Journal de Théorie des Nombres
de Bordeaux 00 (XXXX), 000–000

Extremal Sidon Sets are Fourier Uniform, with Applications to Partition Regularity

par Miquel Ortega et Sean Prendiville

Résumé. En généralisant des résultats d'Erdős-Freud et Lindström, nous prouvons que le plus grand sous-ensemble de Sidon d'un intervalle d'entiers borné est équidistribué dans des voisinages de Bohr. Nous le faisons en montrant que les ensembles de Sidon extrémaux sont Fourier-pseudoaléatoire, dans le sens qu'ils n'ont pas de coefficients de Fourier grands non triviaux. Comme application, nous en déduisons que pour une equation à cinq ou plus variables et régulière sous partitions, toute coloration finie d'un ensemble extrémal de Sidon à une solution monochromatique.

Abstract. Generalising results of Erdős-Freud and Lindström, we prove that the largest Sidon subset of a bounded interval of integers is equidistributed in Bohr neighbourhoods. We establish this by showing that extremal Sidon sets are Fourier-pseudorandom, in that they have no large non-trivial Fourier coefficients. As a further application we deduce that, for any partition regular equation in five or more variables, every finite colouring of an extremal Sidon set has a monochromatic solution.

Contents

Manuscrit reçu le September 22, 2022.
2010 *Mathematics Subject Classification.* 11B30, 11B25.
Mots-clefs. Sidon sets, pseudorandomness, equidistribution, partition regularity.

1. Introduction

A subset S of an additively-written abelian group is *Sidon* if every non-zero x has at most one representation as a difference $x = s_1 - s_2$ with $s_1, s_2 \in S$. There have been a number of works investigating the size of Sidon sets $S \subset \mathbb{Z}$. Erdős and Turán [10] established the well-known bound[1]

$$|(n, n+N] \cap S| \leqslant N^{1/2} + O(N^{1/4}). \tag{1.1}$$

A corresponding lower bound was found by Singer [22], who constructed a Sidon set $S \subset [N] := \{1, 2, \dots, N\}$ of size

$$|S| \geqslant N^{1/2} - O(N^{\alpha/2}), \tag{1.2}$$

where α is a real number for which there is always a prime in $[x - x^{\alpha}, x]$ when x is large (the current record [1] is $\alpha = 0.525$).

Informally, we call a Sidon set $S \subset [N]$ *extremal* if its size is 'close' to $N^{1/2}$ in some sense. There has been speculation on the (im)possibility of characterising such sets [19, 13, 12, 7, 8]. We contribute to this discussion by showing that extremal Sidon sets are Fourier pseudorandom, by which we mean that (after appropriate renormalisation) their Fourier transform behaves essentially like the Fourier transform of the ambient interval.

Definition 1.1 (Fourier transform)**.** For $f : \mathbb{Z} \to \mathbb{C}$ with finite support define $\hat{f} : \mathbb{T} \to \mathbb{C}$ by

$$\hat{f}(\alpha) := \sum_{n \in \mathbb{Z}} f(n) e(\alpha n).$$

Here $e(\beta)$ stands for $e^{2\pi i \beta}$.

Theorem 1.2 (Fourier uniformity)**.** *Let $S \subset [N]$ be a Sidon set. Then*

$$\left\| \hat{1}_S - \tfrac{|S|}{N} \hat{1}_{[N]} \right\|_\infty \ll N^{1/2} \left(\left| \tfrac{|S|}{N^{1/2}} - 1 \right| + N^{-1/6} \right)^{1/2}. \tag{1.3}$$

Remark 1.3. The exponent $-1/6$ appearing in (1.3) can be improved to $-1/4$. This is accomplished by replacing a use of (1.1) in our proof with a sharper estimate of Cilleruelo [6]; see Theorem 6.3.

We note that on combining the Erdős-Turán upper bound (1.1) with Singer's lower bound (1.2), the largest Sidon subset S of $[N]$ satisfies $\left| |S| - N^{1/2} \right| \ll N^{21/80}$, in which case the $N^{-1/6}$ error term dominates (1.3).

Corollary 1.4. *The largest Sidon subset $S \subset [N]$ satisfies*

$$\left\| \hat{1}_S - \tfrac{|S|}{N} \hat{1}_{[N]} \right\|_\infty \ll \left\| \hat{1}_S \right\|_\infty N^{-\frac{1}{12}}. \tag{1.4}$$

---

[1] For a proof see Appendix A.

We are not the first to investigate the uniformity of extremal Sidon sets. Erdős and Freud [9] established that such sets are equidistributed in short intervals[2], whilst Lindström [17] proved equidistribution in arithmetic progressions[3]. We are able to re-prove (quantitatively weaker) versions of these results as a consequence of Theorem 1.2.

Corollary 1.5 (Equidistribution in short intervals)**.** *Let $I \subset [0,1]$ be an interval and $S \subset [N]$ a Sidon set of size*[4] $|S| \geqslant \frac{1}{100}N^{1/2}$. *Then we have the asymptotic*[5]

$$\mathbb{E}_{x\in S}1_I(x/N) = \mathrm{meas}(I) + O_\varepsilon\left(N^\varepsilon\left(\left|\tfrac{|S|}{N^{1/2}} - 1\right| + N^{-1/6}\right)^{1/2}.\right) \tag{1.5}$$

Corollary 1.6 (Equidistribution in residue classes)**.** *For any congruence class $a \pmod q$ and Sidon set $S \subset [N]$ of size $|S| \geqslant \frac{1}{100}N^{1/2}$ we have the asymptotic*

$$\mathbb{E}_{x\in S}1_{q\cdot\mathbb{Z}+a}(x) = q^{-1} + O_\varepsilon\left(N^\varepsilon\left(\left|\tfrac{|S|}{N^{1/2}} - 1\right| + N^{-1/6}\right)^{1/2}.\right) \tag{1.6}$$

In fact, Theorem 1.2 is more general than Corollaries 1.5 and 1.6, yielding equidistribution in a wider class of sets. In the following, we write $\mathbb{T} := \mathbb{R}/\mathbb{Z}$ and $\|\alpha\|_{\mathbb{T}} := \min_{n\in\mathbb{Z}}|\alpha - n|$.

Corollary 1.7 (Equidistribution in smooth Bohr neighbourhoods)**.** *Let $F : \mathbb{T}^d \to [0,1]$ have Lipschitz constant $K \geqslant 1$, in that for any $\alpha, \beta \in \mathbb{T}^d$ we have*

$$|F(\alpha) - F(\beta)| \leqslant K \max_j \|\alpha_j - \beta_j\|_{\mathbb{T}}.$$

Then for any Sidon set $S \subset [N]$ of size $|S| \geqslant \frac{1}{100}N^{1/2}$ and $\alpha \in \mathbb{T}^d$ we have

$$\mathbb{E}_{x\in S}F(\alpha x) = \mathbb{E}_{x\in[N]}F(\alpha x) + O\left(K^{\frac{2}{3}}\left(\left|\tfrac{|S|}{N^{1/2}} - 1\right| + N^{-1/6}\right)^{\frac{1}{8d}}\right).$$

Remark 1.8. We have not striven for quantitative efficiency in the error term of Corollary 1.7, which could be easily improved.

We can drop the smoothness assumption on the Bohr neighbourhood if we are prepared to assume that it is *regular*.

Definition 1.9 (Regular Bohr set)**.** Given $\alpha \in \mathbb{T}^d$ and $\rho > 0$, we say that the Bohr set

$$B(\alpha, \rho) := \left\{x \in [N] : \max_i \|\alpha_i x\| \leqslant \rho\right\}$$

---

[2] There have since been quantitative improvements in this result, see [6].

[3] For quantitative improvements, see [16].

[4] One could replace the factor 1/100 with any positive absolute constant. This assumption makes our conclusions notationally simpler, and is always satisfied in the range of interest, when $|S| = N^{1/2}(1 + o(1))$ with N large.

[5] See (1.7) for the definition of $\mathbb{E}$.

is *regular* if for any $|\kappa| \leqslant \frac{1}{100d}$ we have

$$\left| \frac{|B(\alpha, (1+\kappa)\rho)|}{|B(\alpha, \rho)\,|} - 1 \right| \leqslant 100d|\kappa|.$$

Remark 1.10. Bourgain [2] established that regular Bohr sets are ubiquitous; see Tao and Vu [23, Lemma 4.25].

Corollary 1.11 (Equidistribution in regular Bohr sets)**.** *Let $B = B(\alpha, \rho)$ be a regular Bohr set with $\alpha \in \mathbb{T}^d$. Then for any Sidon set $S \subset [N]$ of size $|S| \geqslant \frac{1}{100} N^{1/2}$ we have*

$$\mathbb{E}_{x\in S} 1_B(x) = \mathbb{E}_{x\in[N]} 1_B(x) + O\left(d\rho^{-1} \left(\left| \tfrac{|S|}{N^{1/2}} - 1 \right| + N^{-1/6} \right)^{\frac{1}{14d}} \right).$$

Remark 1.12. As in Corollary 1.6, the error term in Corollary 1.11 can be improved with a more refined analysis.

1.1. Partition regularity over extremal Sidon sets. We offer a further application of Theorem 1.2 in proving a colouring analogue of a recent result of Conlon, Fox, Sudakov and Zhao [3], this being the original motivation for our paper. See also Prendiville [20] for a proof of their result with similar techniques to the ones we use in this paper.

Informally, call a Sidon subset of $[N]$ *dense* if its cardinality is a positive proportion of $N^{1/2}$, say $\frac{1}{100} N^{1/2}$. The authors of [3] show that any dense Sidon subset of $[N]$ contains a non-trivial[6] solution to the equation

$$a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 = (a_1 + a_2 + a_3 + a_4)x_5.$$

The essential features of this equation are that its coefficients sum to zero and that it has at least five variables. The five variable condition cannot be relaxed: every Sidon set lacks non-trivial solutions to the four-variable equation $x_1 - x_2 - x_3 + x_4 = 0$. The assumption that the coefficients sum to zero is also necessary, as we now show.

Proposition 1.13. *There exists a Sidon subset of $[N]$ with at least $\frac{1}{\sqrt{2}} N^{1/2}(1 + o(1))$ elements, and which has no solutions to the equation $x_1 + x_2 + x_3 + x_4 = x_5$*

Proof. Take an extremal Sidon subset S_0 of $[\frac{N-1}{2}]$ and set $S := 2\cdot S_0 + 1$. $\square$

Remark 1.14. The above construction can be adapted to show that for any homogeneous linear equation whose coefficients do not sum to zero there exists a dense Sidon sets lacking solutions to the equation.

To accommodate equations whose coefficients do not sum to zero, one may speculate on whether a colouring analogue of the results of Conlon,

[6]For instance, with all variables distinct.

Fox, Sudakov and Zhao [3] should hold; cf. Rado's criterion for partition regularity [14, §3.2] versus Roth's criterion for density regularity [21]. Proposition 1.13 indicates that one cannot hope to always find monochromatic solutions to $x_1 + x_2 + x_3 + x_4 = x_5$ in colourings of *dense* Sidon sets. In the following theorem we show that such a result does hold for colourings of *extremal* Sidon sets.

Theorem 1.15 (Partition regularity over extremal Sidon sets)**.** *Let $c_1, \dots, c_s \in \mathbb{Z} \setminus \{0\}$ with $s \geqslant 5$ and suppose that there exists a non-empty index set $I \subset [s]$ satisfying $\sum_{i \in I} c_i = 0$. Let r be a positive integer and $S \subset [N]$ a Sidon set. Then at least one of the following holds:*

- *N is small, in that $N \ll_{c_1,\dots,c_s,r} 1$.*
- *S is not extremal, in that $\big||S| - N^{1/2}\big| \gg_{c_1,\dots,c_s,r} N^{1/2}$.*
- *Partition regularity: For any r-colouring $S = C_1 \cup \dots \cup C_r$, there exists a colour class C_j such that*

$$\sum_{c_1x_1+\dots+c_sx_s=0} 1_{C_j}(x_1) \cdots 1_{C_j}(x_s) \gg_{c_1,\dots,c_s,r} |S|^s N^{-1}$$

Remark 1.16. If $\sum_{i \in I} c_i \neq 0$ for all $\emptyset \neq I \subset [s]$, then there exists a finite colouring of $\mathbb{N}$ with no monochromatic solutions to the equation $c_1x_1 + \dots + c_sx_s = 0$ (see Rado's criterion for partition regularity [14, §3.2]). Hence the assumption that some subset of coefficients sums to zero is necessary.

Conlon, Fox, Sudakov and Zhao [3] derive their results on dense Sidon sets via a removal lemma for C_4-free graphs. Since "extremal" C_4-free graphs are pseudorandom, see [15, Theorem 5.1], the transference approach employed in this paper can surely be combined with a counting lemma from [3, Theorem 3.1], to prove that every finite colouring of an "extremal" C_4-free graph has a monochromatic C_5. It may be interesting to investigate which other monochromatic subgraphs can be guaranteed in this manner, see [4].

Acknowledgements. We thank David Conlon for an informative talk on [3] in the *Webinar in Additive Combinatorics*[7].

Notation.

Standard conventions. We use $[N]$ to denote the set of consecutive integers $\{1, 2, \dots, N\}$. We use counting measure on $\mathbb{Z}$, so that for $f, g : \mathbb{Z} \to \mathbb{C}$, we have

$$\|f\|_p := \left(\sum_x |f(x)|^p\right)^{\frac{1}{p}} \text{ and } (f * g)(x) := \sum_y f(y)g(x-y).$$

[7] `https://sites.google.com/view/web-add-comb/`

Any sum of the form $\sum_x$ is to be interpreted as a sum over $\mathbb{Z}$. The *support* of f is the set $\operatorname{supp}(f) := \{x \in \mathbb{Z} : f(x) \neq 0\}$. For a finite set S and function $f : S \to \mathbb{C}$, denote the average of f over S by

$$\mathbb{E}_{s\in S} f(s) := \frac{1}{|S|}\sum_{s\in S} f(s). \tag{1.7}$$

We use Haar probability measure on $\mathbb{T} := \mathbb{R}/\mathbb{Z}$, so that for integrable $F : \mathbb{T} \to \mathbb{C}$, we have

$$\|F\|_p := \left(\int_{\mathbb{T}} |F(\alpha)|^p \mathrm{d}\alpha\right)^{\frac{1}{p}} = \left(\int_0^1 |F(\alpha)|^p \mathrm{d}\alpha\right)^{\frac{1}{p}}$$

and

$$\|F\|_\infty := \sup_{\alpha\in\mathbb{T}} |F(\alpha)|.$$

Write $\|\alpha\|_{\mathbb{T}} := \min_{n\in\mathbb{Z}} |\alpha - n|$ for the distance from $\alpha \in \mathbb{R}$ to the nearest integer. This remains well-defined on $\mathbb{T}$.

Asymptotic notation. For a complex-valued function f and positive-valued function g, write $f \ll g$ or $f = O(g)$ if there exists a constant C such that $|f(x)| \leq Cg(x)$ for all x. We write $f = \Omega(g)$ if $f \gg g$. The notation $f \asymp g$ means that $f \ll g$ and $f \gg g$. We subscript these symbols if the implicit constant depends on additional parameters.

We write $f = o(g)$ if for any $\varepsilon > 0$ there exists $X \in \mathbb{R}$ such that for all $x \geqslant X$ we have $|f(x)| \leqslant \varepsilon g(x)$.

Local conventions. Up to normalisation, all of the above are widely used in the literature. Next, we list notation specific to our paper. We have tried to minimise this in order to aid the casual reader.

For a real parameter $H \geqslant 1$, we use $\mu_H : \mathbb{Z} \to [0,1]$ to represent the following normalised Fejér kernel

$$\mu_H(h) := \frac{1}{\lfloor H\rfloor}\left(1 - \frac{|h|}{\lfloor H\rfloor}\right)_+ = \frac{(1_{[H]} * 1_{-[H]})(h)}{\lfloor H\rfloor^2}, \tag{1.8}$$

where $[H] = [\lfloor H\rfloor]$. This is a probability measure on $\mathbb{Z}$ with support in the interval $(-H, H)$.

2. Fourier uniformity

Given an extremal Sidon set $S \subset [N]$, our main goal in this section is to show that S is a Fourier uniform subset of $[N]$, by proving Theorem 1.2. Recalling our notation (1.8) for the Fejér kernel, we begin with the following version of van der Corput's inequality.

Lemma 2.1 (van der Corput differencing)**.** *Suppose that* $1 \leqslant H \leqslant N$, $f\colon \mathbb{Z} \to \mathbb{C}$ *and* $\operatorname{supp}(f) \subset [N]$*. Then*

$$\left|\sum_x f(x)\right|^2 \leqslant (N+H)\sum_h \mu_H(h)\sum_x f(x)\overline{f(x+h)}. \tag{2.1}$$

Proof. We have

$$\left|\sum_x f(x)\right|^2 = \left|\mathbb{E}_{[H]}\sum_x f(x+h)\right|^2 = \left|\sum_x \mathbb{E}_{[H]} f(x+h)\right|^2.$$

By Cauchy-Schwarz, the latter quantity is bounded by

$$(N+H)\sum_x \left|\mathbb{E}_{[H]} f(x+h)\right|^2 = (N+H)\sum_x \frac{1}{\lfloor H\rfloor^2}\sum_{h_1,h_2\in[H]} f(x+h_1)\overline{f(x+h_2)}.$$

We obtain the desired inequality on changing variables in x and using (1.8). $\square$

The next lemma tells us that, on taking an appropriately sized H and supposing that S is extremal Sidon, the sum of $\mu_H(h)$ over $S-S$ is nearly 1.

Lemma 2.2. *Let* $S \subset [N]$ *be a Sidon set. Then*

$$\sum_{h\in(S-S)\setminus\{0\}} \mu_H(h) \geqslant \frac{|S|^2}{N+H} - \frac{|S|}{\lfloor H\rfloor}.$$

Proof. Since S is Sidon $1_S * 1_{-S}(x) = 1_{S-S}$ if $x \neq 0$. In addition $1_S * 1_{-S}(0) = |S|$, but this does not require S to be Sidon. Hence

$$\sum_{h\in(S-S)\setminus\{0\}} \mu_H(h) = \sum_h 1_S * 1_{-S}(h)\mu_H(h) - \frac{|S|}{\lfloor H\rfloor}. \tag{2.2}$$

Using (1.8) and expanding convolutions we have

$$\begin{aligned}\sum_h 1_S * 1_{-S}(h)\mu_H(h) &= \frac{1}{\lfloor H\rfloor^2}\sum_h 1_S * 1_{-S}(h) 1_{[H]} * 1_{-[H]}(h)\\ &= \frac{1}{\lfloor H\rfloor^2}\sum_h 1_S * 1_{[H]}(h)^2.\end{aligned}$$

We may now apply Cauchy-Schwarz to this last sum, giving us

$$\sum_h 1_S * 1_{-S}(h)\mu_H(h) \geqslant \frac{1}{(N+H)\lfloor H\rfloor^2}\left(\sum_h 1_S * 1_{[H]}(h)\right)^2 = \frac{|S|^2}{(N+H)},$$

which gives the claimed inequality. $\square$

We are now in a position to prove Theorem 1.2. Our use of van der Corput's inequality is reminiscent of the moving averages argument used by Erdős and Turán [10].

Proof of Theorem 1.2. Let $f_1 := 1_S$, $f_2 := \frac{|S|}{N} 1_{[N]}$ and $f := f_1 - f_2$. Applying van der Corput's inequality (2.1) to $\hat{f}$, with $1 \leqslant H \leqslant N$ to be determined, we obtain

$$|\hat{f}(\alpha)|^2 \leqslant (N+H) \sum_h \mu_H(h) \Big| \sum_x \big[f_1(x) f_1(x+h) - f_1(x) f_2(x+h) - f_2(x) f_1(x+h) + f_2(x) f_2(x+h) \big] \Big|.$$

We claim that $\sum_h \mu_H(h) ||S|^2 N^{-1} - \sum_x f_i(x) f_j(x+h)|$ is small for all choices of i, j, so that main terms cancel and we are left only with error terms.

Since $f_2(x) f_2(x+h) = |S|^2 N^{-2} 1_{[N] \cap ([N]-h)}(x)$ we have

$$\sum_h \mu_H(h) \left| |S|^2 N^{-1} - \sum_x f_2(x) f_2(x+h) \right| = \frac{|S|^2}{N^2} \sum_h \mu_H(h) |h| \leqslant \frac{H|S|^2}{N^2}.$$

We have the identity

$$f_1(x) f_2(x+h) = \frac{|S|}{N} 1_{S \cap ([N]-h)}(x) = \frac{|S|}{N} [1_S(x) - 1_{S \cap I_h}(x)],$$

where $I_h \subset [N]$ is an interval of $|h|$ integers. Since S is a Sidon set, the bound (1.1) gives that

$$|S \cap I_h| \ll \sqrt{|h|}. \tag{2.3}$$

Hence

$$\sum_h \mu_H(h) \left| |S|^2 N^{-1} - \sum_x f_1(x) f_2(x+h) \right| \ll \sum_h \mu_H(h) |S| N^{-1} |h|^{1/2} \ll \frac{|S| H^{1/2}}{N}.$$

By symmetry, the same bound applies to the term involving $f_2(x) f_1(x+h)$.

Note that $f_1(x) f_1(x+h) = 1_{S \cap (S-h)}(x)$. On account of S being Sidon, for $h \neq 0$ we have

$$\left| |S|^2 N^{-1} - \sum_x f_1(x) f_1(x+h) \right| \leqslant 1 - 1_{S-S}(h) + \big| |S|^2 - N \big| N^{-1} = 1 - 1_{S-S}(h) + O\Big(\big| |S| - N^{1/2} \big| N^{-1/2} \Big).$$

Thus by Lemma 2.2 we obtain

$$\sum_h \mu_H(h) \left| |S|^2 N^{-1} - \sum_x f_1(x) f_1(x+h) \right| \leqslant 1 - \frac{|S|^2}{N+H} + O\left(\frac{|S|}{\lfloor H \rfloor} + \frac{\big||S| - N^{1/2}\big|}{N^{1/2}} \right).$$

Putting everything together, we deduce that

$$|\hat{f}(\alpha)|^2 \ll N^{1/2} \Big| |S| - N^{1/2} \Big| + \frac{N^{3/2}}{\lfloor H \rfloor} + N^{1/2} H^{1/2}.$$

Balancing error terms, we set $H := N^{2/3}$ which gives

$$|\hat{f}(\alpha)|^2 \ll N^{1/2} \Big| |S| - N^{1/2} \Big| + N^{5/6}. \qquad \square$$

3. Equidistribution in progressions

In this section we derive Corollaries 1.5 and 1.6 from Theorem 1.2. Both corollaries are immediate consequences of the following.

Theorem 3.1. *Let $S \subset [N]$ be a Sidon set and $P \subset \mathbb{Z}$ a finite arithmetic progression. Then*

$$|S \cap P| = \frac{|[N] \cap P||S|}{N} + O_\varepsilon\left(N^{1/2+\varepsilon} \left(\left| \tfrac{|S|}{N^{1/2}} - 1 \right| + N^{-1/6} \right)^{1/2} \right).$$

Theorem 3.1 follows from Theorem 1.2 by the Erdős-Turán inequality (see, for example, [18, Corollary 1.1]). We prove a cheap version of this inequality that suffices for our purposes. We begin with a standard estimate for the L^1 norm of the Fourier transform of a progression.

Lemma 3.2. *Let $P \subset \mathbb{Z}$ be an arithmetic progression. Then*

$$\int_{\mathbb{T}} \big| \hat{1}_P(\alpha) \big| \, \mathrm{d}\alpha \ll \log(|P|).$$

Proof. With a suitable change of variables, we may assume that $P = \{1, \ldots, |P|\}$. Let us first prove a pointwise bound. Summing the geometric series gives

$$\big| \hat{1}_P(\alpha) \big| \ll \frac{1}{|1 - e(\alpha)|} = \frac{1}{|e(-\alpha/2) - e(\alpha/2)|} = \frac{1}{|2\sin(\pi\alpha)|} \ll \frac{1}{\|\alpha\|},$$

where we have used that $2|\alpha| \leqslant |\sin(\pi\alpha)|$ for $\alpha \in [-1/2, 1/2]$. Thus,

$$\big| \hat{1}_P(\alpha) \big| \ll \min\left(|P|, \frac{1}{\|\alpha\|} \right).$$

We now use dyadic decomposition to complete the proof. Let

$$A_k = \left\{ \alpha \in \mathbb{T} \colon \frac{2^{k-1}}{|P|} \leqslant \|\alpha\| \leqslant \frac{2^k}{|P|} \right\}$$

for $k = \{1, \ldots, \lceil \log_2(|P|) \rceil\}$ and $A_0 = \{\alpha \in \mathbb{T} \colon \|\alpha\| \leqslant 1/|P|\}$. For $k \geqslant 1$ we have

$$\int_{A_k} \left| \hat{1}_P(\alpha) \right| \mathrm{d}\alpha \leqslant \int_{A_k} \frac{1}{\|\alpha\|} \mathrm{d}\alpha \leqslant \frac{|P| \mathrm{meas}(A_k)}{2^{k-1}} \ll 1.$$

On the other hand,

$$\int_{A_0} \left| \hat{1}_P(\alpha) \right| \mathrm{d}\alpha \ll |A_0||P| \ll 1.$$

Since $\mathbb{T} \subset \bigcup_k A_k$, the claimed bound follows. □

We are now able to deduce Theorem 3.1 from Theorem 1.2.

Proof of Theorem 3.1. Using orthogonality we have

$$\begin{aligned} |S \cap P| - \tfrac{|[N] \cap P||S|}{N} &= \sum_x 1_{P \cap [N]}(x) \Big(1_S(x) - \tfrac{|S|}{N} 1_{[N]}(x) \Big) \\ &= \int_{\mathbb{T}} \hat{1}_{P \cap [N]}(\alpha) \Big(\hat{1}_S - \tfrac{|S|}{N} \hat{1}_{[N]} \Big) (-\alpha) \mathrm{d}\alpha. \end{aligned}$$

The result follows from Lemma 3.2 and the Fourier uniformity obtained in Theorem 1.2. □

4. Equidistribution in Bohr neighbourhoods

One can prove Corollaries 1.7 and 1.11 using Erdős–Turán type arguments, see for instance [18, Chapter 1]. We opt for the following cruder trigonometric approximation, a proof of which can be found in Appendix B.

Lemma 4.1 (Trigonometric approximation)**.** *Let $F : \mathbb{T}^d \to [0,1]$ have Lipschitz constant $K \geqslant 1$ with respect to the metric*

$$\max_j \|\alpha_j - \beta_j\|_{\mathbb{T}}. \tag{4.1}$$

For any $\varepsilon > 0$ there exists a trigonometric polynomial $F_\varepsilon : \mathbb{T}^d \to [0,1]$ with $\|F - F_\varepsilon\|_\infty \leqslant \varepsilon$ and such that

$$F_\varepsilon(\alpha) = \sum_{|m_i| \leqslant M} \hat{F}_\varepsilon(m) e(m \cdot \alpha)$$

with $M \ll K^2 \varepsilon^{-3}$.

Deduction of Corollary 1.7. Let $F : \mathbb{T}^d \to [0,1]$ be a 1-bounded function with Lipschitz constant $K \geqslant 1$ with respect to the metric (4.1). We apply Lemma 4.1 to obtain a trigonometric approximation $F_\varepsilon : \mathbb{T}^d \to [0,1]$, with $\varepsilon > 0$ to be determined.

Expanding F_ε in terms of its Fourier coefficients gives

$$\sum_{x\in S} F_\varepsilon(\alpha x) = \sum_{|m_i|\leqslant M} \hat{F}_\varepsilon(m) \sum_{x\in S} e(m\cdot \alpha x).$$

Approximating $\sum_{x\in S} e(m\cdot\alpha x)$ with $\sum_{x\in[N]} e(m\cdot\alpha x)$, we deduce that there exists an absolute constant C such that

$$\Bigg|\sum_{x\in S} F_\varepsilon(\alpha x) - \tfrac{|S|}{N}\sum_{x\in[N]} F_\varepsilon(\alpha x)\Bigg| \leqslant (CK^2/\varepsilon^{-3})^d \big\|\hat{1}_S - \tfrac{|S|}{N}\hat{1}_N\big\|_\infty.$$

Balancing the above error term with $\varepsilon|S|$, we take

$$\varepsilon := \Big(|S|^{-1}\big(CK^2\big)^d \,\big\|\hat{1}_S - \tfrac{|S|}{N}\hat{1}_N\big\|_\infty\Big)^{\frac{1}{3d+1}}$$

to yield the asymptotic

$$\big|\mathbb{E}_{x\in S}F(\alpha x) - \mathbb{E}_{x\in[N]}F(\alpha x)\big| \ll \left(K^{2d}\Big(\Big|\tfrac{|S|}{N^{1/2}} - 1\Big| + N^{-1/6}\Big)^{1/2}\right)^{\frac{1}{3d+1}},$$

on employing Theorem 1.2. □

Deduction of Corollary 1.11. Let $\varepsilon > 0$ be a small quantity to be determined. Let $F_1 : \mathbb{T} \to [0,1]$ be a piecewise linear function with $F_1(\alpha) = 1$ on $[-\rho,\rho]$, $F_1(\alpha) = 0$ on $\mathbb{T}\setminus[-\rho-\varepsilon,\rho+\varepsilon]$ and with Lipschitz constant at most ε^{-1}. By a telescoping identity, the function $F(\alpha) := F_1(\alpha_1)\cdots F_1(\alpha_d)$ has Lipschitz constant at most $d\varepsilon^{-1}$. Hence by Corollary 1.7 we have

$$\begin{aligned}\mathbb{E}_{x\in S}1_B(x) &\leqslant \mathbb{E}_{x\in S}F(\alpha x)\\ &= \mathbb{E}_{x\in[N]}F(\alpha x) + O\left((d/\varepsilon)^{\frac{2}{3}}\Big(\Big|\tfrac{|S|}{N^{1/2}} - 1\Big| + N^{-1/6}\Big)^{\frac{1}{8d}}\right).\end{aligned}$$

Since F is supported on $B(\alpha,\rho+\varepsilon)$, regularity (Definition 1.9) ensures that

$$\sum_{x\in[N]} F(\alpha x) \leqslant |B| + O\big(d\rho^{-1}\varepsilon N\big).$$

Therefore

$$\mathbb{E}_{x\in S}1_B(x) \leqslant \tfrac{|B|}{N} + O\left(d\rho^{-1}\varepsilon + (d/\varepsilon)^{\frac{2}{3}}\Big(\Big|\tfrac{|S|}{N^{1/2}} - 1\Big| + N^{-1/6}\Big)^{\frac{1}{8d}}\right).$$

Choosing ε to balance error terms, then bounding exponents somewhat crudely, we obtain

$$\mathbb{E}_{x\in S}1_B(x) \leqslant \tfrac{|B|}{N} + O\left(d\rho^{-1}\left(\left|\tfrac{|S|}{N^{1/2}} - 1\right| + N^{-1/6}\right)^{\frac{1}{14d}}\right).$$

The corresponding lower bound is proved analogously. □

5. Partition regularity

Our deduction of Theorem 1.15 from Theorem 1.2 requires two additional results, the first being the following transference principle for colourings, a proof of which can be found in Appendix C.

Lemma 5.1 (Dense model lemma). *Let $0 < \varepsilon \leqslant 1$ and $\nu : [N] \to [0,\infty)$ be such that there exist functions $f_i : [N] \to [0,\infty)$ with $f_1 + \cdots + f_r \leqslant \nu$ and such that for any $g_i : [N] \to [0,\infty)$ satisfying $g_1 + \cdots + g_r \leqslant 1_{[N]}$ there exists i with*

$$\left\|\hat{f}_i - \hat{g}_i\right\|_\infty > \varepsilon N.$$

Then

$$\left\|\hat{\nu} - \hat{1}_{[N]}\right\|_\infty \gg_{\varepsilon,r} N. \tag{5.1}$$

The second additional result underlying Theorem 1.15 is a lower bound on the number of monochromatic solutions to a partition regular equation in an interval.

Lemma 5.2 (Counting monochromatic solutions in an interval of integers)**.** *Let $c_1, \dots, c_s \in \mathbb{Z}\setminus\{0\}$ and suppose that there is a non-empty index set $I \subset [s]$ satisfying $\sum_{i\in I} c_i = 0$. Then for any functions $g_1, \dots, g_r : [N] \to [0,\infty)$ with $1_{[N]} = g_1 + \cdots + g_r$, either $N \ll_{c_1,\dots,c_s,r} 1$ or there exists g_j satisfying*

$$\sum_{c_1x_1+\cdots+c_sx_s=0} g_j(x_1)\cdots g_j(x_s) \gg_{c_1,\dots,c_s,r} N^{s-1}. \tag{5.2}$$

Proof. When each g_i is a characteristic function of a set $C_i \subset [N]$, the result is a special case of Frankl, Graham and Rödl [11, Theorem 1]. In general, for each $x \in [N]$ fix an index $i = i(x)$ such that $g_i(x) \geqslant 1/r$. At least one such index exists by the pigeon-hole principle. On setting $C_i = \{x \in [N] : i(x) = i\}$ we obtain a colouring, and the lower bound (5.2) follows on employing [11, Theorem 1]. □

Proof of Theorem 1.15. Let $S \subset [N]$ be a Sidon set and $S = C_1 \cup \cdots \cup C_r$. Writing $\delta = \delta_{c_1,\dots,c_s,r} > 0$ for the implicit constant in (5.2), suppose that

$$\sum_{c_1x_1+\cdots+c_sx_s=0} 1_{C_j}(x_1)\cdots 1_{C_j}(x_s) \leqslant \tfrac{1}{2}\delta|S|^sN^{-1} \qquad (1 \leqslant j \leqslant r). \tag{5.3}$$

Define $f_i := N|S|^{-1}1_{C_i}$, so that $f_1 + \cdots + f_r = N|S|^{-1}1_S$. Then, by Lemma 5.2, we either have $N \ll_{c_1,\ldots,c_s,r} 1$, or for any functions $g_j \geqslant 0$ with $g_1 + \cdots + g_r = 1_{[N]}$ there exists g_j such that

$$\left|\sum_{c_1x_1+\cdots+c_sx_s=0}[g_j(x_1)\cdots g_j(x_s) - f_j(x_1)\cdots f_j(x_s)]\right| \gg_{c_1,\ldots,c_s,r} N^{s-1}.$$

Applying a telescoping identity to the left hand side, there exists $h_1, \ldots, h_s$ belonging to $\{g_j, f_j, f_j - g_j\}$, exactly one of which is equal to $f_j - g_j$, and such that

$$\left|\sum_{c_1x_1+\cdots+c_sx_s=0}[g_j(x_1)\cdots g_j(x_s) - f_j(x_1)\cdots f_j(x_s)]\right| \ll_s \left|\sum_{c_1x_1+\cdots+c_sx_s=0} h_1(x_1)\cdots h_s(x_s)\right|.$$

By orthogonality and Hölder's inequality

$$\left|\sum_{c_1x_1+\cdots+c_sx_s=0} h_1(x_1)\cdots h_s(x_s)\right| = \left|\int_{\mathbb{T}} \hat{h}_1(c_1\alpha)\cdots\hat{h}_s(c_s\alpha)\mathrm{d}\alpha\right| \leqslant \left\|\hat{f}_j - \hat{g}_j\right\|_\infty \max\left\{\left\|\hat{f}_j\right\|_{s-1}, \left\|\hat{g}_j\right\|_{s-1}\right\}^{s-1}. \tag{5.4}$$

Since $s-1 \geqslant 2$ and $0 \leqslant g_j \leqslant 1_{[N]}$, Parseval's identity gives that

$$\left\|\hat{g}_j\right\|_{s-1}^{s-1} \leqslant N^{s-2}.$$

Since $s-1 \geqslant 4$ and $0 \leqslant f_j \leqslant N|S|^{-1}1_S$, orthogonality and the Sidon property give that

$$\left\|\hat{f}_j\right\|_{s-1}^{s-1} \leqslant N^{s-1}|S|^{-4}\sum_{x-x'=y-y'} 1_S(x)1_S(x')1_S(y)1_S(y') \leqslant 2N^{s-1}|S|^{-2}.$$

Supposing that $|S| \geqslant \frac{1}{100}N^{1/2}$, the latter quantity is $O(N^{s-2})$. We may assume that $|S| \geqslant \frac{1}{100}N^{1/2}$ for otherwise $||S| - N^{1/2}| \gg N^{1/2}$.

From the above deliberations, we conclude that if (5.3) holds, then either $N \ll_{c_1,\ldots,c_s,r} 1$, or $||S| - N^{1/2}| \gg N^{1/2}$, or for any $g_1, \ldots, g_r \geqslant 0$ with $g_1 + \cdots + g_r = 1_{[N]}$ there exists g_j such that

$$\left\|\hat{f}_j - \hat{g}_j\right\|_\infty \gg_{c_1,\ldots,c_s,r} N. \tag{5.5}$$

Henceforth we assume that we are not in the situation that $N \ll_{c_1,\ldots,c_s,r} 1$ or $||S| - N^{1/2}| \gg N^{1/2}$. Let $\eta = \eta(c_1, \ldots, c_r, r) > 0$ denote the implicit constant in (5.5). If it is the case that there exists $g_1, \dots, g_r \geqslant 0$ with $g_1 + \cdots + g_r = 1_{[N]}$ such that for all $1 \leqslant j \leqslant r-1$ we have

$$\left\|\hat{f}_j - \hat{g}_j\right\|_\infty \leqslant \tfrac{\eta}{2r} N,$$

then (5.5) holds with $j = r$, so by the triangle inequality

$$\left\|N|S|^{-1}\hat{1}_S - \hat{1}_{[N]}\right\|_\infty \geqslant \tfrac{1}{2}\eta N \gg_{c_1,\dots,c_s,r} N.$$

Let us show that this conclusion also holds when for any $g_1, \dots, g_r \geqslant 0$ with $g_1 + \cdots + g_r = 1_{[N]}$ there exists $1 \leqslant j \leqslant r-1$ such that

$$\left\|\hat{f}_j - \hat{g}_j\right\|_\infty > \tfrac{\eta}{2r} N.$$

Since

$$\begin{aligned}&\left\{(g_1, \dots, g_r) : g_1 + \cdots + g_r = 1_{[N]} \text{ and } g_i \geqslant 0 \text{ for all } i\right\}\\ &\quad = \left\{(g_1, \dots, g_{r-1}, 1_{[N]} - g_r) : g_1 + \cdots + g_{r-1} \leqslant 1_{[N]}, g_i \geqslant 0 \text{ for all } i\right\},\end{aligned}$$

we may apply the dense model lemma (Lemma 5.1) to conclude that

$$\left\|N|S|^{-1}\hat{1}_S - \hat{1}_{[N]}\right\|_\infty \gg_{c_1,\dots,c_s,r} N.$$

Hence by Theorem 1.2 we have

$$N^{1/4}\Big||S| - N^{1/2}\Big|^{1/2} + N^{5/12} \gg_{c_1,\dots,c_s,r} |S|. \tag{5.6}$$

Again assuming $|S| \geqslant \frac{1}{100}N^{1/2}$ (as we may), (5.6) implies that either $N \ll_{c_1,\dots,c_s,r} 1$ or $\left||S| - N^{1/2}\right| \gg_{c_1,\dots,c_s,r} N^{1/2}$. □

6. Improving Fourier uniformity

In previous sections we have seen that the quality of Fourier uniformity dictates the level of equidistribution of an extremal Sidon set. In this section we give a modified proof of Theorem 1.2 with an increased power saving in the quality of Fourier uniformity. This is accomplished by incorporating an estimate of Cilleruelo [6, Theorem 1.1] on the level of equidistribution in short intervals.

Theorem 6.1 (Cilleruelo)**.** *Let $S \subset [N]$ be a Sidon set and $I \subset [N]$ an interval. Then*

$$\left||S \cap I| - \frac{|I||S|}{N}\right| \ll \left(N^{1/4} + |I|^{1/2}N^{-1/8}\right)\left(1 + \left(1 - \tfrac{|S|}{N^{1/2}}\right)_+^{1/2} N^{1/8}\right),$$

where $x_+ = \max(0, x)$.

Corollary 6.2. *Let $S \subset [N]$ be a Sidon set and $I \subset [N]$ an interval with $|I| \leqslant N^{3/4}$. Then*

$$|S \cap I| \ll N^{1/4} + \left|1 - \tfrac{|S|}{N^{1/2}}\right|^{1/2} N^{3/8}.$$

Using this result we can refine the bounds obtained in Theorem 1.2.

Theorem 6.3. *Let $S \subset [N]$ be a Sidon set. Then*

$$\left\| \hat{1}_S - \tfrac{|S|}{N}\hat{1}_{[N]} \right\|_\infty \ll N^{1/2}\left(\left| 1 - \tfrac{|S|}{N^{1/2}} \right| + N^{-1/4} \right)^{1/2}. \tag{6.1}$$

Proof. The proof is identical to that given for Theorem 1.2, albeit taking $H := N^{3/4}$ and replacing our use of (1.1) in (2.3) with Corollary 6.2. This gives

$$|S \cap I_h| \ll N^{1/4} + \left| 1 - \tfrac{|S|}{N^{1/2}} \right|^{1/2} N^{3/8}.$$

Hence when $i \neq j$ we have

$$\sum_h \mu_H(h) \left| |S|^2 N^{-1} - \sum_x f_i(x) f_j(x+h) \right| \ll |S| \left(N^{-3/4} + \left| 1 - \tfrac{|S|}{N^{1/2}} \right|^{1/2} N^{-5/8} \right).$$

Putting everything together, as in the proof of Theorem 1.2, then gives the desired bound. □

Remark 6.4. This second version of our Fourier uniformity bound solves a quirk of the previous proof, where we took $H = N^{2/3}$, whereas in many results on extremal Sidon sets taking $H = N^{3/4}$ appears naturally.

Of course, this new version cannot be used to improve the bounds on uniform distribution in intervals, since it would give a circular argument. However, it may be applied to improve our bounds on distribution in residue classes.

Appendix A. The size of a Sidon set

In this appendix we establish the well-known bound (1.1), again using van der Corput's variant of the Cauchy–Schwarz inequality. Let $S \subset (n, n+N]$ be a Sidon set and H be a positive integer (to be determined). Applying (2.1) to $\sum_x 1_S(x)$ we obtain

$$|S|^2 \leqslant (N + \lfloor H \rfloor) \sum_h \mu_H(h) \sum_x 1_S(x) 1_S(x+h).$$

Using the defining property of Sidon sets, and the fact that μ_H is a probability measure, we deduce that

$$|S|^2 \leqslant (N + H)\left(\frac{|S|}{H} + 1 \right).$$

By the quadratic formula $x^2 \leqslant bx + c$ only if $x \leqslant (b + \sqrt{b^2 + 4c})/2$, which in turn implies that $x \leqslant b + \sqrt{c}$. Hence

$$|S| \leqslant \sqrt{N+H} + (N+H)H^{-1} \leqslant N^{1/2} + HN^{-1/2} + NH^{-1} + 1.$$

The bound (1.1) follows on taking, say, $H = \lceil N^{3/4} \rceil$.

Appendix B. Trigonometric approximation

Definition B.1. Given an integrable function $F\colon \mathbb{T}^d \to [0,1]$, we define its Fourier transform to be the function $\hat{F}\colon \mathbb{Z}^d \to \mathbb{C}$ given by

$$\hat{F}(m) = \int_{\alpha \in \mathbb{T}^d} F(\alpha) e(-m\cdot\alpha),$$

where $m \cdot \alpha = m_1\alpha_1 + \cdots + m_d\alpha_d$.

Proof of Lemma 4.1. Let $\lambda_M(\alpha) = \lambda_M(\alpha_1)\cdots\lambda_M(\alpha_d)$ denote the following renormalised Fourier transform of the Féjer kernel:

$$\sum_m \left(1 - \frac{|m_1|}{M}\right)_+ \cdots \left(1 - \frac{|m_d|}{M}\right)_+ e(m\cdot\alpha)$$
$$= M^{-d}\Big(1_{[M]^d} * 1_{-[M]^d}\Big)\hat{}\,(\alpha) = M^{-d}\left|\hat{1}_{[M]^d}(\alpha)\right|^2.$$

Set

$$F_M(\alpha) := F * \lambda_M(\alpha) = \int_{\mathbb{T}^d} F(\alpha-\beta)\lambda_M(\beta)\mathrm{d}\beta.$$

One can check that

$$F * \lambda_M(\alpha) = \sum_m \left(1 - \frac{|m_1|}{M}\right)_+ \cdots \left(1 - \frac{|m_d|}{M}\right)_+ \hat{F}(m) e(m\cdot\alpha).$$

We utilise the following three properties of the Féjer kernel.

(a) (Non-negativity) $\lambda_M \geqslant 0$.
(b) (Mass one) $\int_{\mathbb{T}^d} \lambda_M = 1$.
(c) (Quantitative decay) $\lambda_M(\alpha) \leqslant M^{-1}\|\alpha_j\|^{-2}\prod_{i\neq j}\lambda_M(\alpha_i)$.

The first two facts ensure that $0 \leqslant F_M \leqslant 1$, since $0 \leqslant F \leqslant 1$. Let us estimate the error $\|F - F*\lambda_M\|_\infty$. By definition of convolution

$$F(\alpha) - F*\lambda_M(\alpha) = \int_{\mathbb{T}^d} (F(\alpha) - F(\alpha-\beta))\,\lambda_M(\beta)\mathrm{d}\beta.$$

Since the Féjer kernel is non-negative and has integral 1, the Lipschitz continuity of F gives that

$$\left|\int_{\max_i |\beta_i| \leqslant \eta} \big(F(\alpha) - F(\alpha-\beta)\big)\lambda_M(\beta)\mathrm{d}\beta\right| \leqslant K\eta.$$

By the quantitative decay estimate

$$\left|\int_{\max_i |\beta_i| > \eta} \big(F(\alpha) - F(\alpha-\beta)\big)\lambda_M(\beta)\mathrm{d}\beta\right| \leqslant \frac{2}{M\eta^2}.$$

Taking $\eta^3 = 1/(KM)$ then gives

$$\|F - F * \lambda_M\|_\infty \leqslant 3K^{2/3}M^{-1/3}.$$

Setting $M = \lceil 27K^2\varepsilon^{-3}\rceil$ we have $\|F - F * \lambda_M\|_\infty \leqslant \varepsilon$.

We note that we may assume that $\varepsilon \leqslant 1$, so that $M \ll K^2\varepsilon^{-3}$, for otherwise the result is immediate on taking $F_\varepsilon = 0$ and $M = 0$. ☐

Appendix C. A dense model lemma

Lemma C.1 (Separating hyperplane theorem)**.** *Let $K \subset \mathbb{R}^n$ be closed and convex and $v \notin K$. Then there exists $\phi \in \mathbb{R}^n$ such that for all $u \in K$ we have $v \cdot \phi > u \cdot \phi$.*

Proof. See `https://en.wikipedia.org/wiki/Hyperplane_separation_theorem`. ☐

Lemma C.2. *For $f, \phi : [N] \to \mathbb{R}$ write*

$$\|f\| := \left\|\hat{f}\right\|_\infty \quad \textit{and} \quad \|\phi\|^* := \sup_{\|f\|\leqslant 1}\left|\sum_x f(x)\phi(x)\right|.$$

Then for any $f, \phi, \psi : [N] \to \mathbb{R}$ we have

- $|\sum_x f(x)\phi(x)| \leqslant \|f\| \, \|\phi\|^*$;
- $\|\phi\psi\|^* \leqslant \|\phi\|^* \, \|\psi\|^*$;
- $\|\phi\|_\infty \leqslant \|\phi\|^*$.

Proof. The first inequality follows from the definition of $\|\cdot\|^*$.

Let e_α denote the map $x \mapsto e(\alpha x)$. Then $\|\cdot\|$ is invariant under multiplication by e_α, so for any $f, \phi : [N] \to \mathbb{R}$ the first inequality gives that

$$\left|\widehat{f\phi}(\alpha)\right| = \left|\sum_x f(x)e_\alpha(x)\phi(x)\right| \leqslant \|fe_\alpha\| \, \|\phi\|^* = \|f\| \, \|\phi\|^* .$$

Hence

$$\left|\sum_x f(x)\phi(x)\psi(x)\right| \leqslant \|f\phi\| \, \|\psi\|^* \leqslant \|f\| \, \|\phi\|^* \, \|\psi\|^* .$$

The second inequality follows.

Suppose that $\|\phi\|_\infty = 1$, so that $|\phi(x)| = 1$ for some $x \in [N]$. Notice that the function $f := 1_{\{x\}}$ has Fourier transform bounded in magnitude by 1. Therefore

$$\|\phi\|^* \geqslant \left|\sum_y f(y)\phi(y)\right| = 1 = \|\phi\|_\infty .$$

The third inequality then follows on renormalising. ☐

Proof of Lemma 5.1. We closely follow Conlon and Gowers [5, Lemma 2.6]. Notice that

$$(1+\tfrac{\varepsilon}{2})^{-1}(f_1,\dots,f_r)$$

is not a member of the closed convex set

$$\left\{(g_1+h_1,\dots,g_r+h_r) : g_i \geqslant 0,\ g_1+\dots+g_r \leqslant 1_{[N]},\ \left\|\hat{h}_i\right\|_\infty \leqslant \tfrac14\varepsilon N\right\}.$$

Hence by the separating hyperplane theorem, there exists $(\phi_1,\dots,\phi_r)$ such that for any $g_i \geqslant 0$ with $g_1+\dots+g_r \leqslant 1_{[N]}$ and $\left\|\hat{h}_i\right\|_\infty \leqslant \frac14\varepsilon N$ we have

$$\left(1+\tfrac12\varepsilon\right)^{-1}\sum_i\sum_x f_i(x)\phi_i(x) > \sum_i\sum_x g_i(x)\phi_i(x) + \sum_i\sum_x h_i(x)\phi_i(x). \tag{C.1}$$

Taking all g_i and h_i zero shows that the left-hand side of (C.1) is positive, so we may renormalise $(\phi_1,\dots,\phi_r)$ to give

$$\sum_i\sum_x f_i(x)\phi_i(x) = (1+\tfrac12\varepsilon)N \tag{C.2}$$

and for all $g_i \geqslant 0$ with $g_1+\dots+g_r \leqslant 1_{[N]}$ and $\left\|\hat{h}_i\right\|_\infty \leqslant \frac14\varepsilon N$ we have

$$\sum_i\sum_x g_i(x)\phi_i(x) + \sum_i\sum_x h_i(x)\phi_i(x) < N. \tag{C.3}$$

Notice that

$$\begin{aligned}\sum_i\sum_x f_i(x)\phi_i(x) &\leqslant \sum_i\sum_x f_i(x)\max\{\phi_1(x),\dots,\phi_r(x),0\}\\ &\leqslant \sum_x \nu(x)\max\{\phi_1(x),\dots,\phi_r(x),0\}.\end{aligned} \tag{C.4}$$

For each $x\in[N]$ fix $i(x)\in[r]$ such that

$$\max\{\phi_1(x),\dots,\phi_r(x)\} = \phi_{i(x)}(x).$$

Define

$$g_i(x) := \begin{cases} 1 & \text{if } i = i(x) \text{ and } \phi_i(x)\geqslant 0,\\ 0 & \text{otherwise.}\end{cases}$$

By substituting this function into (C.3) with $h_i = 0$, and using both (C.2) and (C.4), we deduce that

$$\sum_x \left[\nu(x) - 1_{[N]}(x)\right]\max\{\phi_1(x),\dots,\phi_r(x),0\} > \tfrac12\varepsilon N.$$

Using the notation and content of Lemma C.2, the inequality gives that

$$\|\phi_i\|_\infty \leqslant \|\phi_i\|^* \leqslant 4/\varepsilon. \tag{C.5}$$

By the Stone–Weierstrass theorem[8], there exists a polynomial P_ε with degree and coefficients of size $O_{\varepsilon,r}(1)$ such that for all $|x_i| \leqslant 4/\varepsilon$ we have

$$|\max\{x_1, \dots, x_r, 0\} - P_\varepsilon(x_1, \dots, x_s)| \leqslant \varepsilon/100.$$

Notice that we may assume that $\sum_x \nu(x) \leqslant 2N$, otherwise we are done. Hence

$$\sum_x \left[\nu(x) - 1_{[N]}(x)\right] P_\varepsilon(\phi_1(x), \dots, \phi_r(x)) > \tfrac{1}{4}\varepsilon N.$$

Expanding the polynomial, and applying the pigeon-hole principle, there exist $\psi_1, \dots, \psi_R \in \{\phi_1, \dots \phi_r\}$ with $R \ll_{\varepsilon,r} 1$ such that

$$\left|\sum_x \left[\nu(x) - 1_{[N]}(x)\right] \psi_1(x) \cdots \psi_R(x)\right| \gg_{\varepsilon,r} N.$$

Recalling (C.5) and Lemma C.2 we have $\|\psi_1 \cdots \psi_R\|^* \ll_{\varepsilon,r} 1$. Hence, again applying the first inequality in Lemma C.2, we deduce (5.1). □

References

[1] R. C. Baker, G. Harman, and J. Pintz, *The difference between consecutive primes. II.* Proc. London Math. Soc. **83** (2001), 532–562.

[2] J. Bourgain, *On triples in arithmetic progression.* Geom. Funct. Anal. **9** (1999), 968–984.

[3] D. Conlon, J. Fox, B. Sudakov, and Y. Zhao, *The regularity method for graphs with few 4-cycles.* J. London Math. Soc. **104** (2021), 2376–2401.

[4] D. Conlon, J. Fox, B. Sudakov, and Y. Zhao, *Which graphs can be counted in C_4-free graphs?* ArXiv e-prints, 2021.

[5] D. Conlon and W. T. Gowers, *Combinatorial theorems in sparse random sets.* Ann. of Math., **184** (2016), 367–454.

[6] J. Cilleruelo, *Gaps in dense Sidon sets.* Integers **0** (2000), A11.

[7] S. Eberhard, *The apparent structure of dense Sidon sets.* Talk given at CANT (2021), `https://youtu.be/s4ItIkkUvF4`

[8] S. Eberhard and F. Manners, *The apparent structure of dense Sidon sets.* ArXiv e-prints (2021).

[9] P. Erdős and R. Freud, *On sums of a Sidon-sequence.* J. Number Theory **38** (1991), 196–205.

[10] P. Erdős and P. Turán, *On a problem of Sidon in additive number theory, and on some related problems.* J. London Math. Soc. **16** (1941), 212–215.

[11] P. Frankl, R. L. Graham, and V. Rödl, *Quantitative theorems for regular systems of equations.* J. Combin. Theory Ser. A **47** (1988), 246–261.

[12] A. Forey and E. Kowalski, *Algebraic curves in their Jacobian are Sidon sets.* ArXiv e-prints (2021).

[13] W. T. Gowers, *What are dense Sidon subsets of $\{1, 2, \dots, n\}$ like?* Blogpost available at `https://bit.ly/3xHq2NM` (2012).

[14] R. L. Graham, B. L. Rothschild, and J. H. Spencer, *Ramsey theory.* John Wiley & Sons, 1990.

[15] P. Keevash, B. Sudakov, and J. Verstraëte, *On a conjecture of Erdős and Simonovits: Even cycles.* Combinatorica **33** (2013), 699–732.

[16] M. N. Kolountzakis, *On the uniform distribution in residue classes of dense sets of integers with distinct sums.* J. Number Theory **76** (1999), 147–153.

[8] `https://en.wikipedia.org/wiki/Stone-Weierstrass_theorem`

[17] B. Lindström, *Well distribution of Sidon sets in residue classes.* J. Number Theory **69** (1998), 197–200.
[18] H. L. Montgomery, *Ten lectures on the interface between analytic number theory and harmonic analysis.* American Mathematical Society, 1994.
[19] K. O'Bryant, *A complete annotated bibliography of work related to Sidon sequences.* Electron. J. Combin. **DS11** (2004).
[20] S. Prendiville, *Solving equations in dense Sidon sets.* Math. Proc. Cambridge Phil. Soc. **173** (2022), 25–34.
[21] K. F. Roth, *On certain sets of integers. II.* J. London Math. Soc. **29** (1954), 20–26.
[22] J. Singer, *A theorem in finite projective geometry and some applications to number theory.* Trans. Amer. Math. Soc. **43** (1938), 377–385.
[23] T. Tao and V. Vu, *Additive Combinatorics.* Cambridge University Press, 2006.

Departament de Matemàtiques
Universitat Politècnica de Catalunya
08028, Barcelona, Spain
E-mail : `miquel.ortega.sanchez-colomer@upc.edu`

Department of Mathematics and Statistics
Lancaster University LA1 4YF
UK
E-mail : `s.prendiville@lancaster.ac.uk`