A Structure Theorem for Positive Density Sets
Having the Minimal Number of 3-Term
Arithmetic Progressions

Ernie Croot
February 17, 2022

Abstract
Assuming the well known conjecture that for any $\gamma > 0$ and $x$ sufficiently large the interval $[x, x + x^\gamma]$ always contains a prime number, we prove the following unexpected result: There exist numbers $0 < \rho < 1$ arbitrarily close to 0, and arbitrarily large primes $q$, such that if $S$ is any subset of $\mathbb{Z}/q\mathbb{Z}$ of density at least $\rho$, having the least number of 3-term arithmetic progressions among all such sets $S$ (of density $\geq \rho$), then there exists an integer $1 \leq b \leq q - 1$ and a real number $0 < d < 1$ (depending only on $\rho$) such that

$$|S \cap (S + bj)| = |S| \left(1 - O\left(\frac{1}{|\log \rho|}\right)\right), \text{ for every } 0 \leq j < q^d.$$

This result says that $S$ is “nearly translation invariant” in a very strong sense.

A curious feature of the proof is that F. A. Behrend’s result on large subsets of $\{1, 2, ..., x\}$ containing no 3-term arithmetic progressions is a key ingredient. The proof also uses exponential sums, a result of K. F. Roth, as well as a result of P. Varnavides.

1 Introduction
Given a subset $S$ of $\mathbb{Z}/q\mathbb{Z}$, we will say that $S$ has density $\rho$ if $|S| = \rho q$. A well known result of K. F. Roth [2] asserts that for $q$ sufficiently large, and
for \((\log \log q)^{-1} < \rho \leq 1\), any subset of \(\mathbb{Z}/q\mathbb{Z}\) having density at least \(\rho\), must contain a non-trivial three term arithmetic progression; that is, a triple of numbers, \(a, b, c\), \(a \not\equiv b \pmod{q}\), satisfying
\[
a + b \equiv 2c \pmod{q}.
\]

(1)

E. Szemerédi \cite{6}, R. Heath-Brown \cite{4}, and J. Bourgain \cite{2} have improved considerably on Roth’s result, with the Bourgain’s work being the most recent:

**Theorem 1 (J. Bourgain)** For \(q\) sufficiently large and

\[
B \sqrt{\frac{\log \log q}{\log q}} < \rho \leq 1, \text{ where } B > 0 \text{ is some constant},
\]

any subset of \(\mathbb{Z}/q\mathbb{Z}\) of density \(\rho\) contains a triple \(a, b, c\) (with \(a \neq b \neq c\)) satisfying \((1)\).

As a consequence of Roth’s result mentioned above, Varnavides \cite{7} proved the following

**Theorem 2 (P. Varnavides)** Given \(0 < \rho \leq 1\), and a sufficiently large prime \(q\), if \(S\) is any subset of \(\mathbb{Z}/q\mathbb{Z}\) of density \(\rho\), there exists \(\kappa > 0\) so that
\[
\mu_q(S) = \frac{\# \{a, b, c \in S : a + b \equiv 2c \pmod{q}\}}{q^2} > \kappa.
\]

In fact, one can take
\[
\kappa = \frac{\rho}{16h(\rho/2)^2},
\]
where \(h(\rho)\) is defined to be the least integer such that if \(m \geq h(\rho)\) and \(T\) is any subset of \(\{1, ..., m\}\) with density at least \(\rho\), then \(T\) contains a three term arithmetic progression.

The version of the theorem appearing in \cite{7} does not give such an explicit value for \(\kappa\), and the result is stated in terms of subsets of \(\{1, 2, ..., x\}\), rather
than \( \mathbb{Z}/q\mathbb{Z} \). For these reasons, we give our own proof of this result in section 3.

Although the function \( \mu_q(S) \) counts both non-trivial and trivial solutions to \( a + b \equiv 2c \pmod{q} \) (trivial means \( a \equiv b \equiv c \pmod{q} \)), if \( q \) is large enough in terms of \( \rho \), we can deduce that

\[
\frac{\# \{a, b, c \in S, a \neq b \pmod{q} : a + b \equiv 2c \pmod{q}\}}{q^2} > \kappa, \tag{2}
\]

since there can be at most \( |S| < q \) triples \( a, b, c \in S \) with \( a \equiv b \equiv c \pmod{q} \); and so, the contribution of these trivial solutions to \( \mu_q(S) \) is thus only \( O(1/q) \), which tends to 0 as \( q \) tends to infinity, which thus proves (2).

Combining Varnavides's and Bourgain's results, one can show that if \( S \) has density \( \geq \rho \) modulo \( q \), and if \( q \) is a sufficiently large prime, then

\[
\mu_q(S) \geq \exp \left( -D \frac{|\log \rho|}{\rho^2} \right), \tag{3}
\]

where \( D > 0 \) is an absolute constant.

It is of interest to try to find, for a given density \( 0 < \rho \leq 1 \), the smallest value of \( \kappa \) so that the conclusion of Varnavides's theorem still holds. This smallest \( \kappa \) is given by the following function:

\[
r_q(\rho) = \min_{S \subseteq \mathbb{Z}/q\mathbb{Z}, |S| \geq \rho q} \mu_q(S).
\]

One approach to understanding \( r_q(\rho) \) is to try to understand the structure of critical sets, which are subsets \( S \) of \( \mathbb{Z}/q\mathbb{Z} \) of density \( \rho \) and

\[
\mu_q(S) = r_q(\rho).
\]

It would seem that the problem of understanding the structure of these critical sets is a much more difficult problem than that of understanding the behavior of \( r_q(\rho) \); and so, it would seem that we should try to make progress on the order of growth of \( r_q(\rho) \) by some other method first before trying to tackle questions about such sets. Even so, one might would think that they have very little structure; however, motivated by the main result of this paper (theorem 3), we have the following conjecture, which claims, on the contrary, that these sets have a considerable amount of additive structure:
Conjecture 1 Given $0 < \rho < 1$, and $q$ sufficiently large, there exists $0 < d < 1$ such that if $S$ is a critical set of $\mathbb{Z}/q\mathbb{Z}$ of density $\rho$, then there exists a number $1 \leq b \leq q - 1$ such that

$$|S \cap (S + jb)| > (1 - n(\rho))|S|,$$

for every $0 \leq j < q^d$,

where $n(\rho)$ is a function of $\rho$ only, and which tends to 0 as $\rho$ tends to 0.

The main theorem of the paper is a weakened version of this conjecture. Before we state it, first define

$$r(\rho) = \liminf_{q \geq 3 \text{ prime}} r_q(\rho).$$

We will need the following, famous conjecture from prime number theory:

Conjecture 2 Given $\theta > 0$, we have for all $x$ sufficiently large that $[x, x+x^\theta]$ contains a prime number.

From [1] the conjecture is known to hold for all $\theta > 0.525$.

Our main theorem is as follows:

Theorem 3 Assume conjecture [2] holds. For any $0 < \rho_0 < 1$, there exist numbers $0 < \rho_1 < \rho_0$ and $0 < d < 1$, and infinitely many primes $q$ such that the following holds: If $S \subseteq \mathbb{Z}/q\mathbb{Z}$ has density $\geq \rho_1$, and has the least number of 3-term arithemtic progressions modulo $q$ among all sets of density $\geq \rho_1$ (that is, $\mu_q(S) = r_q(\rho_1)$), then there exists a number $1 \leq b \leq q - 1$ such that

$$|S \cap (S + bj)| \geq |S| \left(1 - O \left(\frac{1}{\log \rho_1}\right)\right),$$

for every $0 \leq j \leq q^d$,

where the implied constant in the big-O is absolute.

If we could prove the following conjecture on the behavior of $r_q(\rho)$, then we could remove the assumption that Conjecture [2] holds in the above theorem:

---

1At the end of the proof of Theorem 3 in section 2 we show how the truth of this conjecture implies this stronger version of our main theorem.
Conjecture 3 There exists a constant $A > 1$ such that for every $0 < \rho \leq 1$, every $0 < \theta < 1$, and $n$ sufficiently large,

$$\frac{r_q(\rho)}{r_n(\rho)} \in \left[\frac{1}{A}, A\right],$$

for every $q \in [n, n + n^\theta]$.

We now briefly mention the main ideas and ingredients in the proof of Theorem 3 by describing how to prove a certain toy version of the theorem: For a given set of integers $T$ having density at least $\rho$ modulo $q$, define the exponential sum

$$f_T(t) = \sum_{s \in T} e(st),$$

where $e(u) = \exp(2\pi i u)$. Using the identity

$$\frac{1}{q} \sum_{j=0}^{q-1} e\left(\frac{ja}{q}\right) = \begin{cases} 1, & \text{if } q \text{ divides } a; \\ 0, & \text{if } q \text{ does not divide } a, \end{cases}$$

one can easily show that

$$q^2 \mu_q(T) = \#\{a, b, c \in T : a + b \equiv 2c \pmod{q}\} = \frac{1}{q} \sum_{|a| < q/2} f_T\left(\frac{a}{q}\right)^2 f_T\left(-\frac{2a}{q}\right). \quad (4)$$

It turns out that to estimate the sum on the right to within an error of size less than $\epsilon|T|q < \epsilon q^2$, we need only sum over values $a$ in a set with at most $(\rho\epsilon^2)^{-1}$ elements. Now, given a critical set $S$ of density $\rho$, we show that we can multiply the set by an integer (not divisible by $q$), to produce a new set $W$ modulo $q$, where we think of $W$ as being a subset of $\{0, 1, ..., q-1\}$, such that if

$$\left|f_W\left(\frac{a}{q}\right)\right| \geq \epsilon|W|, \left|\frac{a}{q}\right| < \frac{1}{2},$$

then

$$\left|\frac{a}{q}\right| < q^{-\epsilon^2}, \text{ and } a \text{ is divisible by } 4. \quad (5)$$

Now suppose that $p$ is a prime number very close to $q/2$. As a consequence of (4), we show that the sum on the right-hand-side of (4) is, for small values
of $\epsilon$,

$$q^2 \mu_\rho(T) \approx \frac{1}{q} \sum_{|a|<p/2} f_W \left( \frac{a}{p} \right)^2 f_W \left( \frac{-2a}{p} \right) = \frac{p}{q} \# \{ a, b, c \in W : a + b \equiv 2c \pmod{p} \}, \quad (6)$$

where, in this context, we will use the notation $u(x) \approx v(x)$ to loosely mean that the functions $u(x)$ is “extremely close” to $v(x)$; and, as $\epsilon \to 0$ and $q \to \infty$, the ratio of these two functions tends to 1.

Next, we show that for certain special values of $\rho$, the proportion of residue classes modulo $p$ occupied by $W$ is very nearly $\rho$; so, most residue classes modulo $p$ contain either 0 or 2 elements of $W$. A consequence of this is that if $V_0$ is the set of all elements of $W$ that are in $[0, p-1]$, and if $V_1$ is the set of all elements of $W$ in $[p, q-1]$, then

$$|V_0 \cap (V_1 - p)| \geq |V_0| \left( 1 - O \left( \frac{1}{|\log \rho|} \right) \right).$$

Using an additional trick that involves rotating the set $W$ modulo $q$ (translating by an integer $k$), we can show that

$$|W \cap (W - p)| \geq |W| \left( 1 - O \left( \frac{1}{|\log \rho|} \right) \right),$$

and so there exists an integer $b$, not divisible by $q$, so that

$$|S \cap (S + b)| \geq |S| \left( 1 - O \left( \frac{1}{|\log \rho|} \right) \right),$$

since the elements of $S$ are multiples of elements of $W$ modulo $q$. The additional parameter $j$ in the intersection $S \cap (S + jb)$ stated in theorem 3 comes about via only a slight generalization of the above argument.

The way we deduce that the proportion of residue classes modulo $p$ occupied by $W$ is nearly the same as the proportion modulo $q$ is as follows. First, we need the following result of F. A. Behrend:

**Theorem 4** For sufficiently large $q$ there exists a subset $S$ of $\mathbb{Z}/q\mathbb{Z}$ with

$$|S| > \frac{q}{\exp(C \sqrt{\log q})},$$

for some constant $C > 0$,

such that $S$ contains no three term arithmetic progressions modulo $q$. 6
As a consequence of this theorem, we show that

**Corollary 1** Given $0 < \rho < 1$, we have that

$$r(\rho) \leq \left( -\frac{1}{C^2} \log^2(4\rho) \right).$$

So, Behrend’s result shows that the function $r(\rho)$ decays quite rapidly compared with $\rho$.

Now, $W$ must occupy at least $\rho$ of the residue classes modulo $p$. As a consequence of the above corollary, one can show that for certain values of $\rho$, if $W$ had density more than

$$\rho \left( 1 + \frac{G}{|\log \rho|} \right),$$

modulo $p$, where $G > 0$ is some particular constant, then the number of solutions $a + b \equiv 2c \pmod{p}$, $a, b, c \in W$, would have to be a large multiple of the number of such solutions modulo $q$, and so we would have that (6) could not hold.

The rest of the paper is organized as follows. In the next section, we give a proof of theorem 3 as well as an indication of how conjecture 3 can be used to replace the assumption that conjecture 2 holds. In several of the sections after the proof of the main theorem, we prove several propositions that are used in the proof of theorem 3 as well as in the proofs of other propositions. Finally, in section 7, we state and prove several technical lemmas and corollaries that appear throughout the paper.

## 2 Proof of Theorem 3.

We first pin down the value of $\rho_1$ for which we will prove that $S$ satisfies the conclusion of the theorem, and the following proposition gives us the answer we seek:

---

2 In my terminology, “proposition” is synonymous with “large lemma” or perhaps “meta-theorem”; that is, a proposition is a result whose proof is either too large, or is too technical to be considered a lemma, yet is not sufficiently general or interesting to be considered a theorem.
Proposition 1 Suppose that $0 < \rho_0 < 1$ and $k > 1$. Then, there exists a number $0 < \rho_1 < \rho_0$ such that for all

$$\rho > \rho_1 \left( 1 + \frac{2C^2 \log k}{\log \rho_1} \right),$$

we will have

$$r(\rho) > k \cdot r(\rho_1),$$

where $C$ is the same constant that appears in Theorem 4 mentioned in the introduction.

The proof of this proposition appears in section 4.

We will assume henceforth that $\rho_1$ is any number satisfying the conclusion of this proposition for $k = 100$. Then, for $0 < \epsilon < 1$, to be chosen later, let $q$ be any sufficiently large prime so that

$$r(\rho_1) > \frac{r_q(\rho_1)}{2}, \quad (7)$$

and so that for every prime

$$p \in \left[ \frac{q}{2}, \frac{q}{2} + q^{\rho_1 \epsilon / 2} \right]$$

we have

$$r_p \left( \rho_1 \left( 1 + \frac{2C^2 \log 100}{\log \rho_1} \right) \right) > \frac{1}{2} \cdot r \left( \rho_1 \left( 1 + \frac{2C^2 \log 100}{\log \rho_1} \right) \right). \quad (8)$$

Note: Since we have assumed conjecture 2, we have that the interval above always contains a prime for sufficiently large $q$. We also note that every sufficiently large prime $p$ will satisfy (8), and infinitely many primes $q$ will satisfy (7). Thus, we will have that both (7) and (8) hold for infinitely many primes $q$.

Define the exponential sum

$$f_s(t) = \sum_{s \in S} e(st),$$
where \( e(u) = \exp(2\pi iu) \). We first claim that

\[
\# \left\{ a \in \mathbb{Z}, \ |a| < \frac{q}{2} : \ |f_s \left( \frac{a}{q} \right)| > \epsilon |S| \right\} \leq \frac{1}{\rho_1 \epsilon^2}.
\]

To see this, we have from Parseval’s identity (lemma 3 in section 7) that

\[
\sum_{|a|<q/2} \left| f_s \left( \frac{a}{q} \right) \right|^2 = q |S|.
\]

Now, if there were more than \((\rho_1 \epsilon^2)^{-1}\) values of \(a\) for which \(|f_s(a/q)| > \epsilon |S|\), then the sum on the left hand side would exceed \(|S|^2/\rho = q |S|\), and so couldn’t equal \(q |S|\) as claimed.

We next need the following lemma:

**Lemma 1** There exists an integer \(1 \leq h \leq q - 1\) such that the set

\[
S' = hS = \{ hs \pmod{q} : s \in S \}
\]

has the property that

\[
\left| f_{S'} \left( \frac{a}{q} \right) \right| > \epsilon |S|, \ |a| \leq q/2 \implies |a| < q^{1-\rho_1 \epsilon^2}. \tag{9}
\]

The proof of this lemma makes use of the pigeonhole principle and can be found in section 7. Now, for integers \(k\) and \(1 \leq v < q^{\rho_1 \epsilon^2/4}\), to be chosen later, let

\[
W = W(k, v) = (4v)^{-1}S' + k = \{ (4v)^{-1}s + k \pmod{q} : s \in S' \}
= \{ (4v)^{-1}hs + k \pmod{q} : s \in S \}. \tag{10}
\]

We note that the number of solutions to \(a + b \equiv 2c \pmod{q}\), \(a, b, c \in W\), is the same as the number of solutions among the elements of \(S\); also, one can easily see that for \(|a| \leq q/2\),

\[
\left| f_{S'} \left( \frac{(4v)^{-1}a}{q} \right) \right| = \left| f_W \left( \frac{a}{q} \right) \right| > \epsilon |S| \implies |a| < 4v q^{1-\rho_1 \epsilon^2}, \text{ and } 4v|a|. \tag{11}
\]

We now require the following proposition:
Proposition 2 Suppose \( p \in [q/2, q/2 + q^{\rho_1^2/2}] \) and that \( 1 \leq v \leq q^{\rho_1^2/4} \) is an integer. Then, for \( W = W(k, v) \), we have

\[
\mu_q(W) = \frac{p^3}{q^3} \mu_p(W) + O \left( \left( \frac{\epsilon}{\rho_1} \right)^{1/3} \rho_1^2 \right),
\]

where the constant in the big-O is absolute.

The proof of this result can be found in section 5.

Now suppose that \( R(k, v) \) is the set of residue classes modulo \( p \) that are occupied by \( W(k, v) \). If

\[
|R(k, v)| > \rho_1 \left( 1 + \frac{2C^2 \log 100}{|\log \rho_1|} \right) p,
\]

then we would have from (7), (8), and proposition 2 that

\[
r(\rho_1) > \frac{r_q(\rho_1)}{2} = \frac{\mu_q(W)}{2} = \frac{p^3 \mu_p(W)}{2q^3} + O \left( \left( \frac{\epsilon}{\rho_1} \right)^{1/3} \rho_1^2 \right) \\
\geq \frac{p^3}{2q^3} \rho_1 \left( 1 + \frac{2C^2 \log 100}{|\log \rho_1|} \right) + O \left( \left( \frac{\epsilon}{\rho_1} \right)^{1/3} \rho_1^2 \right) \\
> \frac{p^3}{4q^3} r(\rho_1) \left( 1 + \frac{2C^2 \log 100}{|\log \rho_1|} \right) + O \left( \left( \frac{\epsilon}{\rho_1} \right)^{1/3} \rho_1^2 \right) \\
> \frac{100p^3}{4q^3} r(\rho_1) + O \left( \left( \frac{\epsilon}{\rho_1} \right)^{1/3} \rho_1^2 \right),
\]

which is impossible once \( \epsilon > 0 \) is small enough. Thus, we conclude

\[
\frac{|S|}{2} \leq |R(k, v)| \leq \rho_1 \left( 1 + \frac{2C^2 \log 100}{|\log \rho_1|} \right) p \\
= |S| \left( \frac{1}{2} + O \left( \frac{1}{|\log \rho_1|} \right) \right);
\]

and so,

\[
|R(k, v)| = |S| \left( \frac{1}{2} + O \left( \frac{1}{|\log \rho_1|} \right) \right). \tag{12}
\]
This tells us that a typical residue class modulo $p$ in $R(k, v)$ contains either 0 or 2 elements of $W(k, v)$; more precisely, it says that the number of progressions modulo $p$ containing only one element of $W(k, v)$ is $O(|S|/|\log \rho_1|)$. Thus, if we let $W_0(k, v)$ denote the integers in $W(k, v)$ that lie in $[0, p-1]$, and if we let $W_1(k, v)$ denote the integers in $W(k, v)$ that lie in $[p, q-1]$, then our equation for $R(k, v)$ says that

$$|W_0(k, v) \cap (W_1(k, v) - p)| = |S| \left(\frac{1}{2} - O\left(\frac{1}{|\log \rho_1|}\right)\right).$$  \hspace{1cm} (13)

We will now show that

$$|W(0, v) \cap W(-p, v)| \geq |S| \left(1 - O\left(\frac{1}{|\log \rho_1|}\right)\right).$$  \hspace{1cm} (14)

Here is the proof: First, suppose that $w \in W_0(0, v) \cap (W_1(0, v) - p)$, and note that $0 \leq w \leq p - 1$. For such $w$ we will have $w \in W(0, v) \cap W(-p, v)$. Next, suppose that $w - p \in W_0(-p, v) \cap (W_1(-p, v) - p)$, and note that $w \geq p$. For these integers $w$ we will also have that $w \in W(0, v) \cap W(-p, v)$. Since the two sets of integers $w$ considered are disjoint, we must have

$$|W(0, v) \cap W(-p, v)| \geq |W_0(0, v) \cap (W_1(0, v) - p)| + |W_0(-p, v) \cap (W_1(-p, v) - p)|,$$

and so (14) follows from this inequality and (13).

From this it follows that

$$|S \cap (S - (4v)h^{-1}p)| = |W(0, v) \cap W(-p, v)| \geq |S| \left(1 - O\left(\frac{1}{|\log \rho_1|}\right)\right),$$

which proves the theorem, since we can take $v$ to be any integer in $[1, q^{\rho_1 c^2/4}]$. Note that here we are thinking of the set $S - (4v)h^{-1}p$ as a set of integers in $\{0, 1, ..., q - 1\}$.

The reason that conjecture 3 allows us to remove the assumption that conjecture 2 holds is as follows: The proof of the above theorem actually shows more than is stated. It shows that if $S$ satisfies the hypotheses of the theorem, and if $q$ is a prime satisfying

$$r(\rho_1) > \frac{r_q(\rho_1)}{2},$$

11
and if \( n \in (q/2, q/2 + q^{\rho_1^2/2}) \) is some integer all of whose non-trivial divisors are greater than \( q^{\rho_1^2/4} \), and which satisfies

\[
  r_n(\rho_2) > F r(\rho_2), \text{ where } 0 < F \leq 1 \text{ does not depend on } \rho_2, \tag{15}
\]

where

\[
  \rho_2 = \rho_1 \left( 1 + \frac{2C^2 \log 100}{|\log \rho_1|} \right),
\]

then the conclusion of the theorem holds if \( q \) is sufficiently large. We know that this last condition (15) holds for all sufficiently large primes \( n \), but it is not clear that it holds for all sufficiently large integers.

Now, if we assume conjecture 3, then for every \( 0 < \rho < 1 \) we will have that there exists \( A \geq 1 \) so that

\[
  r(\rho) = \liminf_{q \geq 3} r_q(\rho) \leq A \liminf_{n \geq 3} r_n(\rho). \tag{16}
\]

To see this, we first note from the result in [1], the interval \([n, n + n^{0.53}]\) contains a prime number. For each such \( n \), let \( p_n \) denote a prime in this interval. Then, from conjecture 3 we get that there exists an \( A > 0 \) (which does not depend on \( n \) or \( \rho_1 \)), such that

\[
  r_{p_n}(\rho) \leq A r_n(\rho),
\]

and this implies the inequality (16) above.

From (16) one now sees that (15) holds for all sufficiently large integers \( n \) (prime or not); and so, our assertion that conjecture 3 can be used to remove the assumption conjecture 2 follows.

### 3 Proof of Theorem 2

Suppose \( q > h(\rho/2) \). Set \( k = \lfloor h(\rho/2) \rfloor + 1 \), and let \( P \) be the set of all \( k \)-term arithmetic progressions \( a, a + d, a + 2d, \ldots, a + (k - 1)d \) modulo \( q \), \( d \) is not \( 0 \) modulo \( q \). We treat the progression \( a + (k - 1)d, \ldots, a \) as distinct from \( a, \ldots, a + (k - 1)d \). Clearly,

\[
  |P| = q(q - 1),
\]
since there are \(q\) choices for \(a\) and \(q - 1\) choices for \(d\).

Given a progression \(h \in P\), we let

\[
R(h) = \frac{|h \cap S|}{k}.
\]

If \(R(h) \geq \rho/2\), then \(h \cap S\) must contain a non-trivial 3-term arithmetic progression, from the way we have defined \(h(\rho)\) (and \(h(\rho/2)\)). When \(q\) is sufficiently large, such a progression can obviously only be a subset of at most \(k^2\) progressions \(h \in P\); and so, for \(q\) sufficiently large,

\[
q^2 \mu_q(S) = \# \{a, b, c \in S : a + b \equiv 2c \pmod{q}\} \geq \frac{\# \{h \in P : R(h) \geq \rho/2\}}{k^2}.
\]  
(17)

To bound this last quantity from below, we first note that

\[
\sum_{h \in P} R(h) = \frac{1}{k} \sum_{h \in P} |S \cap h| = \frac{1}{k} \sum_{s \in S} \sum_{h \in P, s \in h} 1 = \frac{1}{k} \sum_{s \in S} k(q - 1) = |S|(q - 1).
\]

The second to the last line follows since each \(s \in S\) can be in any one of the \(k\) terms of a \(k\)-term progression (hence the factor \(k\)); and, once this term is specified, there are \(q - 1\) choices for the common difference of such a progression (that \(s\) lies in).

Now, we get

\[
\sum_{\substack{h \in P \quad | \quad R(h) \geq \rho/2}} 1 \geq \sum_{h \in P} \left(R(h) - \frac{\rho}{2}\right) \geq |S|(q - 1) - \frac{\rho}{2} |P| = \rho q (q - 1) - \frac{\rho}{2} q(q - 1) \geq \frac{\rho q^2}{4}.
\]

Combining this with (17) now gives

\[
\mu_q(S) \geq \frac{\rho}{4k^2} \geq \frac{\rho}{16r(\rho/2)^2}.
\]
which proves the theorem.

4 Proof of Proposition 1

As a consequence of theorem 4 and corollary 1, both of which appear in the introduction, we have the following lemma, which we will need for our proof:

Lemma 2 For any sequence of numbers \( x_1, x_2, x_3, ... \) in \([0,1]\) that converges to 0, there are infinitely many integers \( n \) such that

\[
\frac{r(x_{n+1})}{r(x_n)} < \exp \left( -\frac{1}{2C^2} \left( \log^2 x_{n+1} - \log^2 x_n \right) \right).
\]

Given \( k \geq 1 \) define the sequence

\[
x_1 = \exp(-2C^2 \log k), \quad x_{n+1} = x_n \left( 1 - \frac{C^2 \log k}{|\log x_n|} \right).
\]

We note that \( x_1 \) was chosen so that the following holds for all \( n \geq 1 \):

\[
x_n < x_{n+1} \left( 1 + \frac{2C^2 \log k}{|\log x_{n+1}|} \right),
\]

for \( x_n \) sufficiently close to 0.

Clearly this sequence tends to 0 and lies in \([0,1]\). Thus, the conditions of the above lemma are satisfied; and so there exist terms \( x_n \) arbitrarily close to 0 such that

\[
\frac{r(x_{n+1})}{r(x_n)} < \exp \left( -\frac{1}{2C^2} \left( \log^2 \left( x_n \left( 1 - \frac{C^2 \log k}{|\log x_n|} \right) \right) - \log^2 x_n \right) \right)
\]

\[
= \exp \left( -\frac{1}{2C^2} \left( \log x_n + \log \left( 1 - \frac{C^2 \log k}{|\log x_n|} \right)^2 - \log^2 x_n \right) \right)
\]

\[
< \exp \left( -\frac{1}{2C^2} \left( \log x_n - \frac{C^2 \log k}{|\log x_n|} \right)^2 - \log^2 x_n \right)
\]

\[
= \exp \left( -\frac{1}{2C^2} \left( 2C^2 \log k + \frac{C^4 \log^2 k}{\log^2 x_n} \right) \right)
\]

\[
< \exp (-\log k) = \frac{1}{k}.
\]
Now, if we let \( \rho_1 = x_{n+1} \), where \( x_n \) and \( x_{n+1} \) satisfy the conclusion of the lemma above, and so that \( x_{n+1} \) lies in \((0, \rho_0)\), then

\[
\rho > \rho_1 \left( 1 + \frac{2C^2 \log k}{|\log \rho_1|} \right),
\]

implies \( r(\rho) \geq r(x_n) \), which implies that

\[
\frac{r(\rho_1)}{r(\rho)} \leq \frac{r(x_{n+1})}{r(x_n)} < \frac{1}{k}.
\]

5 Proof of Proposition 2

We note that

\[
p^2 \mu_p(W) = \#\{x, y, z \in W : x + y \equiv 2z \pmod{p}\} = \frac{1}{p} \sum_{|a| < p/2} f_W \left( \frac{a}{p} \right)^2 f_W \left( \frac{-2a}{p} \right) = \Sigma_1 + \Sigma_2,
\]

where \( \Sigma_1 \) is the contribution of the terms when we sum with \( |a| < 5vq^{1-\rho_1 \epsilon^2} \), and \( \Sigma_2 \) is the contribution coming from terms with \( 5vq^{1-\rho_1 \epsilon^2} \leq |a| < p/2 \).

We will now show that \( \Sigma_1 \) gives a good approximation to the sum in (18). First, we will require the following two results:

**Proposition 3** For

\[
5vq^{-\rho_1 \epsilon^2} \leq |u| \leq \frac{1}{2},
\]

for \( \epsilon > 0 \) sufficiently small, and \( q \) sufficiently large, we have that

\[
|f_W(u)| < 2\pi \left( \frac{\epsilon}{\rho_1} \right)^{1/3} |S|.
\]

**Lemma 3** Suppose \( |2b| < Q \) and that \( p = q/2 + \delta \), where \( Q|\delta| < q/3 \). Then, we have that

\[
f_W \left( \frac{2b}{q} \right) = f_W \left( \frac{b}{p} \right) + O \left( \frac{\delta Q|S|}{q} \right).
\]
Note: The proof of proposition 3 can be found in section 6 and the proof of lemma 3 can be found in section 7.

We now establish that

$$5vq^{-\rho_1\epsilon^2} \leq |a| < p/2 \implies f_W \left( \frac{-2a}{p} \right) = O \left( \left( \frac{\epsilon}{\rho_1} \right)^{1/3} |S| \right). \quad (20)$$

The proof goes as follows: Suppose $5vq^{-\rho_1\epsilon^2} \leq |a| < p/2$, and let $b$ be the number in $(-p/2, p/2)$ that is congruent to $-2a$ modulo $p$. If $|b| < 5vq^{-\rho_1\epsilon^2}$, then $b$ must be odd, and so from lemma 3 we have that

$$f_W \left( \frac{-2a}{p} \right) = f_W \left( \frac{2b}{q} \right) + O \left( \epsilon \rho_1^{1/4} \right) = O(\epsilon |S|),$$

where the last inequality follows from (11) since $b$ is odd. If $|b| \geq 5vq^{-\rho_1\epsilon^2}$, then we have from proposition 3 that $f_W(b/p)$ is $O((\epsilon \rho_1^{-1})^{1/3} |S|)$. Thus, (20) follows.

From (20) and Parseval's identity we get

$$|\Sigma_2| = O \left( \frac{|S|}{p} \left( \frac{\epsilon}{\rho_1} \right)^{1/3} \right) \sum_{5q^{-\rho_1\epsilon^2} \leq |a| < p/2} \left| f_W \left( \frac{a}{p} \right) \right|^2$$

$$= O \left( |S|^2 \left( \frac{\epsilon}{\rho_1} \right)^{1/3} \right). \quad (21)$$

Applying corollary 2, which appears just after the statement of Parseval's identity in section 7 and makes use of Parseval's identity and Cauchy's inequality, together with lemma 3, we get that

$$\Sigma_1 = \frac{1}{p} \sum_{|a| \leq 5vq^{-\rho_1\epsilon^2}} f_W \left( \frac{a}{p} \right)^2 \left( f_W \left( \frac{-4a}{q} \right) + O(\epsilon^{-1/3} \rho_1^{1/4}) \right)$$

$$= \frac{1}{p} \sum_{|a| \leq 5vq^{-\rho_1\epsilon^2}} f_W \left( \frac{a}{p} \right)^2 f_{S''} \left( \frac{-4a}{q} \right) + O \left( \epsilon^{-1/4} \rho_1^{1/4} \right)$$

$$= \frac{1}{p} \sum_{|a| \leq 5vq^{-\rho_1\epsilon^2}} f_W \left( \frac{a}{p} \right) f_W \left( \frac{-4a}{q} \right)$$

$$\times \left( f_W \left( \frac{2a}{q} \right) + O(\epsilon^{-1/3} \rho_1^{1/4}) \right) + O \left( \epsilon^{-1/4} \rho_1^{1/4} \right)$$

16
\[ \sum_{|a| \leq 10vq^{1-\rho_1^2}} f_W\left(\frac{a}{q}\right) f_W\left(-\frac{2a}{q}\right) + O\left(q^{2-\rho_1^2/4}\right) \]

\[ = \frac{1}{p} \sum_{|a| \leq 5vq^{1-\rho_1^2}} f_W\left(\frac{2a}{q}\right) f_W\left(-\frac{4a}{q}\right) \times \left(f_W\left(\frac{2a}{q}\right) + O(q^{1-\rho_1^2/4})\right) + O\left(q^{2-\rho_1^2/4}\right) \]

\[ = \frac{1}{p} \sum_{|a| \leq 5vq^{1-\rho_1^2}} f_W\left(\frac{2a}{q}\right)^2 f_W\left(-\frac{4a}{q}\right) + O\left(q^{2-\rho_1^2/4}\right) \]

\[ = \frac{1}{p} \sum_{|a| \leq 10vq^{1-\rho_1^2}} f_W\left(\frac{a}{q}\right)^2 f_W\left(-\frac{2a}{q}\right) + O(q^{2-\rho_1^2/4}). \tag{22} \]

Let \( J \) be the set of all integers \( a \) where either \(|a| \leq 10vq^{1-\rho_1^2} \) and \( a \) is odd, or where \( 10vq^{1-\rho_1^2} < |a| < q/2 \). Note that this set \( J \) is the set of all integers \( a \) “missing” from the sum in the last line of (22). For each \( a \in J \), let \(|b| < q/2\) be congruent to \(-2a \pmod{q}\). We will show that for each such \( a \in J \),

\[ f_W\left(-\frac{2a}{q}\right) = f_W\left(\frac{b}{q}\right) = O(\varepsilon|S|). \tag{23} \]

To see this, we first consider the case where \( 10vq^{1-\rho_1^2} < |a| < q/2 \). For this case, either \( b \) is odd, or else \(|b| > 10vq^{1-\rho_1^2}\). In either case, we deduce (23) from (11). For the case \(|a| \leq 10vq^{1-\rho_1^2} \), \( a \) odd, we also get from (11) that (23) holds, because \(-2a\) is not divisible by 4.

Now, by Parseval’s identity we have that

\[ \frac{1}{p} \sum_{a \in J} f_W\left(\frac{a}{q}\right)^2 f_W\left(-\frac{2a}{q}\right) = O\left(\frac{|S|}{p} \sum_{|a| < q/2} \left|f_W\left(\frac{a}{q}\right)\right|^2\right) \]

\[ = O(\varepsilon|S|^2). \tag{24} \]

Thus, from (22) and (24) we get

\[ \Sigma_1 = \frac{1}{p} \sum_{|a| < q/2} f_W\left(\frac{a}{q}\right)^2 f_W\left(-\frac{2a}{q}\right) + O\left(\varepsilon|S|^2\right) \]
\[
\frac{q^3}{p} \mu_q(W) + O(\epsilon |S|^2).
\]
Combining this with our estimate for \( \Sigma_2 \) above, as well as (18), we have

\[
p^2 \mu_p(W) = \Sigma_1 + \Sigma_2 = \frac{q^3}{p} \mu_q(W) + O \left( \left( \frac{\epsilon}{\rho_1} \right)^{1/3} |S|^2 \right),
\]
which proves the proposition.

6 Proof of Proposition 3

Suppose that \( u \) satisfies (19), and let \( a \) be any integer so that

\[
\left| u - \frac{a}{q} \right| \leq \frac{1}{2q}.
\]
Since the set \( W \) satisfies (11), we have that

if \( b \in \mathbb{Z} \), \( |b| < \nu q^{1-\rho_1 \epsilon^2} - 1 \), then

\[
\left| f_W \left( \frac{a-b}{q} \right) \right| \leq \epsilon |S|.
\]
One basic consequence of this fact is the following lemma, which is proved in section 7:

\textbf{Lemma 4} If \( N \) and \( H \) are non-negative integers such that \( [N+1, N+H] \subseteq [0, q-1] \), \( u \) satisfies (20), \( \epsilon > 0 \) is sufficiently small, and \( q \) is sufficiently large in terms of \( \rho \) and \( \epsilon \), then we have that

\[
\left| \sum_{s \in W} e \left( sa \right) \right| < 2 |S| \left( \frac{\epsilon H}{\rho_1 q} \right)^{1/3}.
\]
To finish the proof of our proposition, we apply this lemma together with partial summation: Let \( \delta = u - a/q \), and observe that \( |\delta| \leq 1/(2q) \). Let

\[
h(x) = \sum_{s \leq x} e \left( \frac{sa}{q} \right).
\]
Then, we have
\[
|f_W(u)| = |f_W\left(\frac{a}{q} + \delta\right)| = \left|\int_0^q e(\delta x) dh(x)\right|
\]
\[
= \left|e(\delta x)h(x)|_0^q - 2\pi i \delta \int_0^q e(\delta x) h(x) dx\right|
\]
\[
\leq |f_W\left(\frac{a}{q}\right)| + 2\pi \delta \int_0^q |h(x)| dx
\]
\[
\leq |f_W\left(\frac{a}{q}\right)| + 4\pi \delta |S| \left(\frac{e}{\rho_1 q}\right)^{1/3} \int_0^q x^{1/3} dx
\]
\[
\leq \left(\epsilon + 3\pi \delta q \left(\frac{e}{\rho_1}\right)^{1/3}\right) |S|.
\]

Using the fact that $|\delta| < 1/(2q)$, the proposition now follows.

### 7 Technical Lemmas

In this section we will state a few technical lemmas that were used throughout the paper, as well as provide proofs of these and other lemmas appearing in the paper.

**Lemma 5** For $-1/2 \leq t \leq 1/2$, $t \neq 0$, we have
\[
\left|\sum_{j=N+1}^{N+H} e(jt)\right| \leq \min\left(H, \frac{1}{2|t|}\right).
\]

**Lemma 6 (Parseval’s Identity)** If
\[
f(t) = \sum_{j=0}^{q-1} \lambda_j e(jt),
\]
then
\[
\sum_{a=0}^{q-1} \left|f\left(\frac{a}{q}\right)\right|^2 = q \sum_{j=0}^{q-1} |\lambda_j|^2.
\]
An almost immediate corollary of this lemma, which follows by combining it with Cauchy’s inequality, is as follows:

**Corollary 2** Suppose \( W \subseteq \{0, 1, ..., q-1\} \), that \( q/2 < p < 2q \), and that both \( b_1 \) and \( b_2 \) are integers such that \((b_1, q) = (b_2, p) = 1\). Then, we have

\[
\sum_{|a|<q/2} \left| f_W\left(\frac{b_1 a}{q}\right) \right| \left| f_W\left(\frac{b_2 a}{p}\right) \right| = O(q|S|).
\]

The proof of this result appears at the end of this section.

**Proof of Corollary**

Set

\[
L(x) = \exp(C \sqrt{\log x}),
\]

where \( C \) is as given in theorem 4. Let \( x \) be the integer satisfying

\[
4L(x) < \frac{1}{\rho} \leq 4L(x + 1),
\]

and suppose that \( q \) is any prime larger than \( 4x \). Further, let \( S \subseteq \{1, 2, ..., x\} \) be any set of density at least \( L(x)^{-1} \) having only trivial 3-term arithmetic progressions, as given by Theorem 4.

Define the set

\[
T \subseteq \left\{ 0, 1, 2, ..., \frac{q-1}{2} \right\}
\]

as follows:

\[
T = \left\{ s + 2kx : s \in S, \ 0 \leq k \leq K = \left\lfloor \frac{q}{4x} \right\rfloor \right\}.
\]

Note that

\[
\frac{|T|}{q} = \frac{|S|(K + 1)}{q} > \frac{|S|}{4x} > \frac{1}{4L(x)} > \rho,
\]

and so we see that \( T \) contains density \( > \rho \) of the residue classes modulo \( q \).

We note that if \( a, b, c \in T, \ 0 \leq a, b, c \leq q - 1 \), then

\[
a + b = 2c \iff a + b \equiv 2c \pmod{q},
\]

20
since $T$ satisfies (27); also, since $S$ contains only trivial 3-term arithmetic progressions, we have that
\[ a + b = 2c \iff a = s + 2xk, \ b = s + 2x(k + d), \ c = s + 2x(k + d), \]
where $s \in S$. Thus, the number of triples $a, b, c \in T$ satisfying $a + b = 2c$ is at most
\[ |S| \#\{k, d : 0 \leq k < k + d < k + 2d < K\} < |S|K^2 \leq \frac{q^2}{x + 1} \leq \frac{q^2}{\exp\left(\frac{1}{x^2 \log^2(4\rho)}\right)}, \]
which proves the corollary.

**Proof of Lemma 1**

The proof is via the pigeonhole principle: Let $a_1, \ldots, a_t$ be all the integers in $(0, q/2)$ such that
\[ |f\left(\frac{a}{q}\right)| > \epsilon |S|, \quad (28) \]
for $a = a_1, \ldots, a_t$.

We have that $t \leq (\rho_1 \epsilon^2)^{-1}/2$. To see this, first note that $|f_S(a/q)| = |f_S(-a/q)|$, and so the number of integers $a$ in $(0, q/2)$ satisfying (28) is at most half the total number of integers $a$ with $|a| < q/2$ satisfying (28), and this total number we know to be at most $(\rho_1 \epsilon^2)^{-1}$.

Now, we note that to prove the lemma, it suffices to find an integer $1 \leq j \leq q - 1$ such that if $b_1, \ldots, b_t$ are the smallest numbers in absolute value that are congruent to $ja_1, \ldots, ja_t$ modulo $q$, respectively, then
\[ |b_i| \leq q^{1-\rho_1 \epsilon^2}, \text{ for all } i = 1, 2, \ldots, t. \quad (29) \]
For if so, then if we let $h \equiv j^{-1} \pmod{q}$ and $S' = hS$, then we get that
\[ |f_S\left(\frac{hb}{q}\right)| = |f_{S'}\left(\frac{b}{q}\right)| > \epsilon |S| = \epsilon |S'| \]
\[ \iff hb \equiv \pm a_i \pmod{q} \text{ where } i = 1, 2, \ldots, t, \]
or $b \equiv 0 \pmod{q}$;
but then this would mean that either \( b \equiv 0 \mod q \) or
\[
    \begin{align*}
        b \equiv \pm h^{-1}a_i \equiv \pm ja_i \equiv \pm b_i \mod q,
    \end{align*}
\]
which means that the least residue in absolute value of \( b \mod q \) is \( \leq q^{1-\rho_1 \epsilon^2} \).

We now show how to find an integer \( 1 \leq j \leq q - 1 \) so that (29) holds:
Partition the cube \([0, q - 1]^t\) into the sub-cubes
\[
    \begin{align*}
        & [j_1 q^{1-\rho_1 \epsilon^2}, (j_1 + 1) q^{1-\rho_1 \epsilon^2}] \times [j_2 q^{1-\rho_1 \epsilon^2}, (j_2 + 1) q^{1-\rho_1 \epsilon^2}] \\
        & \times \cdots \times [j_t q^{1-\rho_1 \epsilon^2}, (j_t + 1) q^{1-\rho_1 \epsilon^2}],
    \end{align*}
\]
where \( 0 \leq j_1, \ldots, j_t < q^{\rho_1 \epsilon^2} \). Clearly, there are at most
\[
    \left( q^{\rho_1 \epsilon^2} + 1 \right)^t < q
\]
such sub-cubes. Now, consider the sequence
\[
    (ra_1 \mod q, ra_2 \mod q, \ldots, ra_t \mod q), \quad \text{where } 0 \leq r \leq q - 1.
\]
Since this sequence contains \( q \) terms that lie inside the box \([0, q - 1]^t\), we must have that at least two of these terms lie in the same sub-cube. If \( r = r_1 \) and \( r = r_2 \) are two such terms that correspond to points lying in the same sub-cube, then it follows that
\[
    ((r_1 - r_2)a_1 \mod q, \ldots, (r_1 - r_2)a_t \mod q) \in [-q^{1-\rho_1 \epsilon^2}, q^{1-\rho_1 \epsilon^2}],
\]
where here we take the least residue in absolute value for the entries. Letting \( j = |r_1 - r_2| \) satisfies (29). We now have that (9) follows for this choice of \( j \) (and \( h \)).

**Proof of Lemma 2**

If the conclusion of the lemma were false, then there exists \( N \geq 1 \) so that if \( n > N \), then
\[
    \frac{r(x_{n+1})}{r(x_n)} \geq \exp \left( -\frac{1}{2C^2} \left( \log^2 x_{n+1} - \log^2 x_n \right) \right).
\]
Thus, for any \( n > N \),
\[
    r(x_n) = r(x_N) \prod_{j=N+1}^{n} \frac{r(x_j)}{r(x_{j-1})} \geq r(x_N) \exp \left( -\frac{1}{2C^2} \left( \log^2 x_n - \log^2 x_N \right) \right).
\]

\[ (30) \]
Noting here that \( r(x_N) > 0 \), which follows from (3) from the introduction, we arrive at a contradiction, because from proposition \([4]\) we get that

\[
r(x_n) < \exp \left( -\frac{1}{C^2} \log^2 x_n \right),
\]

which cannot be consistent with (30) once \( x_n \) is small enough (that is, once \( n \) is large enough), Thus, the lemma follows.

**Proof of Lemma 3**

If \( |2b| \leq Q \), and if \( 0 \leq s \leq q - 1 \), then we have that

\[
e \left( \frac{2bs}{2p - 2\delta} \right) = e \left( \frac{2bs}{2p} + \frac{2\delta}{2p(2p - 2\delta)} \right) = e \left( \frac{bs}{p} \right) \left( 1 + O \left( \frac{bs\delta}{p(2p - 2\delta)} \right) \right) = e \left( \frac{bs}{p} \right) + O \left( \frac{\delta Q}{q} \right).
\]

So,

\[
f_W \left( \frac{2b}{q} \right) = \sum_{s \in W} e \left( \frac{2bs}{q} \right) = \sum_{s \in W} \left( e \left( \frac{bs}{p} \right) + O \left( \frac{\delta Q}{q} \right) \right) = f_W \left( \frac{b}{p} \right) + O \left( \frac{\delta Q |S|}{q} \right),
\]

which proves the lemma.

**Proof of Lemma 4**

For \([N + 1, N + H] \subseteq [0, q - 1] \), let

\[
g(t) = \sum_{s \in W, s \in [N+1,N+H]} e(st),
\]

set

\[
D(t) = \sum_{j=N+1}^{N+H} e(jt),
\]

23
and let
\[ 0 \leq K < vq^{1-\rho_1\epsilon^2} - 1 < q^{1-3\rho_1\epsilon^2/4} - 1 \tag{31} \]
be some parameter, which is to be chosen later.

Then, we have for \( a \) satisfying (25) and \( u \) satisfying (19),
\[
\left| g \left( \frac{a}{q} \right) \right| = \left| \frac{1}{q} \sum_{|b|<q/2} D \left( \frac{b}{q} \right) f \left( \frac{a-b}{q} \right) \right| \\
\leq \frac{1}{q} \sum_{|b|\leq K} \left| D \left( \frac{b}{q} \right) \right| \left| f \left( \frac{a-b}{q} \right) \right| + \Sigma \\
\leq \frac{2\epsilon KH|S|}{q} + \Sigma, \tag{32}
\]
where
\[
\Sigma = \frac{1}{q} \sum_{K<|b|\leq q/2} \left| D \left( \frac{b}{q} \right) \right| \left| f \left( \frac{a-b}{q} \right) \right|.
\]
We note that the last line of (32) follows from (26).

To bound \( \Sigma \) from above we will use Cauchy’s inequality, Parseval’s identity, and the upper bound for \(|D(t)|\) given by Lemma \( \text{[6]} \) (which appears at the beginning of this section). We have
\[
\Sigma \leq \frac{1}{q} \left( \sum_{K<|b|\leq q/2} \left| D \left( \frac{b}{q} \right) \right|^2 \right)^{1/2} \left( \sum_{K<|b|\leq q/2} \left| f \left( \frac{a-b}{q} \right) \right|^2 \right)^{1/2} \\
< \frac{1}{q} \left( \sum_{|b|>K} \frac{q^2}{4b^2} \right)^{1/2} \left( \sum_{j=0}^{q-1} \left| f \left( \frac{j}{q} \right) \right|^2 \right)^{1/2} \\
< \frac{1}{2} \sqrt{q|S|} \leq \frac{|S|}{2} \sqrt{\frac{1}{\rho_1 K}}.
\]

The value of \( K \) that minimizes the last line of (32) is
\[
K = \left( \frac{q^2}{16\rho_1\epsilon^2H^2} \right)^{1/3}.
\]

24
and we note that for \( q \) sufficiently large and \( \epsilon > 0 \) sufficiently small this will satisfy (31); and, with this choice of \( K \), we get that

\[
\left| g \left( \frac{a}{q} \right) \right| < 2|S| \left( \frac{\epsilon H}{\rho_1 q} \right)^{1/3},
\]

which proves the lemma.

**Proof of Lemma 5.** From the geometric series identity, we have for \( t \neq 0 \),

\[
\left| \sum_{j=N+1}^{N+H} e(jt) \right| = \left| \frac{e(Ht) - 1}{e(t) - 1} \right| \leq \frac{2}{|e(t/2) - e(-t/2)|} = \frac{1}{\sin(\pi |t|)} \leq \frac{1}{2|t|}.
\]

The last inequality follows from the fact that for \( 0 \leq u \leq \pi/2 \),

\[
\sin(u) \geq \frac{2u}{\pi}.
\]

**Proof of Corollary 2.** From Parseval’s identity and Cauchy’s inequality we have:

\[
\sum_{|a|<q/2} \left| f_W \left( \frac{b_1 a}{q} \right) \right| \left| f_W \left( \frac{b_2 a}{p} \right) \right| \\
\leq \left( \sum_{|a|<q/2} \left| f_W \left( \frac{b_1 a}{q} \right) \right|^2 \right)^{1/2} \left( \sum_{|a|<q/2} \left| f_W \left( \frac{b_2 a}{p} \right) \right|^2 \right)^{1/2} \\
\leq \left( \sum_{|a|<q/2} \left| f_W \left( \frac{a}{q} \right) \right|^2 \right)^{1/2} \left( 2 \sum_{|a|<p/2} \left| f_W \left( \frac{a}{p} \right) \right|^2 \right)^{1/2} \\
\leq (q|S|)^{1/2} (8p|S|)^{1/2} = O(q|S|).
\]

(33)
8 Acknowledgements

I would like to thank Ben Green for pointing out the reference [7] below in an email many months ago, related to an earlier paper of mine.

References

[1] R. C. Baker, G. Harman, 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. and Funct. Anal. 9 (1999), 968-984.

[3] F. A. Behrend, On Sets of Integers Which Contain No Three Terms in Arithmetical Progression, Proc. Nat. Acad. Sci. U.S.A. 32 (1946), 331-332.

[4] R. Heath-Brown, Integer Sets Containing No Arithmetic Progressions, J. London Math. Soc. 35 (1987), 385-394.

[5] K. F. Roth, On Certain Sets of Integers, J. London Math. Soc. 28 (1953), 245-252.

[6] E. Szemeredi, Integer Sets Containing No Arithmetic Progressions, Acta Math. Hungar. 56 (1990), 155-158.

[7] P. Varnavides, On Certain Sets of Positive Density, J. Lond. Math. Soc. 34 (1959), 358-360.