Math 216A Notes, Week 6
Scribe: Jesse Benavides

Disclaimer: These notes are not nearly as polished (and quite possibly not nearly as correct) as a published paper. Please use them at your own risk.

1. Ramsey’s Theorem

Let $K_n$ be the complete graph on $n$ vertices. Color the edges of $K_n$ red or blue. We would like to say something along the lines of "If $n$ is “large,” then some color has structure”. To be more precise, the structures we’ll be looking for are subgraphs that are entirely of one color or the other. Let $R(k, \ell)$ be the smallest $n$ (if it exists!) so that if we color edges of $K_n$, we get either red $K_k$ or a blue $K_\ell$.

Example 1 Claim: $R(3, 3) = 6$.

There are two things we need to show to see that this is true:
1) Show that $R(3, 3) \geq 6$. Consider the following:

Then in this case neither red nor blue has a triangle.
2) Show that $R(3, 3) \leq 6$.

Proof: Fix one vertex $v$. $v$ has 5 edges leaving it so there are $\geq 3$ red edges or $\geq 3$ blue edges. Without loss of generality, we may assume that there are 3 red edges leaving.

Opposite vertices of those 3 red edges either:
1) Span 3 blue edges, in which case we have a blue triangle
2) Span at least 1 red edge, in which case we have a red triangle. □

Theorem (Special Case of Ramsey’s Theorem) $R(k, \ell)$ is finite. Furthermore, it satisfies,

$$R(k, \ell) \leq R(k - 1, \ell) + R(k, \ell - 1)$$

Note that we have the trivial starting point $R(2, k) = R(k, 2) = k$. What this is saying is that “If I want to color edges $K_k$ with red or blue, either there is a red edge or everything’s blue.”

To prove the recursion, consider a graph on $R(k - 1, \ell) + R(k - 1, \ell)$ vertices. Pick some $v$. By the Pigeon
Hole Principle, there must be either $R(k - 1, \ell)$ red edges leaving $v$, or $R(k, \ell - 1)$ blue edges doing so. Without loss of generality let us assume the former. Let $S$ be the non-$v$ endpoints of those edges. By the definition of $R(k - 1, \ell)$ we know:
1) $S$ spans a red $K_{k-1}$
2) $S$ spans a blue $K_\ell$.
If blue $K_\ell$, then we are done. If red $K_{k-1}$, add $v$ to get a red $K_k$.

The finiteness follows inductively from the recursion and the base case $\min\{k, \ell\} = 2$.

**Theorem** $R(k, \ell) \leq \binom{k + \ell - 2}{k - 1} = \binom{k + \ell - 2}{\ell - 1}$

**Proof.** If $k$ (or $\ell$) = 2, we are done. Otherwise we induct on $k + \ell$, using our recursion to write

$$R(k, \ell) \leq R(k, \ell - 1) + R(k - 1, \ell)$$

$$\leq \binom{k + \ell - 3}{k - 2} + \binom{k + \ell - 3}{k - 1}$$

$$= \binom{k + \ell - 2}{k - 1} \quad \text{(by Pascal's triangle)}$$

\[\square\]

In particular, this gives us the bound

$$R(k, k) \leq \binom{2k - 2}{k - 1} \leq \frac{2^{2k-2}}{\sqrt{k}}$$

2. **Hypergraph Ramsey Theorem**

Fix $m \geq 2$. Instead of coloring edges of a graph, we now consider a coloring of all $m$-element subsets of \{1, \ldots, n\}. We'll say a $k$-element subset is monochromatic red if every $m$-element subset of it is red. We define $R(k, \ell; m)$ to be the smallest $n$ (again if it exists!) such that if we color the $m$-element subsets of \{1, \ldots, n\} red or blue, we have either a monochromatic red $k$-element subset or a monochromatic blue $\ell$-element subset.

**Theorem** $R(k, \ell; m)$ is finite, and we have

$$R(k, \ell; m) \leq R(R(k - 1, \ell; m) + R(k, \ell - 1; m); m - 1) + 1$$

**Proof.** As before, the first part of the theorem is implied by the recursion. Consider a coloring of $m$-element subsets of \{1, \ldots, n\}, where $n$ is the right hand side of the bound we are trying to prove.

Consider an auxiliary coloring on the $m - 1$ element subsets of \{1, \ldots, n - 1\}, where a subset $S$ red if and only if $S \cup \{n\}$ is red in the initial coloring. By hypothesis, I either have a monochromatic red subset of size $R(k - 1, \ell; m)$ or a monochromatic blue subset of size $R(k, \ell - 1; m)$.

Without loss of generality assume the former. Now consider the original coloring on the monochromatic red subset of size $R(k - 1, \ell; m)$. Again by the definition of $R$ we know it contains either:
1) a monochromatic blue subset of size $\ell$. OR
2) a monochromatic red subset of size $k - 1$. 

2
If 1) is true, then we are done. Otherwise every \( m \) element subset of this \( k - 1 \) element subset is red. Furthermore, by construction every subset formed by adding \( \{n\} \) to an \( m - 1 \) element subset of this set is red. If follows that adding \( \{n\} \) to this subset leads to a monochromatic \( k \)-element subset. □

Place \( N \) points in the plane in general position, i.e. no three points are on a line.

2.1. Convex \( n \)-gons via Ramsey Theory.

**Theorem (Erdős-Szekeres)** For all \( n \), there’s a \( N(n) \) such that if \( N > N(n) \), then the points contain a convex \( n \)-gon.

**Example 2** Any 5 points span a convex quadrilateral. This can be seen by considering the convex hull of those 5 points. If it is a pentagon or a quadrilateral, we are done already, while if it is a triangle we can see by inspection that the interior two points always form a convex \( n \)-gon with one side of the triangle.

**Proof.** (of theorem) Number the points arbitrarily from 1 to \( N \). Now consider the following coloring of the 3-element subsets of \( \{1, \ldots, N\} \). For each subset \( \{i, j, k\} \) with \( i < j < k \), we color it red if the triangle \( ijk \) is traversed clockwise going from \( i \rightarrow j \rightarrow k \)

\[
\begin{array}{c}
\text{i} \\
\text{k} \\
\text{j}
\end{array}
\]

and blue if it is traversed counterclockwise.

\[
\begin{array}{c}
\text{i} \\
\text{j} \\
\text{k}
\end{array}
\]

If \( N \geq R(n, n; 3) \), then we know there exists a monochromatic \( n \)-element subset in this coloring. Without loss of generality we may assume it it is monochromatic red.

**Claim:** This monochromatic set is a convex \( n \)-gon.

**Proof:** Suppose that the convex hull is not an \( n \)-gon, because some point \( d \) is inside triangle \( abc \). Without loss of generality we may assume \( a < b < c \)

\[
\begin{array}{c}
\text{a} \\
\text{d} \\
\text{c}
\end{array}
\]

and
For \(adc\) to be red, we need \(a < d < c\). For \(abd\) to be red, we need \(a < b < d\). This implies \(b < d < c\). So triangle \(bdc\) is traversed counterclockwise, and is thus blue. This is a contradiction since we assumed the hull was monochromatic red.

This proof shows a bound along the lines of

\[
N(n) < R(n; n; 3) \approx 2^n \text{ for some } c > 1
\]

(1935) Erdos-Szekéres proved, \(2^{n-2} + 1 \leq N(n) \leq \binom{2^{n-4}}{n-3} + 2\).

The lower bound is conjectured to be optimal here. After a long time, there finally came some progress in improving the upper bound.

(1989) Chung-Graham: \(N(n) \leq \binom{2^{n-4}}{n-3} + 1\)

The best current bound is due to Tóth and Valtr: \(N(n) \leq \binom{2^{n-5}}{n-3} + 2\).

3. **Upper Bounds on \(R(3, k)\)**

We now return to Ramsey’s Theorem for graphs, and the question of getting a better bound on \(R(k, 3) = R(3, k)\). Ramsey’s bound is that

\[
R(3, k) \leq \left(\frac{k + 1}{2}\right)^2 = \frac{k^2}{2}.
\]

“Every graph on \(n\) vertices has either a triangle or an independent set (i.e. no two vertices adjacent) of size \(c\sqrt{n}\).” Note that we can rewrite \(R(3, k) \leq k^2/2\) as \(R(3, c\sqrt{n}) \leq \frac{n}{2}\).

Question: Is there a better bound than \(c\sqrt{n}\)?

Answer: Yes.

**Theorem (Ajtai-Komlós-Szemerédi, 1980)** There is a \(c > 0\) such that

\[
R(3, k) \leq \frac{ck^2}{\log k}.
\]

Equivalently, every graph on \(n\)-vertices has either a triangle or an independent set of size \(c\sqrt{n \log n}\).

**Proof.** Suppose we have a triangle-free graph on \(n\)-vertices. If some \(v\) has some degree \(\geq k\), we’re done, since \(N(v)\) is independent.

So assume that all vertices have degree at most \(k\).

Let \(W\) be a uniformly chosen independent set on our graph. For example, if \(G\) is the graph below then

\[
P(W = \emptyset) = P(W = \{a\}) = P(W = \{b\}) = P(W = \{a, c\}) = \frac{1}{5}
\]
For each \( v \), let \( X_v = k \cdot |W \cap \{v\}| + |W \cap N(v)| \). This is sort of a score function for the vertex \( v \). If \( v \) is in the set, then \( v \) scores \( k \) points for itself, but also scores one point for \( X_w \) for each \( w \) adjacent to \( v \). By our degree assumption, each vertex in \( W \) gives out at most \( 2k \) points. In other words,

\[
\sum_v X_v \leq 2k|W|.
\]

To show that \( E(W) \) is large, therefore, it is enough to show that the expected sum of the \( X_v \) is large. In fact, we will show that each \( X_v \) is individually large:

**Lemma** \( E(X_v) \geq \frac{\log k}{8} \)

Suppose for a moment this lemma were true. Then we’d have

\[
E(\sum_v X_v) \geq \frac{n \log k}{8},
\]

which would imply

\[
E(|W|) \geq \frac{n \log k}{8k}
\]

This would imply our graph has some independent set which is at at least this large. We are aiming for an independent set of size \( k \), so we’d be done if \( n > \frac{16k^2}{\log k} \)

**Proof.** (of lemma) We first consider the special case where \( G \) is a star centered at \( v \) (meaning that every vertex is adjacent to \( v \)). If \( \text{deg}(v) = x \), then there are \( 2^x + 1 \) independent sets in \( G \): The singleton set \( \{v\} \) together with any subset not including \( v \). It follows that

\[
E(X_v) = \frac{k}{2x + 1} + \frac{2^x}{2x + 1} \cdot \frac{x}{2} \geq \frac{\log k}{8}
\]

where the inequality comes from optimizing over \( x \).

We now consider the general case. Our idea here is to condition on what happens outside of \( v \) and its neighborhood. So let \( T = W \cap (G - \{v\} - N(v)) \) be the elements of \( W \) not adjacent to \( v \). To show the lemma, it is enough to show it is true conditioned on every possible choice of \( T \), that is to say

\[
E(X_v | T = T_0) \geq \frac{\log k}{8}.
\]

Given \( T = T_0 \), we know for certain that \( W \) cannot contain any points in \( N(v) \) that are also adjacent to a vertex in \( T_0 \). But other than that, the choice of \( T_0 \) has no effect on which vertices of \( N(v) \cup \{v\} \) are included in \( T \). In other words, after conditioning on \( T = T_0 \) the behavior of \( X_v \) is the same as the behavior of \( X_v \) on a star consisting of \( v \) together with the vertices of \( N(v) \) not adjacent to \( T_0 \). But we’ve already handled this case, so we are done. \( \square \)
4. **Lower Bounds on** $R(k, k)$

So far we’ve proven the upper bounds

$$R(k, k) < \frac{4k^2}{\sqrt{k}}; \quad R(3, k) < \frac{ek^2}{\log k}.$$ 

We now turn to the question of obtaining lower bounds. To show $R(k, \ell) > n$, it is enough to show that there exists a coloring not having the monochromatic subgraphs in question. This (together with the intuition that the colors should be somehow spread out evenly to avoid holes) suggests that the probabilistic method may be of help here. In fact, this was one of the starting points of the probabilistic method.

**Theorem** (Erdős, 1947)

$$R(k, k) \geq (1 + o(1)) \frac{k(\sqrt{2})^k}{e\sqrt{2}}$$

**Proof.** Consider a uniform random coloring on $n$ vertices. (each edge is independently red with probability 0.5, blue with probability 0.5). Any set $S$ of $k$ vertices spans $\binom{k}{2} = \frac{k^2 - k}{2}$ edges. We therefore have

$$P(S \text{ mono. red }) = P(S \text{ mono. blue }) = 2^{-\binom{k}{2}}.$$ 

So the probability that a set is bad for us is twice this large, and

$$E(\text{No. of mono. } S) = \binom{n}{k} 2^{1 - \binom{k}{2}}.$$ 

If this bound is less than 1, we must have a coloring without monochromatic $S$, giving our bound on $R(k, k)$.

Now it just becomes a matter of running through the asymptotics. Here are a couple useful bounds on binomial coefficients.

- If $k$ is much smaller than $\sqrt{n}$, then

$$\binom{n}{k} = \frac{n(n-1) \cdots (n-k+1)}{k!} \approx \frac{n^k}{k!},$$

, since $(n-k+1) \approx n$ and the error is too small to matter even in the product of $k$ terms.

- Using Stirling’s approximation on $k$ in the above bound gives that for $k$ in this range

$$\binom{n}{k} \approx \left( \frac{ne}{k} \right)^k.$$ 

- It is **always** true that $\binom{n}{k} \leq \left( \frac{ne}{k} \right)^k$.

Since we’re looking for a bound where $n$ is exponentially larger than $k$, the $(ne/k)^k$ bound should be a good one. Using it, we can bound the expected number of monochromatic $S$ by

$$2 \left( \frac{ne}{k} \right)^k 2^{-\frac{k(k-1)}{2}} = 2 \left( \frac{ne}{k^2 \frac{e}{2}} \right)^k.$$ 

If $n$ is chosen so the base of the exponent is less than 1, this tends to 0.

\[ \square \]
It is a (notorious) open problem to improve this lower bound to \( c^k \) for any \( c > \sqrt{2} \), or to improve the upper bound to \( a^k \), for any \( a < 4 \). The best known upper bound is \( o\left( \frac{4^k}{k^c} \right) \) for all \( c > 0 \), due to David Conlon (2007).

It is also open to find a method of giving an explicit graph on \((1.00001)^k\) vertices without a clique of size \( k \) for large \( k \).

5. Lower bounds on \( R(3, k) \)

In this section we’ll again be thinking about the problem as finding a graph with no triangles and no independent sets of size \( k \). Again we want to proceed probabilistically, but we no longer want to take every edge in the graph with probability 0.5 (if we did, we’d get many triangles).

**First Attempt:** Take \( p < 1 \), which depends on \( n \), and consider a graph where every edge independently appears with probability \( p \). Our hope would be to optimize over \( p \) later.

The expected number of triangles in our graph is \( \binom{n}{3}p^3 \approx \frac{n^3}{6} \). So if we want to make this small we better take \( p < \frac{2}{n} \). But now the expected number of independent sets of size \( k \) is

\[
\binom{n}{k}(1-p)^{-\binom{k}{2}} \approx \left( \frac{ne}{k} \right)^k \cdot (1-p)^{\frac{k^2}{2}}
\]

As \( n \) gets large, \( p \) tends to 0, so we have \((1-p) \approx e^{-p} \) (Taylor series at 0). Using this, the approximate expected number of independent sets is

\[
\approx \left( \frac{ne}{k} \right)^k e^{-k^2 p} = \left( \frac{ne}{k e^{kp}} \right)^k
\]

But this doesn’t tend to 0 even if we take \( k = n/2 \), since we’re assuming \( p \) is so small. So this approach won’t give us anything better than \( R(3, k) \geq 2k \), which is far too weak.

**Improving the Bound by Deletion** The idea here is that if we cannot find a graph with no bad sets, then let’s find one with few bad sets.

Consider a graph \( G \), where \( G \) has no independent sets of size \( k \). Suppose \( G \) has three triangles, and \( G \) has \( n \) vertices, where \( n \) is large.

Then we can define another graph by removing 3 vertices; one from each triangle. Our new graph \( G' \) has \( n-3 \) vertices, no triangles, and no independent sets. More generally, if we have a graph with \( m \) bad sets (triangles/independent sets of size \( k \)), we can form a graph on (at least) \( n-m \) vertices with no bad sets.
("at least" here because it may be that deleting some vertices fixes several bad sets at once). If \( m \) is much smaller than \( n \), we’re in good shape.

We again take \( p \in [0, 1] \) to be chosen later. The expected bad sets in a random graph with edge probability \( p \) is

\[
\binom{n}{3} p^3 + \binom{n}{k} (1 - p)^{\binom{k}{2}}.
\]

After deleting one vertex from each bad set, the expected number of vertices which remains is at least

\[
n - \binom{n}{3} p^3 - \binom{n}{k} (1 - p)^{\binom{k}{2}}.
\]

To make sure that we don’t delete too many vertices, it is enough to make sure each of the subtracted terms is at most some small fraction of \( n \). Since \( \binom{n}{3} p^3 \leq \frac{n^3 p^3}{6} \), we can take \( p = n^{-\frac{2}{3}} \) to make the first term at most \( \frac{n}{6} \). For the second, we note

\[
\binom{n}{k} (1 - p)^{\binom{k}{2}} \leq \left( \frac{ne}{k} \right)^{\binom{k}{2}} e^{-\frac{k^2}{2}} \leq \left( \frac{ne}{ke^{\frac{k}{2}}} \right)^{k}.
\]

So we want the base of the exponent to be at most \( n^{1/k} \). This isn’t quite optimal, but it’s certainly enough to take \( k = 2n^{2/3} \log n \) and make the base \( o(1) \). What we’ve shown is that there exists a graph on \( \frac{5n}{6} \) vertices without a triangle or an independent set of size \( k \), i.e.

\[
R(3, 2n^{2/3} \log n) > \frac{5n}{6}
\]

Reversing the bounds, we have

\[
R(3, k) > c \cdot \left( \frac{k}{\log k} \right)^{\frac{3}{2}}.
\]

6. INTRODUCTION TO THE LOCAL LEMMA

So far, we’ve been applying our “probabilistic method” mostly in the following way:

1. There’s a lot of bad events, and we want to make sure none of them happen in our random coloring/graph/whatever.
2. Show that the expected number of bad events < 1 (and actually goes to 0).

Therefore 3. with non zero probability, no event occurs.

Really what we care about here is 3, and there are many situations where 3 happens even though 2 does not.

Example 3 Flip 10,000 coins. Let the bad event \( i \) be that coin \( i \) comes up heads. The expected number of bad events is 5000, but the probability no bad event occurs is \( 2^{-10000} > 0 \).

The key here was that the coin flips were all independent. What we’d like to have is something that lets us say if the bad events are "nearly" independent, there’s still a chance none of them happens. But to do this, we first need to define "nearly" independent. We will do this via what is called a dependency graph.

Definition A dependency graph \( H \) for set of events \( A_1, \ldots, A_n \) is a graph on \( n \) vertices such that \( A_i \) is (mutually) independent from the set \( \{ A_j \mid (i, j) \text{ not an edge } \} \).
"Mutually independent" here means that conditioning on any subset of the events \( A_j \) occurring or not occurring has no impact on the probability of \( A_i \).

**Example 4** \( X \) is a point chosen at random from \((1, 1, 0, 0), (1, 0, 1, 0), (1, 0, 0, 1), (0, 1, 0, 1), (0, 0, 1, 1), (1, 1, 1, 0), (0, 0, 0, 1)\).

Let \( X_i \) = event that the \( i^{th} \) coordinate of \( X \) is 1, and consider the three below graphs

Graph 3 is **not** a dependency graph. If I take \( i = 4 \) and condition on \( A_1 \) occurring but not \( A_2 \) or \( A_3 \), then I'm effectively conditioning on \( X = (1, 0, 0, 1) \), which changes the probability of \( A_4 \) to 1.

Graphs 1 and 2 are dependency graphs. For example, (looking at the first graph) I take \( i = 4 \) and condition on \( A_1 \) occurring but not \( A_2 \), then I'm effectively conditioning \( X \) to be uniform from \{ \((1, 0, 1, 0), (1, 0, 0, 1)\) \}, and the probability of \( A_4 \) remains equal to \( 1/2 \).

Note in particular that dependency graphs are not necessarily uniquely defined!

Armed with this definition, we can state one version of the "nearly independent" result we are looking for. We will prove this result next week as a special case of a more general Local Lemma.

**Theorem (Symmetric Lovász Local Lemma)** Suppose that \( A_1, \ldots, A_n \) are events such that

1. Each \( A_i \) has probability at most \( p \).
2. \( A_i \) have a dependency graph with maximum degree at most \( d \) and \( ep(d + 1) < 1 \).

Then with positive probability, no \( A_i \) occurs.

For example suppose \( A_1, \ldots, A_k \) are subsets of size \( n \).

**Question:** Can we color \( A_1 \cup \cdots \cup A_k \) with two colors such that no \( A_i \) is monochromatic?

Using the first moment method from before, we know the answer is yes if \( 2^{1-n} k < 1 \) (a random coloring works with positive probability. If \( A_i \) are disjoint, then it's trivial (assuming \( n \geq 2 \)). What the Local Lemma does is let us say the answer remains true if the sets are "nearly" disjoint.

Suppose that each set \( A_i \) intersects at most \( d \) other \( A_j \). Let \( H \) be the graph where \( i \) is adjacent to \( j \) (\( i \sim j \)) if and only if \( A_i \cap A_j \neq \emptyset \). Let \( B_i \) be the event that \( A_i \) is monochromatic in a random coloring.

Note that \( H \) is a dependency graph for the \( B_i \) events, since if we look at the points outside \( A_i \), the coloring of those points has no impact on whether \( A_i \) is monochromatic. Then \( P(B_i) = 2^{1-n} \), and by assumption no degree is larger than \( d \). So if \( 2^{1-n} e(d + 1) < 1 \), the Local Lemma guarantees a coloring. This happens if \( d \leq \frac{2^n}{2^{1-n} e(d + 1)} \).