Figure 1: Paul Erdős (1913–1996) is considered the father of the probabilistic method. He published around 1,500 papers during his lifetime, and had more than 500 collaborators. To learn more about Erdős, see his biography *The man who loved only numbers* by Hoffman and the documentary *N is a number*.

1 Introduction

Probabilistic method: to prove that an object exists, show that a random construction works with positive probability

Tackle combinatorics problems by introducing randomness

Theorem 1.0.1. Every graph $G = (V, E)$ contains a bipartite subgraph with at least $|E|/2$ edges.

Proof. Randomly color every vertex of G with black or white, iid uniform

Let E' = edges with one end black and one end white

Then (V, E') is a bipartite subgraph of G

Every edge belongs to E' with probability $\frac{1}{2}$, so by linearity of expectation, $\mathbb{E}[|E'|] = \frac{1}{2}|E|$.

Thus there is some coloring with $|E'| \geq \frac{1}{2}|E|$, giving the desired bipartite subgraph. $\square$

1.1 Lower bounds to Ramsey numbers

Ramsey number $R(k, \ell)$ = smallest n such that in every red-blue edge coloring of K_n, there exists a red K_k or a blue K_ℓ.

e.g., $R(3, 3) = 6$

Ramsey (1929) proved that $R(k, \ell)$ exists and is finite

Figure 2: Frank Ramsey (1903–1930) wrote seminal papers in philosophy, economics, and mathematical logic, before his untimely death at the age of 26 from liver problems. See a recent profile of him in the New Yorker.

1.1.1 Erdős' original proof

The probabilistic method started with:
P. Erdős, Some remarks on the theory of graphs, BAMS, 1947

Remark 1.1.1 (Hungarian names). Typing "Erdős" in LaTeX: `Erd\H{o}s` and *not* `Erd\"os`

Hungarian pronunciations: `s` = `/sh/` and `sz` = `/s/`, e.g., Erdős, Szekeres, Lovász

Theorem 1.1.2 (Erdős 1947)**.** If $\binom{n}{k}2^{1-\binom{k}{2}} < 1$, then $R(k,k) > n$. In other words, there exist a red-blue edge-coloring of K_n without a monochromatic K_k.

Proof. Color edges uniformly at random

For every fixed subset R of k vertices, let A_R denote the event that R induces a monochromatic K_k. Then $\mathbb{P}(A_R) = 2^{1-\binom{k}{2}}$.

$$\mathbb{P}(\text{there exists a monochromatic } K_k) = \mathbb{P}\left(\bigcup_{R\in\binom{[n]}{k}} A_R\right) \le \sum_{R\in\binom{[n]}{k}} \mathbb{P}(A_R) = \binom{n}{k}2^{1-\binom{k}{2}} < 1.$$

Thus, with positive probability, the random coloring gives no monochromatic K_k. □

Remark 1.1.3. By optimizing n (using Stirling's formula) above, we obtain

$$R(k,k) > \left(\frac{1}{e\sqrt{2}} + o(1)\right) k2^{k/2}$$

Can be alternatively phrased as counting: of all $2^{\binom{n}{2}}$ possible colorings, not all are bad (this was how the argument was phrased in the original Erdős 1947 paper.

In this course, we almost always only consider finite probability spaces. While in principle the finite probability arguments can be rephrased as counting, but some of the later more involved arguments are impractical without a probabilistic perspective.

Constructive lower bounds? Algorithmic? Open! “Finding hay in a haystack”

Remark 1.1.4 (Ramsey number upper bounds). Erdős–Szekeres (1935):

$$R(k+1,\ell+1) \le \binom{k+\ell}{k}.$$

Recent improvements by Conlon (2009), and most recently Sah (2020+):

$$R(k+1,k+1) \le e^{-c(\log k)^2}\binom{2k}{k}.$$

All these bounds have the form $R(k,k) \le (4+o(1))^k$. It is a major open problem whether $R(k,k) \le (4-c)^k$ is true for some constant $c > 0$ and all sufficiently large k.

1.1.2 Alteration method

Two steps: (1) randomly color (2) get rid of bad parts

Theorem 1.1.5. For any k, n, we have $R(k,k) > n - \binom{n}{k}2^{1-\binom{k}{2}}$.

Proof. Construct in two steps:
(1) Randomly 2-color the edges of K_n
(2) Delete a vertex from every monochromatic K_k
Final graph has no monochromatic K_k

After step (1), every fixed K_k is monochromatic with probability $2^{1-\binom{k}{2}}$, let X be the number of monochromatic K_k's. $\mathbb{E}X = \binom{n}{k}2^{1-\binom{k}{2}}$.

We delete at most $|X|$ vertices in step (2). Thus final graph has size $\ge n - |X|$, which has expectation $n - \binom{n}{k}2^{1-\binom{k}{2}}$.

Thus with positive probability, the remaining graph has size at least $n - \binom{n}{k}2^{1-\binom{k}{2}}$ (and no monochromatic K_k by construction) $\square$

Remark 1.1.6. By optimizing the choice of n in the theorem, we obtain

$$R(k,k) > \left(\frac{1}{e} + o(1)\right) k 2^{k/2},$$

which improves the previous bound by a constant factor of $\sqrt{2}$.

1.1.3 Lovász local lemma

We give one more improvement to the lower bound, using the Lovász local lemma, which we will prove later in the course

Consider "bad events" $E_1, \dots, E_n$. We want to avoid all.

If all $\mathbb{P}(E_i)$ small, say $\sum_i \mathbb{P}(E_i) < 1$, then can avoid all bad events.

Or, if they are all independent, then the probability that none of E_i occurs is $\prod_{i=1}^n (1 - \mathbb{P}(E_i)) > 0$ (provided that all $\mathbb{P}(E_i) < 1$).

What if there are some weak dependencies?

Theorem 1.1.7 (Lovász local lemma)**.** Let $E_1, \dots, E_n$ be events, with $\mathbb{P}[E_i] \le p$ for all i. Suppose that each E_i is independent of all other E_j except for at most d of them. If

$$ep(d+1) < 1,$$

then with some positive probability, none of the events E_i occur.

Remark 1.1.8. The meaning of "independent of ..." is actually somewhat subtle (and easily mistaken). We will come back to this issue later on when we discuss the local lemma in more detail.

Theorem 1.1.9 (Spencer 1977)**.** If $e\left(\binom{k}{2}\binom{n}{k-2}+1\right)2^{1-\binom{k}{2}} < 1$, then $R(k,k) > n$.

Proof. Random 2-color edges of K_n

For each k-vertex subset R, let E_R be the event that R induces a monochromatic K_k. $\mathbb{P}[E_R] = 2^{1-\binom{k}{2}}$.

E_R is independent of all E_S other than those such that $|R \cap S| \ge 2$

For each R, there are at most $\binom{k}{2}\binom{n}{k-2}$ choices S with $|S| = k$ and $|R \cap S| \ge 2$.

Apply Lovász local lemma to the events $\left\{E_R : R \in \binom{V}{k}\right\}$ and $p = 2^{1-\binom{k}{2}}$ and $d = \binom{k}{2}\binom{n}{k-2}$, we get that with positive probability none of the events E_R occur, which gives a coloring with no monochromatic K_k's. $\square$

Remark 1.1.10. By optimizing the choice of n, we obtain

$$R(k,k) > \left(\frac{\sqrt{2}}{e} + o(1)\right) k 2^{k/2}$$

once again improving the previous bound by a constant factor of $\sqrt{2}$. This is the best known lower bound to $R(k,k)$ to date.

1.2 Set systems

1.2.1 Sperner's theorem

Let $\mathcal{F}$ a collection of subsets of $\{1,2,\dots,n\}$. We say that $\mathcal{F}$ is an **antichain** if no set in $\mathcal{F}$ is contained in another set in $\mathcal{F}$.

Question 1.2.1. What is the maximum number of sets in an antichain?

Example: $\mathcal{F} = \binom{[n]}{k}$ has size $\binom{n}{k}$. Maximized when $k = \lfloor \frac{n}{2} \rfloor$ or $\lceil \frac{n}{2} \rceil$. The next result shows that we cannot do better.

Theorem 1.2.2 (Sperner 1928)**.** If $\mathcal{F}$ is an antichain of subsets of $\{1,2,\dots,n\}$, then $|\mathcal{F}| \le \dbinom{n}{\lfloor n/2 \rfloor}$.

In fact, we will show an even stronger result:

Theorem 1.2.3 (LYM inequality; Bollobás 1965, Lubell 1966, Meshalkin 1963, and Yamamoto 1954)**.** If $\mathcal{F}$ is an antichain of subsets of $[n]$, then

$$\sum_{A \in \mathcal{F}} \frac{1}{\binom{n}{|A|}} \le 1.$$

Sperner's theorem follows since $\binom{n}{|A|} \ge \binom{n}{\lfloor n/2 \rfloor}$.

Proof. Consider a random permutation σ of $\{1,2,\dots,n\}$, and its associated chain of subsets

$$\emptyset, \{\sigma(1)\}, \{\sigma(1),\sigma(2)\}, \{\sigma(1),\sigma(2),\sigma(3)\}, \dots, \{\sigma(1),\dots,\sigma(n)\}$$

where the last set is always equal to $\{1,2,\dots,n\}$. For each $A \subset \{1,2,\dots,n\}$, let E_A denote the event that A is found in this chain. Then

$$\mathbb{P}(E_A) = \frac{|A|!(n-|A|)!}{n!} = \frac{1}{\binom{n}{|A|}}.$$

Since $\mathcal{F}$ is an antichain, if $A, B \in \mathcal{F}$ are distinct, then E_A and E_B cannot both occur. So $\{E_A : A \in \mathcal{F}\}$ is a set of disjoint event, and thus their probabilities sum to at most 1. $\square$

1.2.2 Bollobás two families theorem

Sperner's theorem is generalized by the following celebrated result of Bollobás, which has many more generalizations that we will not discuss here.

Theorem 1.2.4 (Bollobás (1965) "two families theorem")**.** Let $A_1, \dots, A_m$ be r-element sets and $B_1, \dots, B_m$ be s-element sets such that $A_i \cap B_i = \emptyset$ for all i and $A_i \cap B_j \neq \emptyset$ for all $i \neq j$. Then $m \leq \binom{r+s}{r}$.

Remark 1.2.5. The bound is sharp: let A_i range over all r-element subsets of $[r+s]$ and set $B_i = [r+s] \setminus A_i$.

Let us give an application/motivation for Bollobás' two families theorem in terms of transversals.

Given a set family $\mathcal{F}$, say that T is a **transversal** for $\mathcal{F}$ if $T \cap S \neq \emptyset$ for all $S \in \mathcal{F}$ (i.e., T hits every element of $\mathcal{F}$).

Let $\tau(\mathcal{F})$, the **transversal number** of $\mathcal{F}$, be the size of the smallest transversal of $\mathcal{F}$.

Say that $\mathcal{F}$ is **τ-critical** if $\tau(\mathcal{F} \setminus \{S\}) < \tau(\mathcal{F})$ for all $S \in \mathcal{F}$.

Question 1.2.6. What is the maximum size of a τ-critical r-uniform $\mathcal{F}$ with $\tau(\mathcal{F}) = s+1$?

We claim that the answer is $\binom{r+s}{r}$. Indeed, let $\mathcal{F} = \{A_1, \dots, A_m\}$, and B_i an s-element transversal of $\mathcal{F} \setminus \{A_i\}$ for each i. Then the condition is satisfied. Thus $m \leq \binom{r+s}{r}$.

Conversely, $\mathcal{F} = \binom{[r+s]}{r}$ is τ-critcal r-uniform with $\tau(\mathcal{F}) = s+1$. (why?)

Here is a more general statement of the Bollobás' two-family theorem.

Theorem 1.2.7. Let $A_1, \dots, A_m$ and $B_1, \dots, B_m$ be finite sets such that $A_i \cap B_i = \emptyset$ for all i and $A_i \cap B_j \neq \emptyset$ for all $i \neq j$. Then

$$\sum_{i=1}^{m} \binom{|A_i| + |B_i|}{|A_i|}^{-1} \leq 1.$$

Note that Sperner's theorem and LYM inequality are also special cases, since if $\{A_1, \dots, A_m\}$ is an antichain, then setting $B_i = [n] \setminus A_i$ for all i satisfies the hypothesis.

Proof. Consider a uniform random ordering of all elements.

Let X_i be the event that all elements of A_i come before B_i.

Then $\mathbb{P}[X_i] = \binom{|A_i|+|B_i|}{|A_i|}^{-1}$ (all permutations of $A_i \cup B_i$ are equally likely to occur).

Note that the events X_i are disjoint (X_i and X_j both occuring would contradict the hypothesis for A_i, B_i, A_j, B_j). Thus $\sum_i \mathbb{P}[X_i] \le 1$. $\square$

1.2.3 Erdős–Ko–Rado theorem on intersecting families

A family $\mathcal{F}$ of sets is **intersecting** if $A \cap B \neq \emptyset$ for all $A, B \in \mathcal{F}$.

Question 1.2.8. What is the largest intersecting family of k-element subsets of $[n]$?

Example: $\mathcal{F}$ = all subsets containing the element 1. Then $\mathcal{F}$ is intersecting and $|\mathcal{F}| = \binom{n-1}{k-1}$

Theorem 1.2.9 (Erdős–Ko–Rado 1961; proved in 1938)**.** If $n \ge 2k$, then every intersecting family of k-element subsets of $[n]$ has size at most $\binom{n-1}{k-1}$.

Remark 1.2.10. The assumption $n \ge 2k$ is necessary since if $n < 2k$, then the family of all k-element subsets of $[n]$ is automatically intersecting by pigeonhole.

Proof. Consider a uniform random circular permutation of $1, 2, \ldots, n$ (arrange them randomly around a circle)

For each k-element subset A of $[n]$, we say that A is **contiguous** if all the elements of A lie in a contiguous block on the circle.

The probability that A forms a contiguous set on the circle is exactly $n/\binom{n}{k}$.

So the expected number of contiguous sets in $\mathcal{F}$ is exactly $n\,|\mathcal{F}|\,/\binom{n}{k}$.

Since $\mathcal{F}$ is intersecting, there are at most k contiguous sets in $\mathcal{F}$ (under every circular ordering of $[n]$). Indeed, suppose that $A \in \mathcal{F}$ is contiguous. Then there are $2(k-1)$ other contingous sets (not necessarily in $\mathcal{F}$) that intersect A, but they can be paired off into disjoint pairs. Since $\mathcal{F}$ is intersecting, it follows that it contains at most k contiguous sets.

Combining with result from the previous paragraph, we see that $n\,|\mathcal{F}|\,/\binom{n}{k} \le k$, and hence $|\mathcal{F}| \le \frac{k}{n}\binom{n}{k} = \binom{n-1}{k-1}$. $\square$

1.3 2-colorable hypergraphs

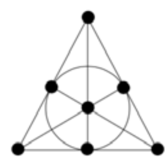
An **k-uniform hypergraph** (or **k-graph**) is a pair $H = (V, E)$, where V (vertices) is a finite set and E (edges) is a set of k-element subsets of E, i.e., $E \subseteq \binom{V}{k}$ (so hypergraphs are really the same concept as set families).

We say that H is **r-colorable** if the vertices can be colored using r colors so that no edge is monochromatic.

Let $m(k)$ denote the minimum number of edges in a k-uniform hypergraph that is not 2-colorable (elsewhere in the literature, "2-colorable" = "property B", named after Bernstein who introduced the concept in 1908)

$m(2) = 3$

$m(3) = 7$. Example: Fano plane (below) is not 2-colorable (the other direction is by exhaustive search)



$m(4) = 23$, proved via exhaustive computer search (Östergård 2014)

Exact value of $m(k)$ is unknown for all $k \geq 5$

The probabilistic method gives a short proof of a lower bound (random coloring):

Theorem 1.3.1 (Erdős 1964)**.** For any $k \geq 2$, $m(k) \geq 2^{k-1}$, i.e., every k-uniform hypergraph with fewer than 2^{k-1} edges is 2-colorable.

Proof. Let there be $m < 2^{k-1}$ edges. In a random 2-coloring, the probability that there is a monochromatic edge is $\leq 2^{-k+1}m < 1$. □

Remark 1.3.2. Later on we will prove an better lower bound $m(k) \gtrsim 2^k \sqrt{k/\log k}$, which is the best known to date.

Perhaps somewhat surprisingly, the state of the art upper bound is also proved using probabilistic method (random construction).

Theorem 1.3.3 (Erdős 1964)**.** $m(k) = O(k^2 2^k)$, i.e., there exists a k-uniform hypergraph with $O(k^2 2^k)$ edges that is not 2-colorable.

Proof. Fix $|V| = n$ to be decided. Let H be the k-uniform hypergraph obtained by choosing m random edges (with replacement) $S_1, \dots, S_m$.

Given a coloring $\chi\colon V \to [2]$, let A_χ denote the event that χ is a proper coloring (i.e., no monochromatic edges). It suffices to check that $\sum_\chi \mathbb{P}[A_\chi] < 1$.

If χ colors a vertices with one color and b vertices with the other color, then the probability that (random) S_1 is monochromatic under (fixed) χ is

$$\frac{\binom{a}{k}+\binom{b}{k}}{\binom{n}{k}} \ge \frac{2\binom{n/2}{k}}{\binom{n}{k}} = \frac{2(n/2)(n/2-1)\cdots(n/2-k+1)}{n(n-1)\cdots(n-k+1)}$$
$$\ge 2\left(\frac{n/2-k+1}{n-k+1}\right)^k = 2^{-k+1}\left(1-\frac{k-1}{n-k+1}\right)^k$$

Setting $n = k^2$, we see that the above quantity is at least $c2^{-k}$ for some constant $c > 0$.

Thus, the probability that χ is a proper coloring (i.e., no monochromatic edges) is at most $(1-c2^{-k})^m \le e^{-c2^{-k}m}$ (using $1+x \le e^x$ for all real x).

Thus, $\sum_\chi \mathbb{P}[A_\chi] \le 2^n e^{-c2^{-k}m} < 1$ for some $m = O(k^2 2^k)$ (recall $n = k^2$). $\square$

1.4 List chromatic number of $K_{n,n}$

Given a graph G, its **chromatic number** $\chi(G)$ is the minimum number of colors required to proper color its vertices.

In **list coloring**, each vertex of G is assigned a list of allowable colors. We say that G is **k-choosable** (also called **k-list colorable**) if it has a proper coloring no matter how one assigns a list of k colors to each vertex.

We write $\mathrm{ch}(G)$, called the **choosability** (also called: **choice number**, **list colorability**, **list chromatic number**) of G, to be the smallest k so that G is k-choosable.

It should be clear that $\chi(G) \le \mathrm{ch}(G)$, but the inequality may be strict.

For example, while every bipartite graph is 2-colorable, $K_{3,3}$ is not 2-choosable. Indeed, no list coloring of $K_{3,3}$ is possible with color lists (check!):

$$\begin{array}{cc} \{2,3\} & \{2,3\} \\ \{1,3\} & \{1,3\} \\ \{1,2\} & \{1,2\} \end{array}$$

Easy to check then that $\mathrm{ch}(K_{3,3}) = 3$.

Question 1.4.1. What is the asymptotic behavior of $\mathrm{ch}(K_{n,n})$?

First we prove an upper bound on $\mathrm{ch}(K_{n,n})$.

Theorem 1.4.2. If $n < 2^{k-1}$, then $K_{n,n}$ is k-choosable.

In other words, $\mathrm{ch}(K_{n,n}) \leq \lfloor \log_2(2n) \rfloor + 1$.

Proof. For each color, mark it either "L" or "R" iid uniformly.

For any vertex of $K_{n,n}$ on the left part, remove all its colors marked R.

For any vertex of $K_{n,n}$ on the right part, remove all its colors marked L.

The probability that some vertex has no colors remaining is at most $2n2^{-k} < 1$. So with positive probability, every vertex has some color remaining. Assign the colors arbitrarily for a valid coloring. □

The lower bound on $\mathrm{ch}(K_{n,n})$ turns out to follow from the existence of non-2-colorable k-uniform hypergraph with many edges.

Theorem 1.4.3. If there exists a non-2-colorable k-uniform hypergraph with n edges, then $K_{n,n}$ is not k-choosable.

Proof. Let $H = (V, E)$ be a k-uniform hypergraph $|E| = n$ edges. Label the vertex of $K_{n,n}$ by v_e and w_e as e ranges over E. View V as colors and assign to both v_e and w_e a list of colors given by the k-element set e.

If this $K_{n,n}$ has a proper list coloring with the assigned colors. Let C be the colors used among the n vertices. Then we get a proper 2-coloring of H by setting C black and $V \setminus C$ white. So if H is not 2-colorable, then this $K_{n,n}$ is not k-choosable. □

Recall from Theorem 1.3.3 that there exists a non-2-colorable k-uniform hypergraph with $O(k^2 2^k)$ edges. Thus $\mathrm{ch}(K_{n,n}) > (1 - o(1)) \log_2 n$.

Putting these bounds together:

Corollary 1.4.4. $\mathrm{ch}(K_{n,n}) = (1 + o(1)) \log_2 n$

It turns out that, unlike the chromatic number, the list chromatic number always grows with the average degree. The following result was proved using the method of **hypergraph containers** (a very important modern development in combinatorics) provides the optimal asymptotic dependence (the example of $K_{n,n}$ shows optimality).

Theorem 1.4.5 (Saxton and Thomason 2015)**.** If a graph G has average degree d, then $\mathrm{ch}(G) > (1 + o(1)) \log_2 d$.

They also proved similar results for the list chromatic number of hypergraphs. For graphs, a slightly weaker result, off by a factor of 2, was proved earlier by Alon (2000).