Canad. J. Math. Vol. 71 (3), 2019 pp. 579–605
http://dx.doi.org/10.4153/CJM-2017-036-1

Monochromatic Solutions to $x + y = z^2$

Ben Joseph Green and Sofia Lindqvist

Abstract. Suppose that $\mathbf{N}$ is 2-coloured. Then there are infinitely many monochromatic solutions to $x + y = z^2$. On the other hand, there is a 3-colouring of $\mathbf{N}$ with only finitely many monochromatic solutions to this equation.

1 Introduction

In this paper we will be concerned with the Ramsey theory of the equation $x + y = z^2$. It was shown relatively recently by Csikvári, Gyarmati, and Sárközy [7] that this equation is not partition regular. Indeed, a 16-colouring of $\mathbf{N}$ is exhibited with no monochromatic solutions to $x + y = z^2$ other than the trivial one $x = y = z = 2$. There remains the question of whether the 16 here is optimal. Our main theorem completely answers this question.

Theorem 1.1 *There is a 3-colouring of $\mathbf{N}$ with no monochromatic solution to $x+y = z^2$ other than the trivial one. On the other hand, every 2-colouring of $\mathbf{N}$ has infinitely many monochromatic solutions to $x + y = z^2$.*

The proof of the first statement is rather simple. It is given in Section 2. By contrast, the proof that every 2-colouring has infinitely many monochromatic solutions to $x + y = z^2$ is complicated and involves a surprisingly large number of tools from additive combinatorics and number theory. It occupies the remaining sections of the paper. We outline the argument now.

If $\mathbf{N} = V \cup W$, then let us assume that there are infinitely many N such that $|V \cap [N, 2N)| \geqslant N/2$. If this is not the case, then a corresponding statement holds for W and we can switch the roles of V and W in what follows. Suppose that there are no solutions to $x + y = z^2$ in either V or W. By a fairly elaborate sequence of arguments involving the arithmetic regularity lemma as well as certain Fourier-analytic and diophantine arguments, as well as a deep result of Lagarias, Odlyzko, and Shearer, we use this to show that for some $q \in N$ and $c > 0$ the set W contains the progression $\mathsf{P}([1, 1 + c]; M, q) := \{n \in \mathbf{Z} : M \leqslant n \leqslant (1 + c)M, n \equiv 0(\mathrm{mod}\ q)\}$ for infinitely many integers M. The details of these arguments can be found in Sections 4 and 5, certain preliminary results having been assembled in Section 3. The proof is concluded in

Received by the editors August 29, 2017.
Published electronically January 17, 2018.
This work was partially supported by a grant from the Simons Foundation (award number 376201 to Ben Green) , and the first author is supported by ERC Advanced Grant AAS 279438. We thank both organisations for their support.
AMS subject classification: 11B75, 05D10.
Keywords: additive combinatorics, Ramsey theory.

Section 7 by performing an iterative argument to get a collection of further progressions inside W, eventually showing that all sufficiently large multiples of q lie in W. An important ingredient here is a result concerning gaps between sums of two squares with certain constraints, proved in Section 6.

The fact that all sufficiently large multiples of q lie in W leads immediately to a contradiction, since W then obviously contains infinitely many solutions to $x + y = z^2$.

We make heavy use of smooth cutoff functions in the latter half of the paper. The properties and constructions of these are recalled in Appendix A.

We remark that our arguments in fact give the following, logically stronger, result: if N is large, then any 2-colouring of $[N, CN^8]$ has a monochromatic solution to $x + y = z^2$. Here, C is an absolute constant that could be computed in principle, but that would be astronomically large due to the application of the regularity lemma. We have found it easier to write the paper in such a way that this result does not immediately follow from our arguments as written, and we leave the interested reader to verify this statement.

Let us remark on the nice work of Khalfallah and Szemerédi [9], which, despite its rather similar title, concerns a somewhat different problem. They show that any finite colouring of $\mathbf{N}$ contains a solution to $x + y = z^2$ with x and y having the same colour (but not necessarily z).

We also remark that for the modular version of the problem the answer is very different. Indeed, the second author [12] has shown that if $p > p_0(k)$ is a prime and if $\mathbf{Z}/p\mathbf{Z}$ is k-coloured, then there are $\gg_k p^2$ monochromatic solutions to $x + y = z^2$.

Notation We collect here some notation used in the paper. Most of it is standard. If X is a finite set, then $\mathbf{E}_{x\in X}$ means $1/|X| \sum_{x\in X}$. For $t \in \mathbf{R}$, we write $e(t) := e^{2\pi i t}$. We write $\mathbf{T} = \mathbf{R}/\mathbf{Z}$ and $\mathbf{T}^d = (\mathbf{R}/\mathbf{Z})^d$. We define a "norm" $\|\cdot\|_{\mathbf{T}^d}\colon \mathbf{T}^d \to [0, \frac{1}{2}]$ by defining $\|x\|_{\mathbf{T}^d} = \|\widetilde{x}\|_{\ell^\infty(\mathbf{R}^d)}$, where $\widetilde{x}$ is the unique element of $(-\frac{1}{2}, \frac{1}{2}]^d$, which projects to x under the natural homomorphism from $\mathbf{R}^d$ to $\mathbf{T}^d$. The notation $X = O(Y)$ and $X \ll Y$ both mean that $X \leqslant CY$ for some constant C. Unless dependence on other parameters is indicated explicitly (for example $X \ll_\varepsilon Y$), C will be an absolute constant.

The notation $\widehat{f}$ always denotes Fourier transform. At various points in the paper, f may be a function on $\mathbf{Z}$, $\mathbf{R}$ or $\mathbf{T}^d$. The definitions we are using are recalled in the text when there is any danger of confusion.

It is convenient to introduce a piece of notation that is less standard, but very useful. If $\Lambda \subset \mathbf{N}$ is a set of integers, then we write $\sqrt{\Lambda} := \{n \in \mathbf{N} : n^2 \in \Lambda\}$ (this is not the same as $\{\sqrt{n} : n \in \Lambda\}$).If $A \subset \mathbf{N}$ is a set, we write $2A = A + A := \{a + a' : a, a' \in A\}$. We will sometimes use notation such as $2\sqrt{2A}$, which means $\sqrt{A + A} + \sqrt{A + A}$.

Finally, as hinted above, when $I \subset \mathbf{R}$ is a closed interval we write $\mathsf{P}(I; N, q) := \{n \in \mathbf{Z} : \frac{n}{N} \in I, q|n\}$.

2 A 3-colouring

In this short section we establish the easy part of Theorem 1.1. That is, we exhibit a 3-colouring of $\mathbf{N}$ for which the only monochromatic solution to $x + y = z^2$ is the trivial solution $x = y = z = 2$. We colour all the points in each dyadic block

$$A_i = \{n \in \mathbf{N} : 2^i \leqslant n < 2^{i+1}\}, \quad i = 0, 1, 2, \dots,$$

in one colour c_i. We assign c_0, c_1, c_2 to be distinct, and then assign the colours c_i, $i \geqslant 3$, inductively in such a way that $c_i \notin \{c_{\lfloor i/2 \rfloor}, c_{\lfloor i/2 \rfloor + 1}\}$. Note that this is possible, since $\lfloor i/2 \rfloor + 1 < i$ for $i \geqslant 3$.

Assume now that $x, y, z \in \mathbf{N}$ have the same colour and that $x + y = z^2$. Without loss of generality we can assume that $x \leqslant y$. Let $i \in \{0, 1, 2, \dots\}$ be such that $y \in A_i$. Then $2^i < x + y < 2^{i+2}$, and hence $2^{i/2} < z < 2^{(i+2)/2}$. Since $i/2 \geqslant \lfloor i/2 \rfloor$ and $(i+2)/2 \leqslant \lfloor i/2 \rfloor + 2$, it follows that $z \in A_{\lfloor i/2 \rfloor} \cup A_{\lfloor i/2 \rfloor + 1}$. By construction, the only way that such a z can have the same colour as y is if $i \in \{0, 1, 2\}$, in which case $x \leqslant y < 8$, and so $z = 2$ or 3. An easy case check confirms that $x = y = z = 2$.

3 Results from the Literature

The rest of the paper is devoted to the harder part of Theorem 1.1. In this section we assemble some basic ingredients from the literature.

We will need a version of Weyl's inequality, which gives a bound for exponential sums $\sum_{n \leqslant N} e(p(n))$ with $p \colon \mathbf{N} \to \mathbf{R}$ a polynomial. The usual proof of Weyl's inequality leads to a factor of $N^{o(1)}$ that renders the result worse than trivial in certain circumstances (the "major arcs"). This is of no consequence in typical applications, which concern minor arc estimates in Waring's problem. Here, however, it is important to have an "ε-free" result. Such results are well known to experts, but it is hard to locate a convenient reference. Wooley [16] discusses the pure power case (that is, sums of the form $\sum_{n \leqslant N} e(\alpha n^k)$), and it is likely that the same methods apply in greater generality, though the verification of this would involve a foray into the inner workings of [15, Chapter 4].

A self-contained source for the purposes of this paper is [6, Lemma 4.4] (described in that paper as a "reformulation" of Weyl's inequality, a slightly inaccurate statement). Here is the statement.

Proposition 3.1 *Let $k \in \mathbf{N}$. Then there is a constant C_k such that the following is true. Let $0 < \delta < 1/2$. Let $g \colon \mathbf{Z} \to \mathbf{R}$ be a polynomial of degree k with leading coefficient α_k (that is, $g(n) = \alpha_k n^k + \dots$). Suppose that $|\mathbf{E}_{n \in I} e(g(n))| \geqslant \delta$, where $I \subset \mathbf{Z}$ is a discrete interval. Then there is some $q \in \mathbf{N}$, $q \leqslant \delta^{-C_k}$, such that $\|q\alpha_k\|_{\mathbf{R}/\mathbf{Z}} \leqslant \delta^{-C_k} |I|^{-k}$.*

We will need this result in the cases where $k = 2$ and $k = 4$. The proof in the latter case is essentially as hard as that of the general case. We remark that in [6, Lemma 4.4] the result is stated with $I = [N]$, but the general case follows trivially from this by translation (which does not affect the leading coefficient α_k).

The following definition is relevant to much of the paper.

Definition 3.2 Suppose that $\theta \in \mathbf{R}^d$. Let $N \geqslant 1$ be an integer and let $A > 0$ be some real parameter. We say that θ is (A, N)-irrational if whenever $\mathbf{r} \in \mathbf{Z}^d \setminus \{0\}$ and $\|\mathbf{r}\|_1 \leqslant A$, we have $\|\mathbf{r} \cdot \theta\|_{\mathbf{T}} \geqslant A/N$.

https://doi.org/10.4153/CJM-2017-036-1 Published online by Cambridge University Press

We record a corollary of Proposition 3.1, phrased in the language of this definition. This corollary is the variant of Weyl's inequality that we have found to be most useful in this paper.

Corollary 3.3 *Let $k, N \in \mathbf{N}$. Suppose that $I \subset \mathbf{Z}$ is a (discrete) interval of length $\leqslant N^{1/k}$. Suppose that $\theta \in \mathbf{R}^d$ is (A, N)-irrational, and suppose that $\mathbf{r} \in \mathbf{Z}^d \smallsetminus \{0\}$. Then*

$$\Big|\sum_{n\in I} e(\mathbf{r}\cdot\theta n^k + \cdots)\Big| \leqslant N^{1/k}\|\mathbf{r}\|_1 A^{-1/C_k}.$$

Here, $\cdots$ denotes polynomial terms in n of degree $k-1$ or lower, and the estimate is uniform in the choice of these terms.

Proof Suppose that the sum is $\geqslant \delta|I|$. Then, by Proposition 3.1 there is some $q \in \mathbf{N}$, $q \leqslant \delta^{-C_k}$, such that $\|q\mathbf{r}\cdot\theta\|_{\mathbf{R}/\mathbf{Z}} \leqslant \delta^{-C_k}|I|^{-k}$. Since θ is (A, N)-irrational, we have either (1) $q\|\mathbf{r}\|_1 \geqslant A$ or (2) $\delta^{-C_k}|I|^{-k} \geqslant A/N$. In case (1), the bound on q implies that $\delta^{-C_k}\|\mathbf{r}\|_1 \geqslant A$. In case (2), we have $\delta^{-C_k} \geqslant A$. Hence, in either case we have $\delta^{-C_k}\|\mathbf{r}\|_1 \geqslant A$, and hence $\delta \leqslant (\|\mathbf{r}\|_1/A)^{1/C_k}$. The result follows (in fact, with $\|\mathbf{r}\|_1$ replaced by the smaller quantity $\|\mathbf{r}\|_1^{1/C_k}$). ∎

Turning to a different type of ingredient of the paper, we require the following estimate.

Proposition 3.4 *Let $S \subset \{1, \dots, N\}$ be any set of squares. For $t \in \mathbf{R}/\mathbf{Z}$, write $\widehat{1_S}(t) := \sum_{n\in S} e(tn)$. Then $\int_0^1 |\widehat{1_S}(t)|^6 dt \ll N^2$.*

Proof It is easy to see that the integral is $\sum_{x\leqslant 3N} r_{3,S}(x)^2$, where $r_{3,S}(x)$ is the number of ways of writing x as $n_1 + n_2 + n_3$ with $n_1, n_2, n_3 \in S$. This quantity is obviously largest when S is the set of all squares $\leqslant N$. In this case, the stated bound is a well-known consequence of the Hardy–Littlewood method. ∎

Remark Using more advanced methods of harmonic analysis (related to the Tomas–Stein restriction theorem) one can show a bound $\int_0^1 |\widehat{1_S}(t)|^q \ll_q N^{q/2-1}$ for any $q > 4$.

Finally, we will also use the following result of Lagarias, Odlyzko, and Shearer [10].

Proposition 3.5 *Suppose that $S \subset \mathbf{Z}/q\mathbf{Z}$, where q is a positive integer, and that $|S| > \frac{11}{32}q$. Then $S + S$ contains a quadratic residue modulo q.*

Remarks The $\frac{11}{32}$ in this theorem is sharp. For our purposes, $\frac{11}{32}$ could be replaced by any constant less than $\frac{1}{2}$. A simpler proof of such a statement could probably be extracted from [10] or the companion paper [11], but we do not know of any argument that could be described as in any way routine.

Instead of the result of Lagarias, Odlyzko, and Shearer, it would suffice to have the following statement: there is some $\eta_k > 0$ such that if $(1-\eta_k)q$ of the elements of $\mathbf{Z}/q\mathbf{Z}$ are k-coloured then there are x, y of the same colour with $x + y$ a square. We believe that such a statement can be established relatively painlessly using a simplified

https://doi.org/10.4153/CJM-2017-036-1 Published online by Cambridge University Press

version of the arguments of Khalfallah and Szemerédi [9]. The second author provides an account of this in an unpublished note [13, Theorem 1.2].

4 Capturing Most of the Squares in a Bohr Set

This section contains the technical heart of the paper. Our aim is to prove the following result. Here, and in what follows, $\mathfrak{S}(b,q)$ denotes the number of solutions to $x^2 \equiv b (\bmod\ q)$ with $x \in \mathbf{Z}/q\mathbf{Z}$.

Proposition 4.1 *Let $\eta > 0$, and let $\Omega\colon \mathbf{N}^3 \to \mathbf{N}$ be a function (which may depend on η), nondecreasing in each variable. Suppose that $N > N_0(\Omega,\eta)$ is sufficiently large, and let $A \subset [N, 2N]$ be a set of size at least $N/2$. Then there are $q, d = O_{\eta,\Omega}(1)$, $\varepsilon \gg_{\eta,\Omega} 1$, $b \in \mathbf{Z}/q\mathbf{Z}$, $x \in [2,4]$, and $\theta, z \in \mathbf{R}^d$ such that*

(i) *b is a quadratic residue modulo q;*

(ii) *θ is $(\Omega(q,d,1/\varepsilon), N)$-irrational;*

(iii) *$A + A$ contains all but at most $\eta\mathfrak{S}(b,q)(2\varepsilon)^{d+1}q^{-1}N^{1/2}$ of the squares in the set $\{n \in \mathbf{N} : n \equiv b(\bmod\ q), |\frac{n}{N} - x|, \|\theta n - z\|_{\mathbf{T}^d} \leqslant \varepsilon\}$.*

Remarks The assumption $|A| \geqslant N/2$ could be weakened to $|A| \geqslant cN$ for any $c > 11/32$, using essentially the same proof. We do not record this explicitly as Proposition 4.1 seems unlikely to be of independent interest. In our applications, η will be an absolute constant that could be specified explicitly if desired ($\eta = 10^{-10}$ should certainly be admissible).

The key tool in the proof of Proposition 4.1 will be the arithmetic regularity lemma, introduced in [4]. The formulation we use here, in a more general guise, is the main result of [5]. That paper is long and quite difficult, but only Sections 1 and 2 of it are relevant to us. Furthermore, that paper establishes a regularity lemma for the Gowers U^{s+1}-norm for general s, whereas we only need the case $s = 1$. This means that the notion of a nilsequence, beyond the abelian case, is not relevant here. A complete, self-contained proof of the arithmetic regularity lemma in the form we need it here can be written up in less than 10 pages. Conveniently, such a writeup has been provided by Eberhard [2].

Here is the arithmetic regularity lemma in the form in which we will need it.

Proposition 4.2 *Suppose we are given $\delta > 0$ and an increasing function $\mathcal{F}\colon \mathbf{N} \to \mathbf{R}_+$. Then there exists $M_{\max} \ll_{\delta,\mathcal{F}} 1$ such that for any function $f\colon [N,\dots,2N) \to [0,1]$ there is an $M \leqslant M_{\max}$ and a decomposition $f = f_{\mathrm{tor}} + f_{\mathrm{sml}} + f_{\mathrm{unf}}$ into functions taking values in $[-1,1]$, where $\sum_{N \leqslant n < 2N} |f_{\mathrm{sml}}(n)| \leqslant \delta N$, $\|\widehat{f_{\mathrm{unf}}}\|_\infty \leqslant N/\mathcal{F}(M)$ and $f_{\mathrm{tor}}(n) = F(n(\bmod q), n/N, \theta n)$ for some $q, d \leqslant M$ and some function $F\colon \mathbf{Z}/q\mathbf{Z} \times [1,2] \times \mathbf{T}^d \to [0,1]$ with Lipschitz constant at most M. Furthermore, θ can be taken to be $(\mathcal{F}(M), N)$-irrational.*

We remark that in the works previously cited the function f_{unf} was controlled in terms of the Gowers U^2-norm, rather than in terms of the supremum norm of the

https://doi.org/10.4153/CJM-2017-036-1 Published online by Cambridge University Press

Fourier transform, defined by

$$\widehat{f_{\mathrm{unf}}}(t) := \sum_{N \leqslant n < 2N} f_{\mathrm{unf}}(n) e(-tn),$$

where $e(x) = e^{2\pi i x}$. However, it is well known (and easy to prove) that for bounded functions, these norms are essentially equivalent.

Moreover, f_{sml} is traditionally controlled in the ℓ^2-norm, rather than the ℓ^1-norm as we have here. However, since f_{sml} is bounded by 1, these two norms are equivalent too. Thus, Proposition 4.2 is equivalent to the arithmetic regularity lemma as usually stated.

Let us now begin the proof of Proposition 4.1 in earnest. Apply Proposition 4.2 with $f = 1_A$, $\delta < \eta$ some small constant ($\delta = 10^{-100}$ would be permissible), and the function $\mathcal{F}$ to be specified later (it will depend on Ω and η). This gives integers $q, d \leqslant M$, $\theta \in \mathbf{R}^d$, and $F \colon \mathbf{Z}/q\mathbf{Z} \times [1,2] \times \mathbf{T}^d \to [0,1]$ and a decomposition

$$1_A = f_{\mathrm{tor}} + f_{\mathrm{sml}} + f_{\mathrm{unf}}$$

with the properties described in the statement of Proposition 4.2 just given.

Lemma 4.3 *Suppose that δ is sufficiently small and that $\mathcal{F}$ grows sufficiently rapidly. Then $\int F d\mu > \frac{9}{20}$, where μ denotes the natural*[1] *measure on $\mathbf{Z}/q\mathbf{Z} \times \mathbf{R} \times \mathbf{T}^d$.*

Remark Here, $\frac{9}{20}$ is simply a convenient fraction less than $\frac{1}{2}$. In fact, $\int F d\mu$ can be made as close to $\frac{1}{2}$ as one wishes by reducing δ and increasing $\mathcal{F}(M)$.

Proof We begin by noting that, by assumption,

$$\mathbf{E}_{N \leqslant n < 2N} 1_A(n) \geqslant \frac{1}{2}. \tag{4.1}$$

If $\delta < \frac{1}{100}$, then

$$|\mathbf{E}_{N \leqslant n < 2N} f_{\mathrm{sml}}(n)| < \frac{1}{100}. \tag{4.2}$$

Also, introducing a smooth majorant ψ for $[N, 2N)$ with $\psi(n) = 1$ for $N \leqslant n < 2N$, we have

$$\begin{aligned} |\mathbf{E}_{N \leqslant n < 2N} f_{\mathrm{unf}}(n)| &= \Big|\frac{1}{N} \sum_n \psi(n) f_{\mathrm{unf}}(n)\Big| \\ &= \Big|\frac{1}{N} \int_0^1 \widehat{\psi}(t) \widehat{f_{\mathrm{unf}}}(t) dt\Big| \leqslant \frac{\|\widehat{\psi}\|_1}{\mathcal{F}(M)}. \end{aligned}$$

With an appropriate choice of ψ (see Lemma A.1 for details) we have $\|\widehat{\psi}\|_1 = O(1)$, and so if $\mathcal{F}(M)$ is sufficiently large, it follows that

$$|\mathbf{E}_{N \leqslant n < 2N} f_{\mathrm{unf}}(n)| < \frac{1}{100}. \tag{4.3}$$

[1] The product of the uniform probability measure on $\mathbf{Z}/q\mathbf{Z}$, Lebesgue measure on $\mathbf{R}$ and normalised Lebesgue measure on $\mathbf{T}^d$.

https://doi.org/10.4153/CJM-2017-036-1 Published online by Cambridge University Press

We also have

$$\mathbf{E}_{N\leqslant n<2N} f_{\mathrm{tor}}(n) = \mathbf{E}_{N\leqslant n<2N} F\Big(n(\mathrm{mod}\ q), \frac{n}{N}, \theta n\Big).$$

However, it was proven[2] in [3, Lemma A.4] that, if $\mathcal{F}$ grows sufficiently rapidly and if N is big enough,

$$|\mathbf{E}_{N\leqslant n<2N} F(n(\mathrm{mod}\ q), \frac{n}{N}, \theta n) - \int F d\mu| < \frac{1}{100}. \tag{4.4}$$

Combining (4.1)–(4.4) concludes the proof. ■

Now let $U \subset \mathbf{Z}/q\mathbf{Z}$ be the set of all $u \in \mathbf{Z}/q\mathbf{Z}$ for which

$$\int_1^2 \int_{\mathbf{T}^d} F(u, x, z) dz dx \geqslant \frac{1}{20} \tag{4.5}$$

and for which

$$\sum_{\substack{N\leqslant n<2N \\ n\equiv u(\mathrm{mod}\, q)}} |f_{\mathrm{sml}}(n)| \leqslant \frac{20\delta}{q} N. \tag{4.6}$$

One should think, informally, of these being the residue classes $(\mathrm{mod}\ q)$ on which A has "significant mass".

Lemma 4.4 *Suppose that δ is sufficiently small and that $\mathcal{F}$ grows sufficiently rapidly. There are elements $u, u' \in U$ such that $u + u'$ is a quadratic residue modulo q.*

Proof Let $U_1 \subset \mathbf{Z}/q\mathbf{Z}$ be the set of all u for which (4.5) fails, and let U_2 be the set of all u for which (4.6) fails. Since $\sum_{N\leqslant n<2N} |f_{\mathrm{sml}}(n)| \leqslant \delta N$, we have $|U_2| \leqslant \frac{q}{20}$.

Furthermore, by Lemma 4.3 we have

$$\frac{9}{20} < \int F d\mu = \frac{1}{q} \sum_{u\in\mathbf{Z}/q\mathbf{Z}} \int_1^2 \int_{\mathbf{T}^d} F(u, x, z) dz dx \leqslant \frac{1}{20} + \frac{1}{q} |(\mathbf{Z}/q\mathbf{Z}) \setminus U_1|.$$

It follows that

$$|U| \geqslant |(\mathbf{Z}/q\mathbf{Z}) \setminus U_1| - |U_2| \geqslant \Big(\frac{9}{20} - \frac{1}{20} - \frac{1}{20}\Big) q > \frac{11q}{32}.$$

The result now follows from Proposition 3.5. ■

Henceforth, we will fix two residue classes $u, u' \in U$ for which $u + u'$ is a quadratic residue modulo q. Define parameters $\varepsilon > \varepsilon' > 0$ by

$$\varepsilon := \frac{\delta}{M} \tag{4.7}$$

and

$$\varepsilon' := \frac{\delta}{dq} (2\varepsilon)^{d+1}. \tag{4.8}$$

[2] This is not an especially difficult argument; roughly, one approximates F by a function with finite Fourier support, then uses the irrationality of θ in estimating the resulting exponential sums.

https://doi.org/10.4153/CJM-2017-036-1 Published online by Cambridge University Press

Note that since $q, d \leqslant M$, we have

$$\varepsilon' \geqslant \frac{\delta}{M^2}\Big(\frac{2\delta}{M}\Big)^{M+1} \gg_{\delta,M} 1. \tag{4.9}$$

(The precise form of this bound is unimportant; what matters is that there is a lower bound depending only on δ and M.)

For $x, x' \in [1,2]$ and $z, z' \in \mathbf{T}^d$, define

$$E_{x,z} := \sum_{\substack{N \leqslant n < 2N \\ n \equiv u (\operatorname{mod} q) \\ |\frac{n}{N} - x| \leqslant \varepsilon \\ \|\theta n - z\|_{\mathbf{T}^d} \leqslant \varepsilon}} |f_{\mathrm{sml}}(n)| \qquad \text{and} \qquad E'_{x',z'} := \sum_{\substack{N \leqslant n < 2N \\ n \equiv u' (\operatorname{mod} q) \\ |\frac{n}{N} - x'| \leqslant \varepsilon' \\ \|\theta n - z'\|_{\mathbf{T}^d} \leqslant \varepsilon'}} |f_{\mathrm{sml}}(n)|.$$

We have

$$\begin{aligned}\int_1^2 \int_{\mathbf{T}^d} E_{x,z}\, dz dx &= \sum_{\substack{N \leqslant n < 2N \\ n \equiv u (\operatorname{mod} q)}} |f_{\mathrm{sml}}(n)| \int_1^2 1_{|\frac{n}{N} - x| \leqslant \varepsilon} dx \int_{\mathbf{T}^d} 1_{\|\theta n - z\|_{\mathbf{T}^d} \leqslant \varepsilon} dz \\ &\leqslant (2\varepsilon)^{d+1} \sum_{\substack{N \leqslant n < 2N \\ n \equiv u (\operatorname{mod} q)}} |f_{\mathrm{sml}}(n)| \leqslant (2\varepsilon)^{d+1} \frac{20\delta}{q} N,\end{aligned}$$

the last step being a consequence of (4.6). It follows from this and (4.5) that

$$\int_1^2 \int_{\mathbf{T}^d} \Big(F(u,x,z) - \frac{q}{800N\delta(2\varepsilon)^{d+1}} E_{x,z} \Big) dz dx \geqslant \frac{1}{40},$$

and so there are specific choices of x, z such that

$$F(u,x,z) - \frac{q}{800N\delta(2\varepsilon)^{d+1}} E_{x,z} \geqslant \frac{1}{40},$$

which implies that

$$F(u,x,z) \geqslant \frac{1}{40} \qquad \text{and} \qquad E_{x,z} \leqslant \frac{800\delta N}{q}(2\varepsilon)^{d+1}. \tag{4.10}$$

Similarly, there are x', z' such that

$$F(u',x',z') \geqslant \frac{1}{40} \qquad \text{and} \qquad E_{x',z'} \leqslant \frac{800\delta N}{q}(2\varepsilon')^{d+1}. \tag{4.11}$$

From now on, we fix these specific choices of x, z, x', z' and set

$$X := \Big\{ n \in \mathbf{N} : n \equiv u (\operatorname{mod} q), \Big| \frac{n}{N} - x \Big|, \|\theta n - z\|_{\mathbf{T}^d} \leqslant \varepsilon \Big\}, \tag{4.12}$$

$$X' := \Big\{ n \in \mathbf{N} : n \equiv u' (\operatorname{mod} q), \Big| \frac{n}{N} - x' \Big|, \|\theta n - z'\|_{\mathbf{T}^d} \leqslant \varepsilon' \Big\}, \tag{4.13}$$

and

$$Y := \Big\{ n \in \mathbf{N} : n \equiv u + u' (\operatorname{mod} q), \Big| \frac{n}{N} - (x + x') \Big|, \|\theta n - (z + z')\|_{\mathbf{T}^d} \leqslant \varepsilon \Big\}. \tag{4.14}$$

Note that with this notation (4.10) and (4.11) imply

$$\sum_{n \in X} |f_{\mathrm{sml}}(n)| \ll \delta(2\varepsilon)^{d+1} q^{-1} N, \qquad \sum_{n \in X'} |f_{\mathrm{sml}}(n)| \ll \delta(2\varepsilon')^{d+1} q^{-1} N. \tag{4.15}$$

Lemma 4.5 *Suppose that* $\mathcal{F}$ *grows sufficiently rapidly, and that N is sufficiently large in terms of* δ, M. *Then the number of squares in* Y *is* $\ll (2\varepsilon)^{d+1} q^{-1} \mathfrak{S}(u+u',q) N^{1/2}$.

Proof Let $\mathscr{A}$ be the set of all $a \in \mathbf{Z}/q\mathbf{Z}$ for which $a^2 \equiv u + u' (\operatorname{mod} q)$. Thus, $|\mathscr{A}| = \mathfrak{S}(u+u',q)$. An upper bound for the number of squares in Y is then

$$\sum_{a\in\mathscr{A}} \sum_{n\in I} 1_{n\equiv a (\operatorname{mod} q)} \psi_\varepsilon^+(\theta n^2 - z - z'),$$

where $I = [(x + x' - \varepsilon)^{1/2} N^{1/2}, (x + x' + \varepsilon)^{1/2} N^{1/2}]$ and ψ_ε^+ is the majorant for the characteristic function of the ball $B_\varepsilon(0)$ in $\mathbf{T}^d$ constructed in Lemma A.2. Fourier expanding

$$1_{n\equiv a(\operatorname{mod} q)} = \frac{1}{q} \sum_{r(\operatorname{mod} q)} e\Big(-\frac{ra}{q}\Big) e\Big(\frac{rn}{q}\Big),$$

$$\psi_\varepsilon^+(t) = \sum_{\mathbf{r}\in\mathbf{Z}^d} \widehat{\psi_\varepsilon^+}(\mathbf{r}) e(\mathbf{r}\cdot t),$$

this can be written as

$$\sum_{a\in\mathscr{A}} \frac{1}{q} \sum_{r(\operatorname{mod} q)} e\Big(-\frac{ra}{q}\Big) \sum_{\mathbf{r}\in\mathbf{Z}^d} \widehat{\psi_\varepsilon^+}(\mathbf{r}) e(-\mathbf{r}\cdot(z+z')) \sum_{n\in I} e\Big(\mathbf{r}\cdot\theta n^2 + \frac{rn}{q}\Big). \tag{4.16}$$

The contribution from $\mathbf{r} = 0$ is

$$\frac{1}{q}\Big(\int \psi_\varepsilon^+\Big) \sum_{a\in\mathscr{A}} \sum_{r(\operatorname{mod} q)} e\Big(-\frac{ra}{q}\Big) \sum_{n\in I} e\Big(\frac{rn}{q}\Big).$$

If $r \neq 0$, the inner sum over n is at most q in magnitude, since the sum of $e(rn/q)$ over any interval of length q is zero. The total contribution from these terms is thus bounded independently of N, and so can be ignored if N is large enough. The contribution from $r = 0$ is $\frac{1}{q}\mathfrak{S}(u+u',q)(\int \psi_\varepsilon^+)|I|$, which is $\ll (2\varepsilon)^{d+1} q^{-1}\mathfrak{S}(u+u',q)N^{1/2}$ by Lemma A.2(i) and the bound $|I| \ll \varepsilon N^{1/2}$. The contribution to (4.16) from $\mathbf{r} \neq 0$ is bounded above by

$$\mathfrak{S}(u+u',q) \sum_{\mathbf{r}\in\mathbf{Z}^d\smallsetminus\{0\}} |\widehat{\psi_\varepsilon^+}(\mathbf{r})| \sup_{r(\operatorname{mod} q)} \Big| \sum_{n\in I} e\Big(\mathbf{r}\cdot\theta n^2 + \frac{r}{q}n\Big)\Big|.$$

By Corollary 3.3 and Lemma A.2(ii), this is

$$\ll q N^{1/2} \mathcal{F}(M)^{-1/C_2} \sum_{\mathbf{r}\in\mathbf{Z}^d\smallsetminus\{0\}} |\widehat{\psi_\varepsilon^+}(\mathbf{r})| \|\mathbf{r}\|_1 \ll_{\delta,M} N^{1/2}\mathcal{F}(M)^{-1/C_2}.$$

(Lemma A.2(ii) gives an implied constant depending on d, ε, but we have $d \leqslant M$ and $\varepsilon = \delta/M$.) Hence if $\mathcal{F}$ is chosen to be sufficiently rapidly-growing, this is smaller than $(\frac{2\delta}{M})^{M+1} M^{-1} N^{1/2}$, which is at most $N^{1/2}(2\varepsilon)^{d+1}q^{-1}N^{1/2}$. ∎

We will also need the following fact, proved using very similar techniques.

Lemma 4.6 *Suppose that* $\mathcal{F}$ *grows sufficiently rapidly, and that N is sufficiently large in terms of* δ, M. *Suppose that* $n \in X$. *Then the number of* $n' \in X'$ *for which* $n + n'$ *is a square is* $\ll (2\varepsilon')^{d+1} q^{-1}\mathfrak{S}(u+u',q)N^{1/2}$, *uniformly in n.*

https://doi.org/10.4153/CJM-2017-036-1 Published online by Cambridge University Press

Proof Once again, write $\mathscr{A}$ for the set of square roots of $u + u'$ in $\mathbf{Z}/q\mathbf{Z}$. Writing $m^2 = n + n'$, an upper bound for the quantity in question is

$$\sum_{a\in\mathscr{A}}\sum_{m\in J} 1_{m\equiv a(\operatorname{mod} q)}\psi^+_{\varepsilon'}(\theta m^2 - \theta n - z'),$$

where $J = [(n+(x'-\varepsilon')N)^{1/2}, (n+(x'+\varepsilon')N)^{1/2}]$ and $\psi^+_{\varepsilon'}$ is the majorant constructed in Lemma A.2 (but now with the smaller parameter ε'). Expanding in Fourier series much as before, this can be written as

$$\sum_{a\in\mathscr{A}}\frac{1}{q}\sum_{r(\operatorname{mod} q)} e\Big(-\frac{ra}{q}\Big)\sum_{\mathbf{r}\in\mathbf{Z}^d}\widehat{\psi^+_{\varepsilon'}}(\mathbf{r})e\big(-\mathbf{r}\cdot\theta(n+z')\big)\sum_{m\in J} e\Big(\mathbf{r}\cdot\theta m^2 + \frac{rm}{q}\Big).$$

Arguing in an essentially identical fashion to the proof of Lemma 4.5, we see that this is bounded by a main term of size $\ll (2\varepsilon')^{d+1}q^{-1}\mathfrak{S}(u+u',q)N^{1/2}$ plus an error of size $\ll_{\delta,M} N^{1/2}\mathcal{F}(M)^{-1/C_2}$. Choosing $\mathcal{F}$ to be sufficiently rapidly-growing, and recalling from (4.9) that $\varepsilon' \gg_{\delta,M} 1$, this can be made $\ll (2\varepsilon')^{d+1}q^{-1}\mathfrak{S}(u+u',q)N^{1/2}$. ∎

Finally, we need yet another fact with a similar proof. Define the set $Y_- \subset Y$ to be

$$\Big\{n\in\mathbf{N} : n \equiv u+u'(\operatorname{mod} q), \Big|\frac{n}{N} - (x+x')\Big|, \|\theta n - (z+z')\|_{\mathbf{T}^d} \leqslant \varepsilon - 2\varepsilon'\Big\}. \tag{4.17}$$

Lemma 4.7 *Suppose that $\mathcal{F}$ grows sufficiently rapidly. Then the number of squares in $Y \setminus Y_-$ is $\ll \delta(2\varepsilon)^{d+1}q^{-1}N^{1/2}$.*

Proof If $n \in Y \setminus Y_-$, then either

$$\varepsilon - 2\varepsilon' < \Big|\frac{n}{N} - (x+x')\Big| < \varepsilon \tag{4.18}$$

or

$$\varepsilon - 2\varepsilon' < \big\|\theta_i n - (z_i + z_i')\big\|_{\mathbf{T}^d} < \varepsilon \tag{4.19}$$

for some $i \in \{1,\dots,d\}$. The number of squares satisfying (4.18) is elementarily seen to be $O(\varepsilon' N^{1/2})$, which[3] is bounded as desired because of the choice of ε' (*cf.* (4.8)).

We now obtain an upper bound for the number of squares satisfying (4.19). By translating the function ψ^+_ε constructed in Lemma A.2 (with $d = 1$ in that lemma) we can obtain a smooth majorant ψ for the interval $\{t \in \mathbf{T} : \varepsilon - 2\varepsilon' < \|t - (z_i + z_i')\|_{\mathbf{T}} < \varepsilon\}$ such that

$$\int \psi \ll \varepsilon', \quad \sum_r |\widehat{\psi}(r)||r| \ll_{\varepsilon'} 1. \tag{4.20}$$

Then the number of squares satisfying (4.19) is bounded above by

$$\sum_{n\leqslant 2N^{1/2}}\psi(\theta_i n^2) = \sum_{r\in\mathbf{Z}}\widehat{\psi}(r)\sum_{n\leqslant 2N^{1/2}} e(r\theta_i n^2).$$

[3] Obviously this bound is rather crude, as we have completely ignored the fact that additionally $n \equiv u + u'(\operatorname{mod} q)$ and $\|\theta n - (z+z')\|_{\mathbf{T}^d} \leqslant \varepsilon$; but this is of little consequence in the grand scheme of the argument.

https://doi.org/10.4153/CJM-2017-036-1 Published online by Cambridge University Press

The term with $r = 0$ is $2N^{1/2}(\int \psi) \ll \varepsilon' N^{1/2}$. By Corollary 3.3 (applied with $d = 1$), the contribution from the terms with $r \neq 0$ is

$$\ll N^{1/2}\mathcal{F}(M)^{-1/C_2} \sum_{r \neq 0} |\widehat{\psi}(r)||r|.$$

By (4.20) this is $\ll_{\varepsilon'} N^{1/2}\mathcal{F}(M)^{-1/C_2}$, which, in view of (4.9), is $O(\varepsilon' N^{1/2})$ provided $\mathcal{F}(M)$ grows sufficiently rapidly. Thus, the total number of n satisfying (4.19) for some $i \in \{1, \dots, d\}$ is $O(\varepsilon' d N^{1/2})$, which is bounded as claimed by the choice of ε'. ■

To complete the proof of Proposition 4.1 it suffices to show that $A + A$ contains all but $\ll \delta(2\varepsilon)^{d+1} q^{-1}\mathfrak{S}(u + u', q)N^{1/2}$ of the squares in Y. Indeed, if δ is chosen small enough, then this will be $\leqslant \eta(2\varepsilon)^{d+1}q^{-1}\mathfrak{S}(u + u', q)N^{1/2}$, the bound claimed. Let $S \subset Y$ be the set of all squares in Y that are not in $A + A$; thus, it suffices to establish the bound

$$|S| \ll \delta(2\varepsilon)^{d+1}q^{-1}\mathfrak{S}(u + u', q)N^{1/2}. \tag{4.21}$$

Recall the definitions (4.12), (4.13) of X, X'. We will need to introduce smoothed approximants χ, χ' to the characteristic functions of X, X', respectively, with the following properties:

(a) χ is a minorant for X, that is to say $0 \leqslant \chi(n) \leqslant 1_X(n)$ for all n;
(b) χ' is a minorant for X', that is to say $0 \leqslant \chi'(n) \leqslant 1_X(n)$ for all n;
(c) $\chi(n) = 1$ on the set $\{n \in \mathbf{N} : n \equiv u (\text{mod } q), |\frac{n}{N} - x|, \|\theta n - z\|_{\mathbf{T}^d} \leqslant \varepsilon - \varepsilon'\}$;
(d) $\int_0^1 |\widehat{\chi}(t)|dt, \int_0^1 |\widehat{\chi'}(t)|dt = O_M(1)$;
(e) $\sum_n \chi'(n) \gg (2\varepsilon')^{d+1}q^{-1}N$.

Such a function is constructed in Lemma A.3 (which must be applied twice, once with parameter ε and once with parameter ε').

In particular, it follows from (4.15) that

$$\sum_n |f_{\text{sml}}\chi(n)| \ll \delta(2\varepsilon)^{d+1}q^{-1}N, \qquad \sum_n |f_{\text{sml}}\chi'(n)| \ll \delta(2\varepsilon')^{d+1}q^{-1}N. \tag{4.22}$$

Our assumption that $A + A$ is disjoint from S implies that

$$\sum_{n \in S} (1_A\chi * 1_A\chi')(n) = 0. \tag{4.23}$$

To investigate this expression, we use the decomposition from the regularity lemma,

$$1_A = f_{\text{tor}} + f_{\text{sml}} + f_{\text{unf}}.$$

The left-hand side of (4.23) can then be expanded as a sum of 9 terms

$$T_{\bullet,\bullet'} := \sum_{n \in S} (f_\bullet \chi * f_{\bullet'}\chi')(n),$$

where $\bullet, \bullet' \in \{\text{tor}, \text{sml}, \text{unf}\}$. Thus

$$|T_{\text{tor},\text{tor}}| \leqslant \sum_{(\bullet,\bullet') \neq (\text{tor},\text{tor})} |T_{\bullet,\bullet'}|. \tag{4.24}$$

We analyse these 9 terms $T_{\bullet,\bullet'}$ separately, beginning with the "main term" $T_{\text{tor},\text{tor}}$.

Writing

$$f_{\text{tor}}(n) = F\Big(n (\text{mod } q), \frac{n}{N}, \theta n\Big),$$

https://doi.org/10.4153/CJM-2017-036-1 Published online by Cambridge University Press

we can expand $T_{\mathrm{tor,tor}}$ as

$$\sum_{n\in S}\sum_m F\Big(m(\mathrm{mod}\ q), \frac{m}{N}, \theta m\Big)\chi(m) \\ \times F\Big(n-m(\mathrm{mod}\ q), \frac{n-m}{N}, \theta(n-m)\Big)\chi'(n-m).$$

Since $\chi(m)$ is supported where $m \equiv u(\mathrm{mod}\ q)$ and $|\frac{m}{N}-x|, \|\theta m - z\|_{\mathbf{T}^d} \leqslant \varepsilon$, and since F is M-Lipschitz, using (4.10), we have that

$$F\Big(m(\mathrm{mod}\ q), \frac{m}{N}, \theta m\Big)\chi(m) = \Big(F(u,x,z) + O(M\varepsilon)\Big)\chi(m) \geqslant \frac{1}{80}\chi(m)$$

if δ is sufficiently small (note, recalling the definition (4.7) of ε, that $M\varepsilon = \delta$). Similarly,

$$F\Big(n-m(\mathrm{mod}\ q), \frac{n-m}{N}, \theta(n-m)\Big)\chi'(n-m) \geqslant \frac{1}{80}\chi'(n-m).$$

It follows that

$$T_{\mathrm{tor,tor}} \gg \sum_{n\in S}\sum_m \chi(m)\chi'(n-m) = \sum_{n\in S}\sum_m \chi(n-m)\chi'(m). \tag{4.25}$$

Recall the definition (4.17) of $Y_- \subset Y$. If $n \in Y_-$ and $m \in \mathrm{Supp}(\chi') \subset X'$ then $n-m \equiv u(\mathrm{mod}\ q)$ and $|\frac{n-m}{N}-x|, \|\theta(n-m)-z\|_{\mathbf{T}^d} \leqslant \varepsilon - \varepsilon'$, and therefore by property (c) of χ we have $\chi(n-m) = 1$. It follows from these observations, (4.25) and point (e) of the properties of χ, χ' that

$$T_{\mathrm{tor,tor}} \gg \sum_{n\in S\cap Y_-}\sum_m \chi'(n-m)\chi'(m) \gg |S\cap Y_-|\sum_m \chi'(m) \\ \gg |S\cap Y_-|(2\varepsilon')^{d+1}q^{-1}N. \tag{4.26}$$

We set this estimate aside for later use.

Next we look at the terms $T_{\bullet,\bullet'}$ in which $\bullet' = \mathrm{sml}$. Here we require the *a priori* bound

$$|S| \ll (2\varepsilon)^{d+1}q^{-1}\mathfrak{S}(u+u',q)N^{1/2}. \tag{4.27}$$

This is, of course, weaker than the result we are trying to prove, but it follows immediately from Lemma 4.5. All of these terms $T_{\bullet,\mathrm{sml}}$ have the form

$$T_{\bullet,\mathrm{sml}} = \sum_{n\in S}(g * f_{\mathrm{sml}}\chi')(n) = \sum_{n\in S}\sum_m g(n-m)f_{\mathrm{sml}}\chi'(m),$$

where g is some function bounded pointwise by 1. Thus,

$$|T_{\bullet,\mathrm{sml}}| \leqslant |S|\sum_m |f_{\mathrm{sml}}\chi'(m)|,$$

and so, by (4.27) and (4.22),

$$T_{\bullet,\mathrm{sml}} \ll \delta(4\varepsilon\varepsilon')^{d+1}q^{-2}\mathfrak{S}(u+u',q)N^{3/2}. \tag{4.28}$$

Next we turn to the bounding of

$$T_{\mathrm{sml,tor}} = \sum_{n\in S}(f_{\mathrm{sml}}\chi * f_{\mathrm{tor}}\chi')(n).$$

This expands as

$$\sum_{n\in S}\sum_m f_{\mathrm{sml}}\chi(n-m)F\Big(m(\mathrm{mod}\ q), \frac{m}{N}, \theta m\Big)\chi'(m).$$

https://doi.org/10.4153/CJM-2017-036-1 Published online by Cambridge University Press

By the Lipschitz property of F and the fact that χ' is supported on X', this is

$$F(u',x',z')\sum_{\substack{m\\ n\in S}} f_{\mathrm{sml}}\chi(n-m)\chi'(m) + O(\varepsilon' M)\sum_{\substack{m\\ n\in S}} |f_{\mathrm{sml}}\chi(n-m)|\chi'(m).$$

Since $\varepsilon' < \varepsilon < 1/M$, it follows that

$$T_{\mathrm{sml,tor}} \ll \sum_{n\in S}\sum_{m} |f_{\mathrm{sml}}\chi(n-m)|\chi'(m) = \sum_{n',m} |f_{\mathrm{sml}}\chi(n')|\chi'(m)1_S(n'+m).$$

By (4.22), this is

$$\ll \delta(2\varepsilon)^{d+1}q^{-1}N \sup_{n'\in\operatorname{Supp}\chi}\sum_{m}\chi'(m)1_S(n'+m).$$

By Lemma 4.6 and the fact that $\operatorname{Supp}\chi \subset X$, $\operatorname{Supp}\chi' \subset X'$, we conclude that

$$T_{\mathrm{sml,tor}} \ll \delta(4\varepsilon\varepsilon')^{d+1}q^{-2}\mathfrak{S}(u+u',q)N^{3/2}. \tag{4.29}$$

In all of the remaining terms $T_{\bullet,\bullet'}$ that we have yet to bound, at least one of $\bullet,\bullet'$ is unf. If $\bullet$ = unf, then such a term has the form

$$T_{\mathrm{unf},\bullet'} = \sum_{n\in S}(f_{\mathrm{unf}}\chi * g)(n),$$

where g is some function bounded pointwise by 1. This can be written in Fourier space as

$$\int_0^1 \widehat{f_{\mathrm{unf}}\chi}(t)\widehat{g}(t)\widehat{1}_S(t)\,dt,$$

where g is a bounded function. By Hölder's inequality, the right-hand side here is bounded above by

$$\|\widehat{f_{\mathrm{unf}}\chi}\|_\infty^{1/3}\Big(\int_0^1|\widehat{f_{\mathrm{unf}}\chi}|^2\Big)^{1/3}\Big(\int_0^1|\widehat{g}|^2\Big)^{1/2}\Big(\int_0^1|\widehat{1}_S|^6\Big)^{1/6}. \tag{4.30}$$

By Parseval's identity and the boundedness of $f_{\mathrm{unf}}, g, \chi$, we have

$$\int_0^1|\widehat{f_{\mathrm{unf}}\chi}|^2,\ \int_0^1|\widehat{g}|^2 \ll N,$$

and Proposition 3.4 tells us that

$$\int_0^1|\widehat{1}_S(t)|^6\,dt \ll N^2.$$

Finally, we note that

$$\widehat{f_{\mathrm{unf}}\chi}(t) = \int_0^1 \widehat{f_{\mathrm{unf}}}(t')\widehat{\chi}(t-t')\,dt',$$

and so by property (d) of χ we have

$$\|\widehat{f_{\mathrm{unf}}\chi}\|_\infty \leqslant \|\widehat{f_{\mathrm{unf}}}\|_\infty\|\widehat{\chi}\|_1 \ll_M N\mathcal{F}(M)^{-1}.$$

Combining all these estimates together gives

$$T_{\mathrm{unf},\bullet'} = \sum_n (f_{\mathrm{unf}}\chi * g)(n)1_S(n) \ll_M N^{3/2}\mathcal{F}(M)^{-1/3}.$$

https://doi.org/10.4153/CJM-2017-036-1 Published online by Cambridge University Press

If the growth of $\mathcal{F}$ is sufficiently rapid, we obtain in view of the fact that $d, q \leqslant M$, $\varepsilon = \delta/M$ and (4.9) that

$$T_{\mathrm{unf},\bullet'} \ll \delta(4\varepsilon\varepsilon')^{d+1}q^{-2}N^{3/2}. \tag{4.31}$$

An almost identical argument (relying instead on the bound $\|\chi'\|_1 = O_M(1)$) yields

$$T_{\bullet,\mathrm{unf}} \ll \delta(4\varepsilon\varepsilon')^{d+1}q^{-2}N^{3/2}. \tag{4.32}$$

Combining (4.26), (4.28), (4.29), (4.31), and (4.32) with (4.24), we obtain

$$|S \cap Y_-|(2\varepsilon')^{d+1}q^{-1}N \ll \delta(4\varepsilon\varepsilon')^{d+1}q^{-2}\mathfrak{S}(u+u',q)N^{3/2},$$

and therefore

$$|S \cap Y_-| \ll \delta(2\varepsilon)^{d+1}q^{-1}\mathfrak{S}(u+u',q)N^{1/2}.$$

Lemma 4.7 provides the bound

$$|S \cap (Y \smallsetminus Y_-)| \ll \delta(2\varepsilon)^{d+1}q^{-1}\mathfrak{S}(u+u',q)N^{1/2}.$$

Combining this with the preceding yields

$$|S| \ll \delta(2\varepsilon)^{d+1}q^{-1}\mathfrak{S}(u+u',q)N^{1/2},$$

which is exactly (4.21). This completes the proof of Proposition 4.1. ∎

5 The Square-root of a Bohr Set

Suppose that $\mathbf{N}$ is partitioned into two colour classes V and W, neither of which has a monochromatic solution to $x + y = z^2$. The main result of the last section, Proposition 4.1, shows that if $V \cap [N, 2N)$ has size at least $N/2$, then $V + V$ contains almost all of the squares in a "Bohr set" $\Lambda := \{n \in \mathbf{N} : n \equiv b (\mathrm{mod}\ q), |\frac{n}{N} - x|, \|\theta n - z\|_{\mathbf{T}^d} \leqslant \varepsilon\}$. This means that most of $\sqrt{\Lambda}$ must lie in W. In this section we examine the additive properties of such square roots $\sqrt{\Lambda}$. (Recall that $\sqrt{\Lambda}$ is by definition the set of *integers* n such that $n^2 \in \Lambda$.)

Here is the main result of the section.

Proposition 5.1 *Let $\eta > 0$. Then there is a function $\Omega: \mathbf{N}^3 \to \mathbf{R}_+$ with the following property. Suppose we have $q, d \in \mathbf{N}$, $\varepsilon > 0$, $x \in [0,3]$, $\theta, z \in \mathbf{T}^d$, and $N \in \mathbf{N}$. Suppose that θ is $(\Omega(q,d,1/\varepsilon), N)$-irrational. Suppose that b is a square modulo q and set*

$$Y := \Big\{ n \in \mathbf{N} : n \equiv b (\mathrm{mod}\ q), \Big|\frac{n}{N} - x\Big|, \|\theta n - z\|_{\mathbf{T}^d} \leqslant \varepsilon \Big\}.$$

Let $Y' \subset Y$ be a set containing all but at most $\eta(2\varepsilon)^{d+1}q^{-1}\mathfrak{S}(b,q)N^{1/2}$ of the squares in Y. Then, for all but at most $O(\eta\varepsilon q^{-1}N^{1/4})$ of the elements $t \in Q$, where

$$Q := \mathsf{P}\Big(\Big[(2x)^{1/4} - \frac{\varepsilon}{100}, (2x)^{1/4} + \frac{\varepsilon}{100} \Big]; N^{1/4}, q \Big),$$

we have $t^2 \in \sqrt{Y'} + \sqrt{Y'}$.

(Recall that $\mathsf{P}(I; N, q) := \{n \in \mathbf{Z} : n/N \in I, q|n\}$.)

https://doi.org/10.4153/CJM-2017-036-1 Published online by Cambridge University Press

The proof of this is a little complicated, so we break it down into a few lemmas. We have $\sqrt{Y} = \bigcup_{a\in\mathscr{A}} Z_+^a \cup Z_-^a$, where

$$Z_\pm^a := \big\{ n \in \mathbf{N} : n \equiv \pm a (\operatorname{mod} q), (x-\varepsilon)^{1/2}N^{1/2} \leqslant n \leqslant (x+\varepsilon)^{1/2}N^{1/2}, \\ \|\theta n^2 - z\|_{\mathbf{T}^d} \leqslant \varepsilon \big\}, \tag{5.1}$$

and $\mathscr{A}$ is the set of square roots of b in $\mathbf{Z}/q\mathbf{Z}$. Define

$$\widetilde{Z}_\pm^a := \sqrt{Y'} \cap Z_\pm^a;$$

then

$$\sum_{a\in\mathscr{A}} |Z_\pm^a \smallsetminus \widetilde{Z}_\pm^a| \ll \eta(2\varepsilon)^{d+1} q^{-1}\mathfrak{S}(b,q)N^{1/2},$$

by assumption. It follows that there is some $a \in \mathscr{A}$ such that

$$|Z_\pm^a \smallsetminus \widetilde{Z}_\pm^a| \ll \eta(2\varepsilon)^{d+1}q^{-1}N^{1/2}. \tag{5.2}$$

Henceforth, we fix this value of a and write $Z_\pm = Z_\pm^a$ for brevity. To orient ourselves we remark that, if Ω grows sufficiently rapidly, then one could prove that

$$|Z_\pm| \sim (2\varepsilon)^{d+1}q^{-1}N^{1/2}$$

(here we are using ~ somewhat informally). We will not need to explicitly prove any statement of this kind separately.

Lemma 5.2 *Suppose that $n_+ \in Z_+$. Then*

$$\#\{n_- \in Z_- : n_- + n_+ = q^2m^2 \textit{ for some } m \in \mathbf{Z}\} \ll (2\varepsilon)^{d+1}q^{-1}N^{1/4},$$

the implied constant being uniform in n_+ and independent of a (recall that $Z_\pm$ depends on a). Similarly, if $n_- \in Z_-$, then

$$\#\{n_+ \in Z_+ : n_- + n_+ = q^2m^2 \textit{ for some } m \in \mathbf{Z}\} \ll (2\varepsilon)^{d+1}q^{-1}N^{1/4},$$

the implied constant being uniform in n_- and in a.

Proof The quantity we are interested in can be written as

$$\sum_{m\in I(n_+)} 1_{\|\theta(q^2m^2-n_+)^2-z\|_{\mathbf{T}^d}\leqslant\varepsilon},$$

where $I(n_+)$ is the interval

$$\frac{1}{q}\big((x-\varepsilon)^{1/2}N^{1/2}+n_+\big)^{1/2} \leqslant m \leqslant \frac{1}{q}\big((x+\varepsilon)^{1/2}N^{1/2}+n_+\big)^{1/2},$$

the cardinality of which satisfies

$$|I(n_+)| \ll \varepsilon q^{-1}N^{1/4} \tag{5.3}$$

uniformly in n_+. To bound this above, take a majorant ψ_ε^+ to the unit ball $B_\varepsilon(0) \subset \mathbf{T}^d$, as in Lemma A.2. Then our quantity is at most

$$\sum_{m\in I(n_+)} \psi_\varepsilon^+\big(\theta(q^2m^2-n_+)^2 - z\big).$$

https://doi.org/10.4153/CJM-2017-036-1 Published online by Cambridge University Press

Fourier expanding ψ_ε^+, this is

$$\sum_{\mathbf{r}\in\mathbf{Z}^d} \widehat{\psi_\varepsilon^+}(\mathbf{r}) \sum_{m\in I(n_+)} e(q^4\mathbf{r}\cdot\theta m^4 + \cdots),$$

where the dots denote terms of degree at most 2 in m (which can depend on $\mathbf{r}, n_+, \theta, z, q$). The contribution from $\mathbf{r} = 0$ is $|I(n_+)|(\int \psi_\varepsilon^+)$, which, by (5.3) and Lemma A.2(i), is $\ll (2\varepsilon)^{d+1}q^{-1}N^{1/4}$. By Corollary 3.3 (and since $|I(n_+)| \leqslant N^{1/4}$), we have

$$\Big|\sum_{m\in I(n_+)} e(q^4\mathbf{r}\cdot\theta m^4 + \cdots)\Big| \leqslant N^{1/4}\Big(\frac{q^4\|\mathbf{r}\|_1}{\Omega(q,d,1/\varepsilon)}\Big)^{1/C_4}.$$

By Lemma A.2(ii), the contribution from $\mathbf{r} \neq 0$ is therefore

$$\ll N^{1/4}\Big(\frac{q^4}{\Omega(q,d,1/\varepsilon)}\Big)^{1/C_4} \sum_{\mathbf{r}\in\mathbf{Z}^d\setminus\{0\}} |\widehat{\psi_\varepsilon^+}(\mathbf{r})|\|\mathbf{r}\|_1$$

$$\ll_{\varepsilon,d} N^{1/4}\Big(\frac{q^4}{\Omega(q,d,1/\varepsilon)}\Big)^{1/C_4},$$

which is also $\ll (2\varepsilon)^{d+1}q^{-1}N^{1/4}$ if Ω is chosen appropriately. ∎

Define progressions P_+, P_- by

(5.4) $$P_\pm := \big\{n \in N : n \equiv \pm a (\mathrm{mod}\ q), (x-\varepsilon)^{1/2}N^{1/2} \leqslant n \leqslant (x+\varepsilon)^{1/2}N^{1/2}\big\},$$

and recall from the statement of Proposition 5.1 the definition of Q, viz.

$$Q := \mathsf{P}\Big(\Big[(2x)^{1/4} - \frac{\varepsilon}{100}, (2x)^{1/4} + \frac{\varepsilon}{100}\Big]; N^{1/4}, q\Big).$$

Observe that if $t \in Q$ then t^2 is a sum $p_+ + p_-$ in $\gg \varepsilon q^{-1}N^{1/2}$ ways. Indeed

$$\Big((2x)^{1/2} - \frac{\varepsilon}{10}\Big)N^{1/2} < t^2 < \Big((2x)^{1/2} + \frac{\varepsilon}{10}\Big)N^{1/2}$$

and $t^2 \equiv 0 (\mathrm{mod}\ q)$, hence for any of the $\gg \varepsilon q^{-1}N^{1/2}$ values of p_+ with

$$(x^{1/2} - \frac{\varepsilon}{10})N^{1/2} < p_+ < (x^{1/2} + \frac{\varepsilon}{10})N^{1/2}$$

and $p_+ \equiv a (\mathrm{mod}\ q)$ we have $t^2 - p_+ \in P_-$.

Note that from (5.1) and (5.4) we have

(5.5) $$Z_\pm = \{n \in P_\pm : \|\theta n^2 - z\|_{\mathbf{T}^d} \leqslant \varepsilon\}.$$

This suggests the intuition behind the arguments that follow, which is that $Z_\pm$ behaves like a "pseudorandom" subset of $P_\pm$ of density $(2\varepsilon)^d$. Thus, it is reasonable to expect that a typical t^2, $t \in Q$, will have $\gg (2\varepsilon)^{2d+1}q^{-1}N^{1/2}$ representations as $z_+ + z_-$ with $z_+ \in Z_+, z_- \in Z_-$.

Lemma 5.3 *Suppose that Ω grows sufficiently rapidly. Write $r(n)$ for the number of representations of n as $z_+ + z_-$ with $z_\pm \in Z_\pm$. Suppose that Ω grows fast enough. Then all but at most $\eta\varepsilon q^{-1}N^{1/4}$ of elements $t \in Q$, have $r(t^2) \gg (2\varepsilon)^{2d+1}q^{-1}N^{1/2}$.*

https://doi.org/10.4153/CJM-2017-036-1 Published online by Cambridge University Press

Proof If the lemma is false, then for any absolute constant c (which we may specify later) there is a set $T \subset Q$, $|T| \geqslant \eta\varepsilon q^{-1}N^{1/4}$, such that

$$\sum_{t\in T} r(t^2) \leqslant c(2\varepsilon)^{2d+1}q^{-1}|T|N^{1/2}. \tag{5.6}$$

We first introduce a smoothed variant of r, defined by $\widetilde{r}(n) = f_+ * f_-(n)$, where

$$f_\pm(n) = 1_{P_\pm}(n)\psi_\varepsilon^-(\theta n^2 - z),$$

where ψ_ε^- is a suitable minorant to $B_\varepsilon(0)$, as constructed in Lemma A.2. From (5.5) we see that $1_{Z_\pm} \geqslant f_\pm$ pointwise, and so $r(n) \geqslant \widetilde{r}(n)$ pointwise. Define

$$g_\pm(n) = 1_{P_\pm}(n)\Big(\psi_\varepsilon^-(\theta n^2 - z) - \int \psi_\varepsilon^-\Big).$$

Fourier expanding ψ_ε^-, we see that

$$\widehat{g_\pm}(t) = \sum_{\mathbf{r}\in\mathbf{Z}^d\smallsetminus\{0\}} \widehat{\psi_\varepsilon^-}(\mathbf{r}) \sum_{n\in P_+} e(\mathbf{r}\cdot\theta n^2 + nt - \mathbf{r}\cdot z).$$

Parametrising $n \in P_+$ as $n = qm + b$ for m in some interval I with $|I| = |P_+| < N^{1/2}$, it follows from Corollary 3.3 that the inner sum is $\ll N^{1/2}\Omega(q,d,1/\varepsilon)^{-1/C_2}\|\mathbf{r}\|_1$. Therefore, by Lemma A.2(ii), we have

$$\begin{aligned} \|\widehat{g_\pm}\|_\infty &\ll N^{1/2}\Omega(q,d,1/\varepsilon)^{-1/C_2} \sum_{\mathbf{r}\in\mathbf{Z}^d} |\widehat{\psi_\varepsilon^-}(\mathbf{r})|\|\mathbf{r}\|_1 \\ &\ll_{\varepsilon,d} N^{1/2}\Omega(q,d,1/\varepsilon)^{-1/C_2}. \end{aligned} \tag{5.7}$$

Now, writing

$$f_\pm = 1_{P_\pm}\int \psi_\varepsilon^- + g_\pm,$$

we may expand $\sum_{t\in T}\widetilde{r}(t^2)$ as a sum of four terms. The "main term" is

$$E_{\text{main}} = \Big(\int \psi_\varepsilon^-\Big)^2 \sum_{t\in T} 1_{P_+} * 1_{P_-}(t^2).$$

The three error terms each have the shape

$$E_{\text{error}} = \sum_{t\in T} g_\pm * h_\mp(t^2),$$

where $h_\mp$ is bounded pointwise by 1 and supported on $P_\mp$.

We have already remarked that if $t \in Q$, then t^2 has $\gg \varepsilon q^{-1}N^{1/2}$ representations as $p_+ + p_-$, and therefore

$$E_{\text{main}} \gg (2\varepsilon)^{2d}\cdot|T|\cdot\varepsilon q^{-1}N^{1/2} \gg \eta(2\varepsilon)^{2d+2}q^{-2}N^{3/4}. \tag{5.8}$$

On the other hand,

$$E_{\text{error}} = \int_0^1 \widehat{g_\pm}(\theta)\widehat{h_\mp}(\theta)\widehat{1_{T^2}}(\theta)d\theta,$$

where $T^2 := \{t^2 : t \in T\}$. Using the same application of Hölder's inequality as in (4.30),

$$E_{\text{error}} \ll \|\widehat{g_\pm}\|_\infty^{1/3}\Big(\int_0^1 |\widehat{g_\pm}|^2\Big)^{1/3}\Big(\int_0^1 |\widehat{h_\mp}|^2\Big)^{1/2}\Big(\int_0^1 |\widehat{1_{T^2}}|^6\Big)^{1/6}.$$

By Parseval and the crude bound $|P_\pm| \ll N^{1/2}$, we have

$$\int_0^1 |\widehat{g_\pm}|^2, \int_0^1 |\widehat{h_\mp}|^2 \ll N^{1/2}.$$

Proposition 3.4 tells us that

$$\int_0^1 |\widehat{1_{T^2}}|^6 \ll N.$$

Putting this together with (5.7) gives

$$E_{\text{error}} \ll \Omega(q, d, 1/\varepsilon)^{-1/3C_2} N^{3/4}.$$

Choosing Ω to grow sufficiently quickly, we see from (5.8) that this can be made less than $\frac{1}{10}$ of E_{main}. It follows from (5.8) that

$$\sum_{t \in T} \widetilde{r}(t^2) \geqslant E_{\text{main}} - 3E_{\text{error}} > \frac{1}{2} E_{\text{main}} \gg (2\varepsilon)^{2d+1} q^{-1} |T| N^{1/2},$$

contrary to (5.6) if c was chosen small enough. ∎

Finally we put Lemmas 5.2 and 5.3 together to establish Proposition 5.1. It is certainly enough (in view of the definitions of $\widetilde{Z}_\pm$) to show that $\widetilde{Z}_+ + \widetilde{Z}_-$ contains t^2 for all but at most $O(\eta\varepsilon q^{-1} N^{1/4})$ of the elements $t \in Q$. By Lemma 5.3, all but at most $\eta\varepsilon q^{-1} N^{1/4}$ elements $t \in Q$ are such that t^2 is *well-represented* in $Z_+ + Z_-$, by which we mean that $r(t^2) \gg (2\varepsilon)^{2d+1} q^{-1} N^{1/2}$, where $r(t^2)$ is the number of representations of t^2 as $z_+ + z_-$. Suppose now that we pass from $Z_\pm$ to $\widetilde{Z}_\pm$. The number of pairs (z_+, z_-) with $z_+ + z_-$ the square of an element in Q that are lost in this way is, by Lemma 5.2, bounded above by $\ll |Z_\pm \setminus \widetilde{Z}_\pm| (2\varepsilon)^{d+1} q^{-1} N^{1/4}$. By (5.2), this is bounded by $\ll \eta(2\varepsilon)^{2d+2} q^{-2} N^{3/4}$. The number of t for which t^2 is well-represented but does not lie in $\widetilde{Z}_+ + \widetilde{Z}_-$ is therefore bounded above by

$$\ll \frac{\eta(2\varepsilon)^{2d+2} q^{-2} N^{3/4}}{(2\varepsilon)^{2d+1} q^{-1} N^{1/2}} = O(\eta\varepsilon q^{-1} N^{1/4}).$$

This completes the proof of Proposition 5.1. ∎

6 Gaps Between Sums of Two Squares

In this section we prove a result, Proposition 6.1, that we will need in the next section. It seems possible that such a result appears in the literature already, but we do not know a reference. We prove a slightly more general result than we actually need, since this is plausibly of independent interest.

Proposition 6.1 *Let $\alpha_1, \beta_1, \gamma_1, \alpha_2, \beta_2, \gamma_2$ be nonnegative reals with $\alpha_1 < \beta_1$, $\alpha_2 < \beta_2$, $\alpha_1^2 + \alpha_2^2 < \gamma_1 < \gamma_2 < \beta_1^2 + \beta_2^2$. Let $q \in \mathbf{N}$ and set $P_i := \mathrm{P}([\alpha_i, \beta_i]; N, q)$ for $i = 1, 2$. Suppose that $\gamma_1 \leqslant n/N^2 \leqslant \gamma_2$. Then there are $n_1 \in P_1$, $n_2 \in P_2$ such that*

$$|n_1^2 + n_2^2 - n| \ll \sqrt{N}.$$

The implied constant may depend on $\alpha_i, \beta_i, \gamma_i, q$ but is independent of n and N.

https://doi.org/10.4153/CJM-2017-036-1 Published online by Cambridge University Press

Remark A well-studied case is that in which $P_1 = P_2 = \{1, \dots, N\}$. Then it is well known that there is a sum of two squares $n_1^2 + n_2^2$ within $O(N^{1/2})$ of any $n \leqslant N^2$. One argument to prove this is very simple: take $n_1 = \lfloor \sqrt{n} \rfloor$, noting that $|n - n_1^2| \ll N$, and then set $n_2 := \lfloor \sqrt{n - n_1^2} \rfloor$. No bound of the form $o(N^{1/2})$ is known, a problem Montgomery [14, Problem 64, p. 208] attributes to Littlewood. The argument just sketched does not adapt to our case, since the n_2 produced is necessarily very small. However, there is another type of argument giving a similar bound and allowing us to take $n_1 \approx n_2$. The idea here is to take

$$n_1(k) = \lfloor \sqrt{n/2} \rfloor + k, \quad n_2(k) = \lfloor \sqrt{n/2} \rfloor - k,$$

where $k \in Z$ is to be specified later. Observe that

$$n_1(k)^2 + n_2(k)^2 = 2\lfloor \sqrt{n/2} \rfloor^2 + 2k^2,$$

and so, in particular,

$$n_1(0)^2 + n_2(0)^2 \leqslant n,$$
$$n_1(k)^2 + n_2(k)^2 \geqslant n - 2\sqrt{n} + 2k^2 > n$$

for $k = \lceil \sqrt{n} \rceil$ and

$$(n_1(k+1)^2 + n_2(k+1)^2) - (n_1(k)^2 - n_2(k))^2 = 4k + 2 \ll \sqrt{n}$$

uniformly for $k \leqslant \lceil \sqrt{n} \rceil$. It follows from the "discrete intermediate value theorem" that there is some k for which $|n_1(k)^2 + n_2(k)^2 - n| \ll \sqrt{n}$.

It turns out that this argument does generalise to allow us to prove Proposition 6.1.

Proof For the duration of this proof, the implied constant in the $O()$ and $\ll, \gg$ notations may depend on $\alpha_i, \beta_i, \gamma_i, q$. We can clearly assume that N is sufficiently large.

For each $\gamma \in [\gamma_1, \gamma_2]$, define I_γ to be the set of all $\lambda \in \mathbf{R}$ for which there exist $t_1, t_2 \in \mathbf{R}$ with $\alpha_1 \leqslant t_1 \leqslant \alpha_2$, $\beta_1 \leqslant t_2 \leqslant \beta_2$, $t_1/t_2 = \lambda$ and $t_1^2 + t_2^2 = \gamma$. Let $\widetilde{I}_\gamma$ be the middle half of I_γ. It is easy to see that I_γ is a closed interval whose length is positive and varies continuously as a function of γ, and is therefore bounded below uniformly in γ. The same is true for $\widetilde{I}_\gamma$. This implies that

(a) There is an absolute $\varepsilon \gg 1$ such that if $\lambda \in \widetilde{I}_\gamma$, then we may find t_1, t_2 with $t_1/t_2 = \lambda$ and

$$\alpha_i + \varepsilon \leqslant t_i \leqslant \beta_i - \varepsilon; \tag{6.1}$$

(b) $\widetilde{I}_\gamma$ contains a rational $a(\gamma)/b(\gamma)$ with $a(\gamma), b(\gamma) = O(1)$ and neither $a(\gamma)$ nor $b(\gamma)$ zero.

Now suppose that n is given satisfying $\gamma_1 \leqslant n/N^2 \leqslant \gamma_2$. Set $\gamma := n/N^2$, and select rationals $a = a(\gamma)$, $b = b(\gamma)$, not both zero, as in (b) above. According to (a), there are t_1, t_2 with $t_1^2 + t_2^2 = \gamma$, $t_1/t_2 = a/b$ and such that (6.1) is satisfied.

Now set

$$n_1(k) := q\left\lfloor \frac{t_1 N}{q} \right\rfloor + qkb, \quad n_2(k) := q\left\lfloor \frac{t_2 N}{q} \right\rfloor - qka.$$

Evidently, $q | n_1(k), n_2(k)$. Moreover, from (6.1), it follows that $\alpha_i \leqslant n_i(k)/N \leqslant \beta_i$ provided $|k| \leqslant cN$ for suitably small $c \gg 1$. Therefore, for k in this range we have $n_i(k) \in P_i$. Observe that

$$n_1(0)^2 + n_2(0)^2 \leqslant (t_1^2 + t_2^2)N^2 = n.$$

Also,

$$\begin{aligned}(6.2)\quad & n_1(k)^2 + n_2(k)^2 \\ &= q^2\Big(\Big\lfloor \frac{t_1N}{q}\Big\rfloor^2 + \Big\lfloor \frac{t_2N}{q}\Big\rfloor^2 + 2k\big(a\{\frac{t_2N}{q}\} - b\{\frac{t_1N}{q}\}\big) + k^2(a^2+b^2)\Big) \\ &\geqslant n - O(N) - O(k) + q^2k^2(a^2+b^2),\end{aligned}$$

and in particular

$$n_1(k)^2 + n_2(k)^2 > n$$

for some $k = O(\sqrt{N})$.

Moreover, from (6.2) again we have

$$\big|(n_1(k+1)^2 + n_2(k+1)^2) - (n_1(k)^2 - n_2(k))^2\big| = O(k).$$

It follows from these properties and a discrete intermediate value argument that there is some $k = O(\sqrt{N})$ for which $|n_1(k)^2 + n_2(k)^2 - n| \ll \sqrt{N}$. The result follows. ∎

7 Proof of the Main Theorem

In Proposition 7.2 we will synthesise the main results of Sections 4 and 5, together with the following small (and well-known) lemma.

Lemma 7.1 *Let $Q \subset \mathbf{N}$ be a finite arithmetic progression of size at least* 100, *and suppose that $S \subset Q$ is a set of size at least $\frac{9}{10}|Q|$. Then $S + S$ contains a subprogression of $Q + Q$ of size at least $|Q|$ with the same common difference as Q.*

Proof By translating we can assume that $Q = \{1, \dots, m\}$. Suppose that $x \leqslant m$. Then the pairs $\{j, x-j\}$, $1 \leqslant j < x/2$, are disjoint. If $S + S$ does not contain x, then S cannot contain both elements of any such pair, and hence $|Q \setminus S| \leqslant \lfloor x/2 \rfloor$. Therefore, $\lfloor x/2 \rfloor \leqslant \frac{m}{10}$, and so $x \leqslant \frac{m}{5} + 2$. A similar argument holds for $x \geqslant m$, with the conclusion now being that $2m - x \leqslant \frac{m}{5} + 2$. Thus, $S + S$ contains the progression $\frac{m}{5} + 2 < x < 2m - \frac{m}{5} - 2$. This is more than m elements if $m \geqslant 100$. ∎

Proposition 7.2 *Suppose that $A \subset [N, 2N)$ is a set of size at least $N/2$. Then*

$$\sqrt{2\sqrt{2\sqrt{2A}}}$$

contains a progression $\mathsf{P}(I; N^{1/8}, q)$ for some interval $I \subset [0.1, 10]$ with $|I| \gg 1$ and for some $q = O(1)$.

Proof Let $\eta > 0$ be a quantity to be specified later. Let $\Omega: \mathbf{N}^3 \to \mathbf{R}_+$ be the growth function appearing in the statement of Proposition 5.1. Apply Proposition 4.1 with this function. Let $q, d, \varepsilon, \theta, z, b$ be as in the conclusion of that proposition. Taking Y as in the statement of Proposition 5.1, Proposition 4.1 then tells us that

https://doi.org/10.4153/CJM-2017-036-1 Published online by Cambridge University Press

$Y' := (A + A) \cap Y = 2A \cap Y$ satisfies the hypotheses of Proposition 5.1. It follows that $2\sqrt{Y'}$, and hence $2\sqrt{2A}$, contains t^2 for all but at most $O(\eta\varepsilon q^{-1}N^{1/4})$ values of $t \in Q = \mathsf{P}([(2x)^{1/4} - \frac{\varepsilon}{100}, (2x)^{1/4} + \frac{\varepsilon}{100}]; N^{1/4}, q)$. Therefore, $\sqrt{2\sqrt{2A}}$ contains all but at most $O(\eta\varepsilon q^{-1}N^{1/4})$, and therefore at least $(1 - C\eta)|Q|$, of the elements of Q. If η is chosen suitably, this is at least $\frac{9}{10}|Q|$ elements of Q, and so by Lemma 7.1 we see that $2\sqrt{2\sqrt{2A}}$ contains a subprogression $Q' \subset Q$ of the form $Q' = \mathsf{P}(I; N^{1/4}, q)$ with $|I| \gg \varepsilon$. Finally, note that $\sqrt{Q'}$ contains a progression of the form $\mathsf{P}(I'; N^{1/8}, q)$ for some $I' \subset [0.1, 10]$ with $|I| \gg \varepsilon$. ■

We are finally ready to complete the proof of Theorem 1.1. Suppose we have a 2-colouring $V \cup W$ of all sufficiently large positive integers, with no monochromatic solution to $x + y = z^2$. Without loss of generality, there are infinitely many N such that $|V \cap [N, 2N)| \geqslant \frac{N}{2}$. Then we have the following chain of inclusions:

$$\sqrt{2V} \subset W,$$
$$\sqrt{2\sqrt{2V}} \subset \sqrt{2W} \subset V,$$
$$\sqrt{2\sqrt{2\sqrt{2V}}} \subset \sqrt{2V} \subset W.$$

It follows from Proposition 7.2 that W contains , for infinitely many N, a progression $\mathsf{P}(I_N; N^{1/8}, q_N)$, where $I_N \subset [0.1, 10]$, $|I_N| \gg 1$ and $q_N = O(1)$, both of these uniformly in N. By pigeonholing in the value of q_N, we can assume that $q_N = q$ does not depend on N. Moreover, taking $M = \lceil 10/\inf|I_N| \rceil$, we see that every I_N contains one of the finite collection of intervals $[\frac{i}{M}, \frac{i+1}{M}]$, $M/10 \leqslant i \leqslant 10M$. Therefore, we can pigeonhole in the choice of interval as well and assume that $I_N = I$ does not depend on N. Thus, W contains $\mathsf{P}(I; N^{1/8}, q)$ for some $I \subset [0, 1, 10]$ and for infinitely many N. Rescaling N, we see that W contains $\mathsf{P}([1, 1 + c]; N, q)$ for infinitely many N and for some $c > 0$.

From now on, this is the only consequence of the elaborate techniques of the earlier parts of the paper that we will require.

Using Proposition 6.1 as a tool, we find longer and longer progressions inside W. The following lemma formalizes this process.

Lemma 7.3 *Let $P_1 = \mathsf{P}([\alpha_1, \beta_1]; N, q)$ and $P_2 = \mathsf{P}([\alpha_2, \beta_2], N, q)$. Suppose that $\gamma_1 > \sqrt{\alpha_1^2 + \alpha_2^2}$ and that $\gamma_2 < \sqrt{\beta_1^2 + \beta_2^2}$. Then if N is large enough (depending on $\alpha_i, \beta_i, \gamma_i, q$), we have*

$$\mathsf{P}([\gamma_1, \gamma_2]; N, q) \subset \sqrt{P_1^2 + P_2^2 - P_1 - P_2}.$$

Remark Here and in what follows, A^2 means $\{a^2 : a \in A\}$ and *not* $a \cdot a' : a, a' \in A$ as one might find in other literature.

Proof Fix $\widetilde{\gamma}_1, \widetilde{\gamma}_2$ with $\gamma_1 > \widetilde{\gamma}_1 > \sqrt{\alpha_1^2 + \alpha_2^2}$ and $\gamma_2 < \widetilde{\gamma}_2 < \sqrt{\beta_1^2 + \beta_2^2}$. By Proposition 6.1, $P_1^2 + P_2^2$ has a point within $O(\sqrt{N})$ of every point of $\mathsf{P}([\widetilde{\gamma}_1^2, \widetilde{\gamma}_2^2]; N^2, q)$. Then $P_1 + P_2$ is a progression of length $\gg N$ consisting of multiples of q, and so it is easy to see that $P_1^2 + P_2^2 - P_1 - P_2$ contains all of $\mathsf{P}([\tilde{\gamma}_1^2, \tilde{\gamma}_2^2]; N^2, q)$ with the possible exception of points within $O(N)$ of the endpoints, and hence it contains $\mathsf{P}([\gamma_1, \gamma_2]; N^2, q)$. ■

Starting from the fact that

$$\mathsf{P}([1, 1+c]; N, q) \subset W \qquad \text{for infinitely many } N, \tag{7.1}$$

we apply Lemma 7.3 iteratively. Observe that if $n_1, n_2, n_3, n_4 \in W$, then $n_1^2 - n_3 \in V$, $n_2^2 - n_4 \in V$, and hence (if it is an integer)

$$\sqrt{n_1^2 + n_2^2 - n_3 - n_4} \in W.$$

Thus, if $P_1, P_2 \subset W$, then $\sqrt{P_1^2 + P_2^2 - P_1 - P_2} \subset W$. Using this observation and repeated applications of Lemma 7.3, we see that for any finite k and any choice of closed intervals $I_i \subset (\sqrt{i}, (1+c)\sqrt{i})$ there is an infinite sequence of Ns such that $\mathsf{P}(I_i; N, q) \subset W$ for $i = 1, 2, \dots, k$.

We claim that there is some $k = k(c)$ and some choice of $I_1, \dots, I_k$ such that $\bigcup_{i=1}^k I_i$ contains an interval of the form $[x, 3x]$. First note that if $i > 1/2c$, then $(1+c)\sqrt{i} > \sqrt{i+1}$, and so the intervals $(\sqrt{i}, (1+c)\sqrt{i})$ and $(\sqrt{i+1}, (1+c)\sqrt{i+1})$ overlap. Thus, if we set $i_0 := \lceil 1/2c \rceil$ and $i_1 := 9i_0$, then $\bigcup_{i_0 \leqslant i \leqslant i_1} (\sqrt{i}, (1+c)\sqrt{i})$ is an interval containing a subinterval of the form $[x, 3x]$.

Thus, W contains $\mathsf{P}([x, 3x]; N, q)$ for infinitely many N, and hence (replacing N by $\lfloor 1.1xN \rfloor$) we see that we have bootstrapped (7.1) to the stronger statement that

$$\mathsf{P}([1, 2]; N, q) \subset W \qquad \text{for infinitely many } N.$$

Pick one such $N = N_0$, sufficiently large. Thus,

$$\mathsf{P}\big([1, 2]; N_0, q\big) \subset W. \tag{7.2}$$

By Lemma 7.3 once more (and the inequalities $\sqrt{2} < \frac{3}{2} < \frac{5}{2} < \sqrt{8}$) we have

$$\mathsf{P}\Big(\Big[\frac{3}{2}, \frac{5}{2}\Big]; N_0, q\Big) \subset W.$$

Together with (7.2), this implies that

$$\mathsf{P}\big([1, 2]; N_0 + 1, q\big) \subset W.$$

Continuing inductively, we obtain

$$\bigcup_{N \geqslant N_0} \mathsf{P}\big([1, 2]; N, q\big) \subset W.$$

This implies that all sufficiently large multiples of q lie in W. But there are arbitrarily large multiples x, y, z of q satisfying $x + y = z^2$, and so at last we obtain a contradiction.

A Some Smooth Cutoff Functions

In the main body of the paper we required various smooth cutoff functions to (characteristic functions of) discrete intervals, balls in the torus $\mathbf{T}^d$, and Bohr sets. In this appendix we prove the existence of functions with required properties.

It is convenient to have a C^∞-function $f: \mathbf{R} \to [0, 1]$ with $\operatorname{Supp}(f) \subset [-1, 1]$ and $\int f(x)\,dx = 1$. Such a function can be constructed with a "trick", for example defining $f(x) = C \exp(\frac{1}{x^2 - 1})$ for an appropriate constant C (for a very elegant analysis of this,

https://doi.org/10.4153/CJM-2017-036-1 Published online by Cambridge University Press

see [1, Lemma 9]), or by convolving an infinite sequence of normalised characteristic functions of intervals $[-\ell_j, \ell_j]$ with $\sum_j \ell_j \leqslant 1$.

Let $g\colon \mathbf{R} \to \mathbf{R}$ be any compactly supported C^∞ function (for example, f). Then, since the M-th derivative $g^{(M)}$ is continuous and supported on $[-1,1]$, we have the bound $\|g^{(M)}\|_\infty = O_M(1)$. By integration by parts this leads to the standard bound

$$|\widehat{g}(\xi)| \ll_M \min(1, |\xi|^{-M}) \tag{A.1}$$

for $\xi \in \mathbf{R}$, where $\widehat{g}(\xi) = \int_{\mathbf{R}} g(x)e(-\xi x)dx$.

Lemma A.1 *Let $N \in \mathbf{N}$. There is a function $\psi = \psi_N\colon \mathbf{N} \to [0,\infty)$ with $\psi(n) = 1$ for $N \leqslant n < 2N$ and $\|\widehat{\psi}\|_1 = O(1)$ (uniformly in N), where the Fourier transform $\widehat{\psi}(t)$ is defined to be $\sum_n \psi(n)e(-tn)$ for $t \in \mathbf{T}$.*

Sketch of the proof Define first a function $g\colon \mathbf{R} \to \mathbf{R}$ via $g = 1_{[0,3]} * f$. It is easy to check that g is C^∞, compactly supported, and that $g(x) = 1$ for $x \in [1,2]$. We can then define $\psi(n) := g(n/N)$. By the Poisson summation formula we have

$$\widehat{\psi}(\theta) = N \sum_{k \in \mathbf{Z}} \widehat{g}(N(k+\theta)),$$

and so

$$\|\widehat{\psi}\|_1 \leqslant N \int_{\infty}^{\infty} |\widehat{g}(Nu)|du = \|\widehat{g}\|_1,$$

where the ℓ^1 norm on the right is taken on $\mathbf{R}$. The bound $\|\widehat{g}\|_1 = O(1)$ follows quickly by taking $M = 2$ in (A.1).

Alternatively, one can take ψ to be a de la Vallée Poussin type kernel as in the figure and proceed quite explicitly using the fact that this is a difference of two Fejér kernels. Details can be found in [8, Section 1.2]. ∎

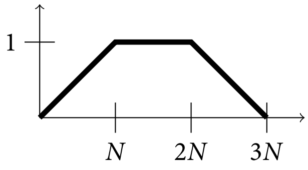

Figure 1: de la Vallée Poussin kernel.

Suppose now that $\varepsilon > 0$ and that $d \in \mathbf{N}$. Let us define $f_\varepsilon\colon \mathbf{T}^d \to [0,\infty)$ by $f_\varepsilon(x) = (2\varepsilon)^{-d}\prod_{i=1}^d f(\widetilde{x}_i/\varepsilon)$, where $\widetilde{x}$ is the unique element of $(-\frac{1}{2}, \frac{1}{2}]^d$ mapping to x under the natural projection. Note that $\int_{\mathbf{T}^d} f_\varepsilon(x)dx = 1$.

Lemma A.2 *There is a majorant ψ_ε^+ and a minorant ψ_ε^- to the ball $B_\varepsilon(0)$ in $\mathbf{T}^d$ satisfying*

(i) $\frac{1}{2} \leqslant (2\varepsilon)^{-d}\int_{\mathbf{T}^d} \psi_\varepsilon^\pm(t)dt \leqslant 2$ *and*

(ii) $\sum_{\mathbf{r} \in \mathbf{Z}^d \setminus \{0\}} |\widehat{\psi_\varepsilon^\pm}(\mathbf{r})| \|\mathbf{r}\|_1 = O_{\varepsilon,d}(1)$.

https://doi.org/10.4153/CJM-2017-036-1 Published online by Cambridge University Press

Proof We construct ψ_ε^+. The construction of ψ_ε^- is very similar and is left to the reader. Set $\varepsilon' := \varepsilon/10d$. For $x \in \mathbf{T}^d$, set

$$\psi_\epsilon^+(x) = 1_{B_{\varepsilon+\varepsilon'}(0)} * f_{\varepsilon'}(x) = \int_{\mathbf{T}^d} f_{\varepsilon'}(x-y)1_{B_{\varepsilon+\varepsilon'}(0)}(y)dy.$$

Since $f_{\varepsilon'}$ is supported on $B_{\varepsilon'}(0)$, $\psi_\varepsilon^+(x) = 1$ for $x \in B_\varepsilon(0)$, and in particular ψ_ε^+ is a majorant to the ball $B_\varepsilon(0)$.

Moreover, ψ_ε is bounded pointwise by 1 and is supported on $B_{\varepsilon+\varepsilon'}(0)$, whence

$$\int_{\mathbf{T}^d} \psi_\varepsilon^+(t)dt \leqslant \mu_{\mathbf{T}^d}(B_{\varepsilon+\varepsilon'}(0)) = (1+\frac{\varepsilon'}{\varepsilon})^d(2\varepsilon)^d \leqslant 2(2\varepsilon)^d.$$

Thus (i) is satisfied.

Next we turn to point (ii). Suppose that $\mathbf{r} \in \mathbf{Z}^d \smallsetminus \{0\}$. Write $\mathbf{r} = (r_1,\dots,r_d)$, and assume without loss of generality that $|r_1| = \|\mathbf{r}\|_\infty$. Performing M integration by parts in the integral

$$\widehat{\psi_\varepsilon^+}(\mathbf{r}) = \int_{\mathbf{T}^d} \psi_\varepsilon^+(x)e(-x\cdot\mathbf{r})dx$$

with respect to x_1, to get that

$$\widehat{\psi_\varepsilon^+}(\mathbf{r}) = \frac{1}{(-2\pi i r_1)^M}\int \frac{\partial^M\psi_\varepsilon^+(x)}{\partial x_1^M}e(-x\cdot\mathbf{r})dx \ll_{\varepsilon,d,M} \|\mathbf{r}\|_\infty^{-M}$$

for any $M \in \mathbf{N}$ (this is essentially the same bound as (A.1)). The ℓ^1 and ℓ^∞ norms of $\mathbf{r}$ are comparable up to factors of $O_d(1)$, and hence

$$\sum_{\mathbf{r}\in\mathbf{Z}^d\smallsetminus\{0\}} |\widehat{\psi_\varepsilon^+}(\mathbf{r})|\|\mathbf{r}\|_1 \ll_{\varepsilon,d,M} \sum_{\mathbf{r}\in\mathbf{Z}^d\smallsetminus\{0\}} \|\mathbf{r}\|_1^{1-M}.$$

Taking $M = d+2$, it is easy to see that the sum on the right converges and is bounded by $O_d(1)$. ■

Finally, we turn to the most complicated of our constructions, a smooth approximant for the Bohr-type set X considered in Section 5.

Lemma A.3 *Let $0 < \varepsilon' < \varepsilon < 1$, $d, q \in \mathbf{N}$, $x \in \mathbf{R}$, and $\theta, z \in \mathbf{T}^d$. Then there is an $A = A(\varepsilon,\varepsilon',d,q)$ with the following property. Suppose that N is sufficiently large in terms of $\varepsilon,\varepsilon',d,q,A$. Set*

$$X = \Big\{n\in\mathbf{N} : n \equiv u(\text{mod } q), \Big|\frac{n}{N}-x\Big|, \|\theta n - z\|_{\mathbf{T}^d} \leqslant \varepsilon\Big\},$$
$$X_- = \Big\{n\in\mathbf{N} : n \equiv u(\text{mod } q), \Big|\frac{n}{N}-x\Big|, \|\theta n - z\|_{\mathbf{T}^d} \leqslant \varepsilon-\varepsilon'\Big\}.$$

Suppose that $\varepsilon' < \varepsilon/10d$ and θ is (A,N)-irrational. Then there exists a function χ satisfying

(i) $1_{X_-}(n) \leqslant \chi(n) \leqslant 1_X(n)$ *for all n;*
(ii) $\|\widehat{\chi}\|_1 = O_{\varepsilon,\varepsilon',q,d}(1)$;
(iii) $\sum_n \chi(n) \geqslant \frac{1}{2}(2\varepsilon)^{d+1}q^{-1}N$.

Proof Let $g\colon \mathbf{R} \to [0,\infty)$ be a C^∞ function with $g(t) = 1$ for $|t-x| \leqslant \varepsilon-\varepsilon'$ and $g(t) = 0$ for $|t-x| > \varepsilon$. Such a function can be obtained by convolving the characteristic function of the interval $\{t : |t-x| \leqslant \varepsilon - \frac{1}{2}\varepsilon'\}$ with the function $\frac{2}{\varepsilon'}f(\frac{2t}{\varepsilon'})$.

https://doi.org/10.4153/CJM-2017-036-1 Published online by Cambridge University Press

Define a function $h\colon \mathbf{T}^d \to [0,\infty)$ by $h := f_{\varepsilon'/2} * 1_{B_{\varepsilon-\varepsilon'/2}(z)}$.

Now define

$$\chi(n) := g\Big(\frac{n}{N}\Big) h(\theta n) 1_{n\equiv u (\operatorname{mod} q)}.$$

The relevant support properties (i) can be easily checked. Turning to point (ii), we begin by noting the expansion

$$1_{n\equiv u(\operatorname{mod} q)} = q^{-1} \sum_{s\in\mathbf{Z}/q\mathbf{Z}} e\Big(\frac{(n-u)s}{q}\Big).$$

This implies that

$$\widehat{\chi}(t) = q^{-1} \sum_{s\in\mathbf{Z}/q\mathbf{Z}} e\Big(-\frac{us}{q}\Big) \widehat{g\Big(\frac{\cdot}{N}\Big) h(\theta\cdot)}\Big(t + \frac{s}{q}\Big). \tag{A.2}$$

Therefore, in order to establish (ii), it suffices to prove that

$$\|\widehat{g\Big(\frac{\cdot}{N}\Big) h(\theta\cdot)}\|_1 = O_{\varepsilon,\varepsilon',d}(1). \tag{A.3}$$

Fourier expanding h and applying Poisson summation, we have

$$\begin{aligned} \widehat{g\Big(\frac{\cdot}{N}\Big) h(\theta\cdot)}(t) &= \sum_n g\Big(\frac{n}{N}\Big) h(\theta n) e(-tn) \\ &= \sum_n g\Big(\frac{n}{N}\Big) \sum_{\mathbf{r}} \widehat{h}(\mathbf{r}) e\big((\mathbf{r}\cdot\theta - t)n\big) \\ &= N \sum_{\mathbf{r}} \widehat{h}(\mathbf{r}) \sum_{k\in\mathbf{Z}} \widehat{g}\big(N(t + k - \mathbf{r}\cdot\theta)\big). \end{aligned} \tag{A.4}$$

Thus,

$$\Big\|\widehat{g\Big(\frac{\cdot}{N}\Big) h(\theta\cdot)}\Big\|_1 \leqslant N \sum_{\mathbf{r}} |\widehat{h}(\mathbf{r})| \int_{-\infty}^{\infty} |\widehat{g}(Nu)| du = \|\widehat{g}\|_1 \|\widehat{h}\|_1,$$

where the ℓ^1 norms are on $\mathbf{Z}^d$ and $\mathbf{R}$, respectively.

That $\|\widehat{g}\|_1 \ll_{\varepsilon,\varepsilon'} 1$ follows immediately from (A.1) with $M = 2$.

By essentially the same reasoning used in the proof of Lemma A.2, we have

$$|\widehat{h}(\mathbf{r})| \ll_{\varepsilon,\varepsilon',d,M} \|\mathbf{r}\|_\infty^{-M}. \tag{A.5}$$

Taking $M = d+1$, we obtain $\|\widehat{h}\|_1 = O_{\varepsilon,\varepsilon',d}(1)$. Putting these facts together completes the proof of (A.3) and hence of (ii).

It remains to verify (iii). Note that we have not yet used the irrationality of θ. From (A.2) we have

$$\sum_n \chi(n) = \widehat{\chi}(0) = q^{-1} \sum_{s\in\mathbf{Z}/q\mathbf{Z}} e\Big(-\frac{us}{q}\Big) \widehat{g\Big(\frac{\cdot}{N}\Big) h(\theta\cdot)}\Big(\frac{s}{q}\Big).$$

By (A.4), it follows that

$$\sum_n \chi(n) = Nq^{-1} \sum_{\mathbf{r}\in\mathbf{Z}^d} \sum_{s\in\mathbf{Z}/q\mathbf{Z}} \sum_{k\in\mathbf{Z}} e\Big(-\frac{us}{q}\Big) \widehat{h}(\mathbf{r}) \widehat{g}\Big(N\Big(\frac{s}{q} + k - \mathbf{r}\cdot\theta\Big)\Big). \tag{A.6}$$

The contribution from $\mathbf{r} = 0$, $s = 0$, $k = 0$ is $Nq^{-1}(\int_{\mathbf{T}^d} h)(\int_{\mathbf{R}} g)$. Since $\varepsilon' < \varepsilon/10d$, we have $\int_{\mathbf{T}^d} h \geqslant \mu_{\mathbf{T}^d}(B_{\varepsilon-\varepsilon'}(0)) \geqslant 0.9(2\varepsilon)^d$, and evidently $\int_{\mathbf{R}} g \geqslant 2(\varepsilon - \varepsilon') > 0.9(2\varepsilon)$. Thus, the contribution from this term is $\geqslant \frac{3}{4}(2\varepsilon)^{d+1} q^{-1} N$. To complete the proof of

(iii) it suffices to show that the contribution of the other terms to (A.6) is at most $\frac{1}{4}(2\varepsilon)^{d+1}q^{-1}N$, to which end it is enough to show that

$$\sum_{\mathbf{r}\in\mathbf{Z}^d}\sum_{s\in\mathbf{Z}/q\mathbf{Z}}\sum_{k\in\mathbf{Z}}|\widehat{h}(\mathbf{r})||\widehat{g}\Big(N\Big(\frac{s}{q}+k-\mathbf{r}\cdot\theta\Big)\Big)|\leqslant\frac{1}{4}(2\varepsilon)^{d+1}, \tag{A.7}$$

where the sum omits the term $\mathbf{r}=0$, $s=0$, $k=0$.

By (A.5) (with $M=d+1$) and (A.1) (with $M=2$), the left-hand side is bounded by

$$O_{\varepsilon,\varepsilon',d}(1)\sum_{\mathbf{r}\in\mathbf{Z}^d}\sum_{s\in\mathbf{Z}/q\mathbf{Z}}\sum_{k\in\mathbf{Z}}\min(1,\|\mathbf{r}\|^{-d-1})\min\Big(1,N^{-2}|k+\frac{s}{q}-\mathbf{r}\cdot\theta|^{-2}\Big). \tag{A.8}$$

If $0<\|\mathbf{r}\|_1\leqslant A/q$, then it follows from the fact that θ is (A,N)-irrational that $|k+\frac{s}{q}-\mathbf{r}\theta|\geqslant\frac{A}{qN}$ (no matter the value of s or k). The same is trivially true when $\mathbf{r}=0$, provided that not both of s,k are zero and that N is sufficiently large. In the inner sum over k in (A.8), the contribution from all but at most one term is

$$\ll N^{-2}\sum_{m\in\mathbf{Z}\setminus\{0\}}|m|^{-2}\ll N^{-2},$$

and so when $\|\mathbf{r}\|_1\leqslant A/q$ the inner sum over k is $\ll\frac{q^2}{A^2}+N^{-2}$, which is $\ll q^2/A^2$ if N is big enough. Therefore,

$$\begin{aligned}\sum_{\substack{\mathbf{r}\in\mathbf{Z}^d\\ \|\mathbf{r}\|\leqslant A/q}}\sum_{s\in\mathbf{Z}/q\mathbf{Z}}\sum_{k\in\mathbf{Z}}&\min(1,\|\mathbf{r}\|^{-d-1})\min(1,N^{-2}|k+\frac{s}{q}-\mathbf{r}\cdot\theta|^{-2})\\ &\ll\frac{q^3}{A^2}\sum_{\mathbf{r}}\|\mathbf{r}\|^{-d-1}\ll_{d,q}A^{-2}.\end{aligned}$$

All other terms in (A.8) have $\|\mathbf{r}\|\geqslant\frac{A}{q}$. Using the trivial bound

$$\sum_{k\in\mathbf{Z}}\min(1,N^{-2}|k+\frac{s}{q}-\mathbf{r}\cdot\theta|^{-2})\ll 1,$$

the contribution from these is bounded by

$$O_{d,\varepsilon,\varepsilon',q}(1)\sum_{\|\mathbf{r}\|\geqslant A/q}\|\mathbf{r}\|^{-d-1}\ll_{d,\varepsilon,\varepsilon',q}A^{-1}.$$

Putting all of this together shows that (A.8) is bounded by $O_{d,\varepsilon,\varepsilon,q}(A^{-1})$, and so (A.7) does indeed hold if A is large enough as a function of $\varepsilon,\varepsilon',d,q$. ■

References

[1] E. Bombieri, J. B. Friedlander, and H. Iwaniec, *Primes in arithmetic progressions to large moduli. II.* Math. Ann. 277(1987), no. 3, 361–393. http://dx.doi.org/10.1007/BF01458321

[2] S. Eberhard, *The abelian arithmetic regularity lemma.* arxiv:1606.09303

[3] S. Eberhard, B. Green and F. Manners, *Sets of integers with no large sum-free subset,* Ann. of Math. (2) **180** (2014), no. 2, 621–652. http://dx.doi.org/10.4007/annals.2014.180.2.5

[4] B. Green, *A Szemerédi-type regularity lemma in abelian groups, with applications.* Geom. Funct. Anal. 15(2005), no. 2, 340–376. http://dx.doi.org/10.4007/annals.2012.175.2.2

[5] B. Green and T. Tao, *An arithmetic regularity lemma, an associated counting lemma, and applications.* In: An irregular mind, Bolyai Soc. Math. Stud., 21, János Bolyai Math. Soc., Budapest, 2010, pp. 261–334. http://dx.doi.org/10.1007/978-3-642-14444-8_7

[6] ———, *The quantitative behaviour of polynomial orbits on nilmanifolds.* Ann. of Math. (2) 175(2012), no. 2, 465–540. http://dx.doi.org/10.4007/annals.2012.175.2.2

https://doi.org/10.4153/CJM-2017-036-1 Published online by Cambridge University Press

[7] K. Gyarmati, P. Csikvári, and A. Sárközy, *Density and Ramsey type results on algebraic equations with restricted solution sets.* Combinatorica **32**(2012), 425–449. http://dx.doi.org/10.1007/s00493-012-2697-9

[8] Y. Katznelson, *An introduction to harmonic analysis.* Second Ed., Dover Publications, Inc., New York, 1976.

[9] A. Khalfallah and E. Szemerédi, *On the number of monochromatic solutions of $x + y = z^2$.* Combin. Probab. Comput. **15**(2006), no. 1–2, 213–227. http://dx.doi.org/10.1017/S0963548305007169

[10] J. C. Lagarias, A. M. Odlyzko, and J. B. Shearer, *On the density of sequences of integers the sum of no two of which is a square. I. Arithmetic progressions.* J. Combin. Theory Ser. A **33**(1982), no. 2, 167–185. http://dx.doi.org/10.1016/0097-3165(82)90005-X

[11] ———, *On the density of sequences of integers the sum of no two of which is a square. II. General sequences.* J. Combin. Theory Ser. A **34**(1983), no. 2, 123–139. http://dx.doi.org/10.1016/0097-3165(83)90051-1

[12] S. Lindqvist, *Partition regularity of generalised Fermat equations.* arxiv:1606.07334

[13] ———, *Monochromatic solutions to $x + y$ a square in $\mathbf{Z}/q\mathbf{Z}$,* http://people.maths.ox.ac.uk/lindqvist/notes/xysumsquare.pdf

[14] H. L. Montgomery, *Ten lectures on the interface between analytic number theory and harmonic analysis.* CBMS Regional Conference Series in Mathematics, 84, Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1994. http://dx.doi.org/10.1090/cbms/084

[15] R. C. Vaughan, *The Hardy-Littlewood method.* Cambridge Tracts in Mathematics, 125, Cambridge University Press, Cambridge, 1997. http://dx.doi.org/10.1017/CBO9780511470929

[16] T. D. Wooley, *On Diophantine inequalities: Freeman's asymptotic formulae.* In: Proceedings of the Session in Analytic Number Theory and Diophantine Equations, Bonner Math. Schriften, 360, Univ. Bonn, Bonn, 2003.

Mathematical Institute, Radcliffe Observatory Quarter, Woodstock Rd, Oxford OX2 6GG
e-mail: ben.green@maths.ox.ac.uk lindqvist.sofia@gmail.com

https://doi.org/10.4153/CJM-2017-036-1 Published online by Cambridge University Press