

10.5.5 Let  $f : Z \rightarrow \mathbf{C}$  be normalized so that  $\|f\|_{L^2(Z)} = \|\hat{f}\|_{l^2(Z)} = 1$ . Suppose that one can cover the set  $\{T^h f : h \in Z\} \subset L^2(Z)$  by  $M$  balls of radius  $\sigma$  in the  $L^2(Z)$  metric. Show that  $f$  is  $(O_{M,\sigma}(1), 4\sigma)$ -almost periodic. (Hint: use the pigeonhole principle and the Fourier transform to establish a lower bound for

$$\mathbf{P}_{h \in Z} \left( \sum_{\xi \in Z} |e(\xi \cdot h) - 1|^2 |\hat{f}(\xi)|^2 < 4\sigma^2 \right).$$

Remove the  $K$  largest Fourier coefficients from  $f$ , for some  $K = O_{M,\sigma}(1)$  to be chosen later, and apply the previous exercise to conclude an upper bound on the  $l^2$  norm of the remaining Fourier coefficients.) This result, combined with Exercise 10.5.2, gives a way to define almost periodicity purely in terms of the precompactness of the orbit  $\{T^h f : h \in Z\}$ , without explicit mention of the Fourier transform.

## 10.6 The Szemerédi regularity lemma

In the original proof of Szemerédi's theorem (Theorem 10.1), Szemerédi introduced an important result in graph theory, the *Szemerédi regularity lemma*. This lemma has since become one of the main tools in discrete mathematics. It asserts, roughly speaking, that any dense large graph can be decomposed into a relatively small number of disjoint subgraphs, most of which behave pseudo-randomly. A more “ergodic” way of viewing the lemma is as an assertion that the indicator function of a graph can be decomposed into a “low-complexity” component and a “pseudo-random” component.

To state the lemma, we need some notation.

**Definition 10.41** ( $\epsilon$ -regularity) Let  $G(V, E)$  be a graph. If  $X, Y$  are disjoint non-empty subsets of  $V$ , we define the *edge density*  $d(X, Y)$  between  $X$  and  $Y$  to be the quantity

$$d(X, Y) := \mathbf{P}_{x \in X, y \in Y} (\{x, y\} \in E).$$

If  $\epsilon > 0$ , we say that the pair  $(X, Y)$  is  $\epsilon$ -regular if we have

$$|d(X', Y') - d(X, Y)| \leq \epsilon$$

whenever  $X' \subseteq X, Y' \subseteq Y$  are such that  $|X'| \geq \epsilon|X|$  and  $|Y'| \geq \epsilon|Y|$ .

A partition  $V = V_1 \cup V_2 \cup \dots \cup V_k$  is *near-uniform* if  $-1 \leq |V_i| - |V_j| \leq 1$ . Szemerédi's Regularity Lemma asserts that given a positive constant  $\epsilon$  and a graph  $G$ , one can find a near-uniform partition of  $V$  in not too many parts so that most of the pairs  $(V_i, V_j)$  are  $\epsilon$ -regular.

**Lemma 10.42** (Regularity Lemma) *Let  $\epsilon$  be a positive constant,  $m \geq 1$  an integer, and  $G = G(V, E)$  a graph. If  $|V|$  is sufficiently large depending on  $\epsilon$  and  $m$ , then there exists a near-uniform partition  $V = V_1 \cup \dots \cup V_k$  for some  $m \leq k \leq O_{\epsilon, m}(1)$ , such that all but at most  $\epsilon k^2$  of the pairs  $(V_i, V_j)$  are  $\epsilon$ -regular.*

**Remark 10.43** The Regularity Lemma does not assert that all pairs  $(V_i, V_j)$  are regular, only that  $(1 - \epsilon)$  of the pairs are. In fact, there are examples showing that one cannot expect regularity of all the pairs (Exercise 10.6.5).

**Remark 10.44** The theorem requires  $|V|$  to be large depending on  $\epsilon$  and  $m$ , or to put it another way, one needs  $\epsilon = o_{|V| \rightarrow \infty; m}(1)$ . The proof of the Regularity Lemma allows us to have  $\epsilon = O_m(\frac{1}{(\log_* |V|)^{1/5}})$ , where  $\log_*$  is the inverse to the tower exponential  $n \mapsto e \uparrow \uparrow n$ , defined recursively by  $e \uparrow \uparrow 1 = e$  and  $e \uparrow \uparrow (n + 1) := \exp(e \uparrow n)$ . Quite amazingly, Gowers shown that this bound is essentially tight, namely, for any sufficiently large  $|V|$ , there are graphs where one cannot find an  $\epsilon$ -regular partition with  $\epsilon$  larger than  $\epsilon = \frac{1}{(\log_* |V|)}$ .

The proof of the regularity lemma can be found in various textbooks on graph theory; in Section 11.6 we shall give a proof of this lemma using “ergodic” techniques similar to that of the previous section. See also [359] for an information-theoretic perspective on the lemma, and [239] for an analytic perspective.

The survey paper [208] contains a wide range of applications of the regularity lemma. In this section, we restrict ourself to a few applications in additive combinatorics, and in particular to Roth’s theorem.

To prove Roth’s theorem via the regularity lemma, it is convenient to first prove some graph-theoretic results. Let  $G = G(V, E)$  be a graph. A set  $\{e_1, \dots, e_k\}$  in  $E$  forms a *matching* if  $e_1, \dots, e_k$  are mutually disjoint. A matching is *induced* if the subgraph spanned by its endpoint does not contain any edge other than those already in the matching.

**Proposition 10.45** [304] *Let  $G = G(V, E)$  be a graph whose edge set is the union of  $|V|$  induced matchings. Then  $|E| = o_{|V| \rightarrow \infty}(|V|^2)$ .*

*Proof* The strategy will be to apply the regularity lemma, combined with the intuitive fact that a dense  $\epsilon$ -regular graph cannot support any large induced matchings.

Assume that the proposition failed. Then one could find an integer  $m \geq 1$  and arbitrarily large graphs  $G(V, E)$  with  $|E| \geq \frac{6}{m}|V|^2$  (say) such that each of the graphs  $G$  was the union of  $|V|$  induced matchings.

Fix one of these large graphs. Applying the regularity lemma (with  $\epsilon := 1/m$ ) we obtain a partition  $V = V_1 \cup \dots \cup V_k$  with  $m \leq k \leq O_m(1)$  with  $|V_i| = \frac{1}{k}|V| + O(1)$  for all  $i, j$ , and such that all but at most  $\frac{1}{m}k^2$  of the pairs  $(V_i, V_j)$  are  $\frac{1}{m}$ -regular.

Call an edge  $e$  of  $G$  *bad* if one of the following three events occurs:

- $e$  is contained in one of the  $V_i$ ;
- $e$  connects  $V_i$  to  $V_j$ , where  $d(V_i, V_j) \leq \frac{2}{m}$ ;
- $e$  connects  $V_i$  to  $V_j$ , where  $(V_i, V_j)$  is not  $\frac{1}{m}$ -regular.

One can easily verify that the total number of bad edges is at most

$$(1 + o_{|V| \rightarrow \infty, m}(1)) \left( k \binom{|V|/k + O(1)}{2} + \frac{2}{m} \binom{k}{2} \frac{|V|^2}{k^2} + \frac{1}{m} k^2 \frac{|V|^2}{k^2} \right) \leq \frac{3}{m} |V|^2,$$

if  $V$  is large enough depending on  $m$ . Thus if we let  $E' \subset E$  be the edges of  $E$  that are not bad, we still have  $|E'| \geq \frac{3}{m} |V|^2$ . By the pigeonhole principle, we can thus find an induced matching  $F$  of  $G$  which contains at least  $\frac{3}{m} |V|$  edges from  $E'$ .

Call a set  $V_i$  *poor* if it contains at most  $\frac{2}{m} |V_i|$  vertices from  $F$ . If we delete all the poor sets  $V_i$  (and their associated edges) from  $F$ , we will have deleted at most  $\frac{2}{m} |V|$  edges in all. Thus the remaining matching  $F'$  will still contain an edge from  $E'$ . By definition, this edge connects two distinct sets  $V_i, V_j$  which are not poor, which have edge density at least  $\frac{2}{m}$ , and is  $\frac{1}{m}$ -regular. If we let  $V_{i,F}$  and  $V_{j,F}$  be the vertices from  $F$  in  $V_i, V_j$  respectively, we thus have

$$d(V_{i,F}, V_{j,F}) \geq d(V_i, V_j) - \frac{1}{m} \geq \frac{1}{m}.$$

On the one hand, since  $F$  is an induced matching, the number of edges in  $V_{i,F}$  and  $V_{j,F}$  cannot exceed  $|V_{i,F}|$ , and so the edge density cannot exceed  $1/|V_{j,F}|$ . We conclude that

$$|V_{j,F}| \leq m.$$

On the other hand, we have  $|V_{j,F}| \geq \frac{2}{m} |V_j|$  (since  $V_j$  is not poor) and  $|V_j| = \frac{1}{k} |V| + O(1)$ . We conclude that  $|V| = O_{m,k}(1) = O_m(1)$ , contradicting the hypothesis that  $V$  could be arbitrarily large depending on  $m$ . The claim follows.  $\square$

There are several equivalent formulations of the above theorem; see the exercises. A slightly stronger version of the theorem is as follows.

**Lemma 10.46 (Triangle removal lemma)** [304] *Let  $G = G(V, E)$  be a graph which contains at most  $\delta |V|^3$  triangles. Then it is possible to remove  $o_{\delta \rightarrow 0}(|V|^2)$  edges from  $G$  to obtain a graph which is triangle-free (it contains no triangles whatsoever).*

Lemma 10.46 can be proven by the same method used to prove Proposition 10.45 and is left as an exercise. In fact one can easily use Lemma 10.46 to deduce Proposition 10.45.

Now we use Proposition 10.45 to give another proof of Roth's theorem, Theorem 10.8.

*Proof* Fix a finite additive group  $Z$  of odd order, and a subset  $A$  of  $Z$  which contains no arithmetic progressions. It suffices to show that  $|A| = o_{|Z| \rightarrow \infty}(|Z|)$ .

We define a bipartite graph  $G$  as follows. The color classes are the sets  $Z \times \{1\}$  and  $Z \times \{2\}$ . We draw an edge between  $(a + r, 1)$  and  $(a + 2r, 2)$  for every  $a \in Z$  and  $r \in A$ . For each  $a \in Z$ , the edges between  $(a + r, 1)$ ,  $(a + 2r, 2)$  for  $r \in A$  form a matching. We claim that this matching is induced. For, if there was another edge connecting  $(a + r, 1)$  with  $(a + 2s, 2)$  for some distinct  $r, s \in A$ , then by construction we would have  $2s - r \in A$ . But then  $r, s, 2s - r$  would be a proper progression of length three in  $A$ , a contradiction. Thus  $G$  is the union of  $|Z|$  induced matchings, and hence has at most  $o_{|Z| \rightarrow \infty}(|Z|^2)$  edges. Since the number of edges in  $G$  is clearly  $|A||Z|$ , the claim follows.  $\square$

In fact, the above methods yield the following stronger form of Roth's theorem.

**Proposition 10.47** [3] *Let  $Z$  be a finite additive group, and let  $A \subset Z \times Z$  be such that  $A$  contains no "right-angled triangles"  $(a, b)$ ,  $(a, b + r)$ ,  $(a + r, b)$  with  $a, b, r \in Z$  and  $r \neq 0$ . Then  $|A| = o_{|Z| \rightarrow \infty}(|Z|^2)$ .*

We leave the proof of Proposition 10.47 (and its connection to Roth's theorem) to the exercises.

It is of interest to obtain more quantitative bounds for the  $o()$  terms in the above results. By using an explicitly quantitative formulation of the regularity lemma, one can sharpen the  $o_{|V| \rightarrow \infty}(|V|^2)$  expression in Proposition 10.45 to  $O(|V|^2/(\log_* |V|)^{1/5})$ , and similarly for Lemma 10.46, Roth's theorem and Proposition 10.47. Thus the quantitative bounds achieved by this method compare poorly to that achieved by the Fourier method. (However, the graph-theoretical method is slightly easier to extend to the case of general  $k$ ; see the next chapter.) Given that the bounds of Roth's theorem are significantly better than what is achieved by the regularity lemma, one is then naturally led to ask the following question:

**Question 10.48** [139] *Prove Proposition 10.45 (or Lemma 10.46) without using the Regularity Lemma. Find a better quantitative bound.*

In the case of Proposition 10.47, there has been some recent progress on this question [381], [314]. In particular, the best known bound here is  $A = O(\frac{|Z|}{\log \log |Z|})$ , due to Shkredov [314].

## Exercises

10.6.1 [304] Show that Proposition 10.45 is equivalent to the following statement: Let  $G(V, E)$  be a graph such that each edge is contained in at most one triangle. Then  $|E| = o_{|V| \rightarrow \infty}(|V|^2)$ .

- 10.6.2 ((6, 3)-theorem) [304] Show that Proposition 10.45 is equivalent to the following statement: let  $G = G(V, E)$  be a 3-uniform hypergraph (thus each “edge” in  $E$  is a collection  $\{x, y, z\}$  of three vertices in  $V$ ) such that there is no set of six vertices in  $V$  which contain three or more edges in  $E$ . Then  $|E| = o_{|V| \rightarrow \infty}(|V|^2)$ .
- 10.6.3 [304] Show that Lemma 10.46 implies Proposition 10.45. (Hint: first reduce to the case of a bipartite graph which is the union of induced matchings. Add  $|V|$  additional vertices to the graph, one for each induced matching, and connect each new vertex to all the vertices in an induced matching. This creates a tripartite graph with rather few triangles, but which requires many edges to be removed in order to make it triangle-free.)
- 10.6.4 [304] Use the regularity lemma to prove Lemma 10.46.
- 10.6.5 [8] Let  $V_1 = \{v_1, \dots, v_n\}$ ,  $V_2 = \{w_1, \dots, w_n\}$  be disjoint collections of vertices, let  $V := V_1 \cup V_2$ , and let  $G = G(V_1, V_2, E)$  be the bipartite graph formed from all those edges  $\{v_i, w_j\}$  for which  $i \leq j$ . Use this to show that even for very simple graphs one must require an exceptional set of pairs  $(V_i, V_j)$  which is not regular.
- 10.6.6 By modifying the proof of Roth's theorem, use Proposition 10.45 to prove Proposition 10.47.
- 10.6.7 [323] Use Lemma 10.46 to prove Proposition 10.47, without going through Proposition 10.45. (Hint: consider a graph whose vertices are the vertical lines  $\{(a, b) : a = \text{const}\}$ , horizontal lines  $\{(a, b) : b = \text{const}\}$  and diagonal lines  $\{(a, b) : a + b = \text{const}\}$  in  $Z^2$ , and with two vertices connected by an edge if their associated lines have distinct orientations and intersect in a point in  $A$ .)
- 10.6.8 Show that Proposition 10.47 implies Roth's theorem. (Hint: if  $A \subseteq Z$ , consider sets of the form  $\{(a, b) \in Z \times Z : a + 2b \in A\}$ .)
- 10.6.9 [136] Let  $V_1, V_2$  be disjoint finite sets, and let  $f_1 : V_1 \rightarrow \{-1, +1\}$  and  $f_2 : V_2 \rightarrow \{-1, +1\}$  be functions. Let  $G = G(V_1, V_2, E)$  be the bipartite graph formed by creating an edge between  $x_1 \in V_1$  and  $x_2 \in V_2$  if and only if  $f_1(x_1) = f_2(x_2)$ . Let  $X_1 \subset V_1$  and  $X_2 \subset V_2$  be non-empty, and let  $0 < \varepsilon < 1$ . Show that if  $(X_1, X_2)$  is  $\varepsilon$ -regular, then

$$|\mathbf{E}_{x_1 \in X_1} f_1(x_1)|, |\mathbf{E}_{x_2 \in X_2} f_2(x_2)| \geq 1 - O(\varepsilon).$$

This shows that any partition of  $V_1$  and  $V_2$  into regular pairs will have to essentially be a refinement of the sets  $\{x_1 \in V_1 : f_1(x_1) = \pm 1\}$  and  $\{x_2 \in V_2 : f_2(x_2) = \pm 2\}$ .

- 10.6.10 [136] Let  $V$  be a large finite set. Show that there exist  $n$  functions  $f_1, \dots, f_n : V \rightarrow \{-1, +1\}$  for some  $n = \Omega(\log |V|)$  with the property