P. Erdős and L. Lovász (1975), Problems and results on 3-chromatic hypergraphs and some related questions, in: Infinite and Finite Sets (A. Hajnal et al., eds.), North-Holland, Amsterdam, pp. 609-628.

P. Erdős and A. Rényi (1960), On the Evolution of Random Graphs, Mat Kutató Int. Közl. *5*, 17-60

P. Erdős and J. Spencer (1991), Lopsided Lovász Local Lemma and Latin Transversals, Disc. Appl. Math. *30*, 151-154

P. Erdős and G. Szekeres (1935), A Combinatorial Problem in Geometry, Compositio Math. *2*, 463-470

P. Erdős and P. Tetali (1990), Representations of integers as the Sum of k Terms, Random Structures and Algorithms *1*, 245-261.

R. Fagin (1976), Probabilities in Finite Models, J. Symbolic Logic *41*, 50-58

Y.V. Glebskii, D.I. Kogan, M.I. Liagonkii and V.A. Talanov (1969), Range and degree of realizability of formulas the restricted predicate calculus, Cybernetics *5*, 142-154 (Russian original: Kibernetica *5*, 1969, 17-27)

L. Harper (1966), Optimal numberings and isoperimetric problems on graphs, J. Combinatorial Theory *1*, 385-394.

S. Janson, T. Łuczak, A. Rucinski (1990), An Exponential Bound for the Probability of Nonexistence of a Specified Subgraph in a Random Graph, *in* Random Graphs '87 (M. Karonski, J. Jaworski, A. Rucinski, eds.), John Wiley, 73-87

R.M. Karp (1990), The transitive closure of a Random Digraph, Random Structures and Algorithms *1*, 73-94

T. Łuczak (1990), Component Behavior near the Critical Point of the Random Graph Process, Random Structures and Algorithms *1*, 287-310.

P. Mani-Levitska and J. Pach (1988), Decomposition problems for multiple coverings with unit balls, to appear.

D.W. Matula (1976), The Largest Clique Size in a Random Graph, Tech. Rep. Dept. Comp. Sci. Southern Methodist University, Dallas

B. Maurey (1979), Construction de suites symétriques, Compt. Rend. Acad. Sci. Paris *288*, 679-681.

V. D. Milman and G. Schechtman (1986), Asymptotic Theory of Finite Dimensional Normed Spaces, Lecture Notes in Mathematics 1200, Springer Verlag, Berlin and New York.

E. Shamir and J. Spencer (1987), Sharp concentration of the chromatic number in random graphs $G_{n,p}$, Combinatorica *7*, 121-130

J. Shearer (1985), On a problem of Spencer, Combinatorica *5*, 241-245.

S. Shelah and J. Spencer (1988), Zero-One Laws for Sparse Random Graphs, J. Amer. Math. Soc. *1*, 97-115

J. Spencer (1977), Asymptotic Lower Bounds for Ramsey Functions, Disc. Math. *20*, 69-76

J. Spencer (1990a), Threshold Functions for Extension Statements, J. Comb. Th. (Ser A) *53*, 286-305

J. Spencer (1990b), Counting Extension, J. Combinatorial Th. (Ser A) *55*, 247-255.

P. Turán (1934), On a theorem of Hardy and Ramanujan, J. London Math Soc. *9*, 274-276

E.M. Wright (1977), The number of connected sparsely edged graphs, Journal of Graph Theory *1*, 317-330.

entries of A. Then A has a Latin Transversal.

Let π be a random permutation of $\{1, 2 \leq n\}$, chosen according to a uniform distribution among all possible $n!$ permutations. Denote by T the set of all ordered fourtuples (i, j, i', j') satisfying $i < i', j \neq j'$ and $a_{ij} = a_{i'j'}$. For each $(i, j, i', j') \in T$, let $A_{iji'j'}$ denote the event that $\pi(i) = j$ and $\pi(i') = j'$. The existence of a Latin transversal is equivalent to the statement that with positive probability none of these events hold. Let us define a symmetric digraph, (i.e., a graph) G on the vertex set T by making (i, j, i', j') adjacent to (p, q, p', q') if and only if $\{i, i'\} \cap \{p, p'\} \neq \emptyset$ or $\{j, j'\} \cap \{q, q'\} \neq \emptyset$. Thus, these two fourtuples are not adjacent iff the four cells $(i, j), (i', j'), (p, q)$ and (p', q') occupy four distinct rows and columns of A. The maximum degree of G is less than $4nk$; indeed, for a given $(i, j, i', j') \in T$ there are $4n$ choices of (p, q) with either $p \in \{i, i'\}$ or $q \in \{j, j'\}$, and for each of these choices of (p, q) there are less than k choices for $(p', q') \neq (p, q)$ with $a_{pq} = a_{p'q'}$. Since $e \cdot 4nk \cdot \frac{1}{n(n-1)} \leq 1$, the desired result follows from the above mentioned strengthening of the symmetric version of the Lovász Local Lemma, if we can show that

$$\Pr(A_{iji'j'} | \bigwedge_S \overline{A_{pqp'q'}}) \leq 1/n(n-1)$$

for any $(i, j, i', j') \in T$ and any set S of members of T which are nonadjacent in G to (i, j, i', j'). By symmetry, we may assume that $i = j = 1, i' = j' = 2$ and that hence none of the p's nor q's are either 1 or 2. Let us call a permutation π *good* if it satisfies $\bigwedge_S \overline{A_{pqp'q'}}$, and let S_{ij} denote the set of all good permutations π satisfying $\pi(1) = i$ and $\pi(2) = j$. We claim that $|S_{12}| \leq |S_{ij}|$ for alll $i \neq j$. Indeed, suppose first that $i, j > 2$. For each good $\pi \in S_{12}$ define a permutation π^* as follows. Suppose $\pi(x) = i$, $\pi(y) = j$. Then define $\pi^*(1) = i, \pi^*(2) = j, \pi^*(x) = 1, \pi^*(y) = 2$ and $\pi^*(t) = \pi(t)$ for all $t \neq 1, 2, x, y$. One can easily check that π^* is good, since the cells $(1, i), (2, j), (x, 1), (y, 2)$ are not part of any $(p, q, p', q') \in S$. Thus $\pi^* \in S_{ij}$, and since the mapping $\pi \to \pi^*$ is injective $|S_{12}| \leq |S_{ij}|$, as claimed. Similarly one can define injective mappings showing that $|S_{12}| \leq |S_{ij}|$ even when $\{i, j\} \cap \{1, 2\} \neq \emptyset$. It follows that $\Pr(A_{1122} \bigwedge \bigwedge_S \overline{A_{pqp'q'}}) \leq Pr(A_{1i2j} \bigwedge \bigwedge_S \overline{A_{pqp'q'}})$ for all $i \neq j$ and hence that $Pr(A_{1122} | \bigwedge \bigwedge_S \overline{A_{pqp'q'}}) \leq 1/n(n-1)$. By symmetry, this implies (6.1) and completes the proof.□

References

N. Alon and Z. Bregman (1988), Every 8-uniform 8-regular hypergraph is 2-colorable, Graphs and Combinatorics *4*, 303-305.

B. Bollobás (1985), *Random Graphs*, Academic Press.

B. Bollobás (1988), *The chromatic number of random graphs*, Combinatorica *8*, 49-55.

B. Bollobás and P. Erdős (1976), Cliques in Random Graphs, Math. Proc. Camb. Phil. Soc. *80*, 419-427

R. B. Boppana and J. H. Spencer (1989), A useful elementary correlation inequality, J. Combinatorial Theory Ser. A, *50*, 305-307.

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

P. Erdős (1956), Problems and results in additive number theory, Colloque sur le Théorie des Nombres (CBRM, Bruselles) 127-137

P. Erdős (1959), Graph theory and probability, Canad. J. Math. *11* (1959), 34-38.

P. Erdős and M. Kac (1940), The Gaussian Law of Errors in the Theory of Additive Number Theoretic Functions, Amer. J. Math. *62*, 738-742

two pairwise disjoint families F_1 and F_2, each being a covering of R^3. Mani and Pach [1988] constructed, for any integer $k \geq 1$, a non-decomposable k-fold covering of R^3 by open unit balls. On the other hand they proved that any k-fold covering of R^3 in which no point is covered by more than $c2^{k/3}$ balls is decomposable. This reveals a somewhat surprising phenomenon that it is more difficult to decompose coverings that cover some of the points of R^3 too often, than to decompose coverings that cover every point about the same number of times. The exact statement of the Mani-Pach Theorem is the following.

Theorem 4.1. Let $F = \{B_i\}_{i \in I}$ be a k-fold covering of the 3 dimensional Euclidean space by open unit balls. Suppose, further, than no point of R^3 is contained in more than t members of F. If

$$e \cdot t^3 2^{18}/2^{k-1} \leq 1$$

then F is decomposable.

Let $\{C_j\}_{j \in J}$ be the connected components of the set obtained from R^3 by deleting all the boundaries of the balls B_i in F. Let $H = (V(H), E(H))$ be the (infinite) hypergraph defined as follows; the set of vertices of H, $V(H)$ is simply $F = \{B_i\}_{i \in I}$. The set of edges of H is $E(H) = \{E_j\}_{j \in J}$, where $E_j = \{B_i : i \in I \text{ and } C_j \subseteq B_i\}$. Since F is a k-fold covering, each edge E_j of H contains at least k vertices. We claim that each edge of H intersects less than $t^3 2^{18}$ other edges of H. To prove this claim, fix an edge E_l, corresponding to the connected component C_l, where $l \in J$. Let E_j be an arbitrary edge of H, corresponding to the component C_j, that intersects E_l. Then there is a ball B_i containing both C_l and C_j. Therefore, any ball that contains C_j intersects B_i. It follows that all the unit balls that contain or touch a C_j, for some j that satisfies $E_j \cap E_l \neq \emptyset$ are contained in a ball B of radius 4. As no point of this ball is covered more than t times we conclude, by a simple volume argument, that the total number of these unit balls is at most $t \cdot 4^3 = t \cdot 2^6$. It is not too difficult to check that m balls in R^3 cut R^3 into less than m^3 connected components, and since each of the above C_j is such a component we have $|\{j : E_j \cap E_l \neq \emptyset\}| < (t \cdot 2^6)^3 = t^3 2^{18}$, as claimed.

Consider, now, any finite subhypergraph L of H. Each edge of L has at least k vertices, and it intersects at most $d < t^3 2^{18}$ other edges of L. Since, by assumption, $e(d+1) \leq 2^{k-1}$, Theorem 2.1 (which is a simple corollary of the local lemma), implies that L is 2-colorable. This means that one can color the vertices of L blue and red so that no edge of L is monochromatic. Since this holds for any finite L, a compactness argument, analogous to the one used in the proof of Theorem 2.2, shows that H is 2-colorable. Given a 2-coloring of H with no monochromatic edges, we simply let F_1 be the set of all blue balls, and F_2 be the set of all red ones. Clearly, each F_i is a covering of R^3, completing the proof of the theorem.$\square$

It is worth noting that Theorem 4.1 can be easily generalized to higher dimensions. We omit the detailed statement of this generalization.

5 Latin Transversals

Following the proof of the Lovász Local Lemma we noted that the mutual independency assumption in this lemma can be replaced by the weaker assumption that the conditional probability of each event, given the mutual non-occurance of an arbitrary set of events, each nonadjacent to it in the dependency digraph, is sufficiently small. In this section we describe an application, from Erdős-Spencer [1991], of this modified version of the lemma. Let $A = (a_{ij})$ be an n of n matrix with, say, integer entries. A permutation π is called a *Latin transversal* (of A) if the entries $a_{i\pi(i)}$ $(1 \leq i \leq n)$ are all distinct.

Theorem 6.1. Suppose $k \leq (n-1)/(4e)$ and suppose that no integer appears in more than k

It is obvious that each event A_S is mutually independent of all the events A_T, but those which satisfy $|S \cap T| \geq 2$, since this is the only case in which the corresponding complete graphs share an edge. We can therefore apply Corollary 1.2 with $p = 2^{1-\binom{k}{2}}$ and $d = \binom{k}{2}\binom{n}{k-2}$ to conclude;

Proposition 3.1. If $e\left(\binom{k}{2}\binom{n}{k-2}+1\right) \cdot 2^{1-\binom{k}{2}} < 1$ then $R(k,k) > n$.

A short computation shows that this gives $R(k,k) > \frac{\sqrt{2}}{e}(1+o(1))k2^{k/2}$, only a factor 2 improvement on the bound obtained by the straightforward probabilistic method. Although this minor improvement is somewhat disappointing it is certainly not surprising; the Local Lemma is most powerful when the dependencies between events are rare, and this is not the case here. Indeed, there is a total number of $K = \binom{n}{k}$ events considered, and the maximum outdegree d in the dependency digraph is roughly $\binom{k}{2}\binom{n}{k-2}$. For large k and much larger n (which is the case of interest for us) we have $d > K^{1-O(1/k)}$, i.e., quite a lot of dependencies. On the other hand, if we consider small sets S, e.g., sets of size 3, we observe that out of the total $K = \binom{n}{3}$ of them each shares an edge with only $3(n-3) \approx K^{1/3}$. This suggests that the Lovász Local Lemma may be much more significant in improving the off-diagonal Ramsey numbers $R(k,l)$, especially if one of the parameters, say l, is small. Let us consider, for example, following Spencer (1977), the Ramsey number $R(k,3)$. Here, of course, we have to apply the nonsymmetric form of the Lovász Local Lemma. Let us 2-color the edges of K_n randomly and independently, where each edge is colored blue with probability p. For each set of 3 vertices T, let A_T be the event that the triangle on T is blue. Similarly, for each set of k vertices S, let B_S be the event that the complete graph on S is red. Clearly $Pr(A_T) = p^3$ and $Pr(B_S) = (1-p)^{\binom{k}{2}}$. Construct a dependency digraph for the events A_T and B_S by joining two vertices by edges (in both directions) iff the corresponding complete graphs share an edge. Clearly, each A_T-node of the dependency graph is adjacent to $3(n-3) < 3n$ $A_{T'}$-nodes and to at most $\binom{n}{k}$ $B_{S'}$-nodes. Similarly, each B_S-node is adjacent to $\binom{k}{2}(n-k) < k^2n/2$ A_T nodes and to at most $\binom{n}{k}$ $B_{S'}$-nodes. It follows from the general case of the Lovász Local Lemma that if we can find a $0 < p < 1$ and two real numbers $0 \leq x < 1$ and $0 \leq y < 1$ such that

$$p^3 \leq x(1-x)^{3n}(1-y)^{\binom{n}{k}}$$

and

$$(1-p)^{\binom{k}{2}} \leq y(1-x)^{k^2n/2}(1-y)^{\binom{n}{k}}$$

then $R(k,3) > n$.

Our objective is to find the largest possible $k = k(n)$ for which there is such a choice of p, x and y. An elementary computation (if you have a spare weekend!) shows that the best choice is when $p = c_1 n^{-1/2}$, $k = c_2 n^{1/2} \log n$, $x = c_3/n^{3/2}$ and $y = c_4 e^{-n^{1/2}\log^2 n}$. This gives that $R(k,3) > c_5 k^2/\log^2 k$. A similar argument gives that $R(k,4) > k^{5/2+o(1)}$. In both cases the amount of computation required is considerable. However, the hard work does pay; the bound $R(k,3) > c_5 k^2/\log^2 k$ matches a lower bound of Erdős proved in 1961 by a highly complicated probabilistic argument. The bound above for $R(k,4)$ is better than any bound for $R(k,4)$ known to be proven without the Local Lemma.

4 A geometric result

A family of open unit balls F in the 3-dimensional Euclidean space R^3 is called a *k-fold covering* of R^3 if any point $x \in R^3$ belongs to at least k balls. In particular, a 1-fold covering is simply called a *covering*. A k-fold covering F is called *decomposable* if there is a partition of F into

is incident with k edges (including f), it follows that f intersects at most $d = k(k-1)$ other edges. The desired result follows, since $e(k(k-1)+1) < 2^{k-1}$ for each $k \geq 9$. This special case has a different proof (see [Alon-Bregman (1988)]), which works for each $k \geq 8$. It is plausible to conjecture that in fact for each $k \geq 4$ each k-uniform k-regular hypergraph is has Property B. The next result we consider, which appeared in the original paper of Erdős and Lovász, deals with k-colorings of the real numbers. For a k-coloring $c : R \to \{1, 2 \ldots k\}$ of the real numbers by the k colors $1, 2 \ldots k$, and for a subset $T \subset R$, we say that T is *multicolored* (with respect to c) if $c(T) = \{1, 2 \ldots k\}$, i.e., if T contains elements of all colors.
Theorem 2.2. Let m and k be two positive integers satisfying

$$e\left(m(m-1)+1\right)k\left(1-\frac{1}{k}\right)^m \leq 1$$

Then, for any set S of m real numbers there is a k-coloring so that each translation $x + S$ (for $x \in R$) is multicolored.

Notice that the condition holds whenever $m > (3 + o(1))k \log k$. There is no known proof of existence of any $m = m(k)$ with this property without using the local lemma.

We first fix a *finite* subset $X \subset R$ and show the existence of a k-coloring so that each translation $x + S$ (for $x \in X$) is multicolored. This is an easy consequence of the Lovász Local Lemma. Indeed, put $Y = \bigcup_{x \in X}(x + S)$ and let $c : Y \to \{1, 2 \ldots k\}$ be a random k-coloring of Y obtained by choosing, for each $y \in Y$, randomly and independently, $c(y) \in \{1, 2 \ldots, k\}$ according to a uniform distribution on $\{1, 2 \ldots k\}$. For each $x \in X$, let A_x be the event that $x + S$ is not multicolored (with respect to c). Clearly $Pr(A_x) \leq k\left(1 - \frac{1}{k}\right)^m$. Moreover, each event A_x is mutually independent of all the other events $A_{x'}$ but those for which $(x + S) \cap (x' + S) \neq \emptyset$. As there are at most $m(m-1)$ such events the desired result follows from Corollary 1.2.

We can now prove the existence of a coloring of the set of all reals with the desired properties, by a standard compactness argument. Since the discrete space with k points is (trivially) compact, Tychanov's Theorem (which is equivalent to the axiom of choice) implies that an arbitrary product of such spaces is compact. In particular, the space of all functions from the reals to $\{1, 2 \ldots k\}$, with the usual product topology, is compact. In this space for every fixed $x \in R$, the set C_x of all colorings c, such that $x + S$ is multicolored is closed. (In fact, it is both open and closed, since a basis to the open sets is the set of all colorings whose values are prescribed in a finite number of places). As we proved above, the intersection of any finite number of sets C_x is nonempty. It thus follows, by compactness, that the intersection of all sets C_x is nonempty. Any coloring in this intersection has the properties in the conclusion of Theorem 2.2.□

Note that it is impossible, in general, to apply the Lovász Local Lemma to an infinite number of events and conclude that in some point of the probability space none of them holds. In fact, there are trivial examples of countably many mutually independent events A_i, satisfying $Pr(A_i) = 1/2$ and $\bigwedge_{i \geq 1} \overline{A_i} = \emptyset$. Thus the compactness argument is essential in the above proof.

3 Lower bounds for Ramsey numbers

The deriviation of lower bounds for Ramsey numbers by Erdős in 1947 was one of the first applications of the probabilistic method. The Lovász Local Lemma provides a simple way of improving these bounds. Let us obtain, first, a lower bound for the diagonal Ramsey number $R(k, k)$. Consider a random 2-coloring of the edges of K_n. For each set S of k vertices of K_n, let A_S be the event that the complete graph on S is monochromatic. Clearly $\Pr(A_S) = 2^{1-\binom{k}{2}}$.

$$\cdot \left(1 - \Pr\left(A_{j_2} | \overline{A_{j_1}} \wedge B\right)\right) \cdots \left(1 - \Pr\left(A_{j_r} | \overline{A_{j_1}} \wedge \ldots \wedge \overline{A_{j_{r-1}}} \wedge B\right)\right)$$
$$\geq (1 - x_{j_1}) \cdots (1 - x_{j_r}) \geq \prod_{(i,j) \in E} (1 - x_j)$$

Substituting we conclude that $\Pr\left(A_i | \bigwedge_{j \in S} \overline{A_j}\right) \leq x_i$, completing the proof of the induction.

The assertion of Lemma 1.1 now follows easily, as

$$\Pr\left(\bigwedge_{i=1}^{n} \overline{A_i}\right) = (1 - Pr(A_1)) \cdot (1 - Pr(A_2|\overline{A_1})) \cdot \ldots \cdot (1 - Pr(A_n| \bigwedge_{i=1}^{n-1} \overline{A_i}) \geq \prod_{i=1}^{n} (1 - x_i)$$

completing the proof. □

Corollary 1.2 (Lovász Local Lemma; Symmetric Case): Let $A_1, A_2 \ldots A_n$ be events in an arbitrary probability space. Suppose that each event A_i is mutually independent of a set of all the other events A_j but at most d, and that $Pr(A_i) \leq p$ for all $1 \leq i \leq n$. If

$$ep(d + 1) \leq 1$$

then $\Pr\left(\bigwedge_{i=1}^{n} \overline{A_i}\right) > 0$.

If $d = 0$ the result is trivial. Otherwise, by the assumption there is a dependency digraph $D = (V, E)$ for the events $A_1 \ldots A_n$ in which for each i $|\{j : (i,j) \in E\}| \leq d$. The result now follows from Lemma 1.1 by taking $x_i = 1/(d + 1)(< 1)$ for all i and using the fact that for any $d \geq 2$, $\left(1 - \frac{1}{d+1}\right)^d > 1/e$.□

It is worth noting that as shown by Shearer in 1985,the constant "e" is the best possible constant in inequality (1.5). Note also that the proof of Lemma 1.1 indicates that the conclusion remains true even when we replace the two assumptions that each A_i is mutually independent of $\{A_j : (i,j) \notin E)$ and that $Pr(A_i) \leq x_i \prod_{(ij) \in E}(1 - x_j)$ by the weaker assumption that for each i and each $S_2 \subset \{1 \ldots n\} - \{j : (i,j) \in E\}$, $Pr\left(x_i | \bigwedge_{j \in S_2} \overline{A_j}\right) \leq x_i \prod_{(i,j) \in E}(1 - x_j)$. This turns out to be useful in certain applications.

In the next few sections we present various applications of the Lovász Local Lemma for obtaining combinatorial results. There is no known proof of any of these results, which does not use the this Lemma.

2 Property B and multicolored sets of real numbers

A hypergraph $H = (V, E)$ is said to have property B if there is a coloring of V by two colors so that no edge $f \in E$ is monochromatic.

Theorem 2.1. Let $H = (V, E)$ be a hypergraph in which every edge has at least k elements, and suppose that each edge of H intersects at most d other edges. If $e(d + 1) \leq 2^{k-1}$ then H has property B.

Color each vertex v of H, randomly and independently, either blue or red (with equal probability). For each edge $f \in E$, let A_f be the event that f is monochromatic. Clearly $Pr(A_f) = 2/2^{|f|} \leq 1/2^{k-1}$. Moreover, each event A_f is clearly mutually independent of all the other events $A_{f'}$ for all edges f' that do not intersect f. The result now follows from Corollary 1.2. □

A special case of Theorem 2.1 is that for any $k \geq 9$, any k-uniform k-regular hypergraph H has property B. Indeed, since any edge f of such an H contains k vertices, each of which

1 The Lemma

In a typical probabilistic proof of a combinatorial result, one usually has to show that the probability of a certain event is positive. However, many of these proofs actually give more and show that the probability of the event considered is not only positive but is large. In fact, most probabilistic proofs deal with events that hold with high probability, i.e., a probability that tends to 1 as the dimensions of the problem grow. On the other hand, there is a trivial case in which one can show that a certain event holds with positive, though very small, probability. Indeed, if we have n mutually independent events and each of them holds with probability at least $p > 0$, then the probability that all events hold simultaneously is at least p^n, which is positive, although it may be exponentially small in n.

It is natural to expect that the case of mutual independence can be generalized to that of rare dependencies, and provide a more general way of proving that certain events hold with positive, though small, proability. Such a generalization is, indeed, possible, and is stated in the following lemma, known as the Lovász Local Lemma. This simple lemma, first proved in [Erdős-Lovász (1975)] is an extremely powerful tool, as it supplies a way for dealing with rare events.

Lemma 1.1 (The Local Lemma; General Case):

Let $A_1, A_2 \dots A_n$ be events in an arbitrary probability space. A directed graph $D = (V, E)$ on the set of vertices $V = \{1, 2 \dots n\}$ is called a *dependency digraph* for the events $A_1 \dots A_n$ if for each i, $1 \le i \le n$, the event A_i is mutually independent of all the events $\{A_j : (i,j) \notin E\}$. Suppose that $D = (V, E)$ is a dependency digraph for the above events and suppose there are real numbers $x_1 \dots x_n$ such that $0 \le x_i < 1$ and $Pr(A_i) \le x_i \prod_{(i,j)\in E}(1 - x_j)$ for all $1 \le i \le n$. Then $Pr\left(\bigwedge_{i=1}^{n} \overline{A_j}\right) \ge \prod\limits_{i=1}^{n}(1 - x_i)$. In particular, with positive probability no event A_i holds.

We first prove, by induction on s, that for any $S \subset \{1 \dots n\}$, $|S| = s < n$ and any $i \notin S$

$$\Pr\left(A_i | \bigwedge_{j\in S} \overline{A_j}\right) \le x_i$$

This is certainly true for $s = 0$. Assuming it holds for all $s' < s$, we prove it for S. Put

$$S_1 = \{j \in S; (i,j) \in E\}, S_2 = S - S_1$$

Then

$$\Pr\left(A_i | \bigwedge_{j\in S} \overline{A_j}\right) = \frac{\Pr\left(A_i \wedge (\bigwedge_{j\in S_1} \overline{A_j}) | \bigwedge_{l\in S_2} \overline{A_l}\right)}{\Pr\left(\bigwedge_{j\in S_1} \overline{A_j} | \bigwedge_{l\in S_2} \overline{A_l}\right)}$$

To bound the numerator observe that since A_i is mutually independent of the events $\{A_l : l \in S_2\}$

$$\Pr\left(A_i \wedge (\bigwedge_{j\in S_1} \overline{A_j}) | \bigwedge_{l\in S_2} \overline{A_l}\right) \le \Pr\left(A_i | \bigwedge_{l\in S_2} \overline{A_l}\right) = Pr(A_i) \le x_i \prod_{(i,j)\in E} (1 - x_j)$$

The denominator, on the other hand, can be bounded by the induction hypothesis. Indeed, suppose $S_1 = \{j_1, j_2 \dots j_r\}$. If $r = 0$ then the denominator is 1, and (1.1) follows. Otherwise, setting $B = \wedge_{l\in S_2} \overline{A_l}$,

$$\Pr\left(\overline{A_{j_1}} \wedge \overline{A_{j_2}} \dots \overline{A_{j_r}} | B\right) = (1 - \Pr(A_{j_1} | B)) \cdot$$

at a time. The value of ϵ_i can only change X by two so direct application of Theorem 4.1 gives $|X_{i+1} - X_i| \leq 2$. But let ϵ, ϵ' be two n-tuples differing only in the i-th coordinate.

$$X_i(\epsilon) = \frac{1}{2}\left[X_{i+1}(\epsilon) + X_{i+1}(\epsilon')\right]$$

so that

$$|X_i(\epsilon) - X_{i+1}(\epsilon)| = \frac{1}{2}\left|X_{i+1}(\epsilon') - X_{i+1}(\epsilon)\right| \leq 1$$

Now apply Azuma's Inequality. □

For a third illustration let ρ be the Hamming metric on $\{0,1\}^n$. For $A \subseteq \{0,1\}^n$ let $B(A,s)$ denote the set of $y \in \{0,1\}^n$ so that $\rho(x,y) \leq s$ for some $x \in A$. ($A \subseteq B(A,s)$ as we may take $x = y$.)

Theorem 5.3. Let $\epsilon, \lambda > 0$ satisfy $e^{-\lambda^2/2} = \epsilon$. Then

$$|A| \geq \epsilon 2^n \Rightarrow |B(A, 2\lambda\sqrt{n})| \geq (1-\epsilon)2^n$$

Proof. Consider $\{0,1\}^n$ as the underlying probability space, all points equally likely. For $y \in \{0,1\}^n$ set

$$X(y) = \min_{x \in A} \rho(x,y)$$

Let $X_0, X_1, \ldots, X_n = X$ be the martingale given by exposing one coordinate of $\{0,1\}^n$ at a time. The Lipschitz condition holds for X: If y, y' differ in just one coordinate then $X(y) - X(y') \leq 1$. Thus, with $\mu = E[X]$

$$\Pr[X < \mu - \lambda\sqrt{n}] < e^{-\lambda^2/2} = \epsilon$$

$$\Pr[X > \mu + \lambda\sqrt{n}] < e^{-\lambda^2/2} = \epsilon$$

But

$$\Pr[X = 0] = |A|2^{-n} \geq \epsilon$$

so $\mu \leq \lambda\sqrt{n}$. Thus

$$\Pr[X > 2\lambda\sqrt{n}] < \epsilon$$

and

$$|B(A, 2\lambda\sqrt{n})| = 2^n \Pr[X \leq 2\lambda\sqrt{n}] \geq 2^n(1-\epsilon) \qquad \square$$

Actually, a much stronger result is known. Let $B(s)$ denote the ball of radius s about $(0,\ldots,0)$. The Isoperimetric Inequality proved by Harper in 1966 states that

$$|A| \geq |B(r)| \Rightarrow |B(A,s)| \geq |B(r+s)|$$

One may actually use this inequality as a beginning to give an alternate proof that $\chi(G) \sim n/2\log_2 n$ and to prove a number of the other results we have shown using martingales.

Lecture 9: The Lovász Local Lemma

where q_{h^*} is the conditional probability that g agrees with h^* on B_{i+1} given that it agrees with h on B_i. (This is because for $h^* \in H[h']$ $w_{h'}$ is also the conditional probability that $g = h^*$ given that $g = h^*$ on B_{i+1}.) Thus

$$\begin{aligned} |X_{i+1}(h) - X_i(h)| &= \left|\sum_{h' \in H} w_{h'}[L(h') - \sum_{h^* \in H[h']} L(h^*) q_{h^*}]\right| \\ &\leq \sum_{h' \in H} w_{h'} \sum_{h^* \in H[h']} |q_{h^*}[L(h') - L(h^*)]| \end{aligned}$$

The Lipschitz condition gives $|L(h') - L(h^*)| \leq 1$ so

$$|X_{i+1}(h) - X_i(h)| \leq \sum_{h' \in H} w_{h'} \sum_{h^* \in H[h']} q_{h^*} = \sum_{h' \in H} w_{h'} = 1 \qquad \square$$

Now we can express Azuma's Inequality in a general form.
Theorem 4.2. Let L satisfy the Lipschitz condition relative to a gradation of length m and let $\mu = E[L(g)]$. Then for all $\lambda > 0$

$$\Pr[L(g) > \mu + \lambda\sqrt{m}] < e^{-\lambda^2/2}$$

$$\Pr[L(g) < \mu - \lambda\sqrt{m}] < e^{-\lambda^2/2}$$

5 Three Illustrations

Let g be the random function from $\{1, \ldots, n\}$ to itself, all n^n possible function equally likely. Let $L(g)$ be the number of values not hit, i.e., the number of y for which $g(x) = y$ has no solution. By Linearity of Expectation

$$E[L(g)] = n\left(1 - \frac{1}{n}\right)^n \sim \frac{n}{e}$$

Set $B_i = \{1, \ldots, i\}$. L satisfies the Lipschitz condition relative to this gradation since changing the value of $g(i)$ can change $L(g)$ by at most one. Thus
Theorem 5.1.

$$\Pr[|L(g) - \frac{n}{e}| > \lambda\sqrt{n}] < 2e^{-\lambda^2/2}$$

Deriving these asymptotic bounds from first principles is quite cumbersome.

As a second illustration let B be any normed space and let $v_1, \ldots, v_n \in B$ with all $|v_i| \leq 1$. Let $\epsilon_1, \ldots, \epsilon_n$ be independent with

$$\Pr[\epsilon_i = +1] = \Pr[\epsilon_i = -1] = \frac{1}{2}$$

and set

$$X = |\epsilon_1 v_1 + \ldots + \epsilon_n v_n|$$

Theorem 5.2.

$$\Pr[X - E[X] > \lambda\sqrt{n}] < e^{-\lambda^2/2}$$

$$\Pr[X - E[X] < -\lambda\sqrt{n}] < e^{-\lambda^2/2}$$

Proof. Consider $\{-1, +1\}^n$ as the underlying probability space with all $(\epsilon_1, \ldots, \epsilon_n)$ equally likely. Then X is a random variable and we define a martingale $X_0, \ldots, X_n = X$ by exposing one ϵ_i

With probability at least $1-\epsilon$ there is a u-coloring of all but at most $c'\sqrt{n}$ vertices. By the Lemma almost always, and so with probability at least $1-\epsilon$, these points may be colored with 3 further colors, giving a $u+3$-coloring of G. The minimality of u guarantees that with probability at least $1-\epsilon$ at least u colors are needed for G. Altogether

$$\Pr[u \leq \chi(G) \leq u+3] \geq 1-3\epsilon$$

and ϵ was arbitrarily small. □

Using the same technique similar results can be achieved for other values of α. For any fixed $\alpha > \frac{1}{2}$ one finds that $\chi(G)$ is concentrated on some fixed number of values.

4 A General Setting

The martingales useful in studying Random Graphs generally can be placed in the following general setting which is essentially the one considered in Maurey [1979] and in Milman and Schechtman [1986]. Let $\Omega = A^B$ denote the set of functions $g: B \rightarrow A$. (With B the set of pairs of vertices on n vertices and $A = \{0,1\}$ we may identify $g \in A^B$ with a graph on n vertices.) We define a measure by giving values p_{ab} and setting

$$\Pr[g(b) = a] = p_{ab}$$

with the values $g(b)$ assumed mutually independent. (In $G(n,p)$ all $p_{1b} = p, p_{0b} = 1-p$.) Now fix a gradation

$$\emptyset = B_0 \subset B_1 \subset \ldots \subset B_m = B$$

Let $L: A^B \rightarrow R$ be a functional. (E.g., clique number.) We define a martingale $X_0, X_1, \ldots, X_m$ by setting

$$X_i(h) = E[L(g)|g(b) = h(b) \text{ for all } b \in B_i]$$

X_0 is a constant, the expected value of L of the random g. X_m is L itself. The values $X_i(g)$ approach $L(g)$ as the values of $g(b)$ are "exposed". We say the functional L satisfies the Lipschitz condition relative to the gradation if for all $0 \leq i < m$

$$h, h' \text{ differ only on } B_{i+1} - B_i \Rightarrow |L(h') - L(h)| \leq 1$$

Theorem 4.1. Let L satisfy the Lipschitz condition. Then the corresponding martingale satisfies

$$|X_{i+1}(h) - X_i(h)| \leq 1$$

for all $0 \leq i < m$, $h \in A^B$.

Proof. Let H be the family of h' which agree with h on B_{i+1}. Then

$$X_{i+1}(h) = \sum_{h' \in H} L(h') w_{h'}$$

where $w_{h'}$ is the conditional probability that $g = h'$ given that $g = h$ on B_{i+1}. For each $h' \in H$ let $H[h']$ denote the family of h^* which agree with h' on all points except (possibly) $B_{i+1} - B_i$. The $H[h']$ partition the family of h^* agreeing with h on B_i. Thus we may express

$$X_i(h) = \sum_{h' \in H} \sum_{h^* \in H[h']} [L(h^*) q_{h^*}] w_{h'}$$

cliques.) G has no k-clique if and only if $Y=0$. Apply Azuma's Inequality with $m=\binom{n}{2}\sim n^2/2$ and $E[Y]\geq \frac{n^2}{2k^4}(1+o(1))$. Then

$$\begin{array}{lll} \Pr[\omega(G)<k] & =\Pr[Y=0] & \leq \Pr[Y-E[Y]\leq -E[Y]] \\ & \leq e^{-E[Y]^2/2\binom{n}{2}} & \leq e^{-(c'+o(1))n^2/k^8} \\ & = e^{-(c+o(1))n^2/\ln^8 n} & \end{array}$$

as desired. □

Here is another example where the martingale approach requires an inventive choice of graphtheoretic function.

Theorem 3.3. Let $p=n^{-\alpha}$ where α is fixed, $\alpha>\frac{5}{6}$. Let $G=G(n,p)$. Then there exists $u=u(n,p)$ so that almost always

$$u\leq \chi(G)\leq u+3$$

That is, $\chi(G)$ is concentrated in four values.

We first require a technical lemma that had been well known.

Lemma 3.4. Let α, c be fixed $\alpha>\frac{5}{6}$. Let $p=n^{-\alpha}$. Then almost always every $c\sqrt{n}$ vertices of $G=G(n,p)$ may be 3-colored.

Proof. If not, let T be a minimal set which is not 3-colorable. As $T-\{x\}$ is 3-colorable, x must have internal degree at least 3 in T for all $x\in T$. Thus if T has t vertices it must have at least $\frac{3t}{2}$ edges. The probability of this occuring for some T with at most $c\sqrt{n}$ vertices is bounded from above by

$$\sum_{t=4}^{c\sqrt{n}}\binom{n}{t}\binom{\binom{t}{2}}{\frac{3t}{2}}p^{3t/2}$$

We bound

$$\binom{n}{t}\leq (\frac{ne}{t})^t \text{ and } \binom{\binom{t}{2}}{\frac{3t}{2}}\leq (\frac{te}{3})^{3t/2}$$

so each term is at most

$$\left[\frac{ne}{t}\frac{t^{3/2}e^{3/2}}{3^{3/2}}n^{-3\alpha/2}\right]^t\leq \left[c_1 n^{1-\frac{3\alpha}{2}}t^{1/2}\right]^t\leq \left[c_2 n^{1-\frac{3\alpha}{2}}n^{1/4}\right]^t=\left[c_2 n^{-\epsilon}\right]^t$$

with $\epsilon=\frac{3\alpha}{2}-\frac{5}{4}>0$ and the sum is therefore $o(1)$.

Proof of Theorem 3.3. Let $\epsilon>0$ be arbitrarily small and let $u=u(n,p,\epsilon)$ be the least integer so that

$$\Pr[\chi(G)\leq u]>\epsilon$$

Now define $Y(G)$ to be the minimal size of a set of vertices S for which $G-S$ may be u-colored. This Y satisfies the vertex Lipschitz condition since at worst one could add a vertex to S. Apply the vertex exposure martingale on $G(n,p)$ to Y. Letting $\mu=E[Y]$

$$\Pr[Y\leq \mu-\lambda\sqrt{n-1}]<e^{-\lambda^2/2}$$

$$\Pr[Y\leq \mu+\lambda\sqrt{n-1}]<e^{-\lambda^2/2}$$

Let λ satisfy $e^{-\lambda^2/2}=\epsilon$ so that these tail events each have probability less than ϵ. We defined u so that with probability at least ϵ G would be u-colorable and hence $Y=0$. That is, $\Pr[Y=0]>\epsilon$. The first inequality therefore forces $c\leq\lambda\sqrt{n-1}$. Now employing the second inequality

$$\Pr[Y\geq 2\lambda\sqrt{n-1}]\leq \Pr[Y\geq \mu+\lambda\sqrt{n-1}]\leq\epsilon$$

Theorem 2.4 (Shamir, Spencer[1987]) Let n, p be arbitrary and let $c = E[\chi(G)]$ where $G \sim G(n,p)$. Then

$$\Pr[|\chi(G) - c| > \lambda\sqrt{n-1}] < 2e^{-\lambda^2/2}$$

Proof. Consider the vertex exposure martingale $X_1, \ldots, X_n$ on $G(n,p)$ with $f(G) = \chi(G)$. A single vertex can always be given a new color so the vertex Lipschitz condition applies. Now apply Azuma's Inequality. □

Letting $\lambda \to \infty$ arbitrarily slowly this result shows that the distribution of $\chi(G)$ is "tightly concentrated" around its mean. The proof gives no clue as to where the mean is.

3 Chromatic Number

We have previously shown that $\chi(G) \sim n/2\log_2 n$ almost surely, where $G \sim G(n,1/2)$. Here we give the original proof of Béla Bollobás using martingales. We follow the earlier notations setting $f(k) = \binom{n}{k}2^{-\binom{k}{2}}$, k_0 so that $f(k_0-1) > 1 > f(k_0)$, $k = k_0 - 4$ so that $k \sim 2\log_2 n$ and $f(k) > n^{3+o(1)}$. Our goal is to show

$$\Pr[\omega(G) < k] = e^{-n^{2+o(1)}},$$

where $\omega(G)$ is the size of the maximum clique of G. We shall actually show in Theorem 3.2 a more precise bound. The remainder of the argument is as given earlier.

Let $Y = Y(H)$ be the maximal size of a family of edge disjoint cliques of size k in H. This ingenious and unusual choice of function is key to the martingale proof.

Lemma 3.1. $E[Y] \geq \frac{n^2}{2k^4}(1+o(1))$

Proof. Let $\mathcal{K}$ denote the family of k-cliques of G so that $f(k) = \mu = E[|\mathcal{K}|]$. Let W denote the number of unordered pairs $\{A, B\}$ of k-cliques of G with $2 \leq |A \cap B| < k$. Then $E[W] = \Delta/2$, with Δ as described earlier, $\Delta \sim \mu^2 k^4 n^{-2}$. Let $\mathcal{C}$ be a random subfamily of $\mathcal{K}$ defined by setting, for each $A \in \mathcal{K}$,

$$\Pr[A \in \mathcal{C}] = q,$$

q to be determined. Let W' be the number of unordered pairs $\{A, B\}$, $A, B \in \mathcal{C}$ with $2 \leq |A \cap B| < k$. Then

$$E[W'] = E[W]q^2 = \Delta q^2/2$$

Delete from $\mathcal{C}$ one set from each such pair $\{A, B\}$. This yields a set $\mathcal{C}^*$ of edge disjoint k-cliques of G and

$$E[Y] \geq E[|\mathcal{C}^*|] \geq E[|\mathcal{C}|] - E[W'] = \mu q - \Delta q^2/2 = \mu^2/2\Delta \sim n^2/2k^4$$

where we choose $q = \mu/\Delta$ (noting that it is less than one!) to minimize the quadratic. □

We conjecture that Lemma 3.1 may be improved to $E[Y] > cn^2/k^2$. That is, with positive probability there is a family of k-cliques which are edge disjoint and cover a positive proportion of the edges.

Theorem 3.2.

$$\Pr[\omega(G) < k] < e^{-(c+o(1))\frac{n^2}{\ln^8 n}}$$

with c a positive constant.

Proof. Let $Y_0, \ldots, Y_m$, $m = \binom{n}{2}$, be the edge exposure martingale on $G(n,1/2)$ with the function Y just defined. The function Y satisfies the edge Lipschitz condition as adding a single edge can only add at most one clique to a family of edge disjoint cliques. (Note that the Lipschitz condition would not be satisfied for the number of k-cliques as a single edge might yield many new

The figure shows why this is a martingale. The conditional expectation of $f(H)$ knowing the first $i-1$ edges is the weighted average of the conditional expectations of $f(H)$ where the i-th edge has been exposed. More generally - in what is sometimes referred to as a Doob martingale process - X_i may be the conditional expectation of $f(H)$ after certain information is revealed as long as the information known at time i includes the information known at time $i-1$.
The Vertex Exposure Martingale. Again let $G(n,p)$ be the underlying probability space and f any graphtheoretic function. Define $X_1,\ldots,X_n$ by

$$X_i(H) = E[f(G)|\text{for } x,y \le i, \{x,y\} \in G \longleftrightarrow \{x,y\} \in H]$$

In words, to find $X_i(H)$ we expose the first i vertices and all their internal edges and take the conditional expectation of $f(G)$ with that partial information. By ordering the edges appropriately the vertex exposure martingale may be considered a subsequence of the edge exposure martingale. Note that $X_1(H) = E[f(G)]$ is constant as no edges have been exposed and $X_n(H) = f(H)$ as all edges have been exposed.

2 Large Deviations

Maurey [1979] applied a large deviation inequality for martingales to prove an isoperimetric inequality for the symmetric group S_n. This inequality was useful in the study of normed spaces; see Milman and Schechtman [1986] for many related results. The applications of martingales in Graph Theory also all involve the same underlying martingale results used by Maurey, which are the following.
Theorem 2.1 (Azuma's Ineqality) Let $0 = X_0,\ldots,X_m$ be a martingale with

$$|X_{i+1} - X_i| \le 1$$

for all $0 \le i < m$. Let $\lambda > 0$ be arbitrary. Then

$$\Pr[X_m > \lambda\sqrt{m}] < e^{-\lambda^2/2}$$

Corollary 2.2 Let $c = X_0,\ldots,X_m$ be a martingale with

$$|X_{i+1} - X_i| \le 1$$

for all $0 \le i < m$. Then

$$\Pr[|X_m - c| > \lambda\sqrt{m}] < 2e^{-\lambda^2/2}.$$

A graph theoretic function f is said to satisfy the *edge Lipschitz condition* if whenever H and H' differ in only one edge then $|f(H) - f(H')| \le 1$. It satisfies the *vertex Lipschitz condition* if whenever H and H' differ at only one vertex $|f(H) - f(H')| \le 1$.
Theorem 2.3 When f satisfies the edge Lipschitz condition the corresponding edge exposure martingale satisfies $|X_{i+1} - X_i| \le 1$. When f satisfies the vertex Lipschitz condition the corresponding vertex exposure martingale satisfies $|X_{i+1} - X_i| \le 1$.

We prove these results in a more general context later. They have the intuitive sense that if knowledge of a particular vertex or edge cannot change f by more than one then exposing a vertex or edge should not change the expectation of f by more than one. Now we give a simple application of these results.

objects) while here it depends only on their sizes. So it seems there should be a probability space whose elements are histories - i.e., the value of $P(\lambda)$ for all real λ - where the change from $P(\lambda)$ to $P(\lambda + d\lambda)$ is governed by these coagulation laws and where further there have to be some appropriate entry laws so that each $P(\lambda)$ has the appropriate distribution. Not that any of this has been done - but in theory there is a theory!

Lecture 8: Martingales

1 Definitions

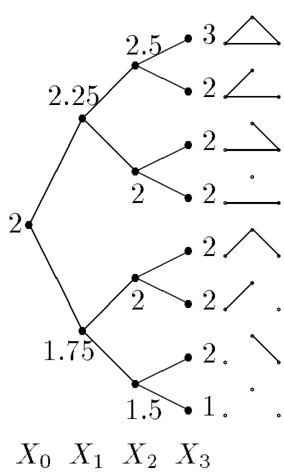
A martingale is a sequence $X_0, \ldots, X_m$ of random variables so that for $0 \leq i < m$,

$$E[X_{i+1}|X_i] = X_i$$

The Edge Exposure Martingale Let the random graph $G(n,p)$ be the underlying probability space. Label the potential edges $\{i,j\} \subseteq [n]$ by $e_1, \ldots, e_m$, setting $m = \binom{n}{2}$ for convenience, in any specific manner. Let f be any graphtheoretic function. We define a martingale $X_0, \ldots, X_m$ by giving the values $X_i(H)$. $X_m(H)$ is simply $f(H)$. $X_0(H)$ is the expected value of $f(G)$ with $G \sim G(n,p)$. Note that X_0 is a constant. In general (including the cases $i = 0$ and $i = m$)

$$X_i(H) = E[f(G)|e_j \in G \longleftrightarrow e_j \in H, 1 \leq j \leq i]$$

In words, to find $X_i(H)$ we first expose the first i pairs $e_1, \ldots, e_i$ and see if they are in H. The remaining edges are not seen and considered to be random. $X_i(H)$ is then the conditional expectation of $f(G)$ with this partial information. When $i = 0$ nothing is exposed and X_0 is a constant. When $i = m$ all is exposed and X_m is the function f. The martingale moves from no information to full information in small steps.



The edge exposure martingale with $n = m = 3, f$ the chromatic number, and the edges exposed in the order "bottom,left,right". The values $X_i(H)$ are given by tracing from the central node to the leaf labelled H.

and letting X^* be the total number of components of size between $an^{2/3}$ and $bn^{2/3}$

$$\lim_{n\to\infty} E[X^*] = \int_a^b e^{-\frac{c^3}{6}-\frac{\lambda^2 c}{2}+\frac{\lambda c^2}{2}} c^{-5/2}(2\pi)^{-1/2} g(c)dc$$

where

$$g(c) = \sum_{l=0}^{\infty} c_l c^{\frac{3}{2}l}$$

a sum convergent for all c, (here $c_0 = 1$). A component of size $\sim cn^{2/3}$ will have probability $c_l c^{\frac{3}{2}l}/g(c)$ of having $l-1$ more edges than vertices, independent of λ. As $\lim_{c\to 0} g(c) = 1$, most components of size $\epsilon n^{2/3}$, $\epsilon << 1$, are trees but as c gets bigger the distribution on l moves inexoribly higher.

An Overview. For any fixed λ the sizes of the largest components are of the form $cn^{2/3}$ with a distribution over the constant. For $\lambda = -10^6$ there is some positive limiting probability that the largest component is bigger than $10^6 n^{2/3}$ and for $\lambda = +10^6$ there is some positive limiting probability that the largest component is smaller than $10^{-6}n^{2/3}$, though both these probabilities are minuscule. The functions integrated have a pole at $c = 0$, reflecting the notion that for any λ there should be many components of size near $\epsilon n^{2/3}$ for $\epsilon = \epsilon(\lambda)$ appropriately small. When λ is large negative (e.g., -10^6) the largest component is likely to be $\epsilon n^{2/3}$, ϵ small, and there will be many components of nearly that size. The nontree components will be a negligible fraction of the tree components.

Now consider the evolution of $G(n,p)$ in terms of λ. Suppose that at a given λ there are components of size $c_1 n^{2/3}$ and $c_2 n^{2/3}$. When we move from λ to $\lambda + d\lambda$ there is a probability $c_1 c_2 d\lambda$ that they will merge. Components have a peculiar gravitation in which the probability of merging is proportional to their sizes. With probability $(c_1^2/2)d\lambda$ there will be a new internal edge in a component of size $c_1 n^{2/3}$ so that large components rarely remain trees. Simultaneously, big components are eating up other vertices.

With $\lambda = -10^6$, say, we have feudalism. Many small components (castles) are each vying to be the largest. As λ increases the components increase in size and a few large components (nations) emerge. An already large France has much better chances of becoming larger than a smaller Andorra. The largest components tend strongly to merge and by $\lambda = +10^6$ it is very likely that a giant component, Roman Empire, has emerged. With high probability this component is nevermore challenged for supremacy but continues absorbing smaller components until full connectivity - One World - is achieved.

An Continuous Model. In discussions at St. Flour it became apparent that there was a continuous model underlying the asymptotic behavior of $G(n,p)$ with $p = n^{-1} + \lambda n^{-4/3}$. The following should be regarded as only tentative steps toward defining of that continuous model. For fixed λ and k arbitrarily large but fixed one can look at the k largest components of $G(n,p)$ and parametrize them $x_1 n^{2/3}, \ldots, x_k n^{2/3}$ in decreasing order. One can give explicitly a limiting distribution function $H(x_1,\ldots,x_k)$ for these values. Now one can go to the limit with k and consider the "state" $P(\lambda)$ at "time" λ to be an infinite sequence $x_1 > x_2 > \ldots$ of decreasing reals. There will be a distribution over the possible sequences. The sequences must be well-behaved; one can show, for example, that the number of x_i bigger than c must be asymptotic to $\frac{2}{3}(2\pi)^{-1/2}c^{-3/2}$ as $c \to 0$. (There is further information concerning the nature of the components - e.g., are they trees, unicyclic,...- that could also be added.) Now the intriguing thing is the "gravity" that defines $P(\lambda + d\lambda)$ in terms of $P(\lambda)$ in an appropriate limiting sense. If $P(\lambda)$ has terms x_i, x_j then with probability $x_i x_j d\lambda$ they will "merge" and form a single term with value $x_i + x_j$. This corresponds to certain coagulation models in physics though in the physical world the probability of coagulation depends on the surface area (and perhaps other invariants) of the so that

$$\sum_{i=1}^{k-1} - \ln(1-\frac{i}{n}) = \frac{k^2}{2n} + \frac{k^3}{6n^2} + o(1) = \frac{k^2}{2n} + \frac{c^3}{6} + o(1)$$

Also

$$p^{k-1} = n^{1-k}(1+\frac{\lambda}{n^{1/3}})^{k-1}$$

$$(k-1)\ln(1+\frac{\lambda}{n^{1/3}}) = (k-1)(\frac{\lambda}{n^{1/3}} - \frac{\lambda^2}{2n^{2/3}} + O(n^{-1})) = \frac{\lambda k}{n^{1/3}} - \frac{\lambda^2 c}{2} + o(1)$$

Also

$$\ln(1-p) = -p + O(n^{-2}) = -\frac{1}{n} - \frac{\lambda}{n^{4/3}} + O(n^{-2})$$

and

$$k(n-k) + \binom{k}{2} - (k-1) = kn - \frac{k^2}{2} + O(n^{2/3})$$

so that

$$[k(n-k) + \binom{k}{2} - (k-1)]\ln(1-p) = -k + \frac{k^2}{2n} - \frac{\lambda k}{n^{1/3}} + \frac{\lambda c^2}{2} + o(1)$$

and

$$E[X] \sim \frac{n^k k^{k-2}}{k^k\sqrt{2\pi k}n^{k-1}}e^A$$

where

$$A = k - \frac{k^2}{2n} - \frac{c^3}{6} + \frac{\lambda k}{n^{1/3}} - \frac{\lambda^2 c}{2} - k + \frac{k^2}{2n} - \frac{\lambda k}{n^{1/3}} + \frac{\lambda c^2}{2} + o(1)$$

$$= -\frac{c^3}{6} - \frac{\lambda^2 c}{2} + \frac{\lambda c^2}{2} + o(1)$$

so that

$$E[X] \sim n^{-2/3}e^{-\frac{c^3}{6}-\frac{\lambda^2 c}{2}+\frac{\lambda c^2}{2}}c^{-5/2}(2\pi)^{-1/2}$$

For any particular such k $E[X] \to 0$ but if we sum k between $cn^{2/3}$ and $(c+dc)n^{2/3}$ we multiply by $n^{2/3}dc$. Going to the limit gives an integral: For any fixed a, b, λ let X be the number of tree components of size between $an^{2/3}$ and $bn^{2/3}$. Then

$$\lim_{n\to\infty} E[X] = \int_a^b e^{-\frac{c^3}{6}-\frac{\lambda^2 c}{2}+\frac{\lambda c^2}{2}}c^{-5/2}(2\pi)^{-1/2}dc$$

The large components are not all trees. E.M. Wright [1977] proved that for fixed l there are asymptotically $c_l k^{k-2+\frac{3}{2}l}$ connected graphs on k points with $k-1+l$ edges, where c_l was given by a specific recurrence. Asymptotically in l,

$$c_l \sim \left(\frac{e}{12l}(1+o(1))\right)^{l/2}.$$

The calculation for $X^{(l)}$, the number of such components on k vertices, leads to extra factors of $c_l k^{\frac{3}{2}l}$ and n^{-l} which gives $c_l c^{\frac{3}{2}l}$. For fixed a, b, λ, l the number $X^{(l)}$ of components of size between $an^{2/3}$ and $bn^{2/3}$ with $l-1$ more edges than vertices satisfies

$$\lim_{n\to\infty} E[X^{(l)}] = \int_a^b e^{-\frac{c^3}{6}-\frac{\lambda^2 c}{2}+\frac{\lambda c^2}{2}}c^{-5/2}(2\pi)^{-1/2}(c_l c^{\frac{3}{2}l})dc$$

Here we use the nontrivial fact, due to Cayley, that there are k^{k-2} possible trees on a given k-set. For c, k fixed

$$E[X] \sim n\frac{e^{-ck}k^{k-2}c^{k-1}}{k!}$$

As trees are strictly balanced a second moment method gives $X \sim E[X]$ almost always. Thus $\sim p_k n$ points lie in tree components of size k where

$$p_k = \frac{e^{-ck}(ck)^{k-1}}{k!}$$

It can be shown analytically that $p_k = \Pr[T = k]$ in the Branching Process with mean c. Let Y_k denote the number of cycles of size k and Y the total number of cycles. Then

$$E[Y_k] = \frac{(n)_k}{2k}(\frac{c}{n})^k \sim \frac{c^k}{2k}$$

for fixed k. For $c < 1$

$$E[Y] = \sum E[Y_k] \to \sum_{k=1}^{\infty} \frac{c^k}{2k}$$

has a finite limit whereas for $c > 1$, $E[Y] \to \infty$. Even for $c > 1$ for any fixed k the number of k-cycles has a limiting expectation and so do not asymptotically affect the number of components of a given size.

3 Inside the Phase Transition

In the evolution of the random graph $G(n, p)$ a crucial change takes place in the vicinity of $p = c/n$ with $c = 1$. The small components at that time are rapidly joining together to form a giant component. This corresponds to the Branching Process when births are Poisson with mean 1. There the number T of organisms will be finite almost always and yet have infinite expectation. No wonder that the situation for random graphs is extremely delicate. In recent years there has been much interest in looking "inside" the phase transition at the growth of the largest components. (See, e.g. Luczak [1990].) The appropriate parametrization is, perhaps surprisingly,

$$p = \frac{1}{n} + \frac{\lambda}{n^{4/3}}$$

When $\lambda = \lambda(n) \to -\infty$ the phase transition has not yet started. The largest components are $o(n^{2/3})$ and there are many components of nearly the largest size. When $\lambda = \lambda(n) \to +\infty$ the phase transition is over - a largest component, of size $>> n^{2/3}$ has emerged and all other components are of size $o(n^{2/3})$. Let's fix λ, c and let X be the number of tree components of size $k = cn^{2/3}$. Then

$$E[X] = \binom{n}{k} k^{k-2} p^{k-1} (1-p)^{k(n-k)+\binom{k}{2}-(k-1)}$$

Watch the terms cancel!

$$\binom{n}{k} = \frac{(n)_k}{k!} \sim \frac{n^k e^k}{k^k\sqrt{2\pi k}} \prod_{i=1}^{k-1}(1 - \frac{i}{n})$$

For $i < k$

$$-\ln(1 - \frac{i}{n}) = \frac{i}{n} + \frac{i^2}{2n^2} + O(\frac{i^3}{n^3})$$

Now assume $c > 1$. For any fixed t, $\lim_{n\to\infty} \Pr[T = t] = \Pr[T^* = t]$ but what corresponds to $T^* = \infty$? For $t = o(n)$ we may estimate $1-(1-p)^t \sim pt$ and $n-1 \sim n$ so that

$$\Pr[Y_t \leq 0] = \Pr[B[n-1, 1-(1-p)^t] \leq t-1] \sim \Pr[B[n, tc/n] \leq t]$$

drops exponentially in t by Large Deviation results. When $t = \alpha n$ we estimate $1-(1-p)^t$ by $1-e^{-c\alpha}$. The equation $1-e^{-c\alpha} = \alpha$ has solution $\alpha = 1-y$ where y is the extinction probability. For $\alpha < 1-y$, $1-e^{-c\alpha} > \alpha$ and

$$\Pr[Y_t \leq 0] \sim \Pr[B[n, 1-e^{-c\alpha}] \leq \alpha n]$$

is exponentially small while for $\alpha > 1-y$, $1-e^{-c\alpha} < \alpha$ and $\Pr[Y_t \leq 0] \sim 1$. Thus almost always $Y_t = 0$ for some $t \sim (1-y)n$. Basically, $T^* = \infty$ corresponds to $T \sim (1-y)n$. Let $\epsilon, \delta > 0$ be arbitrarily small. With somewhat more care to the bounds we may show that there exists t_0 so that for n sufficiently large

$$\Pr[t_0 < T < (1-\delta)n(1-y) \text{ or } \mathrm{T > (1+\delta)n(1-y)}] < \epsilon$$

Pick t_0 sufficiently large so that

$$y - \epsilon \leq \Pr[T^* \leq t_0] \leq y$$

Then as $\lim_{n\to\infty} \Pr[T \leq t_0] = \Pr[T^* \leq 0]$ for n sufficiently large

$$y - 2\epsilon \leq \Pr[T \leq t_0] \leq y + \epsilon$$

$$1 - y - 2\epsilon \leq \Pr[(1-\delta)n(1-y) < T < (1+\delta)n(1-y)] < 1 - y + 3\epsilon$$

Now we expand our procedure to find graph components. We start with $G \sim G(n,p)$, select $v = v_1 \in G$ and compute $C(v_1)$ as before. Then we delete $C(v_1)$, pick $v_2 \in G - C(v_1)$ and iterate. At each stage the remaining graph has distribution $G(m,p)$ where m is the number of vertices. (Note, critically, that no pairs $\{w, w'\}$ in the remaining graph have been examined and so it retains its distribution.) Call a component $C(v)$ small if $|C(v)| \leq t_0$, giant if $(1-\delta)n(1-y) < |C(v)| < (1+\delta)n(1-y)$ and otherwise failure. Pick $s = s(\epsilon)$ with $(y+\epsilon)^s < \epsilon$. (For ϵ small $s \sim K \ln \epsilon^{-1}$.) Begin this procedure with the full graph and terminate it when either a giant component or a failure component is found or when s small components are found. At each stage, as only small components have thus far been found, the number of remaining points is $m = n - O(1) \sim n$ so the conditional probabilities of small, giant and failure remain asymptotically the same. The chance of ever hitting a failure component is thus $\leq s\epsilon$ and the chance of hitting all small components is $\leq (y+\epsilon)^s \leq \epsilon$ so that with probability at least $1-\epsilon'$, where $\epsilon' = (s+1)\epsilon$ may be made arbitrarily small, we find a series of less than s small components followed by a giant component. The remaining graph has $m \sim yn$ points and $pm \sim cy = d$, the conjugate of c as defined earlier. As $d < 1$ the previous analysis gives the maximal components. In summary: almost always $G(n, c/n)$ has a giant component of size $\sim (1-y)n$ and all other components of size $O(\ln n)$. Furthermore, the Duality Principle has a discrete analog.
Discrete Duality Principle. Let $d < 1 < c$ be conjugates. The structure of $G(n, c/n)$ with its giant component removed is basically that of $G(m, d/m)$ where m, the number of vertices not in the giant component, satisfies $m \sim ny$.

The small components of $G(n, c/n)$ can also be examined from a static view. For a fixed k let X be the number of tree components of size k. Then

$$E[X] = \binom{n}{k} k^{k-2} p^{k-1} (1-p)^{k(n-k)+\binom{k}{2}-(k-1)}$$

the number of neutral vertices at time t and show, equivalently,

$$N_t \sim B[n-1,(1-p)^t]$$

This is reasonable since each $w \neq v$ has independent probability $(1-p)^t$ of staying neutral t times. Formally, as $N_0 = n-1$ and

$$\begin{aligned} N_t &= n-t-Y_t &&= n-t-B[n-(t-1)-Y_{t-1},p]-Y_{t-1}+1 \\ & &&= N_{t-1}-B[N_{t-1},p] \\ & &&= B[N_{t-1},1-p] \end{aligned}$$

the result follows by induction. □

We set $p = c/n$. When t and Y_{t-1} are small we may approximate Z_t by $B[n,c/n]$ which is approximately Poisson with mean c. Basically small components will have size distribution as in the Branching Process. The analogy must break down for $c > 1$ as the Branching Process may have an infinite population whereas $|C(v)|$ is surely at most n. Essentially, those v for which the Branching Process for $C(v)$ does not "die early" all join together to form the giant component.

Fix c. Let $Y_0^*, Y_1^*, \ldots, T^*, Z_1^*, Z_2^*, \ldots, H^*$ refer to the Branching Process and let the unstarred $Y_0, Y_1, \ldots, T, Z_1, Z_2, \ldots, H$ refer to the Random Graph process. For any possible history $(z_1, \ldots, z_t)$

$$\Pr[H^* = (z_1, \ldots, z_t)] = \prod_{i=1}^{t} \Pr[Z^* = z_i]$$

where Z^* is Poisson with mean c while

$$\Pr[H = (z_1, \ldots, z_t)] = \prod_{i=1}^{t} \Pr[Z_i = z_i]$$

where Z_i has Binomial Distribution $B[n-1-z_1-\ldots-z_{i-1}, c/n]$. The Poisson distribution is the limiting distribution of Binomials. When $m = m(n) \sim n$ and c, i are fixed

$$\lim_{n\to\infty} \Pr[B[m,c/n] = i] = \lim_{n\to\infty} \binom{m}{z} (\frac{c}{n})^z (1-\frac{c}{n})^{m-z} = e^{-c}c^z/z!$$

hence

$$\lim_{n\to\infty} \Pr[H = (z_1, \ldots, z_t)] = \Pr[H^* = (z_1, \ldots, z_t)]$$

Assume $c < 1$. For any fixed t, $\lim_{n\to\infty} \Pr[T = t] = \Pr[T^* = t]$. We now bound the size of the largest component. For any t

$$\Pr[T > t] \le \Pr[Y_t > 0] = \Pr[B[n-1, 1-(1-p)^t] \ge t] \le \Pr[B[n, tc/n] \ge t]$$

as $1-(1-p)^t \le tp$ and $n-1 < n$. By Large Deviation Results

$$\Pr[T > t] < e^{-\alpha t}$$

where $\alpha = \alpha(c) > 0$. Let $\beta = \beta(c)$ satisfy $\alpha\beta > 1$. Then

$$\Pr[T > \beta \ln n] < n^{-\alpha\beta} = o(n^{-1})$$

There are n choices for initial vertex v. Thus almost always *all* components have size $O(\ln n)$.

$y_{i-1} + z_i - 1$ has $y_i > 0$ for $0 \leq i < t$ and $y_t = 0$. When Z is Poisson with mean λ

$$\Pr[H = (z_1, \ldots, z_t)] = \prod_{i-1}^{t} \frac{e^{-\lambda}\lambda^{z_i}}{z_i!} = \frac{e^{-\lambda}(\lambda e^{-\lambda})^{t-1}}{\prod_{i=1}^{t} z_i!}$$

since $z_1 + \ldots + z_t = t - 1$.

We call $d < 1 < c$ a conjugate pair if

$$de^{-d} = ce^{-c}$$

The function $f(x) = xe^{-x}$ increases from 0 to e^{-1} in [0,1) and decreases back to 0 in $(1, \infty)$ so that all $c \neq 1$ have a uniqe conjugate. Let $c > 1$ and $y = \Pr[T < \infty]$ so that $y = e^{c(y-1)}$. Then $(cy)e^{-cy} = ce^{-c}$ so

$$d = cy$$

Duality Principle. Let $d < 1 < c$ be conjugates. The Branching Process with mean c, conditional on extinction, has the same distribution as the Branching Process with mean d.
Proof. It suffices to show that for every history $H = (z_1, \ldots, z_t)$

$$\frac{e^{-c}(ce^{-c})^{t-1}}{y\prod_{i=1}^{t} z_i!} = \frac{e^{-d}(de^{-d})^{t-1}}{\prod_{i=1}^{t} z_i!}$$

This is immediate as $ce^{-c} = de^{-d}$ and $ye^{-d} = ye^{-cy} = e^{-c}$.

2 The Giant Component

Now let's return to random graphs. We define a procedure to find the component $C(v)$ containing a given vertex v in a given graph G. We are motivated by Karp [1990] in which this approach is applied to random digraphs. In this procedure vertices will be live, dead or neutral. Originally v is live and all other vertices are neutral, time $t = 0$ and $Y_0 = 1$. Each time unit t we take a live vertex w and check all pairs $\{w, w'\}$, w' neutral, for membership in G. If $\{w, w'\} \in G$ we make w' live, otherwise it stays neutral. After searching all neutral w' we set w dead and let Y_t equal the new number of live vertices. When there are no live vertices the process terminates and $C(v)$ is the set of dead vertices. Let Z_t be the number of w' with $\{w, w'\} \in G$ so that

$$Y_0 = 1$$

$$Y_t = Y_{t-1} + Z_t - 1$$

With $G = G(n, p)$ each neutral w' has independent probability p of becoming live. Here, critically, no pair $\{w, w'\}$ is ever examined twice so that the conditional probability for $\{w, w'\} \in G$ is always p. As $t - 1$ vertices are dead and Y_{t-1} are live

$$Z_t \sim B[n - (t-1) - Y_{t-1}, p]$$

Let T be the least t for which $Y_t = 0$. Then $T = |C(v)|$. As in Section 1 we continue the recursive definition of Y_t, this time for $0 \leq t \leq n$.
Claim 2.1 For all t

$$Y_t \sim B[n-1, 1 - (1-p)^t] + 1 - t$$

It is more convenient to deal with

$$N_t = n - t - Y_t$$

With additional work one can prove Theorem 3.2 with $c_1 = K(1-\epsilon'), c_2 = K(1+\epsilon')$ for arbitrarily small ϵ' and K dependent only on ϵ'.

Lecture 7: The Phase Transition

1 Branching Processes

Paul Erdős and Alfred Rényi, in their original 1960 paper, discovered that the random graph $G(n,p)$ undergoes a remarkable change at $p = 1/n$. Speaking roughly, let first $p = c/n$ with $c < 1$. Then $G(n,p)$ will consist of small components, the largest of which is of size $\Theta(\ln n)$. But now suppose $p = c/n$ with $c > 1$. In that short amount of "time" many of the components will have joined together to form a "giant component" of size $\Theta(n)$. The remaining vertices are still in small components, the largest of which has size $\Theta(\ln n)$. They dubbed this phenomenon the *Double Jump*. We prefer the descriptive term Phase Transition because of the connections to percolation (e.g., freezing) in mathematical physics.

To better understand the Phase Transition we make a lengthy detour into the subject of Branching Processes. Imagine that we are in a unisexual universe and we start with a single organism. Imagine that this organism has a number of children given by a given random variable Z. (For us, Z will be Poisson with mean c.) These children then themselves have children, the number again being determined by Z. These grandchildren then have children, etc. As $Z = 0$ will have nonzero probability there will be some chance that the line dies out entirely. We want to study the total number of organisms in this process, with particular eye to whether or not the process continues forever. (The original application of this model was to a study of the -gasp!- male line of British peerage.)

Now lets be more precise. Let $Z_1, Z_2, \ldots$ be independent random variables, each with distribution Z. Define $Y_0, Y_1, \ldots$ by the recursion

$$Y_0 = 1$$

$$Y_i = Y_{i-1} + Z_i - 1$$

and let T be the least t for which $Y_t = 0$. If no such t exists (the line continuing forever) we say $T = +\infty$. The Y_i and Z_i mirror the Branching Process as follows. We view all organisms as living or dead. Initially there is one live organism and no dead ones. At each time unit we select one of the live organisms, it has Z_i children, and then it dies. The number Y_i of live organisms at time i is then given by the recursion. The process stops when $Y_t = 0$ (extinction) but it is a convenient fiction to define the recursion for all t. Note that T is not affected by this fiction since once $Y_t = 0$, T has been defined. T (whether finite or infinite) is the total number of organisms, including the original, in this process. (A natural approach, found in many probability texts, is to have all organisms of a given generation have their children at once and study the number of children of each generation. While we may think of the organisms giving birth by generation it will not affect our model.)

We shall use the major result of Branching Processes that when $E[Z] = c < 1$ with probability one the process dies out ($T < \infty$) but when $E[Z] = c > 1$ then there is a nonzero probability that the process goes on forever ($T = \infty$).

When a branching process dies we call $H = (Z_1, \ldots, Z_T)$ the *history* of the process. A sequence $(z_1, \ldots, z_t)$ is a possible history if and only if the sequence y_i given by $y_0 = 1, y_i =$

$c_2 \log n$ for all sufficiently large n.
Proof. Define S randomly by

$$\Pr[x \in S] = p_x = \min\left[10\left(\frac{\ln x}{x^2}\right)^{1/3}, \frac{1}{2}\right]$$

Fix n. Now $g(n)$ is a random variable and

$$\mu = E[g(n)] = \sum_{x+y+z=n} p_x p_y p_z$$

Careful asymptotics give

$$\mu \sim 10^3 \ln n \int_{x=0}^{1} \int_{y=0}^{1-x} \frac{dxdy}{[xy(1-x-y)]^{2/3}} = K \ln n$$

where K is large. (We may make K arbitrarily large by increasing "10".) We apply the Janson inequality. Here $\epsilon = 1/8$ as all $p_x \le 1/2$. Also

$$\Delta = \sum p_x p_y p_z p_{y'} p_{z'},$$

the sum over all five-tuples with $x + y + z = x + y' + z' = n$. Roughly there are n^3 terms, each $\sim p_n^5 = n^{-10/3+o(1)}$ so that the sum is $o(1)$. Care must be taken that those terms with one (or more) small variables don't contribute much to the sum.

Now we emulate the argument of Theorem 5.3.1. Call F a maximal disjoint family of solutions if F is a family of sets $\{x_i, y_i, z_i\}$ with all x_i, y_i, z_i distinct, all $x_i + y_i + z_i = n$, all $x_i, y_i, z_i \in S$ and so that there is no $x, y, z \in S$ with $x + y + z = n$ and x, y, z distinct from all x_i, y_i, z_i. Let $Z^{(s)}$ denote the number of maximal disjoint families of solutions of size s. As in Theorem 5.3.1 when $s < \log^2 n$

$$E[Z^{(s)}] < \frac{\mu^s}{s!} e^{-\mu(1+o(1))}$$

while for $s \ge \log^2 n$

$$E[Z^{(s)}] < \mu^s / s!$$

so that $\sum^* E[Z^{(s)}] = o(n^{-10})$, say, where $\sum^*$ is over those s with $|s - \mu| > \epsilon\mu$. (Here ϵ is fixed and K must be sufficiently large.) With probability $1 - o(n^{-10})$ there is an F with $|s - \mu| < \epsilon\mu$.

When this occurs $g(n) \ge |F| \ge (1-\epsilon)\mu$ but again we must worry about $g(n)$ being considerably larger than $|F|$. Here we use only that $p = n^{-2/3+o(1)}$. Note that the number of representations of $n = x + y + z$ with a given x is the number of representations $m = y + z$ of $m = n - x$.
Lemma 3.3. Almost surely no sufficiently large m has four (or more) representations as $m = y+z$, $y, z \in S$.
Proof. Here $\mu = \Theta(m^{-1/3})$ so the expected number of 4-tuples of representatives is $O(m^{-4/3})$ and so the probability of having four representatives is $O(m^{-4/3})$. Apply Borel-Cantelli. □

Now almost surely there is a C so that *no* m has more than C representations $m = y + z$. Let S be such that this holds and that all maximal disjoint families of solutions F have

$$K(1-\epsilon)\log n < |F| < K(1+\epsilon)\log n$$

Each triple $x, y, z \in S$ with $x+y+z = n$ must include one of the at most $3K(1+\epsilon)\log n$ elements of sets of F and each such element is in less than C such triples so that $g(n) < 3CK(1+\epsilon)\log n$. Take $c_1 = K(1-\epsilon)$ and $c_2 = 3KC(1+\epsilon)$.

3 Counting Representations

For a given set S of natural numbers let (for every $n \in N$) $f(n) = f_S(n)$ denote the number of representations $n = x + y$, $x, y \in S, x \neq y$. For many years it was an open question whether there existed an S with $f(n) \geq 1$ for all sufficiently large n and yet $f(n) \leq n^{o(1)}$.
Theorem 3.1. (Erdős (1956)) There is a set S for which $f(n) = \Theta(\ln n)$. That is, there is a set S and constants c_1, c_2 so that for all sufficiently large n

$$c_1 \ln n \leq f(n) \leq c_2 \ln n$$

Proof. Define S randomly by

$$\Pr[x \in S] = p_x = \min\left[10\sqrt{\frac{\ln x}{x}}, 1\right]$$

Fix n. Now $f(n)$ is a random variable with mean

$$\mu = E[f(n)] = \sum_{x+y=n} p_x p_y$$

Roughly there are n addends with $p_x p_y > p_n^2 = 100\frac{\ln n}{n}$. We have $p_x p_x = \Theta(\frac{\ln n}{n})$ except in the regions $x = o(n), y = o(n)$ and care must be taken that those terms don't contribute significantly to μ. Careful asymptotics (and first year Calculus!) yield

$$\mu \sim (100 \ln n) \int_0^1 \frac{dx}{\sqrt{x(1-x)}} = 100\pi \ln n$$

The negligible effect of the $x = o(n), y = o(n)$ terms reflects the finiteness of the indefinite integral at poles $x = 0$ and $x = 1$. The possible representations $x + y = n$ are mutually independent events so that from basic Large Deviation results

$$\Pr[|f(n) - \mu| > \epsilon\mu] < 2(1-\delta)^{\mu}$$

for constants ϵ, δ. To be specific we take $\epsilon = .9, \delta = .1$ and

$$\Pr[|f(n) - \mu| > .9\mu] < .9^{314 \ln n} < n^{-1.1}$$

for n sufficiently large. Take $c_1 < .1(100\pi)$ and $c_2 > 1.9(100\pi)$.

Let A_n be the event that $c_1 \ln n \leq f(n) \leq c_2 \ln n$ does *not* hold. We have $\Pr[A_n] < n^{-1.1}$ for n sufficiently large. The Borel Cantelli Lemma applies, almost always all A_n fail for n sufficiently large. Therefore there exists a specific point in the probability space, i.e., a specific set S, for which $c_1 \ln n \leq f(n) \leq c_2 \ln n$ for all sufficiently large n. □

Now for a given set S of natural numbers let $g(n) = g_S(n)$ denote the number of representations $n = x + y + z$, $x, y, z \in S$, all unequal.
Theorem 3.2. (Erdős, Tetali[1990]) There is a set S and a positive constants c_1, c_2 so that

$$c_1 \log n \leq g(n) \leq c_2 \log n$$

for all sufficiently large n.

The full result of Erdős and Tetali was that for each k there is a set S and constants c_1, c_2 so that the number of representations of n as the sum of k terms of S lies between $c_1 \log n$ and

Define a random subset $X \subseteq S$ by

$$\Pr[y \in X] = p_y = 10(\ln y)^{1/4} y^{-1/4}$$

for $y \in S$, $y \geq 10^8$. For definiteness say $\Pr[y \in X] = p_y = 1$ for $y \in S$, $y < 10^8$. Then

$$E[N_X(x)] = \sum_{i=0}^{x^{1/2}} \Pr[i^2 \in X] = O(x^{1/4}(\ln x)^{1/4})$$

and large deviation results give $N_X(x) = O(x^{1/4}(\ln x)^{1/4})$ almost always.

For any given $n \not\equiv 0(\mathrm{mod} 4)$, $n \geq 10^8$, let $\mathcal{F}_n$ denote the family of sets F of four squares adding to n. For each $F \in \mathcal{F}_n$ let A_F be the event $F \subseteq X$. We apply Janson's Inequality to give an upper bound to $\Pr[\wedge_{F \in \mathcal{F}_n} \overline{A}_F]$. Observe that this probability increases when the p_y decrease so, as the function p_y is decreasing in y, we may make the simplifying assumption

$$p_y = p = 10(\ln n)^{1/4} n^{-1/4}$$

for all $y \in S$, $y \leq n$. Then

$$\Pr[A_F] = p^4 = 10^4(\ln n)/n$$

and

$$\mu \geq (1 + o(1))(n/48)10^4(\ln n)/n \geq (100 + o(1))(\ln n)$$

Thus $e^{-\mu} < n^{-100+o(1)}$. The addends of Δ break into two parts, those $\Pr[A_F \wedge A_{F'}]$ with $|F \cap F'| = 1$ and those with $|F \cap F'| = 2$. The bounds on $r_3(n)$ give that there are at most $n^{3/2+o(1)}$ pairs F, F' of the first type and each has

$$\Pr[F \cap F'] = p^7 = n^{-7/4+o(1)}$$

The bounds on $r_2(n)$ give that there are at most $n^{1+o(1)}$ pairs F, F' of the second type and each has

$$\Pr[F \cap F'] = p^6 = n^{-3/2+o(1)}$$

Hence

$$\Delta \leq n^{3/2+o(1)-7/4+o(1)} + n^{1+o(1)-3/2+o(1)} = o(1)$$

Thus

$$\Pr[\wedge_{F \in \mathcal{F}_n} \overline{A}_F] \leq (1 + o(1))e^{-\mu} \leq n^{-100+o(1)}$$

As $\sum n^{-100+o(1)}$ converges the Borel-Cantelli lemma gives that almost always all sufficiently large $n \not\equiv 0(\mathrm{mod} 4)$ will be the sum of four elements of X.

Remark The constant "10" could be made smaller as long as the exponent of n here is less than -1.

Let X be a particular set having the above properties. (As customary, the probabilistic method does not actually "construct" X.) Suppose all $n \geq n_0$, $n \not\equiv 0(\mathrm{mod} 4)$ are the sum of four elements of X. Add to X all squares up to n_0. This does not affect the asymptotics of $N_X(x)$ and now all $n \not\equiv 0(\mathrm{mod} 4)$ are the sum of four elements of X. Finally, replace X by $X \cup 4X \cup 4^2X \cup 4^3X \cup \ldots$. This affects the asyptotics of $N_X(x)$ only by a constant and now *all* integers are the sum of four elements of X.

With p, q distinct primes, $X_p X_q = 1$ if and only if $p|x$ and $q|x$ which occurs if and only if $pq|x$. Hence

$$Cov[X_p, X_q] = E[X_p]E[X_q] - E[X_p X_q] = \frac{\lfloor n/pq \rfloor}{n} - \frac{\lfloor n/p \rfloor}{n}\frac{\lfloor n/q \rfloor}{n}$$

$$\leq \frac{1}{pq} - (\frac{1}{p} - \frac{1}{n})(\frac{1}{q} - \frac{1}{n}) \leq \frac{1}{n}(\frac{1}{p} + \frac{1}{q})$$

Thus

$$\sum_{p \neq q} Cov[X_p, X_q] \leq \frac{1}{n}\sum_{p \neq q} \frac{1}{p} + \frac{1}{q} = \frac{\pi(n) - 1}{n} \sum_p \frac{2}{p}$$

where $\pi(n) \sim \frac{n}{\ln n}$ is the number of primes $p \leq n$. So

$$\sum_{p \neq q} Cov[X_p, X_q] < \frac{(n/\ln n)}{n}(2 \ln\ln n) = o(1)$$

That is, the covariances do not affect the variance, $Var[X] \sim \ln\ln n$ and Chebyschev's Inequality actually gives

$$\Pr[|v(n) - \ln\ln n| > \lambda\sqrt{\ln\ln n}] < \lambda^{-2} + o(1)$$

for any constant λ. □

In a classic paper Paul Erdős and Marc Kac [1940] showed, essentially, that X does behave like a normal distribution with mean and variance $\ln\ln n + o(1)$. Here is their precise result.
The Erdős-Kac Theorem. Let λ be fixed, positive, negative or zero. Then

$$\lim_{n\to\infty} \frac{1}{n}|\{x : 1 \leq x \leq n, v(x) \geq \ln\ln n + \lambda\sqrt{\ln\ln n}\}| = \int_\lambda^\infty \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$$

We do not prove this result here.

2 Four Squares with Few Squares

The classic theorem of Lagrange states that every nonnegative integer n is the sum of four squares. How "sparse" can a set of squares be and still retain the four square property. For any set X of nonnegative integers set $N_X(x) = |\{i \in X, i \leq x\}|$. Let $S = \{0, 1, 4, 9, \ldots\}$ denote the squares. If $X \subseteq S$ and every $n \geq 0$ can be expressed as the sum of four elements of X then how slow can be the growth rate of $N_X(x)$? Clearly we must have $N_X(x) = \Omega(x^{1/4})$. Our object here is to give a quick proof of the following result of Wirsing.
Theorem. There is a set $X \subseteq S$ such that every $n \geq 0$ can be expressed as the sum of four elements of X and

$$N_X(x) = O(x^{1/4}(\ln x)^{1/4})$$

In 1828 Jacobi showed that the number $r_4(n)$ of solutions in integers to $n = a^2 + b^2 + c^2 + d^2$ is given by eight times the sum of those $d|n$ with $d \not\equiv 0(\text{mod} 4)$. In 1801 Gauss found an exact expression for the number $r_2(n)$ of solutions in integers to $n = a^2 + b^2$. We will need only $r_2(n) = n^{o(1)}$ which follows easily from his results. From this the number $r_3(n)$ of solutions to $n = a^2 + b^2 + c^2$ is $O(n^{1/2+o(1)})$. Now suppose $n \not\equiv 0(\text{mod} 4)$. Then $r_4(n) > 8n$ so, excluding order there are at least $n/48$ different solutions to $n = a^2 + b^2 + c^2 + d^2$ in nonnegative integers. From $r_2(n) = n^{o(1)}$ it follows that there are $O(n^{1/2+o(1)})$ solutions with $a = b$. Hence there are at least $(1 + o(1))n/48$ *sets* F of four squares adding to n .

$cl_K(x_1, \ldots, x_r)$ and Spoiler moved Inside. Hence $x_{r+1} \notin H_0$. Since $|H| \leq K \leq a$, H_0 lies inside $cl_a(x_1, \ldots, x_r)$. The isomorphism between $cl_a(x_1, \ldots, x_r)$ and $cl_a(y_1, \ldots, y_r)$ maps H_0 into a copy of itself in the graph G_2.

For any copy of H_0 in G_2, let $N(H_0)$ denote the number of extensions of H_0 to H. From Theorem 3.2 one can show that a.s all $N(H_0) = \Theta(n^{v-\alpha e})$, with $v = v(H_0, H)$, $e = e(H_0, H)$ and $v - \alpha e > 0$. For a given H_0 each y_{r+1} is in only a bounded number of copies of H since all copies of H lie in $cl_b(y_1, \ldots, y_r, y_{r+1})$. Hence there are $\Theta(n^{v-\alpha e})$ vertices y_{r+1} so that $cl_b(y_1, \ldots, y_r, y_{r+1})$ contains H. Arguing as with the first move there a.s. are $\Theta(n^{v-\alpha e})$, hence at least one, y_{r+1} with $cl_b(y_1, \ldots, y_r, y_{r+1}) \cong H$. Duplicator selects such a y_{r+1}.

Lecture 6: A Number Theory Interlude

We take a break from Graph Theory and explore applications of these methods to Number Theory.

1 Prime Factors

The second moment method is an effective tool in number theory. Let $v(n)$ denote the number of primes p dividing n. (We do not count multiplicity though it would make little difference.) The following result says, roughly, that "almost all" n have "very close to" $\ln \ln n$ prime factors. This was first shown by Hardy and Ramanujan in 1920 by a quite complicated argument. We give the proof of Paul Turan [1934] a proof that played a key role in the development of probabilistic methods in number theory.

Theorem 1.1 Let $\omega(n) \to \infty$ arbitrarily slowly. Then the number of x in $\{1, \ldots, n\}$ such that

$$|v(x) - \ln \ln n| > \omega(n)\sqrt{\ln \ln n}$$

is $o(n)$.

Proof. Let x be randomly chosen from $\{1, \ldots, n\}$. For p prime set

$$X_p = \begin{cases} 1 & \text{if } p|x \\ 0 & \text{otherwise} \end{cases}$$

and set $X = \sum X_p$, the summation over all primes $p \leq n$, so that $X(x) = v(x)$. Now

$$E[X_p] = \frac{\lfloor n/p \rfloor}{n}$$

As $y - 1 < \lfloor y \rfloor \leq y$

$$E[X_p] = 1/p + O(1/n)$$

By linearity of expectation

$$E[X] = \sum_{p \leq n} \frac{1}{p} + O(\frac{1}{n}) \sim \ln \ln n$$

Now we bound the variance

$$Var[X] \leq (1 + o(1)) \ln \ln n + \sum_{p \neq q} Cov[X_p, X_q]$$

Now we define the $a_1, \ldots, a_t$ of the lookahead strategy by reverse induction. We set $a_t = 0$. If at the end of the game Duplicator can assure that the 0-types of $x_1, \ldots, x_t$ and $y_1, \ldots, y_t$ are the same then they have the same induced subgraphs and he has won. Suppose, inductively, that $b = a_{r+1}$ has been defined. Let, applying the Lemma, K be a.s. an upper bound on all $cl_b(z_1, \ldots, z_{r+1})$. We then define $a = a_r$ by $a = K + b$.

Now we need show that a.s. this strategy works. Let $G_1 \sim G(n, n^{-\alpha})$, $G_2 \sim G(m, m^{-\alpha})$ and suppose Duplicator tries to play the $(a_1, \ldots, a_t)$ lookahead strategy on $EHR(G_1, G_2, t)$.

Set $a = a_1$ and consider the first move. Spoiler will select, say, $y = y_1 \in G_2$. Duplicator then must play $x = x_1 \in G_1$ with $cl_a(x) \cong cl_a(y)$. Can he always do so - that is, do a.s. G_1 and G_2 have the same values of $cl_a(x)$? The size of $cl_a(x)$ is a.s. bounded so it suffices to show for any potential H that either there almost surely is an x with $cl_a(x) \cong H$ or there almost surely is no x with $cl_a(x) \cong H$.

Let H have v vertices and e edges. Suppose H has a subgraph H' (possibly H itself) with v' vertices, e' edges and $v' - \alpha e' < 0$. The expected number of copies of H' in G_1 is

$$\Theta(n^{v'} p^{e'}) = \Theta(n^{v' - \alpha e'}) = o(1)$$

so a.s. G_1 contains no copy of H', hence no copy of H, hence no x with $cl_a(x) \cong H$. If this does not occur then (since, critically, α is irrational) all $v' - \alpha e' > 0$ so the expected number of copies of all such H' approaches infinity. From Theorem 1.4.5 a.s. G_1 has $\Theta(n^{v-\alpha e})$ copies of H. For x in appropriate position in such a copy of H we cannot deduce $cl_a(x) \cong H$ but only that $cl_a(x)$ contains H as a subgraph. (Essentially, x may have additional extension properties.) For each such x as $cl_a(x)$ is bounded, $cl_a(x)$ contains only a bounded number of copies of H. Hence there are $\Theta(n^{v-\alpha e})$ *different* $x \in G_1$ so that $cl_a(x)$ contains H as a subgraph.

Let H' be a possible value for $cl_a(x)$ that contains H as a subgraph. Let H' have v' vertices and e' edges. As (x, H') is rigid, (H, H') is dense and so

$$(v' - v) - \alpha(e' - e) < 0$$

There are $\Theta(n^{v' - \alpha e'})$ different x with $cl_a(x)$ containingH' but since $v' - \alpha e' < v - \alpha e$ this is $o(n^{v-\alpha e})$. Subtracting off such x for all the boundedly many such potential H' there a.s. remain $\Theta(n^{v-\alpha e})$, hence at least one, x with $cl_a(x) \cong H$.

Now, in general, consider the $(r+1)$-st move. We set $b = a_{r+1}$, $a = a_r$ for notational convenience and recall $a = K + b$ where K is an upper bound on $cl_b(z_1, \ldots, z_{r+1})$. Points $x_1, \ldots, x_r \in G_1$, $y_1, \ldots, y_r \in G_2$ have been selected with

$$cl_a(x_1, \ldots, x_r) \cong cl_a(y_1, \ldots, y_r)$$

Spoiler picks, say, $x_{r+1} \in G_1$. We distinguish two cases. We say Spoiler has moved Inside if

$$x_{r+1} \in cl_K(x_1, \ldots, x_r)$$

Otherwise we say Spoiler has moved Outside.

Suppose Spoiler moves Inside. Then

$$cl_b(x_1, \ldots, x_r, x_{r+1}) \subseteq cl_{K+b}(x_1, \ldots, x_r) = cl_a(x_1, \ldots, x_r)$$

The isomorphism from $cl_a(x_1, \ldots, x_r)$ to $cl_a(y_1, \ldots, y_r)$ sends x_{r+1} to some y_{r+1} which Duplicator selects.

Suppose Spoiler moves Outside. Set $H = cl_b(x_1, \ldots, x_r, x_{r+1})$. Let H_0 be the union of all rigid extensions of any size of $x_1, \ldots, x_r$ in H. If $x_{r+1} \in H_0$ then, as $|H| \leq K$, $x_{r+1} \in$

will show that he almost always wins with perfect play - it only indicates that the strategy used need be more complex. Now let us fix $\alpha \in (0,1)$, α irrational.

Now recall our notion of rooted graphs (R,H) but this time from the perspective of a particular $p = n^{-\alpha}$. We say (R,H) is *dense* if $v - e\alpha < 0$ and *sparse* if $v - e\alpha > 0$. The irrationality of α assures us that all (R,H) are in one of these categories. We call (R,H) *rigid* if for all S with $R \subseteq S \subset V(H)$, (S,H) is dense.

For any r,t there is a finite list (up to isomorphism) of rigid rooted graphs (R,H) containing r roots and with $v(R,H) \le t$. In any graph G we define the t-closure $cl_t(x_1,\ldots,x_r)$ to be the union of all $y_1,\ldots,y_v$ with (crucially) $v \le t$ which form an (R,H) extension where (R,H) is rigid. If there are no such sets we define the default value $cl_t(x_1,\ldots,x_r) = \{x_1,\ldots,x_r\}$. We say two sets $x_1,\ldots,x_r$ and $x_1',\ldots,x_r'$ have the same t-type if their t-closures are isomorphic. (To be precise, these are ordered r-tuples and the isomorphism must send x_i into x_i'.)
Example. Taking $\alpha \sim .51$ (but irrational, of course), $cl_1(x_1,x_2)$ consists of x_1,x_2 and all y adjacent to both of them. $cl_3(x_1,x_2)$ has those points and all y_1,y_2,y_3 which together with x_1 form a K_4 (note that this gives an (R,H) with $v=3, e=6$) and a finite number of other possibilities.

We can already describe the nature of Duplicator's strategy. At the end of the r-th move, with $x_1,\ldots,x_r$ and $y_1,\ldots,y_r$ having been selected from the two graphs, Duplicator will assure that these sets have the same $a_r - type$. We shall call this the $(a_1,\ldots,a_t)$ *lookahead strategy.* Here a_r must depend only on t, the total number of moves in the game and α. We shall set $a_t = 0$ so that at the end of the game, if Duplicator can stick to the $(a_1,\ldots,a_t)$ lookahead strategy then he has won. If, however, Spoiler picks, say, x_r so that there is no corresponding y_r with $x_1,\ldots,x_r$ and $y_1,\ldots,y_r$ having the same a_r-type then the strategy fails and we say that Spoiler wins. The values a_r give the "lookahead" that Duplicator uses but before defining them we need some preliminary results.
Lemma 4.6 Let α, $r,t > 0$ be fixed. Then there exists $K = K(\alpha,r,t)$ so that in $G(n,n^{-\alpha})$ a.s.

$$|cl_t(x_1,\ldots,x_r)| \le K$$

for *all* $x_1,\ldots,x_r \in G$.
Proof. Set $K = r + t(L-1)$. If $X = \{x_1,\ldots,x_r\}$ has t-closure with more than K points then there will be L sets $Y^1,\ldots,Y^L$ disjoint from X, all $|Y^j| \le t$ so that each $(X, X\cup Y^j)$ forms a rigid extension and with each Y^j having at least one point not in $Y^1 \cup \ldots Y^{j-1}$. Begin with X and add the Y^j in order. Adding Y^j will add, say, v_j vertices and e_j edges. Since $(X, X\cup Y^j)$ was *rigid*, $(X\cup Y^1\cup\ldots\cup Y^{j-1}, X\cup Y^1\cup\ldots\cup Y^j)$ is dense and so $v_j - e_j\alpha < 0$. As $v_j \le t$ there are only a finite number of possible values of $v_j - e_j\alpha$ and so there is an $\epsilon = \epsilon(\alpha,r,t)$ so that all $v_j - e_j\alpha \le -\epsilon$. Pick L (and therefore K) so that $r - L\epsilon < 0$. The existence of a t-closure of size greater than K would imply the existence in $G(n,n^{-\alpha})$ of one of a finite number of graphs that would have some $r + v_1 + \ldots + v_L$ vertices and at least $e_1 + \ldots + e_L$ edges. But the probability of G containing such a graph is bounded by

$$\begin{aligned} n^{r+v_1+\ldots+v_L}p^{e_1+\ldots+e_L} &= n^{r+v_1+\ldots+v_L-\alpha(e_1+\ldots+e_L)} \\ n^{r+(v_1-\alpha e_1)+\ldots+(v_L-\alpha e_L)} &\le n^{r-L\epsilon} \\ = o(1) \end{aligned}$$

so a.s. no such t-closures exist. □
Remark. The value of K given by the above proof depends strongly on how close α may be approximated by rationals of denominator at most t. This is often the case. If, for example, $\frac{1}{2} + \frac{1}{s+1} < \alpha < \frac{1}{2} + \frac{1}{s}$ then a.s. there will be two points $x_1,x_2 \in G(n,n^{-\alpha})$ having s common neighbors so that $|cl_1(x_1,x_2)| = s+2$.

We say that a graph G has the full level s extension property if for every distinct $u_1, \ldots, u_a$ and $v_1, \ldots, v_b$ i G with $a + b \leq s$ there is an $x \in V(G)$ with $\{x, u_i\} \in E(G)$, $1 \leq i \leq a$ and $\{x, v_j\} \notin V(G)$, $1 \leq j \leq b$. Suppose that G, H both have the full level $s-1$ extension property. Then Duplicator wins $EHR[G, H, s]$ by the following simple strategy. On the i-th round, with $x_1, \ldots, x_{i-1}, y_1, \ldots, y_{i-1}$ already selected, and Spoiler picking, say, x_i, Duplicator simply picks y_i having the same adjacencies to the $y_j, j < i$ as x_i has to the $x_j, j < i$. The full extension property says that such a y_i will surely exist.

Theorem 4.5 For any fixed p, $0 < p < 1$, and any s, $G(n, p)$ almost always has the full level s extension property.

Proof. For every distinct $u_1, \ldots, u_a, v_1, \ldots, v_b, x \in G$ with $a + b \leq s$ let $E_{u_1,\ldots,u_a,v_1,\ldots,v_b,x}$ be the event that $\{x, u_i\} \in E(G)$, $1 \leq i \leq a$ and $\{x, v_j\} \notin V(G)$, $1 \leq j \leq b$. Then

$$\Pr[E_{u_1,\ldots,u_a,v_1,\ldots,v_b,x}] = p^a(1-p)^b$$

Now define

$$E_{u_1,\ldots,u_a,v_1,\ldots,v_b} = \wedge_x \overline{E_{u_1,\ldots,u_a,v_1,\ldots,v_b,x}}$$

the conjunction over $x \neq u_1, \ldots, u_a, v_1, \ldots, v_b$. But these events are mutually independent over x since they involve different edges. Thus

$$\Pr[\wedge_x \overline{E_{u_1,\ldots,u_a,v_1,\ldots,v_b,x}}] = [1 - p^a(1-p)^b]^{n-a-b}$$

Set $\epsilon = \min(p, 1-p)^s$ so that

$$\Pr[\wedge_x \overline{E_{u_1,\ldots,u_a,v_1,\ldots,v_b,x}}] \leq (1-\epsilon)^{n-s}$$

The key here is that ϵ is a fixed (dependent on p, s) positive number. Set

$$E = \vee E_{u_1,\ldots,u_a,v_1,\ldots,v_b}$$

the disjunction over all distinct $u_1, \ldots, u_a, v_1, \ldots, v_b \in G$ with $a + b \leq s$. There are less than $s^2 n^s$ such choices as we can choose a, b and then the vertices. Thus

$$\Pr[E] \leq s^2 n^s (1-\epsilon)^{n-s}$$

But

$$\lim_{n \to \infty} s^2 n^s (1-\epsilon)^{n-s} = 0$$

and so E holds almost never. Thus $\neg E$, which is precisely the statement that $G(n, p)$ has the full level s extension property, holds almost always. □

But now we have proven Theorem 4.1. For any $p \in (0, 1)$ and any fixed s as $m, n \to \infty$ with probability approaching one both $G(n, p)$ and $H(m, p)$ will have the full level s extension property and so Duplicator will win $EHR[G(n, p), H(m, p), s]$.

Why can't Duplicator use this strategy when $p = n^{-\alpha}$? We illustrate the difficulty with a simple example. Let $.5 < \alpha < 1$ and let Spoiler and Duplicator play a three move game on G, H. Spoiler thinks of a point $z \in G$ but doesn't tell Duplicator about it. Instead he picks $x_1, x_2 \in G$, both adjacent to z. Duplicator simply picks $y_1, y_2 \in H$, either adjacent or not adjacent dependent on whether $x_1 \sim x_2$. But now wily Spoiler picks $x_3 = z$. $H \sim H(m, m^{-\alpha})$ does not have the full level 2 extension property. In particular, most pairs y_1, y_2 do not have a common neighbor. Unless Duplicator was lucky, or shrewd, he then cannot find $y_3 \sim y_1, y_2$ and so he loses. This example does not say that Duplicator will lose with perfect play - indeed, we that *at* a threshold function the Zero-One Law will not hold and so to say that $p(n)$ satisfies the Zero-One Law is to say that $p(n)$ is not a threshold function - that it is a boring place in the evolution of the random graph, at least through the spectacles of the First Order language. In stark terms: What happens in the evolution of $G(n,p)$ at $p = n^{-\pi/7}$? The answer: Nothing!

Our approach to Zero-One Laws will be through a variant of the Ehrenfeucht Game, which we now define. Let G, H be two vertex disjoint graphs and t a positive integer. We define a perfect information game, denoted $EHR[G,H,t]$, with two players, denoted Spoiler and Duplicator. The game has t rounds. Each round has two parts. First the Spoiler selects either a vertex $x \in V(G)$ or a vertex $y \in V(H)$. He chooses which graph to select the vertex from. Then the Duplicator must select a vertex in the other graph. At the end of the t rounds t vertices have been selected from each graph. Let $x_1, \ldots, x_t$ be the vertices selected from $V(G)$ and $y_1, \ldots, y_t$ be the vertices selected from $V(H)$ where x_i, y_i are the vertices selected in the i-th round. Then Duplicator wins if and only if the induced graphs on the selected vertices are order-isomorphic: i.e., if for all $1 \le i < j \le t$

$$\{x_i, x_j\} \in E(G) \longleftrightarrow \{y_i, y_j\} \in E(H)$$

As there are no hidden moves and no draws one of the players must have a winning strategy and we will say that that player wins $EHR[G,H,t]$.
Lemma 4.3 For every First Order A there is a $t = t(A)$ so that if G, H are any graphs with $G \models A$ and $H \models \neg A$ then Spoiler wins $EHR[G,H,t]$.

A detailed proof would require a formal analysis of the First Order language so we give only an example. Let A be the property $\forall_x \exists_y [x \sim y]$ of not containing an isolated point and set $t = 2$. Spoiler begins by selecting an isolated point $y_1 \in V(H)$ which he can do as $H \models \neg A$. Duplicator must pick $x_1 \in V(G)$. As $G \models A$, x_1 is not isolated so Spoiler may pick $x_2 \in V(G)$ with $x_1 \sim x_2$ and now Duplicator cannot pick a "duplicating" y_2.
Theorem 4.4 A function $p = p(n)$ satisfies the Zero-One Law if and only if for every t, letting $G(n, p(n)), H(m, p(m))$ be independently chosen random graphs on disjoint vertex sets

$$\lim_{m,n\to\infty} \Pr[\text{ Duplicator wins} EHR[G(n,p(n)), H(m,p(m)), t]] = 1$$

Remark. For any given choice of G, H somebody must win $EHR[G,H,t]$. (That is, there is no random play, the play is perfect.) Given this probability distribution over (G,H) there will be a probability that $EHR[G,H,t]$ will be a win for Duplicator, and this must approach one.
Proof. We prove only the "if" part. Suppose $p = p(n)$ did not satisfy the Zero-One Law. Let A satisfy

$$\lim_{n\to\infty} \Pr[G(n,p(n)) \models A] = c$$

with $0 < c < 1$. Let $t = t(A)$ be as given by the Lemma. With limiting probability $2c(1-c) > 0$ exactly one of $G(n,p(n)), H(n,p(n))$ would satisfy A and thus Spoiler would win, contradicting the assumption. This is not a full proof since when the Zero-one Law is not satisfied $\lim_{n\to\infty} \Pr[G(n,p(n)) \models A]$ might not exist. If there is a subsequence n_i on which the limit is $c \in (0,1)$ we may use the same argument. Otherwise there will be two subsequences n_i, m_i on which the limit is zero and one respectively. Then letting $n, m \to \infty$ through n_i, m_i respectively, Spoiler will win $EHR[G,H,t]$ with probability approaching one. $\square$

Theorem 4.4 provides a bridge from Logic to Random Graphs. To prove that $p = p(n)$ satisfies the Zero-One Law we now no longer need to know anything about Logic - we just have to find a good strategy for the Duplicator.

extension. Set $\mu = \binom{n-r}{v}p^e$, the expected value of N in $G(n,p)$.
Theorem 3.2. Let (R,H) be strictly balanced. Then for all $\epsilon > 0$ there exists K so that if $p = p(n)$ is such that $\mu > K \log n$ then almost surely

$$|N(x_1,\ldots,x_r) - \mu| < \epsilon\mu$$

for all $x_1,\ldots,x_r$.

In particular if $\mu \gg \log n$ then almost surely all $N(x_1,\ldots,x_r) \sim \mu$.

4 Zero-One Laws

In this section we restrict our attention to graph theoretic properties expressible in the First Order theory of graphs. The language of this theory consists of variables $(x,y,z,\ldots)$, which always represent vertices of a graph, equality and adjacency $(x = y, x \sim y)$, the usual Boolean connectives $(\wedge, \neg, \ldots)$ and universal and existential quanfication $(\forall_x, \exists_y)$. Sentences must be finite. As examples, one can express the property of containing a triangle

$$\exists_x \exists_y \exists_z [x \sim y \wedge x \sim z \wedge y \sim z]$$

having no isolated point

$$\forall_x \exists_y [x \sim y]$$

and having radius at most two

$$\exists_x \forall_y [\neg(y = x) \wedge \neg(y \sim x) \longrightarrow \exists_z [z \sim y \wedge y \sim x]]$$

For any property A and any n,p we consider the probability that the random graph $G(n,p)$ satisfies A, denoted

$$\Pr[G(n,p) \models A]$$

Our objects in this section will be the theorem of Glebskii et.al. [1969] and independently Fagin[1976]
Theorem 4.1 For any fixed p, $0 < p < 1$ and any First Order A

$$\lim_{n\to\infty} \Pr[G(n,p) \models A] = 0 \text{ or } 1$$

and that of Shelah and Spencer[1988]
Theorem 4.2 For any *irrational* α, $0 < \alpha < 1$, setting $p = p(n) = n^{-\alpha}$

$$\lim_{n\to\infty} \Pr[G(n,p) \models A] = 0 \text{ or } 1$$

Both proofs are only outlined.

We shall say that a function $p = p(n)$ satisfies the Zero-One Law if the above equality holds for every First Order A.

The Glebskii/Fagin Theorem has a natural interpretation when $p = .5$ as then $G(n,p)$ gives equal weight to every (labelled) graph. It then says that any First Order property A holds for either almost all graphs or for almost no graphs. The Shelah/Spencer Theorem may be interpreted in terms of threshold functions. For example, $p = n^{-2/3}$ is a threshold function for containment of a K_4. That is, when $p \ll n^{-2/3}$, $G(n,p)$ almost surely does not contain a K_4 whereas when $p \gg n^{-2/3}$ it almost surely does contain a K_4. In between, say at $p = n^{-2/3}$, the probability is between 0 and 1, in this case $1 - e^{-1/24}$. The (admittedly rough) notion is

3 All Vertices in nearly the same number of Triangles

Returning to the example of §1, let $N(x)$ denote the number of triangles containing vertex x. Set $\mu = \binom{n-1}{2}p^3$ as before.

Theorem 3.1. For every $\epsilon > 0$ there exists K so that if $p = p(n)$ is such that $\mu = K \log n$ then almost surely

$$(1-\epsilon)\mu < N(x) < (1+\epsilon)\mu$$

for *all* vertices x.

We shall actually show that for a given vertex x

$$\Pr[|N(x) - \mu| > \epsilon\mu] = o(n^{-1})$$

If the distribution of $N(x)$ were Poisson with mean μ then this would follow by Large Deviation results and indeed our approach will show that $N(x)$ is closely approximated by the Poisson distribution.

We call F a maximal disjoint family of extensions if F consists of pairs $\{x_i, y_i\}$ such that all x, x_i, y_i are triangles in $G(n,p)$, the x_i, y_i are all distinct, and there is no $\{x', y'\}$ with x, x', y' a triangle and x', y' both distinct from all the x_i, y_i. Let $Z^{(s)}$ denote the number of maximal disjoint families of size s. Lets restrict $0 \le s \le \log^2 n$ (a technical convenience) and bound $E[Z^{(s)}]$. There are $\sim \binom{n-1}{2}^s/s!$ choices for F. Each has probability $(p^3)^s$ that all x_i, y_i do indeed give extensions. We further need that the $n-1-2s \sim n$ other vertices contain no extension. The calculation of §1 may be carried out here to show that this probability is $\sim e^{-\mu}$. All together

$$E[Z^{(s)}] \le (1+o(1))\frac{\mu^s}{s!}e^{-\mu}$$

But now the right hand side is asymptotically the Poisson distribution so that we can choose K so that

$$\sum * E[Z^{(s)}] = o(n^{-1}) \qquad (*)$$

where $\sum^*$ is over $s < \log^2 n$ with $|s - \mu| > \epsilon\mu$.

When $s > \log^2 n$ we ignore the condition that F be maximal so that $E[Z^{(s)}] < \mu^s/s! = o(n^{-10})$, say. Thus $(*)$ holds with $\sum^*$ over all s with $|s-\mu| > \epsilon\mu$. Thus with probability $1 - o(n^{-1})$ all maximal disjoint families of extensions F have $|s - \mu| < \epsilon\mu$. But there *must* be some maximal disjoint family of extensions. Thus with probability $1 - o(n^{-1})$ there is a maximal disjoint family of extensions F with $|s-\mu| < \epsilon\mu$. As F consists of extensions

$$\Pr[N(x) < (1-\epsilon)\mu] = o(n^{-1})$$

To complete the upper bound we need show that $N(x)$ will not be much larger than $|F|$. Here we use only that $p = n^{-2/3+o(1)}$. There is $o(n^{-1})$ probability that $G(n,p)$ has an edge $\{x, x'\}$ lying in ten triangles. There is a $o(n^{-1})$ that $G(n,p)$ has a vertex x with $u_i, v_i, w_i, 1 \le i \le 7$ all distinct and all x, u_i, v_i and x, v_i, w_i triangles. When these do not occur $N(x) \le |F| + 70$ for any maximal disjoint family of extensions $|F|$ and so for any $\epsilon' > \epsilon$

$$\Pr[N(x) > (1+\epsilon')\mu] < o(n^{-1}) + \Pr[\text{some } |F| > (1+\epsilon)\mu] = o(n^{-1})$$

With some additional work one can find K so that the conclusions of the theorem hold for any $p = p(n)$ with $\mu > K \log n$. The general result is stated in terms of rooted graphs. For a given rooted graph (R, H) let $N(x_1, \ldots, x_r)$ denote the number of $(y_1, \ldots, y_v)$ giving an (R, H)

where $x_1,\ldots,x_r$ are some particular vertices. But

$$C_{x_1} \wedge \ldots \wedge C_{x_r} = \wedge \overline{B_{x_i y z}},$$

the conjunction over $1 \leq i \leq r$ and all y,z. We apply Janson's Inequality to this conjunction. Again $\epsilon = p^3 = o(1)$. The number of $\{x_i, y, z\}$ is $r\binom{n-1}{2} - O(n)$, the overcount coming from those triangles containing two (or three) of the x_i. (Here it is crucial that r is fixed.) Thus

$$\sum \Pr[B_{x_i y z}] = p^3\left(r\binom{n-1}{2} - O(n)\right) = r\mu + O(n^{-1+o(1)})$$

As before Δ is p^5 times the number of pairs $x_i yz \sim x_j y'z'$. There are $O(rn^3) = O(n^3)$ terms with $i=j$ and $O(r^2n^2) = O(n^2)$ terms with $i \neq j$ so again $\Delta = o(1)$. Therefore

$$\Pr[C_{x_1} \wedge \ldots \wedge C_{x_r}] \sim e^{-r\mu}$$

and

$$E[X^{(r)}/r!] \sim \frac{(ne^{-\mu})^r}{r!} = \frac{c^r}{r!}$$

Hence X has limiting Poisson distribution, in particular $\Pr[X=0] \to e^{-\mu}$. $\Box$

2 Rooted Graphs

The above result was only a special case of a general result of Spencer[1990] which we now state. By a *rooted graph* is meant a pair (R,H) consisting of a graph $H = (V(H), E(H))$ and a specified proper subset $R \subset V(H)$ of vertices called the roots. For convenience let the vertices of H be labelled $a_1,\ldots,a_r,b_1,\ldots,b_v$ with $R = \{a_1,\ldots,a_r\}$. In a graph G we say that vertices $y_1,\ldots,y_v$ make an (R,H)-extension of vertices $x_1\ldots,x_r$ if all these vertices are distinct; y_i, y_j are adjacent in G whenever b_i, b_j are adjacent in H; and x_i, y_j are adjacent in G whenever a_i, b_j are adjacent in H. So G on $x_1,\ldots,x_r,y_1,\ldots,y_v$ gives a copy of H which may have additional edges – except that edges between the x's are not examined. We let $Ext(R,H)$ be the property the for all $x_1,\ldots,x_r$ there exist $y_1,\ldots,y_v$ giving an (R,H) extension. For example, when H is a triangle and R one vertex $Ext(R,H)$ is the statement that every vertex lies in a triangle. When H is a path of length t and R the endpoints $Ext(R,H)$ is the statement that every pair of vertices lie on a path of length t. When $R = \emptyset$ $Ext(\emptyset, H)$ is the already examined statement that there exists a copy of H. As in that situation we have a notion of balanced and strictly balanced. We say (R,H) has type (v,e) where v is the number of nonroot vertices and e is the number of edges of H, not counting edges with both vertices in R. For every S with $R \subset S \subseteq V(H)$ let (v_S, e_S) be the type of $(R, H|_S)$. We call (R,H) balanced if $e_S/v_S \leq e/v$ for all such S and we call (R,H) strictly balanced if $e_S/v_S < e/v$ for all proper $S \subset V(H)$. We call (R,H) nontrivial if every root is adjacent to at least one nonroot.

Theorem 2.1. Let (R,H) be a nontrivial strictly balanced rooted graph with type (v,e) and $r = |R|$. Let c_1 be the number of graph automorphism $\sigma : V(H) \to V(H)$ with $\sigma(x) = x$ for all roots x. Let c_2 be the number of bijections $\sigma : R \to R$ which are extendable to some graph automorphism $\lambda : V(H) \to V(H)$. Let $\mu > 0$ be arbitrary and fixed. Let $p = p(n)$ satisfy

$$\frac{n^v p^e}{c_1} = \ln\left(\frac{n^r}{c_2\mu}\right)$$

Then

$$\lim_{n\to\infty} \Pr[G(n,p) \models Ext(R,H)] = e^{-\mu}$$

While the counting of automorphisms leads to some technical complexities the proof is essentially that of the "every vertex in a triangle" case.

Lecture 5: Counting Extensions and Zero-One Laws

The threshold behavior for the existence of a copy of H in $G(n,p)$ is well understood. Now we turn to what, in a logical sense, is the next level which we call *extension statements*. We want $G(n,p)$ to have the property that every $x_1,\ldots,x_r$ belong to a copy of H. For example ($r=1$), every vertex lies in a triangle. We find the fine threshold behavior for this property and further show - continuing this example - that for p a bit larger almost surely every vertex lies in about the same number of triangles.

1 Every Vertex in a Triangle

Let A be the property that every vertex lies in a triangle.
Theorem 1.1. Let $c>0$ be fixed and let $p=p(n)$, $\mu=\mu(n)$ satisfy

$$\binom{n-1}{2}p^3=\mu$$

$$e^{-\mu}=\frac{c}{n}$$

Then

$$\lim_{n\to\infty}\Pr[G(n,p)\models A]=e^{-c}$$

Proof. First fix $x\in V(G)$. For each unordered $y,z\in V(G)-\{x\}$ let B_{xyz} be the event that $\{x,y,z\}$ is a triangle of G. Let C_x be the event $\wedge\overline{B_{xyz}}$ and X_x the corresponding indicator random variable. We use Janson's Inequality to bound $E[X_x]=\Pr[C_x]$. Here $p=o(1)$ so $\epsilon=o(1)$. $\sum Pr[B_{xyz}]=\mu$ as defined above. Dependency $xyz\sim xuv$ occurs if and only if the sets overlap (other than in x). Hence

$$\Delta=\sum_{y,z,z'}\Pr[B_{xyz}\wedge B_{xyz'}]=O(n^3)p^5=o(1)$$

since $p=n^{-2/3+o(1)}$. Thus

$$E[X_x]\sim e^{-\mu}=\frac{c}{n}$$

Now define

$$X=\sum_{x\in V(G)}X_x,$$

the number of vertices x not lying in a triangle. Then from Linearity of Expectation

$$E[X]=\sum_{x\in V(G)}E[X_x]\to c$$

We need show that the Poisson Paradigm applies to X. To do this we show that all moments of X are the same as for the Poisson distribution. Fix r. Then

$$E[X^{(r)}/r!]=S^{(r)}=\sum\Pr[C_{x_1}\wedge\ldots\wedge C_{x_r}],$$

the sum over all sets of vertices $\{x_1,\ldots,x_r\}$. All r-sets look alike so

$$E[X^{(r)}/r!]=\binom{n}{r}\Pr[C_{x_1}\wedge\ldots\wedge C_{x_r}]\sim\frac{n^r}{r!}\Pr[C_{x_1}\wedge\ldots\wedge C_{x_r}]$$

3 Some Very Low Probabilities

Let A be the property that G does not contain K_4 and consider $\Pr[G(n,p) \models A]$ as p varies. (Results with K_4 replaced by an arbitrary H are discussed at the end of this section.) We know that $p = n^{-2/3}$ is a threshold function so that for $p \gg n^{-2/3}$ this probability is $o(1)$. Here we want to estimate that probability. Our estimates here will be quite rough, only up to a $o(1)$ additive factor in the hyperexponent, though with more care the bounds differ by "only" a constant factor in the exponent. If we were to consider all potential K_4 as giving mutually independent events then we would be led to the estimate $(1-p^6)^{\binom{n}{4}} = e^{-n^{4+o(1)}p^6}$. For p appropriately small this turns out to be correct. But for, say, $p = \frac{1}{2}$ it would give the estimate $e^{-n^{4+o(1)}}$. This must, however, be way off the mark since with probability $2^{-\binom{n}{2}} = e^{-n^{2+o(1)}}$ the graph G could be empty and hence trivially satisfy A.

Rather than giving the full generality we assume $p = n^{-\alpha}$ with $\frac{2}{3} > \alpha > 0$. The result is:

$$\Pr[G(n,p) \models A] = e^{-n^{4-6\alpha+o(1)}}$$

for $\frac{2}{3} > \alpha \geq \frac{2}{5}$ and

$$\Pr[G(n,p) \models A] = e^{-n^{2-\alpha+o(1)}}$$

for $\frac{2}{5} \geq \alpha > 0$.

The upper bound follows from the inequality

$$\Pr[G(n,p) \models A] \geq \max\left[(1-p^6)^{\binom{n}{4}}, (1-p)^{\binom{n}{2}}\right]$$

This is actually two inequalities. The first comes from the probability of G not containing a K_4 being at most the probability as if all the potential K_4 were independent. The second is the same bound on the probability that G doesn't contain a K_2 - i.e., that G has no edges. Calculation shows that the "turnover" point for the two inequalities occurs when $p = n^{-2/5+o(1)}$.

The upper bound follows from the Janson inequalities. For each four set α of vertices B_α is that that 4-set gives a K_4 and we want $\Pr[\wedge \overline{B_\alpha}]$. We have $\mu = \Theta(n^4p^6)$ and $-\ln M \sim \mu$ and (as shown in Lecture 1) $\Delta = \Theta(\mu\Delta^*)$ with $\Delta^* = \Theta(n^2p^5 + np^3)$. With $p = n^{-\alpha}$ and $\frac{2}{3} > \alpha > \frac{2}{5}$ we have $\Delta^* = o(1)$ so that

$$\Pr[\wedge \overline{B_\alpha}] \leq e^{-\mu(1+o(1))} = e^{-n^{4-6\alpha+o(1)}}$$

When $\frac{2}{5} > \alpha > 0$ then $\Delta^* = \Theta(n^2p^5)$ (somewhat surprisingly the np^3 never is significant in these calculations) and the extended Janson inequality gives

$$\Pr[\wedge \overline{B_\alpha}] \leq e^{-\Theta(\mu^2/\Delta)} = e^{-\Theta(\mu/\Delta^*)} = e^{-n^{2-\alpha}}$$

The general result has been found by T. Łuczak, A. Rucinski and S. Janson. Let H be any fixed graph and let A be the property of not containing a copy of H. For any subgraph H' of H the correlation inequality gives

$$\Pr[G(n,p) \models A] \leq e^{-E[X_{H'}]}$$

where $X_{H'}$ is the number of copies of H' in G. Now let $p = n^{-\alpha}$ where we restrict to those α for which p is past the threshold function for the appearance of H. Then

$$\Pr[G(n,p) \models A] = e^{n^{o(1)}} \min_{H'} e^{-E[X_{H'}]}$$

so that

$$f(k) > n^{3+o(1)}$$

Now we use the Generalized Janson Inequality to estimate $\Pr[\omega(G) < k]$. Here $\mu = f(k)$. (Note that Janson's Inequality gives a lower bound of $2^{-f(k)} = 2^{-n^{3+o(1)}}$ to this probability but this is way off the mark since with probability $2^{-\binom{n}{2}}$ the random G is empty!) The value Δ was examined in Lecture 2 and we showed

$$\frac{\Delta}{\mu^2} = \frac{\Delta^*}{\mu} = \sum_{i=2}^{k-1} g(i)$$

There $g(2) \sim k^4/n^2$ and $g(k-1) \sim 2kn2^{-k}/\mu$ were the dominating terms. In our instance $\mu > n^{3+o(1)}$ and $2^{-k} = n^{-2+o(1)}$ so $g(2)$ dominates and

$$\Delta \sim \frac{\mu^2 k^4}{n^2}$$

Hence we bound the *clique* number probability

$$\Pr[\omega(G) < k] < e^{-\mu^2(1+o(1))/2\Delta} = e^{-(n^2/k^4)(1+o(1))} = e^{-n^{2+o(1)}}$$

as $k = \Theta(\ln n)$. (The possibility that G is empty gives a lower bound so that we may say the probability is $e^{-n^{2+o(1)}}$, though a $o(1)$ in the hyperexponent leaves lots of room.)
Theorem 2.1. (Bollobás [1988]) Almost always

$$\chi(G)) \sim \frac{n}{2\log_2 n}$$

Proof. The argument that

$$\chi(G) \geq \frac{n}{\alpha(G)} \geq \frac{n}{2\log_2 n}(1+o(1))$$

almost always was given in Lecture 2.

The reverse inequality was an open question for a full quarter century! Set $m = \lfloor n/\ln^2 n \rfloor$. For any set S of m vertices the restriction $G|_S$ has the distribution of $G(m, 1/2)$. Let $k = k(m) = k_0(m) - 4$ as above. Note

$$k \sim 2\log_2 m \sim 2\log_2 n$$

Then

$$\Pr[\alpha[G|_S] < k] < e^{-m^{2+o(1)}}$$

There are $\binom{n}{m} < 2^n = 2^{m^{1+o(1)}}$ such sets S. Hence

$$\Pr[\alpha[G|_S] < k \text{ for some } m\text{-set } S] < 2^{m^{1+o(1)}} e^{-m^{2+o(1)}} = o(1)$$

That is, almost always *every* m vertices contain a k-element independent set.

Now suppose G has this property. We pull out k-element independent sets and give each a distinct color until there are less than m vertices left. Then we give each point a distinct color. By this procedure

$$\chi(G) \leq \lceil \frac{n-m}{k} \rceil + m \leq \frac{n}{k} + m = \frac{n}{2\log_2 n}(1+o(1)) + o(\frac{n}{\log_2 n})$$

$$= \frac{n}{2\log_2 n}(1+o(1))$$

and this occurs for almost all G. □

Observe that for this n the left hand side is $1+o(1)$. Note that $\binom{n}{k}$ grows, in n, like n^k. For any $\lambda \in (-\infty, +\infty)$ if

$$n = n_0(k)[1 + \frac{\lambda + o(1)}{k}]$$

then

$$\binom{n}{k} 2^{-\binom{k}{2}} = [1 + \frac{\lambda + o(1)}{k}]^k = e^{\lambda} + o(1)$$

and so

$$\Pr[\omega(G(n,p)) < k] = e^{-e^{\lambda}} + o(1)$$

As λ ranges from $-\infty$ to $+\infty$, $e^{-e^{\lambda}}$ ranges from 1 to 0. As $n_0(k+1) \sim \sqrt{2} n_0(k)$ the ranges will not "overlap" for different k. More precisely, let K be arbitrarily large and set

$$I_k = [n_0(k)[1 - \frac{K}{k}], n_0(k)[1 + \frac{K}{k}]]$$

For $k \geq k_0(K)$, $I_{k-1} \cap I_k = \emptyset$. Suppose $n \geq n_0(k_0(K))$. If n lies between the intervals (which occurs for "most" n), which we denote by $I_k < n < I_{k+1}$, then

$$\Pr[\omega(G(n,p)) < k] \leq e^{-e^{K}} + o(1),$$

nearly zero, and

$$\Pr[\omega(G(n,p)) < k+1] \geq e^{-e^{-K}} + o(1),$$

nearly one, so that

$$\Pr[\omega(G(n,p)) = k] \geq e^{-e^{-K}} - e^{-e^{K}} + o(1),$$

nearly one. When $n \in I_k$ we still have $I_{k-1} < n < I_{k+1}$ so that

$$\Pr[\omega(G(n,p)) = k \textit{ or } k-1] \geq e^{-e^{-K}} - e^{-e^{K}} + o(1),$$

nearly one. As K may be made arbitrarily large this yields the celebrated two point concentration theorem on clique number given as Corollary 2.1.2. Note, however, that for most n the concentration of $\omega(G(n, 1/2))$ is actually on a single value!

2 Chromatic Number

Again fix $p = 1/2$ (there are similar results for other p) and let $G \sim G(n, \frac{1}{2})$. We shall find bounds on the chromatic number $\chi(G)$. The original proof of Bollobás used martingales and will be discussed later. Set

$$f(k) = \binom{n}{k} 2^{-\binom{k}{2}}$$

Let $k_0 = k_0(n)$ be that value for which

$$f(k_0 - 1) > 1 > f(k_0)$$

Then $n = \sqrt{2}^{k(1+o(1))}$ so for $k \sim k_0$,

$$f(k+1)/f(k) = \frac{n}{k} 2^{-k}(1 + o(1)) = n^{-1+o(1)}$$

Set

$$k = k(n) = k_0(n) - 4$$

There are $O(n^{2v-j})$ choices of α, β intersecting in j points since α, β are determined, except for order, by $2v - j$ points. For each such α, β

$$\Pr[B_\alpha \wedge B_\beta] = p^{|A_\alpha \cup A_\beta|} = p^{2e - |A_\alpha \cap A_\beta|} \leq p^{2e - f_j}$$

Thus

$$\Delta = \sum_{j=2}^{v} O(n^{2v-j}) O(n^{-\frac{v}{e}(2e - f_j)})$$

But

$$2v - j - \frac{v}{e}(2e - f_j) = \frac{vf_j}{e} - j < 0$$

so each term is $o(1)$ and hence $\Delta = o(1)$. By Janson's Inequality

$$\lim_{n \to \infty} \Pr[\wedge \overline{B}_\alpha] = \lim_{n \to \infty} M = exp[-c^e/a]$$

completing the proof. $\square$

The fine threshold behavior for the appearance of an arbitrary graph H has been worked out but it can get quite complicated.

Lecture 4: The Chromatic Number Resolved!

The centerpiece of this lecture is the result of Béla Bollobás that, with $G \sim G(n, \frac{1}{2})$, the chromatic number $\chi(G)$ is asymptotically $n/(2 \log_2 n)$ almost surely.

1 Clique Number Revisited

In this section we fix $p = 1/2$, (other values yield similar results), let $G \sim G(n, p)$ and consider the clique number $\omega(G)$. For a fixed $c > 0$ let $n, k \to \infty$ so that

$$\binom{n}{k} 2^{-\binom{k}{2}} \to c$$

As a first approximation

$$n \sim \frac{k}{e\sqrt{2}} \sqrt{2}^k$$

and

$$k \sim \frac{2 \ln n}{\ln 2}$$

Here $\mu \to c$ so $M \to e^{-c}$. The Δ term was examined earlier. For this k, $\Delta = o(E[X]^2)$ and so $\Delta = o(1)$. Therefore

$$\lim_{n,k \to \infty} \Pr[\omega(G(n,p)) < k] = e^{-c}$$

Being more careful, let $n_0(k)$ be the minimum n for which

$$\binom{n}{k} 2^{-\binom{k}{2}} \geq 1.$$

We set

$$p = \frac{\mu(1-\epsilon)}{\Delta}$$

so as to maximize this quantity. The added assumption of Theorem 1.2 assures us that the probability p is at most one. Then

$$E\left[-\ln[\Pr[\wedge_{i\in S}\overline{B_i}]\right] \geq \frac{\mu^2(1-\epsilon)}{2\Delta}$$

Therefore there is a specific $S \subset I$ for which

$$-\ln[\Pr[\wedge_{i\in S}\overline{B_i}] \geq \frac{\mu^2(1-\epsilon)}{2\Delta}$$

That is,

$$\Pr[\wedge_{i\in S}\overline{B_i}] \leq e^{-\frac{\mu^2(1-\epsilon)}{2\Delta}}$$

But

$$\Pr[\wedge_{i\in I}\overline{B_i}] \leq \Pr[\wedge_{i\in S}\overline{B_i}]$$

completing the proof. $\square$

3 Appearance of Small Subgraphs Revisited

Generalizing the fine threshold behavior for the appearance of K_4 we find the fine threshold behavior for the appearance of any strictly balanced graph H.

Theorem 3.1 Let H be a *strictly* balanced graph with v vertices, e edges and a automorphisms. Let $c > 0$ be arbitrary. Let A be the property that G contains no copy of H. Then with $p = cn^{-v/e}$,

$$\lim_{n\to\infty} \Pr[G(n,p) \models A] = exp[-c^e/a]$$

Proof. Let $A_\alpha, 1 \leq \alpha \leq \binom{n}{v} v!/a$, range over the edge sets of possible copies of H and let B_α be the event $G(n,p) \supseteq A_\alpha$. We apply Janson's Inequality. As

$$\lim_{n\to\infty} \mu = lim_{n\to\infty}\binom{n}{v} v!p^e/a = c^e/a$$

we find

$$\lim_{n\to\infty} M = exp[-c^e/a]$$

Now we examine (similar to Theorem 1.4.2)

$$\Delta = \sum_{\alpha\sim\beta} \Pr[B_\alpha \wedge B_\beta]$$

We split the sum according to the number of *vertices* in the intersection of copies α and β. Suppose they intersect in j vertices. If $j = 0$ or $j = 1$ then $A_\alpha \cap A_\beta = \emptyset$ so that $\alpha \sim \beta$ cannot occur. For $2 \leq j \leq v$ let f_j be the maximal $|A_\alpha \cap A_\beta|$ where $\alpha \sim \beta$ and α, β intersect in j vertices. As $\alpha \neq \beta$, $f_v < e$. When $2 \leq j \leq v-1$ the critical observation is that $A_\alpha \cap A_\beta$ is a subgraph of H and hence, as H is strictly balanced,

$$\frac{f_j}{j} < \frac{e}{v}$$

from the Correlation Inequality. Thus

$$Pr[B_i| \wedge_{1\leq j<i} \overline{B_j}] \geq \Pr[B_i] - \sum_{j=1}^{d} Pr[B_j \wedge B_i]$$

Reversing

$$Pr[\overline{B_i}| \wedge_{1\leq j<i} \overline{B_j}] \leq \Pr[\overline{B_i}] + \sum_{j=1}^{d} Pr[B_j \wedge B_i]$$

$$\leq \Pr[\overline{B_i}] \left(1 + \frac{1}{1-\epsilon}\sum_{j=1}^{d} Pr[B_j \wedge B_i]\right)$$

since $\Pr[\overline{B_i}] \geq 1-\epsilon$. Employing the inequality $1+x \leq e^x$,

$$Pr[\overline{B_i}| \wedge_{1\leq j<i} \overline{B_j}] \leq \Pr[\overline{B_i}] e^{\frac{1}{1-\epsilon}\sum_{j=1}^{d} Pr[B_j\wedge B_i]}$$

For each $1 \leq i \leq m$ we plug this inequality into

$$\Pr[\wedge_{i\in I}\overline{B_i}] = \prod_{i=1}^{m} Pr[\overline{B_i}| \wedge_{1\leq j<i} \overline{B_j}]$$

The terms $\Pr[\overline{B_i}]$ multiply to M. The exponents add: for each $i, j \in I$ with $j < i$ and $j \sim i$ the term $Pr[B_j \wedge B_i]$ appears once so they add to $\Delta/2$. $\square$
Proof of Theorem 1.2 As discussed earlier, the proof of Theorem 1.1 gives

$$\Pr[\wedge_{i\in I}\overline{B_i}] \leq e^{-\mu + \frac{1}{1-\epsilon}\frac{\Delta}{2}}$$

which we rewrite as

$$-\ln[\Pr[\wedge_{i\in I}\overline{B_i}]] \geq \sum_{i\in I}\Pr[B_i] - \frac{1}{2(1-\epsilon)}\sum_{i\sim j}\Pr[B_i \wedge B_j]$$

For any set of indices $S \subset I$ the same inequality applied only to the $B_i, i \in S$ gives

$$-\ln[\Pr[\wedge_{i\in S}\overline{B_i}]] \geq \sum_{i\in S}\Pr[B_i] - \frac{1}{2(1-\epsilon)}\sum_{i,j\in S, i\sim j}\Pr[B_i \wedge B_j]$$

Let now S be a random subset of I given by

$$\Pr[i \in S] = p$$

with p a constant to be determined, the events mutually independent. (Here we are using probabilistic methods to prove a probability theorem!) Each term $\Pr[B_i]$ then appears with probability p and each term $\Pr[B_i \wedge B_j]$ with probability p^2 so that

$$E\left[-\ln[\Pr[\wedge_{i\in S}\overline{B_i}]\right] \geq E\left[\sum_{i\in S}\Pr[B_i]\right] - \frac{1}{2(1-\epsilon)}E\left[\sum_{i,j\in S, i\sim j}\Pr[B_i \wedge B_j]\right]$$

$$= p\mu - \frac{1}{1-\epsilon}p^2\frac{\Delta}{2}$$

Theorem 1.2 (The Generalized Janson Inequality). Under the assumptions of Theorem 1.1 and the further assumption that $\Delta \geq \mu(1-\epsilon)$

$$\Pr[\wedge_{i\in I}\overline{B_i}] \leq e^{-\frac{\mu^2(1-\epsilon)}{2\Delta}}$$

Theorem 1.2 (when it applies) often gives a much stronger result than Chebyschev's Inequality as used earlier. We can bound $Var[X] \leq \mu + \Delta$ so that

$$\Pr[\wedge_{i\in I}\overline{B_i}] = \Pr[X=0] \leq \frac{Var[X]}{E[X]^2} \leq \frac{\mu+\Delta}{\mu^2}$$

Suppose $\epsilon = o(1)$, $\mu \to \infty$, $\mu \ll \Delta$, and $\gamma = \frac{\mu^2}{\Delta} \to \infty$. Chebyschev's upper bound on $\Pr[X=0]$ is then roughly γ^{-1} while Janson's upper bound is roughly $e^{-\gamma}$.

2 The Proofs

The original proofs of Janson are based on estimates of the Laplace transform of an appropriate random variable. The proof presented here follows that of Boppana and Spencer [1989]. We shall use the inequalities

$$\Pr[B_i| \wedge_{j\in J}\overline{B_j}] \leq \Pr[B_i]$$

valid for all index sets $J \subset I, i \notin J$ and

$$\Pr[B_i|B_k \wedge \bigwedge_{j\in J}\overline{B_j}] \leq \Pr[B_i|B_k]$$

valid for all index sets $J \subset I, i,k \notin J$. The first follows from general Correlation Inequalities. The second is equivalent to the first since conditioning on B_k is the same as assuming $p_r = \Pr[r \in R] = 1$ for all $r \in A_k$.
Proof of Theorem 1.1 The lower bound follows immediately. Order the index set $I = \{1,\ldots,m\}$ for convenience. For $1 \leq i \leq m$

$$Pr[B_i| \wedge_{1\leq j<i}\overline{B_j}] \leq \Pr[B_i]$$

so

$$Pr[\overline{B_i}| \wedge_{1\leq j<i}\overline{B_j}] \geq Pr[\overline{B_i}]$$

and

$$\Pr[\wedge_{i\in I}\overline{B_i}] = \prod_{i=1}^{m} Pr[\overline{B_i}| \wedge_{1\leq j<i}\overline{B_j}] \geq \prod_{i=1}^{m} Pr[\overline{B_i}]$$

Now the upper bound. For a given i renumber, for convenience, so that $i \sim j$ for $1 \leq j \leq d$ and not for $d+1 \leq j < i$. We use the inequality $\Pr[A|B\wedge C] \geq \Pr[A\wedge B|C]$, valid for any A,B,C. With $A = B_i$, $B = \overline{B_1} \wedge \ldots \wedge \overline{B_d}$, $C = \overline{B_{d+1}} \wedge \ldots \wedge \overline{B_{i-1}}$

$$Pr[B_i| \wedge_{1\leq j<i}\overline{B_j}] = \Pr[A|B\wedge C] \geq Pr[A\wedge B|C] = Pr[A|C]Pr[B|A\wedge C]$$

From the mutual independence $Pr[A|C] = Pr[A]$. We bound

$$\Pr[B|A\wedge C] \geq 1 - \sum_{j=1}^{d} Pr[B_j|B_i \wedge C] \geq 1 - \sum_{j=1}^{d} Pr[B_j|B_i]$$

When the B_i are "mostly" independent the Janson Inequalities allow us, sometimes, to say that these two quantities are "nearly" equal.

Let Ω be a finite universal set and let R be a random subset of Ω given by

$$\Pr[r \in R] = p_r,$$

these events mutually independent over $r \in \Omega$. (In application to $G(n,p)$, Ω is the set of pairs $\{i,j\}$, $i,j \in V(G)$ and all $p_r = p$ so that R is the edge set of $G(n,p)$.) Let $A_i, i \in I$, be subsets of Ω, I a finite index set. Let B_i be the *event* $A_i \subseteq R$. (That is, each point $r \in \Omega$ " flips a coin" to determine if it is in R. B_i is the event that the coins for all $r \in A_i$ came up "heads".) Let X_i be the indicator random variable for B_i and $X = \sum_{i \in I} X_i$ the number of $A_i \subseteq R$. The event $\wedge_{i \in I} \overline{B_i}$ and $X = 0$ are then identical. For $i,j \in I$ we write $i \sim j$ if $i \neq j$ and $A_i \cap A_j \neq \emptyset$. Note that when $i \neq j$ and not $i \sim j$ then B_i, B_j are independent events since they involve separate coin flips. Furthermore, and this plays a crucial role in the proofs, if $i \notin J \subset I$ and not $i \sim j$ for all $j \in J$ then B_i is mutually independent of $\{B_j | j \in J\}$, i.e., independent of any Boolean function of those B_j. This is because the coin flips on A_i and on $\cup_{j \in J} A_j$ are independent. We define

$$\Delta = \sum_{i \sim j} \Pr[B_i \wedge B_j]$$

Here the sum is over ordered pairs so that $\Delta/2$ gives the same sum over unordered pairs. (This will be the same Δ as in Lecture 1. We set

$$M = \prod_{i \in I} \Pr[\overline{B_i}],$$

the value of $\Pr[\wedge_{i \in I} \overline{B_i}]$ if the B_i were independent.
Theorem 1.1 (The Janson Inequality). Let $B_i, i \in I, \Delta, M$ be as above and assume all $\Pr[B_i] \leq \epsilon$. Then

$$M \leq \Pr[\wedge_{i \in I} \overline{B_i}] \leq M e^{\frac{1}{1-\epsilon}\frac{\Delta}{2}}$$

Now set

$$\mu = E[X] = \sum_{i \in I} \Pr[B_i]$$

For each $i \in I$

$$\Pr[\overline{B_i}] = 1 - \Pr[B_i] \leq e^{-\Pr[B_i]}$$

so, multiplying over $i \in I$,

$$M \leq e^{-\mu}$$

It is often more convenient to replace the upper bound of Theorem 1.1 with

$$\Pr[\wedge_{i \in I} \overline{B_i}] \leq e^{-\mu + \frac{1}{1-\epsilon}\frac{\Delta}{2}}$$

As an example, set $p = cn^{-2/3}$ and consider the probability that $G(n,p)$ contains no K_4. The B_i then range over the $\binom{n}{4}$ potential K_4 - each being a 6-element subset of Ω. Here, as is often the case, $\epsilon = o(1)$, $\Delta = o(1)$ (as calculated previously) and μ approaches a constant, here $k = c^6/24$. In these instances $\Pr[\wedge_{i \in I} \overline{B_i}] \rightarrow e^{-k}$. Thus we have the fine structure of the threshold function of $\omega(G) = 4$.

As Δ becomes large the Janson Inequality becomes less precise. Indeed, when $\Delta \geq 2\mu(1-\epsilon)$ it gives an upper bound for the probability which is larger than one. At that point (and even somewhat before) the following result kicks in.

then $R(k,l) > n$.
Proof. Let $G \sim G(n,p)$ and color the edges of G red and the other edges of K_n blue. Then the left hand side above is simply the expectation of the number of red K_k plus the number of blue K_l. For some G this is zero and that G gives the desired coloring. □

Dealing with the asymptotics of this result can be quite tricky. For example, what does this imply about $R(k,2k)$?

5 High Girth and High Chromatic Number

Many consider the following one of the most pleasing uses of the probabilistic method, as the result is surprising and does not appear to call for nonconstructive techniques. The *girth* of a graph G is the size of its smallest circuit.
Theorem 5.1(Erdős [1959]). For all k,l there exists a graph G with $girth(G) > l$ and $\chi(G) > k$.
Proof. Fix $\theta < 1/l$ and let $G \sim G(n,p)$ with $p = n^{\theta-1}$. Let X be the number of circuits of size at most l. Then

$$E[X] = \sum_{i=3}^{l} \frac{(n)_i}{2i} p^i \leq \sum_{i=3}^{l} \frac{n^{\theta i}}{2i} = o(n)$$

as $\theta l < 1$. In particular

$$\Pr[X \geq \frac{n}{2}] = o(1)$$

Set $x = \lceil \frac{3}{p} \ln n \rceil$ so that

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

Let n be sufficiently large so that both these events have probability less than .5. Then there is a specific G with less than $n/2$ cycles of length less than l and with $\alpha(G) < 3n^{1-\theta} \ln n$. Remove from G a vertex from each cycle of length at most l. This gives a graph G^* with at least $n/2$ vertices. G^* has girth greater than l and $\alpha(G^*) \leq \alpha(G)$. Thus

$$\chi(G^*) \geq \frac{|G^*|}{\alpha(G^*)} \geq \frac{n/2}{3n^{1-\theta} \ln n} = \frac{n^\theta}{6 \ln n}$$

To complete the proof, let n be sufficiently large so that this is greater than k.

Lecture 3: The Poisson Paradigm

When X is the sum of many rare indicator "mostly independent" random variables and $\mu = E[X]$ we would like to say that X is close to a Poisson distribution with mean μ and, in particular, that $\Pr[X = 0]$ is nearly $e^{-\mu}$. We call this rough statement the Poisson Paradigm. We give a number of situations in which this Paradigm may be rigorously proven.

1 The Janson Inequalities

In many instances we would like to bound the probability that none of a set of bad events $B_i, i \in I$ occur. If the events are mutually independent then

$$\Pr[\wedge_{i \in I} \overline{B_i}] = \prod_{i \in I} \Pr[\overline{B_i}]$$

For such p, $n(1-p)^{n-1} \sim \mu = e^{-c}$ and by the above argument the probability that X has no isolated vertices approaches $e^{-\mu}$. If G has no isolated vertices but is not connected there is a component of k vertices for some $2 \le k \le \frac{n}{2}$. Letting B be this event

$$\Pr[B] \le \sum_{k=2}^{n/2} \binom{n}{k} k^{k-2} p^{k-1} (1-p)^{k(n-1)-\binom{k}{2}}$$

The first factor is the choice of a component set $S \subset V(G)$. The second factor is a choice of tree on S. The third factor is the probability that those tree pairs are in $E(G)$. The final factor is that there be no edge from S to $V(G) - S$. Some calculation (which we omit but note that $k = 2$ provides the main term) gives that $\Pr[B] = o(1)$ so that $X \neq 0$ and connectivity have the same limiting probability. $\Box$

4 The Probabilistic Method

In 1947 Paul Erdős started what is now called the Probabilistic Method with a three page paper in the Bulletin of the American Mathematical Society. The Ramsey function $R(k,l)$ is defined as the least n such that if the edges of K_n are colored Red and Blue then there is either a Red K_k or a Blue K_l. The existence of such an n is a consequence of Ramsey's Theorem and will not concern us here. Rather, we are interested in lower bounds on the Ramsey function. To unravel the definition $R(k,l) > n$ means that there *exists* a Red-Blue coloring of K_n with neither Red K_k nor Blue K_l. In his 1947 paper Erdős considered the case $k = l$.
Theorem 4.1. If

$$\binom{n}{k} 2^{1-\binom{k}{2}} < 1$$

then $R(k,l) > n$.
Proof. Let $G \sim G(n, \frac{1}{2})$ and consider the random two-coloring given by coloring the edges of G red and the other edges of K_n blue. Let X be the number of monochromatic K_k. Then the left hand side above is simply $E[X]$. With $E[X] < 1$, $\Pr[X = 0] > 0$. Hence there is a point in the probability space - i.e., a graph G, whose coloring has $X = 0$ monochromatic K_k. $\Box$

Note here a subtle (for some) point. With positive probability $G(n, \frac{1}{2})$ has the desired property and therefore there must - absolutely, positively - exist a G with the desired property. Random Graphs and the Probabilistic Method are closely related. In Random Graphs we study the probability of $G(n,p)$ having certain properties. In the Probabilistic Method our goal is to prove the existence of a G having certain properties. We create a probability space in which the probability of the random G having these properties is positive, and from that it follows that some such G must exist.

Applying some simple asymptotics to the theorem yields that $R(k,k) > \sqrt{2}^{n(1+o(1))}$. In 1935 Erdős and George Szekeres found the upper bound $R(k,k) < 4^{n(1+o(1))}$ by nonrandom means. While there have been improvements in lower order terms, these bounds remain the best known up to $(1+o(1))^n$ terms. It is also interesting that no exponential lower bound is known by constructive means.

A general lower bound is the following.
Theorem 4.2. If there exists $p \in [0,1]$ with

$$\binom{n}{k} p^{\binom{k}{2}} + \binom{n}{l} (1-p)^{\binom{l}{2}} < 1$$

For the lower bound (which is not best possible) we outline an analysis of the following "greedy algorithm". We find an independent set C on G as follows. Set $S_0 = V(G)$, $a_1 = 1$ and S_1 equal the set of vertices not adjacent to a_1. Having determined $a_1, \ldots, a_i$ and S_i let a_{i+1} be the least vertex of S_i and let S_{i+1} be those $x \in S_i - \{a_i\}$ not adjacent to a_{i+1}. Continue until $S_t = \emptyset$ and set $C = \{a_1, \ldots a_t\}$. A fairly straightforward analysis gives that $|C| \sim \log_2 n$ almost surely, and moreover that the probability (for any given $\epsilon > 0$ that $|C| < (log_2 n)(1-\epsilon)$ is $o(n^{-1})$. Call this one pass of the algorithm. Now we give all points of C color "one", remove vertices C from G and iterate. Let G^1 be G with C removed. Critically, it finding C we only "exposed" edges involving C so that we can consider G^1 to have distribution $G(n_1, \frac{1}{2})$, where $n_1 = n - |C|$ is the number of vertices. Letting n_j be the number of vertices remaining after the j-th pass, almost surely we have $n_{j+1} < n_j - (1-\epsilon)\log_2 n_j$ so that the algorithm is completed using less than $\frac{n}{\log_2 n}(1+\epsilon')$ colors. (Actually, to avoid end effects we can stop the algorithm when there are $o(n/\log n)$ vertices remaining and simply give each such vertex a separate color.)

It is tempting to improve the lower bound as follows. We know that almost surely G contains an independent set of size $\sim 2\log_2 n$. Let C be that set, remove C from G giving G^1 and iterate. The problem is, of course, that G^1 no longer has distribution $G(n_1, \frac{1}{2})$ and no proof has been found along these lines of the true result that $\chi(G) \sim \frac{n}{2\log_2 n}$ almost surely.

3 Connectivity

In this section we give a relatively simple example of what we call the Poisson Paradigm: the rough notion that if there are many rare and nearly independent events then the number of events that hold has approximately a Poisson distribution. This will yield one of the most beautiful of the Erdős- Rényi results, a quite precise description of the threshold behavior for connectivity. A vertex $v \in G$ is *isolated* if it is adjacent to no $w \in V$. In $G(n,p)$ let X be the number of isolated vertices.

Theorem 3.1. Let $p = p(n)$ satisfy $n(1-p)^{n-1} = \mu$. Then

$$\lim_{n\to\infty} \Pr[X = 0] = e^{-\mu}$$

We let X_i be the indicator random variable for vertex i being isolated so that $X = X_1 + \ldots + X_n$. Then $E[X_i] = (1-p)^{n-1}$ so by linearity of expectation $E[X] = \mu$. Now consider the r-th factorial moment $E[(X)_r]$ for any fixed r. By the symmetry $E[(X)_r] = (n)_r E[X_1 \cdots X_r]$. For vertices $1, \ldots, r$ to all be isolated the $r(n-1) - \binom{r}{2}$ pairs $\{i, x\}$ overlapping $1, \ldots, r$ must all not be edges. Thus

$$E[(X)_r] = (n)_r (1-p)^{r(n-1)-\binom{r}{2}} \sim n^r (1-p)^{r(n-1)} \sim \mu^r$$

(That is, the dependence among the X_i was asymptotically negligible.) As all the moments of X approach those of $P(\mu)$, X approaches $P(\mu)$ in distribution and in particular the theorem holds. □

Now we give the Erdős-Rényi famous "double exponential" result.

Theorem 3.2. Let

$$p = p(n) = \frac{\log n}{n} + \frac{c}{n} + o(\frac{1}{n})$$

Then

$$\lim_{n\to\infty} \Pr[G(n,p) \text{ is connected}] = e^{-e^{-c}}$$

where we set

$$g(i) = \frac{\binom{k}{i}\binom{n-k}{k-i}}{\binom{n}{k}} 2^{\binom{i}{2}}$$

Observe that $g(i)$ may be thought of as the probability that a randomly chosen T will intersect a fixed S in i points times the factor increase in $\Pr[A_T]$ when it does. Setting $i = 2$,

$$g(2) = 2\frac{\binom{k}{2}\binom{n-k}{k-2}}{\binom{n}{k}} \sim \frac{k^4}{n^2} = o(1)$$

At the other extreme $i = k - 1$

$$g(k-1) = \frac{k(n-k)2^{-(k-1)}}{\binom{n}{k}2^{-\binom{k}{2}}} \sim \frac{2kn2^{-k}}{E[X]}$$

As $k \sim 2\log_2 n$ the numerator is $n^{-1+o(1)}$. The denominator approaches infinity and so $g(k-1) = o(1)$. Some detailed calculation (which we omit) gives that the remaining $g(i)$ are also negligible so that Corollary 1.3.5 applies. $\Box$

Theorem 1.1 leads to a strong concentration result for $\omega(G)$. For $k \sim 2\log_2 n$

$$\frac{f(k+1)}{f(k)} = \frac{n-k+1}{k+1}2^{-k} = n^{-1+o(1)} = o(1)$$

Let $k_0 = k_0(n)$ be that value with $f(k_0) \geq 1 > f(k_0+1)$. For "most" n the function $f(k)$ will jump from a large $f(k_0)$ to a small $f(k_0+1)$. The probability that G contains a clique of size k_0+1 is at most $f(k_0+1)$ which will be very small. When $f(k_0)$ is large Theorem 1.1 implies that G contains a clique of size k_0 with probability nearly one. Together, with very high probability $\omega(G) = k_0$. For some n one of the values $f(k_0), f(k_0+1)$ may be of moderate size so this argument does not apply. Still one may show a strong concentration result found independently by Bollobás, Erdős [1976] and Matula [1976].
Corollary 1.2 There exists $k = k(n)$ so that

$$\Pr[\omega(G) = k \text{ or } k+1] \to 1$$

2 Chromatic Number

Again let us fix $p = \frac{1}{2}$ and this time we consider the chromatic number $\chi(G)$ with $G \sim G(n,p)$. Our results in this section will be improved in Lecture 4.
Theorem 2.1. Almost surely

$$\frac{n}{2\log_2 n}(1+o(1)) \leq \chi(G) \leq \frac{n}{log_2 n}(1+o(1))$$

For the lower bound we use the general bound

$$\chi(G) \geq n/\omega(\overline{G})$$

which is true since each color class must be a clique in $\overline{G}$ and so can be used at most $\omega(\overline{G})$ times. But $\overline{G}$ has the same distribution as G so almost surely $\omega(\overline{G}) \leq (2\log_2 n)(1+o(1)$. This will turn out to be the right asymptotic answer.

Our assumption $p \gg n^{-v/e}$ implies $E[X] \to \infty$. It suffices therefore to show $\Delta^* = o(E[X])$. Fixing $x_1, \ldots, x_v$,

$$\Delta^* = \sum_{(y_1,\ldots,y_v)\sim(x_1,\ldots,x_v)} \Pr[A_{(y_1,\ldots,y_v)}|A_{(x_1,\ldots,x_v)}]$$

There are $v!/a = O(1)$ terms with $\{y_1, \ldots, y_v\} = \{x_1, \ldots, x_v\}$ and for each the conditional probability is at most one (actually, at most p), thus contributing $O(1) = o(E[X])$ to Δ^*. When $\{y_1, \ldots, y_v\} \cap \{x_1, \ldots, x_v\}$ has i elements, $2 \leq i \leq v-1$ the argument of Theorem 4.2 gives that the contribution to Δ^* is $o(E[X])$. Altogether $\Delta^* = o(E[X])$ and we apply Corollary 3.5 □

Theorem 4.5 Let H be *any* fixed graph. For every subgraph H' of H (including H itself) let $X_{H'}$ denote the number of copies of H' in $G(n,p)$. Assume p is such that $E[X_{H'}] \to \infty$ for every H'. Then

$$X_H \sim E[X_H]$$

almost always.

Proof. Let H have v vertices and e edges. As in Theorem 4.4 it suffices to show $\Delta^* = o(E[X])$. We split Δ^* into a finite number of terms. For each H' with w vertice and f edges we have those $(y_1, \ldots, y_v)$ that overlap with the fixed $(x_1, \ldots, x_v)$ in a copy of H'. These terms contribute, up to constants,

$$n^{v-w}p^{e-f} = \Theta\left(\frac{E[X_H]}{E[X_{H'}]}\right) = o(E[X_H])$$

to Δ^*. Hence Corollary 3.5 does apply. □

Lecture 2: More Random Graphs

1 Clique Number

Now we fix edge probability $p = \frac{1}{2}$ and consider the clique number $\omega(G)$. We set

$$f(k) = \binom{n}{k} 2^{-\binom{k}{2}},$$

the expected number of k-cliques. The function $f(k)$ drops under one at $k \sim 2\log_2 n$. (Very roughly, $f(k)$ is like $n^k 2^{-k^2/2}$.)

Theorem 1.1 Let $k = k(n)$ satisfy $k \sim 2\log_2 n$ and $f(k) \to \infty$. Then almost always $\omega(G) \geq k$.

Proof. For each k-set S let A_S be the event "S is a clique" and X_S the corresponding indicator random variable. We set

$$X = \sum_{|S|=k} X_S$$

so that $\omega(G) \geq k$ if and only if $X > 0$. Then $E[X] = f(k) \to \infty$ and we examine Δ^*. Fix S and note that $T \sim S$ if and only if $|T \cap S| = i$ where $2 \leq i \leq k-1$. Hence

$$\Delta^* = \sum_{i=2}^{k-1} \binom{k}{i}\binom{n-k}{k-i} 2^{\binom{i}{2}-\binom{k}{2}}$$

and so

$$\frac{\Delta^*}{E[X]} = \sum_{i=2}^{k-1} g(i)$$

common edges - i.e., if and only if $|S \cap T| = i$ with $2 \le i \le v-1$. Let S be fixed. We split

$$\Delta^* = \sum_{T \sim S} \Pr[A_T|A_S] = \sum_{i=2}^{v-1} \sum_{|T \cap S|=i} \Pr[A_T|A_S]$$

For each i there are $O(n^{v-i})$ choices of T. Fix S, T and consider $\Pr[A_T|A_S]$. There are $O(1)$ possible copies of H on T. Each has - since, critically, H is balanced - at most $\frac{ie}{v}$ edges with both vertices in S and thus at least $e - \frac{ie}{v}$ other edges. Hence

$$\Pr[A_T|A_S] = O(p^{e-\frac{ie}{v}})$$

and

$$\Delta^* = \sum_{i=2}^{v-1} O(n^{v-i}p^{e-\frac{ie}{v}}) = \sum_{i=2}^{v-1} O((n^v p^e)^{1-\frac{i}{v}})$$

$$= \sum_{i=2}^{v-1} o(n^v p^e) = o(E[X])$$

since $n^v p^e \to \infty$. Hence Corollary 3.5 applies. $\square$

Theorem 4.3 In the notation of Theorem 4.2 if H is *not* balanced then $p = n^{-v/e}$ is *not* the threshold function for A.

Proof. Let H_1 be a subgraph of H with v_1 vertices, e_1 edges and $e_1/v_1 > e/v$. Let α satisfy $v/e < \alpha < v_1/e_1$ and set $p = n^{-\alpha}$. The expected number of copies of H_1 is then $o(1)$ so almost always $G(n,p)$ contains no copy of H_1. But if it contains no copy of H_1 then it surely can contain no copy of H. $\square$

The threshold function for the property of containing a copy of H, for general H, was examined in the original papers of Erdős and Rényi. Let H_1 be that subgraph with maximal density $\rho(H_1) = e_1/v_1$. (When H is balanced we may take $H_1 = H$.) They showed that $p = n^{-v_1/e_1}$ is the threshold function. This will follow fairly quickly from the methods of theorem 4.5.

We finish this section with two strengthenings of Theorem 4.2.

Theorem 4.4 Let H be strictly balanced with v vertices, e edges and a automorphisms. Let X be the number of copies of H in $G(n,p)$. Assume $p \gg n^{-v/e}$. Then almost always

$$X \sim \frac{n^v p^e}{a}$$

Proof. Label the vertices of H by $1, \ldots, v$. For each ordered $x_1, \ldots, x_v$ let $A_{x_1,\ldots,x_v}$ be the event that $x_1, \ldots, x_v$ provides a copy of H in that order. Specifically we define

$$A_{x_1,\ldots,x_v} : \{i,j\} \in E(H) \Rightarrow \{x_i, x_j\} \in E(G)$$

We let $I_{x_1,\ldots,x_v}$ be the corresponding indicator random variable. We define an equivalence class on v-tuples by setting $(x_1, \ldots, x_v) \equiv (y_1, \ldots, y_v)$ if there is an automorphism σ of $V(H)$ so that $y_{\sigma(i)} = x_i$ for $1 \le i \le v$.Then

$$X = \sum I_{x_1,\ldots,x_v}$$

gives the number of copies of H in G where the sum is taken over one entry from each equivalence class. As there are $(n)_v/a$ terms

$$E[X] = \frac{(n)_v}{a} E[I_{x_1,\ldots,x_v}] = \frac{(n)_v p^e}{a} \sim \frac{n^v p^e}{a}$$

When $p(n) \ll n^{-2/3}$, $E[X] = o(1)$ and so $X = 0$ almost surely.

Now suppose $p(n) \gg n^{-2/3}$ so that $E[X] \to \infty$ and consider the Δ^* of Corollary 3.5. (All 4-sets "look the same" so that the X_S are symmetric.) Here $S \sim T$ if and only if $S \neq T$ and S, T have common edges - i.e., if and only if $|S \cap T| = 2$ or 3. Fix S. There are $O(n^2)$ sets T with $|S \cap T| = 2$ and for each of these $\Pr[A_T|A_S] = p^5$. There are $O(n)$ sets T with $|S \cap T| = 3$ and for each of these $\Pr[A_T|A_S] = p^3$. Thus

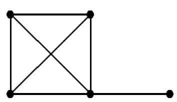
$$\Delta^* = O(n^2p^5) + O(np^3) = o(n^4p^6) = o(E[X])$$

since $p \gg n^{-2/3}$. Corollary 3.5 therefore applies and $X > 0$, i.e., there *does* exist a clique of size 4, almost always. □

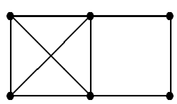
The proof of Theorem 4.1 appears to require a fortuitous calculation of Δ^*. The following definitions will allow for a description of when these calculations work out.

Definitions. Let H be a graph with v vertices and e edges. We call $\rho(H) = e/v$ the *density* of H. We call H *balanced* if every subgraph H' has $\rho(H') \leq \rho(H)$. We call H *strictly balanced* if every proper subgraph H' has $\rho(H') < \rho(H)$.

Examples. K_4 and, in general, K_k are strictly balanced. The graph



is not balanced as it has density $7/5$ while the subgraph K_4 has density $3/2$. The graph



is balanced but not strictly balanced as it and its subgraph K_4 have density $3/2$.

Theorem 4.2 Let H be a balanced graph with v vertices and e edges. Let $A(G)$ be the event that H is a subgraph (not necessarily induced) of G. Then $p = n^{-v/e}$ is the threshold function for A.

Proof. We follow the argument of Theorem 4.1 For each v-set S let A_S be the event that $G|_S$ contains H as a subgraph. Then

$$p^e \leq \Pr[A_S] \leq v!p^e$$

(Any particular placement of H has probability p^e of occuring and there are at most $v!$ possible placements. The precise calculation of $\Pr[A_S]$ is, in general, complicated due to the overlapping of potential copies of H.) Let X_S be the indicator random variable for A_S and

$$X = \sum_{|S|=v} X_S$$

so that A holds if and only if $X > 0$. Linearity of Expectation gives

$$E[X] = \sum_{|S|=v} E[X_S] = \binom{n}{v} \Pr[A_S] = \Theta(n^v p^e)$$

If $p \ll n^{-v/e}$ then $E[X] = o(1)$ so $X = 0$ almost always.

Now assume $p \gg n^{-v/e}$ so that $E[X] \to \infty$ and consider the Δ^* of Corollary 3.5 (All v-sets look the same so the X_S are symmetric.) Here $S \sim T$ if and only if $S \neq T$ and S, T have

Note that when $i \sim j$

$$Cov[X_i, X_j] = E[X_i X_j] - E[X_i]E[X_j] \leq E[X_i X_j] = \Pr[A_i \wedge A_j]$$

and that when $i \neq j$ and not $i \sim j$ then $Cov[X_i, X_j] = 0$. Thus

$$Var[X] \leq E[X] + \Delta$$

Corollary 3.4. If $E[X] \to \infty$ and $\Delta = o(E[X]^2)$ then $X > 0$ almost always. Furthermore $X \sim E[X]$ almost always.

Let us say $X_1, \ldots, X_m$ are *symmetric* if for every $i \neq j$ there is an automorphism of the underlying probability space that sends event A_i to event A_j. Examples will appear in the next section. In this instance we write

$$\Delta = \sum_{i \sim j} \Pr[A_i \wedge A_j] = \sum_i \Pr[A_i] \sum_{j \sim i} \Pr[A_j | A_i]$$

and note that the inner summation is independent of i. We set

$$\Delta^* = \sum_{j \sim i} \Pr[A_j | A_i]$$

where i is any fixed index. Then

$$\Delta = \sum_i \Pr[A_i]\Delta^* = \Delta^* \sum_i \Pr[A_i] = \Delta^* E[X]$$

Corollary 3.5. If $E[X] \to \infty$ and $\Delta^* = o(E[X])$ then $X > 0$ almost always. Furthermore $X \sim E[X]$ almost always.

The condition of Corollary 3.5 has the intuitive sense that conditioning on any specific A_i holding does not substantially increase the expected number $E[X]$ of events holding.

4 Appearance of Small Subgraphs

What is the threshold function for the appearance of a given graph H. This problem was solved in the original papers of Erdős and Rényi. We begin with an instructive special case.
Theorem 4.1 The property $\omega(G) \geq 4$ has threshold function $n^{-2/3}$.
Proof. For every 4-set S of vertices in $G(n,p)$ let A_S be the event "S is a clique" and X_S its indicator random variable. Then

$$E[X_S] = \Pr[A_S] = p^6$$

as six different edges must all lie in $G(n,p)$. Set

$$X = \sum_{|S|=4} X_S$$

so that X is the number of 4-cliques in G and $\omega(G) \geq 4$ if and only if $X > 0$. Linearity of Expectation gives

$$E[X] = \sum_{|S|=4} E[X_S] = \binom{n}{4} p^6 \sim \frac{n^4 p^6}{24}$$

precisely the conjunction $\wedge \overline{B}_S$ over all S. If the B_S were mutually independent then we *would* have

$$\Pr[\wedge \overline{B}_S] = \prod[\overline{B}_S] = (1-p^3)^{\binom{n}{3}} \sim e^{-\binom{n}{3}p^3} \to e^{-c^3/6}$$

The reality is that the B_S are not mutually independent though when $|S \cap T| \leq 1$, B_S and B_T are mutually independent. This is quite a typical situation in the study of random graphs in which we must deal with events that are "almost", but not precisely, mutual independent.

3 Variance

Here we introduce the Variance in a form that is particularly suited to the study of random graphs. The expressions Δ and Δ^* defined in this section will appear often in these notes.

Let X be a nonnegative integral valued random variable and suppose we want to bound $\Pr[X = 0]$ given the value $\mu = E[X]$. If $\mu < 1$ we may use the inequality

$$\Pr[X > 0] \leq E[X]$$

so that if $E[X] \to 0$ then $X = 0$ almost always. (Here we are imagining an infinite sequence of X dependent on some parameter n going to infinity.) But now suppose $E[X] \to \infty$. It does *not* necessarily follow that $X > 0$ almost always. For example, let X be the number of deaths due to nuclear war in the twelve months after reading this paragraph. Calculation of $E[X]$ can make for lively debate but few would deny that it is quite large. Yet we may believe - or hope - that $Pr[X \neq 0]$ is very close to zero. We can sometimes deduce $X > 0$ almost always if we have further information about $Var[X]$.

Theorem 3.1

$$\Pr[X = 0] \leq \frac{Var[X]}{E[X]^2}$$

Proof. Set $\lambda = \mu/\sigma$ in Chebyschev's Inequality. Then

$$\Pr[X = 0] \leq \Pr[|X - \mu| \geq \lambda\sigma] \leq \frac{1}{\lambda^2} = \frac{\sigma^2}{\mu^2} \ \Box$$

We generally apply this result in asymptotic terms.

Corollary 3.2

$$\text{If } Var[X] = o(E[X]^2) \text{ then } X > 0 \text{ a.a.}$$

The proof of the Theorem actually gives that for any $\epsilon > 0$

$$\Pr[|X - E[X]| \geq \epsilon E[X]] \leq \frac{Var[X]}{\epsilon^2 E[X]^2}$$

and thus in asymptotic terms we actually have the following stronger assertion:

Corollary 3.3

$$\text{If } Var[X] = o(E[X]^2) \text{ then } X \sim E[X] \text{ a.a.}$$

Suppose again $X = X_1 + \ldots + X_m$ where X_i is the indicator random variable for event A_i. For indices i, j write $i \sim j$ if $i \neq j$ and the events A_i, A_j are not independent. We set (the sum over ordered pairs)

$$\Delta = \sum_{i \sim j} \Pr[A_i \wedge A_j]$$

graph of all "on" edges has distribution $G(n,p)$. As p increases $G(n,p)$ *evolves* from empty to full.

In their original paper Erdős and Rényi let $G(n,e)$ be the random graph with n vertices and precisely e edges. Again there is a dynamic model: Begin with no edges and add edges randomly one by one until the graph becomes full. Generally $G(n,e)$ will have very similar properties as $G(n,p)$ with $p \sim \frac{e}{\binom{n}{2}}$. We will work on the probability model exclusively.

2 Threshold Functions

The term "the random graph" is, strictly speaking, a misnomer. $G(n,p)$ is a probability space over graphs. Given any graph theoretic property A there will be a probability that $G(n,p)$ satisfies A, which we write $\Pr[G(n,p) \models A]$. When A is monotone $\Pr[G(n,p) \models A]$ is a monotone function of p. As an instructive example, let A be the event "G is triangle free". Let X be the number of triangles contained in $G(n,p)$. Linearity of expectation gives

$$E[X] = \binom{n}{3} p^3$$

This suggests the parametrization $p = c/n$. Then

$$\lim_{n\to\infty} E[X] = \lim_{n\to\infty} \binom{n}{3} p^3 = c^3/6$$

We shall see that the distribution of X is asymptotically Poisson. In particular

$$\lim_{n\to\infty} \Pr[G(n,p) \models A] = \lim_{n\to\infty} \Pr[X=0] = e^{-c^3/6}$$

Note that

$$\lim_{c\to 0} e^{-c^3/6} = 1$$

$$\lim_{c\to\infty} e^{-c^3/6} = 0$$

When $p = 10^{-6}/n$, $G(n,p)$ is very unlikely to have triangles and when $p = 10^6/n$, $G(n,p)$ is very likely to have triangles. In the dynamic view the first triangles almost always appear at $p = \Theta(1/n)$. If we take a function such as $p(n) = n^{-.9}$ with $p(n) \gg n^{-1}$ then $G(n,p)$ will almost always have triangles. Occasionally we will abuse notation and say, for example, that $G(n, n^{-.9})$ contains a triangle - this meaning that the probability that it contains a triangle approaches 1 as n approaches infinity. Similarly, when $p(n) \ll n^{-1}$, for example, $p(n) = 1/(n \ln n)$, then $G(n,p)$ will almost always not contain a triangle and we abuse notation and say that $G(n, 1/(n \ln n))$ is trianglefree. It was a central observation of Erdős and Rényi that many natural graph theoretic properties become true in a very narrow range of p. They made the following key definition.
Definition. $r(n)$ is called a *threshold function* for a graph theoretic property A if
(i) When $p(n) \ll r(n), \lim_{n\to\infty} \Pr[G(n,p) \models A] = 0$
(ii) When $p(n) \gg r(n), \lim_{n\to\infty} \Pr[G(n,p) \models A] = 1$
or visa versa.

In our example, $1/n$ is a threshold function for A. Note that the threshold function, when one exists, is not unique. We could equally have said that $10/n$ is a threshold function for A.

Lets approach the problem of $G(n, c/n)$ being trianglefree once more. For every set S of three vertices let B_S be the event that S is a triangle. Then $\Pr[B_S] = p^3$. Then "trianglefreeness" is

NINE LECTURES ON RANDOM GRAPHS

Joel Spencer

Courant Institute of Mathematical Sciences
251 Mercer St., New York, NY 10012

Graph Theory Preliminaries A graph G, formally speaking, is a pair $(V(G), E(G))$ where the elements $v \in V(G)$ are called vertices and the elements of $E(G)$, called edges, are two element subsets $\{v, w\}$ of $V(G)$. When $\{v, w\} \in E(G)$ we say v, w are adjacent. (In standard graph theory terminology our graphs are undirected and have no loops and no multiple edges.) Pictorially, we often display the $v \in V(G)$ as points and draw an arc between v and w when they are adjacent. We call $V(G)$ the vertex set of G and $E(G)$ the edge set of G. (When G is understood we shall write simply V and E respectively. We also often write $v \in G$ or $\{v, w\} \in G$ instead of the formally correct $v \in V(G)$ and $\{v, w\} \in E(G)$ respectively.) A set $S \subseteq V$ is called a *clique* if all pairs $x, y \in S$ are adjacent. The clique number, denoted by $\omega(G)$, is the largest size of a clique in G. An independent set S is one for which no pairs $x, y, \in S$ are adjacent, the largest size of an independent set is called the independence number and is denoted $\alpha(G)$. A k-coloring of G is a map $f : V \to \{1, \ldots, k\}$ such that if x, y are adjacent then $f(x) \neq f(y)$. The minimal k for which a k-coloring exists is called the chromatic number of G and is denoted $\chi(G)$. Note $\omega(G) \leq \chi(G)$ since all vertices of a clique much receive distinct colors. The complement of G, denoted $\overline{G}$, has the same vertex set as G and x, y are adjacent in $\overline{G}$ if and only if x, y are not adjacent in G. The complete graph on k vertices, denoted by K_k, consists of a vertex set of size k with all pairs x, y adjacent. The empty graph on k vertices, denoted by I_k, consists of a vertex set of size k with no pairs x, y adjacent.
References. Theorem 2.3.4 refers to Lecture 2, Section 3, theorem 4. Double indexing, such as Theorem 3.4, refers to Section 3, theorem 4 in the current lecture.

Lecture 1: Basics

1 What is a Random Graph

Let n be a positive integer, $0 \leq p \leq 1$. The random graph $G(n, p)$ is a probability space over the set of graphs on the vertex set $\{1, \ldots, n\}$ determined by

$$\Pr[\{i, j\} \in G] = p$$

with these events mutually independent.

Random Graphs is an active area of research which combines probability theory and graph theory. The subject began in 1960 with the monumental paper *On the Evolution of Random Graphs* by Paul Erdős and Alfred Rényi. The book *Random Graphs* by Béla Bollobás is the standard source for the field.

There is a compelling dynamic model for random graphs. For all pairs i, j let $x_{i,j}$ be selected uniformly from $[0, 1]$, the choices mutually independent. Imagine p going from 0 to 1. Originally, all potential edges are "off". The edge from i to j (which we may imagine as a neon light) is turned on when p reaches $x_{i,j}$ and then stays on. At $p = 1$ all edges are "on". At time p the