[2] P. Erdős, Problems and results in additive number theory, in *Colloque sur la Théorie des Nombres (CBRM)*, Bruxelles, 1955, 127-137.

[3] P. Erdős, Graph Theory and Probability, *Canad. J. Math.* 11 (1959), 34-38.

[4] P. Erdős, Graph Theory and Probability II., *Canad. J. Math.* 13 (1961), 346-352.

[5] P. Erdős, On circuits and subgraphs of chromatic graphs, *Mathematika* 9(1962), 170-175.

[6] P. Erdős, On a combinatorial problem I., *Nordisk. Mat. Tidskr.* 11 (1963), 5-10.

[7] P. Erdős, On a combinatorial problem II., *Acta. Math. Acad. Sci. Hungar.* 15 (1964), 445-447.

[8] P. Erdős and J. W. Moon, On sets of consistent arcs in a tournament, *Canad. Math. Bull.* 8 (1965), 269-271.

[9] P. Erdős and A. Rényi, On the evolution of random graphs, *Mat. Kutató Int. Közl.* 5 (1960), 17-60.

[10] P. Erdős and J. Spencer, Imbalances in k-colorations, *Networks* 1 (1972), 379-385.

[11] P. Erdős and G. Szekeres, A combinatorial problem in geometry, *Compositio Math.* 2(1935), 463-470.

[12] W. F. de la Vega, On the maximal cardinality of a consistent set of arcs in a random tournament, *J. Combinatorial Theory, Series B* 35 (1983), 328-332.

7 1965: Unrankable Tournaments

Let T be a tournament with players $1, \ldots, n$ (each pair play one game and there are no ties) and σ a ranking of the players, technically a permutation on $\{1, \ldots, n\}$. Call game $\{i, j\}$ a nonupset if i beats j and $\sigma(i) < \sigma(j)$; an upset if i beats j but $\sigma(j) < \sigma(i)$. The fit $f(T, \sigma)$ is the number of nonupsets minus the number of upsets. One might have thought - in preprobabilistic days! - that every tournament T had a ranking σ with a reasonably good fit. With J.W. Moon, Erdős [?] easily destroyed that conjecture.
Theorem. There is a T so that for all σ

$$f(T, \sigma) \leq n^{3/2}(\ln n)^{1/2}$$

Thus, for example, there are tournaments so that under any ranking at least 49% of the games are upsets. Erdős and Moon take the random tournament, for each pair {i,j} one "flips a fair coin" to see who wins the game. For any fixed σ each game is equally likely to be upset or nonupset and the different games are independent. Thus $f(T, \sigma) \sim S_m$, where $m = \binom{n}{2}$ and S_m is the number of heads minus the number of tails in m flips of a fair coin. Large deviation theory gives

$$\Pr[S_m > \alpha] < e^{-\frac{\alpha^2}{2m}}$$

One now uses *very* large deviations. Set $\alpha = n^{3/2}(\ln n)^{1/2}$ so that the above probability is less than $n^{-n} < n!^{-1}$. This supersmall probability is used because there are $n!$ possible σ. Now with positive probability no σ have $f(T, \sigma) > \alpha$. Thus there is a T with no σ having $f(T, \sigma) > \alpha$.

The use of extreme large deviations has become a mainstay of the Probabilistic Method. But I have a more personal reason for concluding with this example. Let $g(n)$ be the least integer so that every tournament T on n players has a ranking σ with $f(T, \sigma) \geq g(n)$. Then $g(n) \leq n^{3/2}(\ln n)^{1/2}$. Erdős and Moon showed $g(n) > cn$, leaving open the asymptotics of $g(n)$. In my doctoral dissertation I showed $g(n) > c_1 n^{3/2}$ and later (but see delaVega [?] for the "book proof") that $g(n) < c_2 n^{3/2}$. Though at the time I was but an ϵ Paul responded with his characteristic openness and soon [?] I had an Erdős number of one. Things haven't been the same since.

References

[1] P. Erdős, Some remarks on the theory of graphs, *Bull. Amer. Math. Soc.* 53 (1947), 292-294.

Proof: Color Ω randomly. Each A_i has probability 2^{1-n} of being monochromatic, the probability some A_i is monochromatic is then at most $m2^{1-n} < 1$ so with positive probability no A_i is monochromatic. Take that coloring.

In 1964 Erdős [?]showed this result was close to best possible.
Theorem. There exists a family $\mathcal{A}$ with $m = cn^2 2^n$ which is not 2-colorable.

Here Erdős turns the original probability argument inside out. Before the sets were fixed and the coloring was random, now, essentially, the coloring is fixed and the sets are random. He sets $\Omega = \{1, \ldots, u\}$ with u a parameter to be optimized later. Let $A_1, \ldots, A_m$ be random n-sets of Ω. Fix a coloring χ with a red points and $b = u - a$ blue points. As A_i is random

$$\Pr[\chi(A_i) \text{ constant}] = \frac{\binom{a}{n} + \binom{b}{n}}{\binom{u}{n}} \leq \frac{2\binom{u/2}{n}}{\binom{u}{n}}$$

The second inequality, which follows from the convexity of $\binom{x}{n}$, indicates that it is the equicolorings that are the most troublesome. As the A_i are independent

$$\Pr[\text{no } A_i \text{ monochromatic}] \leq \left[1 - \frac{2\binom{u/2}{n}}{\binom{u}{n}}\right]^m$$

Now suppose

$$2^u \left[1 - \frac{2\binom{u/2}{n}}{\binom{u}{n}}\right]^m < 1$$

The expected number of χ with no A_i monochromatic is less than one. Therefore there is a choice of $A_1, \ldots, A_m$ for which no such χ exists, i.e., $\mathcal{A}$ is not 2-colorable. Solving, one may take

$$m = \left\lceil \frac{m \ln 2}{-\ln\left[1 - \frac{2\binom{u/2}{n}}{\binom{u}{n}}\right]} \right\rceil$$

Estimating $-\ln(1-\epsilon) \sim \epsilon$ this is roughly $cu\binom{u}{n}/\binom{u/2}{n}$. This leads to an interesting calculation problem (as do many problems involving the Probabilistic Method!) – find u so as to m. The answer turns out to be $u \sim n^2/2$ at which value $m \sim (e \ln 2) n^2 2^{n-2}$.

Erdős has defined $m(n)$ as the least m for which there is a family of n-sets which cannot be 2-colored. His results give $\Omega(2^n) = m(n) = O(n^2 2^n)$. Beck has improved the lower bound to $\Omega(n^{1/3} 2^n)$ but the actual asymptotics of $m(n)$ remain elusive.

conjecture that if every, say, $n/(\ln n)$ vertices could be 3-colored then G could be 4-colored. This theorem disproves that conjecture.

We examine the random graph $G \sim G(n,p)$ with $p = c/n$. As in the 1957 paper

$$\Pr[\alpha(G) \geq x] < \binom{n}{x}(1-p)^{\binom{x}{2}} < \left[(ne/x)e^{-p(x-1)/2}\right]^x$$

When c is large and, say, $x = 10n(\ln c)/c$, the bracketed quantity is less than one so the entire quantity is $o(1)$ and a.s. $\alpha(G) \leq x$ and so $\chi(G) \geq c/(10 \ln c)$. Given k Erdős may now simply select c so that, with $p = c/n$, $\chi(G) > k$ a.s.

Now for the local coloring. If some set of $\leq \epsilon n$ vertices cannot be 3-colored then there is a minimal such set S with, say, $|S| = i \leq \epsilon n$. In the restriction $G|_S$ every vertex v must have degree at least 3 - otherwise one could 3-color $S - \{v\}$ by minimality and then color v differently from its neighbors. Thus $G|_S$ has at least $3i/2$ edges. The probability of G having such an S is bounded by

$$\sum_{i=4}^{\epsilon n} \binom{n}{i}\binom{\binom{i}{2}}{3i/2} p^{3i/2} \leq \sum_{i=4}^{\epsilon n} \left[\frac{ne}{i}\left(\frac{ei}{3}\right)^{3/2}\left(\frac{c}{n}\right)^{3/2}\right]^i$$

employing the useful inequality $\binom{a}{b} \leq \left(\frac{ea}{b}\right)^b$. Picking $\epsilon = \epsilon(c)$ small the bracketed term is always less than one, the entire sum is $o(1)$, a.s. no such S exists, and a.s. every ϵn vertices may be 3-colored.

Erdős's monumental study with Alfred Rényi "On the Evolution of Random Graphs" [?] had been completed only a few years before. The behavior of the basic graph functions such as chromatic and clique number were fairly well understood throughout the evolution. The argument for local coloring required a "new idea" but the basic framework was already in place.

6 1963/4: Coloring Hypergraphs

Let $A_1, \ldots, A_m$ be n-sets in an arbitrary universe Ω. The family $\mathcal{A} = \{\mathcal{A}_\infty, \ldots, \mathcal{A}_{\Updownarrow}\}$ is 2-colorable (Erdős used the term "Property B") if there is a 2-coloring of the underlying points Ω so that no set A_i is monochromatic. In 1963 Erdős gave perhaps the quickest demonstration of the Probabilistic Method.

Theorem[?]: If $m < 2^{n-1}$ then $\mathcal{A}$ is 2-colorable.

Each $d(t)$ has Binomial Distribution $B(x,p)$ and so expectation $xp = \Theta(\ln n)$ so that one can get fairly easily $E[Z] = \Theta(n \ln^2 n)$. Note this is the same order as x^2. It is definitely not easy to show that for appropriate A, c (Erdős takes $c = A^{-1/2}$ and A large) $Z < \frac{1}{2}\binom{x}{2}$ with high probability. The requirement "with high probability" is quite severe. But note, at least, that this is a pure probability statement. Lets accept it and move on. Call a pair $\{i,j\} \subset S$ soiled if it lies in a triangle with third vertex outside of S. At most Z pairs are soiled so with high probability at least $\frac{1}{2}\binom{x}{2}$ pairs are unsoiled. Now we expose the edges of G inside S. If any of the unsoiled pairs are in G then G is good and so the failure probability at most

$$(1-p)^{\frac{1}{2}\binom{x}{2}} < e^{-\Omega(px^2)} = o\left(\binom{n}{x}^{-1}\right)$$

and so G is good with high probability.

Sound complicated. Well, it is complicated and it is simultaneously a powerful application of the Probabilistic Method and a technical *tour de force*. The story has a coda: the Lovász Local Lemma, developed in the mid-1970s, gave a new sieve method for showing that a set of bad events could simultaneously not hold. This author applied it to the random graph $G(n,p)$ with $p = cn^{-1/2}$ with the bad events being the existence of the various potential triangles and the independence of the various x-sets. The conditions of the Local Lemma made for some calculations but it was relatively straightforward to duplicate this result. Still, the ideas behind this proof, the subtle extension of the Deletion Method notion, are too beautiful to be forgotten.

5 1962: No Local Coloring

With his 1957 paper previously discussed Erdős had already shown that chromatic number cannot be considered simply a local phenomenon. With this result he puts the nail in the coffin.

Theorem[**?**]. For any $k \geq 3$ there is an $\epsilon > 0$ so that the following holds for all sufficiently large n: There exists a graph G on n vertices which cannot be k-colored and yet the restriction of G to any ϵn vertex subgraph can be 3-colored.

Often probabilistic theorem are best understood as negative results, as counterexamples to natural conjectures. A priori, for example, one might

ever since. We have already spoken of his 1947 paper on $R(k,k)$. In his 1961 paper Erdős [?] proves

$$R(3,k) > c\frac{k^2}{\ln^2 k}$$

The upper bound $R(3,k) = O(k^2)$ was already apparent from the original Szekeres proof so the gap was relatively small. Only in 1994 was the correct order $R(3,k) = \Theta(\frac{k^2}{\ln k})$ finally shown.

Erdős shows that there is a graph on n vertices with no triangle and no independent set of size x where $x = \lceil An^{1/2} \ln n \rceil$, and A is a large absolute constant. This gives $R(3,x) > n$ from which the original statement follows easily. We'll ignore A in our informal discussion. He takes a random graph $G(n,p)$ with $p = cn^{-1/2}$. The probability that some x-set is independent is at most

$$\binom{n}{x}(1-p)^{x(x-1)/2} < \left[ne^{-p(x-1)/2}\right]^x$$

which is *very* small. Unfortunately this G will have *lots* ($\Theta(n^{3/2})$) of triangles. One needs to remove an edge from each triangle without making any of the x-sets independent.

The Erdős method may be thought of algorithmically. Order the edges $e_1, \ldots, e_m$ of $G \sim G(n,p)$ arbitrarily. Consider them sequentially and reject e_i if it would make a triangle with the edges previously accepted, otherwise accept e_i. The graph G^- so created is certainly trianglefree. What about the sets of x vertices. Call a set S of x vertices *good* (in G, not G^-) if it contains an edge e which cannot be extended to a triangle with third vertex outside of S. Suppose S is good and let e be such an edge. Then S cannot be independent in G^-. If e is accepted we're clearly OK. The only way e could be rejected is if e is part of a triangle e, e^1, e^2 where the other edges have already been accepted. But then e_1, e_2 must (as S is good) lie in S and again S is not independent.

Call S bad if it isn't good. Erdős shows that almost always there are no bad S. Lets say something occurs with high probability if its failure probability is $o(\binom{n}{x})^{-1})$. It suffices to show that a given $S = \{1, \ldots, x\}$ is good with high probability. This is the core of the argument. We expose (to use modern terminology) G in two phases. First we examined the pairs $\{s,t\}$ with $s \in S, t \notin S$. For each $t \notin S$ let $d(t)$ be the number of edges to S. Set

$$Z = \sum_{t \notin S} \binom{d(t)}{2}$$

u was then

$$\sum_{i=1}^{n} \binom{\binom{u}{2}}{l} \binom{\binom{n}{2} - \binom{u}{2}}{m-l} < (n+1) \binom{\binom{u}{2}}{n} \binom{\binom{n}{2} - \binom{u}{2}}{m} <$$

$$< u^{2n} \binom{\binom{n}{2} - \binom{m}{2}}{m} < \binom{\binom{n}{2}}{m} u^{2n} \left(1 - \frac{\binom{u}{2}}{\binom{n}{2}}\right)^{m} <$$

$$< \binom{\binom{n}{2}}{m} u^{2n} \left(1 - \frac{u^2}{n^2}\right)^{m} < \binom{\binom{n}{2}}{m} u^{2m} e^{-mu^2/n^2}$$

Now the number of possible choices for the u points is

$$\binom{n}{u} < n^u < u^n$$

and so the number of graphs without the desired property is

$$\binom{\binom{n}{2}}{m} u^{3n} e^{-n^{1+\epsilon-2\eta}} = o\left(\binom{\binom{n}{2}}{m}\right)$$

as desired. Today, with large deviation results assumed beforehand, the proof can be given in one relatively leisurely page.

Many consider this one of the most pleasing applications of the Probabilistic Method as the result seems not to call for probability in the slightest and earlier attempts had been entirely constructive. The further use of large deviations and the introduction of the Deletion Method greatly advanced the Probabilistic Method. And, most important, the theorem gives an important truth about graphs. In a rough sense the truth is a negative one: chromatic number cannot be determined by local considerations only.

4 1961: Ramsey $R(3,k)$

Ramsey Theory was one of Paul Erdős's earliest interests. The involvement can be dated back to the winter of 1932/33. Working on a problem of Esther Klein, Erdős proved his famous result that in every sequence of n^2+1 real numbers there is a monotone subsequence of length $n+1$. At the same time, and for the same problem, George Szekeres rediscovered Ramsey's Theorem. Both arguments appeared in their 1935 joint paper[**?**]. Bounds on the various Ramsey functions, particularly the function $R(l,k)$, have fascinated Erdős

number and arbitrarily high girth – i.e. no small cycles. To many graph theorists this seemed almost paradoxical. A graph with high girth would locally look like a tree and trees can easily be colored with two colors. What reason could force such a graph to have high chromatic number? As we'll see, there is a global reason: $\chi(G) \geq n/\alpha(G)$. To show $\chi(G)$ is large one "only" has to show the nonexistence of large independent sets.

Erdős [**?**] proved the existence of such graphs by probabilistic means. Fix l, k, a graph is wanted with $\chi(G) > l$ and no cycles of size $\leq k$. Fix $\epsilon < \frac{1}{k}$, set $p = n^{\epsilon - 1}$ and consider $G \sim G(n,p)$ as $n \to \infty$. There are small cycles, the expected number of cycles of size $\leq k$ is

$$\sum_{i=3}^{k} \frac{(n)_i}{2i} p^i = \sum_{i=3}^{k} O((np)^i) = o(n)$$

as $k\epsilon < 1$. So almost surely the number of edges in small cycles is $o(n)$. Also fix positive $\eta < \epsilon/2$. Set $\lfloor u = n^{1-\eta} \rfloor$. A set of u vertices will contain, on average, $\mu \sim u^2 p/2 = \Omega(n^\alpha)$ edges where $\alpha = 1 + \epsilon - 2\eta > 1$. Further, the number of such edges is given by a Binomial Distribution. Applying large deviation results, the probability of the u points having fewer than half their expected number of edges is $e^{-c\mu}$. As $\alpha > 1$ this is smaller than exponential, so $o(2^{-n})$ so that almost surely *every* u points have at least $\mu/2$ edges. We need only that $\mu/2 > n$.

Now Erdős introduces what is now called the Deletion Method. This random graph G almost surely has only $o(n)$ edges in small cycles and every u vertices have at least n edges. Take a specific graph G with these properties. Delete all the edges in small cycles giving a graph G^-. Then certainly G^- has no small cycles. As fewer than n edges have been deleted every u vertices of G^-, which had more than n edges in G, still have an edge. Thus the independence number $\alpha(G^-) \leq u$. But

$$\chi(G^-) \geq \frac{n}{\alpha(G^-)} \geq \frac{n}{u} \sim n^\eta$$

As n can be arbitrarily large one can now make $\chi(G^-) \geq k$, completing the proof.

The use of counting arguments became a typographical nightmare. Erdős considered all graphs with precisely m edges where $m = \lfloor n^{1+\epsilon} \rfloor$. He needed that almost all of them had the property that every u vertices (u as above) had more than n. The number of graphs failing that for a given set of size

events being mutually independent over integers x, setting

$$p_x = K\left(\frac{\ln x}{x}\right)^{1/2}$$

where K is a large absolute constant. (For the finitely many x for which this is greater than one simply place $x \in S$.) Now $f(n)$ becomes a random variable. For each $x < y$ with $x + y = n$ let I_{xy} be the indicator random variable for $x, y \in S$. Then we may express $f(n) = \sum I_{xy}$. From Linearity of Expectation

$$E[f(n)] = \sum E[I_{xy}] = \sum p_x p_y \sim K' \ln n$$

by a straightforward calculation.

Lets write $\mu = \mu(n) = E[f(n)]$. The key ingredient is now a *large deviation* result. One shows, say, that

$$\Pr[f(n) < \frac{1}{2}\mu] < e^{-c\mu}$$

$$\Pr[f(n) > 2\mu] < e^{-c\mu}$$

where c is a positive absolute constant, not dependent on n, K or μ. This makes intuitive sense: as $f(n)$ is the sum of mutually independent rare indicator random variables it should be roughly a Poisson distribution and such large deviation bounds hold for the Poisson. Now pick K so large that K' is so large that $c\mu > 2\ln n$. Call n a failure if either $f(n) > 2\mu$ or $f(n) < \mu/2$. Each n has probability less than $2n^{-2}$ failure probability. By the Borel-Cantelli Lemma (as $\sum n^{-2}$ converges) almost surely there are only a finite number of failures and so almost surely this random S has the desired properties.

While the original Erdős proof was couched in different, counting, language the use of large deviation bounds can be clearly seen and, on this count alone, this paper marks a notable advance in the Probabilistic Method.

3 1959: High Girth, High Chromatic Number

Tutte was the first to show the existence of graphs with arbitrarily high chromatic number and no triangles, this was extended by Kelly to arbitrarily high chromatic number and no cycles of sizes three, four or five. A natural question occured – could graphs be found with arbitrarily high chromatic

Erdős used a counting argument above, in the more modern language we would speak of the random graph $G \sim G(n,p)$ with $p = \frac{1}{2}$. The probability that G contains a complete graph of order k is less than

$$\binom{N}{k} 2^{-k(k-1)/2} < \frac{N^k}{k!} 2^{-k(k-1)/2} < \frac{1}{2}$$

(calculations as in the original paper) and so the probability that G or G' contains a complete graph is less than one so that with positive probability G doesn't have this property and therefore there exists a G as desired. Erdős has related that after lecturing on his result the probabilist J. Doob remarked "Well, thats very nice but it really is a counting argument." For this result the proofs are nearly identical, the probabilistic proof having the minor advantage of avoiding the annoying $2^{N(N-1)/2}$ factors. Erdős writes interchangably in the two styles. As the methodology has progressed the probabilistic ideas have become more subtle and today it is quite rare to see a paper written in the counting style. We'll take the liberty of translating Erdős's later results into the more modern style.

The gap between $2^{k/2}$ and 4^k for $R(k,k)$ remains one of the most vexing problems in Ramsey Theory and in the Probabilistic Method. All improvements since this 1947 paper have been only to smaller order terms so that even today $\lim R(k,k)^{1/k}$ could be anywhere from $\sqrt{2}$ to 4, inclusive. Even the existence of the limit has not been shown!

2 1955: Sidon Conjecture

Let S be a set of positive integers. Define $f(n) = f_S(n)$ as the number of representations $n = x + y$ where x, y are distinct elements of S. We call S a *basis* if $f(n) > 0$ for all sufficiently large n. Sidon, in the early 1930s, asked if there existed "thin" bases, in particular he asked if for all positive ϵ there existed a basis with $f(n) = O(n^\epsilon)$. Erdős heard of this problem at that time and relates that he told Sidon that he thought he could get a solution in "a few days". It took somewhat longer. In 1941 Erdős and Turán made the stronger conjecture that there exists a basis with $f(n)$ bounded from above by an absolute constant – a conjecture that remains open today. In 1955 Erdős [**?**] resolved the Sidon conjecture with the following stronger result. Theorem: There exists S with $f(n) = \Theta(\ln n)$.

The proof is probabilistic. Define a random set by $\Pr[x \in S] = p_x$, the

THE ERDŐS EXISTENCE ARGUMENT

Joel Spencer

The Probabilistic Method is now a standard tool in the combinatorial toolbox but such was not always the case. The development of this methodology was for many years nearly entirely due to one man: Paul Erdős. Here we reexamine some of his critical early papers. We begin, as all with knowledge of the field would expect, with the 1947 paper [**?**] giving a lower bound on the Ramsey function $R(k,k)$. There is then a curious gap (certainly *not* reflected in Erdős's overall mathematical publications) and our remaining papers all were published in a single ten year span from 1955 to 1965.

1 1947: Ramsey $R(k,k)$

Let us repeat the key paragraph nearly verbatim. Erdős defines $R(k,l)$ as the least integer so that given any graph G of $n \geq R(k,l)$ vertices then either G contains a complete graph of order k or the complement G' contains a complete graph of order l.
Theorem. Let $k \geq 3$ Then

$$2^{k/2} < R(k,k) \leq \binom{2k-2}{k-1} < 4^{k-1}$$

Proof. The second inequality was proved by Szekeres thus we only consider the first one. Let $N \leq 2^{n/2}$. Clearly the number of graphs of N vertices equals $2^{N(N-1)/2}$. (We consider the vertices of the graph as distinguishable.) The number of different graphs containing a complete graph of order k is less than

$$\binom{n}{k}\frac{2^{N(N-1)/2}}{2^{k(k-1)/2}} < \frac{N^k}{k!}\frac{2^{N(N-1)/2}}{2^{k(k-1)/2}} < \frac{2^{N(N-1)/2}}{2} \quad (*)$$

since by a simple calculation for $N \leq 2^{k/2}$ and $k \geq 3$

$$2N^k < k!2^{k(k-1)/2}$$

But it follows immediately from (*) that there exists a graph such that neither it nor its complementary graph contains a complete subgraph of order k, which completes the proof of the Theorem.