SETS WHOSE DIFFERENCE SET IS SQUARE-FREE

J. WOLF

Abstract. The purpose of this note is to give an exposition of the best-known bound on the density of sets whose difference set contains no squares which was first derived by Pintz, Steiger and Szemerédi in [PSS88]. We show how their method can be brought in line with the modern view of the energy increment strategy employed in problems such as Szemerédi's Theorem on arithmetic progressions, and explore the extent to which the particularities of the method are specific to the set of squares.

1. Introduction

Results about the types of arithmetic structures one is guaranteed to find inside dense sets of integers have been around since the 1950s when Roth [Rot53] first proved that any subset of the integers of positive upper density contains a three-term arithmetic progression. Szemerédi [Sze75], and independently, Furstenberg [Fur77] extended this result to longer progressions. Much of what drives additive combinatorics these days is closely related to the search for better bounds for this problem. Another type of structure mathematicians have always been fascinated by is that of perfect squares. Sárközy [Sár78a] proved the following beautiful theorem in 1978.

Theorem 1.1. *Any subset $A \subseteq [N] = \{1, 2, \ldots, N\}$ which contains no square difference has density*

$$\alpha \ll \frac{(\log\log N)^{2/3}}{(\log N)^{1/3}}.$$

Throughout this article, we shall take the symbol "$\ll$" to mean "is bounded above by a constant times". Results of this type can, by very similar methods, be extended to polynomial structures other than the squares, more precisely, any polynomial that has an integer root. For example, it is true for x^2-1 (for a simple argument in the spirit of [Gre02], see [Wol03]) but not x^2+1 (since there are no squares congruent to 2 mod 3, we can take

Date: September 2005.

the set of all multiples of 3). The general polynomial result is known as the Bergelson-Leibmann Theorem, and was first proved by ergodic theoretic methods [BL96]. Although these are extremely natural and beautiful, no quantitative bounds can be obtained.

Let us also briefly mention that one can ask whether the sets of differences of a dense set necessarily contains an element which is a prime minus 1. Again, the answer is yes and the interested reader is referred to [Sár78c]. Observe that this problem is of no interest for differences of the form $p-k$ with $k \neq 1$: If k is prime, the difference set always contains 0 which is of the form $p-k$. If k is composite, the set of all multiples of k is very dense and contains no differences of the form $p-k$. If $k=0$, we can take all multiples of any composite number to give us a dense counterexample. The methods of [PSS88] were recently applied to the shifted (by 1) primes by [Luc07], but the bounds are superseded by recent work of Ruzsa and Sanders [RS07]. We shall briefly discuss these matters in the final section.

Finally, let us remark that the corresponding problem for squares in sumsets was settled in [JLS82] by graph theoretic methods. In this case it is possible to find a set of density 11/32 whose sumset is square-free.

In this article we are mainly concerned with outlining the main steps leading to a proof of the best known bound on the density of sets whose difference sets contain no squares, which was published in [PSS88], with a subsequent extension of the result to kth powers in [BPPS91].

Theorem 1.2. *There exist constants c and c' such that for all N sufficiently large, any set $A \subseteq [N]$ whose difference set is square-free has density*

$$\alpha \leq c(\log N)^{-c' \log\log\log\log N}.$$

This bound is quite extraordinary in the sense that it is by far superior to any bound known for a similar problem concerning arithmetic structures in sets of integers. In particular, the best known bound for the existence of 3-term arithmetic progressions was very recently improved by Bourgain [Bou06] to

$$\alpha \ll (\log\log N)^2 (\log N)^{-2/3}.$$

See [GT06] for the currently best known bounds for progressions of length 4. For progressions of length $k > 4$, the best known bound is due to Gowers [Gow01] and of the form $(\log\log N)^{-c}$, where the constant c can be taken to be $2^{-2^{k+9}}$.

In fact, the squares bound is good enough to give us information about the existence of arithmetic structure in the prime numbers, which have asymptotic density $(\log N)^{-1}$. We cannot draw similar conclusions from the bounds on Roth's Theorem, although the existence of arithmetic progressions in the primes is now known by other methods [GT04].
In [Sár78a] Sárközy conjectures that $\alpha \ll N^{-1/2+\epsilon}$ for any positive ϵ. He also showed in Part II [Sár78b] of his impressive series of papers that $\alpha - 1/2 > q(p)/2$ for all primes $p \equiv 1 \bmod 4$, where $q(p)$ is the least positive quadratic non-residue of p, so the conjecture would imply that $q(p) = O(p^{\epsilon})$ for all $p \equiv 1 \bmod 4$, which is believed to be out of reach of the currently known techniques in analytic number theory.
The conjecture should also be compared with the best known construction which is due to Ruzsa [Ruz84]. He constructs a subset of $[N]$ of density

$$\alpha \geq N^{-1/2(1-\log 7/\log 65)},$$

where the exponent is approximately equal to -0.266923.
Let us first recall the comparatively simple iteration argument used by Green [Gre02] to tackle this question, which yields the bound $\alpha \ll (\log\log N)^{-1/11}$. At the ith step we have a set A_i of density α_i whose difference set is square-free. The latter property ensures the existence of a large Fourier coefficient, which in turn can be used to establish that A_i has increased density on an arithmetic progression (of a certain length, N_{i+1} say). After rescaling, we obtain a set A_{i+1} of increased density $\alpha_{i+1} \geq \alpha_i(1+\alpha_i^{12})$, whose difference set is again square-free. If α were $\gg (\log\log N)^{-1/11}$, we could repeat this process until the density has increased beyond 1, which is clearly nonsense.
It has been shown in several instances that it can be more efficient to use a collection of large Fourier coefficients rather than a single one. This is what we shall refer to as the *energy increment strategy*, which originated in the work of Szemerédi [Sze90] and was also deployed around the same time by [HB87].
In order to obtain the much better bound stated above in Theorem 1.2, Pintz, Steiger and Szemerédi employ a further iteration, sitting inside the one just described, which aims to build up a relatively large collection of large Fourier coefficients. By nature of the set of squares, we should be able to locate these large Fourier coefficients near rationals with small denominator. Either we can increase the number of such intervals with a large Fourier coefficient at each step of the iteration significantly, and we end up with large total L^2-mass (which gives a good bound on α by Parseval), or we fail to do so at some point. Using combinatorial properties of the rational numbers, the latter case implies a

lower bound on the L^2-mass of Fourier coefficients near rationals with a specific (although unspecified) denominator, and as usual this allows us to pass down to a subprogression on which A has increased density.

Throughout the proof, we may assume that $\alpha \geq c\,(\log N)^{-c' \log\log\log\log N}$ for suitable constants c and c'. We shall use the letter A to denote the characteristic function of the set A, and for ease of notation we set $L = \log N, l = \log\log N, \log_i N = \log\log ... \log N$, where the logarithm is always taken to base e. Put $k = e^{2l}, K = e^{l^2}$.

We shall be using Fourier analysis on $\mathbb{Z}$, and define the Fourier coefficient of a set $A \subseteq [N]$ at $\theta \in \mathbb{T}$ via the formula

$$\hat{A}(\theta) := \sum_{x \in \mathbb{Z}} A(x) e(\theta x).$$

Also, write $I(a/q, \eta)$ for the interval of length η around a/q, and let

$$F_i(q,\eta) = \frac{1}{|A|N} \sum_{\frac{t}{N} \in \bigcup_{\substack{a \leq q \\ (a,q)=1}} I(\frac{a}{q},\eta)} |\hat{A}_i(t/N)|^2,$$

that is, $F_i(q,\eta)$ is the sum of squares of Fourier coefficients near rationals with denominator q. Parseval's identity takes the form

$$\sum_{t=1}^{N} |\hat{A}(t/N)|^2 = N|A|,$$

which implies that $F_i(q,\eta)$ as defined above is bounded by 1. The convolution of two functions $f, g : \mathbb{Z} \to \mathbb{C}$ will be defined as

$$f * g(x) = \sum_y f(x)\overline{g(y-x)},$$

which has the well-known property that its Fourier transform is the pointwise product of the Fourier transforms of f and g.

Let us briefly outline the structure of the remainder of this article. Section 2 is devoted to describing the (pretty standard) iteration, which already gives some improvement over previously known bounds. Section 3 and 4 contain the details of the inner iteration, whereas in Section 5 we will be concerned with working out bounds. Then we will be in a position to discuss the merits of the method in Section 6. An appendix is included for readers who are not familiar with traditional circle method estimates, although we do take some pre-exposure to Fourier analysis for granted.

2. The outer iteration

At step i, we are given a set A_i of density α_i whose difference set is square-free. From now on we fix i, and dropping the index we write $A = A_i$, $\alpha = \alpha_i$, $F(q,\eta) = F_i(q,\eta)$. We shall set $N_1 = N/2$, and without loss of generality assume that A has density at least $\alpha/2$ on $[N_1]$. Write $B = A \cap [N_1]$, and let the function S be defined by $S(x) = 2\sqrt{x/N_1}T(x)$, where T denotes the characteristic function of the set of squares less than N_1. Following the lines of the usual argument, we have

$$\sum_{x\in\mathbb{Z}} A * B(x)S(x) = 0, \tag{1}$$

from which it follows that

$$\sum_{t\neq 0} |\hat{A}(t/N)\hat{B}(t/N)\hat{S}(t/N)| \gg |A|^2|T|. \tag{2}$$

Working with a weighted version of the squares makes them uniformly distributed on $[N_1]$, a process which does not harm the validity of (1) but improves the major arc estimates for $\hat{S}(t/N)$ significantly. Note that this strategy corresponds to replacing the characteristic function with the von-Mangoldt function in the corresponding problem for primes, a standard procedure in analytic number theory.

We shall see that (2) implies that $\hat{A}(t/N)$ takes rather large values rather frequently. By Hölder's Inequality, we can neglect those values of t for which $|\hat{A}(t/N)|$ or $|\hat{B}(t/N)| \leq |A|/K$ provided that $\alpha \gg K^{-2/5}$. Indeed, we have

$$\sum_{\text{these } t} |\hat{A}(t/N)\hat{B}(t/N)\hat{S}(t/N)| \ll \max_{\text{these } t} |\hat{A}(t/N)|^{1/3} \left(\sum_{t=1}^{N} |\hat{A}(t/N)|^2\right)^{5/6} \left(\sum_{t=1}^{N} |\hat{S}(t/N)|^6\right)^{1/6}.$$

The L^6-estimate for $\hat{S}(t/N)$ (see Lemma A.5 in the appendix) implies that this expression is bounded above by a small constant times $|A|^2|T|$.

It turns out that we can also neglect the values of t which belong to the minor arcs (again, for a precise definition see appendix). That is, we need only consider those t for which $t/N \in I(a/q, (qQ)^{-1})$ for $q \leq R$. Indeed, for t/N close to rationals with denominator greater than R, Lemma A.4 implies that

$$\sum_{\text{these } t} |\hat{A}(t/N)\hat{B}(t/N)\hat{S}(t/N)| \ll \frac{N|A||T|}{\sqrt{K/L}}.$$

This quantity is negligible provided that $\alpha \gg (K/L)^{-1/2}$.

In order to perform dyadic averaging over the remaining ranges of parameters, define for $1 \leq b \leq r \leq R$ with $(b,r) = 1$ the *A-special major arcs* as

$$\tau(b,r) = \left\{ t \neq 0 : \frac{t}{N} \in I\left(\frac{b}{r}, \frac{1}{rQ}\right), |\hat{A}(t/N)| \geq \frac{|A|}{K} \right\},$$

where $Q = N/K$ throughout.

Lemma 2.1 (Bound on A-special very major arcs)**.** *Let $1 \leq b \leq r \leq R$ with $(b,r) = 1$. Then we have*

$$\sum_{t \in \tau(b,r)} |\hat{S}(t/N)| \ll \frac{l^3|T|}{\sqrt{r}}$$

with $\tau(b,r)$ defined as above.

Proof. We use the exponential sum estimates in the appendix (Lemma A.1 and Lemma A.2) to obtain

$$\begin{aligned} \sum_{t \in \tau(b,r)} |\hat{S}(t/N)| &\ll \sum_{t \in \tau(b,r)} \left(\frac{\sqrt{\log r}}{\sqrt{r}} |F_S(t/N - b/r)| + \sqrt{r \log r}(1 + |t/N - b/r|N) \right) \\ &\ll \frac{\sqrt{\log r}}{\sqrt{r}} |T| \log K + \sqrt{r \log r} K^2, \end{aligned}$$

where by our choice of K the first term is bounded by $l^3|T|/\sqrt{r}$ and the second term is clearly negligible. □

It follows easily that

$$|A|^2|T| \ll \sum_{r \leq R} \sum_{\substack{b \leq r \\ (b,r)=1}} |\hat{A}(t/N)\hat{B}(t/N)\hat{S}(t/N)| \ll \sum_{r \leq R} \sum_{\substack{b \leq r \\ (b,r)=1}} \max_{t \in \tau(b,r)} |\hat{A}(t/N)| \max_{t \in \tau(b,r)} |\hat{B}(t/N)| \frac{l^3|T|}{\sqrt{r}}.$$

Next we shall partition the set of relevant fractions $\frac{b}{r}$ into sets

$$\mathcal{L}_{X,V} = \left\{ \frac{b}{r} : X < r \leq 2X, \frac{|A|}{V} < \max_{t \in \tau(b,r)} |\hat{A}(t/N)| \leq \frac{2|A|}{V}, \frac{|A|}{V} < \max_{t \in \tau(b,r)} |\hat{B}(t/N)| \leq \frac{2|A|}{V} \right\}$$

for integers $X \leq R, V \leq K$. There are $\log R \log K$ of these sets. Hence there exist parameters $X \leq R, V \leq K$ such that

$$\frac{|A|^2|T|}{\log R \log K} \ll |\mathcal{L}_{X,V}| \frac{|A|^2}{V^2} \frac{l^3|T|}{\sqrt{X}},$$

which in turn immediately implies that

$$|\mathcal{L}_{X,V}| \gg \frac{V^2\sqrt{X}}{l^3 \log R \log K}.$$

But we know more: By definition, $|\mathcal{L}_{X,V}| \leq \alpha^{-1}XV^2 \max_{X<r\leq 2X} F(r,(rQ)^{-1})$, and by Parseval $|\mathcal{L}_{X,V}| \leq \alpha^{-1}V^2$. Putting everything together, we obtain

$$\max_{X<r\leq 2X} F(r,\frac{1}{rQ}) \gg \frac{\alpha^2}{(l^3 \log R \log K)^2}. \tag{3}$$

By our choice of the parameters R and K, we will always have $\log R = O(l^2) = \log K$ so that the denominator is always $O(l^{14})$. The bound (3) will be useful in conjunction with the following standard lemma, which says that we can obtain a density increment of size about $F(q,(qQ)^{-1})$ on a progression of common difference q^2 and length at least $Q/(qL)$.

Lemma 2.2 (Density increment on an arithmetic progression)**.** *Let $q > 1, N' = \lfloor (\eta q^2 L)^{-1} \rfloor$, and let $A \subset [N]$ have density α. Then we can find a set $A' \subset [N']$ of density*

$$\alpha' \geq \alpha + \frac{1}{8}F(q,\eta),$$

with the additional property that if $A - A$ was square-free, so is $A' - A'$.

Proof. We shall show that under the assumption that A has large Fourier mass near rationals with denominator q, A has large intersection with some translate of an arithmetic progression of common difference q^2 which is not too short. Let this progression be $P = \{q^2 k : 1 \leq k \leq |P|\}$ with $|P| = N'$, and consider

$$J := \frac{1}{N}\sum_{t=1}^{N} |\widehat{A * P}(t/N)|^2 = \sum_x |A * P(x)|^2 = \sum_x |A \cap (P+x)|^2, \tag{4}$$

which is the quantity we are trying to find a lower bound for. Now if $t/N \in I(a/q,\eta)$, then $q^2 kt/N = aqk + O(\eta q^2 |P|)$, so that $e(q^2 kt/N) = 1 + O(L^{-1})$ and hence

$$|\hat{P}(t/N)| = |\sum_{k=1}^{|P|} e(q^2 kt/N)| = |P|\left(1 + O\left(L^{-1}\right)\right).$$

It follows from this and 4 that

$$J = \frac{1}{N}\sum_{t=1}^{N} t|\hat{A}(t/N)|^2 |\hat{P}(t/N)|^2 \geq \frac{|A|^2|P|^2}{N}\left(1 + O\left(L^{-1}\right)\right)(1 + \alpha^{-1}F(q,\eta)).$$

We therefore find that there exists an x such that

$$|A \cap (P+x)| \gg |P|(\alpha + \frac{1}{8}F(q,\eta)),$$

and the statement of the Lemma follows. □

The argument so far shows that we can get a density increase of $\alpha \mapsto \alpha + F/8$ with $F \gg \alpha^2 l^{-14}$ at each step, and the length of the progression to which we scale after d steps is $N \mapsto N/(KRL)^d = \Omega(N/L^{cld})$, which means we can iterate $d \ll L/l^2$ times. This gives rise to the condition $L/l^2 \ll \alpha^{-1} \log \alpha^{-1} l^{14}$, which in turn results in a bound on the density of

$$\alpha \ll \frac{l^{17}}{L}.$$

Using a further iteration, which we are about to describe in more detail, we shall be able to raise the exponent of the denominator in the above bound from 1 to a function of N tending (slowly) to infinity.

3. The inner iteration

At the mth step of what we from now on call the *inner iteration*, we inherit a set of large Fourier coefficients near rationals with denominator bounded by X_m,

$$\mathcal{P}^{(m)}_{X_m,V_m} = \left\{ u : \frac{u}{N} \in I\left(\frac{a}{q}, \frac{m}{Q}\right),\ 1 \le a \le q \le X_m,\ (a,q) = 1,\ |\hat{A}(u/N)| \ge \frac{|A|}{V_m} \right\},$$

where X_m and V_m are the parameters maximizing the expression $|\mathcal{P}_{X,V}|V^{-2}$. Since trivially, $\max_{1\le X,1\le V} |\mathcal{P}_{X,V}|V^{-2} \ge 1$, we may assume that $V_m \le X_m$. Let $\mathcal{R}^{(m)}_{X_m,V_m}$ be the corresponding set of centres of intervals a/q.

For fixed $u \in \mathcal{P}^{(m)}$, write $B_u(x) = e(ux/N)B(x)$. We now consider the expression

$$\sum_{x\in\mathbb{Z}} A * -B_u(x)S(x),$$

which is again zero under the assumption that $A - A$ is square-free. Observe that this is where we make definite use of that fact that $A - A$ contains no squares, as opposed to relatively few. It follows that for fixed $u \in \mathcal{P}^{(m)}$, we have

$$\sum_{t\neq 0} |\hat{A}(t/N)\hat{B}((u+t)/N)\hat{S}(t/N)| \gg \frac{|A|^2|T|}{V_m}.$$

Just as before, by a simple use of Hölder's Inequality we can neglect values of t for which one of $\hat{A}(t/N), \hat{A}((u+t)/N)$ or $\hat{S}(t/N)$ is small in modulus. Indeed, if t is such that $|\hat{A}(t/N)|$ or $|\hat{B}((u+t)/N)| \le |A|/K$, then the contribution from these t is bounded by

$$\max_{\text{these } t} |\hat{A}(t/N)|^{1/3} \left(\sum_{t=1}^{N} |\hat{A}(t/N)|^2\right)^{5/6} \left(\sum_{t=1}^{N} |\hat{S}(t/N)|^6\right)^{1/6} \ll \frac{|A|^2|T|}{\alpha^{5/6}K^{1/3}},$$

which is negligible compared with $|A|^2|T|V_m^{-1}$ provided that $\alpha \gg (X_m K^{-1/3})^{6/5}$.

On the other hand, minor and major arc estimates for $\hat{S}(t/N)$ imply that for t to be taken into account, t/N needs to be close to a rational with small denominator $r < X_{m+1}/X_m$. For if t is near a rational with denominator between $X_{m+1}/X_m = X_m^3 X_1$ and K, which corresponds to the fairly major arcs, Lemma A.3 yields

$$\sum_{\text{these } t} |\hat{A}(t/N)\hat{B}((u+t)/N)\hat{S}(t/N)| \leq \max_{X_{m+1}/X_m < r \leq K} \frac{N|A|T}{r^{1/3}} \leq \frac{\alpha^{-1}|A|^2|T|}{X_m X_1^{1/3}},$$

which is $\ll |A|^2|T|V_m^{-1}$ provided that $\alpha \gg X_1^{-1/3}$.
Similarly, for t on the minor arcs (Lemma A.4), that is $K \leq r \leq Q$, we have

$$\sum_{\text{these } t} |\hat{A}(t/N)\hat{B}((u+t)/N)\hat{S}(t/N)| \leq \frac{\alpha^{-1}|A|^2|T|}{\sqrt{K/L}},$$

which is $\ll |A|^2|T|V_m^{-1}$ provided that $\alpha \gg X_m(K/L)^{-1/2}$.
We again perform dyadic averaging over the remaining ranges of parameters. To this end, for $u \in \mathcal{P}^{(m)}$, $1 \leq b \leq r \leq Q$, we define the *A-special major arcs with respect to* u as $(b,r) = 1$, let

$$\tau(b,r,u) = \left\{ t \neq 0 : \frac{t}{N} \in I\left(\frac{b}{r}, \frac{1}{rQ}\right), |\hat{A}(t/N)| \geq \frac{|A|}{K}, |\hat{B}((u+t)/N)| \geq \frac{|A|}{K} \right\}.$$

With this definition we have that for each $u \in \mathcal{P}^{(m)}$,

$$\begin{aligned} \frac{|A|^2|T|}{V_m} &\ll \sum_{r \leq \frac{X_{m+1}}{X_m}} \sum_{\substack{b \leq r \\ (b,r)=1}} \sum_{t \in \tau(b,r,u)} |\hat{A}(t/N)\hat{B}((u+t)/N)\hat{S}(t/N)| \\ &\ll \sum_{r \leq \frac{X_{m+1}}{X_m}} \sum_{\substack{b \leq r \\ (b,r)=1}} \max_{t \in \tau(b,r,u)} |\hat{A}(t/N)| \max_{t \in \tau(b,r,u)} |\hat{B}((u+t)/N)| \sum_{t \in \tau(b,r,u)} |\hat{S}(t/N)| \end{aligned}$$

But as before, we have $\sum_{t \in \tau(b,r,u)} |\hat{S}(t/N)| \ll l^3|T|r^{-1/2}$ by Lemma 2.1. Hence for each $u \in \mathcal{P}^{(m)}$, we can choose integers $1 \leq V_u \leq K, 1 \leq W_u \leq K, 1 \leq X_u \leq X_{m+1}/X_m$ such that the set $\mathcal{L}_u$ given by

$$\left\{ \frac{b}{r} : X_u < r \leq 2X_u,\ \frac{|A|}{V_u} < \max_{t \in \tau(b,r,u)} \left|\hat{A}\left(\frac{t}{N}\right)\right| \leq \frac{2|A|}{V_u},\ \frac{|A|}{W_u} < \max_{t \in \tau(b,r,u)} \left|\hat{B}\left(\frac{u+t}{N}\right)\right| \leq \frac{2|A|}{W_u} \right\}$$

has size

$$|\mathcal{L}_u| \gg \frac{V_u W_u \sqrt{X_u}}{l^3 (\log K)^2 V_m \log(X_{m+1}/X_m)}.$$

When splitting the sum into dyadic ranges, the number of choices for V_u, W_u, X_u is bounded above by $(\log K)^2 \log(X_{m+1}/X_m)$. Hence we can make the same choice of V_u, W_u, X_u for at least $|\mathcal{P}^{(m)}|/(\log K)^2 \log(X_{m+1}/X_m)$ different $u \in \mathcal{P}^{(m)}$. Let us denote the set of such u by $\tilde{\mathcal{P}}^{(m)}$, using parameters $\tilde{V}, \tilde{W}, \tilde{X}$.

Observe that for each $u \in \tilde{\mathcal{P}}^{(m)}$, we have found a $w \in \tau(b, r, u)$ with the property that $|\hat{A}((u+w)/N)| \geq |A|/\tilde{W}$. We would like to count the number of distinct $u+w$ in order to determine whether we can achieve a significant increase in total L^2-mass. For the sake of clarity, the technical details of this counting argument as well as the rough explanation of why we should expect it to work have been postponed until the next section. Writing $F^{(m)} = \max_{\tilde{X}<r\leq 2\tilde{X}} F(r, \frac{1}{rQ})$ and $\tau = \max_{q\leq X_m} \tau(q)$, we find by Lemma 4.1 that there are at least

$$\frac{\alpha^2}{F^{(m)}} \frac{|\tilde{\mathcal{P}}^{(m)}_{X_m,V_m}|}{\tau^4 \tilde{X} \log \tilde{X} \tilde{V}^2} \min |\mathcal{L}_u|^2 \gg \frac{\alpha^2 \tilde{W}^2}{F^{(m)}} \frac{|\tilde{\mathcal{P}}^{(m)}_{X_m,V_m}|}{V_m^2} \frac{1}{\tau^4 (\log K)^{4+2} (\log(X_{m+1}/X_m))^{2+1+1} (l^3)^2}$$

different $u+w$ with the property that $(u+w)/N \in I(c/s, (m+1)/Q)$ and $|A|/\tilde{W} < |\hat{A}((u+w)/N)| \leq 2|A|/\tilde{W}$. This allows us to define the set $\mathcal{P}^{(m+1)}_{X_{m+1},V_{m+1}}$, where we choose parameter $V_{m+1} = \tilde{W}$ and $X_{m+1} = X_m^4 X_1$ (before passing to the next step of the iteration, we will need to reset them so they correspond to the maximum of the expression $|\mathcal{P}_{X,V}|V^{-2}$). Thus we have just shown that

$$\frac{|\mathcal{P}^{(m+1)}_{X_{m+1},V_{m+1}}|}{V_{m+1}^2} \geq \frac{|\mathcal{P}^{(m)}_{X_m,V_m}|}{V_m^2} \frac{\alpha^2}{F^{(m)}} \frac{1}{\tau^4 (\log K)^6 (\log(X_{m+1}/X_m))^4 l^6} = \frac{|\mathcal{P}^{(m)}_{X_m,V_m}|}{V_m^2} \frac{\alpha^2 E}{F^{(m)}}.$$

When choosing our main parameters X_1 and M we shall ensure that $E = \Omega(L^{-1/2})$ always. Now we are faced with two possible cases:

(1) Suppose $\alpha^2 E/F^{(m)} \geq L^{1/2}$ for all $m \leq M$, then by Parseval we have $\alpha \leq L^{-M/2}$, and we will have completed the proof without leaving the inner iteration, simply by building up a collection of Fourier coefficients with large total L^2-mass.

(2) Otherwise, there exists $m \leq M$ such that $\alpha^2 E/F^{(m)} \leq L^{1/2}$, i.e. $F^{(m)} \geq \alpha E/L^{1/2}$. This lower bound on $F^{(m)}$ enables us to pass down to a subprogression on which A has increased density.

4. Combinatorics of rational numbers

We had established that for all $u \in \tilde{\mathcal{P}}$ (recall that $u/N \in I(a/q, m/Q)$), there is a set $\mathcal{L}_u$ defined by

$$\left\{\frac{b}{r} : X_u < r \le 2X_u,\ \frac{|A|}{V_u} < \max_{t\in\tau(b,r,u)}\left|\hat{A}\left(\frac{t}{N}\right)\right| \le \frac{2|A|}{V_u},\ \frac{|A|}{W_u} < \max_{t\in\tau(b,r,u)}\left|\hat{B}\left(\frac{u+t}{N}\right)\right| \le \frac{2|A|}{W_u}\right\}.$$

Since $X_m \le X_1^{4^m}$, the intervals $I(\frac{a}{q}, \frac{m}{Q})$ are disjoint whenever $m \le Q/X_1^{4^m}$ (which is yet another condition we have to satisfy when choosing our parameters), so that counting the number of distinct $u + w$ is equivalent to counting the number of distinct $\frac{a}{q} + \frac{b}{r}$.
In lowest terms, $\frac{a}{q} + \frac{b}{r}$ can be expressed as

$$\frac{\frac{ar'+bq'}{f}}{\frac{r'q'd}{f}},$$

where $d = (q, r), q = dq', r = dr', f = (ar'+bq', d)$, and we immediately note that $(q', r') = 1$ and $(f, q') = (f, r') = 1$.
For fixed $\frac{a}{q}$ we associate a pair $\{d, f\}$ with every $\frac{b}{r} \in \mathcal{L}_{a/q} = \mathcal{L}_u$, where u is the unique element in $\tilde{\mathcal{P}}$ associated with $\frac{a}{q}$. For each $\frac{a}{q}$, there exists a pair $\{d, f\}$ associated with lots of $\frac{b}{r} \in \mathcal{L}_{a/q}$, say all $\frac{b}{r} \in \tilde{\mathcal{L}}_{a/q}$. By averaging, we find that $|\tilde{\mathcal{L}}_{a/q}| \ge \tau(q)^{-2}|\mathcal{L}_{a/q}|$. Similarly, for each q, there exists $\{d, f\}$ associated with lots of $\frac{a}{q}$, say all $\frac{a}{q}$ with $a \in \tilde{A}(q)$. Again, by averaging, we must have $|\tilde{A}(q)| \ge \tau(q)^{-2}|A(q)|$, while $\sum_{q\le X_m}|A(q)| = \tilde{\mathcal{P}}$.
Now fix $\frac{c}{s}$, and count the number of solutions to the equation

$$\frac{c}{s} = \frac{a}{q} + \frac{b}{r} \tag{5}$$

with $\frac{a}{q} \in \tilde{\mathcal{Q}} = \{\frac{a}{q} : q \le \tilde{X}, a \in \tilde{A}(q)\}$ and $\frac{b}{r} \in \tilde{\mathcal{L}}_{a/q}$.
Write $s = q'r'e$, then choose f, which immediately determines d, q, r. It is clear that $a \bmod q'$ is determined by $ar' + bq' = cf$. Denote the number of distinct $a \mod f$ by $r(q)$. By the Chinese Remainder Theorem, we deduce that there are $r(q)\frac{q}{q'f}$ choices for a, which in turn automatically determines b. We conclude that the number of solutions to (5) is $\le \sum_{q=q'r'e}\sum_{f\le d\le r\le\tilde{X}} r(efq')\frac{d}{f}$, so we have an upper bound on the number of solutions provided we have an upper bound for $r(q)$.

Fix q, and the associated popular pair $\{d, f\}$. The crucial observation is that $\tilde{\mathcal{L}}_{a_1/q}$ and $\tilde{\mathcal{L}}_{a_2/q}$ are disjoint if $a_1 \not\equiv a_2 \mod f$. Then

$$r(q)\min|\tilde{\mathcal{L}}_{a/q}| \le |\bigcup_{a\in\tilde{A}(q)}\tilde{\mathcal{L}}_{a/q}| \le |\{\frac{b}{r} : \frac{b}{r}\in\cup\mathcal{L}_{a/q}\}| \le \sum_{r\le R, d|r}|\{b : \frac{b}{r}\in\cup\mathcal{L}_{a/q}\}| \le \frac{\tilde{X}}{d}B_r,$$

where B_r is the number of distinct numerators b such that $\frac{b}{r}\in\cup\mathcal{L}_{a/q}$, so that the number of solutions to (5) is bounded above by $\frac{\tilde{X}\log\tilde{X}B_r}{\min|\tilde{\mathcal{L}}_{a/q}|}$. It follows immediately that the number of distinct $\frac{a}{q}+\frac{b}{r}$ with $\frac{a}{q}\in\tilde{\mathcal{R}}^{(m)}, \frac{b}{r}\in\mathcal{L}_{a/q}$ is

$$\ge \frac{\sum_{q\le\tilde{X}}\sum_{a\in\tilde{A}(q)}|\tilde{\mathcal{L}}_{a/q}|}{\text{no. of sols to (5)}} \gg \sum_{q\le X_m}|\tilde{A}(q)|\min|\tilde{\mathcal{L}}_{a/q}|\left(\frac{\tilde{X}\log\tilde{X}B_r}{\min|\tilde{\mathcal{L}}_{a/q}|}\right)^{-1} \gg \frac{\min|\tilde{\mathcal{L}}_{a/q}|^2|\tilde{\mathcal{R}}^{(m)}|}{\tau^4\tilde{X}\log\tilde{X}B_r}.$$

But B_r and the L^2-mass of Fourier coefficients near rationals with denominator r are by sheer definition related via the inequality $F^{(m)}(r,\frac{1}{rQ}) \ge \frac{\alpha B_r}{\tilde{V}^2}$. Thus we have proved:

Lemma 4.1 (Combinatorics of rational numbers)**.** *Let $\tilde{\mathcal{R}}^{(m)}$ be the set of centres of intervals corresponding to $\tilde{\mathcal{P}}^{(m)}$, with parameters $\tilde{V}, \tilde{W}, \tilde{X}$ as specified in the preceding section. For $u\in\tilde{\mathcal{P}}$, let*

$$\mathcal{L}_u = \left\{\frac{b}{r} : X_u < r \le 2X_u,\ \frac{\alpha}{V_u} < \max_{t\in\tau(b,r,u)}|\hat{A}(t)| \le \frac{2\alpha}{V_u},\ \frac{\alpha}{W_u} < \max_{t\in\tau(b,r,u)}|\hat{A}(u+t)| \le \frac{2\alpha}{W_u}\right\}.$$

Then the number of distinct $\frac{a}{q}+\frac{b}{r}$ with $\frac{a}{q}\in\tilde{\mathcal{R}}^{(m)}, \frac{b}{r}\in\mathcal{L}_{a/q}$ is

$$\gg \frac{|\tilde{\mathcal{R}}^{(m)}|}{\tilde{V}^2}\frac{\alpha^2}{F^{(m)}}\frac{\min|\mathcal{L}_{a/q}|^2}{\tau^4\tilde{X}\log\tilde{X}},$$

where $F^{(m)} = \max_{\tilde{X}<r\le 2\tilde{X}}F^{(m)}(r,\frac{1}{rQ})$ and $\tau = \max_{q\le X_m}\tau(q)$.

Let us summarize what this section has achieved: We were trying to assess whether we could increase the L^2-mass of the large Fourier coefficients, and for this purpose we counted how many of them there are, that is we counted the number of distinct new intervals with centres $\frac{a}{q}+\frac{b}{r}$. The obvious way of doing this is to divide the number of all relevant fractions of the form $\frac{a}{q}+\frac{b}{r}$, that is $\sum_{\text{appropriate } a,q}|\mathcal{L}_{a/q}|$, by the number of solutions to $\frac{c}{s} = \frac{a}{q}+\frac{b}{r}$ with $\frac{a}{q}\in\mathcal{R}^{(m)}, \frac{b}{r}\in\mathcal{L}_{a/q}$.
$F^{(m)}(r,\frac{1}{rQ}) \ge \frac{\alpha B_r}{\tilde{V}^2}$ immediately gives us the desired connection between the L^2-mass near denominator r and the number of distinct numerators b such that $\frac{b}{r}\in\cup\mathcal{L}_{a/q}$, B_r. The upshot is that either we have lots of these for some r, i.e. $B = \max_r B_r$ is large, in which case we have (by definition) large L^2-mass near a specific denominator and we can scale. If

not, i.e. if B is small, then by the above counting argument we obtain lots of new intervals so that the total L^2 mass increases significantly.

5. Working out bounds

It's a bit of a mess: To make combinatorics of rationals work (and this is where the real restrictions of this method lie), we need $E = \Omega(L^{-1/2})$ as remarked above, that is we need

$$\tau = \max_{q \le X_m} \tau(q) \ll L^C \tag{6}$$

for some small constant C and

$$\log \frac{X_{m+1}}{X_m} \le l^2 l \tag{7}$$

It is a well-known number-theoretic fact that

$$\log \tau(X_m) \ll \frac{\log X_m}{\log\log X_m} \ll \frac{4^m \log X_1}{m + \log\log X_1} \le Cl,$$

and we therefore choose

$$M = \frac{1}{2}\log_4 N$$

as well as

$$X_1 = L^{(\log_3 N)^{1/4}}$$

in order to satisfy both (6) and (7). We can then check that all the other conditions are satisfied: We also needed

$$\alpha \gg X_1^{-1/3}, \;\; \alpha \gg \frac{X_m}{\sqrt{K/L}}, \;\; \alpha \gg \left(\frac{X_m}{K^{1/3}}\right)^{6/5}$$

to ensure that we can neglect the contributions from the non-arithmetic part, that is the fairly major and the minor arcs, and

$$m \le \frac{Q}{X_1^{4^m}}$$

to force the intervals $I(\frac{a}{q}, \frac{m}{Q})$ to be disjoint. We have made no attempts to optimize the constants involved here.

6. Concluding remarks

The method which we have discussed was extended to cover the case of kth powers in [BPPS91]. Only minor modifications to the argument are necessary, and these occur almost exclusively through the Hardy-Littlewood type estimates in the appendix.

It should also be clear that similar progress can be made for polynomial differences such as x^2-1. Very recently, Lucier [Luc07] applied the method to the shifted primes to obtain a bound of

$$\left(\frac{(\log_3 N)^4}{\log\log N}\right)^{\log_5 N}$$

on the density of the set which avoids the set of all $p=1$, p a prime. However, it should be noted that the currently best-known bound for this problem obtained in [RS07] is of the form

$$\exp(-c\sqrt[4]{\log N})$$

and does not use this technique. Indeed, it is relatively straightforward to obtain a density increment of size in the case of primes, which cannot be bettered by the technique described in this article. (For comparison, the straightforward density increase in the case of squares is of size α^3, and can be improved to α^2 using combinatorics of rationals.)
Given the fact that the application to the primes is slightly bogus, it would be very interesting to find a genuinely new and useful application of this method.

A. Appendix: Major and minor arcs estimates for weighted squares

The material in this section is entirely standard and we give barely enough detail to make this exposition self-contained. For an introduction to the circle method, see [Vau81].
By Dirichlet's Theorem, $t/N \in I(a/q,(qQ)^{-1})$ for some $1\le a\le q\le Q, (a,q)=1$. Call the set of those t for which $q\le R$ the *major arcs* and the set of those t for which $R<q\le Q=N/K$ the *minor arcs*. It is a typical feature of the Hardy-Littlewood method that the exact values of the boundaries between the arcs need to be determined in the course of the proof. We define the generating function of the weighted squares by

$$F_S(\theta)=\sum_{x^2\le N_1}\frac{2x}{\sqrt{N_1}}e(x^2\theta).$$

Note that $F_S(\theta)$ coincides with our earlier definition of $\hat{S}(\theta)$ used throughout the proof.

Lemma A.1 (Weighted exponential sums for squares)**.** *Let θ belong to the interval $I(a/q,\eta)$. Then we have the bound*

$$|F_S(\theta)|\ll\frac{\sqrt{\log q}}{\sqrt{q}}|F_S(\eta)|+\sqrt{q\log q}(1+|\eta|N).$$

Proof. Consider the truncated version $F_S(\theta,m)=\sum_{x\le m}2xe(x^2\theta)/\sqrt{N_1}$ of F_S, as well as the Gauss sum $B(a/q,m)=\sum_{x\le m}e(x^2a/q)$. If $m\le q$, we have $B(a/q,m)\ll\sqrt{q\log q}$

and using Abel's Inequality (which says that if g is monotone, then $|\sum_{x\le m} g(x)f(x)|$ is bounded above by $\max_{x\le m}|g(x)|\max_{j\le m}|\sum_{x\le j} f(x)|$), we conclude that $F_S(a/q,m) \ll m\sqrt{q\log q/N}$. It follows that $F_S(a/q) \ll \sqrt{q\log q}$. In the case where $m > q$, we find $F_S(a/q,m) = B(a/q,q)m^2/(q\sqrt{N}) + O(m\sqrt{q\log q/N})$ by splitting into segments of length q, and so $F_S(a/q) = B(a/q,q)\sqrt{N}/q + O(\sqrt{q\log q})$. Now let $\theta = a/q+\eta$ with $(a,q)=1$. By partial summation, we obtain $F_S(\theta,m) - B(a/q,q)F_S(\eta,m)/q = O(m\sqrt{q\log q/N}(1+|\eta|m^2))$, whence $F_S(\theta) = B(a/q,q)F_S(\eta)/q + O(\sqrt{q\log q}(1+|\eta|N))$. □

For small values of η, we can give a fairly good estimate for $F_S(\eta)$.

Lemma A.2. *Let $\frac{1}{10} < h = \eta N \le H = N^{1/8}$. Then*

$$|F_S(\eta)| \ll \frac{|T|}{|h|}.$$

Note that without weighting the exponential sum, we would have a bound of $|T||h|^{-1/2}$ here, which isn't good enough for the purposes of this paper.

Proof. Let us split the range of summation for F_S into intervals

$$R_{ij} = \{x : x^2 \in [N(i+j/H)/h, N(i+(j+1)/H)/h)\}.$$

Now break up the sum

$$F_S(h/N) = \sum_{i=1}^{\lfloor h/2\rfloor -1}\sum_{j=0}^{H}\sum_{x\in R_{ij}} 2xe(x^2h/N)/|T| + O(|T|/|h|).$$

On R_{ij}, x^2h/N is equal to an integer plus a small remainder of at most H^{-1}, so the sum becomes

$$\sum_{i=1}^{h/2}\sum_{j=0}^{T} e(j/H)/|T| \sum_{x\in R_{ij}} 2x + \sum_{x^2\le N_1} 2x/(H|T|).$$

It is easily shown that $\sum_{x\in R_{ij}} 2x = N/(Hh) + O(|T|)$, and hence the sum is bounded by $O(hH + |T|/H)$ □

Lemma A.3 (Major arcs)**.** *For $t \in I(a/q,(qQ)^{-1})$ with $q \le R$, we have*

$$|F_S(t/N)| \ll \frac{|T|}{\sqrt[3]{q}}.$$

Proof. If $q \ll K$, then $h > 1/10$ and putting together the previous two lemmas yields $|F_S(t/N)| = \sqrt{\log q/q}|T|/|h| + O(\sqrt{q\log q}N/(qQ))$. The first term clearly dominates and thus, if $q \le R$, we have $F_S(t/N) \ll |T|q^{-1/3}$. □

Lemma A.4 (Minor arcs)**.** *For $t \in I(a/q, (qQ)^{-1})$ with $R < q \leq Q$, we have*

$$|F_S(t/N)| \ll \frac{|T|}{\sqrt{K/L}}.$$

Proof. If q ranges between R and $N^{1/8}$, the result follows from methods used above. For $q > N^{1/8}$, it follows from Weyl's Inequality that $|F_S(t/N)| \ll \sqrt{N \log N}(q^{-1/2} + \sqrt{Q/N})$, which is clearly bounded above by $\sqrt{QL}$ provided that $q \gg K$. □

We also need the following consequence of Hua's Lemma:

Lemma A.5 (L^6-bound for weighted squares)**.**

$$\sum_{t=1}^{N} |F_S(t/N)|^6 \ll |T|^6.$$

We omit the proof but point out that the lemma corresponds to (a weighted version of) the well-known fact that the number of representations of an integer n as the sum of six squares is asymptotic to n^2.

References

[BL96] V. Bergelson and A. Leibman. Polynomial extensions of Van der Waerden's and Szemerédi's theorems. *J. Amer. Math. Soc.*, 9:725–753, 1996.

[Bou06] J. Bourgain. Roth's theorem on progressions revisited. Preprint, 2006.

[BPPS91] A. Balog, J. Pelikan, J. Pintz, and E. Szemerédi. Difference sets without kth powers. *DIMACS Technical Report*, 64, 1991.

[Fur77] H. Furstenberg. Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions. *J. Analyse Math.*, 31:204–256, 1977.

[Gow01] W.T. Gowers. A new proof of Szemerédi's theorem. *GAFA*, 11:465–588, 2001.

[Gre02] B.J. Green. On arithmetic structure in dense sets of integers. *Duke Math. Journal*, 114:215–238, 2002.

[GT04] B.J. Green and T. Tao. There are arbitrarily long arithmetic progressions in the primes. Available at arXiv:math.NT/0404188, 2004.

[GT06] B.J. Green and T. Tao. New bounds for Szemerédi's theorem, II: A new bound for $r_4(N)$. Submitted. Available at arXiv:math.NT/0610604, 2006.

[HB87] D.R. Heath-Brown. Integer sets containing no arithmetic progressions. *J. London Math. Soc. (2)*, 35:385–394, 1987.

[JLS82] A.M. Odlyzko J.C. Lagarias 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. Comb. Theory, Series A.*, 33:167–185, 1982.

[Luc07] J. Lucier. Difference sets and shifted primes. Preprint. Available at arxiv:math.NT/0705.3749, 2007.

[PSS88] J. Pintz, W.L. Steiger, and E. Szemerédi. On sets of natural numbers whose difference set contains no squares. *J. London Math. Soc. (2)*, 37:219–231, 1988.

[Rot53] K.F. Roth. On certain sets of integers. *J. London Math. Soc.*, 28:104–109, 1953.

[RS07] I.Z. Ruzsa and T. Sanders. Difference sets and the primes. Preprint. Available at arxiv:math.NT/, 2007.

[Ruz84] I.Z. Ruzsa. Difference sets without squares. *Periodica Math. Hungar.*, 15:205–209, 1984.

[Sár78a] A. Sárközy. On difference sets of sequences of integers I. *Acta Math. Acad. Sci. Hungar.*, 31:125–149, 1978.

[Sár78b] A. Sárközy. On difference sets of sequences of integers II. *Ann. Univ. Sci. Budapest*, 21:45–53, 1978.

[Sár78c] A. Sárközy. On difference sets of sequences of integers III. *Acta Math. Acad. Sci. Hungar.*, 31:355–386, 1978.

[Sze75] E. Szemerédi. On integer sets containing no k elements in arithmetic progression. *Acta Arith.*, 27:199–245, 1975.

[Sze90] E. Szemerédi. Integer sets containing no arithmetic progressions. *Acta Math. Hungar.*, 56(1-2):155–158, 1990.

[Vau81] R.C. Vaughan. *The Hardy-Littlewood Method*. Cambridge University Press, 1981.

[Wol03] J. Wolf. Arithmetic structure in difference sets. Part III Essay. University of Cambridge., 2003.

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

E-mail address: J.Wolf@dpmms.cam.ac.uk