[12] J.E. Graver, J. Yackel, Some graph theoretic results associated with Ramsey's theorem. *J. Combinatorial Theory* 4 (1968), 125-175

[13] S. Janson, T. Luczak and A. Rucinski. An exponential bound for the probability of nonexistence of specified subgraphs of a random graph, in *Procedings of Random Graphs '87*, M. Karonski et. al. eds, J. Wiley 1990, 73-87

[14] J.H. Kim, The Ramsey Number $R(3,t)$ has Order of Magnitude $t^2/\log t$, (preprint, 1994)

[15] M. Krivelevich, Bounding Ramsey numbers through large deviation inequalities, (preprint, 1994)

[16] C. McDiarmid, On the method of bounded differences, in J. Siemons, ed., *Surveys in Combinatorics 1988*, Cambridge Univ. Pr. 1989, 148-188

[17] N. Pippenger, J. Spencer, Asymptotic behavior of the chromatic number for hypergraphs, *J. Combinatorial Theory (Ser. A)*, 51 (1989), 24-42

[18] V. Rödl, On a packing and covering problem, *European Journal of Combinatorics*, 5 (1985), 69-78

[19] E. Shamir, J. Spencer, Sharp concentration of the chromatic number on random graphs $G_{n,p}$. *Combinatorica* 7 (1987), 121-129

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

[21] N. Wormald, Differential Equations for Random Processes and Random Graphs, (preprint 1994)

Author Address:
Courant Institute
251 Mercer St.
New York, NY 10012
USA
spencer@cs.nyu.edu

$t \to \infty$. Kim shows that the solutions $a(t), b(t), c(t)$ given above remain asymptotically valid up to $t = n^\gamma$, for γ a small but absolute constant. To be sure, this is considerably more difficult then showing validity for a finite time interval. The first method is not strong enough for this, he uses (basically) the second approach. Keeping bounds on the error terms brought on by the randomness requires mastery of the martingale inequalities. At $t = n^\gamma$ some $cn^{3/2} \ln^{1/2} n$ edges have been accepted to G. A random graph with this many edges has no independent set of size $k = Cn^{1/2} \ln^{1/2} n$. To be sure, this graph is anything but random. Still Kim shows that for any k-set S the probability that S remains independent is basically what it would be were G random. This yields the solution to a sixty year old problem, the asymptotics of $R(3, k)$.

References

[1] M. Ajtai, J. Komlós, E. Szemerédi, A note on Ramsey numbers, *J. Combinatorial Theory (Ser A)*, 29 (1980), 354-360

[2] N. Alon, J. Spencer, *The Probabilistic Method*, John Wiley, New York, 1991

[3] B. Bollobás, *Random Graphs*, Academic Press, London, 1985

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

[5] R. Boppana, J. Spencer, A Useful Elementary Correlation Inequality, *J. Combinatorial Theory - Ser. A* 50 (1989), 305-307

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

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

[8] P. Erdős and H. Hanani, On a limit theorem in combinatorial analysis, *Publ. Math. Debrecen*, 10 (1963), 10-13

[9] P. Erdős and A. Rényi. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kut. Int. Közl* 5 (1960), 17-61

[10] P. Erdős, P. Tetali, Representations of integers as the sum of k terms, *Random Structures & Algorithms*, 1 (1990), 245-261

[11] D. Grable, More-Than-Nearly-Perfect Packings and Partial Designs, (preprint 1994)

denote the number of triangles containing e in S, and for $e = \{i, j\} \in S$ and designated i let $N(e, i)$ denote the number of k with $\{i, k\} \in S$ and $\{j, k\} \in G$. (In this case we call $\{i, j\}, \{i, k\}$ a cherry – if one is born the other dies.) Suppose that for every i

$$\deg_S(i) \sim a(t)n$$

and for every $e = \{i, j\} \in S$

$$\deg_\Delta(e) \sim b(t)n$$

$$N(e, i) \sim c(t)n^{1/2}$$

Now add an infinitesmal time dt and consider expectations. Each surviving e is in $\sim 2c(t)n^{1/2}$ cherries (half from each end) so with probability $2c(t)dt$ one of the other edges will be born and so e dies with probability $2c(t)dt$. Of the $a(t)n$ edges containing a given i an expected $2a(t)c(t)ndt$ die. Thus

$$a'(t) = -2a(t)c(t)$$

Similarly of the $b(t)n$ triangles containing e an expected $2b(t)(2c(t)dt)$ will be "destroyed" in that one of their edges will be born so

$$b'(t) = -4b(t)c(t)$$

Of the $c(t)n^{1/2}$ cherries containing e at i, $2c^2(t)n^{1/2}dt$ will be lost by having the other edge die but $b(t)n^{1/2}dt$ new cherries are created when an old triangle containing e has the edge not through i born. Thus

$$c'(t) = -2c^2(t) + b(t)$$

Further at time zero S is the complete graph so we have initial conditions $a(0) = 1 = b(0)$, $c(0) = 0$, yielding a unique solution. This has a nice solution in terms of $G(t)$ given by 18. Then

$$a(t) = e^{-G(t)^2} = G'(t) \qquad b(t) = a(t)^2, \qquad c(t) = G(t)a(t) \tag{19}$$

At time t proportion $a(t)$ of the pairs are surviving so $\frac{1}{2}n^{1/2}a(t)dt$ pairs are accepted by time $t + dt$ so the expected total edges in G_t is $\frac{1}{2}n^{3/2}\int_0^t a(x)dx = \frac{1}{2}n^{3/2}G(t)$.

All this is lead in to our exciting finish. Jeong Han Kim [14] has found (up to constants) the asymptotics of $R(3, k)$. He has improved the Erdős lower bound to $R(3, k) > c\frac{k^2}{\ln k}$. Reversing parameters he shows the existance of a trianglefree graph G on n vertices with no independent set of size $k = Cn^{1/2}\ln^{1/2} n$. The method (at least from this author's vantagepoint) is to consider dynamically the random trianglefree graph as described above. At time t it has $\sim \frac{1}{2}n^{3/2}G(t)$ edges. From 18 one sees $G(t) \sim \ln^{1/2} t$ as with $\{i,k\},\{j,k\}$ already in G – so that it would be accepted if born now. Let $S = S_t$ be the graph of surviving pairs. Let $g_n(c)$ be the probability that any particular e (they all look alike) is surviving at time t.

We define a continuous time branching process that will mirror the fate of e above. Begin at time c with a single "Eve". Split $[0,c]\times[0,c]$ into infinitesmal squares $[x,x+dx]\times[y,y+dy]$. With probability $dx\cdot dy$ Eve gives birth to twins with birthtimes x,y. Equivalently, Eve gives birth to X pairs of twins with X having Poisson distribution with mean c^2 and given the number of births all birthdates are independent (twins are not born at the same time) and uniform in $[0,c]$. A child born at time x then gives birth by the same process in $[0,x]\times[0,x]$. A rooted tree T is thus generated and it can be shown that with probability one T is finite. We call vertices of T surviving or dying as follows. All childless vertices are surviving. A vertex is dying if and only if it has a (at least one) birth where both twins are surviving. Working up from the leaves of the tree every vertex of T is so designated. Let $g(c)$ denote the probability that the root Eve survives.

We give a rough argument that $\lim_{n\to\infty} g_n(c) = g(c)$. For $e=\{v,w\}$ we look at those u for which $x_{vu}\le c$ and $x_{wu}\le c$. There are $n-2$ potential u, each independently has this property with probability $(cn^{-1/2})^2$, so the number is asymptotically Poisson with mean c^2. Given that, the actual birthtimes are uniform in $[0,c]$. We then consider uv, uw twins of $e=vw$. We continue this process building up a tree. The analogy fails if some edge is child to two edges but this can be shown to occur with probability $o(1)$. Working backwards from the leaves one sees that an edge f is placed in G exactly when, considered as a vertex of T, it survives as described.

We find $g(c)$ by a differential equation. The difference $g(t)-g(t+dt)$ is the probability that Eve has no twins both born before t (probability $g(t)$) then has a pair of twins one of which is born in $[t,t+dt)$ (probability $2t\cdot dt$) and then they both survive. The twin born in $[t,t+dt)$ has probability $\sim g(t)$ to survive. The other is born uniformly in $[0,t]$ so its expected probability to survive is the average of $g(x)$ over the interval. This yields the differential equation

$$g'(t) = -2tg^2(t)\frac{1}{t}\int_0^t g(x)dx \tag{16}$$

or, setting $G(t)=\int_0^t g(x)dx$,

$$G''(t) = -2(G'(t))^2G(t) \tag{17}$$

With initial conditions $G(0)=0$ and $G'(0)=g(0)=1$ this has a unique solution given implicitly by

$$t = \int_0^{G(t)} e^{x^2}dx \tag{18}$$

Here is a second approach to the same result. At a given time t let $\deg_S(i)$ denote the number of neighbors of vertex i in S, for $e=\{i,j\}\in S$ let $\deg_\Delta(e)$

A's mutually independent we replace this with $\sum \Pr[A_{i_1}]\cdots\Pr[A_{i_s}]$. This sum over all $i_1,\ldots,i_s \in \{1,\ldots,m\}$ is precisely μ^s and each desired term has been counted $s!$ times. $\Box$.

Krivelevich's Proof: Let $G \sim G(n,p)$ with $p = \epsilon n^{-1/2}$ and set $k = Kn^{1/2}\ln n$ with $\epsilon = .1$ and $K = 10^6$ for definiteness though any moderately small ϵ and very large K would do. Let F be a (any) maximal family of edge disjoint triangles of G and let $G^* = G - \bigcup F$, i.e., G with all edges of all triangles of F removed. G^* is certainly trianglefree. For any k-set of vertices S the number X_S of edges of $G|_S$ has Binomial Distribution $B(\binom{k}{2},p)$. Elementary large deviation results give

$$\Pr[X_S \le \frac{1}{2}p\binom{k}{2}] < (0.9)^{\binom{k}{2}p}$$

and since (estimating $\binom{k}{2} \sim \frac{1}{2}k^2$ and $\binom{n}{k} \le n^k$)

$$\binom{n}{k}(0.9)^{\binom{k}{2}p} < \left[n(0.9)^{pk/2}\right]^k \ll 1$$

almost surely *all* S have at least $\frac{1}{4}pk^2 \sim \frac{1}{4}\epsilon K^2 n^{1/2}\ln^2 n$ edges. Again fix S and consider all potential triangles efg (listing the edges) with $e \subset S$. For each let A_{efg} be the event that they all lie in G so that $\Pr[A_{efg}] = p^3$. There are $\binom{k}{2}(n-k) + \binom{k}{3} \sim \frac{1}{2}k^2 n$ such events so

$$\mu = \sum_{e\subset S}\Pr[A_{efg}] \sim \frac{1}{2}k^2 np^3 = k\left(\frac{1}{2}\epsilon^3 K^2 \ln n\right)$$

Events A_{efg} are mutually independent when their edge sets are disjoint. From the Lemma above the probability that there are 3μ edge-disjoint triangles efg in G with $e \subset S$ is less than $\mu^{3\mu}/(3\mu)!$ and as

$$\binom{n}{k}\frac{\mu^{3\mu}}{(3\mu)!} < n^k(0.95)^{3\mu} < \left[n(0.95)^{\frac{3}{2}\epsilon^2 K^2 \ln n}\right] \ll 1$$

almost surely for *every* S there are less than 3μ such triangles and therefore $\bigcup F$ will have less than 9μ edges in S. Having picked ϵ, K so that $9\mu < \frac{1}{2}p\binom{k}{2}$ the elimination of these edges makes no S independent. Thus $R(3,k) > n$ or, reversing variables, $R(3,k) = \Omega(k^2/\ln^2 k)$. $\Box$

We improve this classic result by thinking *dynamically*.

Consider the following random dynamic process to form a trianglefree graph G on vertices $1,\ldots,n$. To each pair $e = \{i,j\}$ assign a birthtime $x_e \in [0, n^{1/2}]$, independently and uniformly. At time zero G is empty. When e is born it is added to G if its addition does not create a triangle. We say e is accepted in that case, rejected otherwise. Let G_t denote G at time t. A pair $\{i,j\}$ is surviving at time t if it has not been born and there is no k

Certainly the above argument needs work to be made formally correct. But suppose its correctness and now consider a vertex v. Let $g(T)$ be the probability that $v \in L_T$. We get 14 as before so that $g(T) \rightarrow 0$. But in each infinitesmal time interval $[t, t+dt)$ the probabilities that v is in a newly born E and that v is killed off are in the ratio y_v to $1 - y_v$ – i.e., conditioning on $v \in L_{t+dt} - L_t$, $\Pr[v \in E \in P_{t+dt}] \sim y_v$. Thus $\Pr[v \in E \in P_T] \sim y_v(1-g(T))$. Summing over all vertices v, the expected size

$$E[|\bigcup P_T|] \sim (1 - g(T)) \sum_v y_v = (Q+1)(1-g(T)) \sum_E x_E$$

so that as $T \rightarrow \infty$ the expected size of P_T approaches $\sum_E x_E$ as desired.

2.2 Ramsey $R(3, k)$

The Ramsey function $R(l, k)$ is defined as the minimal n so that any graph on n vertices must contain either a clique of size l or an independent set of size k. Existence of such n is Ramsey's Theorem itself. Asymptotics of the Ramsey function (and its numerous generalizations) have been closely linked with probabilistic methods from the beginning.
Theorem (Erdős (1947)[6]:

$$\binom{n}{k} 2^{-\binom{k}{2}} < 1 \Rightarrow R(k, k) > n \tag{15}$$

Proof. Take the random graph $G \sim G(n, p)$ with $p = \frac{1}{2}$. Then $\binom{n}{k} 2^{-\binom{k}{2}}$ is the expected number of cliques and independent sets of size k. When this number is less than one then with positive probability it is zero so that $R(k, k) > n$. □

Here we concentrate on $l = 3$ and the asymptotics as $k \rightarrow \infty$. The basic upper bound, from the proof of Ramsey's Theorem, was $R(3, k) \leq \binom{k+1}{2}$ which was lowered to $O(k^2 \frac{\ln \ln k}{\ln k})$ by Graver and Yackel [12] in 1968 and then to $O(\frac{k^2}{\ln k})$ by Ajtai, Komlós and Szemerédi [1] in 1980. A lower bound $R(3, k) > n$ means that there exists a trianglefree graph G on n vertices with no independent k set. After a number of "false starts" a lower bound $R(3, k) = \Omega(\frac{k^2}{\ln^2 k})$ was shown by Erdős [7] in 1961. This paper displays a remarkable combination of insight and technical skill. Over the decades, as new techniques have emerged, a number of authors have reproven this result. My own effort [20] in 1977 used the Lovász Local Lemma. Perhaps the most elementary proof is due to Krivelevich[15], we repeat it here in essentially complete form. We use an elementary and quite useful lemma from [10].
Lemma: Let $A_1, \ldots, A_m$ be events with $\sum \Pr[A_i] = \mu$. Then the probability that there exist s of the events, say $A_{i_1}, \ldots, A_{i_s}$ which are mutually independent events and which all hold is at most $\mu^s / s!$.
Proof. We bound $\sum \Pr[A_{i_1} \wedge \ldots \wedge A_{i_s}]$ over all such $i_1, \ldots, i_s$. With the

require all $x_E \in \{0,1\}$ this yields the packing $P = \{E : x_E = 1\}$ and the solution, denoted by ν, is the size of the maximal packing. Thus $\nu \leq \nu^*$.
Kahn's Theorem: For all Q and all $\epsilon > 0$ there exists $\delta > 0$ so that if x_E is a feasible solution to the above system with

$$\sum_{v,w \in E} x_E < \delta$$

for all distinct v, w then there is a packing P with

$$|P| \geq (1-\epsilon) \sum_{E \in H} x_E$$

Roughly speaking, Kahn's Theorem says that under appropriate side conditions $\nu \sim \nu^*$. We shall indicate an argument for Kahn's Theorem by creating an appropriate continuous time process. We are given the hypergraph H and the values x_E. It will be convenient to set

$$y_v = \sum_{v \in E} x_E$$

so that all $y_v \in [0,1]$. Give each E independently a birthdate t_E such that given E has not been born by t its probability of being born in the next infinitesimal time dt is $x_E dt$. Formally this is the exponential distribution, $\Pr[t_E > c] = e^{-cx_E}$.

As with Pippenger's theorem we dynamically keep a set $L = L_t$ and let S_t (the surviving edges) be the restriction of H to L_t. Again if $E \in S_t$ is born in infinitesimal time interval $[t, t+dt)$ it is added to the packing P so that $P_{t+dt} = P_t \cup \{E\}$. Now, however, we introduce the possibility of killing a vertex v. For each $v \in L_t$ define $y_v(t) = \sum_{v \in E \in S_t} x_E$. We think of this as the weighted degree of v at time t. Set $f(t) = (1+Qt)^{-1}$ as before. Now we kill v in the time interval $[t, t+dt)$ with probability $(f(t) - y_v(t))dt$. (If this is negative v is not killed.) Killing v means v is removed from L so $L_{t+dt} = L_t - \{y\}$ and therefore all $E \in S_t$ containing y are no longer surviving. The claim now is that at time t most v have

$$y_v(t) \sim f(t) y_v$$

As $f(0) = 1$ this holds for $t = 0$, now assume it holds for t. Any $w \in L_t$ is part of a newly born (in $[t, t+dt)$) E with probability $y_w(t)dt$ and is killed with the compensating probability $(f(t) - y_w(t))dt$ so it is removed from L with probability $f(t)dt$. Given that v itself remains in L each of its edges E has probability $Qf(t)dt$ of having a vertex lost, which would subtract x_E from $\deg(v)$. Then the expected total loss in the weighted degree is $y_v(t)(Qf(t)dt) \sim Qy_v f^2(t)dt$. Since $f(t)$ satisfies 12 the expected new value of the weighted degree is $\sim y_v f^2(t+dt)$ as desired.

Now let $g(t)$ be the probability that $v \in L_t$. Given $v \in L_t$ it has probability $\sim f(t)dt$ of being in an edge now placed in P so that $g(t+dt) \sim g(t)(1-f(t)dt)$. Letting $h(t) = \ln g(t)$ we have $h(t+dt) \sim h(t) - f(t)dt$ and $h(0) = 0$ so that

$$g(c) = e^{h(c)} = e^{-\int_0^c f(t)dt} = (1+Qc)^{-1/Q} \tag{14}$$

matching the previous results.

This approach has advantages and disadvantages. The main disadvantage is the difficulty of proving its validity. As the random process continues there will be more and more variance from expected behavior. It must be shown that the accumulated errors do not overwhelm the actual values. In essense we are dealing here with a stochastic differential equation. Indeed, proofs that the solution to this equation accurately portrays the situation look much like the original Rödl nibble. To examine the situation at $t = c$ we split the time interval into $c\epsilon^{-1}$ intervals of some very small length ϵ. (Each time interval ϵ is a nibble.) With ϵ very small the solution to the corresponding difference equation is close to the solution of the differential equation. Suppose in time ϵ the expected change of a degree is some αD with α, ϵ comparable. Roughly speaking the variance in that change will go like $(\alpha D)^{1/2}$. We need $D\epsilon \gg 1$ so that the variance is small compared to the expected change.

A big advantage of this approach is that it can be extended past any finite time. Consider the Erdős-Hanani as a $(Q+1)$-uniform hypergraph on $v = \Theta(n^k)$ vertices , regular of degree $D = \Theta(n^{k-l})$ with all codegrees $O(D/n)$. At finite time T the proportion of uncovered vertices (l-sets in the original formulation) is $O(T^{-1/Q})$ for T large. Now suppose the differential equation can be shown to remain valid up to time n^γ for some positive constant γ. Then we get an improvement on Rödl's Theorem. The dynamic algorithm then gives a packing so that the proportion of uncovered l-sets is $O(n^{\gamma'})$ for a calculatable positive constant γ'. It isn't so easy – extending the range of validity of the differential equation to T a function of n requires great care with the errors introduced from all sources. However, this approach has been used with success by N. Wormald [21] and, independently, D. Grable [11].

A generalization of Pippenger's Theorem has been given by J. Kahn in unpublished work. Let H be a $(Q+1)$-uniform hypergraph. Consider the following linear programming problem on real variables x_E, E ranging over the edges of H.

$$\begin{gathered} \text{maximize } \textstyle\sum_{E \in H} x_E \\ \text{given } \textstyle\sum_{v \in E} x_E \leq 1 \text{ for all } v \in V(H) \\ \text{and all } 0 \leq x_E \leq 1 \end{gathered}$$

A feasible solution to the above system is called a fractional packing. We let ν^* denote the solution to this linear programming problem. If we also

Together with the initial condition $g(0) = 1$ this has the unique solution

$$g(c) = (1 + Qc)^{-1/Q} \tag{11}$$

As $\lim_{c\to\infty} g(c) = 0$ this gives a proof of Rödl's Theorem.

One can put these results into a more general (and perhaps more natural) context. Let H be a $Q+1$-uniform hypergraph on v vertices. Suppose H is nearly regular in the sense that $\deg(e) \sim D$ for every vertex e where $D \to \infty$. Define the *codegree* of e, f to be the number of edges containing them both and assume that all codegrees are $o(D)$. (Formally, we may consider an infinite sequence of such hypergraphs, Q fixed, with asymptotics defined as the structures become bigger.) N. Pippenger (as given in [17]) has shown that under these circumstances there exists a packing P of $\sim v/(Q+1)$ disjoint edges. To translate the Erdős-Hanani situation into this context create a hypergraph H_n whose vertices are the l-element subsets of $\Omega = \{1, \dots, n\}$ and whose edges are $\{e \subset E : |e| = l\}$ for each k-set $E \subset \Omega$. Then $Q + 1 = \binom{k}{l}$, H is regular with $D = \binom{n-l}{k-l}$ and the codegrees are all $O(n^{k-l-1}) = o(D)$. The proof of Pippenger's generalization via continuous time branching processes is essentially as before. Now we give each edge E a birthdate unifromly distributed in $[0, D]$. Again given a vertex e and time c we generate a tree T to determine if e is covered. Again the analogy may fail and the condition on the codegrees turns out to be precisely what is neded to show that this occurs with probabiity $o(1)$.

Now, sticking with the hypergraph format, we describe another means toward the same end. Again we have a $Q+1$-uniform hypergraph with all $\deg(v) \sim D$ and all codegrees $o(D)$, to each edge E we assign a uniformly distributed birthdate $x_E \in [0, D]$ and let $P = P_t$ be the packing at time t. Let $L = L_t$ be the complement of $\bigcup P$. Let $S = S_t$ (for surviving) be the restriction of H to L_t. Let $\deg_t(v)$ denote the degree of v in S_t, defined for $v \in L_t$. The idea now is to find a function $f(t)$ so that almost surely most $\deg_t(v) \sim f(t)D$.

Suppose there is such a function $f(t)$. Consider the evolution of S from t to $t+dt$ with respect to a vertex $v \in L_t$. Most $w \in L_t$ lie in $\sim f(t)D$ edges of S_t and each edge is born with probability dt/D (and if born it is added to P) so with probability $\sim f(t)dt$ w is removed from L. Consider an edge $E \in S_t$ containing v. Conditioning on v itself remaining in L there is probability $\sim Qf(t)dt$ that $E \notin S_{t+dt}$ as any of the other vertices could be removed. Thus the degree of v will drop by an expected amount $(f(t)D)(Qf(t)dt)$, giving $\deg_{t+dt}(v) \sim D(f(t) - Qf^2(t)dt)$. This yields the differential equation

$$f'(t) = -Qf^2(t) \tag{12}$$

for f which, given the initial condition $f(0) = 1$, has the unique solution

$$f(t) = (1 + Qt)^{-1} \tag{13}$$

birth are called wombmates. (Littermates is the biologist's term for animals but English lacks a word for humans except when $Q = 2$. Note "siblings" is quite different.) The children are born mature and have births by the same random process as do their children and so forth. A rooted tree T is thus generated and it can be shown that with probability one T is finite. We call vertices of T surviving or dying as follows. All childless vertex are surviving. A vertex is dying if and only if it has a (at least one) birth all of whose wombmates are surviving. Working up from the leaves of the tree every vertex of T is so designated. Let $g(c)$ denote the probability that the root Eve survives.

We claim that the limit (as $n \to \infty$) of the probability e is not covered is $g(c)$. To see (informally) the mirror fix an l-set e and start at time c, time going backwards. Identify e with Eve. When a k-set $A \supset e$ is born consider this a birth of the Q l-sets $f \subset A$, $f \neq e$. There are $\binom{n-l}{k-l}$ potential births so in infinitesmal time dt there is probability dt of having such a birth. (The total number of births is given by a Binomial distribution and a central aspect of the asymptotics is estimation of the Binomial by the Poisson.) Once f has been born the birth of a k-set $B \supset f$ is considered as a birth of the Q l-sets $g \subset B$, $g \neq f$. A tree T is thus generated. (Actually, it *may* happen that a k-set A is born which contains two (or more) l-sets f in which case our analogy fails. This, however, can be shown to occur with probability $o(1)$.) T determines if e is covered. If T consists only of root e then no $A \supset e$ were born so e is not covered. We show by induction on the size of T that e is not covered if and only if it survives in T as defined above. If e survives then for every $A \supset e$ there is an $f \subset A$, $f \neq e$ that does not survive. The rooted subtree at f is the tree generated starting at f at time x_A. By induction f did not survive so there was a $B \supset f$ with $x_B < x_A$ that was added to P. Then at time x_A A was not added to P. This holds for all A so e is not covered at time c. Inversely if e does not survive then there is an $A \supset e$ so that all $f \subset A$, $f \neq e$, do survive. By induction at time x_A no such f has been covered. Either e is already covered or A is now placed in P, so e is covered by or at time x_A, either way e is covered by time c.

Now we can focus attention on $g(c)$, a totally continuous problem for which the pesky n has disappeared. We consider this a function of c and compare $g(c)$ with $g(c + dc)$ for infinitesmal dc. The difference in Eve's survival chances are if she had no births up to time c for which all children survived but then has a birth in the time interval $[c, c + dc)$ for which all children survive. Thus $g(c) - g(c + dc)$ is roughly $g(c)(dc)g(c)^Q$, reflecting the three factors. This can be made precise (we've skipped the necessary first step, showing that $g(c)$ is continuous) and g can be shown to satisfy the differential equation

$$g'(c) = -g(c)^{Q+1} \tag{10}$$

2 Dynamic Algorithms

2.1 Asymptotic Packing

For $2 \leq l < k < n$ let $m(l,k,n)$ be the maximal size of a family of P of k-element subsets of an n-set Ω such that every l points lie in *at most* one $A \in P$. Such P are naturally called packings. Our concern will be for l, k fixed, $n \to \infty$. Elementary counting gives

$$m(l,k,n) \leq \frac{\binom{n}{l}}{\binom{k}{l}} \tag{8}$$

with equality if and only if there is a design with every l points lying in a unique $A \in P$. For $l = 2, k = 3$ these are the famous Steiner Triple Systems and for $l = 2$ and any fixed k now classic results of R. Wilson give asymptotic necesary and sufficient conditions for the existence of such designs. The situation for $l > 2$ is much less well understood. In 1961 Paul Erdős and Haim Hanani [8] asked whether 8 holds asymptotically – i.e.:

$$\lim_{n\to\infty} m(l,k,n)\frac{\binom{k}{l}}{\binom{n}{l}} = 1 \tag{9}$$

This conjecture was proven by V. Rödl [18] in 1985 by a technique often called the Rödl nibble.

Recent years have seen a reevaluation of Rödl's Theorem from the viewpoint of random dynamic algorithms. Take all $\binom{n}{k}$ k-sets and order them randomly. Now create a packing P dynamically, beginning with $P = \emptyset$. We consider the k-sets E in order. We add E to P if possible. More precisely, E is added to P if and only if there is no F already in P overlapping E in at least l points. This certainly will create a packing P but the real result is that P will have expected size as desired.

We turn this into a continuous time dynamic process as follows. To each of the $\binom{n}{k}$ k-sets E we assign a birthtime x_E. The x_E are chosen independently, each a uniformly chosen real number in $[0, \binom{n-l}{k-l}]$. (This choice of interval length will make calculations convenient shortly.) Time starts at zero with $P = \emptyset$. When E is born it is added to P if possible, as before. Of course, the E are considered in random order so that the *final* value of P has the same distribution as before. We consider P_c, the value of P at time c, where c is a fixed real. We say an l-set e is covered at time c if $e \subset E$ for some $E \in P_c$. We now want the probability e is so covered.

We define a continuous time branching process that mirrors the fate of e above. Begin at time c with a single "Eve". Time goes continuously backwards to zero. Eve gives birth with a unit density Poisson process – in infinitesmal time dt she has probability dt of giving birth. All births are to precisely Q children where $Q = \binom{k}{l} - 1$, the children in the same of H. Then the distribution $\chi(H_p)$ satisfies the concentration 4. Somewhat more generally suppose for each $1 \leq i < j \leq n$ there is a $p_{ij} \in [0,1]$ and consider the random graph G with i,j adjacent with probability p_{ij}, the adjacencies mutually independent events. Again, for any choice of p_{ij} and any vertex Lipschitz X the random variable $X(G)$ satisfies the concentration 4.

We call graph function X edge Lipschitz if changing any single edge (from in to out or out to in) can change $X(G)$ by at most one. Set $m = \binom{n}{2}$. We can decompose $G \sim G(n,p)$ as the product of m Binary choices so that Azuma's Inequality gives

$$\Pr[|X(G) - \mu| \geq \lambda m^{1/2}] < 2e^{-\lambda^2/2} \tag{5}$$

where $\mu = E[X]$. Bollobás [4] used this to give a remarkable bound on the clique number $\omega(G)$. Fix $p = \frac{1}{2}$ for definiteness. Set Y_k equal the number of k-cliques and

$$f(k) = E[Y_k] = \binom{n}{k} 2^{-\binom{k}{2}}$$

Elementary analysis shows that $f(k_0) > 1 > f(k_0+1)$ for $k_0 \sim 2\log_2 n$ and, for $k \sim k_0$, $f(k+1)/f(k) = n^{-1+o(1)}$. As $\Pr[\omega(G) \geq k] \leq E[Y_k]$ almost surely $\omega(G) < k_0+2$. Now set $k^- = k_0 - 3$ so that $f(k^-) > n^{3-o(1)}$. Bollobás showed

$$\Pr[\omega(G) < k^-] < 2^{cn^2 \ln^{-8} n} \tag{6}$$

This is "near" best possible in that with probability 2^{-cn^2} the graph has no edges whatsoever. To prove this set X equal to the maximal size of a family of edge-disjoint cliques of size k^-. Note $X = 0$ if and only if $\omega(G) < k^-$ and that X is edge Lipschitz. From less modern (though nontrivial) probabilistic methods one can show $\mu = E[X] > cn^2 \ln^{-4} n$. Now 6 follows from 5 by setting $\lambda = \mu m^{-1/2}$. Its interesting to note that the same result (with a different power of $\ln n$ in the exponent) can be derived directly from the Extended Janson Inequality. From this Bollobás showed that the chromatic number $\chi(G)$ was almost surely $\sim 2\log_2 n$.

We conclude with a variant of Azuma's Inequality used in the work J.H. Kim discussed in §2.2. Let I_i, $1 \leq i \leq m$, be mutually independent identically distributed indicator random variables with $E[I_i] = p$. (E.g., I_i is the indicator for the i-th edge in $G(n,p)$.) Let X be a function of the I_i (e.g., a graph function) such that changing I_i can change X by at most c_i. Set $\sigma = [p \sum c_i^2]^{1/2}$. (If $X = \sum c_i I_i$ then $Var[X] < \sigma^2$ and σ is like a standard deviation.) Then

$$\lambda \max(c_i) \leq 2\sigma^2 \ln 2 \Rightarrow \Pr[|X - E[X]| \geq \lambda] < 2e^{-\lambda^2/4\sigma^2} \tag{7}$$

For $p = o(1)$ this is much tighter that the basic Azuma bound and the use of the c_i allows a clear sense of the weighting of influences of different potential edges.

1.2 Martingale Inequalities

Martingales have a long history in probability theory but their usefulness in our context is quite new. We refer to Colin McDiarmid's excellent survey [16] at this meeting for a more detailed examination. For our purposes we consider a martingale to be a sequence $X_0, \ldots, X_m$ of random variables (on a common space) so that for any $0 \leq i < m$ and value a $E[X_{i+1}|X_i = a] = a$. We further assume $X_0 = \mu$, a constant. Then $\mu = E[X_i]$ for all i.
Azuma's Inequality: Let $\mu = X_0, X_1, \ldots, X_m = X$ be a martingale in which $|X_{i+1} - X_i| \leq 1$. Then

$$\Pr[X > \mu + a] < \exp(-a^2/2m) \tag{3}$$

In application we use an isoperimetric version. Let $\Omega = \prod_{i=1}^{m} \Omega_i$ be a product probability space and X a random variable on it. Call X *Lipschitz* if whenever $\omega, \omega' \in \Omega$ differ on only one coordinate $|X(\omega) - X(\omega')| \leq 1$. Set $\mu = E[X]$.
Azuma's Perimetric Inequality: $\Pr[X \geq \mu + a] < e^{-a^2/2m}$.

The connection is via the Doob Martingale, $X_i(\omega)$ being the conditional expectation of X given the first i coordinates of ω. The same inequality holds for $\Pr[X \leq \mu - a]$.

E. Shamir and this author [19] applied this result to the chromatic number $\chi(G)$ of the random graph $G \sim G(n,p)$. (Again [2] gives a general discussion.) Let Ω be the probability space, whose vertices, i.e., graphs, may be thought of as Boolean arrays of length $\binom{n}{2}$. Let $X : \Omega \to Z$ be chromatic number. For $2 \leq i \leq n$ let Ω_i be the restriction of the graph to the pairs $\{j, i\}, 1 \leq j < i$. We may think of Ω_i as $i - 1$ values of the full Boolean array or as the "information" about vertex i looking to the "left". Now X is Lipschitz since we can make any change to the edges involving vertex i and it can only increase X by at most one since we can always give i a new color. This yields a strong concentration result:

$$\Pr[|\chi(G) - \mu| \geq \lambda(n-1)^{1/2}] < 2e^{-\lambda^2/2} \tag{4}$$

so that, roughly, the chromatic number is concentrated within $n^{1/2}$ of its expectation. An oddity of this method is that it does not by itself give the value of the expectation, it only deduces that the random variable is tightly concentrated around its expectation.

We can generalize 4 considerably. We call a graph function X vertex Lipschitz if changing the edges at one vertex can only change $X(G)$ by at most one. Then 4 holds with χ replaced by any vertex Lipschitz X. Further we can alter the probability measure (holding to the set of graphs on $\{1, \ldots, n\}$ as our objects) as long as the component parts Ω_i are mutually independent. For example, let H be a fixed (not necessarily random) graph on $\{1, \ldots, n\}$ and let H_p denote the random subgraph given by selecting edges from H with independent probability p and selecting no edges outside with $\mu = \binom{n}{3}p^3 \sim c^3/6$. We bound Δ by noting that we only need consider terms of the form $A_{ijk} \wedge A_{ijl}$ as otherwise the edge sets do not overlap. There are $O(n^4)$ choices of such i, j, k, l. For each the event $A_{ijk} \wedge A_{ijl}$ is that a certain five edges (ij, ik, jk, il, jl) belong to $G(n, p)$, which occurs with probability p^5. Hence

$$\Delta = \sum \Pr[A_{ijk} \wedge A_{ijl}] = O(n^4 p^5)$$

With $p = c/n$ we have $\epsilon = O(n^{-3}) = o(1)$ and $\Delta = o(1)$ so that the Janson Inequality gives an *asyptotic formula*

$$\Pr[TF] \sim M \sim e^{-\frac{c^3}{6}}$$

This much could already be done in the original work of Erdős and Rényi by calculation of moments. But the Janson Inequalities allow us to proceed beyond $p = \Theta(1/n)$. The calculation $\Delta = o(1)$ had plenty of room. For any $p = o(n^{-4/5})$ we have $\Delta = o(1)$ and therefore an asymptotic formula $\Pr[TF] \sim M$. For example, if $p = \Theta\left((\ln n)^{1/3}/n\right)$ this yields that $G(n, p)$ has polynomially small probability of being trianglefree. Once p reaches $n^{-4/5}$ the value Δ becomes large and we no longer have an asymptotic formula. But as long as $p = o(n^{-1/2})$ we have $\Delta = O(n^4 p^5) = o(n^3 p^3) = o(\mu)$ and so we get the *logarithmically asymptotic* formula

$$\Pr[TF] = e^{-\mu(1+o(1))} = e^{-\frac{n^3 p^3}{6}(1+o(1))}$$

Once p reaches $n^{-1/2}$ we lose this formula. But now the Extended Janson Inequality comes into play. We have $\mu = \Theta(n^3 p^3)$ and $\Delta = \Theta(n^4 p^5)$ so for $p \gg n^{-1/2}$

$$\Pr[TF] < e^{-\Omega(\mu^2/\Delta)} = e^{-\Omega(n^2 p)}$$

The Extended Janson Inequality gives, in general, only an upper bound. In this case, however, we note that $\Pr[TF]$ is at least the probability that $G(n, p)$ has no edges whatsoever and so, for $n^{-1/2} \ll p \ll 1$

$$\Pr[TF] > (1-p)^{\binom{n}{2}} = e^{-\Omega(n^2 p)}$$

With a bit more care, in fact, one can estimate $\Pr[TF]$ up to a constant in the logarithm for all p. These methods do not work just for trianglefreeness. In a remarkable paper Andrzej Rucinski, Tomasz Luczak and Svante Janson [13] have examined the probability that $G(n, p)$ does not contain a copy of H, where H is any particular fixed graph, and they estimate this probability, up to a constant in the logarithm, for the entire range of p. Their paper was the first and still one of the most exciting applications of the Janson Inequality.

We make the following *correlation assumptions*:
• for all i, S with $i \notin S$

$$\Pr[A_i | \wedge_{j \in S} \overline{A}_j] \leq \Pr[A_i]$$

• for all i, k, S with $i, k \notin S$

$$\Pr[A_i \wedge A_k | \wedge_{j \in S} \overline{A}_j] \leq \Pr[A_i \wedge A_k]$$

Finally, let ϵ be such that $\Pr[A_i] \leq \epsilon$ for all i.
The Janson Inequality: Under the above assumptions

$$M \leq \Pr[\wedge \overline{A}_i] \leq M e^{\frac{(1-\epsilon)\Delta}{2}} \tag{1}$$

We set

$$\mu = \sum \Pr[A_i],$$

the expected number of A_i that occur. As $1 - x \leq e^{-x}$ for all $x \geq 0$ we may bound $M \leq e^{-\mu}$ and then rewrite the upper bound in the somewhat weaker but quite convenient form

$$\Pr[\wedge \overline{A}_i] \leq e^{-\mu + \frac{(1-\epsilon)\Delta}{2}}$$

In most applications $\epsilon = o(1)$ and the pesky factor of $1-\epsilon$ is no real trouble. Indeed just assuming all $\Pr[A_i] \leq \frac{1}{2}$ is plenty for all cases we know of. In many cases we also have $\Delta = o(1)$. Then the Janson Inequality gives an asymptotic formula for $\Pr[\wedge \overline{A}_i]$. When $\Delta \gg \mu$, as also occurs in some important cases, the above gives an upper bound for $\Pr[\wedge \overline{A}_i]$ which is bigger than one. In those cases we sometimes can use the following:
The Extended Janson Inequality: Under the assumptions of the Janson Inequality and the additional assumption that $\Delta(1-\epsilon) \geq \mu$

$$\Pr[\wedge \overline{A}_i] \leq e^{-\frac{\mu^2(1-\epsilon)}{2\Delta}} \tag{2}$$

In our application the underlying probability space will be the random graph $G(n,p)$. The events A_α will all be of the form that $G(n,p)$ contains a particular set of edges E_α. The correlation assumptions are then an example of far more general result called the *FKG inequalities*. We have a natural dependency graph by making A_α, A_β adjacent exactly when $E_\alpha \cap E_\beta \neq \emptyset$.

Let us parametrize $p = c/n$ and consider the property , call it TF, that G is triangle free. Let A_{ijk} be the event that $\{i,j,k\}$ is a triangle in G. Then

$$TF = \wedge \overline{A}_{ijk},$$

the conjunction over all triples $\{i,j,k\}$. We calculate

$$M = \left(1 - p^3\right)^{\binom{n}{3}} \sim e^{-\mu}$$

Modern Probabilistic Methods in Combinatorics
Joel Spencer

The *probabilistic method* is a means to prove the existence of configurations by showing that an appropriately defined random configuration has a positive probability of having the desired property. The method is approaching its golden anniversary, its beginning generally considered a three-page paper by Paul Erdős [6] in 1947. Closely aligned is the study of *random graphs*, more generally random configurations, in which problems about probabilities concerning random graphs are considered for their own sake. This topic began in 1961 with the monumental study of Paul Erdős and Alfred Rényi [9], "On the evolution of random graphs". For many years the uses of probability in these twin topics was surprisingly elementary, linearity of expectation, variance and the Chernoff bounds could take a fledgling researcher a long long way. Recent years have seen more sophisticated uses of probability and our emphasis here will be on the newer probabilistic methodologies and how they are applied to these topics. We give our recent book [2] and the book of Bollobás [3] on Random Graphs as a general references.

1 Exponential Haystacks

1.1 Janson Inequalities

Let $A_1, \ldots, A_m$ be events in a probability space. Set

$$M = \prod_{i=1}^{m} \Pr[\overline{A_i}]$$

The Janson Inequality allows us, sometimes, to estimate $\Pr[\wedge \overline{A_i}]$ by M, the probability if the A_i were mutually independent. The original proof by Svante Janson is in [13]. See [5] for a more "elementary" proof and [2] for general discussion. We let G be a dependency graph for the events – i.e., the vertices are the indices $i \in [m]$ and each A_i is mutually independent of all A_j with j not adjacent to i in G. (This notion was first used with the Lovász Local Lemma. While the dependency graph is not uniquely defined there is usually a clear candidate.) We write $i \sim j$ when i, j are adjacent in G. We set

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