# Finding arithmetic progressions in sumsets

Bora Çalım

Abstract

Following [5], we prove that the sumset of sufficiently large subsets of $\{1, 2, ..., N\}$ contain an arithmetic progression of length at least $\exp(c(\log N)^{1/2})$ for some absolute constant $c$. The proof uses Fourier analysis and some probabilistic estimates, and Chang's lemma on the large spectrum of a set.

## 1. *Introduction*

Since convolution is in some sense a "smoothing" operation and the sumset operation is essentially taking the convolution of the indicator functions of two sets, it seems reasonable to expect sumsets to contain more structure than arbitrary sets. Bogolyubov-type theorems, for instance, make this intuition precise. As a manifestation of this phenomenon, Green [5] proved the following:

Theorem 1.1. *(Green) There are absolute constants $c, c' > 0$ such that the following holds: Let $C, D \subseteq \{1, 2, ..., N\}$ with $|C| = \gamma N, |D| = \delta N$. Then the sumset $C + D$ contains an arithmetic progression of length at least* $\exp(c((\gamma\delta \log N)^{1/2} - c' \log \log N))$.

Remark 1. Earlier, Bourgain [1] had proved the same result with 1/3 in place of 1/2. Later, Croot, Łaba and Sisask [3] proved the same result as Green [5] by the machinery of almost-periodicity which later found many more applications in additive combinatorics.

Remark 2. Due to a construction of Ruzsa [9], one cannot hope for a larger number than 2/3 in place of 1/2. As far as I know, the exponent 1/2 has not been improved.

### 1.1. *Understanding the "numerology"*

In this short section we give a non-rigorous interpretation of the bound given by the main theorem.

It can be observed that we can "ignore" the $-\log \log N$ term. Assume $N$ has $k$ digits. Then a short calculation (remembering that $e^k$ has, up to a constant factor, $k$ digits) gives an arithmetic progression whose *length* has $c(\gamma\delta)^{1/2}k^{1/2}$ digits. Regarding $\gamma$ and $\delta$ as constants, this means that in the sumset of sufficiently dense sets (we need this because of the $-\log \log N$ term, this issue will be dealt with later in a more precise manner) in $\{1, 2, ..., 10^k\}$, we can find APs of length $c10^{\sqrt{k}}$.

### 1.2. *Definitions, notation and basic facts of Fourier analysis in $\mathbb{Z}/n\mathbb{Z}$*

Here we fix some notation and record several facts about Fourier analysis in $\mathbb{Z}/n\mathbb{Z}$ (henceforth denoted by $G$), mainly to fix the summing/averaging conventions. In particular, we always average in the physical space and sum in the Fourier space.

For convenience, we regard $N$ as fixed and write $\omega = e^{2\pi i/N}$. We write $\mathbb{E}$ as shorthand for $\frac{1}{|G|}\sum$. If confusion can arise, we use the latter (this will be the case in the part where we prove Rudin's inequality).

We define the averaging and summing inner products as $\langle f, g\rangle_{L^2} = \mathbb{E}_{x\in G} f(x)\overline{g(x)}$, $\langle f, g\rangle_{\ell^2} = \sum_{r\in G} f(x)\overline{g(x)}$. As noted before, we use the former mainly in the physical space and the latter mainly in the Fourier space.

Throughout, we identify sets with their characteristic functions. We define the Fourier transform as $\hat{f}(r) = \mathbb{E}_{x\in G} f(x)\omega^{-rx}$, and the convolution of two functions as $f * g(x) = \mathbb{E}_{y\in G} f(y)g(x-y)$. We now record several facts.

Proposition 1.2. *(Fourier analysis in $G$) Let $f, g : G \to \mathbb{C}$ be two functions. Then the following statements hold:*

(i) *(Fourier inversion)* $f(x) = \sum_{r\in G} \hat{f}(r)\omega^{rx}$
(ii) *(Plancherel)* $\mathbb{E}_{x\in G} f(x)\overline{g(x)} = \sum_{r\in G} \hat{f}(r)\overline{\hat{g}(r)}$
(iii) *(Convolution identity)* $\widehat{f * g}(r) = \hat{f}(r)\hat{g}(r)$

We record the special case of Plancherel with $f = g$ often, which we call Parseval, in the form $\|f\|_{L^2} = \|\hat{f}\|_{\ell^2}$. With the same notation we define the p-norms $\|f\|_{L^p}$ and $\|\hat{f}\|_{\ell^p}$. We may shorten $\|f\|_{L^p}$ to $\|f\|_p$ for convenience.

For a set $\Gamma \subseteq G$, we define the Bohr set $B(\Gamma, \epsilon)$ with width $\epsilon$ as $\{x \in G : \|\frac{rx}{N}\|_{\mathbb{R}/\mathbb{Z}} \le \epsilon$ for all $r \in \Gamma\}$, where the "norm" denotes the distance to the nearest integer. Bohr sets contain large APs, and also are somewhat rigid under linear combinations with small coefficients because of the triangle inequality. The latter heuristic fact is used in the last step of the proof of the first.

Lemma 1.3. *(Bohr set contains a long AP) If $|\Gamma| = d$, then $B(\Gamma, \epsilon)$ contains an AP with length at least $\epsilon N^{1/d}$.*

*Proof.* This is a simple pigeonholing argument. We give the details for completeness. Consider $\mathbb{R}^d/N\mathbb{Z}^d$ and let $x = (\gamma_1, \gamma_2, \dots, \gamma_d)$ (where $\gamma_j$ are all the elements of $\Gamma$) in this space. Consider closed cubes of side length $N^{1-1/d}$ centered at the points $0, x, 2x, \dots, (N-1)x$. The volume of these cubes sum to $N^d$. Assume for contradiction that the cubes are pairwise disjoint. Then, since they are compact, they are separated, so the volume of $\mathbb{R}^d/N\mathbb{Z}^d$ becomes larger than $N^d$, which is absurd. Therefore there is a pair of cubes which intersect, which means there is a pair $n, m$ and a point $u = (u_1, \dots, u_d)$ such that $|u_j - nx_j|, |u_j - mx_j| \le N^{1-1/d}/2$ for all $j$ (where distances are taken in $\mathbb{R}/\mathbb{N}$), so by triangle inequality $|(n-m)x_j| \le N^{1-1/d}$ for all $j$. Remembering the definition of $x$, this means there exists some $r = n - m \neq 0 \in G$ such that $\|\frac{r\gamma_i}{N}\|_{\mathbb{R}/\mathbb{Z}} \le N^{-1/d}$ for all $i$.

Let $M = \lfloor \epsilon N^{1/d} \rfloor$, $P = \{-Mr, \dots, -r, 0, r, \dots, Mr\}$. We claim that $P \subseteq B(\Gamma, \epsilon)$. Indeed, for any $\gamma \in \Gamma$ and $kr \in P$, we have $\|\frac{kr\gamma}{N}\|_{\mathbb{R}/\mathbb{Z}} \le \lfloor \epsilon N^{1/d} \rfloor \|\frac{r\gamma_i}{N}\|_{\mathbb{R}/\mathbb{Z}} \le \lfloor \epsilon N^{1/d} \rfloor N^{-1/d} \le \epsilon$. Since $P$ is an AP and $|P| = 2M + 1 \ge \epsilon N^{1/d}$, this completes the proof. □

### 1.3. *Brief outline of the argument and preliminary reductions*

The main technical result of [**5**] states that if all subsets of a set have a large nontrivial Fourier coefficient, then the complement of this set contains a long AP. This is connected to Theorem 1.1 by a short Fourier analytic argument showing that the complement of a sumset satisfies this condition, so the sumset contains a long AP.

The proof of the technical result goes roughly as follows: If a subset of appropriate size of the set in question minimizes the size of its largest nontrivial Fourier coefficient, unless the complement of the set contains a long AP, we can "play with the subset randomly" and get a subset with smaller largest nontrivial Fourier coefficient, which is a contradiction.

Actually this is not exactly how it is done. Rather, we assume we can "play with the subset randomly" to get a smaller largest nontrivial FC, and show that this implies there is no long AP in the complement. However it can be seen that these two approaches are equivalent, since they are contrapositives of each other. I am not sure how one would think of approaching this problem in this way.

We now give a precise statement of the main technical result, and prove Theorem 1.1 assuming it. A set $A \subseteq G$ is called $\alpha$-hereditarily non-uniform ($\alpha$-HNU for short) if for all subsets $S$ of $A$, we have $\sup_{r \neq 0} |\hat{S}(r)| \geq \alpha |S|/N$. Since Fourier coefficients of $S$ are bounded by $|S|/N$, this heuristically means *all* subsets of $A$ are "at least $\alpha$ fraction of being as non-uniform as possible". The result concerning these sets is as follows:

Theorem 1.4. *There exists absolute constants $c, c' > 0$ such that the following holds: Assume that $A \subseteq G$ is $\alpha$-HNU with $\alpha \geq c \log\log N/(\log N)^{1/2}$. Then $A^c$ (the complement of A) contains an AP of length at least $e^{c'\alpha\sqrt{\log N}}$.*

Remark 3. Carrying out the calculations carefully, some admissible pair of $c, c'$ is in the range $10^8 < c < 10^{10}$, $10^{-8} > c' > 10^{-10}$.

We now give the short Fourier analytic argument connecting 1.4 and sumsets.

Lemma 1.5. *Let $C, D \subseteq G$ with $|C| = \gamma N$, $|D| = \delta N$. Then $(C + D)^c$ is $\sqrt{\gamma\delta}$-HNU.*

*Proof.* Let $S \subseteq (C + D)^c$. Then, since $S$ and $C + D$ are disjoint, $\mathbb{E}_{x \in G} S(x)(C * D)(x) = 0$. By Plancherel and convolution identity, we obtain $\sum_{r \in G} \hat{S}(r)\hat{C}(r)\hat{D}(r) = 0$. Separating this into $r = 0$ and $r \neq 0$ and using the fact that the trivial Fourier coefficient of an indicator function of a set is its density (also noting that it is real-valued and positive, so it equals its conjugate), we obtain

$$\begin{aligned}
\frac{|S|\gamma\delta}{N} = \frac{|S||C||D|}{N^3} = \hat{S}(0)\hat{C}(0)\hat{D}(0) = \left|\sum_{r \neq 0} \hat{S}(r)\overline{\hat{C}(r)\hat{D}(r)}\right| &\leq \sum_{r \neq 0} |\hat{S}(r)||\hat{C}(r)||\hat{D}(r)|. \\
&\leq \sup_{r \neq 0} |\hat{S}(r)| \sum_{r} |\hat{C}(r)||\hat{D}(r)| \\
&\leq \sup_{r \neq 0} |\hat{S}(r)| \|\hat{C}\|_{\ell^2} \|\hat{D}\|_{\ell^2} \\
&= \sup_{r \neq 0} |\hat{S}(r)| \|C\|_{L^2} \|D\|_{L^2} \\
&= \sup_{r \neq 0} |\hat{S}(r)| \sqrt{\gamma\delta},
\end{aligned}$$

where we used Cauchy-Schwarz and Parseval in the third-to-last and second-to-last lines. Rearranging, this gives the desired result. □

*Proof of Theorem 1.1 assuming Theorem 1.4 and Lemma 1.5.* Embed $C$ and $D$ in $\mathbb{Z}/(6N+1)\mathbb{Z}$. Then the densities of $C$ and $D$ are $\geq \gamma/7$, $\delta/7$. By Lemma 1.5, $(C + D)^c$ is $\gamma\delta/7$-HNU. Now Theorem 1.4 gives us the following dichotomy: Either $\gamma\delta < 7c(\log\log 7N)^2/\log 7N$, or $C + D$ contains an AP of length at least $e^{c'\sqrt{\gamma\delta \log 7N/7}}$. Since $\log 7N$ and $\log N$ are of the same order, we can replace the $7N$'s appearing with $N$'s at the expense of slightly increasing $c$ and decreasing $c'$.

Now let us investigate the bounds we obtained more carefully. If $\sqrt{\gamma\delta \log N} < c \log\log N$, then we do not have any nontrivial information, and if this is not the case, then we have an AP of length $\geq e^{c'\sqrt{\gamma\delta \log N}}$. Putting these together without the condition on $\gamma$ and $\delta$ requires us to offset the first case (and lose some lower-order terms in the process), and this is exactly what the statement of Theorem 1.1 does: If $\gamma\delta$ is too small, we do not get anything, and if it is large, we get a long AP.

Theorem 1.1 follows from observing that an AP in $C + D \subseteq \{1, 2, \ldots, 2N\} \subseteq \mathbb{Z}/(6N+1)\mathbb{Z}$ is an AP in $\mathbb{Z}$ by Freiman isomorphism considerations.

□

## 2. *Probabilistic preliminaries: Bernstein's inequality*

In this section we establish a "large deviation inequality", which originates from Bernstein (I have not been able to precisely locate where it was first proven, because of language barriers), although we will follow the proof in [**7**].

The inequality heuristically states that the sum of independent random variables is close to their mean with high probability. Precisely,

Theorem 2.1. *(Bernstein) Assume that $X_1, \ldots, X_n$ are $\mathbb{R}$-valued independent random variables with $\mathbb{E}X_i = 0$, $\mathbb{E}|X_i|^2 = \sigma_i^2$, $\sigma^2 = \sigma_1^2 + \ldots + \sigma_n^2$, $|X_i| \leq 1$ uniformly in $i$, then*

$$\mathbb{P}(|\overline{X}| \geq t) \leq 2\exp\left( - \frac{nt^2}{\frac{2\sigma^2}{n} + 2t/3}\right).$$

We will use the following consequence of this inequality:

Corollary 2.2. *Assume that $X_1, \ldots, X_n$ are $\mathbb{C}$-valued independent random variables with $\mathbb{E}X_i = 0$, $\mathbb{E}|X_i|^2 = \sigma_i^2$, $\sigma^2 = \sigma_1^2 + \ldots + \sigma_n^2$, $|X_i| \leq 1$ uniformly in $i$, and $\sigma^2 \geq 6nt$, then*

$$\mathbb{P}(|\overline{X}| \geq t) \leq 4\exp\left( - \frac{n^2t^2}{8\sigma^2}\right).$$

*Proof of Corollary 2.2 assuming Theorem 2.1.* $\operatorname{Re}(X_1), \ldots, \operatorname{Re}(X_n)$, and $\operatorname{Im}(X_1), \ldots, \operatorname{Im}(X_n)$ are $\mathbb{R}$-valued independent random variables with mean zero and are uniformly bounded by 1. Let $\mathbb{E}|\operatorname{Re}(X_i)|^2 = \sigma_{i,1}^2$ and $\mathbb{E}|\operatorname{Im}(X_i)|^2 = \sigma_{i,2}^2$. Observe that $\sigma_i^2 = \sigma_{i,1}^2 + \sigma_{i,2}^2$ by linearity of expectation. Let $\sigma_{re}^2 = \sum \sigma_{i,1}^2$ and $\sigma_{im}^2 = \sum \sigma_{i,2}^2$. Then $\sigma^2 = \sigma_{re}^2 + \sigma_{im}^2$. By simple geometric considerations, $\mathbb{P}(|\overline{X}| \geq t) \leq \mathbb{P}(|\operatorname{Re}(\overline{X})| \geq t/\sqrt{2}) + \mathbb{P}(|\operatorname{Im}(\overline{X})| \geq t/\sqrt{2})$, so by Bernstein, we have

$$\mathbb{P}(|\overline{X}| \geq t) \leq 2\exp\left( - \frac{nt^2}{\frac{4\sigma_{re}^2}{n} + 4t/3}\right) + 2\exp\left( - \frac{nt^2}{\frac{4\sigma_{im}^2}{n} + 4t/3}\right). \tag{2.1}$$

Let $\sigma_{max} = \max(\sigma_{re}, \sigma_{im})$, so that $\sigma_{max}^2 \geq 3nt$. Then $4t/3 \leq 4\sigma_{max}^2/9n$, so (after some calculation)

$$(2.1) \leq 4\exp\left( - \frac{9n^2t^2}{40\sigma_{max}^2}\right) \leq 4\exp\left( - \frac{9n^2t^2}{40\sigma^2}\right) \leq 4\exp\left( - \frac{n^2t^2}{8\sigma^2}\right), \tag{2.2}$$

as claimed. □

We now prove Theorem 2.1. I have to admit I do not know exactly what is going on in this proof, but the calculations work. We begin with a preliminary lemma.

Lemma 2.3. *Let $X$ be a real-valued random variable bounded by 1 with mean zero and variance $\sigma^2$. Then for any $z > 0$, we have $\mathbb{E}(e^{zX}) \leq e^{\sigma^2(e^z - z - 1)}$.*

*Proof.* Let $F = \sum_{r\geq 2} \frac{z^{r-2}\mathbb{E}(X^r)}{r!\sigma^2}$. Then, by Taylor, $\mathbb{E}(e^{zX}) = \mathbb{E}(1 + zX + z^2\sigma^2 F)$ $\leq 1 + z^2\sigma^2 F \leq \exp(z^2\sigma^2 F)$. For $r \geq 2$, since $X$ is bounded by 1, $\mathbb{E}(X^r) = \mathbb{E}(X^{r-2}X^2) \leq \sigma^2$. Therefore $F \leq \sum_{r\geq 2} \frac{z^{r-2}}{r!} = \frac{e^z - 1 - z}{z^2}$. Rearranging, we obtain the desired result. □

*Proof of Bernstein's inequality.* Observe that it suffices to prove that

$$\mathbb{P}(\overline{X} \geq t) \leq \exp\left( - \frac{nt^2}{\frac{2\sigma^2}{n} + 2t/3}\right),$$

since then we can apply this to $-\overline{X}$ and obtain Theorem 2.1.

By the previous lemma, we have $\mathbb{E}(e^{zX_i}) \leq \exp \sigma_i^2(e^z - z - 1)$. Thus, by Markov's inequality and independence we have

$$\begin{aligned}\mathbb{P}(\overline{X} \geq t) = \mathbb{P}(\sum X_i \geq nt) &= \mathbb{P}(\exp(z \sum X_i) \geq e^{nzt}) \leq e^{-nzt}\mathbb{E}\exp(z\sum X_i) \\ &= e^{-nzt} \prod \mathbb{E}\exp(zX_i) \leq e^{-nzt} \prod \exp(\sigma_i^2(e^z - z - 1)) \\ &= e^{-nzt}\exp(\sigma^2(e^z - z - 1)).\end{aligned}$$

This holds for all $z > 0$. We take $z = \log(1 + tn/\sigma^2)$. Expanding out everything, we see that $\mathbb{P}(\overline{X} \geq t) \leq \exp(-\sigma^2 h(nt/\sigma^2))$, where $h(u) = \log(1 + u) + u\log(1 + u) - u$. Observing the graphs, we see that $h(u) \geq u^2/(2 + 2u/3)$, which finishes the proof after expanding out everything again.

□

## 3. *Analytic and combinatorial preliminaries: Rudin's inequality and Chang's lemma*

In this section we prove Chang's lemma on the large spectrum, which heuristically states that the set of large Fourier coefficients of a small set is highly structured, in the sense that its largest "unstructured" subset is quite small (this can be extracted from its proof).

We call a subset $A = \{a_1, \dots, a_n\}$ of an abelian group *dissociated* if $\epsilon_i \in \{-2, -1, 0, 1, 2\}$ and $\sum \epsilon_i a_i = 0$ implies $\epsilon_i = 0$ for all $i$. In other words, there is no nontrivial additive relation (i.e., solutions to $a_1 + \ldots + a_j = a'_1, \ldots + a'_k$) between elements of $A$. Thus, a dissociated set has basically no additive structure. Chang's lemma [**2**] states the following:

Theorem 3.1. *(Chang) Let $A \subseteq G$, $|A| = \delta N$, $\Gamma = \{r \in G : |\hat{A}(r)| \geq \rho\delta\}$. Then there is a $\Lambda \subseteq \Gamma$ with $|\Lambda| \leq 6000\rho^{-2}\log(1/\delta)$ such that each element of $\Gamma$ can be written as a linear combination of elements of $\Lambda$ with coefficients $\pm 1$ or $0$.*

I think it is instructive to think about what the bounds mean. By Parseval (dividing the sum into large and small Fourier coefficients), $\Gamma \leq \rho^{-2}\delta^{-1}$. As $\delta$ goes to zero, $\log(1/\delta)$ is much smaller than $\delta^{-1}$. From the proof, it will turn out that $\Lambda$ is obtained from (modulo a small technicality) the largest dissociated subset of $\Gamma$. Therefore the largest "unstructured" subset of $\Gamma$ is much smaller than $|\Gamma|$, so the heuristic interpretation $\Gamma$ is highly structured makes sense.

In view of the fact that Bohr sets are rigid under linear combinations with small coefficients, it can be seen that how this result will be useful: In a Bohr neighborhood of the large spectrum of a set, we can find another Bohr set with much smaller dimension, which gives a longer arithmetic progression than we would get if we only used Lemma 1.3.

In order to prove Chang's lemma, we will first prove a special case of Young's convolution inequality and Khintchine's inequality, then we will use these to prove Rudin's inequality, which we will use in proving Chang's lemma.

Lemma 3.2. *(Young) If $f, g : G \to \mathbb{C}$ and $p \geq 1$, then $\|f * g\|_p \leq \|f\|_1\|g\|_p$.*

*Proof.* This proof is probably standard, and unfortunately I do not remember where I took it from. Let $p' = p/(p-1)$ so that $1/p + 1/p' = 1$. For all $x$, by Hölder we have

$$\begin{aligned}|f * g(x)| &\leq \mathbb{E}_y|f(x-y)g(y)| = \mathbb{E}_y|f(x-y)|^{1/p'}|f(x-y)|^{1/p}|g(y)| \\ &\leq \Big(\mathbb{E}_y|f(x-y)|\Big)^{1/p'}\Big(\mathbb{E}_{y'}|f(x-y')||g(y')|^p\Big)^{1/p},\end{aligned}$$

and noting that the first term on the RHS is $\|f\|_1^{1/p'}$, we have

$$\begin{aligned}\|f*g\|_p &= \Big(\mathbb{E}_x|f*g(x)|^p\Big)^{1/p} \le \|f\|_1^{1/p'}\Big(\mathbb{E}_x\mathbb{E}_y|f(x-y)||g(y)|^p\Big)^{1/p}\\ &= \|f\|_1^{1/p'}\Big(\mathbb{E}_y\mathbb{E}_x|f(x-y)||g(y)|^p\Big)^{1/p} = \|f\|_1^{1/p'}\Big(\mathbb{E}_y|g(y)|^p\mathbb{E}_x|f(x-y)|\Big)^{1/p}\\ &= \|f\|_1^{1/p'}\|g\|_p\|f\|_1^{1/p} = \|f\|_1\|g\|_p, \qquad (3.1)\end{aligned}$$

□

We now prove (one side of) Khintchine's inequality, which heuristically states that the $L^p$ norm of a sum of random variables with random signs is comparable to its $L^2$ norm. We follow the proof in [**11**].

Lemma 3.3. *(Khintchine) Let $p > 1$. Then there is a constant $C_p$ such that the following holds: Let $\epsilon_i$ be independent random variables taking values $\pm 1$ with equal probability (known as Rademacher random variables) and $a_1, \dots, a_n \in \mathbb{C}$. Then $\mathbb{E}|\sum \epsilon_i a_i|^p \le C_p\Big(\sum |a_i|^2\Big)^{p/2}$.*

Note that $C_p$ depends only on $p$, in particular, it does not depend on the sequence $a_n$ or even the length $n$ of the sequence. Observe that by expanding out the square and using linearity of expectation, we can write the sum on the RHS as $\|\sum \epsilon_i a_i\|_2^2$. So another formulation of the inequality would be $\|\sum \epsilon_i a_i\|_p \le C_p\|\sum \epsilon_i a_i\|_2$ (with a different $C_p$). Here, we can take $C_p = 8\sqrt{p}$.

It can also be proved that $\|\sum \epsilon_i a_i\|_p \ge c_p\|\sum \epsilon_i a_i\|_2$ (justifying the word 'comparable' in the heuristic interpretation above), but we will not need this.

*Proof.* By Taylor expansion and noting that $n!2^n \le (2n)!$, we have $e^x + e^{-x} \le 2e^{x^2/2}$. Thus, for real $a_i$, for any $t$ we have

$$\mathbb{E}e^{t\sum \epsilon_i a_i} = \prod \mathbb{E}e^{t\epsilon_i a_i} \le e^{t^2/2\sum a_i^2}.$$

By Markov, for any $\lambda > 0$, $\mathbb{P}(\sum \epsilon_i a_i \ge \lambda) = \mathbb{P}(\exp(t\sum \epsilon_i a_i) \ge e^{t\lambda}) \le \mathbb{E}\exp(t\sum \epsilon_i a_i - t\lambda)$. Taking $t = \frac{\lambda}{\sum a_i^2}$ and using the inequality above, we obtain

$$\mathbb{P}(\sum \epsilon_i a_i \ge \lambda) \le \exp\Big(-\frac{\lambda^2}{2\sum a_i^2}\Big).$$

Therefore, by symmetry, $\mathbb{P}(|\sum \epsilon_i a_i| \ge \lambda) \le 2\exp\Big(-\frac{\lambda^2}{2\sum a_i^2}\Big)$ for real $a_i$.

Assume now that $a_j$ are complex, with $a_j = u_j + iv_j$ for real $u_j$, $v_j$ (this is the only time we use $i$ for the imaginary unit). Then $|\sum \epsilon_i a_i|^2 = |\sum \epsilon_i u_i|^2 + |\sum \epsilon_i v_i|^2$. Observe that $\sum |a_i|^2 = \sum u_i^2 + \sum v_i^2$, so $|\sum \epsilon_i a_i| \ge \lambda$ implies either $|\sum \epsilon_i u_i| \ge \lambda\Big(\frac{\sum u_i^2}{\sum |a_i|^2}\Big)^{1/2}$ or $|\sum \epsilon_i v_i| \ge \lambda\Big(\frac{\sum v_i^2}{\sum |a_i|^2}\Big)^{1/2}$. Thus, invoking the result for the real case we have

$$\begin{aligned}\mathbb{P}(|\sum \epsilon_i a_i| \ge \lambda) &\le \mathbb{P}\Big(|\sum \epsilon_i u_i| \ge \lambda\Big(\frac{\sum u_i^2}{\sum |a_i|^2}\Big)^{1/2}\Big) + \mathbb{P}\Big(|\sum \epsilon_i v_i| \ge \lambda\Big(\frac{\sum v_i^2}{\sum |a_i|^2}\Big)^{1/2}\Big)\\ &\le 4\exp\Big(-\frac{\lambda^2}{2\sum |a_i|^2}\Big).\end{aligned}$$

We can express the $L^p$ norm of a function on an arbitrary $\sigma$-finite probability space $(X, \mathcal{A}, \mu)$ as an integral over the real line as follows (this proof is in [**4**]):

$$\begin{aligned}\mathbb{E}|f|^p &= \int_X |f(x)|^p d\mu(x) = \int_X \int_0^{|f(x)|} p\alpha^{p-1} d\alpha d\mu\\ &= p\int_0^\infty \alpha^{p-1}\int_X 1_{\{|f|\ge\alpha\}} d\mu d\alpha = p\int_0^\infty \alpha^{p-1}\mathbb{P}(|f| \ge \alpha)d\alpha.\end{aligned}$$

Taking $f = \sum \epsilon_i a_i$ and making the substitution $\alpha = (2\sum |a_i|^2 x)^{1/2}$, $d\alpha = \frac{(2\sum |a_n|^2)^{1/2}}{2\sqrt{x}}dx$, we obtain

$$\begin{aligned}\mathbb{E}(|\sum \epsilon_i a_i|^p) &= p\int_0^\infty (2\sum |a_i|^2 x)^{(p-1)/2}\mathbb{P}\Big(|\sum \epsilon_i a_i| \geq (2\sum |a_i|^2 x)^{1/2}\Big)\frac{(2\sum |a_n|^2)^{1/2}}{2\sqrt{x}}dx \\ &= p(2\sum |a_i|^2)^{p/2}\int_0^\infty x^{(p-2)/2}\mathbb{P}\Big(|\sum \epsilon_i a_i| \geq (2\sum |a_i|^2 x)^{1/2}\Big)dx \\ &\leq 2p(2\sum |a_i|^2)^{p/2}\int_0^\infty x^{(p-2)/2}e^{-x}dx = 2p(2\sum |a_i|^2)^{p/2}\Gamma(p/2). \\ &= 2^{(p+2)/2}p\Gamma(p/2)(\sum |a_i|^2)^{p/2},\end{aligned}$$

which completes the proof. To obtain the precise estimate for $C_p$, we take $1/p$'th powers and observe by looking at the graphs of the relevant functions that $2^{(p+2)/2p} \leq 4$, $p^{1/p} \leq 2$, $\Gamma(p/2) \leq \sqrt{p}$. Therefore $\|\sum \epsilon_i a_i\|_p \leq 8\sqrt{p}\|\sum \epsilon_i a_i\|_2$, as claimed. □

We now prove Rudin's inequality (often [**8**] is cited for this, but I have failed to find the exact statement in there, although there seem to be some related ideas), following the proof in [**10**]. Heuristically, the inequality states that if a function has Fourier transform with dissociated support, then all its $L^p$ norms are controlled by its $L^2$ norm. Recalling that Khintchine states the same for Rademacher random variables, this can be interpreted as saying that characters from a dissociated set behave similarly to Rademacher random variables.

Lemma 3.4. *(Rudin) Let $p > 1$, $S \subseteq G$ be a dissociated set, $f : G \to \mathbb{C}$ with* $\mathrm{supp}(\hat{f}) \subseteq S$. *Then $\|f\|_p \leq 16\sqrt{p}\|f\|_2$.*

By Fourier inversion, this is equivalent to

$$\Big(\mathbb{E}_{x\in G}|\sum_{r\in S}\hat{f}(r)\omega^{rx}|^p\Big)^{1/p} \leq 16\sqrt{p}\Big(\mathbb{E}_{x\in G}|\sum_{r\in S}\hat{f}(r)\omega^{rx}|^2\Big)^{1/2}.$$

The method of proof is unlike any analysis proof I have ever seen. It is proved that when we randomize the sign of the Fourier coefficients, on average the inequality holds. Then it is proved that the $L^p$ norm of these randomized functions are not much larger than of $f$, proving the theorem. I do not know how one would come up with this proof.

*Proof.* For $r \in S$, let $\epsilon_r$ be independent Rademacher random variables. Define $f_\epsilon(x) = \sum_{r\in S}\epsilon_r\hat{f}(r)\omega^{rx}$ (note that the value of $f_\epsilon$ at each $x$ is a random variable). For all $x \in G$, by Khintchine (with $a_i = \hat{f}(i)\omega^{ix}$), noting that $|\epsilon_r| = |\omega^{rx}| = 1$, and Parseval we have

$$\begin{aligned}\mathbb{E}|f_\epsilon(x)|^p = \mathbb{E}|\sum_{r\in S}\epsilon_r\hat{f}(r)\omega^{rx}|^p &\leq (8\sqrt{p})^p\Big(\sum_{r\in S}|\hat{f}(r)|^2\Big)^{p/2} \\ &= (8\sqrt{p})^p\|f\|_2^p,\end{aligned}$$

where the expectation is taken over the choices for $\epsilon_r$. Note that RHS does not depend on $x$. Averaging over $x$ and using linearity of expectation, we obtain $\mathbb{E}\|f_\epsilon\|_p^p \leq (8\sqrt{p})^p\|f\|_2^p$. Therefore, there is some assignment of ±1 to $\epsilon_r$ such that $\|f_\epsilon\|_p \leq 8\sqrt{p}\|f\|_2$. From now on we fix $\epsilon$ as this assignment. We will do a long computation to obtain $\|f\|_p \leq 2\|f_\epsilon\|_p$.

Let

$$\begin{aligned}p_\epsilon(x) = \prod_{r\in S}(1 + \frac{\epsilon_r}{2}\omega^{rx} + \frac{\epsilon_r}{2}\omega^{-rx}) &= \sum_{\substack{X,X'\subseteq S\\ X,X'\, disjoint}}\prod_{r\in X}\frac{\epsilon_r}{2}\omega^{rx}\prod_{r'\in X'}\frac{\epsilon_r'}{2}\omega^{-r'x} \\ &= \sum_{\substack{X,X'\subseteq S\\ X,X'\, disjoint}}\Big(\prod_{r\in X\cup X'}\frac{\epsilon_r}{2}\Big)\omega^{(\sum_{r\in X} r - \sum_{r'\in X'} r')x}.\end{aligned}$$

Observe that $p_\epsilon$ is nonnegative, so $\|p_\epsilon\|_1 = \mathbb{E}p_\epsilon(x)$ (note that our setting is deterministic now, so expectation symbols denote averaging). Thus,

$$\|p_\epsilon\|_1 = \sum_{\substack{X,X' \subseteq S \\ X,X'\, disjoint}} \left( \prod_{r \in X \cup X'} \frac{\epsilon_r}{2} \right) \mathbb{E}_{x \in G} \omega^{(\sum_{r \in X} r - \sum_{r' \in X'} r')x}.$$

The inner expectation is one or zero depending on whether $\sum_{r \in X} r = \sum_{r' \in X'} r'$ or not. However, by dissociativity, the only way this can happen is $X = X' = \varnothing$, so $\|p_\epsilon\| = 1$.

We claim that $f = 2f_\epsilon * p_\epsilon$. By Fourier inversion and convolution identity it suffices to prove that $\hat{f} = 2\hat{f_\epsilon}\hat{p_\epsilon}$. By hypothesis $\hat{f}(r_0) = 0$ unless $r_0 \in S$. Also $\hat{f_\epsilon}(r_0) = 2\epsilon_{r_0}\hat{f}(r_0)$ if $r_0 \in S$ and zero otherwise. Now consider $\hat{p_\epsilon}(r_0)$ for $r_0 \in S$. By definition, this is equal to

$$\sum_{\substack{X,X' \subseteq S \\ X,X'\, disjoint}} \left( \prod_{r \in X \cup X'} \frac{\epsilon_r}{2} \right) \mathbb{E}_{x \in G} \omega^{(\sum_{r \in X} r - \sum_{r' \in X'} r' - r_0)x}.$$

By orthogonality and dissociativity, the only contribution to this sum are when $X = \{r_0\}$ and $X' = \varnothing$, so the sum is $\epsilon_{r_0}/2$. Note that this is why we also exclude coefficients ±2 in our definition of dissociativity, as the sum multiplying $x$ may include a coefficient $-2$ for $r_0$. If we used the other definition (only excluding ±1) of dissociativity, there could be other contributions to this sum. Consider for example $S = \{1, 2\}$. Then $1 = 2 - 1$, so we get unwanted contributions.

Now, combining everything we have, $2\hat{f_\epsilon}(r_0)\hat{p_\epsilon}(r_0) = 2\epsilon_{r_0}\hat{f}(r_0)\epsilon_{r_0}/2 = \hat{f}(r_0)$, as claimed. Therefore, using Young's inequality and commutativity of convolution, $\|f\|_p \leq 2\|f_\epsilon * p_\epsilon\|_p \leq 2\|f_\epsilon\|_p\|p_\epsilon\|_1 = 2\|f_\epsilon\|_p \leq 2 \cdot 8\sqrt{p}\|f\|_2 = 16\sqrt{p}\|f\|_2$, which completes the proof. □

*Proof of Chang's lemma.* Let $E \subseteq \Gamma$ be a dissociated subset. Define $a_r = \frac{\hat{A}(r)}{\sqrt{\sum_{r \in E} |\hat{A}(r)|^2}}$ and $g(x) = \sum_{r \in E} a_r \omega^{rx}$. Then, for any $p > 1$, by Rudin and Parseval, we have $\|g\|_p \leq 16\sqrt{p}\|g\|_2 = 16\sqrt{p}\|\hat{g}\|_{\ell^2} = 16\sqrt{p}$. Write $q = p/(p-1)$ so that $1/p + 1/q = 1$. Then $\|A\|_q = \delta^{1/q}$. Then, since $E \subseteq \Gamma$, by Plancherel and Hölder we have

$$\begin{aligned} \rho\delta\sqrt{|E|} &\leq \sqrt{\sum_{r \in E} |\hat{A}(r)|^2} = \frac{\sum_{r \in E} |\hat{A}(r)|^2}{\sqrt{\sum_{m \in E} |\hat{A}(m)|^2}} = |\sum_{r \in E} \hat{A}(r)\overline{a_r}| = |\langle \hat{A}, \hat{g} \rangle_{\ell^2}| = |\langle A, g \rangle_{L^2}| \\ &\leq \|A\|_q\|g\|_p \leq 16\sqrt{p}\delta^{1/q}. \end{aligned}$$

Taking $p = \log(1/\delta)$, we obtain $|E| \leq 256\rho^{-2}\delta^{-2}\delta^{-2/\log(1/\delta)} = 256e^2\rho^{-2}\log(1/\delta)$. Note that $256e^2 \leq 2000$.

Let $E_0$ be a maximal dissociated subset of $\Gamma$. Then for any $s \in \Gamma \setminus E_0$, $E_0 \cup \{s\}$ is not dissociated. Therefore, if $s \in \Gamma \setminus E_0$, there is a nontrivial relation of the form $\epsilon_s s + \sum \epsilon_i e_i = \epsilon'_s s + \sum \epsilon'_i e_i$, where $e_i$ are the elements of $E_0$ and $\epsilon_i$ are zero or ±1. Since $E_0$ is dissociated, $\epsilon_s \neq \epsilon'_s$. Thus $s$ is a linear combination of the elements of $E_0$ with coefficients 0, ±1/2,±1 or ±2. Taking $\Lambda = E_0/2 \cup E_0 \cup 2E_0$, we finish the proof. □

There is another proof of Chang's lemma due to Ruzsa, not relying on Rudin's inequality (in fact Rudin's inequality can be deduced from it), given in [**6**], which seems to be not very well-known. I have not attempted to understand it, but it seems to use less heavy machinery.

## 4. *Proof of Theorem 1.4*

### 4.1. *Brief outline of proof (again)*

Let us recall what we are going to do (in a bit more detail than the brief outline in the first section). Let $A$ be our $\alpha$-HNU set. We fix some density $\beta$ depending on $\alpha$ (we will have $\beta = \exp(-c\alpha\sqrt{\log N})$ for some suitable constant $c$), consider the subset $B \subseteq A$ of size $\beta N$ minimizing the largest nontrivial Fourier coefficient (a very small technicality: $\beta$ need not be a multiple of $1/N$, but when $N$ is large enough (even, say, $N > 100$), which is the regime we are interested in, $\lfloor \beta N \rfloor$ is within 0.99 and 1.01 times $\beta N$, and the argument will still work. Thus, for convenience we will assume $\beta$ is a multiple of $1/N$. It might be the case that slight alteration of the constants appearing is needed to make this fully rigorous, but surely it will work.), and let $\eta\beta$ be this minimum (since $A$ is $\alpha$-HNU, $\eta \geq \alpha$). We use Chang's lemma to find a long AP $P$ in a Bohr neighborhood of the large spectrum of B. Most translates of $P$ do not have too large intersection with $B$. We let $C$ be the set of such translations.

Since $C$ is very large, it will not have large nontrivial Fourier coefficients. Then, taking elements at random, we can find a small subset $D$ of $C$ which also does not have large nontrivial Fourier coefficients. By the same probabilistic argument there will be a subset $X$ of $B$ with the same size as $D$ and with Fourier coefficients about the same size as $B$. It turns out that however we translate the elements of $D$ by elements of $P$ to obtain a set $D'$, the multiset (i.e., we allow multiple elements, so this corresponds to adding and subtracting indicator functions) $D' \cup B \setminus X$ has smaller largest Fourier coefficient than $\eta\beta$, so if it were a genuine set, we would obtain a contradiction.

It follows that $A^c \setminus B$ contains a very large proportion of a translate of $P$, which, after some calculations and optimization of parameters, gives an AP of desired length in $A^c$. Now we make this discussion rigorous.

## 4.2. *Using Chang's lemma*

Lemma 4.1. *Let $R = \{r \in G : |\hat{B}(r)| \geq \eta\beta/2\}$. Then $B(R, \eta/64)$ contains an AP $P$ of length $\geq \eta^3 N^{\eta^2/24000 \log(1/\beta)}/2^{21} \log(1/\beta)$.*

*Proof.* By Chang, each element of $R$ is in the $\pm 1$-span of some $\Lambda$ of size $m \leq 6000(\frac{\eta}{2})^{-2} \log(1/\beta) = 24000\eta^{-2} \log(1/\beta)$. By triangle inequality, $B(\Lambda, \eta/64m) \subseteq B(R, \eta/64)$. By Lemma 1.3, $B(\Lambda, \eta/64m)$ contains an AP $P$ of size at least $\eta N^{1/m}/64m \geq \eta^3 N^{\eta^2/24000 \log(1/\beta)}/2^{21} \log(1/\beta)$ (since $24000 < 2^{15}$). □

## 4.3. *Finding the set C*

Lemma 4.2. *For at least $(1 - \eta/16)N$ values of $x \in G$, we have $|(x + P) \cap B| \leq 16\beta|P|/\eta$.*

*Proof.* If not, then we would have $|P||B| = \sum_{x \in G} |(x + P) \cap B| > \frac{\eta}{16} N \cdot 16\frac{\beta}{\eta}|P| \geq |P||B|$, which is absurd. □

Let $C$ be the set of $(1 - \eta/16)N$ values of $x$ in this lemma. Then, since $\hat{1}(r) = 0$ for all $r \neq 0$ and $|C|/N \geq 1/2$, we have $|\hat{C}(r)| \leq |C^c|/N \leq \eta/16 \leq \eta|C|/8N$ for $r \neq 0$. Since the largest possible size of a Fourier coefficient of $C$ is $|C|/N$, this means all nontrivial Fourier coefficients of $C$ are at most $\eta/8$ times as large as possible, so they are quite small. Remember that we are aiming to eventually construct a (multi-)set whose all nontrivial Fourier coefficients are less than $\eta$.

## 4.4. *Using Bernstein's inequality and finding the sets D and X*

Lemma 4.3. *Let $3|C|/4 \geq t > 15000\eta^{-2} \log N$. Then there is a subset $D \subseteq C$ of size $t$ such that $|\hat{D}(r)| \leq \eta t/4N$ for all $r \neq 0$.*

*Proof.* Construct a set $E \subseteq C$ randomly as follows: For each $x \in C$, include $x$ in $E$ independently with probability $t/|C|$. Consider the rescaled Fourier coefficients $N\hat{E}(r)$. Each of these is the sum of $|C|$ independent random variables $E(x)\omega^{-rx}$ with mean $tC(x)/|C|$, variance $\frac{t}{|C|} \cdot (1 - \frac{t}{|C|}) \leq t/|C|$. By Bernstein (with $t = \eta t/(24|C|)$), provided that (after some straightforward computation) $t \leq |C|(1 - \eta/4)$, we have

$$\mathbb{P}\left(\left|\hat{E}(r) - \frac{t\hat{C}(r)}{|C|}\right| \geq \eta t/24N\right) \leq 4\exp(-n^2 t/4608) \leq 4\exp(-n^2 t/5000). \tag{4.1}$$

Observe that, if $r \neq 0$, since $|\hat{C}(r)| \leq \eta|C|/8N$,

$$|\hat{E}(r)| \geq \eta t/6N \text{ implies } \left|\hat{E}(r) - \frac{t\hat{C}(r)}{|C|}\right| \geq \eta t/24N.$$

Writing $r = 0$ in (4.1), we obtain $\mathbb{P}(||E| - t| \geq \eta t/24) \leq 4\exp(-n^2 t/5000)$. We claim that if $t \geq 15000\eta^{-2}\log N$, then $\mathbb{P}(||E| - t| \geq \eta t/24)$ and $|\hat{E}(r)| \geq \eta t/6N$ for all $r \neq 0$ with positive probability. By the union bound, the probability of all these happening is $\geq 1 - N \cdot 4N^{1-15000/5000} = 1 - 4/N > 0$ (for $N > 4$). Therefore there exists a set $E_0$ with size $t(1 - \eta/24) \leq |E_0| \leq t(1 + \eta/24)$ and $|\hat{E}_0(r)| \leq \eta t/6N$ for all $r \neq 0$.

Now, to obtain $D$, we can add or delete elements from $E_0$ arbitrarily to get the correct size, since $|E \Delta D| = k$ implies $|\hat{E}(r) - \hat{D}(r)| \leq k/N$. Since $6 + 1/24 \leq 1/4$, we are done. □

The following lemma has the exact same proof as the previous, so we omit its proof.

Lemma 4.4. *Let $\beta N \geq t \geq 15000\eta^{-2}\log N$. Then there exists a subset $X \subseteq B$ of size $t$ such that $\left|\hat{X}(r) - \frac{t}{|B|}\hat{B}(r)\right| \leq \eta t/12N$.*

### 4.5. *Finding a contradictory multiset $S'$*

Lemma 4.5. *Let $S$ be the multiset $D \cup B \setminus X$. Then, for all $r \in R$, we have $|\hat{S}(r)| \leq \eta|B|/N - \eta t/6N$, and for all $r \notin R \cup \{0\}$, we have $|\hat{S}(r)| \leq \eta|B|/2N + \eta t/3N$.*

Note that $|S| = |B|$.

*Proof.* By linearity, $\hat{S}(r) = \hat{B}(r) - \hat{X}(r) + \hat{D}(r)$. Using the two lemmas in the previous section, we obtain $\hat{S}(r) = (1 - t/|B|)\hat{B}(r) + Q(r)$, where $|Q(r)| \leq \eta t/3N$. If $r \in R$, then, by definition of $R$ we have $|\hat{B}(r)|/|B| \geq \eta/2N$, so $|\hat{S}(r)| \leq \eta|B|/N - \eta t/6N$.

If $r \notin R \cup \{0\}$, then $|\hat{S}(r)| \leq \eta|B|/2 + \eta t/3$, again by definition of $R$. □

Let $D = \{d_1, \ldots, d_t\}$ and $D'$ be any set obtained by replacing each $d_i$ with $d_i + x_i$ for some $x_i \in P$. Let $S' = D' \cup B \setminus X$ (as a multiset). If $t$ is small enough, we will see that $S'$ will be a contradictory multiset. Note again that $|S'| = |B|$.

Lemma 4.6. *Assume $t \leq \eta\beta N/10 = \eta|B|/10$. Then for any $r \neq 0$, we have $|\hat{S}'(r)| < \eta|B|/N$.*

*Proof.* Since $|S \Delta S'| = |D \Delta D'| \leq 2t$, if $r \notin R \cup \{0\}$ we have $|\hat{S}'(r)| \leq |\hat{S}(r)| + 2t/N \leq \eta|B|/2N + \eta t/3N + 2t/N \leq \eta|B|/2N + 3t/N$, which is $\leq \eta|B|/N$ provided that $t \leq \eta\beta N/6$.

If $r \in R$, then, since $P \subseteq B(R, \eta/64)$, we have $|\hat{S}'(r) - \hat{S}(r)| \leq \frac{1}{N}\sum_{j=1}^{t} |\omega^{r(d_j + x_j)} - \omega^{rd_j}| \leq \frac{t}{N} \cdot 2\pi\eta/64 \leq \eta t/8N$. Combining with the previous lemma, we obtain $|\hat{S}(r)| \leq \eta|B|/N$. □

Now, if $S'$ were a genuine set (i.e., all its elements are distinct), then this would contradict the minimality of $\eta$. It follows that there is no choice of $x_i$'s making $S'$ a set.

### 4.6. *Finding a large proportion of a long AP in $A^c$*

Lemma 4.7. *There is some $j$ such that $d_j + P \subseteq B \cup A^c$, except for at most $t$ elements.*

*Proof.* Assume for contradiction that the statement is false. Then for all $j$, $|(d_j + P) \cap (A \setminus B)| \geq t$. Pick $x_1 \in (d_1 + P) \cap (A \setminus B)$, $x_2 \in (d_2 + P) \cap (A \setminus B)$ not equal to $x_1$, and so on, with each $x_i$ in $(d_i + P) \cap (A \setminus B)$ and not equal to $x_1, \ldots, x_{i-1}$. This is possible because of our assumption.

However, $d_i + x_i$ are all distinct, picking $D'$ as the set of these numbers and noting that $D' \cap B = \varnothing$, we see that this gives a $S'$ which is a genuine set, which is absurd. □

Pick some $j$ satisfying the conclusion of the above lemma. Since $d_j \in D \subseteq C$, we have $|(d_j + P) \cap B| \leq 16\beta|P|/\eta$.

Thus, $|(d_j + P) \cap A^c| \geq (1 - 16\beta/\eta)|P| - t$. If we choose $t$ so that $t \leq 16\beta|P|/\eta$, then $A^c$ contains at least $|(1 - 32\beta/\eta)|P|$ elements of $d_j + P$. We see that this implies $A^c$ contains an AP of length at least $\eta/64\beta$ provided that $|P| \geq 32\beta/\eta$, by using the following combinatorial lemma:

Lemma 4.8. *If $P$ is an arithmetic progression, $\epsilon \geq 1/|P|$ and $S$ contains at least $(1 - \epsilon)|P|$ elements of $P$. Then $S$ contains an arithmetic progression of length at least $1/2\epsilon$.*

*Proof.* Partition $S \cap P$ into consecutive runs $P_1, \ldots, P_k$. By hypothesis, $k \leq \epsilon|P|$. We have $\sum_{i=1}^{k} |P_i| \geq (1 - \epsilon)|P|$. If the lemma were false, then we would have $(1 - \epsilon)|P| < \epsilon|P|/2\epsilon = |P|/2$, so $\epsilon > 1/2$, but in that case the lemma is true trivially, since $1/2\epsilon \leq 1$. □

### 4.7. *Finishing the proof and choosing parameters*

In this section we carry out a sketch (just the last steps of each computation) of very unpleasant computations to finish the proof. It is likely that there are slight numerical inaccurracies in this section, but they should not be very serious. Recall the following: We require $t \leq 16\beta|P|/\eta$, $t \geq 15000\eta^{-2}\log N$, $t \leq \eta\beta N/10$, $|P| \geq \eta/32\beta$.

Pick $t = 15000\eta^{-2}\log N$, $\beta = \exp(-c\alpha\sqrt{\log N})$, with $c = 1/15000 \cdot 2^{17}$. Recall that $\eta \geq \alpha \geq c' \log\log N/\sqrt{\log N}$.

If $\beta \geq \alpha$, then $\eta$ is undefined, but in this case $A^c$ contains a progression of length at least $1/2\beta \geq \eta/64\beta$ by the above lemma.

To prove the first requirement, taking logs and rearranging we find that $(c'(15000 \cdot 2^{17}/24000 - c) - 3/2)\log\log N \geq 0$, which is true for $c' > 1$.

The second requirement is trivial. Again by taking logs, the third can be seen to be implied by $(5/2 + cc')\log\log N \leq 3\log c' + 3\log\log\log N + \log N$, which is always true when $c' > 1$. The fourth requirement is immediate from the first one, since $t > 1/2$.

It remains to show that $\log(\eta/64\beta) \geq c''\alpha\sqrt{\log N}$ for some $c''$. Expanding everything, we see that this is implied by $\log\log\log N + \log(c'/64) \geq (c'(c'' - c) + 1/2)\log\log N$, so taking $c'' = c/2$ and $c' = 1/c$, we complete the proof.

## *References*

**1.** J. Bourgain, On arithmetic progressions in sums of sets of integers, in "A Tribute to Paul Erdős", CUP (1990), 105–109.

**2.** M.-C. Chang, A polynomial bound in Freiman's theorem, Duke Math J. 113(3) (2002), 399–419.

**3.** E. Croot, I. Laba, and O. Sisask, Arithmetic progressions in sumsets and $L^p$-almost periodicity, Combinatorics, Probability and Computing 22(3) (2013), 351–365.

**4.** L. Grafakos, Classical Fourier Analysis (Third edition), Graduate Texts in Mathematics vol. 249, Springer, 2014.

**5.** B. Green, Arithmetic progressions in sumsets, Geom. Funct. Anal. GAFA 12(3) (2002), 584–597.

**6.** B. Green, Structure Theory of Set Addition, notes from ICMS Instructional Conference in Combinatorial Aspects of Mathematical Analysis, 2002.

**7.** J. Lafferty, H. Liu, and L. Wasserman, Excerpt from book in preparation "Statistical Machine Learning", available at https://www.stat.cmu.edu/ larry/=sml/Concentration.pdf

**8.** W. Rudin, Fourier Analysis on Groups, Wiley, 1990 (reprint of 1962 original)

**9.** I. Z. Ruzsa, Arithmetic progressions in sumsets, Acta Arithmetica 60(2) (1991) 191–202.

**10.** F. Tyrrell, Almost-Periodicity in Additive Number Theory, MMath dissertation, Oxford, 2023.

**11.** T. H. Wolff, Lectures in Harmonic Analysis, AMS University Lecture Series vol. 29, 2002.

*Bora Çalım*
*Koç University*

bcalim21@ku.edu.tr
bcalim06@gmail.com