10.2 The small torsion case

We now use the above Fourier-analytic methods and the density increment argument to prove the following simple special case of Roth's theorem.

Proposition 10.12 (Roth's theorem for p-torsion groups) *[248] Let Z be a p-torsion group (thus $px = 0$ for all $x \in Z$) for some odd prime p. Then*

$$r_3(Z) < \frac{3}{\log_p |Z|}|Z|.$$

Remark 10.13 Define a *capset* to be a subset of the vector space $\mathbf{Z}_3^n$ which contains no lines. Then the above proposition implies that capsets have density less than $3/n$. Rather amazingly, this simple bound is essentially the best known (other than improving the constant 3); in the converse direction, the best lower bound known on the density of capsets in $\mathbf{Z}_3^n$ is $(0.724581\ldots + o(1))^n$; see [75]. Any improvement of the upper bound to $o(1/n)$, or the lower bound to $(1 - o(1))^n$, would be a significant advance in our understanding of the Erdős–Turán conjecture.

Remark 10.14 A useful heuristic is that the cyclic group $\mathbf{Z}_N$ (or the interval $[1, N]$) should behave roughly like the p-torsion group $\mathbf{Z}_p^n$ whenever $N \sim p^n$. Using this heuristic and the above proposition, one would expect that $r_3([1, N])$ and $r_3(\mathbf{Z}_N)$ should be $O(N/\log N)$. Such a bound would essentially be equivalent to the Erdős–Turán conjecture (Conjecture 10.6) in the $k = 3$ case. Unfortunately the direct analog of the above argument gives $r_3([1, N]), r_3(\mathbf{Z}_N) = O(N\sqrt{\frac{\log\log N}{\log N}})$, see Theorem 10.30. In general, the p-torsion groups are somewhat easier to analyze than general groups, due to their vector space structure over the field F_p. To extend the p-torsion arguments to more general settings, one needs some additional machinery, in particular the theory of Bohr sets.

We now begin the proof of Proposition 10.12. We may view Z as a vector space over F_p. Assume for contradiction that we can find a set $A \subset Z$ of density $\mathbf{P}_Z(A) \geq \frac{3}{\log_p |Z|}$ which has no proper progressions of length 3. From Corollary 10.10 we already know that A must exhibit linear bias, thus $\|A\|_u$ is large. To use this fact, we need to convert linear bias to a more useful structural property. This is achieved as follows.

Lemma 10.15 (Non-uniformity implies density increment) *Let Z be a vector space over a finite field F_p of prime order, and let $f : Z \to \mathbf{R}$ be a function with mean zero, $\mathbf{E}_Z(f) = 0$. Then there exists a subspace Z' of Z of codimension 1 over F_p, and a point $x_0 \in Z$, such that*

$$\mathbf{E}_{x \in x_0 + Z'} f(x) \geq \frac{1}{2}\|f\|_{u^2(Z)}.$$

Proof Without loss of generality we may take $Z = F_p^n$, and use the bilinear form in Example 4.2.

By definition of $\|f\|_{u^2(Z)}$ and the mean zero hypothesis, we can find a non-zero $\xi \in Z$ and a phase $\theta \in \mathbf{R}/\mathbf{Z}$ such that

$$\mathrm{Re}\, \mathbf{E}_{y \in Z} f(y) e(\xi \cdot y + \theta) = \|f\|_{u^2(Z)},$$

where e is the exponential map defined by equation (4.1). Applying the mean zero hypothesis again, we conclude

$$\mathrm{Re}\, \mathbf{E}_{y \in Z} f(y)(e(\xi \cdot y + \theta) + 1) = \|f\|_{u^2(Z)}$$

Let $Z' := \{\xi\}^\perp = \{x \in Z : \xi \cdot x = 0\}$ be the orthogonal complement of ξ; then Z' is a subspace of Z of codimension 1, and the function $y \mapsto e(\xi \cdot y + \theta) + 1$ is constant on every coset of Z'. Making the change of variables $y = x_0 + x$ for each $x \in Z'$, and then averaging over x, we conclude

$$\begin{aligned}\mathrm{Re}\, \mathbf{E}_{y \in Z} f(y)(e(\xi \cdot y + \theta) + 1) &= \mathbf{E}_{x \in Z'} \mathbf{E}_{x_0 \in Z} f(x_0 + x) \mathrm{Re}(e(\xi \cdot x + \theta) + 1) \\ &= \mathbf{E}_{x_0 \in Z} (\mathbf{E}_{x \in x_0 + Z'} f(x)) \mathrm{Re}(e(\xi \cdot x_0 + \theta) + 1).\end{aligned}$$

By the pigeonhole principle there must therefore exist a coset $x_0 + Z'$ such that

$$(\mathbf{E}_{x \in x_0 + Z'} f(x)) \mathrm{Re}(e(\xi \cdot x_0 + \theta) + 1) \geq \|f\|_{u^2(Z)}.$$

Since $\mathrm{Re}(e(\xi \cdot x_0 + \theta) + 1) \leq 2$, the claim follows. □

Remark 10.16 The reason to add 1 to $e(\xi \cdot y + \theta)$ is to make sure that $\mathrm{Re}(e(\xi \cdot y + \theta) + 1)$ is non-negative. We will use this trick repeatedly in this chapter.

We can now prove Proposition 10.12, by using the density increment argument of Roth.

Proof of Theorem 10.12 By Corollary 3.8 we may take $Z = F_p^n$, with the standard bilinear form in Example 4.2. We induce on n. The claim is trivial when $n \leq 3$, so suppose $n > 3$. Suppose for contradiction that $r_3(F_p^n) \geq 3/n$, then we can find a set $A \subset Z$ with density $\mathbf{P}_Z(A) \geq 3/n$ containing no proper progressions of length 3. Then by Lemma 10.15 (applied to $f := 1_A - \mathbf{P}_Z(A)$) we have a coset $x_0 + Z'$ of Z of codimension one such that

$$\mathbf{P}_{x_0 + Z'}(A) \geq \mathbf{P}_Z(A) + \frac{1}{2}\|A\|_u.$$

Applying Corollary 10.10 we conclude

$$\begin{aligned}\mathbf{P}_{x_0+Z'}(A) &\geq \frac{3}{n} + \frac{1}{2}\frac{9}{n^2} - \frac{1}{2|Z|} \\ &\geq \frac{3}{n} + \frac{4}{n^2} \\ &\geq \frac{3}{n-1}\end{aligned}$$

since $|Z| = p^n \geq n^2$ and $n \geq 3$. By the induction hypothesis, the set $(A - x_0) \cap Z'$ thus contains a proper arithmetic progression of length 3, and hence A does also, which gives the desired contradiction. □

A very similar argument also establishes Varnavides' theorem in this setting:

Proposition 10.17 (Varnavides's theorem for p-torsion groups) *Let Z be a p-torsion group for some odd prime p, and let $f : Z \to \mathbf{R}^+$ be such that $0 \leq f(x) \leq 1$ for all $x \in Z$. Then*

$$\Lambda_3(f, f, f) \geq p^{-6/\mathbf{E}_Z(f)}.$$

Proof We induce on $n := \lfloor 3/\mathbf{E}_Z(f) \rfloor$. When $n \leq 3$ the claim is trivial, so suppose $n > 3$ and the claim has already been proven for $n - 1$. We may again view Z as a vector space over F_p, with a standard bilinear form. Write $f = f_{U^\perp} + f_U$, where $f_{U^\perp} := \mathbf{E}_Z(f)$ and $f_U := f - f_{U^\perp}$. Observe that

$$\Lambda_3(f_{U^\perp}, f_{U^\perp}, f_{U^\perp}) = \mathbf{E}_Z(f)^3.$$

If we had

$$\Lambda_3(f, f, f) \geq \mathbf{E}_Z(f)^3/9$$

(say) then we would be done (since $\mathbf{E}_Z(f)^3/9 \geq p^{-6/\mathbf{E}_Z(f)}$), so let us assume instead that

$$|\Lambda_3(f, f, f) - \Lambda_3(f_{U^\perp}, f_{U^\perp}, f_{U^\perp})| \geq 8\mathbf{E}_Z(f)^3/9.$$

We can rewrite the left-hand side as the telescoping sum of three terms,

$$|\Lambda_3(f_U, f, f) + \Lambda_3(f_{U^\perp}, f_U, f) + \Lambda_3(f_{U^\perp}, f_{U^\perp}, f_U)|.$$

From their definitions, we see that f_U has mean zero, and $f_{U^\perp}$ is constant. Thus one can easily verify that the latter two terms vanish. Hence

$$|\Lambda_3(f_U, f, f)| \geq 8\mathbf{E}_Z(f)^3/9.$$

Since f is bounded by 1, we have

$$\|f\|_{L^2(Z)}^2 = \mathbf{E}_Z(f^2) \leq \mathbf{E}_Z(f)$$

and hence by Proposition 10.11 we have

$$\|f_U\|_{u^2(Z)} \geq 4\mathbf{E}_Z(f)^2/9.$$

Applying Lemma 10.15, we can find a subspace Z' of Z of codimension 1, such that

$$\mathbf{E}_{x\in x_0+Z'} f(x) \geq \mathbf{E}_Z(f) + 4\mathbf{E}_Z(f)^2/9.$$

If we let $g : Z' \to \mathbf{R}$ be the function $g(x) := f(x + x_0)$, then g ranges between 0 and 1 and we have

$$\mathbf{E}_Z(g) \geq \mathbf{E}_Z(f) + 4\mathbf{E}_Z(f)^2/9;$$

this in particular forces $\mathbf{E}_Z(f) \leq 3/4$, and then from elementary algebra one concludes

$$\frac{6}{\mathbf{E}_Z(g)} \leq \frac{6}{\mathbf{E}_Z(f)} - 2.$$

By the induction hypothesis we then have

$$\Lambda_3(g, g, g) \geq p^2 p^{-6/\mathbf{E}_Z(f)},$$

while from definition of g and positivity of f we have $\Lambda_3(f, f, f) \geq p^{-2}\Lambda_3(g, g, g)$. This completes the induction. □

A remarkable phenomenon is that lower bounds of the above type still persist when the boundedness condition $f \leq 1$ is replaced by a more general condition $f \leq \nu$, providing that the enveloping weight ν is sufficiently *pseudo-random*. This phenomenon (essentially first observed in [212], [147]) was made more explicit in [158], when a *transference principle* was formulated. This principle was aimed at studying progressions of arbitrary length k and was phrased in an ergodic theory language, but a parallel Fourier-analytic principle in $k = 3$ exists, and was developed in [159]. We give a simplified formulation of this result below, in the special contexts of random subsets of p-torsion groups. Specifically, we shall prove

Theorem 10.18 (Roth's theorem in random subsets of torsion groups) *Let Z be a finite p-torsion group for some odd prime p, let $|Z|^{-0.01} \leq \tau \leq 1$, and let B be a random subset of Z with the events $x \in B$ being independent with probability $\mathbf{P}(x \in B) = \tau$. Then with probability $1 - o_{|Z|\to\infty;p}(1)$ we have $r_3(B) = o_{|Z|\to\infty;p}(|B|)$.*

Remark 10.19 The point of this theorem is that it allows us to detect arithmetic progressions in subsets of Z of density as low as $|Z|^{-0.01}$, which is well beyond the reach of Proposition 10.12, provided that those sets have large *relative* density compared to a random set. A modification of the proof given below can be

used to establish that any subset of the primes of positive relative density contains infinitely many arithmetic progressions of length 3; see [147], [159]; the point was that the primes were contained in a set of "almost primes" which was very uniform (or "pseudo-random") and thus behaved very much like a random set in a certain Fourier-analytic sense. By replacing the Fourier-analytic methods with ergodic theory methods (and replacing linear uniformity with the notion of Gowers uniformity, which could be obtained for the almost primes by some number-theoretic arguments of Goldston and Yildirim), this result was then extended to cover arithmetic progressions of arbitrary length; see [158]. Note that the original proof in [212] relied on the Szemerédi regularity lemma (Lemma 10.42 below) instead of Fourier-analytic methods (and has weaker bounds as a consequence); on the other hand, it works for an arbitrary finite additive group Z of odd order, and allows the density τ to approach $|Z|^{-1/2}$, which is the optimal value (Exercise 10.2.3).

We now begin the proof of Theorem 10.18. We shall need the following extension of Proposition 10.17, in which f is not bounded by 1, but is instead bounded by a "pseudo-random measure", and also enjoys some Fourier bounds.

Theorem 10.20 *[159] Let Z be a finite p-torsion group for some odd prime p, and let $f : Z \to \mathbf{R}_{\geq 0}$ be a non-negative function such that*

$$\|\hat{f}\|_{l^q(Z)} \leq M \tag{10.7}$$

for some $2 < q < 3$ and $0 < M < \infty$. Suppose also that we have the bound $f \leq \nu$ where $\nu : Z \to \mathbf{R}_{\geq 0}$ obeys the pseudo-randomness condition

$$|\hat{\nu}(\xi) - \mathbf{I}(\xi = 0)| \leq \eta \tag{10.8}$$

for some $0 < \eta < 1$. Then we have

$$\Lambda_3(f, f, f) \geq 8p^{-12/\mathbf{E}_Z(f)} - 7M^3 \log_p^{1-3/q} \frac{1}{\eta}.$$

Note that Proposition 10.17 corresponds to the case $\nu = 1$, in which case we can take $\eta = 0$ (and q, M are irrelevant). More generally, this theorem is useful when η is very small compared to δ and M. The constants can be improved somewhat but this will not concern us here.

Proof We may assume Z is a vector space over F_p, with a bilinear form as in Example 4.2. Let $\alpha := M/\log_p^{1/q} \frac{1}{\eta}$. We recall the spectrum $\mathrm{Spec}_\alpha(f) \subseteq Z$, defined as

$$\mathrm{Spec}_\alpha(f) := \{\xi \in Z : |\hat{f}(\xi)| \geq \alpha\}.$$

From the hypothesis (10.7) and Chebyshev's inequality we have

$$|\mathrm{Spec}_\alpha(f)| \leq M^q/\alpha^q = \log_p \frac{1}{\eta}. \tag{10.9}$$

Thus if we let $V = \mathrm{Spec}_\alpha(f)^\perp$ be the orthogonal complement to $\mathrm{Spec}_\alpha(f)$, then V is a subspace of Z and[1]

$$|V^\perp| \leq p^{|\mathrm{Spec}_\alpha(f)|} \leq \frac{1}{\eta}. \tag{10.10}$$

We split $f = f_U + f_{U^\perp}$, where $f_U := f - f * \frac{1_V}{\mathbf{P}_Z(V)}$ is the "uniform" component of f and $f_{U^\perp} := f * \frac{1_V}{\mathbf{P}_Z(V)}$ is the "anti-uniform" component. This allows us to split $\Lambda_3(f, f, f)$ into eight terms,

$$\Lambda_3(f, f, f) = \Lambda_3(f_U, f_U, f_U) + \cdots + \Lambda_3(f_{U^\perp}, f_{U^\perp}, f_U) + \Lambda_3(f_{U^\perp}, f_{U^\perp}, f_{U^\perp})$$

The idea is to use Proposition 10.17 to obtain lower bounds on the last term, and (10.6) to obtain magnitude bounds on the remaining seven terms.

We begin by controlling $f_{U^\perp}$. Since f is bounded pointwise by ν, we can use the Poisson summation formula (Exercise 4.1.7) and (10.10), (10.8) to obtain

$$\begin{aligned} f * \frac{1_V}{\mathbf{P}_Z(V)}(x) &= \nu * \frac{1_V}{\mathbf{P}_Z(V)}(x) \\ &= \sum_{\xi \in V^\perp} \hat{\nu}(\xi) e(\xi \cdot x) \\ &\leq 1 + |V^\perp| \sup_{\xi \in V^\perp \backslash 0} |\hat{\nu}(\xi)| \\ &\leq 1 + \frac{1}{\eta}\eta = 2. \end{aligned}$$

We thus see that $f_{U^\perp}$ is bounded above by 2. Also it is non-negative and $\mathbf{E}_Z(f_{U^\perp}) = \mathbf{E}_Z(f)$ thanks to (4.10). Thus by Proposition (10.17) (applied to $f_{U^\perp}/2$) we have

$$\Lambda_3(f_{U^\perp}, f_{U^\perp}, f_{U^\perp}) \geq 8p^{-12/\mathbf{E}_Z(f)}.$$

Now we consider the other terms. From the Poisson summation formula again we have

$$\hat{f}_{U^\perp} = \hat{f} 1_{V^\perp} \text{ and } \hat{f}_U = \hat{f}(1 - 1_{V^\perp}).$$

In particular we have

$$\|\hat{f}_U\|_{l^q(Z)}, \|\hat{f}_{U^\perp}\|_{l^q(Z)} \leq M.$$

Furthermore, since $V^\perp$ contains $\mathrm{Spec}_\alpha(f)$, we see that

$$\sup_{\xi \in Z} |\hat{f}_U(\xi)| \leq \alpha.$$

[1] This is extremely crude. It is likely that one can use the machinery of dissociated sets as in Lemma 4.36 to do better here.

Applying (10.6) and Hölder's inequality we obtain

$$|\Lambda_3(f_U, f_{U^\perp}, f_{U^\perp})| \leq M^q \alpha^{3-q} = M^3 \log_p^{1-3/q} \frac{1}{\eta}$$

and similarly for the other six $\Lambda_3()$ expressions to be estimated. The claim follows. □

Remark 10.21 The strategy of the above transference argument was to identify a fairly coarse partition of Z (in this case, into cosets of V) to average against in order to produce a well-behaved approximant $f_{U^\perp}$ to f, with the error f_U between f and $f_{U^\perp}$ being so uniform (in the Fourier sense) as to be negligible. This philosophy was developed in a quantitative manner in [150], in which an arithmetic version of the Szemerédi regularity lemma was obtained.

The hypothesis (10.7) in this Corollary may seem to be restrictive, but in many cases one can control the l^q norm of $\hat{f}$, or at least the spectrum $\mathrm{Spec}_\alpha(f)$ of f, by exploiting the pseudo-randomness properties of ν. For instance, one has

Lemma 10.22 (Tomas–Stein argument) *Let Z be a finite additive group, and let $\nu : Z \to \mathbf{R}^+$ and $f : Z \to \mathbf{C}$ be such that (10.8) holds for some η, and such that $|f(x)| \leq \nu(x)$ for all $x \in Z$. For any $\alpha > 0$ let $\mathrm{Spec}_\alpha(f) := \{\xi \in Z : |\hat{f}(\xi)| \geq \alpha\}$. Then we have*

$$|\mathrm{Spec}_\alpha(f)| \leq 4/\alpha^2$$

for all $\alpha \geq 2\eta^{1/2}$.

Remark 10.23 This estimate should be compared with (4.37); the point is that no L^2 bound on f is assumed, otherwise this type of estimate would follow from Plancherel's theorem. The orthogonality argument used here plays a fundamental role in the restriction theory of the Fourier transform, see for instance [356] for a survey. It is also closely related to the *large sieve inequality* in analytic number theory.

Proof For each $\xi \in \mathrm{Spec}_\alpha(f)$ let $c(\xi) := \mathrm{sgn}(\hat{f}(\xi))$. Then we have

$$\left| \sum_{\xi \in \mathrm{Spec}_\alpha(f)} \hat{f}(\xi)\overline{c(\xi)} \right| = \sum_{\xi \in \mathrm{Spec}_\alpha(f)} |\hat{f}(\xi)| \geq \alpha |\mathrm{Spec}_\alpha(f)|.$$

But the left-hand side can be rewritten as

$$\mathbf{E}_Z \left(f \overline{\sum_{\xi \in \mathrm{Spec}_\alpha(f)} c(\xi) e_\xi} \right).$$

Since $f \leq \nu$, we may use Cauchy–Schwarz and conclude that

$$\alpha|\mathrm{Spec}_\alpha(f)| \leq \mathbf{E}_Z(\nu)^{1/2}\mathbf{E}_Z\left(\nu\left|\sum_{\xi\in\mathrm{Spec}_\alpha(f)} c(\xi)e_\xi\right|^2\right)^{1/2}.$$

Since $\mathbf{E}_Z(\nu) = \hat{\nu}(0) \leq 1 + \eta \leq 2$, we thus conclude that

$$\mathbf{E}_Z\left(\nu\left|\sum_{\xi\in\mathrm{Spec}_\alpha(f)} c(\xi)e_\xi\right|^2\right) \geq \frac{1}{2}\alpha^2|\mathrm{Spec}_\alpha(f)|^2.$$

We can expand the left-hand side as

$$\sum_{\xi,\xi'\in\mathrm{Spec}_\alpha(f)} c(\xi)\overline{c(\xi')}\mathbf{E}_Z(\nu e_\xi\overline{e_{\xi'}}) = \sum_{\xi,\xi'\in\mathrm{Spec}_\alpha(f)} c(\xi)\overline{c(\xi')}\hat{\nu}(\xi-\xi').$$

But since $|c(\xi)| = 1$ and $|\hat{\nu}(\xi-\xi')| \leq \eta + \mathbf{I}(\xi-\xi'=0)$, we conclude that

$$\begin{aligned}\frac{1}{2}\alpha^2|\mathrm{Spec}_\alpha(f)|^2 &\leq \sum_{\xi,\xi'\in\mathrm{Spec}_\alpha(f)} \eta + \mathbf{I}(\xi-\xi'=0)\\ &\leq \eta|\mathrm{Spec}_\alpha(f)|^2 + |\mathrm{Spec}_\alpha(f)|.\end{aligned}$$

Since $\alpha \geq 2\eta^{1/2}$, we have $\eta|\mathrm{Spec}_\alpha(f)|^2 \leq \frac{1}{4}\alpha^2|\mathrm{Spec}_\alpha(f)|^2$, and the claim follows. □

We can now prove Theorem 10.18.

Proof of Theorem 10.18 We may assume that $|Z|$ is sufficiently large depending on δ, p since the claim is vacuous otherwise. We shall abbreviate $o_{|Z|\to\infty;p}(1)$ simply as $o(1)$. From Corollary 1.9 we have $\mathbf{P}_Z(B) = \tau + O(|Z|^{-1/5})$ (say) with probability $1 - o(1)$; in particular B is non-empty. Also, if we set $\nu := 1_B/\tau$, then by Lemma 4.16 (with A replaced by Z) we have

$$\sup_{\xi\in Z\setminus 0} |\hat{\nu}(\xi)| = O\big(|Z|^{-1/5}\big)$$

again with probability $1 - o(1)$. Combining this with our density bound on $\mathbf{P}_Z(B)$, we thus have

$$\sup_{\xi\in Z} |\hat{\nu}(\xi) - \mathbf{I}(\xi=0)| = O\big(|Z|^{-1/5}\big) \tag{10.11}$$

with probability $1 - o(1)$. Henceforth we shall condition on these events.

Let $\delta = \delta(|Z|, p) < 1$ be a small quantity decaying to zero very slowly as $|Z| \to \infty$ (i.e. $\delta = o(1)$); it will suffice to show that for δ sufficiently slowly decaying, and conditioning on the previous events, every subset A of B with relative density $|A|/|B| \geq \delta$ will contain a proper arithmetic progression of length 3.