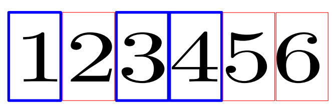
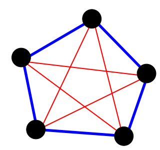
Lecture 4

Linear Algebra Methods in Combinatorics

with Applications to Geometry and CS



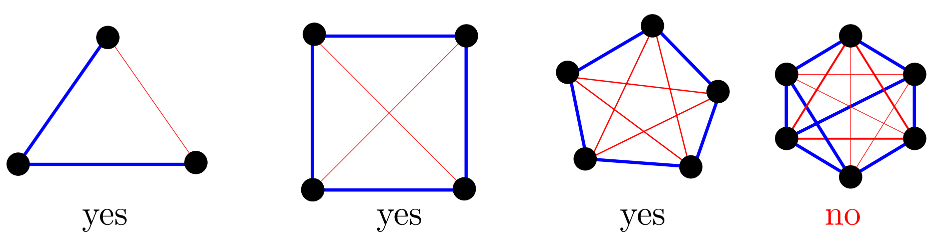


Tools from the previous lectures

Recall the **Odd**town problem and its upper bound.

1 Ramsey Theorem (and Ramsey graphs)

Color the edges of the complete graph over n nodes using two colors (red and blue) without creating a monochromatic triangle:

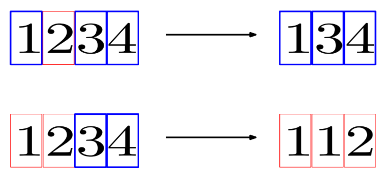


Exercise 1. *Prove that if you color the complete graph with **six nodes** with **two colors**, then you cannot avoid a **monochromatic triangle** (i.e., three nodes whose three edges get all the same color).*

We could increase the number of colors or the "size of the triangle", but eventually such monochromatic component must appear:

> **Ramsey Theorem for graphs (informal).** Fix the number of colors c and some parameter k. For n large enough, you cannot color the complete graph with n (or more) nodes without creating a complete subgraph of size k with all edges getting the same color.

Now try to **color numbers** instead, and try to avoid triples a, b, $a+b$ that get the same color (also for $a=b$):



> **Schur Theorem (informal)** No matter how we fix the number c of colors, if we color the integers $\{1,2,...,n\}$ with c colors, and n is large enough, then there exist integers a,b and $a+b$ that get the same color.

1.1 More about two colors

A nice trick to prove the Ramsey Theorem for two colors is to consider a generalization where we specify **different** size for the monochromatic components of the two colors:

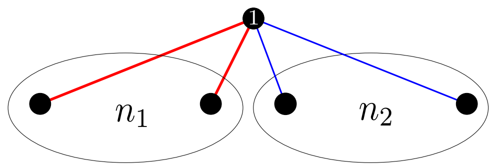
> No blue component of size s and no red component of size t.

Theorem 1 (Ramsey Theorem for graphs – two-color version)**.** *For every natural numbers s and t there exists a natural number $n = R(s,t)$ such that, if we color the complete graph with n or more nodes using colors red and blue, there must be either a red complete subgraph of size s or a blue complete subgraph of size t.*

Proof. We show by induction on $s+t$ that the number $R(s,t)$ exists (and is finite). The main idea is to prove the following:

$$R(s,t) \leq R(s-1,t) + R(s,t-1) \tag{1}$$

This inequality means that, if we color the graph with $n = R(s-1, t)+R(s, t-1)$ nodes, then either there is a monochromatic red component of size s or a monochromatic blue component of size t. By contradiction, suppose we can color this graph without creating these monochromatic components (red of size s or blue of size t). Take any node (say node 1) and look at the neighbors (all other nodes):



The coloring of these edges splits these nodes into two groups (as shown in the figure). The key observation is that the n_1 nodes on the left cannot form a red component of size $s - 1$ (otherwise this component plus node 1 yields a red component of size s). Of course we still "forbid" a blue component of size t among these nodes. So, by inductive hypothesis, we have that

$$n_1 < R(s - 1, t)$$

and by a symmetric argument

$$n_2 < R(s, t - 1)$$

This is not possible because $n = 1 + n_1 + n_2$, while these inequalities tell us $n_1 \leq R(s - 1, t) - 1$ and $n_2 \leq R(s, t - 1) - 1$. To conclude the proof, we observe that the **base case** for the induction is given by

$$R(1, 1) = R(1, t) = R(s, 1) = 1$$

because a single node is a "monochromatic" component of size 1 (all of the edges get the same color...just because there are no edges to color).[1] □

The proof can be turned into a concrete upper bound:

$$R(s, t) \leq \binom{s + t - 2}{s - 1} \quad \text{and} \quad R(t, t) < 4^t \tag{2}$$

[1]If you find this strange, you can start with a base case $R(2, t) = t + 1 = R(t, 2)$.

Exercise 2. *Prove (2) by using the identity* $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$.

The two colors can be viewed as "red $\equiv$ remove the edge":

Two-colors revisited:

- Every "sufficiently large" graph must contain a **clique** of size t or an **independent set** of size t (Ramsey Theorem restated).
- t**-Ramsey graph** is a graph which does not contain a clique of size t nor an independent set of size t (a coloring with no monochromatic component of size t).

$$\underbrace{LB(t) \leq}_{\text{find a construction}} \qquad R(t,t) \qquad \underbrace{\leq UB(t)}_{\text{prove the theorem}}$$

How large can be a t-Ramsey graph?

- No t-Ramsey graph of size 4^t (proof above).
- There are t-Ramsey graphs of size $2^{t/2}$ (probabilistic method – below)

which in other words means

$$\left(\sqrt{2}\right)^t \leq R(t,t) \leq 4^t \tag{3}$$

1.2 Probabilistic method

There exist t-Ramsey graphs of size $2^{t/2}$.

Color each edge **blue** with probability 1/2 and **red** with probability 1/2. If this is done **independently** for each edge, then we can show that the probability of a monochromatic component is **less than** 1 (if the graph is "not too big").

- Every subset S of size t has $\binom{t}{2}$ edges and the probability that **this** subset is monochromatic is

$$\Pr[MON_S] = 2 \cdot 1/2^{\binom{t}{2}}$$

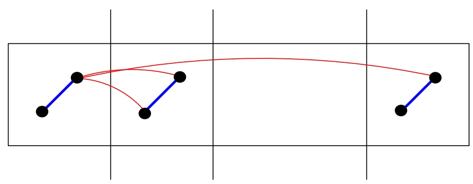
- The **union bound**[2] implies that the probability of some monochromatic subset is at most

$$\Pr\left[\bigcup_S MON_S\right] \leq \binom{n}{t} \cdot 1/2^{\binom{t}{2}-1}$$

- For $n \leq 2^{t/2}$ this probability is < 1 and thus there must be one coloring with no monochromatic component of size t.

2 Explicit constructions

Quadratic construction. Partition all nodes into blocks of the same size (except maybe the last block). Color edges "inside the same block" blue, and edges across blocks red:



The largest monochromatic **blue** component is at most the size of a block. The largest monochromatic **red** component is at most the number of blocks.

Exercise 3. *Prove that this construction yields t-Ramsey graphs of size* $\Theta(t^2)$.

Cubic construction. Each node corresponds to a subset $S_i \subset [n]$ of size 3 (so we have $\binom{n}{3} = \Theta(n^3)$ nodes). The color of each edge (S_i, S_j) depends on the cardinality of the intersection

$$|S_i \cap S_j| \in \{0, 1, 2\}$$

Suppose we color as follows:

- **Red** if $|S_i \cap S_j|$ is even;
- **Blue** if $|S_i \cap S_j|$ is odd.

[2] $\Pr[E_1 \cup \cdots \cup E_k] \leq \Pr[E_1] + \cdots + \Pr[E_k]$

Exercise 4. *Prove that the largest* ***red*** *complete subgraph has size at most n. (**Hint:** use the Oddtown problem.)*

For the **blue** complete subgraphs we use the following result (which is interesting by itself):

Theorem 2 (Fisher inequality)**.** *If subsets $S_1, \ldots, S_m \subseteq [n]$ satisfy*

$$|S_i \cap S_j| = c \quad \textit{for all } i \neq j$$

then $m \leq n$.

Proof. We first prove the theorem for the case in which all subsets have the same cardinality larger than c, i.e., $|S_i| > c$. Consider the 0/1-vectors corresponding to these subsets. Then the condition on the intersections becomes

$$v_i \cdot v_j = c \quad \text{for all } i \neq j$$

We show that these vectors must be linearly independent. If

$$\lambda_1 v_1 + \cdots + \lambda_m v_m = \mathbf{0}$$

then

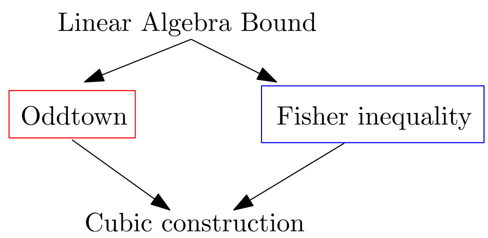
$$\begin{aligned} 0 &= (\lambda_1 v_1 + \cdots + \lambda_m v_m)(\lambda_1 v_1 + \cdots + \lambda_m v_m) && (4)\\ &= \sum_i \lambda_i^2 |S_i| + \sum_{i,j \neq i} \lambda_i \lambda_j c && (5)\\ &= \sum_i (\lambda_i^2 |S_i| - \lambda_i^2 c) + \sum_{i,j} \lambda_i \lambda_j c && (6) \end{aligned}$$

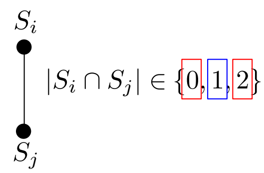
The first summation is positive if at least one $\lambda_i \neq 0$ (here we use $|S_i| > c$). The second summation is equal to $c(\lambda_1 + \cdots + \lambda_m)^2$ which is nonnegative. We conclude that it must be $\lambda_1 = \cdots = \lambda_m = 0$. These vectors are linearly independent and the linear algebra bound yields $m \leq n$.

Finally, for the general case, we observe that it cannot happen that $|S_i| < c$. If one $|S_i| = c$ then a simple argument shows that $m \leq n$ (**Exercise**!). □

Fisher inequality says that, in our construction, the largest **blue** complete subgraph has size at most n (**Exercise!**) and thus:

> There is an explicit construction of $(t+1)$-Ramsey graphs of size $\binom{t}{3}$.





3 What to remember and where to look

The idea of the cubic construction is to color edges according to the "distances" between the endpoints: here distance means **size** of the intersection. The idea is that there cannot be a "too large" red component because this would give a "large" solution for Oddtown (and we know already that this cannot happen). For the blue component, we introduced a similar theorem (Fisher inequality). In the next lecture we will use this idea, but with more general theorems so we can increase the size of the graph that we construct (from cubic to superpolynomial).

The proofs of the Ramsey Theorem (Theorem 1) and of Fisher inequality (Theorem 2) are from [Juk01]. The cubic construction of Ramsey graphs is in [BF92, Sect. 4.2].

More about Ramsey theorems (if you want to know more). It is possible to generalize the result in Theorem 1 to this case: color r-regular **hypergraphs**, that is, color all subsets of size r of $[n]$ using any number c of colors (see e.g. [Juk01]). The exact values of the Ramsey Numbers $R(t) = R(t,t)$ are known only for $R(3) = 6$ and $R(4) = 18$. For $t = 5$ we only know that $43 \leq R(5) \leq 48$ (two years ago!!), while $798 \leq R(10) \leq 23556$ (big gap!!) You can find these bounds in [Rad94]. To see how a 5-Ramsey graph is constructed look at [Exo89] (less than two pages!!).

References

[BF92] L. Babai and P. Frankl. *Linear algebra methods in combinatorics with applications to geometry and computer science.* The University of Chicago, 1992.

[Exo89] G. Exoo. A lower bound for $R(5,5)$. *Journal of graph theory*, 13(1):97–98, 1989.

[Juk01] S. Jukna. *Extremal combinatorics: with applications in computer science*. Springer Verlag, 2001.

[Rad94] S.P. Radziszowski. Small ramsey numbers. *Electronic Journal of Combinatorics*, 1:28, 1994.

Exercises
(during next exercise class - 15.3.2018)

We shall discuss and solve together the following exercise:

Exercise 5. *Here is an improvement of the cubic construction of Ramsey graphs. Each node corresponds to a subset* $S_i \subset [n]$ *of size* 4 *(so we have* $\binom{n}{4} = \Theta(n^4)$ *nodes). The color of each edge* (S_i, S_j) *depends on the cardinality of the intersection*

$$|S_i \cap S_j| \in \{0, 1, 2, 3\}$$

Suppose we color as follows:

- ***Red*** *if* $|S_i \cap S_j|$ *is even;*
- ***Blue*** *if* $|S_i \cap S_j|$ *is odd.*

Claim: *The largest **red** complete subgraph has size at most* n*. Same for the largest **blue** complete subgraph.*

*The claim is **false** and the construction does not give any improvement. Explain why. Try the same with subsets of size* 5 *instead. Does it give an improvement over the cubic construction?*