1 The Erdős-Rényi Phase Transition

In their great 1960 work *On the Evolution of Random Graphs*, Paul Erdős and Alfred Rényi expressed a special interest in the behavior of $\Gamma_{n,N(n)}$, the random graph with n vertices and $N(n)$ edges, when $N(n)$ was near $\frac{n}{2}$:

> Thus the situation may be summarized as follows: the largest component of $\Gamma_{n,N(n)}$ is of order $\log n$ for $\frac{N(n)}{n} \to c < \frac{1}{2}$, of order $n^{2/3}$ for $\frac{N(n)}{n} \to c \sim \frac{1}{2}$, and of order n for $\frac{N(n)}{n} \to c > \frac{1}{2}$. This double "jump" of the size of the largest component when $\frac{N(n)}{n}$ passes the value $\frac{1}{2}$ is one of the most striking facts concerning random graphs.

Striking, indeed. The past half century have certainly confirmed the excitement that Erdős and Rényi expressed in their discovery.

1.1 An Overview

We favor the more modern viewpoint, examining the random graph $G(n,p)$. The behavior of Erdős and Rényi's $\Gamma_{n,N(n)}$ then corresponds to that of $G(n,p)$ with $p = N(n)/\binom{n}{2}$. We shall assume $p = \Theta(n^{-1})$ throughout this chapter.

We shall call

$$p = \frac{c}{n} \tag{1}$$

the *coarse* parametrization. The value $\frac{1}{2}$ in the Erdős-Rényi formulation corresponds to the value $c = 1$ in our parametrization. Values $c < 1$ and $c > 1$ give $G(n,p)$ that are essentially different. We shall call

$$p = \frac{1}{n} + \lambda n^{-4/3} \tag{2}$$

the *fine* parametrization. The importance of this parametrization is not *a priori* at all obvious. Indeed, its "discovery" was one of the great advances in the field. In §1.7 we give a heuristic argument why this is the appropriate fine parametrization. Along with the fine parametrization we also define

$$\epsilon = \lambda n^{-1/3} \text{ so that } p = \frac{1+\epsilon}{n} \tag{3}$$

We shall express various results in terms of either λ or ϵ (or both), whichever best illustrates the result. We shall think of ϵ, λ as functions of n. To avoid negative numbers we shall sometimes parametrize $p = \frac{1-\epsilon}{n}$ with $\epsilon = \lambda n^{-1/3}$.

This includes functions such as $p = \frac{1}{n} - 100n^{0.01}n^{-4/3}$. Of course, for n small this would give $p < 0$ and so would be nonsense. For n sufficiently large we will have $p \in [0,1]$. As our results are always asymptotic we shall allow this slight abuse of notation and consider $G(n,p)$ defined only for n appropriately large.

In describing the nature of $G(n,p)$ we shall refer to the complexity of components, as defined below. Observe that complexity zero and one correspond to tree components and unicyclic components respectively.

Definition 1 *A connected component of a graph G with v vertices and e edges is said to have* complexity $e-v+1$. *Components with complexity zero or one are called* simple*; components with complexity greater than one are called* complex.

Let $C(v)$ denote the component containing a given vertex v. Its size $|C(v)|$ has a distribution. From the symmetry of $G(n,p)$ the distribution of all $|C(v)|$ are the same. We shall be concerned with the sizes of the largest components. We shall let C_i denote the i-th largest component and L_i denote its number of vertices. Thus $L_1 = \max_v |C(v)|$. We shall be particularly interested in L_1, L_2 and whether or not they are close together.

The study of $G(n,p)$ when $p = \Theta(n^{-1})$ splits into five regions. We describe them in order of increasing p, thus giving some sense of the evolution.
`Very Subcritical.` Here we employ the coase parametrization $p = \frac{c}{n}$ and assume c is a constant with $c < 1$. Example: $p = \frac{1}{2n}$.

1. All components are simple.
2. $L_1 = \Theta(\ln n)$.
3. $L_k \sim L_1$ for all fixed k.

`Barely Subcritical.` Here we employ the fine parametrization. $p = \frac{1-\epsilon}{n}$ with $\epsilon = \lambda n^{-1/3}$. We assume $\epsilon = o(1)$. We assume that $\lambda \to \infty$. Example: $p = \frac{1}{n} - n^{-4/3}n^{0.01}$.

1. All components are simple.
2. $L_1 = \Theta(\epsilon^{-2}\ln(\lambda)) = \Theta(n^{2/3}\lambda^{-2}\ln(\lambda))$.
3. $L_k \sim L_1$ for all fixed k.

`The Critical Window.` Here λ is a real constant. Example: $p = \frac{1}{n} - 2n^{-4/3}$. The value $\lambda = 0$, perhaps surprisingly, has no special status. The largest k components (k fixed) all have size $L_k = \Theta(n^{-4/3})$. Parametrizing $L_k = c_k n^{2/3}$ and letting d_k denote the complexity of C_k there is a nontrivial joint distribution for $c_1, \ldots, c_k, d_1, \ldots, d_k$.
`Barely Supercritical.` Here we employ the fine parametrization. We assume $\epsilon, \lambda > 0$. We assume $\epsilon = o(1)$. We assume that $\lambda \to +\infty$. Example: $p = \frac{1}{n} + n^{-4/3}n^{0.01}$.

1. $L_1 \sim 2\epsilon n = 2\lambda n^{2/3}$.
2. The largest component has complexity approaching infinity.

3. All other components are simple

4. $L_2 = \Theta(\epsilon^{-2} \ln(\lambda)) = \Theta(n^{2/3}\lambda^{-2} \ln(\lambda))$.

Note that the ratio L_1/L_2 goes to infinity. For this reason, in this regime we call the largest component the *dominant component*.

`Very Supercritical.` We employ the coarse parametrization and assume $c > 1$.

1. $L_1 \sim yn$ where $y = y(c)$ is that positive real satisfying the implicit equation
$$e^{-cy} = 1 - y \tag{4}$$

2. The largest component has complexity approaching infinity.

3. All other components are simple.

4. $L_2 = \Theta(\ln n)$.

Following the terminology made famous by Erdős and Rényi, we call the largest component the *giant component*.

We shall give arguments for only some of the above statements, and then often in limited form. Other results are given in the exercises. Full arguments for these results, and much much more, can be found in the classic texts of Bollobas [**?**] and of Janson, Łuczak and Rucinski [**?**].

1.2 Three Processes

We place here in concise form three classes of probability spaces that we shall contrast and analyze.

• The Poisson Branching Model

Parameter: Nonnegative real c

Underlying Space: An infinite sequence Z_t, $t = 1, 2, \ldots$ of independent identically distributed random variables, each having Poisson Distribution with mean c.

Auxilliary Y_t, $t \geq 0$, given by initial value $Y_0 = 1$ and recursion $Y_t = Y_{t-1} + Z_t - 1$.

Auxilliary T: T is that minimal t with $Y_t = 0$. If no such t exists we write $T = \infty$.

Nomenclature: Z_t is the number nodes born at time t, Y_t is the queue size at time t, T is the total size.

Interpretation: T is the total size of a Galton-Watson process, as described in §1.3, using a Poisson distribution with mean c.

• The Binomial Branching Model

Parameters: Positive Integer m, Real $p \in [0, 1]$.

Underlying Space: An infinite sequence Z_t, $t = 1, 2, \ldots$ of independent identically distributed random variables, each having Binomial Distribution $B[m, p]$

Auxilliary Y_t, $t \geq 0$, given by initial value $Y_0 = 1$ and recursion $Y_t = Y_{t-1} + Z_t - 1$.

Auxilliary T: T is that minimal t with $Y_t = 0$. If no such t exists we write

$T = \infty$.
Nomenclature: Z_t is the number of nodes born at time t, Y_t is the queue size at time t, T is the total size.
Interpretation: T is the total size of a Galton-Watson process, as descirbed in §1.3 using a Binomial distribution with parameters m, p.
• The Graph Branching Model
Parameters: Positive Integer n, Real $p \in [0, 1]$.
Underlyng Space: A sequence $Z_1, \ldots, Z_n$. Z_t has Binomial Distribution with parameters N_{t-1}, p, with N_{t-1} as given below.
Auxilliary Y_t, $t \geq 0$, given by initial value $Y_0 = 1$ and recursion $Y_t = Y_{t-1} + Z_t - 1$.
Auxillary N_t, $t \geq 0$, given by initial value $N_0 = n - 1$ and recursion $N_t = N_{t-1} - Z_t$. Equivalently: $N_t = n - t - Y_t$.
Auxilliary T: T is that minimal t with $Y_t = 0$ or, equivalently, $N_t = n - t$. $1 \leq T \leq n$ always.
Nomenclature: Z_t is the number of nodes born at time t, Y_t is the queue size at time t, N_t is number of neutral vertices at time t, T is total size.
Interpretation: T is the size of the component $C(v)$ of a given vertex v in $G(n, p)$, as found by the BFS process described in §1.5

We use the superscripts *po* (Poisson), *bin* (Binomial), and *gr* (graph) to distinguish these three processes when necessary.

1.3 The Galton-Watson Branching Process

Let Z be a distribution over the nonnegative integers. The Galton-Watson process begins with a single root rode, we can call her Eve. Eve has Z children. Each of her children (if there are any) now independently have Z children. The process continues, each new offspring having an independent number Z children. Let T be the total number of nodes (including Eve herself) created in the process. It is possible that the process goes on forever, in which case we write $T = \infty$.

Our analysis of the Galton-Watson process uses fictional continuation. Let Z_i, $i = 1, 2, \ldots$, be a countable sequence of independent identically distributed variables, each having distribution Z. This defines our probability space. We think of the children being born in a Breadth First Search manner. That is: Eve has her children which are ordered in some way. Now the children, in order, have children. Each child's children are ordered in some way and this gives an ordering of Eve's grandchildren. Now the grandchildren have children in order, and the process continues. We count Eve as node number 1, her children have node numbers $2, \ldots, 1 + Z_1$ and, more generally, each node is given a distinct poistive integer as its node number. We let Z_i be the number of children of the i-th node. Since the Z_i are independent and have distribution Z this corresponds to the Galton-Watson process. Imagine the i-th node having Z_i children and then dying. By time t we mean the process after the t-th node has had her children and died. Let Y_t be the number of living children at time t. We set initial value $Y_0 = 1$, corresponding to the node Eve. We have the recursion

$$Y_t = Y_{t-1} + Z_t - 1 \text{ for all } t \geq 1 \tag{5}$$

There are two essentially different cases:
• $Y_t > 0$ for all $t \geq 0$. In this case the Galton-Watson process goes on forever and $T = \infty$.
• $Y_t = 0$ for some $t \geq 0$. In this case let T be the *least* integer for which $Y_T = 0$. Then the Galton-Watson process stops with the death of the T-th node and T is the total number of nodes in the process.

Our fictional continuation enables us to consider the Y_t as an infinite random walk, with step size $Z - 1$. When $c < 1$ the walk has negative drift and so tends to minus infinity. When $c > 1$ the walk has positive drift and tends to plus infinity. The process when $c < 1$ is called *subcritical* and the process when $c > 1$ is called *supercritical.* When $c = 1$ the walk has zero drift and the situation is especially delicate.

The above is quite general. When Z is Poisson or Binomial (the only cases of interest to us) this yields the Poisson branching process and the Binomial branching process of §1.2.

1.4 Analysis of the Poisson Branching Process

In this section we study $T = T_c^{po}$. We often drop the value c and the superscript *po* for notational simplicity.

Theorem 1.1 *If $c < 1$, T is finite with probability one. If $c = 1$, T is finite with probability one. If $c > 1$ then T is infinite with probability $y = y(c)$ where y is that unique positive real satisfying the equation (4)*

`Proof:` Suppose $c < 1$. If $T > t$ then $Y_t > 0$ so that $Z_1 + \ldots + Z_t \geq t$. Chernoff bounds give that $\Pr[Y_t > 0] < e^{-kt}$ for a constant k. In particular, $\Pr[Y_t > 0] \to 0$ so that $\Pr[T > t] \to 0$ and T is finite with probability one.

Suppose $c \geq 1$. Set $z = 1 - y = \Pr[T < \infty]$. Given that Eve has i children the probability that the branching process is finite is z^i as all i branches must be finite. Thus

$$z = \sum_{i=0}^{\infty} \Pr[Z_1 = i] z^i = \sum_{i=0}^{\infty} e^{-c} \frac{c^i z^i}{i!} = e^{c(z-1)}$$

Setting $y = 1 - z$ gives the equation (4). For $c = 1$, $e^{-y} > 1 - y$ for $y > 0$ so the solution must be $y = 0$. For $c > 1$ the function $f(y) = 1 - y - e^{-cy}$ has $f(0) = 1$, $f(1) < 0$ and $f'(0) = c - 1 > 0$ so there is a $y \in (0, 1)$ with $f(y) = 0$. Further, as f is convex, there is precisely one y. We have shown that either $\Pr[T < \infty] = 1$ or $\Pr[T < \infty] = y > 0$. The argument that $\Pr[T < \infty] \neq 1$ (not surprising as the walk has positive drift) is left for the Exercises.

Theorem 1.2 *For any positive real c and any integer k, setting $T = T_c^{po}$,*

$$\Pr[T = k] = \frac{e^{-ck}(ck)^{k-1}}{k!}$$

We defer the proof of this classic result to §1.6 when we will give a probabilistic proof!

When $c = 1$ Stirling's Formula gives

$$\Pr[T_1 = k] = \frac{e^{-k}k^{k-1}}{k!} \sim \frac{1}{\sqrt{2\pi}}k^{-3/2} \tag{6}$$

This perforce approaches zero but it does so only at polynomial speed. In general

$$\Pr[T_c = k] \sim \frac{1}{\sqrt{2\pi}}k^{-3/2}c^{-1}(ce^{-c})^k$$

For any $c \neq 1$ (whether larger or smaller than one) $ce^{-c} < 1$ and therefore $\Pr[T_c = k]$ approaches zero at exponential speed. This gives a bound on the tail distribution

$$\Pr[T_c \geq u] < e^{-u(\alpha+o(1))} \tag{7}$$

where $\alpha = c - \ln c > 0$.

We are particularly interested in the Poisson branching process when c is *near* one. Let us parametrize

$$c = 1 + \epsilon$$

When $\epsilon > 0$, $\Pr[T_{1+\epsilon} = \infty]$ is that $y = y(\epsilon) \in (0,1)$ satisfying $f(y) = 1 - y - e^{-(1+\epsilon)y} = 0$. Some fun Calculus gives

$$\Pr[T_{1+\epsilon} = \infty] \sim 2\epsilon \text{ as } \epsilon \to 0^+ \tag{8}$$

Suppose $c \to 1^+$ so that $\epsilon \to 0^+$. We have

$$\ln(ce^{-c}) = \ln(1+\epsilon) - \epsilon \sim -\frac{\epsilon^2}{2}$$

Thus

$$\Pr[T_{1+\epsilon} = u] \sim \frac{1}{\sqrt{2\pi}}u^{-3/2} \text{ for } u = o(\epsilon^{-2})$$

Note that $\Pr[T_{1+\epsilon} = u] \sim \Pr[T_1 = u]$ in this range. When u reaches order ϵ^{-2} there is a change. For $u = A\epsilon^{-2}$ and fixed A:

$$\Pr[T_{1+\epsilon} = A\epsilon^{-2}] \sim \frac{1}{\sqrt{2\pi}}\epsilon^3 A^{-3/2}e^{-A/2}$$

When $A \to \infty$ we absorb smaller factors into the exponential term:

$$\Pr[T_{1+\epsilon} = A\epsilon^{-2}] = \epsilon^3 e^{-(1+o(1))A/2}$$

When c is slightly less than one we can write $c = 1 - \epsilon$ where $\epsilon \to 0^+$. We have $\ln(ce^{-c}) \sim -\frac{1}{2}\epsilon^2$, the same as for $c = 1 + \epsilon$. Indeed when $u = o(\epsilon^{-3})$:

$$\Pr[T_{1-\epsilon} = u] \sim \Pr[T_{1+\epsilon} = u]$$

For $A \to \infty$:

$$\Pr[T_{1-\epsilon} = A\epsilon^{-2}] = \epsilon^3 e^{-(1+o(1))A/2}$$

The Poisson branching processes with means $1 + \epsilon$ and $1 - \epsilon$ look almost the same, with the (important!) distinction is that the mean $1 + \epsilon$ process is sometimes infinite while the mean $1 - \epsilon$ process never is.

In short: the Poisson branching process with mean $1 \pm \epsilon$ acts as if it had mean 1 until reaching size on the order ϵ^{-2}. Until then $\Pr[T_{1\pm\epsilon}] = u$ is dropping at a polynomial rate. Upon reaching order ϵ^{-2} $\Pr[T_{1\pm\epsilon}] = u$ drops exponentially in u.

We are particularly interested in the tail distribution. For $\epsilon \to 0^+$ and $A \to \infty$

$$\Pr[T_{1-\epsilon} > A\epsilon^{-2}] < e^{-(1+o(1))A/2}\epsilon \tag{9}$$

The same holds for the *finite* part of $T_{1+\epsilon}$

$$\Pr[\infty > T_{1+\epsilon} > A\epsilon^{-2}] < e^{-(1+o(1))A/2}\epsilon \tag{10}$$

When $A \to \infty$ this quantity is $o(\epsilon)$ so (8) gives

$$\Pr[T_{1+\epsilon} > A\epsilon^{-2}] \sim 2\epsilon \text{ when } \epsilon \to 0^+ \text{ and } A \to \infty \tag{11}$$

1.5 The Graph Branching Model

`Abbreviation:` We use BFS as an abbreviation for Breadth First Search. BFS algorithms are a mainstay of Computer Science and central to our approach.

Let $C(v)$ denote the component, in $G(n,p)$, containing a designated vertex v. We generate $C(v)$ using the (standard) BFS algorithm to find $C(v)$. We begin with root v. In this procedure all vertices will be live, dead, or neutral. The live vertices will be contained in a queue. Initially, at time zero, v is live, the queue consists of one vertex, v itself, and all other vertices are neutral. At each time t we remove a live vertex w from the top of the queue (in Computer Science parlance we "pop the queue") and check all pairs $\{w, w'\}$, w' neutral, for adjacency in G. The popped vertex w is now dead. Those neutral w' (if any) adjacent to w are added to the bottom of the queue and are now live. (They can be placed in any particular order.) The procedure ends when the queue is empty. We let T denote that time. At time T all vertices are neutral or dead and the set of dead vertices is precisely the component $C(v)$. That is, $T = |C(v)|$.

Let Z_t denote the number of vertices added to the queue at time t. Let Y_t denote the size of the queue at the conclusion of time t. We set $Y_0 = 1$, reflecting the initial size of the queue. At time t we remove one vertex and add Z_t vertices to the queue so we have the recursion $Y_t = Y_{t-1} + 1 - Z_t$ Let N_t denote the number of neutral vertices at time t. As Z_t vertices switched from neutral to live at time t, N_t satisfies the recursion $N_0 = n-1$, $N_t = N_{t-1} - Z_t$. Equivalently, as there are t dead and Y_t live vertices at time t, $Z_t = n - t - Y_t$. Z_t is found by checking N_{t-1} pairs for adjacency. As these pairs have not yet been examined they remain adjacent with independent probability p. That is,

$$Z_t \sim B[N_{t-1}, p] \sim B[n - (t-1) - Y_{t-1}, p] \tag{12}$$

The Graph Branching Process of §1.2 mirrors the above analysis until time T and then continues until time n. This fictional continuation shall be useful in the analysis of $C(v)$. The graph branching process is similar

to a Binomial branching process in that the Z_t have binomial distributions but dissimilar in that the parameter N_{t-1} in the graph branching process depends on previous values Z_i.

As $N_t = N_{t-1} - Z_t$, (12) yields $N_t \sim B[N_{t-1}, 1-p]$. By induction we find the distributions

$$N_t \sim B[n-1, (1-p)^t] \text{ for } 0 \leq t \leq n \tag{13}$$

If $T = t$ it is necessary (though not sufficient, due to fictitious continuation) that $N_t = n - t$. This yields the useful inequalities:

Theorem 1.3 *In* $G(n,p)$

$$\Pr[|C(v)| = t] \leq \Pr[B[n-1, (1-p)^t] = n - t \tag{14}$$

or, equivalently,

$$\Pr[|C(v)| = t] \leq \Pr[B[n-1, 1-(1-p)^t] = t - 1 \tag{15}$$

`An Alternate Analysis` The following analysis of $C(v)$ on $G(n,p)$ has been explored by van der Hofstad and Spencer [**?**]. Each $w \neq v$ flips a coin, heads with probability p, reapeatedly until getting a head. Let X_w denote that flip on which w gets a head. Suppose $X_w = j$. Then w enters the BFS at time j. (However, it may have missed the boat if the BFS has already terminated.) This reverses the usual randomness, we are here imagining the $w \neq v$ trying to get into the BFS tree, rather than the BFS tree trying to expand by finding neutral vertices. Suppose $t = |C(v)|$. Every $w \neq v$ which is in $C(v)$ must have entered by time t so $X_w \leq t$. Every $w \neq v$ which is not in $C(v)$ had t opportunities to enter $C(v)$ and so $X_w > t$. Thus $\Pr[|C(v)| = t]$ is at most the probability that $X_w \leq t$ for *precisely* $t-1$ $w \neq v$. For each $w \neq v$, $\Pr[X_w = t] = 1 - (1-p)^t$ and these events are independent over w, yielding (15). In (**??**) this analysis is extended to give more precise bounds on $\Pr[|C(v)| = t]$.

1.6 The Graph and Poisson Processes Compared

Set $p = \frac{c}{n}$. A key observation is that $Z_1 \sim B[n-1, \frac{c}{n}]$ approaches (in n) the Poisson distribution with mean c. Further, in a more rough sense, the same holds for Z_t as long as $N_{t-1} = o(n)$ or, equivalently, the number of live and dead vertices is $o(n)$. That is, *the generation of* $C(v)$ *mimics the Poisson branching process with mean* c *as long as the number of vertices found is not too large.* This allows for a very accurate description in the Very Subcritical regime $c < 1$. But in the Very Supercritical regime $c > 1$ the relationship between the generation of $C(v)$ and the Poisson branching process breaks down. As the number N_{t-1} of neutral vertices drops so does the expected number $E[Z_t]$ of vertices added to the queue. Eventually the drift of the walk Y_t lowers from positive to negative, and this eventually causes the process to halt. We call this phenomenon *the ecological limitation.* Indeed, there must be an ecological limitation. The Poisson branching process becomes infinite with positive probability, the component $C(v)$ tautologically cannot be greater than n.

Theorem 1.4 *For any positive real c and any fixed integer k*

$$\lim_{n\to\infty} \Pr[|C(v)| = k \text{ in } G(n, c/n)] = \Pr[T_c = k]$$

Proof: Let Z_t^{po}, T^{po} and $Z_t^{gr}, Z_t^{gr}, T^{gr}$ denote the values in the Poisson branching process with parameter c and the graph branching process with parameters n, p repsectively. Let Γ denote the set of k-tuples $\vec{z} = (z_1, \ldots, z_k)$ of nonnegative integers such that the recursion $y_0 = 1, y_t = y_{t-1} + z_t - 1$ has $y_t > 0$ for $t < k$ and $y_k = 0$. Then

$$\Pr[T^{gr} = k] = \sum \Pr[Z_i^{gr} = z_i, 1 \le i \le k]$$

$$\Pr[T^{po} = k] = \sum \Pr[Z_i^{po} = z_i, 1 \le i \le k]$$

where both sums are over $\vec{z} \in \Gamma$. Fix such a $\vec{z}$.

$$\Pr[Z_i^{gr} = z_i, 1 \le i \le k] = \prod_{i=1}^{k} \Pr[B[Z_{i-1}^{gr}, p] = z_i]$$

As i, y_{i-1}, z_i are fixed $Z_{i-1} = n - O(1)$ and

$$\lim_{n\to\infty} \Pr[B[Z_{i-1}, p] = z_i] = \Pr[Z_i^* = z_i]$$

Further, as the products are of a fixed number of terms

$$\lim_{n\to\infty} \Pr[Z_i^{gr} = z_i, 1 \le i \le k] = \Pr[Z_i^{po} = z_i, 1 \le i \le k]$$

giving the theorem.

Now we prove Theorem 1.2. From Theorem 1.4,

$$\Pr[T_c^{po} = k] = \lim_{n\to\infty} \Pr[|C(v)| = k]$$

where the second probability is in $G(n, p)$ with $p = \frac{c}{n}$ and v is an arbitrary vertex of that graph. There are $\binom{n}{k-1}$ choices for $S := C(v)$. On any particular S there is probability $O(p^n) = O(n^{-n})$ that $G(n, p)$ has more than $n-1$ edges. If $G(n, p)$ has precisely $n-1$ edges on S they must form a tree. There are k^{k-2} such trees. Each occurs with probability $p^{k-1}(1-p)^{\binom{k}{2}-k+1} \sim p^{k-1} = c^{k-1}n^{1-k}$. Thus the total probability that $G(n, p)$ restricted to S forms a connected graph is $\sim k^{k-2}c^{k-1}n^{1-k}$. For $S = C(v)$ we must further have no edges between S and its complement, this has probability $(1-p)^k(n-k) \sim e^{-ck}$. Thus

$$\Pr[C(v) = k] \sim \binom{n}{k-1} k^{k-2}c^{k-1}n^{1-k}e^{-ck} \to \frac{e^{-ck}(ck)^{k-1}}{k!}$$

as desired.

The graph branching process can be compared to the Binomial branching process in both directions. An important cautionary note: the event $T_{n-1,p}^{bin} \ge u$ in Theorem 1.5 (and similarly $T_{n-u,p}^{bin} \ge u$ in Theorem 1.6) includes the possibity that the Binomial branching process is infinite. Indeed, in application this will be the critical term.

Theorem 1.5 *For any u*

$$\Pr[T^{gr}_{n,p} \geq u] \leq \Pr[T^{bin}_{n-1,p} \geq u]$$

Proof: We modify the graph branching process by constantly replenishing the supply of neutral vertices. That is, when we pop the vertex w and there are $n-1-s$ neutral vertices we create s fictional vertices w' and allow w, w' to be adjacent with probability p. This gives a component of size $T^{bin}_{n-1,p}$, the actual $C(v)$ will be a subset of it. Thus $T^{bin}_{n-1,p}$ dominates $T^{gr}_{n,p}$

Theorem 1.6 *For any u*

$$\Pr[T^{gr}_{n,p} \geq u] \geq \Pr[T^{bin}_{n-u,p} \geq u]$$

Proof: We halt the graph branching process if and when the number of found (live plus dead) vertices reaches u. This does not effect the probability of finding at least u vertices. In this truncated graph process we diminish the number of neutral vertices to $n-u$. That is, when we pop the vertex w and there are $n-1-s \geq n-u$ neutral vertices we select $n-u$ of them and only allow adjacencies w, w' to them. The truncated graph process dominates this truncated Binomial $n-u, p$ process and so has a greater or equal probability of reaching u.

`The Poisson Approximation.` We are working in the range $p = \Theta(n^{-1})$. There the Binomial $B[n-1, p]$ distribution and the Poisson distribution with mean np are very close. The Poisson branching process is precisely understood and, we feel, the "purest" branching process. Our goal in this chapter is to give the reader a picture for the "why" of the various regimes. To do this we shall often avoid the technical calculations and simply assume that the Binomial $n-1, p$ branching process is very close to the Poisson branching process with mean np.

1.7 The Parametrization Explained

The question most frequently asked about the Erdős-Rényi phase transition is: Why is the exponent $-\frac{1}{3}$? Here is a heuristic that may (or may not!) be helpful.

Parametrize $p = \frac{1+\epsilon}{n}$ with $\epsilon = \epsilon(n)$ positive and approaching zero. We look for the following picture. Consider the Poisson branching process $T = T^{po}_{1+\epsilon}$. It is infinite with probability $\sim 2\epsilon$, otherwise its probability of exceeding $A\epsilon^{-2}$ drops exponentially in A. The graph branching process mimics the Poisson branching process as long as it is not too successful. The cases when the Poisson branching process is finite are mimicked, yielding components of size up to roughly ϵ^{-2}. The cases when the Poisson branching process is infinite are mimicked by components that "escape" until the ecological limitation sets in. These components all join together. They form a single component, the dominant component, of size $2\epsilon n$.

In order for the above (admittedly rough) picture to hold there needs be a distinction between the small components, up to size ϵ^{-2}, and the dominant component of size $2\epsilon n$. That is, we need $2\epsilon n \gg \epsilon^{-2}$. This heuristic leads us to $\epsilon = n^{-1/3}$ as the breakpoint. When $\epsilon \gg n^{-1/3}$ we have the distinction

between small and dominant and are in the supercritical regime. When $\epsilon = O(n^{-1/3})$ there is no effective analogy to the Poisson branching process being infinite, and there is no dominant component.

1.8 The Subcritical Regions

Let $p = \frac{c}{n}$ with $c < 1$. Theorem 1.5 gives

$$\Pr[T^{gr}_{n,p} \geq u] \leq \Pr[T^{bin}_{n-1,p} \geq u]$$

With the Poisson approximation, $\Pr[|C(v)| \geq u] \leq (1 + o(1)) \Pr[T_c \geq u]$. From (7) this drops exponentially in u. Taking $u = K \ln n$ for appropriately large K, $\Pr[|C(v)| \geq u] < n^{-1.01}$. As this holds for each of the n vertices v, the probability that *any* v v has $|C(v)| \geq u$ is less than $nn^{-1.01} \to 0$. That is, $L_1 = O(\ln n)$ with probability tending to one.

Lets push this argument into the barely subcritical regime $p = \frac{1-\epsilon}{n}$ with $\epsilon = \lambda n^{-1/3}$. Let I_v be the indicator random variable for $C(v)$ having at least u vertices, u to be determined below. As above Theorem 1.5 and our Poisson approximation give the bound

$$\Pr[|C(v)| \geq u] \leq (1 + o(1)) \Pr[T_{1-\epsilon} \geq u]$$

We now parametrize

$$u = K\epsilon^{-2} \ln \lambda = Kn^{2/3}\lambda^{-2} \ln \lambda$$

For an appropriately high constant K the bound (9) gives

$$\Pr[T_{1-\epsilon} \geq u] \leq \epsilon e^{-3.1\lambda} = \epsilon \lambda^{-3.1}$$

Let $X = \sum_v I_v$ be the number of vertices v in components of size at least u and let Y be the number of components of $G(n,p)$ of size at least u. Linearity of Expectation gives

$$E[X] = nE[I_v] \leq n\epsilon\lambda^{-3.1} = n^{2/3}\lambda^{-2.1}$$

As $Y \leq Xu^{-1}$,

$$E[Y] \leq u^{-1}E[X] \leq K^{-1}\lambda^{-0.1} \to 0$$

With probability approaching one, $Y = 0$ and so

$$L_1 \leq u = K\epsilon^{-2} \ln \lambda = Kn^{2/3}\lambda^{-2} \ln \lambda$$

1.9 The Supercritical Regimes

In the supercritical regimes there are two salient points about the giant or dominant component. First, it exists. Secondly, it is unique. Neither is trivial.

We start with the very supercritical region, $p = \frac{c}{n}$, with $c > 1$ constant. The ideas here will carry into the barely supercritical region. Let $y = y(c)$ be the positive real solution of the equation $e^{-cy} = 1 - y$. Let δ be an arbitrarily small constant and let K be an appropriately large constant. Set

$S = K \ln n$, $L^- = (y-\delta)n$, $L^+ = (y+\delta)n$. Call a component $C(v)$ and its size $|C(v)|$ small if $|C(v)| < S$, giant if $L^- < |C(v)| < L^+$, awkward otherwise.

`No Middle Ground.` We claim that the probability of having any awkward component is $o(n^{-20})$. (We could make 20 arbitrarily large by changing K.) There are n choices for v and n choices for $t = |C(v)|$. Thus it suffices to show that for any v and for any awkward t that $\Pr[|C(v)| = t] = o(n^{-18})$. From Theorem 1.3 it suffices to bound $\Pr[B[n-1, 1-(1-\frac{c}{n})^t] = t-1]$. We indicate the technical calculations. When $t = o(n)$ $1-(1-\frac{c}{n})^t \sim \frac{cn}{t}$ and $c > 1$ so $\Pr[B[n-1, 1-(1-\frac{c}{n})^t] \leq t-1]$ is exponentially small in t. As $t \geq K \ln n$ this is polynomially small in n. When $t \sim xn$, $1-(1-\frac{c}{n})^t \sim 1-e^{-cx}$. For $x \neq y$, $1-e^{-cx} \neq x$ so the mean of the Binomial is not near t and the probability that it is equal to t is exponentially small in n. In all cases the bounds on $\Pr[|C(v)| = t]$ follow from basic Chernoff bounds.

`Escape Probability.` Set $\alpha = \Pr[C(v)$ not small$]$. (When this happens we like to think that the BFS on $G(n,p)$ starting with root v has escaped an early death.) Theorems 1.5,1.6 sandwich

$$\Pr[T^{bin}_{n-S,p} \geq S] \leq \alpha \leq \Pr[T^{bin}_{n-1,p} \geq S]$$

From our Poisson approximation both distributions are the upper and lower bounds are asymptotic to $\Pr[T_c \geq S]$. Thus $\alpha \sim \Pr[T_c \geq S]$. As c is assummed fixed and $S \to \infty$,

$$\alpha \sim \Pr[T_c \geq S] \sim \Pr[T_c = \infty] = y$$

with y as in (4).

Because there is no middle ground, not small is the same as giant. $C(v)$ is giant with probability $\sim y$. Thus the expected number of vertices in giant components is $\sim yn$. Each giant component has size between $(y-\delta)n$ and $(y+\delta)n$. Our goal is a single giant component of size $\sim yn$. We are almost there. But maybe with probability one half there are two giant components.

`Sprinkling.` Set $p_1 = n^{-3/2}$. (Any p_1 with $n^{-2} \ll p_1 \ll n^{-1}$ would do here.) Let $G_1 \sim G(n, p_1)$ be selected independently from $G \sim G(n,p)$ on the same vertex set and let $G^+ = G \cup G_1$ so that $G^+ \sim G(n, p^+)$ with $p^+ = p + p_1 - pp_1$. (We "sprinkle" the relatively few edges of G_1 on G to make G^+.) Suppose $G(n,p)$ had more than one giant component and let V_1, V_2 be the vertex sets of two of those components. There are $\Omega(n^2)$ pairs $\{v_1, v_2\}$ with $v_1 \in V_1$, $v_2 \in V_2$. We have selected p_1 large enough so that with probability $1-o(1)$ at least one of these pairs is in the sprinkling G_1. Adding this edge merges components V_1, V_2 into a component of size at least $2y(1-\delta)n$ in G^+. We have selected p_1 small enough so that $p^+ \sim p = \frac{c}{n}$. The probability that G^+ has a component so large is therefore $o(1)$. Thus the probability that G had had more than one giant component is $o(1)$.

Finally, we make δ arbitrarily small. $G(n,p)$ has an expected number $\sim yn$ of points in giant components and giant components all have size $\sim yn$. Further, by the sprinkling argument, there is $o(1)$ probability of having more than one giant component. Thus with probability $1-o(1)$ there is precisely one giant component. This gives the salient features of the

very supercritical phase. There is a giant component so $L_1 \sim yn$. There is only one giant component and no middle ground so $L_2 \leq S = O(\ln n)$.

The Sprinkling for Complexity argument given below in the Barely Supercritical Phase can be easily modified to show that the giant component has high complexity, indeed, complexity $\Omega(n)$.

`The Barely Supercritical Phase.` Set $p = \frac{1+\epsilon}{n}$ with $\epsilon = \lambda n^{-1/3}$ and $\lambda \to \infty$. Note $\epsilon^{-2} = \lambda^{-2} n^{2/3} \ll 2\epsilon n$. The analysis of the barely supercritical region becomes more difficult as $\lambda = \lambda(n)$ approaches infinity more slowly. We shall add the simplifying assumption that $\lambda \gg \ln n$. Further we shall find somewhat weaker bounds than stated on L_2.

Let δ be an arbitrarily small constant and let K be an appropriately large constant. Set $S = K\epsilon^{-2} \ln n$, $L^- = (1-\delta)2\epsilon n$, $L^+ = (1+\delta)2\epsilon n$. Call a component $C(v)$ and its size $|C(v)|$ small if $|C(v)| < S$, dominant if $L^- < |C(v)| < L^+$, awkward otherwise.

`No Middle Ground.` We claim that the probability of having any awkward component is $o(n^{-20})$. (We could make 20 arbitrarily large by changing K.) There are n choices for v and n choices for $t = |C(v)|$. Thus it suffices to show that for any v and for any awkward t that $\Pr[|C(v)| = t] = o(n^{-18})$. Again we bound $\Pr[B[n-1, 1-(1-p)^t] = t-1]$. We indicate the technical calculations. Let μ and σ^2 denote the mean and of the Binomial. Then $\mu = (n-1)(1-(1-p)^t)$ and, in this range, $\sigma^2 \sim \mu$. When $t = o(n\epsilon)$ we estimate $1-(1-p)^t$ by $pt = t + t\epsilon$. Then $\mu - t \sim -t\epsilon$ and $\sigma^2 \sim t$. This probability is roughly $\exp[-(t\epsilon)^2/2t] = \exp[-t\epsilon^2/2]$. As $t \geq S$ the exponent is $o(n^{-18})$ for $K > 36$. (To push S down to $K\epsilon^{-2}\ln(\lambda)$ requires a finer bound on $\Pr[|C(v)| = t]$.) Now suppose $t \sim xn\epsilon$ where $x \neq 2$. The ecological limitation now has an effect and we estimate $1-(1-p)^t$ by $pt - \frac{1}{2}p^2t^2$ so

$$\mu - t \sim t\epsilon - \frac{1}{2}t^2 n^{-2} \sim (n\epsilon)(x - \frac{1}{2}x^2)$$

(Observe that when $x = 2$ the mean of the binomial is very close to t and so we do not get a small bound on $\Pr[|C(v)| = t$. This is natural when we consider that there will be a dominant component of size $\sim 2\epsilon n$.) Again $\sigma^2 \sim t$ so the probability is $\exp[-\Omega((n\epsilon)^2/t)]$ which is extremely small. When $t \gg n\epsilon$ the probability is even smaller.

`Escape Probability.` Set $\alpha = \Pr[C(v)$ not small$]$. Theorems 1.5,1.6 sandwich

$$\Pr[T^{bin}_{n-S,p} \geq S] \leq \alpha \leq \Pr[T^{bin}_{n-1,p} \geq S]$$

The Poisson approximation for $T^{bin}_{n-1,p}$ is $T_{1+\epsilon}$. As $S \gg \epsilon^{-2}$, bound (11) gives

$$\alpha \leq \Pr[T_{1+\epsilon} \geq S] \sim \Pr[T_{1+\epsilon} = \infty] \sim 2\epsilon$$

Replacing $n-1$ by $n-S$ lowers the mean by $\sim Sn^{-1}$. But $Sn^{-1}/\epsilon \sim (\ln n)/(n\epsilon^3) = \lambda^{-3}\ln n$ and we have made λ large enough that this is $o(1)$. That is, $Sn^{-1} = o(\epsilon)$. Therefore $T^{bin}_{n-S,p}$ is approximated by $T_{1+\epsilon-o(\epsilon)}$ and

$$\alpha \geq \Pr[T_{1+\epsilon+o(\epsilon)} \geq S] \sim \Pr[T_{1+\epsilon+o(\epsilon)} = \infty] \sim 2\epsilon$$

α has been sandwiched and $\alpha \sim 2\epsilon$.

Because there is no middle ground, not small is the same as dominant. $C(v)$ is dominant with probability $\sim 2\epsilon$. Thus the expected number of vertices in dominant components is $\sim 2n\epsilon$. Each giant component has size between $(1-\delta)2n\epsilon$ and $(1+\delta)2n\epsilon$. As in the very supercritical case, we need worry about having more than one giant component.

`Sprinkling.` Set $p_1 = n^{-4/3}$. Let $G_1 \sim G(n, p_1)$ be selected independently from $G \sim G(n,p)$ on the same vertex set and let $G^+ = G \cup G_1$ so that $G^+ \sim G(n, p^+)$ with $p^+ = p + p_1 - pp_1 = 1 + \epsilon + o(\epsilon)$. Suppose $G(n,p)$ had more than one giant component and let V_1, V_2 be the vertex sets of two of those components. There are $\gg n^{4/3}$ pairs $\{v_1, v_2\}$ with $v_1 \in V_1$, $v_2 \in V_2$. With probability $1 + o(1)$ at least one of these pairs is in the sprinkling G_1. Adding this edge merges components V_1, V_2 into a component of size at least $(1-\delta)4\epsilon n$ in G^+. The probability G^+ has such a large component is $\exp[-\Omega((n\epsilon)^2/t)]$ which is certainly $o(n^{-10})$. Thus the probability G had had two (or more) dominant components is $o(n^{-10})$. Taking δ arbitrarily small G has with probability $1 - o(1)$ precisely one dominant component. Thus $L_1 \sim 2n\epsilon$ and, as there is no middle ground, $L_2 \leq K\epsilon^{-2}\ln n$.

`Sprinkling for Complexity.` Take $p_1 = (1 + \frac{\epsilon}{2})/n$ and $p_2 \sim \frac{\epsilon}{2}/n$ so that $p_1 + p_2 - p_1p_2 = (1+\epsilon)/n$. Let $G_1 \sim G(n, p_1)$, $G_2 \sim G(n, p_2)$, and $G_3 = G_1 \cup G_2$ so that $G_3 \sim G(n, (1+\epsilon)/n)$. G_1, G_3 will have a dominant components V_1, V_3 of sizes $\sim n\epsilon$ and $\sim 2n\epsilon$. As G_3 has "no middle ground" in its component sizes, $V_1 \subseteq V_3$. Now the sprinkling G_2 adds $\sim p_2\binom{n\epsilon}{2} \sim n\epsilon^3/2$ edges internal to V_1. Thus V_3 will have complexity at least $n\epsilon^3/2 = \lambda^3/2$ which approaches infinity.

1.10 The Critical Window

We now fix a real λ and set $p = \frac{1}{n} + \lambda n^{-4/3}$. There has been massive study of this critical window, [**?**] and the monumental [**?**] being only two examples. Calculations in this regime are remarkably delicate.

Fix $c > 0$ 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!

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

For $i < k$,

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

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 $p^{k-1} = n^{k-1}(1+\lambda n^{-1/3})^{k-1}$ and expanding $\ln(1+\epsilon) = \epsilon - \frac{1}{2}\epsilon^2 + O(\epsilon^3)$:

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

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

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

Here $A = k - \frac{k^2}{2n} - \frac{c^3}{6} + \frac{\lambda k}{n^{1/3}} - \frac{\lambda^2 c}{2} - k + \frac{k^2}{2n} - \frac{\lambda k}{n^{1/3}} + \frac{\lambda c^2}{2} + o(1)$. Absorbing the $o(1)$ into the asymptotics we can give A the intriguing form

$$A = A(c)\frac{(\lambda - c)^3 - \lambda^3}{6}.$$

Stirling's Formula then yields

$$E[X] \sim n^{-2/3} e^{A(c)} c^{-5/2} (2\pi)^{-1/2}.$$

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

$$\lim_{n\to\infty} E[X] = \int_a^b e^{A(c)} c^{-5/2} (2\pi)^{-1/2}\ dc.$$

The large components are not all trees. [**?**] proved that for fixed $l \geq 0$ there are asymptotically $c_l k^{k-2+\frac{3}{2}l}$ connected graphs on k points with $k-1+l$ edges, where $c_0 = 1$, $c_1 = \sqrt{\pi/8}$ and c_l was given by a specific recurrence. Asymptotically in l, $c_l = l^{-l/2(1+o(1))}$. 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

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

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

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

where

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

a sum convergent for all c. A component of size $\sim cn^{2/3}$ will have probability $c_l c^{\frac{3}{2}l}/g(c)$ of having complexity l, independent of λ. As $\lim_{c\to 0} g(c) = 1$, most components of size $\epsilon n^{2/3}$, $\epsilon \ll 1$, are trees but as c gets bigger the distribution on l moves inexorably 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. This distribution has support the positive reals. Thus, for example, For $\lambda = -4$ there is some positive limiting probability that the largest component is bigger than $10n^{2/3}$ and for $\lambda = +4$ there is some positive limiting probability that the largest component is smaller than $0.1n^{2/3}$, though both these probabilities are miniscule. The $c^{-5/2}$ term dominates the integral as $c \to 0^+$, reflecting the notion that for any fixed 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., $\lambda = -4$) 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. When λ is large positive (e.g., $\lambda = +4$) the dominant component will have begun to emerge. The largest component is likely to be $\sim 2\lambda n^{2/3}$ and of moderately high (not zero nor one) complexity and the second largest component will be considerably smaller and simple.

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 $\lambda = -4$, 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 $\lambda = +4$ it is very likely that a dominant component, a 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.

1.11 Analogies to Classical Percolation Theory

The study of percolation has involved the intense efforts of both mathematicians and physicists for many years. A central object of that study has been bond percolation on Z^d, as described below. Here we explore, without proofs, the fruitful analogies between that percolation and the Erdős-Rényi phase transition. Grimmett's [**?**] *Percolation* is a classic text in this field and we shall follow its treatment.

Let $d \geq 2$. (All parameters below shall depend on the choice of d.) Let Z^d, as usual, represent the set of $\vec{a} = (a_1, \ldots, a_d)$ with a_i integers. The d-dimensional cubic lattice, written L^d, is that graph with vertices Z^d, two

vertices $\vec{a}, \vec{b}$ being adjacent if they agree on $d-1$ coordinates and differ by one on the other coordinate. Let $p \in [0,1]$. The random subgraph $L^d(p)$ contains each edge of L^d (and no others) with independent probability p. We let $C(\vec{a})$ denote the connected component of $L^d(p)$ containing the vertex $\vec{a}$. We generally examine $C(\vec{0})$ as, by symmetry, all $C(\vec{a})$ look the same. (In [**?**] and elsewhere the edges of L^d are called bonds and they are open with probability p and closed otherwise. The word cluster is used in place of connected component.) Naturally, as p becomes larger $L^d(p)$ will have more adjacencies. There is a critical probability, denoted by p_c, at which $L^d(p)$ undergoes a macroscopic change.

• For $p < p_c$, the subcritical region, all connected components are finite.

• For $p > p_c$, the supercritical region, there is precisely one infinite component.

• For $p = p_c$, at the critical point, the situation is particularly delicate, as discussed below.

The constant probabilities of bond percolation correspond to probabilities parametrized $p = c/n$ in the Erdős-Rényi $G(n,p)$. The value $c = 1$ is then the critical probability in the Erdős-Rényi model.

• The infinite component in the bond percolation model is analogous to giant components, components of size $\Omega(n)$, in the Erdős-Rényi model.

• The finite components in the bond percolation model are analogous to components of size $O(\ln n)$ in the Erdős-Rényi model.

The uniqueness of the infinite component in bond percolation was an open question (though the physicists "knew" it was true!) for many years. This corresponds to the uniqueness of the giant component in $G(n,p)$.

In the bond percolation model there are only three choices for p, it can be less than, greater than, or equal to p_c. The barely subcritical and barely supercritical phases of the Erdős-Rényi model correspond to an asymptotic study of the bond percolation model as p approaches p_c from below and from above respectively. This study is done through the use of *critical exponents* as described below.

Set $\theta(p) = \Pr[C(\vec{0}) \text{ is infinite}]$. For $p < p_c$, $\theta(p) = 0$ as with probability one there are no infinite components. For $p > p_c$, $\theta(p) > 0$. This corresponds to the infinite component having positive density, strengthening the analogy to the giant components of the Erdős-Rényi model. When p is barely greater than p_c there will be an infinite component but its density will be very small. The critical exponent β is that real number so that

$$\theta(p) = (p - p_c)^{\beta + o(1)} \text{ as } p \to p_c^+$$

(Conceivably, $\theta(p)$ could behave erratically as $p \to p_c^+$ and β might not exist. Indeed, for all the critical exponents discusses we should add the caveat "if they exist.") Analogously, in the Erdős-Rényi model $\theta(c)$ is the proportion of points in the giant component, that $y = y(c) > 0$ satisfying (4). From (8), $y(1+\epsilon) \sim 2\epsilon$ as $\epsilon \to 0^+$. Therefore, $\beta = 1$.

The susceptibility, denoted by $\chi(p)$ (not to be confused with chromatic number) is given by $\chi(p) = E[|C(\vec{0})|]$. For $p > p_c$, $\chi(p) = \infty$ as with positive probability $C(\vec{0})$ is infinite. For $p < p_c$, $\chi(p)$ is finite and $\chi(p) \to \infty$ as $p \to p_c^-$. That susceptibility approaches infinity at the same critical value

for which an infinite component appears is not at all obvious, and was one of the great developments of the field. When p is barely less than p_c, $\chi(p)$ will be finite but large. The critical number γ is that real number so that

$$\chi(p) = (p_c - p)^{-\gamma + o(1)} \text{ as } p \to p_c^-$$

Analogously, in the Erdős-Rényi model we examine $E[|C(v)|]$ in $G(n, \frac{1-\epsilon}{n})$. In the subcritical region this is well mirrored by $T_{1-\epsilon}$, the total size of a subcritical Poisson branching process. We find $E[T_{1-\epsilon}]$ by looking at each generation. There is one root Eve, who has an expected number $1-\epsilon$ children. They behave similarly and so Eve has an expected number $(1-\epsilon)^2$ grandchildren. This continues, there are an expected number $(1-\epsilon)^i$ nodes in the i-th generation so that

$$E[T_{1-\epsilon}] = \sum_{i=0}^{\infty} (1-\epsilon)^i = \epsilon^{-1}$$

precisely. Therefore $\gamma = 1$.

While $\chi(p)$ is infinite in the supercritical region we can examine the "finite portion" of $L^d(p)$. The finite susceptibility χ^f is given by

$$\chi^f(p) = E[|C(\vec{0})| \text{ conditional on } C(\vec{0}) \text{ being finite}]$$

When p is barely greater than p_c, $\chi^f(p)$ will be finite but large. The critical number γ' is that real number so that

$$\chi^f(p) = (p - p_c)^{-\gamma' + o(1)} \text{ as } p \to p_c^+$$

The Erdős-Rényi analogue is $E[|C(v)|]$ in $G(n, \frac{1+\epsilon}{n})$, conditional on v not being in the giant component. In $G(n, \frac{1+\epsilon}{n})$, $|C(v)|$ has basically distribution $T^{po}_{1+\epsilon}$, with the value $T^{po}_{1+\epsilon} = \infty$ corresponding to being in the giant component. The finite analouge then corresponds to $T^{po}_{1+\epsilon}$, conditional on it being finite. The probability $T^{po}_{1+\epsilon}$ is finite approaches one as $\epsilon \to 0^+$. The Poisson branching processes $T^{po}_{1+\epsilon}$,$T^{po}_{1-\epsilon}$ have nearly the same finite distribution. Conditioning on v not being in the giant component, $|C(v)|$ then behaves like $T^{po}_{1-\epsilon}$. Therefore $\gamma' = 1$.

At the critical value $p = p_c$, all components are finite. The distribution of $|C(\vec{0})|$ will have a heavy tail. The critical number δ is that real number so that at $p = p_c$

$$\Pr[|C(\vec{0})| \geq s] = s^{-1/\delta + o(1)} \text{ as } s \to \infty$$

For the Erdős-Rényi analogue we consider $|C(v)|$ in $G(n, \frac{1}{n})$. One needs be cautious about the double limit. For any *fixed* s,

$$\lim_{n \to \infty} \Pr[|C(v)| \geq s] = \Pr[T_1^{po} \geq s] = \Theta(s^{-1/2})$$

from (6). Therefore $\delta = 2$.

We further examine the gap exponent, denoted by Δ. In the subcritical region the distribution of $|C(\vec{0})|$ drops off exponentially. For each $k \geq 1$ it has a finite k-th moment. The hypothetical quantity Δ is such that

$$\frac{E[|C(\vec{0})|^{k+1}]}{E[|C(\vec{0})|^k]} = (p_c - p)^{-\Delta + o(1)}$$

The belief is that Δ does not depend on the choice of k. In the supercritical region the belief is that the same asymptotics hold when the infinite component is erased. More precisely, that

$$\frac{E[|C(\vec{0})|^{k+1} \text{ given } C(\vec{0}) \text{ is finite}]}{E[|C(\vec{0})|^{k}] \text{ given } C(\vec{0}) \text{ is finite}]} = (p_c - p)^{-\Delta + o(1)}$$

for all $k \geq 1$. In the Erdős-Rényi analogue the distribution of $C(v)$ in $G(n, \frac{1-\epsilon}{n})$ mirrors that of $T^{po}_{1-\epsilon}$. (The supercritical $G(n, \frac{1+\epsilon}{n})$, with its giant component erased, behaves similarly.) From §1.4, $\Pr[T^{po}_{1-\epsilon} = s]$ drops like $s^{-3/2}$ until k reaches $\Theta(\epsilon^{-2})$ when it begins its exponential drop off. The region of exponential drop off has negligible effect on the finite moments. The k-th moment of $T^{po}_{1-\epsilon}$ is basically the sum of $s^{-3/2}s^k$ for $s = O(\epsilon^{-2})$, which is of order $(\epsilon^{-2})^{k+\frac{1}{2}}$, or ϵ^{-2k-1}. The ratio of the $k+1$-st and k-th moments is then $\Theta(\epsilon^{-2})$. Therefore $\Delta = 2$.

Hara and Slade [**?**] have shown that the critical exponent values $\beta = 1$, $\gamma = \gamma' = 2$, $\delta = 2$, $\Delta = 2$ hold in the bond percolation model for sufficiently high dimensions d. (Their argument works for $d \geq 19$ and the statements are believed to hold for all $d \geq 6$.) Mathematical physicists have a term *mean field* which, quoting Grimmett, "permits several interpretations depending on context." A commonly held requirement is that the critical exponents have the values given above. Thus bond percolation for $d \geq 19$ is regarded as exhibiting mean field behavior. Using the analogues described above it seems reasonable to say that the Erdős-Rényi model exhibits mean field behavior.