LTCC Course: Graph Theory January-February 2011
§5 Ramsey Theory and Regularity

Graham Brightwell

February 14, 2011

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 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 build 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.

There is a straightforward inductive proof of Ramsey's Theorem, which at the same time provides an upper bound on how large n has to be in terms of k and r. See, for instance, Theorem 9.1.1 of Diestel, or Section VI.1 of Bollobás.

We'll follow a different route (partly because students are less likely to have seen this before, and partly because there are some instructive points to the argument).

We'll start by establishing the *infinite* version of Ramsey's Theorem, as originally proved by Ramsey in 1930.

For a set X, let $X^{(r)}$ denote the collection of r-subsets (subsets of size r) of X. So, if $r = 2$, this is the set of edges of a complete graph with vertex set X.

Theorem 2.1. *Let c and r be given positive integers, and let $f : \mathbb{N}^{(r)} \to [c]$ be a c-colouring of the r-subsets of $\mathbb{N}$. Then there exists an infinite subset M of $\mathbb{N}$ such that f is monochromatic on $M^{(r)}$, i.e., there is some $j \in [c]$ such that $f(A) = j$ for all $A \in M^{(r)}$.*

Proof. See (for instance) Diestel, Theorem 9.1.2. □

For $r = 2$, this result says that, if we have a c-colouring of the edges of an infinite complete graph, then we can find an infinite set of vertices spanning a monochromatic clique.

How do we deduce the finite version of Ramsey's Theorem from the infinite version? There are various ways, based on different forms of "compactness" argument.

Theorem 2.2. *Let k, c and r be given positive integers. Then there exists $n = n(k, c, r)$ such that, for all c-colourings $f : [n]^{(r)} \to [c]$, there exists a set $M \subseteq [n]$ of size k such that f is monochromatic on $M^{(r)}$.*

Proof. Suppose the result is false for some k, c and r. That means that, for every n, there is a c-colouring $f_n : [n]^{(r)} \to [c]$ with no monochromatic $M^{(r)}$ of size k.

We'll construct, from the f_n, a c-colouring f of the r-subsets of $\mathbb{N}$ that still has no monochromatic $M^{(r)}$ of size k: this will contradict the infinite version of Ramsey's Theorem.

We enumerate the r-subsets of $\mathbb{N}$ as $A_1, A_2, \ldots$. Now we choose a colour $f(A_1)$: we choose a colour that occurs infinitely often as $f_n(A_1)$. We discard all the values of n such that $f_n(A_1) \neq f(A_1)$, and proceed. Next we choose for $f(A_2)$ a colour that occurs infinitely often among the non-discarded values of n. Etc.

At the end, we have coloured all the r-subsets of $\mathbb{N}$. For any k-subset M, there are only finitely many r-subsets of M, so eventually we've passed them all in the enumeration, and at that point f still agrees with infinitely many of the f_n. So f is not monochromatic on $M^{(r)}$. □

This proof can be viewed as a demonstration of "sequential compactness".

Alternatively, one can argue as follows: the set of all c-colourings of $\mathbb{N}^{(r)}$ is formally $[c]^{\mathbb{N}^{(r)}}$: the set $[c]$ has the discrete topology, and $[c]^{\mathbb{N}^{(r)}}$ the product topology. By Tychonov's Theorem (The Cartesian product of any collection of compact sets, with the product topology, is compact), this space is compact. For each M of size k, the set C_M of colourings that are monochromatic on M is an open set. Ramsey's Theorem says that these sets cover the whole of $[c]^{\mathbb{N}^{(r)}}$. By compactness, some finite collection $\{C_{M_1}, \ldots, C_{M_t}\}$ of these sets also covers $[c]^{\mathbb{N}^{(r)}}$. This implies that every c-colouring of $\bigcup_{i=1}^{t} M_i$ is monochromatic on at least one of the M_i, which in turn implies the finite version of Ramsey's Theorem.

Yet another way to see this proof is as an illustration of the *compactness theorem for first-order logic*: a set $\mathcal{S}$ of first-order sentences has a model if and only if every finite subset of $\mathcal{S}$ has a model. Roughly speaking, we can write "f is a c-colouring of $\mathbb{N}^{(r)}$ that is not monochromatic on any $M^{(r)}$ with $|M| = k$" as an infinite collection of first-order sentences. The infinite version of Ramsey's theorem implies that this collection has no model; if some finite subcollection has no model, then the finite version fails.

We are starting to wander from graph theory a little. Let's wander in a different direction to give a couple of sample applications of Ramsey Theory.

Theorem 2.3 (Schur 1912)**.** *For $c \in \mathbb{N}$, there exists $n = n(c) \in N$ such that, for any c-colouring $f : [n] \to [c]$ of $[n]$, there are $x, y \in [n]$ such that x, y and $x + y$ all have the same colour.*

Proof. We choose n large enough so that any c-colouring of the edges of K_n contains a monochromatic triangle. Now, given $f : [n] \to [c]$, 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 $y = k - j$. □

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.

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. □

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 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.

One property that a random bipartite graph is likely to have is that, if $A \subseteq V_1$ and $B \subseteq V_2$ are *large* subsets, then the number $e(A, B)$ of edges between A and B is about $p|A|\,|B|$. We can write this without explicit reference to p as saying that $e(A, B) \simeq e(V_1, V_2)|A|\,|B|/|V_1|\,|V_2|$. Such a statement can't hold for "small" subsets (by Ramsey's Theorem). Also, we can't expect $e(A, B)$ to be always *exactly* what it's supposed to be. So, in order to say what we mean, we introduce an "accuracy" parameter $\varepsilon > 0$.

Definition 3.1. *Let G be a graph, and fix $\varepsilon > 0$. A pair V_1, V_2 of subsets of $V(G)$ is called an* ε-regular pair *if, for every pair $A \subseteq V_1$, $B \subseteq V_2$, with $|A| \geq \varepsilon|V_1|$, $|B| \geq \varepsilon|V_2|$, we have*

$$\left| \frac{e(A, B)}{|A|\,|B|} - \frac{e(V_1, V_2)}{|V_1|\,|V_2|} \right| < \varepsilon.$$

Notice that the "accuracy parameter" is playing two roles here. It will be burdened even further in what follows.

If a bipartite graph is "sparse", then (exercise) the *density* $e(A, B)/|A|\,|B|$ will be close to zero for all large $|A|$ and $|B|$, and the graph is automatically ε-regular. So the definition is only meaningful for graphs that aren't too sparse (and aren't too dense).

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".

This turns out to be too much to ask for; instead we ask that *most* pairs be ε-regular, where, if there are k classes, "most" means all but at most $\varepsilon\binom{k}{2}$.

Next, let's turn to the idea that all graphs have a suitable partition into a bounded number of vertex classes. What does "bounded" mean here? It's too much to ask that there is some absolute bound on the number of classes. On the other hand, to allow the bound to depend on the number of vertices of the graph is asking too little. What we do is to allow the bound on the number of vertex classes to depend on the accuracy ε we are requiring: the less tolerant we are of deviations from the ideal structure, the more classes we have to permit. So the upper bound k_1 on the number of classes is a function $k_1 = k_1(\varepsilon)$.

It also turns out to be useful to demand that the number of classes be at least some pre-specified value k_0, and then k_1 needs to depend on k_0 as well (if only to ensure that $k_1 \geq k_0$).

Finally, if $n = |V(G)|$ is a large prime, we can't partition $|V(G)|$ *exactly* into at most k_1 sets of equal size. This isn't really a crucial point, and in any case it turns out to be convenient to allow ourselves to ignore a small set V_0 of "bad" vertices. How small is small? No larger than εn.

Now we are in a position to state the Regularity Lemma.

Definition 3.2. *For a graph G, an* ε-regular partition *of G is a partition of $V(G)$ into classes $V_0, V_1, \ldots, V_k$, where $|V_0| < \varepsilon|V(G)|$ and $|V_1| = |V_2| = \cdots = |V_k|$, such that all but at most $\varepsilon\binom{k}{2}$ pairs (V_i, V_j) are ε-regular.*

Theorem 3.3 (Szemerédi's Regularity Lemma)**.** *For every $\varepsilon > 0$ and $k_0 \in \mathbb{N}$, there exists $k_1 = k_1(\varepsilon, k_0) \in \mathbb{N}$ such that, for all graphs G with $|V(G)| \geq k_0$, there is an ε-regular partition of $V(G)$ into some number k of classes with $k_0 \leq k \leq k_1$.*

Notice that the Lemma says nothing of interest about graphs with fewer than k_1 vertices: such graphs can be partitioned into singleton sets, and that partition is automatically ε-regular. One should think of the number of vertices of the graph as being much larger than k_1, so that the classes are all large. The key point is that the upper bound k_1 on the number of classes doesn't depend on the number of vertices of G.

Also, the Regularity Lemma says nothing about what happens *inside* one of the classes. This is one of the reasons why it's useful to be able to set a lower bound k_0 on the number of classes, so that the total number of edges inside the classes is guaranteed to be relatively small (at most $\frac{1}{k_0}\binom{n}{2}$).

For a proof of the Regularity Lemma, see Diestel or Bollobás. For such a powerful and subtle result, the proof is not all that hard.

We'll see how and why this is a useful result in the next section. Let's conclude this section with a note of caution. We'd like to apply the result with a reasonably small value of ε, but then how large does $k_1 = k_1(\varepsilon, 2)$ (say) have to be? The standard proof of the Regularity Lemma shows that we can take k_1 to be a tower of twos of height at most ε^{-5}. Moreover, Gowers showed that this is really the truth: a tower of twos of height at least $\varepsilon^{-1/16}$ is needed. Remember: the Regularity Lemma tells us nothing for graphs on fewer than k_1 vertices.

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.

Another key tool that typically forms part of a Regularity Lemma proof is the so-called *Embedding Lemma*. The setting here is that we have a graph G and vertex classes $V_1, \ldots, V_r$, each of size t, each pair of which is ε-regular with density not too close to zero. If the graphs across each pair were genuinely random bipartite graphs, all chosen independently of each other, then it's not hard to show that there are, with probability tending to 1 as $t \to \infty$, many copies of any given r-partite graph. In fact, the same is true if we only assume that the pairs are ε-regular.

The following version of the embedding lemma is taken from Bollobás, IV.6, Theorem 31.

Theorem 4.1 (Embedding Lemma)**.** *Fix $s \geq 2$, $r \geq 2$, $0 < \delta < 1/r$, and set $\varepsilon = \delta^s$. Let $V_1, \dots, V_r$ be disjoint subsets of vertices of a graph G, such that $|V_i| \geq \varepsilon^{-1}$ for each i, and all pairs (V_i, V_j) are ε-regular with density at least $\delta + \varepsilon$. Then G contains every r-colourable graph of order s.*

We'll show how to use this result to prove the celebrated Erdős-Stone Theorem. The result goes back to 1946, although it was first stated in this form, in terms of the chromatic number, by Erdős and Simonovits in 1966.

Theorem 4.2 (Erdős-Stone-Simonovits)**.** *Let H be a graph with chromatic number r, and let $ex(n, H)$ be the maximum number of edges of a graph G on n vertices not containing H as a subgraph. Then*

$$\mathrm{ex}(n, H) = \left(\frac{r-2}{r-1}\right)\frac{n^2}{2} + o(n^2).$$

For instance, suppose $H = K_r$, with $r \geq 3$. In the last lecture, we proved Turán's Theorem, which says that $\mathrm{ex}(n, K_r) = t_{n,r-1}$, the number of edges of the Turán graph $T_{n,r-1}$, a complete $(r-1)$-partite graph with parts as equal as possible in size. In $T_{n,r-1}$, the fraction of non-adjacent pairs is about $1/(r-1)$, which is in line with the Erdős-Stone-Simonovits Theorem.

More generally, the Turán graph $T_{n,r-1}$ doesn't contain any graph with chromatic number at least r, so the lower bound in the Erdős-Stone-Simonovits Theorem is immediate. It's too much to expect that the Turán graph will always be the "extremal" graph – maybe adding one edge to it is not enough to create a copy of H – but the theorem says (at least for $r \geq 3$) that there is no graph with *substantially* more edges than the Turán graph that contains no copy of H.

We give a proof of the Erdős-Stone-Simonovits Theorem based on the Regularity Lemma.

Proof. Let H be a graph of order s, with chromatic number r. Fix some ε with $0 < \varepsilon < r^{-s}$, and take $k_0 = 1/\varepsilon$. Set $\delta = \varepsilon^{1/s} < 1/r$. Let G be a graph with $n \geq k_0$ vertices, not including a copy of H.

Take an ε-regular partition of G, into $k+1$ classes $V_0, V_1, \dots, V_k$. Let's count how many edges G can have. There are at most $\frac{1}{k_0}\binom{n}{2} = \varepsilon\binom{n}{2}$ edges inside the classes. There are at most $n|V_0| \leq \varepsilon n^2$ incident with V_0. There are at most $\varepsilon\binom{n}{2}$ across pairs (V_i, V_j) that are *not* regular. There are at most $(\delta+\varepsilon)\binom{n}{2}$ across pairs that *are* regular, but with density at most $\delta+\varepsilon$. So far, we've accounted for at most $(5\varepsilon + \delta)\binom{n}{2}$ edges. The remaining edges (presumably most of them) are across regular pairs with density at least $\delta + \varepsilon$.

Construct an auxiliary "cluster graph" R on the vertex set $[k]$, joining i and j if (V_i, V_j) is an ε-regular pair of density at least $\delta + \varepsilon$. Let $e(R)$ be the number of edges in R. So the total number of edges in the graph G is at most $(5\varepsilon + \delta)\binom{n}{2} + e(R)|V_i|^2$.

But how large can $e(R)$ be? The embedding lemma tells us that, provided the $|V_i|$ are large enough, if R contains a clique of size r, then G contains a copy of H. So R doesn't contain a clique of size r, and by Turán's Theorem, $e(R) \leq t_{k,r-1} \leq \frac{r-2}{r-1}\frac{k^2}{2}$.

As $|V_i|^2\frac{k^2}{2} \leq \frac{n^2}{2}$, we have

$$e(G) \leq (5\varepsilon + \delta)\binom{n}{2} + \frac{r-2}{r-1}\frac{k^2}{2}|V_i|^2 \leq \frac{n^2}{2}\left(\frac{r-2}{r-1} + 5\varepsilon + \varepsilon^{1/s}\right).$$

Now we are done. Given any $\eta > 0$, we choose $\varepsilon > 0$ so that $5\varepsilon + \varepsilon^{1/s} < \eta$, and conclude that, if n is sufficiently large, then any graph G on n vertices with no copy of H as a subgraph has at most $\frac{n^2}{2}\left(\frac{r-2}{r-1} + \eta\right)$ edges, which is as required. □

There is one small catch. To apply the Embedding Lemma, we need to ensure that the classes have size at least ε^{-1}. The Regularity Lemma indeed does give us that, provided G has a little more than $\varepsilon^{-1}k_1(\varepsilon, k_0)$ vertices. With $\varepsilon = r^{-2s}$, which is what we need to get the error term η down around $1/r^2$, we need the number of vertices to be at least a tower of twos of height r^{10s}.

The original proof of the theorem, while rather more complicated, does give results for rather smaller graphs.

The type of proof we saw here is just the start. There are further tools (the *Counting Lemma*, the *Removal Lemma* and 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. For instance, 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. This is a Ramsey-theoretic result. There is a strengthening of this result: if a set S of integers has *positive upper density* (i.e., $\limsup_{n\to\infty} |S \cap [n]|/n > 0$), 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 $R_k(3) \le k(R_{k-1}(3) - 1) + 2$. [Hint: if you don't know the classic proof that $R_2(3) \le 6$, find and read that first.]

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

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 $\varepsilon > 0$, $p > 0$,

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

(You may need some estimates for tails of Binomial random variables.)
(b) Show that, for any bipartite graph H, and any fixed $p > 0$,

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

[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 $\{x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n\}$. 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.