1. Roth's theorem on progressions of length 3

In this chapter our aim is to prove the following theorem of Roth from 1953.

Theorem 1 (Roth's theorem)**.** *There is an absolute constant C such that any subset $A \subseteq \{1, \dots, N\}$ with cardinality at least $CN/(\log\log N)^{1/5}$ contains a nontrivial three-term arithmetic progression (that is to say, a triple $x, x+d, x+2d$ with $d \neq 0$).*

Note, in particular, that $1/(\log\log N)^{1/5}$ is eventually smaller than any fixed positive constant. Thus, for example, if N is large enough then every subset of $\{1, \dots, N\}$ of cardinality at least $N/100$ contains a three-term progression. Roth actually proved his theorem with $\log\log N$ in place of $(\log\log N)^{1/5}$, a stronger result. We will present the weaker bound stated because it ties in with generalisations to longer progressions in a more obvious fashion. On the example sheet you will find some pointers to a proof of Roth's original bound.

1. The density increment strategy

Roth's theorem proceeds via the so-called *density increment strategy*, and the key proposition which drives this is the following.

Proposition 1. *Suppose that $0 < \alpha < 1$ and that $N > C\alpha^{-C}$. Suppose that $P \subseteq \mathbb{Z}$ is an arithmetic progression of length N and that $A \subseteq P$ is a set with cardinality at αN. Then one of the following two alternatives holds:*

(i) *A contains at least $\frac{1}{10}\alpha^3 N^2$ nontrivial three-term progressions and in particular at least one;*
(ii) *There is an arithmetic progression P' of length $N \geqslant N^{1/3}$ such that, writing $A' := A \cap P'$ and $\alpha' := |A'|/|P'|$, we have $\alpha' > \alpha + C\alpha^6$.*

Theorem 1 follows by iterating this proposition. Set $P_0 := \{1, \dots, N\}$ and let us suppose that we have a set $A \subseteq P_0$ with $|A| = \alpha N$ and containing no nontrivial 3-term progression. Then we attempt to use Proposition 1 repeatedly to obtain a sequence $P_0, P_1, P_2, \dots$ of progressions together with sets $A_i := A \cap P_i$. The length of P_i will be $N_i \geqslant N^{(1/3)^i}$ and the densities $\alpha_i := |A_i|/|P_i|$ will satisfy $\alpha_{i+1} > \alpha_i + C\alpha_i^6$.

Now this iteration cannot last too long: after C/α^5 steps the density has already doubled, after a further $C/32\alpha^5$ steps it has doubled again, and so on. Since no set can have density greater than one, there can be more more than C'/α^5 steps in total. We conclude that our applications of Proposition 1 must have been invalid, which can only mean that the condition $N_i > C\alpha_i^{-C}$ was violated. Since

$$N_i > N^{(1/3)^i} \geqslant N^{(1/3)^{C/\alpha^5}}$$

and (very crudely)

$$\alpha_i \geqslant \alpha,$$

we infer the bound

$$N^{(1/3)^{C/\delta^5}} \leqslant C\alpha^{-C}.$$

Rearranging gives

$$\log\log N \leqslant \log\log(C\alpha^{-C}) + \frac{C}{\alpha^5} \leqslant \frac{C'}{\alpha^5},$$

which immediately gives the claimed bound. □

Remark. The most important parameter by far is the number of times we performed the iteration, which was roughly C/α^5.

2. A dichotomy involving the Gowers U^2-norm

It remains, of course, to establish Proposition 1. Suppose that P is a progression of length N and that $A \subseteq P$ is a set of size αN. By linearly rescaling (which does not affect either the density α or the number of 3-term progressions in A) we may assume that $P = \{1, \dots, N\}$. We shall now employ a technical device which is convenient but, in my opinion, a little ugly and unnatural: set $N' := 2N + 1$, write $G = \mathbb{Z}/N'\mathbb{Z}$ and regard A as a subset of (half of) G in the "obvious" way. Write $\tilde{A}$ for this set, just for now: we shall shortly drop the tilde. The key point to note is that the number of three-term progressions in $\tilde{A}$ is precisely the same as that in A, because the size of N' ensures that there are no wraparound issues.

Let us, then, identify $\tilde{A}$ with A. In order to count 3-term progressions, we introduce the 3-term progressions operator AP_3. Given functions $f_1, f_2, f_3 : \mathbb{Z}/N'\mathbb{Z} \to \mathbb{R}$, set

$$\mathrm{AP}_3(f_1, f_2, f_3) := \mathbb{E}_{x,d \in \mathbb{Z}/N'\mathbb{Z}} f_1(x) f_2(x+d) f_3(x+2d).$$

The quantity

$$\mathrm{AP}_3(1_A, 1_A, 1_A)$$

counts progressions in A; in fact it is $1/N'^2$ times the number of such progressions, including the trivial ones x, x, x. We shall compare this with

$$\mathrm{AP}_3(\alpha 1_{[N]}, \alpha 1_{[N]}, \alpha 1_{[N]}).$$

The function $\alpha 1_{[N]}(x)$ featuring here is defined to be α if $1 \leqslant x \leqslant N$ and zero if $x \in \mathbb{Z}/N'\mathbb{Z} \setminus \{1, \dots, N\}$.

To compute the difference between the two we introduce the *balanced function* $f := 1_A - \alpha 1_{[N]}$. Since AP_3 is multilinear, we may expand $\mathrm{AP}_3(1_A, 1_A, 1_A)$ as a "main term" $\mathrm{AP}_3(\alpha 1_{[N]}, \alpha 1_{[N]}, \alpha 1_{[N]}) = \alpha^3 \mathrm{AP}_3(1_{[N]}, 1_{[N]}, 1_{[N]})$ plus seven other terms $\mathrm{AP}_3(g_1, g_2, g_3)$ with at least one of the g_i being equal to f.

Lemma 1. *Suppose that $N > C\alpha^{-C}$ and that A contains fewer than $\frac{1}{10}\alpha^3 N^2$ nontrivial 3-term progressions. Then there are 1-bounded functions g_1, g_2, g_3, at least one of which is equal to the balanced function f, such that $|T(g_1, g_2, g_3)| \geqslant c\alpha^3$.*

Proof. We proceed as suggested in the preceding paragraph. The "main term" $\alpha^3 \mathrm{AP}_3(1_{[N]}, 1_{[N]}, 1_{[N]})$ can be calculated quite explicitly (see the example sheet)

but it is fairly clearly at least $\frac{1}{9}\alpha^3(N/N')^2$ since whenever one chooses $x, d \leqslant N/3$ the progression $x, x+d, x+2d$ is wholly contained within $\{1, \dots, N\}$, giving at least $N^2/9$ progressions in total. Taking account of the αN trivial progressions x, x, x we have, on the other hand, the bound

$$\mathrm{AP}_3(1_A, 1_A, 1_A) \leqslant \frac{1}{10}\alpha^3(\frac{N}{N'})^2 + \frac{\alpha N}{N'^2}.$$

If N satisfies the condition $N > C\alpha^{-C}$ for appropriately large C then the second term is negligible and this is at most $\frac{2}{19}\alpha^3(N/N')^2$. Comparing these two pieces of information, it follows that the sum of the seven terms involving f must have magnitude at least $(\frac{1}{9} - \frac{2}{19})\alpha^3(N/N')^2$. Since $N' \leqslant 3N$, one of these terms must be larger than $c\alpha^3$, where $c > 0$ is some absolute constant, as claimed. □

The balanced function f has average value 0 by construction, and so it is surprising that $\mathrm{AP}_3(f, *, *)$ (say) is large. To handle this information we introduce the Gowers U^2-norm.

Definition 1 (Gowers U^2-norm)**.** Suppose that $f : \mathbb{Z}/N'\mathbb{Z} \to \mathbb{C}$ is a function. Then we define

$$\|f\|_{U^2} := \big(\mathbb{E}_{x,h_1,h_2 \in G} f(x)\overline{f(x+h_1)f(x+h_2)}f(x+h_1+h_2)\big)^{1/4}.$$

Remarks. At this point it is not particularly important to know that $\|f\|_{U^2}$ is indeed a norm: we have indicated the proof on the example sheet. It is not, in fact, immediately obvious that the quantity whose fourth root is to be taken is real and non-negative, but this will follow from the proof of the next lemma. The U^2-norm belongs to a whole family of norms, the U^k-norms, and we will study them later. For now let us remark that the U^k-norm plays much the same role in the theory of $(k-1)$-term progressions as the U^2-norm does here.

The relationship between the progression operator AP_3 and the Gowers U^2-norm is provided by Lemma 3 below, known as a "generalized von Neumann theorem". The proof of it involves two applications of the Cauchy-Schwarz inequality in the following guise.

Lemma 2 (Cauchy-Schwarz inequality)**.** *Let X, Y be any finite sets. Suppose that $b : X \to \mathbb{C}$ is a 1-bounded function, and let $F : X \times Y \to \mathbb{C}$ be any function. Then*

$$|\mathbb{E}_{x\in X, y\in Y} b(x)F(x,y)|^2 \leqslant \mathbb{E}_{x\in X}\mathbb{E}_{y,y'\in Y} F(x,y)\overline{F(x,y')}.$$

Proof. The usual Cauchy-Schwarz inequality may be rephrased using expectation notation in the form

$$|\mathbb{E}_{x\in X}\alpha(x)\beta(x)|^2 \leqslant \big(\mathbb{E}_{x\in X}|\alpha(x)|^2\big)\big(\mathbb{E}_{x\in X}|\beta(x)|^2\big),$$

this being valid for any functions $\alpha, \beta : X \to \mathbb{C}$. To obtain the lemma, apply this with $\alpha(x) = b(x)$ and $\beta(x) = \mathbb{E}_{y\in Y}F(x,y)$. □

Lemma 3 (Generalized von Neumann theorem)**.** *Suppose that f_1, f_2, f_3 are 1-bounded complex-valued functions. Then*

$$|\,\mathrm{AP}_3(f_1, f_2, f_3)| \leqslant \|f_i\|_{U^2}$$

for $i = 1, 2, 3$.

Proof. We indicate the proof for $i = 1$; the other proofs are extremely similar. All expectations are over G. We begin by observing that

$$\mathrm{AP}_3(f_1, f_2, f_3) = \mathbb{E}_{x,y} f_1(2x - y) f_2(x) f_3(y).$$

This is because the triple $2x - y, x, y$ ranges over all 3-term progressions as x, y range over G: here we have used the fact that $2 \nmid N'$.

Applying the Cauchy-Schwarz inequality together with the bound $|f_2(x)| \leqslant 1$ we obtain

$$|\mathrm{AP}_3(f_1, f_2, f_3)|^2 \leqslant \mathbb{E}_x \mathbb{E}_{y,y'} f_1(2x - y)\overline{f_1(2x - y')} f_3(y)\overline{f_3(y')}.$$

Applying the same inequality once more, this time together with the bound $|f_3(y)\overline{f_3(y')}| \leqslant 1$, we obtain

$$|\mathrm{AP}_3(f_1, f_2, f_3)|^4 \leqslant \mathbb{E}_{x,x'} \mathbb{E}_{y,y'} f_1(2x - y)\overline{f_1(2x - y') f_1(2x' - y)} f_1(2x' - y').$$

The right-hand side here, however, is simply a rewriting of $\|f_1\|_{U^2}^4$. □

Remark. Note, incidentally, that we have confirmed the positivity of the quantity whose fourth root was required to be taken in the definition of the U^2-norm.

It is a very simple matter to combine Lemma 1 with Lemma 3 to obtain the following corollary, which is a significant staging post *en route* to Proposition 1.

Corollary 1 (Gowers norm dichotomy)**.** *Let α, $0 < \alpha < 1$, be a real number. Suppose that $N > C\alpha^{-C}$ and that A is a subset of $\{1, \dots, N\}$ with cardinality αN and fewer than $\frac{1}{10}\alpha^3 N^2$ nontrivial 3-term progressions. Let $f : G \to \mathbb{R}$ be the balanced function of A. Then $\|f\|_{U^2} \geqslant c\alpha^3$.*

3. Inverse results for the Gowers U^2-norm

To conclude the proof of Proposition 1 and hence of Roth's theorem itself, we must study 1-bounded functions f for which $\|f\|_{U^2} \geqslant \delta$. The tool used to do this is the (discrete) Fourier transform.

For $r \in \mathbb{Z}/N'\mathbb{Z}$ we write

$$\widehat{f}(r) := \mathbb{E}_{x \in G} f(x) e(-rx/N').$$

Recall that $e(\theta)$ is the same thing as $e^{2\pi i\theta}$. The following lemma encodes the properties of the Fourier transform that we need here. There is more to the theory – for example we are not mentioning the inversion formula at this stage. We shall develop the theory further in Chapter **??**.

Lemma 4 (Some properties of the Fourier transform)**.** *Suppose that $f, g : G \to \mathbb{C}$ are two functions.*

(i) *Write* $\|f\|_2 := \big(\mathbb{E}_{x\in G}|f(x)|^2\big)^{1/2}$ *and* $\|\widehat{f}\|_2 := \big(\sum_r |\widehat{f}(r)|^2\big)^{1/2}$. *Then* $\|f\|_2 = \|\widehat{f}\|_2$.

(ii) *Define the convolution of* f *and* g, $f * g$, *by*

$$f * g(x) := \mathbb{E}_{y\in G} f(y)g(x-y).$$

Then $(f*g)^\wedge(r) = \widehat{f}(r)\widehat{g}(r)$ *for all* r.

(iii) *We have* $\|f\|_{U^2} = \|\widehat{f}\|_4$, *where* $\|\widehat{f}\|_4 := \big(\sum_r |\widehat{f}(r)|^4\big)^{1/4}$.

Proof. All three parts of this lemma may be verified by straightforward calculation. Parts (i) and (iii) require the orthogonality relation

$$\mathbb{E}_{x\in G} e((r-s)x/N) = \delta_{r,s},$$

which may be verified using the formula for the sum of a geometric series. Though it can be proved by direct computation, the most natural way to prove (iii) is to note that, in view of (i) and (ii), it suffices to prove that $\|f\|_{U^2}^4 = \|f*f\|_2^2$. To see this, expand the right hand side as

$$\mathbb{E}_x \mathbb{E}_{y,y'} f(y)f(x-y)\overline{f(y')f(x-y')}$$

and note that the quadruple $(y, y', x-y', x-y)$ ranges uniformly over the 2-dimensional parallelepipeds in the definition of the U^2-norm as x, y, y' range over G. □

Comment for those interested. Our definition of the norms $\|\cdot\|_2$ and $\|\cdot\|_4$ looks rather haphazard, sometimes involving $\mathbb{E}$ and other times $\sum$. This reflects the fact that the function f is defined on the group $G = \mathbb{Z}/N'\mathbb{Z}$, whereas its Fourier transform $\widehat{f}$ is defined on the *dual* $\widehat{G}$, defined abstractly as the group of characters or homomorphisms from G to $\mathbb{C}^\times$. Though this dual group $\widehat{G}$ is isomorphic to G, it turns out that the normalised counting measure on G is dual to the counting measure on $\widehat{G}$. Thus integration with respect to the normalised counting measure, for which we have been using the symbol $\mathbb{E}$, is dual to $\sum$. We tend to use this notation in the subject not because we are purists, but rather to avoid the necessity to take account of normalising factors of N' throughout the proofs.

We are now in a position to prove an "inverse theorem" for the U^2-norm. In words, this states that a function with large Gowers U^2-norm correlates with a linear phase function. Here is the precise statement.

Theorem 2 (Inverse theorem for the U^2-norm)**.** *Suppose that* $f : G \to \mathbb{C}$ *is a* 1*-bounded function with* $\|f\|_{U^2} \geqslant \delta$. *Then there is some* r *such that*

$$|\mathbb{E}_{x\in G} f(x)e(-rx/N')| \geqslant \delta^2.$$

Proof. The conclusion is equivalent to the assertion that $\|\widehat{f}\|_\infty \geqslant \delta^2$. To prove this, we use Lemma 4, obtaining the following chain of equalities and inequalities:

$$\delta^4 \leqslant \|f\|_{U^2}^4 = \|\widehat{f}\|_4^4 \leqslant \|\widehat{f}\|_2^4 \leqslant \|\widehat{f}\|_2^2\|\widehat{f}\|_\infty^2 = \|f\|_2^2\|\widehat{f}\|_\infty^2 \leqslant \|\widehat{f}\|_\infty^2.$$

The middle bound in an instance of the inequality

$$\sum_{i=1}^{n} a_i^2 \leqslant \big(\max_{i=1,\dots,n} a_i\big) \sum_{i=1}^{n} a_i,$$

valid for nonnegative real numbers a_i and applied on this occasion with the a_i equal to the squared Fourier coefficients $|\widehat{f}(r)|^2$. □

By combining this with Corollary 1 we can show that if a set $A \subseteq \{1, \dots, N\}$ of cardinality αN contains fewer than $\frac{1}{10}\alpha^3 N^2$ 3-term progressions then

$$|\mathbb{E}_{x \in G} f(x) e(-rx/N')| \geqslant c\alpha^6$$

for some $r \in G$. Since the function f is supported on $\{1, \dots, N\}$ and $N' = 2N + 1$, this immediately implies the following proposition in which there is no mention of the group G, which has now served its purpose.

Proposition 2. *Suppose that α is a real number with $0 < \alpha < 1$, that $N > C\alpha^{-C}$ and that $A \subseteq \{1, \dots, N\}$ is a set with cardinality αN. Suppose that A contains fewer than $\frac{1}{10}\alpha^3 N^2$ 3-term APs and let $f = 1_A - \alpha$ be its balanced function, considered now as a function on $\{1, \dots, N\}$. Then there is some $\theta \in [0, 1]$ such that*

$$|\sum_{x \in N} f(x) e(\theta x)| \geqslant c\alpha^6 N.$$

4. An application of Dirichlet's principle of the pigeons

To complete the proof of Proposition 1, the density increment step, and hence of Roth's theorem itself, we must do the following. We are required to take the conclusion of Proposition 2, viz

$$|\sum_{x \in N} f(x) e(\theta x)| \geqslant c\alpha^6 N, \tag{1}$$

and use it to find a progression $P \subseteq \{1, \dots, N\}$ of length at least $N^{1/3}$ on which A has density at least $\alpha + c\alpha^6$, or in other words

$$\sum_{x \in P} f(x) \geqslant c\alpha^6 |P|. \tag{2}$$

Our task, then, is to remove a modulus sign and an $e(\theta x)$.

The latter task is accomplished by partitioning $\{1, \dots, N\}$ into progressions on which $e(\theta x)$ is roughly constant. Let us recall first a lemma of Dirichlet.

Lemma 5 (Dirichlet)**.** *Suppose that $\theta \in \mathbb{R}$ and that $0 < \delta < 1$. Then there is a positive integer $d \leqslant 1/\delta$ such that $\|d\theta\|_{\mathbb{R}/\mathbb{Z}} \leqslant \delta$.*

Proof. Consider the numbers $0, \theta, 2\theta, \dots, m\theta$ where $m = \lfloor 1/\delta \rfloor$. By the pigeonhole principle some pair of these, let us say $j\theta$ and $j'\theta$, have fractional parts differing by at most δ. We may then take $d = |j - j'|$. □

Lemma 6. *Let $0 < \eta < 1$ and suppose that $\theta \in \mathbb{R}$. Suppose that $N > C\eta^{-6}$ and that $\theta \in \mathbb{R}$. Then it is possible to subdivide $\{1, \dots, N\}$ into progressions P_i, $i = 1, \dots, k$, each of length at least $N^{1/3}$, such that $\sup_{x,x' \in P_i} |e(\theta x) - e(\theta x')| \leqslant \eta$ for each i.*

Proof. Take $\delta = \frac{1}{20}\eta N^{-1/3}$ in the previous lemma and select a $d \leqslant 20N^{1/3}/\eta$ for which $\|\theta d\|_{\mathbb{R}/\mathbb{Z}} \leqslant \frac{1}{20}\eta N^{-1/3}$. If P is any progression with common difference d and length at most $2N^{1/3}$ then, by the triangle inequality,

$$\sup_{x,x' \in P} |e(\theta x) - e(\theta x')| \leqslant 2N^{1/3}|e(\theta d) - 1|.$$

Using the inequality $|e(t) - 1| = 2|\sin \pi t| \leqslant 2\pi \|t\|_{\mathbb{R}/\mathbb{Z}}$ it follows that

$$\sup_{x,x' \in P} |e(\theta x) - e(\theta x')| \leqslant 4\pi N^{1/3} \|\theta d\|_{\mathbb{R}/\mathbb{Z}} \leqslant \eta.$$

If $N > C\eta^{-6}$ then d is at most $\sqrt{N}$, and it is clear (though ever-so-slightly tedious to write down properly) that $\{1, \dots, N\}$ may be partitioned into progressions P_i of common difference d and length between $N^{1/3}$ and $2N^{1/3}$. □

Recall that our task was to go between (1) and (2). For the rest of the argument c is the same absolute constant as appears in (1). Let us apply the previous lemma with $\eta = c\alpha^6/2$; we shall take P to be one of the progressions P_i. Equation (1) manifestly implies that

$$\sum_{i=1}^{k} |\sum_{x \in P_i} f(x)e(x\theta)| \geqslant c\alpha^6 \sum_{i=1}^{k} |P_i|.$$

By the triangle inequality and the bound $|f(x)| \leqslant 1$, the left-hand side is at most

$$\sum_{i=1}^{k} |\sum_{x \in P_i} f(x)| + \sum_{i=1}^{k} |P_i| \sup_{x,x' \in P_i} |e(\theta x) - e(\theta x')|,$$

and so

$$\sum_{i=1}^{k} |\sum_{x \in P_i} f(x)| \geqslant \tfrac{1}{2}c\alpha^6 \sum_{i=1}^{k} |P_i|.$$

To get rid of the modulus signs we employ a rather dirty trick, which is to note that

$$\sum_{i=1}^{k} \sum_{x \in P_i} f(x) = 0.$$

Adding this to the preceding inequality and applying the pigeonhole principle, we see that there is some i such that

$$|\sum_{x \in P_i} f(x)| + \sum_{x \in P_i} f(x) \geqslant \tfrac{1}{2}c\alpha^6 |P_i|,$$

which can only mean that

$$\sum_{x \in P_i} f(x) \geqslant \tfrac{1}{4}c\alpha^6 |P_i|.$$

This is (2), and so the proof of Proposition 1 and hence of Roth's theorem is now complete. □

5. Further remarks and reading*

One traditionally writes $r_3(N)$ for the cardinality of the largest subset of $\{1, \dots, N\}$ not containing three distinct elements in arithmetic progression. We have shown that $r_3(N) \leqslant CN(\log\log N)^{-1/5}$; a major open problem in additive combinatorics is to obtain reasonably close upper and lower bounds for $r_3(N)$. As we remarked at the start of the chapter, Roth originally proved the rather better bound $r_3(N) \leqslant CN(\log\log N)^{-1}$. This was subsequently improved to $r_3(N) \leqslant CN(\log N)^{-c}$, for a small constant $c > 0$, by Szemerédi [**?**]. Bourgain [**?**] showed that any $c < 1/2$ is acceptable. Later, in a technical *tour de force*, he improved this to obtain $r_3(N) \leqslant CN(\log N)^{-c}$ for any $c < 2/3$.

The most obvious nontrivial lower bound is that $r_3(N) \geqslant CN^{\log 2/\log 3}$, obtained by noting that the set

$$S = \{\sum_i \epsilon_i 3^i : \epsilon_i \in \{0, 1\}\}$$

consisting of numbers whose base three expansion contains only zeros and ones is free of three-term progressions. In particular, contestants at the 1983 International Mathematical Olympiad were asked whether or not there are 1983 numbers less than 100000: that the answer is yes may easily be shown using this construction.

A better construction is that of Behrend [**?**] from 1946, which has not subsequently been improved in any substantial way. Here is a sketch of the argument, which is a sketch only in that we have been rather carefree in our use of the $\approx$ notation. The reader should have little trouble in filling in the details.

The key observation is that a sphere in $\mathbb{R}^d$ does not contain any nontrivial 3-term progressions, on account of its being convex. We will choose an appropriate d later; let $L > 1$ be a further parameter to be chosen later. If a point $(x_1, \dots, x_d) \in \mathbb{R}^d$ lies in the box $[1, L]^d$ then the square of its distance from the origin is a positive integer less than or equal to dL^2. It follows by the pigeonhole principle that there is sphere containing at least L^{d-2}/d of the integer points in $[1, L]^d$. Let $S \subseteq \mathbb{Z}^d$ be the set of points on this sphere, which is of course free of three-term progressions.

To turn this into a subset of $\mathbb{Z}$, consider the map $\psi : [1, L]^d \to \mathbb{Z}$ defined by

$$\psi(x_1, \dots, x_d) = x_1 + (2L)x_2 + \cdots + (2L)^{d-1}x_d.$$

The set $A = \psi(S)$ is a subset of $\{1, \dots, N\}$, where $N = (2L)^d$, and it is also free of 3-term progressions (since there are no "carries" – we leave the details as an exercise). Now

$$\frac{|A|}{N} = \frac{1}{d2^d L^2} \approx \frac{1}{2^d L^2}.$$

The fact that $N = (2L)^d$ implies that $\log N \approx d \log L$. Therefore

$$\frac{|A|}{N} \approx 2^{-(\frac{\log N}{\log L} + \log L)},$$

at which point it becomes clear that we should choose $\log L \approx \sqrt{\log N}$ to get a bound of the shape $|A| \geqslant Ne^{-C\sqrt{\log N}}$. $\square$