

LTCC Course: Graph Theory October 2013

§5 Ramsey Theory and Regularity

Peter Allen*

October 28, 2013

1 References

There is a wealth of material in Chapters 7 and 9 of Diestel, and a treasurehouse in Chapters IV and VII of Bollobás. Both books will also provide pointers to other sources.

2 Ramsey Theory

Ramsey theory is about results in the following style: no matter how “chaotic” the overall structure is, if we look at the right (usually small) piece of the structure, we will find a pattern.

The most standard version of this type of result is the *graph Ramsey theorem*. Here, the structure we examine is a colouring of the edges of a graph (not necessarily a proper colouring), and the pattern we are seeking is a set of k vertices such that the $\binom{k}{2}$ edges they span all have the same colour – this is a *monochromatic* copy of K_k . The theorem says that, whatever finite number c of colours we are provided with, if n is large enough, then in *every* colouring of the edges of K_n with c colours, there is a monochromatic copy of K_k .

Theorem 2.1. *If $n \geq 4^k$, then every 2-colouring of $E(K_n)$ contains a monochromatic copy of K_k .*

Proof. Let G be a two-coloured complete graph on $[n]$. We first construct a list of integers and colours a_1, \dots, a_{2k-1} and c_1, \dots, c_{2k-1} as follows. We let $a_1 = 1$ and c_1 be the majority colour on edges incident to a_1 (we break a tie arbitrarily). Now for each $2 \leq j \leq 2k-1$ sequentially, let S_j be the set of vertices joined to each a_i by colour c_i for $1 \leq i \leq j$. Let a_j be the smallest vertex of S_j , and c_j the majority colour on edges leaving a_j in S_j . It's trivial to check $|S_j| \geq |S_{j-1}|/2$ for each j , so that we can construct the list. Now there is a majority colour among c_1, \dots, c_{2k-1} which corresponds to a monochromatic clique of size at least k among the a_1, \dots, a_{2k-1} . \square

The first result we saw last time shows that we cannot improve the condition $n \geq 4^k$ to (roughly) $n \geq \sqrt{2}^k$. But this is essentially all we know: we cannot prove either that 3.99^k is sufficient or that 1.42^k is not sufficient.

Another branch of Ramsey theory deals with colourings of the natural numbers and monochromatic sets which satisfy some ‘arithmetic’ condition. The original example is:

Theorem 2.2 (Schur, 1912). *For $c \in \mathbb{N}$, there exists $n = n(c) \in \mathbb{N}$ such that, for any c -colouring $f : [n] \rightarrow [c]$ of $[n]$, there are $x, y \in [n]$ such that x, y and $x + y$ all have the same colour.*

*These notes are based on Graham Brightwell's notes for LTCC Graph Theory Courses in 2009-12.

Proof. We choose n large enough so that any c -colouring of the edges of K_n contains a monochromatic triangle. Now, given $f : [n] \rightarrow [c]$, we construct a colouring g of the edges of K_n by the rule $g(ij) = f(|i - j|)$.

By the choice of n , there is a monochromatic triangle ijk , with $i < j < k$. Now set $x = j - i$ and $y = k - j$. □

And a rather harder result is:

Theorem 2.3 (van der Waerden, 1927). *For each $c, r \in \mathbb{N}$, there exists $n = n(c, r)$ such that, for any c -colouring of $[n]$, there exists a monochromatic arithmetic progression $a, a + d, \dots, a + (r - 1)d$, where $d \geq 1$.*

This is the beginning of a rich theory. We say that the equation $z = x + y$ has a *monochromatic solution* in any c -colouring of $[n]$ for $[n]$ sufficiently large. Which equations, or systems of equations, have this property? This is the subject of Rado's *partition calculus*. See Bollobás.

In yet another direction, there are geometric Ramsey statements.

Theorem 2.4 (Erdős and Szekeres, 1935). *For $k \in \mathbb{N}$, there exists $n = n(k) \in \mathbb{N}$ such that, whenever X is a collection of n points in the plane, in general position, there is a set of k of the points that form the corners of a convex k -gon.*

Proof. (Sketch) Given a set X of points in the plane, in general position, we colour the 4-tuples of points “red” if the points form a convex 4-gon, and “blue” if not, i.e., if one of the points is inside the convex hull of the other three.

Ramsey's Theorem now says that we can either find k points such that all the 4-subsets are red – in which case we are done, as these points form a convex k -gon – or we can find k points such that all the 4-subsets are blue. But the latter is not possible, as there is no way to place even 5 points in the plane in general position without forming a convex 4-gon. □

The theorems we saw all identify some ‘small part’ of a structure which has a nice pattern. But it doesn't have to be this way:

Theorem 2.5 (Gyárfás, 1983). *For any $n \in \mathbb{N}$, if the edges of K_n are coloured with red and blue, there exists a pair of vertex-disjoint paths which cover the vertices of K_n , one using only red edges and the other only blue edges.*

Proof. We construct the two paths P_1 and P_2 as follows. We start with both paths empty. Now for each $1 \leq j \leq n$ in succession, we apply the following algorithm. If we can add j to the end of P_1 and maintain the property that P_1 uses only red edges, we do so. If not, but we can add j to P_2 , we do so. If we can do neither, then let u be an end of P_1 and v an end of P_2 . If uv is red, we remove v from P_2 and add v, j to P_1 , while if uv is blue we remove u from P_1 and add u, j to the end of P_2 . □

3 Regularity

Szemerédi's Regularity Lemma has revolutionised Graph Theory. The purpose of this section is to give a short introduction to what the Lemma says, and how it is used.

Let's start with a very loose and vague (and also false) statement of the Regularity Lemma. Basically, it says:

All graphs can be partitioned into a bounded number of vertex classes of the same size, so that the graph between any pair of classes resembles a random bipartite graph.

Let's try and make some sense of this. First of all, suppose B is a bipartite graph on the two equal vertex classes V_1 and V_2 (so every edge of B has one edge in each class). What does it mean to say that B 'resembles a random bipartite graph'?

First of all, what is a random bipartite graph? We fix some $p \in [0, 1]$, and, for each pair of vertices $u \in V_1, v \in V_2$, we put an edge between u and v with probability p , all choices made independently.

Given $\varepsilon > 0$, using Chernoff's inequality¹ we can check that it is likely that (if $|V_1| = |V_2|$ is large enough) the following holds for *every* pair of sets $A \subseteq V_1$ and $B \subseteq V_2$. The number of edges between A and B is $p|A||B| \pm \varepsilon|V_1||V_2|$.

We will say a balanced bipartite graph B with partition classes V_1 and V_2 is ε -regular with density p if it 'looks like' a random graph in this sense, i.e. if for any pair of sets $A \subseteq V_1$ and $B \subseteq V_2$, we have $p|A||B| \pm \varepsilon|V_1||V_2|$ edges between A and B . Note that the main term is only bigger than the error term if we have $|A| \geq \varepsilon|V_1|$ and $|B| \geq \varepsilon|V_2|$.

It's convenient to formulate all this in terms of *edge densities*. Given a pair of disjoint vertex sets A and B , we write $d(A, B)$ for $e(A, B)/(|A||B|)$. If V_1, V_2 are a pair of disjoint vertex sets in a graph G , and $d(V_1, V_2) = d$, the pair V_1, V_2 is ε -regular (in G) if the bipartite graph induced by V_1 and V_2 is ε -regular, which means that (for some ε' not equal to ε !) any pair of subsets A, B of V_1, V_2 each covering at least ε' -fraction of V_1, V_2 (respectively) has density $d(A, B) = d \pm \varepsilon'$.

Now, let's return to our vague statement of the regularity lemma, and consider our demand that the partition of the vertex set into classes is such that 'the graph between any pair of classes resembles a random bipartite graph', or 'every pair of classes is an ε -regular pair'. It turns out that this is too much to ask for (you'll see an example in the exercises). But we can replace 'every' with 'all but an ε -fraction of'. Note that this ε is the same ε as in ε -regular. This isn't 'necessary': we could have one 'error parameter' for ε -regular and a different one here. But we wouldn't gain anything by doing this, and it simplifies notation to have one error parameter.

We can also ask what we mean by saying that all the classes have the same size—obviously there are not many choices if for example G has a prime number of vertices. It turns out to be useful to allow ourselves to ignore a 'small' set. This leads to the following definition.

Definition 3.1. *For a graph G , an ε -regular partition of G is a partition of $V(G)$ into classes V_0, V_1, \dots, V_k , where $|V_0| < \varepsilon|V(G)|$ and $|V_1| = |V_2| = \dots = |V_k|$, such that all but at most $\varepsilon \binom{k}{2}$ pairs (V_i, V_j) are ε -regular.*

Finally, let's turn to the idea that all graphs have an ε -regular partition into a bounded number of vertex classes. It's easy to check that for any graph G , the partition of $V(G)$ into singletons is a partition in which every pair of parts is ε -regular (for any ε), but it's not useful. To exclude this (at least for large graphs), we impose an upper bound on the number of classes which *can* and *certainly does* depend on ε , but *does not* depend on the number of vertices of G . Another 'useless partition' is the trivial partition $\{V(G)\}$: to exclude this (and because having a reasonably large number of classes turns out to help quite often) we impose a lower bound ε^{-1} on the number of classes. We come to the following theorem.

Theorem 3.2 (Szemerédi's Regularity Lemma). *For every $\varepsilon > 0$, there exists $K = K(\varepsilon) \in \mathbb{N}$ such that every graph G has an ε -regular partition of $V(G)$ into some number k of classes with $\varepsilon^{-1} \leq k \leq K$.*

¹Actually this doesn't follow from the 'Chernoff's inequality' written in Lecture 4 - why? But it does follow from a more general 'Chernoff's inequality' you can find in e.g. Bollobás.

Note that this is trivial for graphs with less than K vertices, because we will get the partition into singletons. But for graphs on much more than K vertices, this is a very strong statement. When we use the Regularity Lemma, we are usually trying to prove a result ‘for graphs on n vertices, for all sufficiently large n ’, in which case we can think of each partition class as being a big set of vertices.

Also, the Regularity Lemma says nothing about what happens *inside* any one of the classes. This is one of the reasons why it’s useful to be able to set a lower bound on the number of classes, so that the total number of edges inside the classes is guaranteed to be relatively small (at most $\varepsilon \binom{n}{2}$).

We will only sketch the proof here. For details (of this approach), see Diestel or Bollobás.

Proof. (Sketch)

The basic idea of the proof is that we will describe an algorithm which starts with an initial partition into ε^{-1} classes (we only need to ask that the classes are all close to the same size) and then successively refines the partition, at each step by a bounded amount. Then we need to show that the algorithm terminates after a bounded number of steps. To do this we use a parameter called ‘mean square density’ or ‘energy’. Given a partition \mathcal{P} of the vertices of the n -vertex graph G , we define

$$\mathcal{E}(\mathcal{P}) = \sum_{X \neq Y \in \mathcal{P}} d(X, Y)^2 |X||Y|/n^2.$$

This parameter behaves a little like the second moment $\mathbb{E}A^2$ of a random variable A —it measures the extent to which the densities between pairs of parts in \mathcal{P} fluctuate. It has two properties we need.

First, if \mathcal{P}' refines \mathcal{P} , then $\mathcal{E}(\mathcal{P}') \geq \mathcal{E}(\mathcal{P})$: this is a fairly easy application of the Cauchy-Schwarz inequality.

Second, suppose V_1, V_2 is not ε -regular. Then there are sets W_1 and W_2 , of sizes at least $\varepsilon|V_1|$ and $\varepsilon|V_2|$, such that $d(W_1, W_2) \neq d(V_1, V_2) \pm \varepsilon$. Now the contribution of (V_1, V_2) to $\mathcal{E}(\mathcal{P})$ is $d(V_1, V_2)^2 |V_1||V_2|/n^2$. If we refine \mathcal{P} by splitting V_1 into W_1 and $V_1 \setminus W_1$, and V_2 similarly, we can easily check that the energy is increased by at least $\varepsilon^4 |V_1||V_2|/n^2$. This is an easy calculation.

Suppose \mathcal{P} is a partition in which all parts are (approximately) the same size, but it is not ε -regular. Then we can imagine performing one refinement as above for each pair which is not ε -regular to get \mathcal{P}' (in fact we need to be a little more careful, but this is not conceptually important). We will split each class of \mathcal{P} into at most $2^{|\mathcal{P}|-1}$ classes, and get an increase in energy of about $\varepsilon^4/(|\mathcal{P}|^2)$ for each pair we refined—of which there are at least $\varepsilon \binom{|\mathcal{P}|}{2}$ —so we will get an increase in energy of at least $\varepsilon^5/4$. Of course \mathcal{P}' may well have parts of very different sizes, but using the first property we can refine further to a partition \mathcal{P}'' in which all classes are about the same size². We have

$$\mathcal{E}(\mathcal{P}) + \varepsilon^5/4 \leq \mathcal{E}(\mathcal{P}') \leq \mathcal{E}(\mathcal{P}'')$$

where the second inequality follows from the first property above. Now either \mathcal{P}'' is the desired ε -regular partition, or we can repeat the refinement. But we cannot repeat this more than $4\varepsilon^{-5}$ times, otherwise we would reach a partition whose energy is larger than one, which is impossible as we always have $d(X, Y) \leq 1$. \square

We’ll see how and why this is a useful result in the next section. But here is a note of caution. We’d like to apply the result with a reasonably small value of ε , but then how large does $K = K(\varepsilon)$ (say) have to be? This proof of the Regularity Lemma shows that we can take K to be a tower of

²Obviously this is a lie—there could be classes with very few (say \sqrt{n}) vertices in \mathcal{P}' , which we have to get rid of and we cannot do that by refining a bounded amount. But they turn out only to be a small technical complication.

twos of height at most $4\varepsilon^{-5}$. Moreover, Gowers showed that this is really something like the truth: a tower of twos of height at least $\varepsilon^{-1/16}$ is needed! So as soon as we prove something using the Regularity Lemma, we introduce unpleasantly large constants.

There are a lot of reasonable questions one can ask about the Regularity Lemma and (this) proof. Why did we take mean *square* density? The answer is that we could actually work with any strictly convex function (replacing Cauchy-Schwarz with Jensen's inequality), but the calculations are easiest with the square. Is it important that we actually never referred to single edges but only edge densities? The answer to this is that it is definitely important: the whole proof goes through taking (measurable) partitions of $[0, 1]$ to deduce 'structure' for any (measurable) $f : [0, 1]^2 \rightarrow I$, where I can be any bounded interval. This leads to a recent topic called 'Graph Limits'. Is there any 'other' proof of the Regularity Lemma? There is: one can show that, given $\varepsilon > 0$, if (a very large but bounded number of) vertices v_1, \dots, v_K of a graph G are picked independently and uniformly at random, then the partition of $V(G)$ given by the Venn diagram of the neighbourhoods of v_1, \dots, v_K is very likely to have most of its pairs ε -regular. It won't have the property that the classes are all about the same size, but this can be 'fixed'.

4 Sample applications of the Regularity Lemma

Many many modern proofs in graph theory start with the phrase 'take an ε -regular partition of G '. The point is that, once we've taken an ε -regular partition, we know a lot about the structure of the graph already.

One important tool that goes together with the Regularity Lemma is the Counting Lemma. We'll just state this for triangles:

Lemma 4.1 (Counting Lemma: triangles). *Suppose X, Y, Z are pairwise disjoint sets of vertices in a graph G . If all three pairs are ε -regular with densities d_{XY}, d_{XZ} and d_{YZ} , then the number of triangles in G with one vertex in each set is $(d_{XY}d_{XZ}d_{YZ} \pm 8\sqrt{\varepsilon})|X||Y||Z|$.*

Proof. Consider the set B of vertices of X with fewer than $(d_{XY} - \sqrt{\varepsilon})|Y|$ neighbours in Y . The set B has size at most $\sqrt{\varepsilon}|X|$, otherwise the pair B, Y violates the ε -regularity of X, Y . Similarly, at most $\sqrt{\varepsilon}|X|$ vertices of X have 'too many' neighbours in Y , and the same holds replacing Y with Z . We conclude that at least $(1 - 4\sqrt{\varepsilon})$ vertices of X have $(d_{XY} \pm \sqrt{\varepsilon})|Y|$ neighbours in Y , and similarly in Z . Each of these vertices (by ε -regularity of YZ) lies in

$$(d_{XY} \pm \sqrt{\varepsilon})|Y| \cdot (d_{XZ} \pm \sqrt{\varepsilon})|Z| \pm \varepsilon|Y||Z|$$

triangles, while the remaining at most $4\sqrt{\varepsilon}|X|$ vertices of X lie in at most $|Y||Z|$ triangles each. \square

The full 'Counting Lemma' provides a similar result for counting copies of *any* graph. Now one way we can represent the partition of a graph G provided by the Regularity Lemma is to draw a *reduced graph* $R(G)$ whose nodes are the partition classes V_1, \dots, V_k and whose edges are given a 'weight' in $[0, 1]$ corresponding to the edge density between the pair of classes. If we define the 'triangle density' of G to be the probability that a randomly selected triple of vertices of G form a triangle, then the Counting Lemma for triangles says that we can approximate the triangle density of G just by looking at $R(G)$ (and in fact just from the 'weighted triangle density' in $R(G)$, think about how this has to be defined). And the full Counting Lemma says we can do this for density of any graph (if ε is small enough!). So $R(G)$ is a 'model' of G whose size is bounded. This turns out to be useful in many proofs.

Here is a nice application of the Regularity Lemma.

Theorem 4.2 (Thomassen, 2000; Łuczak, 2006). *For each $\eta > 0$ there exists $C = C(\eta)$ with the following property. If G is an n -vertex triangle-free graph whose minimum degree is at least $(\frac{1}{3} + \eta)n$, then G has chromatic number at most C .*

This theorem is best possible in the sense that there is a construction (due to Hajnal, 1973) of triangle-free graphs on n vertices with minimum degree $n/3 - o(n)$ and chromatic number tending to infinity as n tends to infinity.

Proof. Given $\eta > 0$, we set $d = \eta/10$ and $\varepsilon = d^3/2$. The Regularity Lemma returns a constant $K(\varepsilon)$, and we set $C(\eta) = 2^{K(\varepsilon)}$. Now let G be any triangle-free n -vertex graph with minimum degree at least $(\frac{1}{3} + \eta)n$. Let V_0, V_1, \dots, V_k be an ε -regular partition of G with $\varepsilon^{-1} \leq k \leq K(\varepsilon)$, as is guaranteed to exist by the Regularity Lemma.

Now we define a second partition of $V(G)$ as follows. For each $I \subseteq [k]$, we let

$$X_I = \left\{ v \in V(G) : \text{for } i \in [k], |N(v) \cap V_i| \geq d|V_i| \text{ if and only if } i \in I \right\}.$$

This partition has $2^k \leq 2^{K(\varepsilon)} = C$ parts, which is independent of n . We claim that all its parts are independent, i.e. it witnesses that $\chi(G) \leq C(\eta)$. We split the proof into two cases.

Case 1: $|I| \geq 2k/3$.

In this case, the set $U = \bigcup_{i \in I} V_i$ has size at least $2n/3 - \varepsilon n$. Thus every vertex of G has at least $(\eta - \varepsilon)n$ neighbours in U , and in particular the average density between pairs from $\{V_i : i \in I\}$ is at least $\eta - 2\varepsilon$ (we ‘lose’ the edges which lie within classes, but there are few such because the classes are small). Now there are three sorts of pair contributing to this average: pairs which are not ε -regular (of which there are few), pairs whose density is smaller than d , and pairs which are ε -regular and of density at least d . The choice of d and ε is such that at least one pair of the latter type occurs, say V_p, V_q . Let x be any vertex of X_I . Then by definition x has at least $d|V_p|$ neighbours in V_p and $d|V_q|$ neighbours in V_q , and we conclude by ε -regularity of V_p, V_q that x is in at least $d^3|V_p||V_q| - \varepsilon|V_p||V_q| > 0$ triangles. This is a contradiction to triangle-freeness of G : we conclude that X_I is empty (and so trivially independent).

Case 2: $|I| < 2k/3$.

In this case the set $U = \bigcup_{i \in I} V_i$ has size at most $2n/3$. Since any vertex of X_I has at most $|V_0| \leq \varepsilon n$ neighbours in V_0 , and at most $d|V_i|$ neighbours in any set V_i with $i \notin I$, we conclude that any vertex of X_I has at least $(\frac{1}{3} + \eta - d - \varepsilon)n$ neighbours in U . This is more than half of $|U|$, and we conclude that any two vertices in X_I have a common neighbour in U . But then any edge in $G[X_I]$ lies in a triangle. Since G is triangle-free, we conclude that X_I is an independent set in G as desired. \square

The type of proof we saw here is just the start. There are further tools (especially the *Blow-Up Lemma*) that enable more sophisticated applications, including the resolution of a number of what were important open problems in graph theory. Moreover, versions of the Regularity Lemma have been proved for sparse graphs, and for hypergraphs, that enable even more applications.

There are also applications outside graph theory. Recall that van der Waerden’s Theorem asserts that there exists $n = n(c, r)$ such that, in any c -colouring of $[n]$, there is a monochromatic *arithmetic progression* of length r . Erdős conjectured a stronger statement: if a set S of integers has *positive upper density* (i.e., $\limsup_{n \rightarrow \infty} |S \cap [n]|/n > 0$), then S contains arbitrarily long arithmetic progressions. This is the theorem proved by Szemerédi in 1975 for which he needed a graph-theoretic lemma ...

5 Exercises

Note: Exercises 2 and 3 are from Bollobás.

1. Fill in the (geometric) details in the proof of Theorem 2.4.
2. Let S be an infinite set of points in the plane. Show that there is an infinite subset A of S such that either no three points of A are on a line, or all points of A are on a line.
3. The Ramsey number $R_k(3)$ is the minimum number n of vertices such that, if the edges of K_n are coloured with k colours, there is always a monochromatic triangle. Show that $R_k(3) \leq k(R_{k-1}(3) - 1) + 2$. [Hint: if you don't know the classic proof that $R_2(3) \leq 6$, find and read that first.]

Deduce that $R_k(3) \leq \lfloor ek! \rfloor + 1$.

4. Let $B_{n,p}$ be a random bipartite graph, with two vertex classes V_1 and V_2 each of size n . So each pair of vertices in different classes is joined by an edge with probability p .
(a) Show that, for all $\varepsilon > 0$, $p > 0$,

$$\mathbb{P}((V_1, V_2) \text{ is an } \varepsilon\text{-regular pair in } B_{n,p}) \rightarrow 1 \text{ as } n \rightarrow \infty.$$

(You may need some estimates for tails of Binomial random variables.)

- (b) Show that, for any bipartite graph H , and any fixed $p > 0$,

$$\mathbb{P}(B_{n,p} \text{ contains a copy of } H \text{ as a subgraph}) \rightarrow 1 \text{ as } n \rightarrow \infty.$$

[Hint: a crude but straightforward approach starts by taking many disjoint subsets of the vertex set, each of size $|V(H)|$.]

5. Suppose G is a bipartite graph, with vertex classes V_1 and V_2 , each of size n . Suppose also that the maximum degree of G is at most $\varepsilon^2 n$. Show that the pair (V_1, V_2) is ε -regular.
6. Let G_n be the following bipartite graph. The vertex set of G_n is $\{x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n\}$. The edges are given by $x_i y_j \in E(G_n)$ if $i < j$.

Fix $\varepsilon > 0$. For each value of n , find an explicit ε -regular partition of G_n into at least three and at most (say) $10/\varepsilon$ parts.