorname{codim}_V U = O(p\epsilon^{-2})$ such that*

$$\|f * \mu_U - f\|_{L_p(\mu_V)} \leqslant \epsilon\|f\|_A.$$

Proof. Apply [CŁS11, Corollary 3.6] to f to get some $k = O(p\epsilon^{-2})$, $r_1, \dots, r_k \in G$, and complex numbers of unit modulus $\omega_1, \dots, \omega_k$ such that[3]

$$\left\| f - \frac{\|f\|_A}{k}\left(\omega_1(-1)^{r_1^t\cdot} + \cdots + \omega_k(-1)^{r_k^t\cdot}\right)\right\|_{L_p(\mu_V)} \leqslant \epsilon\|f\|_A.$$

Let $U := V \cap \bigcap_{i=1}^k \{x : r_i^t \cdot x = 0\}$ and note that the given sum is invariant under translation by elements of U. The result follows by the triangle inequality on rescaling ϵ. □

Proof of Proposition 2.3. We produce subspaces $U_i \leqslant V$ iteratively; initialise with $U_0 := V$. At stage i suppose that

$$\|f - f * \mu_{U_i}\|_{L_p(\mu_W)} > \epsilon\|f\|_A \text{ for some } W \in G/U_i. \tag{3.1}$$

By translation we may suppose that $W = U_i$. Apply Lemma 3.1 to f with parameter $\epsilon_0 := \frac{1}{2}\epsilon$ to get a space $Z_0 \leqslant U_i$ with

$$\operatorname{codim}_{U_i} Z_0 = O(\epsilon^{-2}p) \text{ and } \|f * \mu_{Z_0} - f\|_{L_p(\mu_U)} \leqslant \frac{1}{2}\epsilon\|f\|_A.$$

By the triangle inequality we have

$$\frac{1}{2}\epsilon\|f\|_A < \|f - f * \mu_{U_i}\|_{L_p(\mu_{U_i})} - \|f - f * \mu_{Z_0}\|_{L_p(\mu_{U_i})} \leqslant \|f * \mu_{U_i} - f * \mu_{Z_0}\|_{L_p(\mu_{U_i})}. \tag{3.2}$$

At this point we might use a simply upper bound this last term by the algebra norm and find that we have a large ℓ_1-mass of $\widehat{f}$ on $Z_0^\perp \setminus U_i^\perp$. This could then be iterated. We do a little bit better by a dyadic decomposition.[4]

Let $\epsilon_j := 2^{j-1}\epsilon$ and at stage $j > 0$ let $Z_j \leqslant Z_{j+1} \leqslant U_i$ be a space of minimal codimension such that

$$\|f * \mu_{Z_j} - f * \mu_{Z_{j+1}}\|_{L_p(\mu_{U_i})} \leqslant \epsilon_{j+1}\|f * \mu_{Z_j^\perp}\|_A.$$

[3] Here $(-1)^{r_i^t\cdot}$ denotes the functions $x \mapsto (-1)^{r_i^t x}$.

[4] This is the same idea as is discussed after [GK09, Lemma 4.1] where a power of $\frac{1}{4}$ is improved to $\frac{1}{3}$.

By Lemma 3.1 applied to $f * \mu_{Z_j}$ with parameter ϵ_{j+1} we get a space $Y_{j+1} \leqslant U_i$ with

$$\operatorname{codim}_{U_i} Y_{j+1} = O(\epsilon_{j+1}^{-2}p) \text{ and } \|f * \mu_{Z_j} - f * \mu_{Z_j} * \mu_{Y_{j+1}}\|_{L_p(\mu_{U_i})} \leqslant \epsilon_{j+1}\|f * \mu_{Z_j^\perp}\|_A.$$

It follows that we can take $Z_{j+1} = Z_j + Y_{j+1}$ and have $\operatorname{codim}_{U_i} Z_{j+1} \leqslant \operatorname{codim}_{U_i} Y_{j+1}$ and $Z_0 \leqslant Z_1 \leqslant \cdots$. On the other hand if $J = \lceil \log \epsilon^{-1} \rceil + 2$ then $\epsilon_J \geqslant 2$ and we can certainly take $Z_J = U_i$ by the triangle inequality.

By the triangle inequality it follows that

$$\|f * \mu_{U_i} - f * \mu_{Z_0}\|_{L_p(\mu_{U_i})} \leqslant \sum_{j=0}^{J} \|f * \mu_{Z_j} - f * \mu_{Z_{j+1}}\|_{L_p(\mu_{U_i})} \leqslant \sum_{j=0}^{J} \epsilon_{j+1}\|f * \mu_{Z_j}\|_A,$$

and so by averaging and (3.2) there is some j such that

$$\frac{1}{2(J+1)}\epsilon\|f\|_A \leqslant \epsilon_{j+1}\|f * \mu_{Z_j}\|_A.$$

But then $\operatorname{codim}_{U_i} Z_j = O(\epsilon_{j+1}^{-2}p)$; set $U_{i+1} := Z_j$.

Returning to our main iteration, and writing $d_{i+1} := \operatorname{codim}_{U_i} U_{i+1}$ we then have

$$1 \leqslant d_{i+1} = O(\epsilon_{j+1}^{-2}p) \text{ and } \|f * \mu_{U_{i+1}} - f * \mu_{U_i}\|_A = \Omega\left(\frac{\epsilon\|f\|_A}{\log \epsilon^{-1}}\sqrt{\frac{d_{i+1}}{p}}\right).$$

Since $U_0 \geqslant U_1 \geqslant U_2 \geqslant \dots$ we have

$$\|f\|_A \geqslant \sum_i \|f * \mu_{U_{i+1}} - f * \mu_{U_i}\|_A = \sum_i \Omega\left(\frac{\epsilon\|f\|_A}{\log \epsilon^{-1}}\sqrt{\frac{d_{i+1}}{p}}\right),$$

and since $d_i \geqslant 1$ for all $i > 0$ this sum must involve a finite number of summands and the iteration must terminate with the failure of (3.1) – exactly the conclusion we want. But then

$$\operatorname{codim}_V U \leqslant \sum_i d_i \leqslant \left(\sup_i \sqrt{d_i}\right) \sum_i \sqrt{d_i} = O(\epsilon^{-1}\sqrt{p}) \cdot O(\epsilon^{-1}\sqrt{p}\log \epsilon^{-1}).$$

This gives the result. □

4. From small Algebra norm to small doubling

The aim of this section is to prove Proposition 2.1. Throughout this section it is most natural to use counting measure, and for convenient if $f \in \ell_1(G)$ and $r \in \mathbb{N}$ we define

$$f^{(r)}(x) := \sum_{x_1 + \cdots + x_r = x} f(x_1) \cdots f(x_r) \text{ and } f^{(0)} := 1_{\{0_G\}}.$$

The key result of the section is the following. The essentially uses a refined version of arithmetic connectivity [GS08b, Definition 5.2] and is closely related to ideas of Méla from [Mél82].

Proposition (Proposition 2.1). *There is an absolute $C > 0$ such that the following holds. Suppose that f is ϵ-almost integer-valued; and $\|f\|_A \leqslant M$ with $\epsilon \leqslant \exp(-CM)$. Then there is a set $A \subset \operatorname{supp} f_{\mathbb{Z}}$ such that $|A+A| \leqslant \exp(O(M\log M))|A|$ and $|A| \geqslant \exp(-O(M\log M))|\operatorname{supp} f_{\mathbb{Z}}|$.*

Proof. Let $R := \operatorname{supp} f_{\mathbb{Z}}$ and take l and m to be natural numbers to be chosen shortly. Suppose that there is some $x \in R^m$ such that for any $S \subset [m]$ with $2 \leqslant |S| \leqslant l$ odd we have $f_{\mathbb{Z}}(\sum_{s\in S} x_s) = 0$. For $1 \leqslant i \leqslant m$ put $\omega_i := \operatorname{sgn}(f_{\mathbb{Z}}(x_i))$ and define h by $\widehat{h}(r) = \frac{1}{m}\sum_{j=1}^m \omega_i(-1)^{r^tx_j}$ so that $\|h\|_{\ell_1(G)} \leqslant 1$.

Then for every $1 \leqslant k \leqslant l$ we have

$$\begin{aligned}\left|\langle \widehat{h}^{2k+1}, \widehat{f_{\mathbb{Z}}}\rangle_{L_2(\mu_G)}\right| &= m^{-(2k+1)}\left|\mathbb{E}_r\left(\sum_{i=1}^m \omega_i(-1)^{r^tx_i}\right)^{2k+1}\overline{\widehat{f_{\mathbb{Z}}}(r)}\right|\\ &\leqslant m^{-(2k+1)}\sum_{\sigma\in[m]^{2k+1}} |f_{\mathbb{Z}}(x_{\sigma_1}+\cdots+x_{\sigma_{2k+1}})|\\ &\leqslant m^{-(2k+1)}\sum_{\substack{\sigma\in[m]^{2k+1}\\ |\{\sigma_i : i\in[m]\}|\leqslant k+1}} \|f_{\mathbb{Z}}\|_{\ell_\infty(G)}\\ &\leqslant m^{-(2k+1)}\cdot(M+1)\cdot(2k+1)\cdot m\cdot O(mk)^k = Mm^{-k}O(k)^{k+1}.\end{aligned}$$

Moreover, by Young's inequality $\|h^{(2k+1)}\|_{\ell_1(G)} \leqslant 1$ and so by Plancherel's theorem we see that

$$\left|\langle \widehat{h}^{2k+1}, \widehat{f_{\mathbb{Z}}}\rangle_{L_2(\mu_G)} - \langle \widehat{h}^{2k+1}, \widehat{f}\rangle_{L_2(\mu_G)}\right| = \left|\langle h^{(2k+1)}, f_{\mathbb{Z}} - f\rangle_{\ell_2(G)}\right| \leqslant \|f - f_{\mathbb{Z}}\|_{L_\infty(G)} \leqslant \epsilon$$

for all $0 \leqslant k \leqslant l$.

Let $P(X) = a_1X + a_3X^3 + \cdots + a_{2l+1}X^{2l+1}$ be the Chebychev polynomial (of the first kind[5]) of degree $2l+1$ so that $a_{2r+1} = (-1)^{l-r}2^{2r+1}\binom{l+r+1}{l-r}$ and $\|P\|_{L_\infty([-1,1])} \leqslant 1$. Note that

$$\left|\langle P(\widehat{h}), \widehat{f} - \widehat{f_{\mathbb{Z}}}\rangle_{L_2(\mu_G)}\right| \leqslant \epsilon\sum_i |a_i| \leqslant \epsilon\sum_{k=1}^l 2^{2k+1}\binom{l+k+1}{l-k} \leqslant \epsilon\exp(O(l)).$$

But we also have

$$\begin{aligned}|\langle P(\widehat{h}), \widehat{f_{\mathbb{Z}}}\rangle_{L_2(\mu_G)}| &\geqslant (2l+1)|\langle \widehat{h}, \widehat{f_{\mathbb{Z}}}\rangle_{L_2(\mu_G)}| - \sum_{k=1}^l |a_{2k+1}||\langle \widehat{h}^{2r+1}, \widehat{f_{\mathbb{Z}}}\rangle_{L_2(\mu_G)}|\\ &\geqslant (2l+1) - M\sum_{k=1}^l \binom{l+k+1}{l-k}O(k)^{k+1}m^{-k}\\ &\geqslant (2l+1) - MO\left(\frac{l^3}{m}\exp(O(l^2/m))\right).\end{aligned}$$

[5]See [ZKR03, S6.10.6] for details.

Since $-1 \leqslant \widehat{h}(r) \leqslant 1$ we have $|P(\widehat{h})| \leqslant 1$, and so $|\langle P(\widehat{h}), \widehat{f}\rangle_{L_2(\mu_G)}| \leqslant M$. Combining the results so far using the triangle inequality gives

$$M \geqslant (2l+1) - MO\left(\frac{l^3}{m}\exp(O(l^2/m))\right) - \epsilon \exp(O(l)).$$

It follows that if $\epsilon \leqslant \exp(C_1 l)$ for some sufficiently large $C_1 > 0$ and $m \geqslant C_2 l^3$ for some sufficiently large $C_2 > 0$ then for $l = C_3 M$ we obtain a contradiction, and the supposition at the start of the proof does not hold. In view of this we see that

$$\begin{aligned}|T|^m &\leqslant \sum_{x \in T^m} \sum_{\substack{S \subset [m]\\ 2 \leqslant |S| \leqslant 2l+1\\ |S| \equiv 1 \pmod 2}} 1_R\left(\sum_{s \in S} x_S\right)\\ &= \sum_{x \in T^m} \sum_{k=1}^{l} \sum_{\substack{S \subset [m]\\ |S| = 2k+1}} 1_R\left(\sum_{s \in S} x_S\right) = \sum_{r=1}^{l} \binom{m}{2k+1} \langle 1_T^{(2k+1)}, 1_R\rangle_{\ell_2(G)} |T|^{m-(2k+1)}.\end{aligned}$$

By averaging there is some $1 \leqslant k \leqslant l$ such that

$$1_T^{(4k+2)}(0_G)^{1/2}|R|^{1/2} \geqslant \langle 1_T^{(2k+1)}, 1_R\rangle_{\ell_2(G)} \geqslant \frac{1}{m\binom{m}{2k+1}}|R|^{2k+1}.$$

Since $1_T^{(4k+2)}(0_G) \leqslant |T|^{4k-2} 1_T^{(4)}(0_G)$ we can apply the Balog-Szemerédi-Gowers [TV06, Theorem 2.29] and we are done. □

5. A variant of Freiman's theorem

In this section we prove the following result which may be of independent interest.

Proposition (Proposition 2.2)**.** *Suppose that $A \subset G$ has $|A + A| \leqslant K|A|$. Then there is some $V \leqslant G$ with $|V| \geqslant \exp(-\log^{3+o(1)} K)|A|$ and $|A \cap V| \geqslant \exp(-\log^{1+o(1)} K)|V|$.*

It may be worth taking a moment to compare this with existing work. We consider results of the following form: if $|A + A| \leqslant K|A|$ then there is a subspace V with $\alpha := |A \cap V|/|V|$ and $\delta := |V|/|A|$.

[San13, Theorem 1.4] gives $\alpha, \delta \geqslant \exp(-\log^{3+o(1)} K)$ which is not enough for our purposes. If one wanted $\alpha \geqslant K^{-O(1)}$ then this is guaranteed by [San12, Theorem A.1]) though only with $\delta \geqslant \exp(-O(\log^4 K))$. Indeed, it actually seems likely that those methods could be used to get $\alpha = (1-\eta)K^{-1}$ and $\delta \geqslant \exp(-O_\eta(\log^4 K))$; the bound on α is close to optimal.

It is a celebrated conjecture of Marton [Ruz99] called the Polynomial Freiman-Ruzsa conjecture [TV06, Conjecture 5.34] that we can take $\alpha, \delta \geqslant K^{-O(1)}$.

Proof of Proposition 2.2. Apply [San13, Proposition 8.5] to get $r = \log^{o(1)} K$ and sets S and T with

$$2A \subset S \subset 2rA, |S + T| = O(|S|) \text{ and } |T| \geqslant \exp(-\log^{3+o(1)} K)|S|.$$

By Plünnecke's inequality we have $|2S| \leqslant \exp(\log^{1+o(1)} K)|S|$. Apply the Croot-Sisask Lemma (for example, in the form [San13, Proposition 8.3]) with a parameter m to be optimised shortly to get a set X with

$$|X| \geqslant \exp(-O(m^2 \log^{1+o(1)} K))|T| \text{ and } mX \subset 4S.$$

Let l be a natural to be optimised shortly and note by Plünnecke's inequality again that

$$\begin{aligned}|mlX| \leqslant |4lS| &\leqslant \exp(l \log^{1+o(1)} K)|S| \\ &\leqslant \exp(l \log^{1+o(1)} K + m^2 \log^{1+o(1)} K + \log^{3+o(1)} K)|X|.\end{aligned}$$

Put $k = ml$ and see that for K sufficiently large we can choose $m = \log^{1+o(1)} K$ (and with $m \geqslant r \log K$) and $l = \log^{2+o(1)} K$ such that $|(3k+1)X| < 2^k|X|$. We apply Chang's covering lemma [San13, Lemma 5.2] to get a set $T \subset X$ of size $k = \log^{3+o(1)} K$ such that $3X \subset \langle T \rangle + X$. It follows that $U := (k+1)X$ is a group and since $U + A \subset (16rl+1)A$, we can use averaging and Plünnecke's inequality to get that $\mu_U(X) \geqslant \exp(-\log^{3+o(1)} K)$.

So far the argument is essentially the same as the proof of [San13, Proposition 2.5], and if we stopped here then we would have the conclusion but with a weaker bound on the relative density of A in V. We shall bootstrap what we have into the better result by using the argument behind Chang's theorem [TV06, Theorem 4.41] for two different sets. First, note that

$$\prod_{s=0}^{m-1} \frac{|A + sX + X|}{|A + sX|} \leqslant \frac{|mX + A|}{|A|} \leqslant \frac{|(8r+1)A|}{|A|} \leqslant K^{O(r)},$$

and so by the the pigeon-hole principle there is a set $L = A + sX$ for some $0 \leqslant s \leqslant m-1$ such that $|X + L| \leqslant K^{O(rm^{-1})}|L| = O(|L|)$. (Here we use that $m \geqslant r \log K$.) Since the cosets of U partition G, by averaging there is some translate of U, call it W, such that $|X + W \cap L| \leqslant \exp(\log^{o(1)} L)|W \cap L|$; put

$$B := L \cap W \text{ and } D := \frac{|B + X|}{|B|} = O(1).$$

We may translate so that $W = U$ and now work with the Fourier transform on U. By Chang's Lemma [TV06, Lemma 4.35] there is a set $\Lambda \subset \widehat{U}$ with

$$|\Lambda| = O(D \log \mu_U(X)^{-1}) = O(\log^{3+o(1)} K) \text{ and } \operatorname{Spec}_{\frac{1}{2\sqrt{D}}}(1_X) \subset \langle \Lambda \rangle.$$

It follows that if $x \in \operatorname{Spec}_{\frac{1}{2\sqrt{D}}}(1_X)^{\perp}$ then

$$\begin{aligned}|1_B * 1_B * 1_X * 1_X(x) - 1_B * 1_B * 1_X * 1_X(0_G)| &= \left|\sum_r |\widehat{1_B}(r)|^2 |\widehat{1_X}(r)|^2((-1)^{r^t x} - 1)\right| \\ &\leqslant 2 \sum_{r \notin \operatorname{Spec}_{\frac{1}{2\sqrt{K}}}(1_X)} |\widehat{1_B}(r)|^2 |\widehat{1_X}(r)|^2 \\ &\leqslant \mu_U(B) \frac{1}{2D} \mu_U(X)^2,\end{aligned}$$

by Parseval's theorem. On the other hand, by the Cauchy-Schwarz inequality we have

$$1_B * 1_B * 1_X * 1_X(0_G) \geqslant \frac{(\mu_U(B)\mu_U(X))^2}{\mu_U(X+B)} \geqslant \frac{1}{D}\mu_U(B)\mu_U(X)^2.$$

We conclude that $V := \langle \Lambda \rangle^\perp \subset B + B + X + X \subset 2mX + 2A \subset (16r + 2)A$. Finally, we have

$$\|1_A * \mu_V\|_{L_\infty(G)}\mu_G(A) \geqslant \|1_A * \mu_V\|_{L_2(G)}^2 \geqslant \frac{\mu_G(A)^2}{\mu_G(A+V)} \geqslant \mu_G(A)K^{-O(r)},$$

by the Cauchy-Schwarz and Plünnecke inequalities. It follows that $|A \cap (x+V)| \geqslant K^{O(r)}|V|$ for some $x \in G$ and we get the result by enlarging V if necessary to be the group generated by $x + V$ – this is at most twice as big. □

Acknowledgment

The author should like to thank Ben Green for suggesting that the writing of this paper might make the work of [San16] more accessible.

References

[AFH12] A. Ada, O. Fawzi, and H. Hatami. Spectral norm of symmetric functions. In Anupam Gupta, Klaus Jansen, José Rolim, and Rocco Servedio, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 338–349, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[CŁS11] E. S. Croot, I. Łaba, and O. Sisask. Arithmetic progressions in sumsets and L^p-almost-periodicity. 2011, arXiv:1103.6000.

[GK09] B. J. Green and S. V. Konyagin. On the Littlewood problem modulo a prime. *Canad. J. Math.*, 61(1):141–164, 2009.

[Gro97] V. Grolmusz. On the power of circuits with gates of low L_1 norms. *Theoret. Comput. Sci.*, 188(1-2):117–128, 1997.

[GS08a] B. J. Green and T. Sanders. Boolean functions with small spectral norm. *Geom. Funct. Anal.*, 18(1):144–162, 2008, arXiv:math/0605524.

[GS08b] B. J. Green and T. Sanders. A quantitative version of the idempotent theorem in harmonic analysis. *Ann. of Math. (2)*, 168(3):1025–1054, 2008, arXiv:math/0611286.

[KM93] E. Kushilevitz and Y. Mansour. Learning decision trees using the fourier spectrum. *SIAM Journal on Computing*, 22(6):1331–1348, 1993, https://doi.org/10.1137/0222080.

[Man94] Y. Mansour. Learning Boolean functions via the Fourier transform. *Theoretical Advances in Neural Computation and Learning*, pages 391–424, 1994.

[Mél82] J.-F. Méla. Mesures ε-idempotentes de norme bornée. *Studia Math.*, 72(2):131–149, 1982.

[Ruz99] I. Z. Ruzsa. An analog of Freĭman's theorem in groups. *Astérisque*, (258):xv, 323–326, 1999. Structure theory of set addition.

[San12] T. Sanders. On the Bogolyubov-Ruzsa lemma. *Anal. PDE*, 5(3):627–655, 2012, arXiv:1011.0107.

[San13] T. Sanders. The structure theory of set addition revisited. *Bull. Amer. Math. Soc.*, 50:93–127, 2013, arXiv:1212.0458.

[San16] T. Sanders. Bounds in the idempotent theorem. *ArXiv e-prints*, October 2016, 1610.07092.

[STV14] A. Shpilka, A. Tal, and B. Volk. On the structure of Boolean functions with small spectral norm. In *Proceedings of the 5th Conference on Innovations in Theoretical Computer Science*, ITCS '14, pages 37–48, New York, NY, USA, 2014. ACM.

[TV06] T. C. Tao and H. V. Vu. *Additive combinatorics*, volume 105 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006.

[TWXZ13] H.-Y. Tsang, C. Wong, N. Xie, and S. Zhang. Fourier sparsity, spectral norm, and the log-rank conjecture. In *Proceedings of the 2013 IEEE 54th Annual Symposium on Foundations of Computer Science*, FOCS '13, pages 658–667, Washington, DC, USA, 2013. IEEE Computer Society.

[ZKR03] D. Zwillinger, S. G. Krantz, and K. H. Rosen, editors. *CRC standard mathematical tables and formulae*. CRC Press, Boca Raton, FL, 31st edition, 2003.

Mathematical Institute, University of Oxford, Radcliffe Observatory Quarter, Woodstock Road, Oxford OX2 6GG, United Kingdom

Email address: tom.sanders@maths.ox.ac.uk