Lee's proof of the Burr–Erdős is far too complicated to cover in this course, but we will try to see a few ideas in its direction. The Burr–Erdős conjecture has a long history, with many important partial results. The first major breakthrough in this direction was a theorem of Chvatál, Rödl, Szemerédi, and Trotter, which established the Burr–Erdős conjecture under the stronger assumption that H has bounded maximum degree.

Theorem 15.12 (Chvatál–Rödl–Szemerédi–Trotter)**.** *Graphs of bounded maximum degree have linear Ramsey numbers.*

More precisely, for every $\Delta \geqslant 1$, there exists $C \geqslant 1$ such that the following holds. If an n-vertex graph H has maximum degree at most Δ, then $r(H) \leqslant Cn$.

This result was extremely important, and so was the proof technique they introduced; this theorem is the first result in Ramsey theory to be proved via the so-called *regularity method*, whose basis is the fundamental *regularity lemma* of Szemerédi. This method has become one of the most important techniques in Ramsey theory and in extremal graph theory more broadly. However, let us remark that this proof technique gives truly enormous bounds on how large C has to be as a function of Δ; namely their proof showed that Theorem 15.12 is true for

$$C = \left.2^{2^{\cdot^{\cdot^{\cdot^{2}}}}}\right\} 2^{100\Delta}.$$

This enormous bound is one of several reasons why many researchers attempted to find alternative proofs of Theorem 15.12.

There are now (at least) two other techniques known for proving Theorem 15.12, both of which are very important in their own right. One is the *dependent random choice* technique, which you've seen a glimpse of on the homework, and which is also the main technique underlying Lee's proof of Conjecture 15.9. The other is the *greedy embedding technique*, which was developed in this context by Graham, Rödl, and Ruciński, although it goes back in some form at least to much earlier work of Erdős and Hajnal. We will unfortunately not have time to discuss this technique in detail in this course, but let us see a high-level overview of how it works.

Proof sketch of Theorem 15.12 using greedy embedding. Let H be an n-vertex graph of maximum degree at most Δ, and let $N = Cn$ for a large constant $C = C(\Delta)$ that we choose appropriately. Fix a red/blue coloring of $E(K_N)$. Our goal is to attempt to find a red copy of H in a greedy manner; we'll then show that, if we fail, we will be able to find a blue copy of H.

Let us label the vertices of H as $v_1, \dots, v_n$. Define $V_1 = V_2 = \cdots = V_n = V(K_N)$. We think of V_i as the set of candidate vertices for v_i, and will attempt to embed the vertices of H one at a time, at each step updating the set of candidate vertices. We fix some small parameter $\varepsilon > 0$.

Note that if we pick where to embed v_i into V_i, we need to update our candidate sets. Indeed, since our goal is to build a red copy of H, if we choose where to place v_i, we need to shrink each V_j, for all j such that $v_iv_j \in E(H)$, to only include the red neighbors of the chosen embedding of v_i. Let us call a vertex $w \in V_i$ *prolific* if it has the following property:

if we choose to embed v_i as w, then each candidate set shrinks by at most a factor of ε. In other words, w is prolific if its red neighborhood in V_j has size at least $\varepsilon|V_j|$, for every j such that $v_i v_j \in E(H)$.

Our embedding rule is now as follows. If there is a prolific vertex in V_1, we embed v_1 there and update all the candidate sets appropriately. If there is now a prolific vertex in V_2, we embed v_2 there and update all the candidate sets. We continue in this way as long as we can.

If this process gets to the end, that is, if we embed v_n into V_n, then we have found a red copy of H. So we may assume that the process gets stuck at some step i. Note that every candidate set shrinks at most Δ times, since H has maximum degree at most Δ, and moreover every time it shrinks it does so by at most a factor of ε. Thus, when we get stuck, we still have that $|V_j| \geqslant \varepsilon^{\Delta} N$ for all j. In particular, $|V_i| \geqslant \varepsilon^{\Delta} N$. Moreover, since we got stuck, there is no prolific vertex in V_i. That is, for every vertex $w \in V_i$, there is some j such that the red neighborhood of w in V_j has size less than $\varepsilon|V_j|$. There are at most Δ options for this choice of j, so by the pigeonhole principle, there is some fixed $j \in [\![n]\!]$ and some set $W_i \subseteq V_i$ with $|W_i| \geqslant \frac{1}{\Delta}|V_i|$ such that every $w \in W_i$ has a red neighborhood in V_j of size less than $\varepsilon|V_j|$.

We have thus proved the following lemma. If this greedy embedding procedure ever gets stuck, we find two sets W_i, V_j with $|W_i| \geqslant \frac{1}{\Delta}\varepsilon^{\Delta} N$ and $|V_j| \geqslant \varepsilon^{\Delta} N$, and with the property that the density of red edges between W_i and V_j is less than ε. In other words, we have found two sets A_1, A_2 with $|A_1|, |A_2| \geqslant \frac{1}{\Delta}\varepsilon^{\Delta} N$, and such that the density of *blue* edges between A_1 and A_2 is at least $1-\varepsilon$.

We now iterate this lemma, as follows. Inside A_1, we run the same procedure to attempt to greedily embed H in red. If we succeed, we are done. If we fail, we find two sets $A_{11}, A_{12} \subseteq A_1$ with blue density between them at least $1-\varepsilon$, where $|A_{11}|, |A_{12}| \geqslant (\frac{1}{\Delta}\varepsilon^{\Delta})^2 N$. We also run the same procedure inside A_2 to find two such sets A_{21}, A_{22}. Moreover, since the blue density between A_1 and A_2 was at least $1-\varepsilon$, we can ensure[9] that the blue density between A_{1i} and A_{2j} is at least $1-\varepsilon$, for all $i, j \in [\![2]\!]$.

In other words, we've now found *four* sets, each of size at least $(\frac{1}{\Delta}\varepsilon^{\Delta})^2 N$, such that the blue density between every pair is at least $1-\varepsilon$. Continuing in this manner k times, we can find 2^k such sets, each with size at least $(\frac{1}{\Delta}\varepsilon^{\Delta})^k N$, and with all pairwise blue densities at least $1-\varepsilon$. We now do this until $2^k \geqslant \Delta+1$ (i.e. pick $k = \lceil \log(\Delta+1) \rceil$), and we thus obtain at least $\Delta+1$ sets, which we rename $B_1, \ldots, B_{\Delta+1}$.

Since H has maximum degree at most Δ, it is $(\Delta+1)$-colorable, i.e. it can be partitioned into $\Delta+1$ independent sets $C_1, \ldots, C_{\Delta+1}$. Note that

$$|B_i| \geqslant \left(\frac{1}{\Delta}\varepsilon^{\Delta}\right)^k N \geqslant n,$$

[9]There is some subtlety in doing this step correctly; since A_{1i} and A_{2j} are rather small subsets of A_1, A_2, one needs an extra argument to ensure that the blue density remains high when we restrict to them. The trick to do this is to apply, essentially, Lemma 10.3 to always ensure that the minimum blue degree is high before shrinking.

where we can ensure the final inequality by picking C sufficiently large as a function of Δ and ε (and thus k, which is itself a function of Δ). Thus, each set B_i is large enough to accommodate embedding C_i. Moreover, one can check that if ε is sufficiently small (e.g. $\varepsilon = \Delta^{-2}$ suffices), then the greedy embedding strategy we tried for red is now *guaranteed* to work in blue. Namely, we greedily embed H in blue, ensuring that all vertices of C_i get embedded into B_i, and updating all candidate sets at every step. The strong density conditions we know about blue imply that we will never get stuck. □

Examining the proof sketch above, we see that it gives a bound of the form $C \leqslant 2^{O(\Delta(\log \Delta)^2)}$. Moreover, in case H is bipartite, the iteration step is unnecessary, and we can simply take $k = 1$ in the proof above, and thus obtain a bound of $C \leqslant 2^{O(\Delta \log \Delta)}$. In other words, the greedy embedding technique allowed Graham, Rödl, and Ruciński to prove the following more refined version of Theorem 15.12.

Theorem 15.13 (Graham–Rödl–Ruciński)**.** *There exists an absolute constant $M > 0$ such that the following holds. If H is an n-vertex graph with maximum degree at most Δ, then*

$$r(H) \leqslant 2^{M\Delta(\log \Delta)^2} n.$$

Moreover, if H is bipartite, we have the stronger bound

$$r(H) \leqslant 2^{M\Delta \log \Delta} n.$$

Remarkably, Graham, Rödl, and Ruciński also proved that their upper bound is nearly tight, even for bipartite graphs.

Theorem 15.14 (Graham–Rödl–Ruciński)**.** *There exists an absolute constant $c > 0$ such that the following holds. For every $n > \Delta > 1$, there is an n-vertex bipartite graph H with maximum degree Δ which satisfies*

$$r(H) \geqslant 2^{c\Delta} n.$$

Looking back at the greedy embedding proof sketch above, one might be struck by the fact that the colors play such asymmetrical roles; we keep trying, insistently, to embed H in red, and only when we have failed many times do we relent and succeed in embedding it in blue. This asymmetry is in fact a weakness of the proof technique, and Conlon, Fox, and Sudakov were able to improve the bound of Theorem 15.13 to $r(H) \leqslant 2^{O(\Delta \log \Delta)} n$ for every n-vertex graph H with maximum degree Δ, by modifying the greedy embedding technique so that both colors play roughly the same role. Unfortunately, it is still not known if this technique can be used to remove the final logarithmic factor, and thus match the lower bound of Theorem 15.14.

Moreover, this discussion hints at another, more fundamental, weakness of the greedy embedding technique, which is that it is tailor-made for the two-color case. Indeed, the entire upshot of the technique is that *failing* to find H in red tells us something about the blue edges. In case there are three or more colors, it is not at all clear how to obtain useful information from the failure of the first embedding. As far as I am aware, no one has been able to use the greedy embedding technique to prove any results on $r(H; q)$ for any H and any $q \geqslant 3$.

16 Canonical Ramsey theorems

This section covers two somewhat disparate topics, which nonetheless share a thematic connection. The *extremely* high-level idea is the following. Most mathematical objects are endowed with a notion of sub-objects (e.g. subsets, subgraphs, subgroups, subspaces, subschemes, subterfuges...). Certain objects are *canonical*, in the sense that all of their sub-objects "look like" the original object. For example, an elementary result in group theory is that all subgroups of a cyclic group are cyclic; a more pronounced version of the same fact is that any subgroup of $\mathbb{Z}$ is isomorphic to $\mathbb{Z}$. A substantially deeper and more difficult statement along the same lines is that any subgroup of a free group is again free.

One question we are interested in is a full classification of such examples: for any given notion of mathematical object, what is a complete list of the canonical ones? Having accomplished this task (which requires formalizing what we mean by "looking like" the original object), one can turn to proving a Ramsey-theoretic statement, along the lines of "any sufficiently large object must contain an arbitrarily large canonical sub-object".

We can view Ramsey's theorem as an instance of this general philosophy. Indeed, consider the class of graphs, endowed with the sub-object relation of induced subgraphs. Then complete graphs and empty graphs are examples of canonical objects, since any induced subgraph of a complete graph is again complete, and any induced subgraph of an empty graph is empty. Moreover, Ramsey's theorem implies that every sufficiently large graph contains an arbitrarily large complete or empty induced subgraph.

16.1 Monotone sequences

Consider a sequence $a_1, \ldots, a_k$ of distinct real numbers. A natural definition for a "canonical" sequence is a monotone sequence (that is, a sequence which either strictly increasing or strictly decreasing), since any subsequence of an increasing sequence is again increasing, and the same holds for decreasing sequences.

As you might expect, there is a Ramsey-theoretic statement, asserting that every sequence of distinct real numbers contains a long monotone subsequence; this was proved in the same seminal paper of Erdős and Szekeres.

Theorem 16.1 (Erdős–Szekeres)**.** *Given $k \geqslant 2$, let $N = (k-1)^2 + 1$. Then any sequence $a_1, \ldots, a_N$ of distinct real numbers contains a monotone subsequence of length k. That is, there exist indices $1 \leqslant i_1 < \cdots < i_k \leqslant N$ such that*

$$a_{i_1} < a_{i_2} < \cdots < a_{i_k} \qquad \textit{or} \qquad a_{i_1} > a_{i_2} > \cdots > a_{i_k}.$$

There are many known proofs of this theorem; we will show a particularly elegant proof discovered by Seidenberg.

Proof of Theorem 16.1 (Seidenberg). For an index $m \in [\![N]\!]$, let $\delta(m)$ denote the length of the longest decreasing subsequence ending at a_m, and let $\iota(m)$ denote the length of the longest increasing sequence ending at a_m. We wish to prove that $\delta(m) \geqslant k$ or $\iota(m) \geqslant k$

for some $m \in [\![N]\!]$. So suppose for contradiction that this is not the case, that is, that $1 \leqslant \delta(m), \iota(m) \leqslant k-1$; note that we have a lower bound of 1 on both functions, since we can always view a_m itself as both an increasing and a decreasing subsequence ending at a_m.

This means that there are at most $(k-1)^2$ possible values for the pair $(\delta(m), \iota(m))$. Since $N = (k-1)^2+1$, the pigeonhole principle implies that there exists two indices $1 \leqslant \ell < m \leqslant N$ such that $(\delta(\ell), \iota(\ell)) = (\delta(m), \iota(m))$. Since the elements of our sequence are distinct, we have $a_\ell < a_m$ or $a_\ell > a_m$. Suppose first that $a_\ell < a_m$. Then any increasing sequence ending in a_ℓ can be extended by one to obtain an increasing sequence ending at a_m, implying that $\iota(m) > \iota(\ell)$, a contradiction. Similarly, if $a_\ell > a_m$, then $\delta(m) > \delta(\ell)$, another contradiction. In either case we are done. □

It is not hard to show (as you will do on the homework) that this bound is tight, in that there exist sequences of $(k-1)^2$ distinct real numbers with no monotone subsequence of length k.