ON PHASE TRANSITIONS OF SPECIAL GRAPH PROPERTIES IN RANDOM GRAPHS

DIETER YANG

Abstract

Random Graphs may sound like they has special properties which show up randomly, but it's actually more complicated than that. Like in chemistry, these properties show up in phase transitions. In this paper, we will explore with what probability, p, would these properties appear. Such properties will include connectivity, perfect matching, and Hamilton cycles.

1. Introduction

Random Graphs were first defined by Paul Erdős and Alfréd Rényi in their 1958 paper *On Random Graphs*. However, their paper focused primarily on $\mathbb{G}_{n,m}$, which is based on the number of edges m in such a graph. It was Edgar Gilbert who developed one of the more commonly studied random graph model $\mathbb{G}_{n,p}$. This model is of great use in this paper as it heavily simplifies the computation that would have resulted from the former model. This field was later expanded in the 80's by Béla Bollobás. In this paper, most of the theorems and lemmas arise from these mathematicians which are considered as the "fathers" of the field. These Random Graphs are extremely applicable to a wide variety of topics from Ramsey Theory to Networking and can be used to answer questions about properties of typical graphs. We first provide some background information by including preliminaries and introducing notation that will be used in the paper. In terms of Random Graphs, we will begin by focusing on first on the Rado Graph, which is a great introduction to the field, before switching focus to graph properties. Graph Properties believe it or not show up spontaneously, which we will call phase transitions. We first begin by setting a foundation on our Random Graph models, specifically establish a connection between $\mathbb{G}_{n,m}$ and $\mathbb{G}_{n,p}$. Then, we will shift our focus into specific graph properties such as connectivity, perfect matching, and Hamilton cycles. In these graph properties, we will determine the thresholds in which these properties will occur.

2. Acknowledgements

This paper was written under the 2023 Euler Circle Independent Research and Paper Writing program. The author would like to thank Kishan Jani for mentorship and Hanna Yang for some helpful advice.

Date: July 15, 2023.

3. Preliminaries

Here is the notation and theorems needed for the rest of the paper. First, we explain some graph notation. Let G be a graph. Then we have $G = (V, E)$, where V is the vertex set, E is the edge set of G with $e(G) = |E(G)|$, meaning the number of edges. $S \subseteq E$, then $e_G(S) = |e \in E : e \subseteq S|$ or informally the number of edges in the subset S. We let $N(S) = N_G(S)$ denote the set of neighboring vertices of S or formally for some w we have: $N(S) = \{w \notin S : \exists v \in S : \{v, w\} \in E\}$, with $d_G(S)$ equal to the size of $N(S)$.
When we talk about Random Graphs, there are various models we will use. We let $[n] = \{1, 2, ..., n\}$. $\mathscr{G}_{n,m}$ is the family of all graphs with $V = [n]$ and containing m edges. Similarly, we have $\mathbb{G}_{n,m}$ simply be a random graph chosen from the family $\mathscr{G}_{n,m}$. Another model is $\mathbb{G}_{n,p}$, which also has n vertices, $V = [n]$, but each edge is picked with probability p.
In the event we have a bipartite graph, we have $\mathbb{G}_{n,n,p}$, denote two disjoint vertex sets $[n]$ with probability p that an edge connects two vertices of different sets.
We will also use asymptotic notation to state our results so it is a good idea to familiarize yourself with it. In addition, if something occurs with high probability ($\lim_{n\to\infty} \mathbb{P}(\mathscr{A}_\ltimes) = 1$), we say w.h.p.. More notation will be defined as we get deeper into the topics.

3.1. Inequalities.

Now, we will talk about two inequalities, which will be used frequently in bounding. The first inequality is Markov's Inequality:

Theorem 3.1. *(Markov's Inequality)* $\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}(X)}{t}$, *for some random variable X.*

Proof. A proof of this simply takes some indicator I_A which will equal 1 if some event occurs, or 0 otherwise. We know $X = XI_{\{X\geq t\}} + XI_{\{X<T\}} \geq XI_{\{X\geq t\}} \geq tI_{\{X\geq t\}}$ as t is less than X. As expectation is linear, we have $\mathbb{E}X \geq t\mathbb{E}I_{\{X\geq t\}} = t\mathbb{P}(x \geq t)$ which simplifies to our desired result. ∎

A useful corollary of this occurs when we look at an integer valued random variable that is non negative. If we simply let $t = 1$, note that we get $\mathbb{P}(X \geq 1) \leq \mathbb{E}(X)$ or $\mathbb{P}(X > 0) \leq \mathbb{E}(X)$. This is also known as the First Moment method and will be used frequently throughout the paper. Another theorem that follows from Markov is Chebyshev's Inequality, which states

Theorem 3.2. *(Chebyshev's Inequality)* $\mathbb{P}(|X - \mathbb{E}X| \geq t) \leq \frac{\operatorname{Var} X}{t^2}$

Note that Var is simply the variance of some random variable.$\operatorname{Var} X$ can also be written as $\mathbb{E}X^2 - (\mathbb{E}X)^2$ or $\mathbb{E}(X - \mu)^2$

Proof. A proof of this follows due to Markov's Inequality. Let some random variable $Y = (X - \mu)^2$ and $t = a^2$. Then by Markov, we see that $\mathbb{P}(Y \geq a^2) \leq \frac{\mathbb{E}Y}{a^2}$. By the definition of variance, we have that $\mathbb{E}Y = \mathbb{E}(X - \mu)^2 = \operatorname{Var} X$. In addition, notice how $(X - \mu)^2 \geq a^2$ is the same as $|x - \mu| \geq a$. Thus we get that $\mathbb{P}(|x - \mu| \geq a) \leq \frac{\mathbb{E}Y}{a^2} = \frac{\operatorname{Var} X}{a^2}$. ∎

Again, observe that if $t = \mathbb{E}X$, where X is some non-negative integer valued random variable, then we get a neat corollary. $\mathbb{P}(X = 0) \leq \frac{\operatorname{Var} X}{(\mathbb{E}X)^2} = \frac{\mathbb{E}X^2}{(\mathbb{E}X)^2} - 1$. This is known as the Second moment method.
A stronger version of this inequality actually results from Cauchy-Schwarz.

Proof. We again let X be some non-negative integer valued random variable and then set $X = X \cdot I_{\{X \geq 1\}}$. Then we have as a result

$$(\mathbb{E}X)^2 = (\mathbb{E}(X \cdot I_{\{X \geq 1\}})^2 \leq \mathbb{E}I^2_{\{X \geq 1\}}\mathbb{E}X^2 = \mathbb{P}(X \geq 1)\mathbb{E}X^2.$$

This results in $\mathbb{P}(X = 0) \leq \frac{\operatorname{Var} X}{\mathbb{E}X^2} = 1 - \frac{(\mathbb{E}X)^2}{\mathbb{E}X^2}$ ■

This bound is stronger as $\mathbb{E}X^2 \geq (\mathbb{E}X)^2$ but both bounds are both very useful and can be applied when $\frac{\operatorname{Var} X}{(\mathbb{E}X)^2} \to 0$, or $\frac{\mathbb{E}X^2}{(\mathbb{E}X)^2} \to 1$ as $n \to \infty$.These inequalities will help us discover bounds for all kinds of graph properties.

4. The Rado Graph

What if you and your friend each take a infinitely large piece of paper containing a countably infinite number of vertices on it. Now with probability $0 \leq p \leq 1$, both of you independently draw a line on your own paper, connecting a pair of vertices and repeating this for each pair of vertices. What is the probability you and your friend draw isomorphic graphs? Believe it or not, the probability is actually 1.
To show this, we will require a lemma.

Lemma 4.1 (Extension Property)**.** *For every two disjoint finite sets of vertices, U and V, there exist a vertex x outside of U and V that is connected to all vertices in U but contains no neighbors of V.*

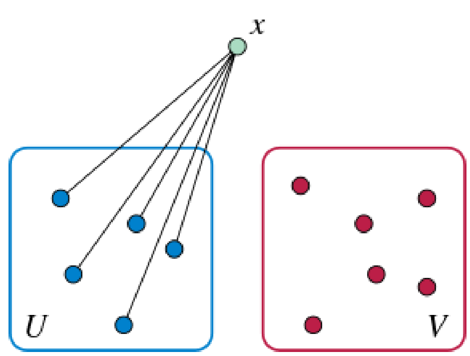


Figure 1. Depiction of Extension Property [DE123]

Proof. Let there be x_1 vertices in U and x_2 vertices in V. Then the probability some vertex x exists is $p^{x_1}(1-p)^{x_2}$. Due to there being an infinite number of vertices, the probability that no x exists is $[1 - p^{x_1}(1-p)^{x_2}]^{\infty}$ which is 0, so some x exists. ■

We will now establish a bijection between you and your friend's graph.
Take some graph G with this property. Label the vertices $1, 2, 3, ...$ for both the Rado graph and our graph G. We proceed with an algorithm alternating between G and the Rado Graph. First, take the smallest unmatched vertex in the Rado Graph (starting with 1).[Note that when we say match, we are matching some vertex from the Rado graph with a vertex from G.Then find it's copy in G]. Then find the smallest unmatched vertex in G and find it's

copy in the Rado Graph. You know you can find it's copy thanks to 4.1. Repeat this. Every vertex will eventually be connected and we find a bijection that preserves edges which shows that G and the Rado Graph are isomorphic. This means that there is only one Random Graph or that $\mathbb{G}_{\infty,p}$ is always the same graph.

5. Relationships, Models, and Thresholds

In this section we will begin to introduce the asymptotic properties of Random Graphs. First, we start by claiming a Random Graph $\mathbb{G}_{n,p}$ given that it's number of edges is m is the same as randomly picking a graph with n vertices and m edges. Formally, this is

Lemma 5.1. *A random graph $\mathbb{G}_{n,p}$ given that it has m edges is equally likely to be one of the $\mathbb{G}_{n,m}$ graphs.*

Proof. We know that there are a total of $\binom{\binom{n}{2}}{m}$ graphs containing n vertices and m edges. We also know that $\mathbb{P}(\mathbb{G}_{n,p} = G_0 | e(\mathbb{G}) = m) = \frac{p^m(1-p)^{\binom{n}{2}-m}}{\binom{\binom{n}{2}}{m} p^m(1-p)^{\binom{n}{2}-m}} = \binom{\binom{n}{2}}{m}^{-1}$. This means that each graph will show up with equal probabilities. ∎

In addition, we gain some intuition that these models are actually similar when m is equal to the expected number of edges in $\mathbb{G}_{n,p}$ or when $m = \binom{n}{2}p$ especially as n grows large. One very useful technique which will be used for many of the proofs below is known as the "coupling technique" which generates a random graph $\mathbb{G}_{n,p}$ in two independent steps.
Suppose we have $p_1 < p$. We define p_2 by the equation

$$p = p_1 + p_2 - p_1 p_2.$$

Observe that this equation is actually a representation of $\mathbb{G}_{n,p} = \mathbb{G}_{n,p_1} \cup \mathbb{G}_{n,p_2}$ with the graphs $\mathbb{G}_{n,p_1}$ and $\mathbb{G}_{n,p_2}$ being independent. This means that we can superimpose $\mathbb{G}_{n,p_1}$ and $\mathbb{G}_{n,p_2}$ to create $\mathbb{G}_{n,p}$ by replacing double edges with just one edge. A similar idea can be applied to $\mathbb{G}_{n,m}$.
Now lets shift our focus into special graph properties. Let $\mathscr{P}$ represent some special graph property, which we formally define as the set of all subsets of graphs on our vertex set $[n]$. Some examples of special graph properties that we will later look at include connectivity and perfect matchings.

Lemma 5.2. *Let $\mathscr{P}$ be any graph property. Then when $p = \frac{m}{\binom{n}{2}}$ where $m = m(n) \to \infty$, and $\binom{n}{2} - m \to \infty$. Then for large n,*

$$\mathbb{P}(\mathbb{G}_{n,m} \in \mathscr{P}) \leq 10 m^{1/2} \mathbb{P}(\mathbb{G}_{n,p} \in \mathscr{P})$$

Proof. First, we proceed with the law of total probability.

$$\mathbb{P}(\mathbb{G}_{n,p} \in \mathscr{P}) = \sum_{k=0}^{\binom{n}{2}} \mathbb{P}(\mathbb{G}_{n,p} \in \mathscr{P} | |E_{n,p}| = k) \mathbb{P}(|E_{n,p}| = k)$$

Then, it follows that

$$\mathbb{P}(\mathbb{G}_{n,p} \in \mathscr{P} | |E_{n,p}| = k) = \frac{\mathbb{P}(\mathbb{G}_{n,p} \in \mathscr{P} \ \& \ |E_{n,p}| = k)}{\mathbb{P}(|E_{n,p}| = k)}$$

Then, calculating probabilities, we see that our numerator and denominator actually differ by a factor of $\binom{\binom{n}{2}}{k}$. The summation of this factor actually gives us $\mathbb{P}(\mathbb{G}_{n,k} \in \mathscr{P})$.

$$= \mathbb{P}(\mathbb{G}_{n,p} \in \mathscr{P}) = \sum_{k=0}^{\binom{n}{2}} \mathbb{P}(\mathbb{G}_{n,k} \in \mathscr{P}) \mathbb{P}(|E_{n,p}| = k)$$

Removing our summation and replacing k with m, we get a lower bound of

$$\leq \mathbb{P}(\mathbb{G}_{n,m} \in \mathscr{P})\mathbb{P}(|E_{n,p}| = m).$$

Note that the number of edges for $|E_{n,p}$ for a random graph follows a binomial distribution as you are choosing 2 vertices each time. This means that if we apply this into Stirling's Formula, we get the following result:

$$k! = (1 + o(1))(\frac{k}{e})^k \sqrt{2\pi k}.$$

Now, let us substitute this to find a rough order on our factorials. We get the following result:

$$\mathbb{P}(|E_{n,p}| = m) = \binom{\binom{n}{2}}{m} p^m (1-p)^{\binom{n}{2}-m}$$

Which will become

$$(1 + o(1)) \frac{\binom{n}{2}^{\binom{n}{2}} \sqrt{2\pi \binom{n}{2}} p^m (1-p)^{\binom{n}{2}-m}}{m^m (\binom{n}{2} - m)^{\binom{n}{2}-m} 2\pi \sqrt{(m(\binom{n}{2} - m)}}$$

Simplifying the numerator and denominators, we get

$$= (1 + o(1)) \sqrt{\frac{\binom{n}{2}}{2\pi m(\binom{n}{2} - m})}$$

Now just doing some rough approximations, we get that

$$\mathbb{P}(|E_{n,p}| = m) \leq \frac{1}{10\sqrt{m}}$$

Now let's substitute this back into our original equation which gives us

$$\mathbb{P}(\mathbb{G}_{n,p} \in \mathscr{P}) \leq 10 m^{1/2} \mathbb{P}(\mathbb{G}_{n,m} \in \mathscr{P}).$$

■

Definition 1. *(Monotone Increasing) A graph property $\mathscr{P}$ as monotone increasing if $G \in \mathscr{P}$, then no matter what edge is added, our new graph will still contain this graph property.*

Examples include whether or not there is a triangle in our random graph. This is as if there already is a triangle, we can add more edges without affecting our triangle. On the contrary,

Definition 2. *(Monotone Decreasing) A graph property to be monotone decreasing if $G \in \mathscr{P}$ and if we remove any edge, our new graph still contains this property.*

One such example would be if a graph was not connected. Removing edges would not make a unconnected graph into a connected one.
Now if we use our coupling argument, we observe that $\mathbb{G}_{n,p} \leq \mathbb{G}_{n,p_1}$ and $\mathbb{G}_{n,m} \leq \mathbb{G}_{n,m_1}$ for $p < p_1$ and $m < m_1$. Using this, we can strengthen our bound from Lemma 5.2 for monotone increasing graph properties.

Lemma 5.3. *Let $\mathscr{P}$ be any monotone increasing graph property with $p = \frac{m}{N}$ Then for large n and $p = o(1)$ such that Np,$N(1-p)/(Np)^{1/2} \to \infty$,*

$$\mathbb{P}(\mathbb{G}_{n,m} \in \mathscr{P}) \leq 3\mathbb{P}(\mathbb{G}_{n,p} \in \mathscr{P})$$

Proof. Let $p = \frac{m}{N}$, where $N = \binom{n}{2}$. Then we have the following result from 5.2:

$$\mathbb{P}(\mathbb{G}_{n,p} \in \mathscr{P}) = \sum_{k=0}^{N} \mathbb{P}(\mathbb{G}_{n,p} \in \mathscr{P}||E_{n,p}| = k)\mathbb{P}(|E_{n,p}| = k)$$

Then, to allow us to apply the coupling property, we replace our starting term to m.

$$\leq \sum_{k=m}^{N} \mathbb{P}(\mathbb{G}_{n,p} \in \mathscr{P}||E_{n,p}| = k)\mathbb{P}(|E_{n,p}| = k)$$

As we know by the coupling property, when $k \geq m$, the following result is true:

$$\mathbb{P}(\mathbb{G}_{n,k} \in \mathscr{P}) \geq \mathbb{P}(\mathbb{G}_{n,m} \in \mathscr{P})$$

From here, due to the binomial nature of the number of edges, we get the following result:

$$\mathbb{P}(\mathbb{G}_{n,p} \in \mathscr{P}) \geq \mathbb{P}(\mathbb{G}_{n,m} \in \mathscr{P}) \sum_{k=m}^{N} \mathbb{P}(|E_{n,p}| = k)$$

Let $u_k = \binom{N}{k} p^k (1-p)^{N-k}$, which is the probability our random graph has k edges. Thus, substituting this probability in, we get

$$\mathbb{P}(\mathbb{G}_{n,m} \in \mathscr{P}) \sum_{k=m}^{N} \mathbb{P}(|E_{n,p}| = k) = \mathbb{P}(\mathbb{G}_{n,m} \in \mathscr{P}) \sum_{k=m}^{N} u_k).$$

Now, we again apply Stirling's formula which for u and simplify giving the following result:

$$u_m = (1+o(1))\frac{N^N p^m (1-p)^{N-m}}{m^m (N-m)^{N-m} (2\pi m)^{1/2}} = \frac{1+o(1)}{(2\pi m)^{1/2}}.$$

Now, an interesting result actually occurs when $k = m + t$, for some $0 \leq t \leq m^{1/2}$.

$$\frac{u_{k+1}}{u_k} = \frac{(N-k)p}{(k+1)(1-p)} = \frac{1 - \frac{t}{N-M}}{1 + \frac{t}{N-m}} \leq exp\{\frac{-t}{N-m-t} - \frac{t+1}{m}\}.$$

Notice that the inequality comes from the fact $1 + x \leq e^x$ and that $1 - x \geq e^{-x/(1-x)}$. Now we can observe that

$$u_{m+t} \geq \frac{1+o(1)}{(2\pi m)^{1/2}} exp\{-\sum_{s=0}^{t-1}(\frac{s}{N-m-s} - \frac{s+1}{m})\} \geq \frac{exp\{-\frac{t^2}{2m} - o(1)\}}{(2\pi m)^{1/2}}$$

as we know that $m = o(N)$. Now, as a direct consequence, we can now convert this discovery into a summation as follows:

$$\sum_{k=m}^{m+m^{1/2}} u_k \geq \frac{1-o(1)}{(2\pi)^{1/2}} \int_0^1 e^{-x^2/2}\, dx \geq 1/3$$

Which we can compute to find a lower bound of 1/3. Now we can simply substitute this back into 5.18, and we are done. ■

These lemmas are crucial to the simplicity of this paper as it is much easier to compute probabilities in $\mathbb{P}(\mathbb{G}_{n,p})$ in comparison to that of $\mathbb{P}(\mathbb{G}_{n,m})$. This allows us to make the observation that if $\mathbb{P}(\mathbb{G}_{n,p}) \to 0$, then $\mathbb{P}(\mathbb{G}_{n,m}) \to 0$ for $n \to \infty$. This brings us to a theorem which is proved by Łuczak which notes the specific conditions needed to be fulfilled for the asymptotic equivalence of random graphs for $\mathbb{G}_{n,p}$ and $\mathbb{G}_{n,m}$.

Theorem 5.4. *Let* $0 \leq p_0 \leq 1$, $s(n) = n\sqrt{p(1-p)} \to \infty$ *and* $\omega(n) \to \infty$ *arbitrarily slowly as* $n \to \infty$

1. *Suppose that* $\mathscr{P}$ *is a graph property such that* $\mathbb{P}(\mathbb{G}_{n,m} \in \mathscr{P}) \to p_0$ *for all*
$$m \in [\binom{n}{2}p - \omega(n)s(n), \binom{n}{2}p + \omega(n)s(n)]$$
Then $\mathbb{P}(\mathbb{G}_{n,p} \in \mathscr{P}) \to p_0$ *as* $n \to \infty$.
2. *Let* $p_- = p - \omega(n)s(n)/n^2$ *and* $p_+ = p + \omega(n)s(n)/n^2$. *Suppose that* $\mathscr{P}$ *is a monotone graph property. Then if* $\mathbb{P}(\mathbb{G}_{n,p_-} \in \mathscr{P}) \to p_0$ *and* $\mathbb{P}(\mathbb{G}_{n,p_+} \in \mathscr{P}) \to p_0$ *we have*
$$\mathbb{P}(\mathbb{G}_{n,m} \in \mathscr{P}) \to \infty$$
as $n \to \infty$ *and* $m = \lfloor \binom{n}{2}p \rfloor$

A proof of this theorem will not be provided, though if you are interested, it can be found here [BB98]. Now, equipped with these theorems and lemmas, lets proceed onto observing patterns in which our properties will appear.

5.1. **Thresholds.** Random Graphs have a very interesting nature of when properties will appear and disappear. They actually show up some what spontaneously and vanish in the same regards. We will now formally define this property as a threshold.

Definition 3. *A function* $m^* = m^*(n)$ *is a threshold for a monotone increasing graph property* $\mathscr{P}$ *in a random graph* $\mathbb{G}_{n,m}$ *if*
$$\lim_{x\to\infty} \mathbb{P}(\mathbb{G}_{n,m} \in \mathscr{P}) = \begin{cases} 0 & \text{if } m/m^* \to 0 \\ 1 & \text{if } m/m^* \to \infty \end{cases}$$
as $n \to \infty$

We can also define this idea for p in $\mathbb{G}_{n,p}$.

Definition 4. *A function* $p^* = p^*(n)$ *is a threshold for a monotone increasing graph property* $\mathscr{P}$ *in a random graph* $\mathbb{G}_{n,p}$ *if*
$$\lim_{x\to\infty} \mathbb{P}(\mathbb{G}_{n,p} \in \mathscr{P}) = \begin{cases} 0 & \text{if } p/p^* \to 0 \\ 1 & \text{if } p/p^* \to \infty \end{cases}$$
as $n \to \infty$

We can also have a more specific definition for a sharp threshold.

Definition 5. *A function* $p^* = p^*(n)$ *is a sharp threshold for a monotone increasing graph property* $\mathscr{P}$ *in a random graph* $\mathbb{G}_{n,p}$ *if for every* $\epsilon > 0$
$$\lim_{x\to\infty} \mathbb{P}(\mathbb{G}_{n,p} \in \mathscr{P}) = \begin{cases} 0 & \text{if } p/p^* \leq 1 - \epsilon \\ 1 & \text{if } p/p^* \geq 1 + \epsilon \end{cases}$$
as $n \to \infty$

And again, we have a similar definition for m in $\mathbb{G}_{n,m}$

Definition 6. *A function $m^* = m^*(n)$ is a sharp threshold for a monotone increasing graph property $\mathscr{P}$ in a random graph $\mathbb{G}_{n,m}$ if for every $\epsilon > 0$*

$$\lim_{x\to\infty} \mathbb{P}(\mathbb{G}_{n,m} \in \mathscr{P}) = \begin{cases} 0 & \textit{if } p/p^* \leq 1-\epsilon \\ 1 & \textit{if } p/p^* \geq 1+\epsilon \end{cases}$$

as $n \to \infty$

Notice that saying p^* is a threshold for a property is identical to saying that a graph does not contain the said property w.h.p. if $p \ll p^*$ and will contain the property w.h.p. if $p \gg p^*$ Now let's move onto one of the most important theorems of the field proved by Bollobàs and Thomason.

Theorem 5.5. *Every non-trivial monotone graph property has a threshold*

Proof. Without loss of generality, we assume that $\mathscr{P}$ is a monotone increasing graph property. Let $0 < \epsilon < 1$. We will define $p(\epsilon)$ as follows:

$$\mathbb{P}(\mathbb{G}_{n,p(\epsilon)} \in \mathscr{P}) = \epsilon$$

But why does $p(\epsilon)$ exist?

$$\mathbb{P}(\mathbb{G}_{n,p} \in \mathscr{P}) = \sum_{G\in\mathscr{P}} p^{|E(G)|}(1-p)^{N-|E(G)|}$$

is a polynomial that increases from 0 to 1 due to our property being monotone increasing. Therefore the higher the probability i.e. the edges, the greater chance of our property occurring. Now let us show that $p^* = p(1/2)$ is a threshold for $\mathscr{P}$ by the coupling argument. Let us create independent copies of $\mathbb{G}_{n,p}$ and number them from 1 to k. Then the graph of $G_1 \cup G_2 \cup ... \cup G_k$ is distributed as the random graph $\mathbb{G}_{n,1-(1-p)^k}$. Notice that $1-(1-p)^k \leq kp$ so therefore coupling results in

$$\mathbb{G}_{n,1-(1-p)^k} \subseteq \mathbb{G}_{n,kp}.$$

Note that the contrary $\mathbb{G}_{n,kp} \notin \mathscr{P}$ implies that non of our independent copies contain the graph property either as it is monotone. Thus we get the following result:

$$\mathbb{P}(\mathbb{G}_{n,kp} \notin \mathscr{P}) \leq [\mathbb{P}(\mathbb{G}_{n,p} \in \mathscr{P})]^k.$$

Now, let ω be a function of n, where $\omega \to \infty$ arbitrarily slowly as $n \to \infty$, with $\omega \ll \log\log n$. We also assume that $p = p^* = p(1/2)$ and $k = \omega$. Then we see

$$\mathbb{P}(\mathbb{G}_{n,\omega p^*} \notin \mathscr{P}) \leq 2^{-\omega} = o(1)$$

Meanwhile, if $p = p^*/\omega$, we observe that

$$\frac{1}{2} = \mathbb{P}(\mathbb{G}_{n,p^*} \notin \mathscr{P}) \leq [\mathbb{P}(\mathbb{G}_{n,p^*/\omega} \notin \mathscr{P})]^\omega$$

Then by simplification, we see that

$$[\mathbb{P}(\mathbb{G}_{n,p^*/\omega} \notin \mathscr{P})] \geq 2^{-1/\omega} = 1 - o(1)$$

■

Given these lemmas and theorems, we are now ready to tackle some basic graph properties. Arguably one of the simplest graph properties is if a graph contains an edge or not. Seems simple enough right? We want to find out at what point is a random graph gonna have an edge w.h.p..

Theorem 5.6. *Let $\mathscr{P} = \{$non edge-less sets of labeled graphs $\mathbb{G}_{n,p}\}$ Then*

$$\lim_{x\to\infty} \mathbb{P}(\mathbb{G}_{n,p} \in \mathscr{P}) = \begin{cases} 0 & \text{if } p \ll n^{-2} \\ 1 & \text{if } p \gg n^{-2} \end{cases}$$

Proof. Let X be a random variable which counts the number of edges in $\mathbb{G}_{n,p}$. We note that $\mathbb{E}X = \binom{n}{2}p$ by linearity of expectation. Similarly, the variance $\operatorname{Var} X = \binom{n}{2}p(1-p) = (1-p)\mathbb{E}X$ as it follows a binomial distribution. Proving the first part of our statement is pretty easy with Markov's inequality via the First Moment Method. In our case, we see that

$$\mathbb{P}(X > 0) \leq \frac{n^2}{2}p \to 0$$

when $p \ll n^{-2}$ for $n \to \infty$. Now, if we want to show the 2nd part of our statement, we now require the use of Chebyshev and the Second Moment Method. We will use this inequality to show that $X \approx \mathbb{E}X$ w.h.p.. The reason behind this is as the probability will approach 1 when $\operatorname{Var} X/(\mathbb{E}X)^2 \to 0$. Thus let us observe what happens to $\operatorname{Var} X/(\mathbb{E}X)^2 \to 0$. Substituting gives

$$\operatorname{Var} X/(\mathbb{E}X)^2 = \frac{1-p}{\mathbb{E}X} \to 0$$

when $n \to \infty$ so this shows that our second statement is true and that $\frac{1}{n^2}$ is indeed our threshold for $\mathbb{G}_{n,m}$ ■

Now that we got done with this example, let's try and do this but with a triangle instead.

Theorem 5.7. *If $m/n \to \infty$ then w.h.p. $\mathbb{G}_{n,m}$ contains at least one triangle.*

Proof. Let us begin with noting that a triangle is clearly a monotone increasing property. Thus, let us prove the result first for $\mathbb{G}_{n,p}$ and then use one of our past lemmas to convert this to $\mathbb{G}_{n,m}$. Here, we will assume that $np \to \infty$.
Let $np = \omega \leq \log n$, where $\omega = \omega(n) \to \infty$. Let our random variable Z count the number of triangles in our random graph. Thus we have the following result:

$$\mathbb{E}Z = \binom{n}{3}p^3 \geq (1 - o(1))\frac{\omega^3}{6} \to \infty$$

However, just because the absolute value approaches infinity, it is not sufficient.
Now let us define $T_1, T_2, ..., T_M$ for $M = \binom{n}{3}$ represent the possible triangles of our random graph. We want to apply Chebyshev, so let's try and calculate $\mathbb{E}Z^2$ to calculate our variance.

$$\mathbb{E}Z^2 = \sum_{i,j=1}^{M} \mathbb{P}(T_i, T_j \in \mathbb{G}_{n,p})$$

Now, let us break apart this summation into two separate summations.

$$\mathbb{E}Z^2 = \sum_{i=1}^{M} \mathbb{P}(T_i \in \mathbb{G}_{n,p}) \sum_{j=1}^{M} \mathbb{P}(T_j \in \mathbb{G}_{n,p} | T_i \in \mathbb{G}_{n,p})$$

Now notice what the first summation is.

$$\sum_{i=1}^{M} \mathbb{P}(T_i \in \mathbb{G}_{n,p}) = M\mathbb{P}(T_1 \in \mathbb{G}_{n,p})$$

This is as the probability is always the same as it is symmetric. And now notice that this is just $\mathbb{E}Z$. Now lets observe some sort of bound for the right summation. We define σ_j as edges T_j and T_1 share.

$$\sum_{j=1}^{M} \mathbb{P}(T_j \in \mathbb{G}_{n,p}|T_1 \in \mathbb{G}_{n,p}) = 1+\sum_{j:\sigma_j=1} \mathbb{P}(\mathbb{P}(T_j \in \mathbb{G}_{n,p}|T_1 \in \mathbb{G}_{n,p}))+\sum_{j:\sigma_j=0} \mathbb{P}(\mathbb{P}(T_j \in \mathbb{G}_{n,p}|T_1 \in \mathbb{G}_{n,p}))$$

Now you may be wondering where all of these summations came from. We break it down into 3 cases, if the triangles share 1,2, or no edges. When they share 2, then the triangle is fixed already so we have probability 1. Now lets calculate the other probabilities and then bound it.

$$= 1 + 3(n-3)p^2 + (\binom{n}{3} - 3n + 8)p^3 \leq 1 + \frac{3^\omega}{n} + \mathbb{E}Z.$$

Now, we use our definition of variance prior to get

$$\operatorname{Var} Z \leq (\mathbb{E}Z)(1 + \frac{3^\omega}{n} + \mathbb{E}Z) - (\mathbb{E}Z)^2 \leq 2\mathbb{E}Z$$

We are now ready to apply Chebyshev inequality.

$$\mathbb{P}(Z = 0) \leq \mathbb{P}(|Z - \mathbb{E}Z| \geq \mathbb{E}Z) \leq \frac{\operatorname{Var} Z}{(\mathbb{E}Z)^2} \leq \frac{2}{\mathbb{E}Z} = o(1).$$

This means that the distribution is essentially zero and that we have proved our theorem. ■

This concludes our basic intro in random graphs. We will now shift our focus into more discrete topics.

6. Phase Transitions

In this section we will focus on the structure of our graphs as $p \to 1$ or $m \to \binom{n}{2}$. One thing to note is that computation is much easier for $\mathbb{G}_{n,p}$ than it is for $\mathbb{G}_{n,m}$. This is why Lemmas 5.1,5.2,5.3 were established prior. This allows us to prove properties for $\mathbb{G}_{n,p}$ and converting them to $\mathbb{G}_{n,m}$, which will greatly simplify our computations.

6.1. **Sub-Critical Phase.**

We begin our journey with Sub-Critical Phase, or the first phase of the Random Graph. At this point, we have a relatively low p and m, and for this reason, our random graphs consist mostly of trees and other small components. Lets begin with some simple properties to get a feel of this first phase.

Theorem 6.1. *If $m \ll n$, then $\mathbb{G}_m$ is a forest w.h.p..*

Proof. We let $m = n/\omega$ and $N = \binom{n}{2}$. Then $p = m/N \leq 3/(\omega n)$. Now we define our random variable X as the number of cycles in $\mathbb{G}_{n,p}$. Then we have

$$\mathbb{E}X = \sum_{k=3}^{n} \binom{n}{k} \frac{(k-1)!}{2} p^k \leq \sum_{k=3}^{n} \frac{n^k}{k!} \frac{(k-1)!}{2} p^k \leq \sum_{k=3}^{n} \frac{n^k}{2k} \frac{3^k}{(\omega n)^k} p^k = O(\omega^{-3}) \to 0$$

We arrive at this result through rough approximations that allow us to cancel the numerators and denominators nicely. As a result we are left with a result that approaches 0. Now we may apply the First Moment Method, which tells us

$$\mathbb{P}(\mathbb{G}_{n,p} \text{ is not a forest}) = \mathbb{P}(X \geq 1) \leq \mathbb{E}X = o(1),$$

This implies that

$$\mathbb{P}(\mathbb{G}_{n,p} \text{ is a forest}) \to 1 \text{ as } n \to \infty$$

Now we notice that a forest is a monotone decreasing property, so by Lemma 5.3, we see that

$$\mathbb{P}(\mathbb{G}_m \text{is a forest}) \to 1 \text{ as } n \to \infty$$

The lemma here was actually to show that the probabilities they are not forests is 0 as that is the monotone property, but we can see that the compliment would also be true. ■

Now, let us examine the point of which $\mathbb{G}_m$ is consisting of single edges or isolated vertices, i.e. no paths of length 2, w.h.p.. This theorem was first proved by Erdős in a series of mass publications.

Theorem 6.2. *If $m \ll n^{1/2}$, then $\mathbb{G}_m$ is the union of isolated vertices and edges w.h.p.*

Proof. We let $p = m/N$, $m = n^{1/2}/\omega$ and let our random variable X be the number of paths of length two in our random graph $\mathbb{G}_{n,p}$. Now let us use the First Moment Method

$$\mathbb{P}(X > 0) \leq \mathbb{E}X = 3\binom{n}{3}p^2 \leq \frac{n^4}{2N^2\omega^2} \to 0, \text{as } n \to 0$$

Therefore we see that

$$\mathbb{P}(\mathbb{G}_{n,p} \text{ contains a path of length two}) = o(1)$$

And now, since our property is monotone increasing, we can apply 5.3

$$\mathbb{P}(\mathbb{G}_m \text{ contains a path of length two}) = o(1)$$

And we are done. ■

And now it may follow from this observation that the contrary is most likely going to be true as well.

Theorem 6.3. *If $m \gg n^{1/2}$, then $\mathbb{G}_m$ contains a path of at least length 2 w.h.p..*

Proof. Let $p = \frac{m}{n}$, $m = \omega n^{1/2}$ and X be the random variable counting the number of paths of length two in $\mathbb{G}_{n,p}$. Then

$$\mathbb{E}X = 3\binom{n}{3}p^2 \approx 2\omega^2 \to \infty$$

However, again this is not enough to imply that $X > 0$ w.h.p. We must now apply the Second Moment Method.
We define $\mathscr{P}_2$ to be the set of all paths of length 2 in the complete graph, with $\hat{X}$ as the number of isolated paths of length 2 in $\mathbb{G}_{n,p}$. Our objective is to now show w.h.p. $\mathbb{G}_{n,p}$ contains this path.

$$\hat{X} = \sum_{P \in \mathscr{P}_2} I_{P \in_i \mathbb{G}_{n,p}}$$

Here, the notation $\in_i$ means that P is isolated. Having a path of length two is a monotone increasing property which allows us to assume that $m = o(n)$ and therefore $np = o(1)$.

$$\mathbb{E}\hat{X} = 3\binom{n}{3}p^2(1-p)^{3(n-3)+1} \geq (1-o(1))\frac{n^3}{2}\frac{4\omega^2}{n}n^4(1-3np) \to \infty.$$

Now let us try and use the Second Moment Method. We observe that

$$\hat{X}^2 = \sum_{P\in\mathscr{P}_2}\sum_{Q\in\mathscr{P}_2} I_{P\in_i\mathbb{G}_{n,p}} I_{Q\in_i\mathbb{G}_{n,p}} = \sum_{P,Q\in\mathscr{P}}^{*} I_{P\in_i\mathbb{G}_{n,p}} I_{Q\in_i\mathbb{G}_{n,p}}$$

The last summation is taken over $P, Q \in \mathscr{P}_2$ s.t. either $P = Q$ or P and Q are vertex disjoint. Now we are able to use the isolated paths we introduced before to calculate the expectation.

$$\mathbb{E}\hat{X}^2 = \sum_P\{\sum_Q \mathbb{P}(Q \subseteq_i \mathbb{G}_{n,p}|P \subseteq_i \mathbb{G}_{n,p})\}\mathbb{P}(P \subseteq_i \mathbb{G}_{n,p}).$$

Now notice that the expression is actually the same for all P so we can now simplify to get

$$\mathbb{E}\hat{X}^2 = \mathbb{E}\hat{X}\left(1 + \sum_{Q\cap P_{(1,2,3)}=\emptyset} \mathbb{P}(Q \subseteq_i \mathbb{G}_{n,p}|P_{(1,2,3)\subseteq_i\mathbb{G}_{n,p}})\right)$$

where we have $P(1, 2, 3)$ denote the path on the vertex set $1, 2, 3$. Now notice that if we assume that $P_{1,2,3}$ is a component of $\mathbb{G}_{n,p}$, then all nine edges between Q and $P_{1,2,3}$ will be missing. Therefore we can bound as follows:

$$\mathbb{E}\hat{X}^2 \leq \mathbb{E}\hat{X}(1 + 3\binom{n}{3}p^2(1-p)^{3(n-6)+1}) \leq \mathbb{E}\hat{X}(1 + (1-p)^{-9}\mathbb{E}\hat{X}).$$

Now, we can finally apply the Second Moment Method which gives:

$$\mathbb{P}(\hat{X} > 0) \geq \frac{(\mathbb{E}\hat{X})^2}{\mathbb{E}\hat{X}(1 + (1-p)^{-9}\mathbb{E}\hat{X})} = \frac{1}{(1-p)^{-9} + [\mathbb{E}\hat{X}]^{-1}} \to 1$$

as $n \to \infty$ due to $p \to 0$ and $\mathbb{E}\hat{X} \to \infty$. Therefore

$$\mathbb{P}(\mathbb{G}_{n,p} \text{ contains an isolated path of length two}) \to 1$$

Again recall that this is a monotone increasing property so

$$\mathbb{P}(\mathbb{G}_m \text{ contains an isolated path of length two}) \to 1$$

for $m \ll n^{1/2}$ and we are done. ■

6.2. **Super-Critical Phase.** We begin our next section, where we now look at the next phase known as the Super-Critical Phase. The structure of our Random Graphs actually become shockingly different as m is around the same order of $n/2$. Again due to the difficult computation of $\mathbb{G}_m$, we will be using $\mathbb{G}_{n,p}$ instead where $p = c/n$ for some constant $c > 1$ and presenting these results in terms of $\mathbb{G}_m$.

Theorem 6.4. *If $m = cn/2$, for some constant $c > 1$, then w.h.p. $\mathbb{G}_m$ consists of a a unique giant component, with $(1 - \frac{x}{c} + o(1))n$ vertices and $(1 - \frac{x^2}{c^2} + o(1))\frac{cn}{2}$ edges. Here $0 < x < 1$ is the solution of the equation $xe^{-x} = ce^{-c}$. The remaining components of our graph are at most of order $O(\log n)$.*

In order to prove this theorem however, many lemmas will be required. We will provide the lemmas without proof, though if you are interested, click here and check out the 2nd chapter. [FK16]

Lemma 6.5. *If $p = c/n$, where $c \neq 1$ is a constant, then in $\mathbb{G}_{n,p}$ w.h.p. the number of vertices in components with exactly one cycle, is $O(\omega)$ for any growing function ω.*

Lemma 6.6. *Let $p = c/n$, where again $c \neq 1$ is a constant, and $\alpha = c - 1 - logc$, and $\omega = \omega(n) \to \infty$, $\omega = o(log \log n)$. Then*

(1) *w.h.p. there exists an isolated tree of order*

$$k_{-} = \frac{1}{\alpha}(longn - \frac{5}{2}loglongn) - \omega$$

(2) *w.h.p. there is no isolated tree of order at least*

$$k_{+} = \frac{1}{\alpha}(\log n - \frac{5}{2}log \log n) + \omega$$

These two lemmas actually prove an interesting result that

$$x = x(c) = \begin{cases} \text{c} & \text{if } c \leq 1 \\ \text{The solution } (0,1) \text{ to } xe^{-x} = ce^{-c} & \text{if } c > 1 \end{cases}$$

This actually shows the existence and uniqueness of x as xe^{-x} is continuously increasing as x increases from 0 to 1. Lastly, we have

Lemma 6.7. *If $c > 0$, $c \neq 1$ is a constant, and $x = x(c)$ is defined above, then*

$$\frac{1}{x}\sum_{k=1}^{\infty}\frac{k^{k-1}}{k!}(ce^{-c})^k = 1.$$

We are now ready to dive into our theorem.

Proof. Let's begin by defining Z_k as the number of components of order k in the random graph $\mathbb{G}_{n,p}$. Then we can bound the number of components by the number of trees with k vertices to give us the following expected value:

$$\mathbb{E}Z_k \leq \binom{n}{k}k^{k-2}p^{k-1}(1-p)^{k(n-k)} \leq A(\frac{ne}{k})^k k^{k-2}(\frac{c}{n})^{k-1}e^{-ck+ck^2/n} \leq \frac{An}{k^2}(ce^{1-c+ck/n})^k$$

Now let us define $\beta_1 = \beta_1(c)$ as some small enough function such that

$$ce^{1-c+c\beta_1} < 1,$$

and similarly, let $\beta_0 = \beta_0(c)$ be some large enough function such that

$$(ce^{1-c+o(1)})^{\beta_0 logn} < \frac{1}{n^2}.$$

If we choose β_1 and β_0 as above, then it follows that w.h.p. there is no component of order $k \in [\beta_0 \log n, \beta_1 n]$.
Now let us try to estimate the number of vertices on small components. To do this, we first estimate the total number of vertices on small tree components.
We proceed by assuming that $1 \leq k \leq k_0$, where $k_0 = \frac{1}{2\alpha}\log n$, where α is from Lemma

6.6. It is shown that as a consequence of this lemma, and Stirling's Distribution that the following is true

$$\mathbb{E}(\sum_{k=1}^{k_0} kX_k) \approx \frac{n}{c}\sum_{k=1}^{k_0}\frac{k^{k-1}}{k!}(ce^{-c})^k \approx \frac{n}{c}\sum_{k=1}^{\infty}\frac{k^{k-1}}{k!}(ce^{-c})^k$$

Now notice that if we bound $k^{k-1}/k! < e^k$ and $ce^{-c} < e^{-1}$ for $c \neq 1$, it allows us to extend the summation above to infinity.

Now let us substitute $\epsilon = 1/\log n$ into the following expression which arises from Chebyshev and Lemma 6.6.

$$\mathbb{P}(|X_k - \mathbb{E}X_k| \le \epsilon\mathbb{E}X_k) \le \frac{1}{\epsilon^2\mathbb{E}X_k} + \frac{2ck^2}{\epsilon^2 n} = o(1)$$

to get

$$\sum_{k=1}^{k_0}[\frac{(\log n)^2}{n^{1/2=o(1)} + O(\frac{(\log n)^4}{n}}] = o(1)$$

Now, if we have $x = x(c), 0 < x < 1$ be the unique solution of the equation $xe^{-x} = ce^{-c}$, then we observe w.h.p.,

$$\sum_{k=1}^{k_0} kX_k \approx \frac{n}{c}\sum_{k=1}^{\infty}\frac{k^{k-1}}{k!}(xe^{-x})^k = \frac{nx}{c}$$

Which is a direct consequence of Lemma 6.7. Now going back we have

$$\mathbb{E}(\sum_{k=k_0+1}^{\beta_0\log n} kX_k) \le \frac{n}{c}\sum_{k=k_0+1}^{\beta_0\log n}(ce^{1-c+ck/n})^k = O(n(ce^{1-c})^{k_0}) = O(n^{1/2+o(1)}.$$

Now we see through Markov Inequality, w.h.p.

$$\sum_{k=k_0+1}^{\beta_0\log n} kX_k = o(n).$$

Now let us consider the number of non-tree components with k vertices, which we label as Y_k, where $1 \le k \le \beta_0 \log n$.

$$\mathbb{E}(\sum_{k=1}^{\beta_0\log n} kY_k) \le \sum_{k=1}\beta_0\log n\binom{n}{k}k^{k-1}\binom{k}{2}(\frac{c}{n})^k(1-\frac{c}{n})^k n - k \le \sum_{k=1}\beta_0\log nk(ce^{1-c+ck/n})^k = O(1)$$

Now, we apply Markov again to see that w.h.p.

$$\sum_{k=1}^{\beta_0\log n} kY_k = o(n).$$

Now to summarize what we have done so far, we have shown that with high probability, there are approximately $\frac{nx}{c}$ vertices on components of order k, and that the remaining giant components are at least of size $\beta_1 n$.

To finish off our proof, we must now show that this giant component is unique.

Let us define $c_1 = c - \frac{\log n}{n}$ and $p_1 = \frac{c_1}{n}$, and p_2 as $1 - p = (1-p_1)(1-p_2)$. You may see where we are going with this, as a similar idea was brought up prior.

We have that $\mathbb{G}_{n,p} = \mathbb{G}_{n,p_1} \cup \mathbb{G}_{n,p_2}$.

Now if $x_1 e^{-x_1} = c_1 e^{-c_1}$, then $x_1 \approx x$, therefore implying that w.h.p. $\mathbb{G}_{n,p_1}$ has no components

with number of vertices in the range $[\beta_0 \log n, \beta_1 n]$.
Now we suppose there are some components $C_1, ...C_l$, where $|C_i| > \beta_1 n$. Here $l \leq 1/\beta_1$. We will now add the edges now of our two random graphs. We have the following result

$$\mathbb{P}(\exists\ i,j:\ \text{no edge in } \mathbb{G}_{n,p_2} \text{ joins } C_i \text{ with } C_j) \leq \binom{l}{2}(1-p_2)^{(\beta_1 n)^2} \leq l^2 e^{-\beta_1^2 \log n} = o(1).$$

Therefore, we see that $\mathbb{G}_{n,p}$ has a unique component with more that $\beta_0 \log n$ vertices and it has $\approx (1-\frac{x}{c})n$ vertices.
Now let us consider the number of edges in C_0. We will now switch to using $G = \mathbb{G}_{n,m}$. We suppose that the edges of G are $e_1, ...e_m$ in some order. Now we will estimate that the probability $e = e_m = \{x, y\}$ is an edge of the giant. We let G_1 be the graph induced by the following edge set $e_1, ...e_{m-1}$. G_1 is distributed as $\mathbb{G}_{n,m-1}$ and therefore, with high probability, G_1 has a unique giant C_1 and other components of size $O(\log n)$. This means that the probability of the edge of a giant is $o(1)$ plus the probability that x or y is a vertex of this giant C_1 component.
We get the following result:

$$\mathbb{P}(e \notin C_0 | |C_1| \approx n(1-\frac{x}{c})) = \mathbb{P}(e \cap C_1 = \emptyset | |C_1| \approx n(1-\frac{x}{c})) = (1-\frac{|C_1|}{n})(1-\frac{|C_1|+1}{n}) \approx (\frac{x}{c})^2.$$

Now it follows that the expected number of edges in the giant is as claimed. Now to prove the concentration, we can apply Chebyshev. We will fix $i, j \leq m$ and let C_2 denote the unique giant component.

$$\mathbb{P}(e_i, e_j \subseteq C_0) = o(1) + \mathbb{P}(e_j \cap C_2 \neq \emptyset)\mathbb{P}(e_i \cap C_2 \neq \emptyset) = (1+o(1))\mathbb{P}(e_i \subseteq C_0)\mathbb{P}(e_j \subseteq C_0).$$

From here, we can apply Chebyshev again to show that the number of edges is concentrated and we are done. ■

The theorem above and the results from the last section show that when $m = cn/2$ and c passes the critical value equal to 1, and that the typical structure of a random graph changes from a scattered collection of small trees into a giant component, which we call the **phase transition**.

6.3. **Commentary on Phase Transitions.** These past two sections looked into the sub and super critical phases of random graphs. There have been **simple** components in which they contain exactly one cycle, and **complex** components in which they contain more than one cycle. In the sub-critical phase, there were mainly several small, simple components. However, in the super-critical phase, or when $m \geq n/2$, we see a giant complex component in tandem with several simple components. It is due to this drastic change of structure in a random graph that we call this phenomenon a **phase transition**. One could think about this like in chemistry, where water freezes below 0 degrees and boils at 100 degrees Celsius. Aside from these temperatures, there really is no difference between water appearance wise; anything below 0 degrees would just be ice ect... If you would like to read up more about phase transitions for random graphs, please read here [KS13].

7. Connectivity

Now let us dive into one of the most well known graph properties. Connectivity, where you can go from any vertex to another, creating a connected graph is extremely fascinating. Our first theorem was shown by Erdős and Rényi.

Theorem 7.1. *Let $m = \frac{1}{2}n(\log n + c_n)$. Then we have that*

$$\lim_{n\to\infty} \mathbb{P}(\mathbb{G}_m \textit{ is connected}) = \begin{cases} 0 & \textit{if } c_n \to -\infty \\ e^{-e^{-c}} & \textit{if } c_n \to c \textit{ (a constant)} \\ 1 & \textit{if } c_n \to \infty. \end{cases}$$

Proof. We once again consider a random graph $\mathbb{G}_{n,p}$. Our plan is to show that when $p = \frac{\log n + c}{n}$,

$$\mathbb{P}\mathbb{G}_{n,p} \text{ is connected} \to e^{-e^{-c}}.$$

Then we will use Theorem 5.4 to translate this to $\mathbb{G}_{\gtrdot}$ and use the idea of this being a monotone property to show it for $c_n \to \infty$

We define $X_k = X_{k,n}$ as the number of components with k vertices in our random graph. Now we consider the complement of the event that our graph is connected. We see that

$$\mathbb{P}(\mathbb{G}_{n,p} \text{ is not connected}) = \mathbb{P}(\bigcup_{k=1}^{n/2} (\mathbb{G}_{n,p} \text{ has a component of order k}) = \mathbb{P}(\bigcup_{k=1}^{n/2} \{X_k > 0\})$$

We can simplify this by considering the case $k = 1$, which is simply an isolated vertex. Thus we get

$$\mathbb{P}(X_1 > 0) \leq \mathbb{P}(\mathbb{G}_{n,p} \text{ is not connected} \leq \mathbb{P}(X_1 > 0) + \sum_{k=2}^{n/2} \mathbb{P}(X_k > 0.$$

Now observe that our summation can be bounded as follows:

$$\sum_{k=2}^{n/2} \mathbb{P}(X_k > 0 \leq \sum_{k=2}^{n/2} \mathbb{E}X_k \leq \sum_{k=2}^{n/2} \binom{n}{k} k^{k-2} p^{k-1} (1-p)^{k(n-k)} = \sum_{k=2}^{n/2} u_k.$$

Now when $2 \leq k \leq 10$, we see

$$u_k \leq e^k n^k (\frac{\log n + c}{n})^{k-1} e^{-k(n-10)\frac{\log n + c}{n}} \leq (1 + o(1)) e^{k(1-c)} (\frac{\log n}{n})^{k-1}$$

And if we look at $k > 10$, we have

$$u_k \leq (\frac{ne}{k})^k k^{k-2} (\frac{\log n + c}{n})^{k-1} e^{-k(\log n + c)/2} \leq n(\frac{e^{1-c/2+O(1)} \log n}{n^{1/2}})^k.$$

Therefore combining this, we observe

$$\sum_{k=2}^{n/2} u_k \leq (1 + o(1)) \frac{e^{-c} \log n}{n} + \sum_{k=10}^{n/2} n^{1+o(1)-k/2} = O(n^{o(1)-1}).$$

Therefore, we see that

$$\mathbb{P}(\mathbb{G}_{n,p} \text{ is connected}) = \mathbb{P}(X_1 = 0) + o(1).$$

Then, it is known that when $p = \frac{\log n + c}{n}$, the isolated vertices is asymptotically Poisson distributed and has probability $e^{-e^{-c}}$.

This proves our theorem. ■

Now, lets create a more precise result of a random graph by using the idea that a random graph will become connected exactly at the moment when the last isolated vertex disappears.

Theorem 7.2. *Consider the random graph process $\{\mathbb{G}_m\}$. Let*

$$m_1^* = min\{m : \delta(\mathbb{G}_m) \geq 1\}$$
$$m_c^* = min\{m : \mathbb{G}_m is\ connected\}$$

Then w.h.p.

$$m_1^* = m_c^*$$

Quick note that $\delta(\mathbb{G}_m)$ simply means the minimum degree of some vertex. Informally this theorem states that the minimum edges required such that the degree of each vertex is at least 1 is equal to the minimum of edges needed for a connected random graph with high probability.

Proof. We let $m_\pm = 1/2n \log n \pm 1/2n \log \log n$ and $p_\pm = \frac{m_\pm}{N} \approx \frac{\log n \pm \log \log n}{n}$.
We will first show that w.h.p.

(1) G_{m_-} consists of a giant connected component plus a set V_1 of at most $2 \log n$ isolated vertices,
(2) G_{m_+} is connected

First let us show that these conditions are sufficient.
We will assume that (1) and (2) are true. Then it follows that w.h.p. $m_- \leq m_1^* \leq m_c^* \leq m_+$. We create $\mathbb{G}_{m_+}$ by taking $\mathbb{G}_{m_-}$ and adding $m_+ - m_-$ edges on random. Also the case of equality for $m_1^* = m_c^*$ occurs if none of the edges are contained by the set of isolated vertices.

$$\mathbb{P}(m_1^* < m_c^*) \leq o(1) + (m_+ - m_-)\frac{1/2|V_1|^2}{N - m_+} \leq o(1) + \frac{2n((\log n)^2) \log \log n}{1/2n^2 - O(n \log n)} = o(1)$$

This shows that these conditions will be sufficient to prove the theorem. Now let

$$p_- = \frac{m_-}{N} \approx \frac{\log n - \log \log n}{n},$$

Let X_1 be the random variable for isolated vertices in $\mathbb{G}_{n,p_-}$. Then

$$\mathbb{E}X_1 = n(1 - p_-)^{n-1} \approx ne^{-np_-} \approx \log n.$$

In addition,

$$\mathbb{E}X_1^2 = \mathbb{E}X_1 + n(n-1)(1-p_-)^{2n-3} \leq \mathbb{E}X_1 + (\mathbb{E}X_1)^2(1-p_-)^{-1}.$$

Thus we can now calculate the variance.

$$\text{Var}\, X_1 \leq \mathbb{E}X_1 + 2(\mathbb{E}X_1)^2 p_-,$$

It follows

$$\mathbb{P}(X_1 \geq 2 \log n) = \mathbb{P}(|X_1 - \mathbb{E}x_1| \geq (1 + o(1))\mathbb{E}X_1) \leq (1 + o(1))(\frac{1}{\mathbb{E}X_1} + 2p_-) = o(1)$$

Since having at some amount of isolated vertices is a monotone property, then w.h.p. $\mathbb{G}_{m_-}$ has less then $2 \log n$ vertices.
Now we will show that the rest of our random graph is a single connected component. We again define X_k as the random variable counting the number of components with k vertices in $\mathbb{G}_{p_-}$
We will use the calculations from the last Theorem to see that

$$\mathbb{E}(\sum_{k=2}^{n/2} X_k) = O(n^{o(1)-1})$$

We let $\mathscr{E} = \{\exists$ component of order $2 \leq k \leq n/2\}$.
Then we see that $\mathbb{P}(\mathbb{G}_{m_-} \in \mathscr{E}) \leq O(\sqrt{n}\mathbb{P}(\mathbb{G}_{p_-} \in \mathscr{E} = o(1)$ Which completes the proof of (1).
To prove (2), we again use 7.1. We know that by simplification $\mathbb{G}_m$ is connected w.h.p. if $n\frac{m}{n} - \log n \to \infty$. Therefore know that
$\frac{nm_+}{N} = \frac{n(1/2n\log n + 1/2n\log\log n)}{N} \approx \log n + \log\log n$ ■

These ideas can also be applied to find a threshold for k-connectivity of a random graph, which is simply that any k-1 vertices can be removed without disconnecting our graph. However for the scope of this paper, it will not be covered. Here is a theorem, which we will later use to help find a threshold for a Hamilton Cycle.

Theorem 7.3. *Let* $m = \frac{1}{2}n(\log n + (k-1)\log\log n + c_n), k = 1, 2, ...$ *Then*

$$\lim_{n\to\infty} \mathbb{P}(\mathbb{G}_m \textit{ is k-connected}) = \begin{cases} 0 & \textit{if } c_n \to -\infty \\ e^{\frac{-e^{-c}}{(k-1)!}} & \textit{if } c_n \to c \\ 1 & \textit{if } c_n \to \infty \end{cases}$$

The proof is very complicated so it will not be covered.

8. Perfect Matchings

Now that we established a threshold for connectivity, lets move on to another popular graph property, known as Perfect Matchings. A perfect matching is when there exists some matching of vertices so that each vertex is joined to precisely one other vertex. Formally, a perfect matching is a collection of independent edges covering all of the vertices of a graph. First let us define some notation. Let $\mathbb{G}_{n,n,p}$ be the random bipartite graph with vertex bipartition $V = (A, B)$, where $A = [1, n]$ and $B = [n+1, 2n]$, with each of the n^2 edges appearing independently with probability p.
Before we are able to determine any thresholds for perfect matching, we first must introduce two theorems. Hall's Marriage Theorem and Tutte's Theorem, both of which provide necessary and sufficient conditions for a perfect matching to exist.

Theorem 8.1. *(Hall's Marriage Theorem) Let* $G = (X, Y, E)$ *be a bipartite graph with bipartite vertex sets* X, Y *and edge set* E*. Then, for any subset* W *of* X*, there is a perfect matching from* X *to* Y *if and only if*

$$|W| \leq |N_G(W)|$$

Informally, this just means that the size of the neighbors of any subset of vertices in X must at least be the size of the subset itself. A proof is relatively straight forward. This condition is necessary as follows: if we find some subset of vertices in which $|W| > |N_G(W)|$, then at least one vertex in W cannot be matched to a neighbor by pigeonhole principle. A proof for sufficiency is much more difficult and is also unrelated to our exploration of random graphs and therefore is left out. If you are curious of a proof, both brilliant.org and wikipedia contain good proofs of this theorem.
Another theorem crucial to our exploration of perfect matchings is an analogue to Hall's.

Theorem 8.2. *(Tutte's Thoerem) A graph* G *has a perfect matching iff*

$$\forall S \in V, o(G\backslash S) \leq |S|$$

Where $o(G \backslash S)$ represents the number of odd components that get generated if S is removed from G.

Again as with Hall's, we will not provide a proof as it is somewhat complex and unrelated to our exploration of thresholds for perfect matchings. However, if interested in a proof, click here for a great article about Tutte's theorem. [W+01]
Now equipped with two strong theorems about perfect matchings, we are ready to find thresholds in our random graphs. Lets begin with Bipartite graphs.

Theorem 8.3. *Let $\omega = \omega(n)$, $c > 0$ be a constant, and $p = \frac{\log n + \omega}{n}$. Then*

$$\lim_{n\to\infty} \mathbb{P}(\mathbb{G}_{n,n,p} \text{ has a perfect matching}) = \begin{cases} 0 & \text{if } \omega \to -\infty \\ e^{-2e^{-c}} & \text{if } \omega \to c \text{ (a constant)} \\ 1 & \text{if } \omega \to \infty. \end{cases}$$

Moreover,
$\lim_{n\to\infty} \mathbb{P}(\mathbb{G}_{n,n,p}$ *has a perfect matching*$) = \lim_{n\to\infty} \mathbb{P}(\delta(\mathbb{G}_{n,n,p}) \geq 1)$.

Proof. We will be using Hall's condition in order to determine whether or not our graph has a perfect matching. We can make a minor adjustment to Hall's below:

$$\forall S \subseteq A, |S| \leq \frac{n}{2}, |N(S)| \geq |S|$$

And

$$\forall T \subseteq B, |T| \leq \frac{n}{2}, |N(T)| \geq |T|.$$

The reason behind this idea is that if $|S| > n/2$ and $|N(S)| < |S|$, then $T = B \backslash N(S)$ which is a contradiction.
We can now focus on S and T which should satisfy $|S| = |T| + 1$ and each vertex in T having at least 2 neighbors in S. Take some pair S, T with $|S| + |T|$ as small as possible. If the minimum degree $\delta \geq 1$, then $|S| \geq 2$

(1) If $|S| > |T| + 1$, we can remove $|S| - |T| - 1$ vertices from $|S|$, a contradiction.
(2) Suppose $\exists\ w \in T$ such that w has less than 2 neighbors in S. We will remove w and its unique neighbor in $|S|$, a contradiction.

Now it follows that

$$\mathbb{P}(\exists v : \text{v is isolated}) \leq \mathbb{P}(\nexists \text{ a perfect matching})$$

which is

$$\leq \mathbb{P}(\exists v : \text{v is isolated}) + 2\mathbb{P}(\exists S \subseteq A, T \subseteq B, 2 \leq k = |S| \leq n/2, |T| = k-1, N(S) \subseteq T \wedge e(S:T) \geq 2k-2).$$

Where $e(S : T)$ denotes the number of edges between the two sets and is at least $2k - 2$.
Now, let $p = \frac{\log n + c}{n}$ for some constant c. Then we define Y to be the random variable counting the number of sets S and T not satisfying the conditions we listed prior. We have:

$$\mathbb{E}Y \leq 2 \sum_{k=2}^{n/2} \binom{n}{k} \binom{n}{k-1} \binom{k(k-1)}{2k-2} p^{2k-2} (1-p)^{k(n-k)}$$

Now if we do a bit of approximating with the binomials, we get

$$\leq 2\sum_{k=2}^{n/2}(\frac{ne}{k})^k(\frac{ne}{k-1})^{k-1}(\frac{ke(\log n+c)}{2n})^{2k-2}e^{-npk(1-k/n)}$$

And with a bit more simplification observe that it becomes

$$\leq 2\sum_{k=2}^{n/2}n(\frac{e^{O(1)}n^{k/n}(\log n)^2}{n^{1-1/k}} = 2\sum_{k=2}^{n/2}u_k$$

We now break this into two cases

(1) Case 1: $2 \leq k \leq n^{3/4}$ We see that

$$2\sum_{k=2}^{n^{3/4}}u_k = O(\frac{1}{n^{1/2-o(1)}})$$

(2) Case 2: $n^{3/4} < k \leq n/2$ Here we see that

$$\leq 2\sum_{k=n^{3/4}}^{n/2}u_k = O(n^{-n^{3/4}/3})$$

Now we arrive at the conclusion that $\mathbb{P}(\exists v:\ \text{v is isolated}) + o(1) = \mathbb{P}(\nexists \text{ a perfect matching})$ Then if we let X_0 count the number of isolated vertices in $\mathbb{G}_{n,n,p}$, we have

$$\mathbb{E}X_0 = 2n(1-p)^n \approx 2e^{-c}$$

Then through Principle Inclusion Exclusion, we see that $\mathbb{P}(X_0 = 0) \approx e^{-2e^{-c}}$ If we want to prove the case for $|\omega| \to \infty$, we can use the fact this property is monotone increasing and that the value $e^{-2e^{-c}}$ approaches 0 when c is a very large negative number and infinity when c is very large positive number. ■

Now you might be wondering what would happen if we had a Non-Bipartite Graph? Now we could replace Hall's with Tutte's theorem, which was what Erdős and Rényi did. However, a simpler approach was actually found by Bollobás and Frieze. The proof is very complex, so we will leave it out, but we will provide several lemmas that make up the proof and are good to know for the next section.

Theorem 8.4. *Let $\omega = \omega(n)$, $c > 0$ be a constant, and let $p = \frac{\log n + c_n}{n}$. Then for even n,*

$$\lim_{n\to\infty}\mathbb{P}(\mathbb{G}_{n,p} \text{ has a perfect matching}) = \begin{cases} 0 & \text{if } c_n \to -\infty \\ e^{-e^{-c}} & \text{if } c_n \to c \text{ (a constant)} \\ 1 & \text{if } c_n \to \infty. \end{cases}$$

Moreover,
$\lim_{n\to\infty}\mathbb{P}(\mathbb{G}_{n,p}$ *has a perfect matching*$) = \lim_{n\to\infty}\mathbb{P}(\delta(\mathbb{G}_{n,n,p}) \geq 1)$.

The first lemma is:

Lemma 8.5. *Let G be a graph without a perfect matching and let M be a maximum matching and v be a vertex isolated by M. Then $|N_G(A(v,M))| < |A(v,M)|$*

As a reminded, a maximum matching is a matching that contains the most amount of edges possible.

Proof. Suppose that $x \in N_G(A(v,M))$ and that $f = u, x \in E$ where $u \in A(v,M)$. There exists some y such that $e = \{x,y\} \in M$, or else $x \in S_0(N) \subseteq A(v,M)$. Now we make the argument that $y \in A(v,M)$. This way, every neighbor of $A(v,M)$ is the neighbor via an edge of M.
Suppose that $y \notin A(v,M)$. We define M' to be a maximum matching that 1. isolates u and 2. is obtainable from M via a sequence of flips. Now $e \in M'$ because if e has been flipped out, then either x or y is placed in $A(v,M)$. But then, we can perform another flip with M', e and the edge $f = \{u,x\}$, placing $y \in A(v,M)$, which is a contradiction. ∎

If you let $p = \frac{\log n + \theta \log\log n + \omega}{n}$, where $\omega = o(\log\log n)$, a new lemma will arise.

Lemma 8.6. *Let* $M = 100(\theta+7)$. *w.h.p.* $S \subseteq [n]$, $|S| \leq \frac{n}{2e(\theta+5)M}$ *implies* $|N(S)| \geq (\theta+1)|S|$, *where* $N(S) = N_{\mathbb{G}_{n,p_1}}(S)$.

These lemmas and theorem will contribute a lot of ideas to our proofs of Hamilton Cycles, the focus of our next section.

9. Hamilton Cycles

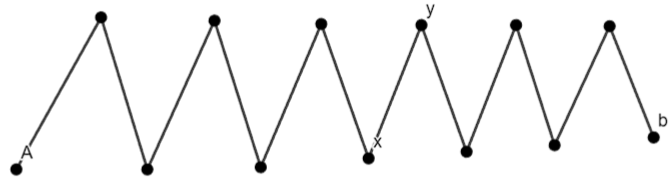
At first glance, you may be wondering what does a Hamilton Cycle have to do with perfect matchings? Well a perfect matching could be viewed as half a Hamilton cycle. This allows us to prove thresholds for the existence of perfect matchings and Hamilton cycles in very similar fashion. A Hamilton cycle for those who do not know is a cycle that goes through every vertex once and ends back at the start vertex.
The following theorem can be credited

Theorem 9.1. *Let* $p = \frac{\log n + log \log n + c_n}{n}$. *Then*

$$\lim_{n\to\infty} \mathbb{P}(\mathbb{G}_{n,p} \text{ has a Hamilton cycle}) = \begin{cases} 0 & \text{if } c_n \to -\infty \\ e^{-e^{-c}} & \text{if } c_n \to c \text{ (a constant)} \\ 1 & \text{if } c_n \to \infty. \end{cases}$$

Moreover,
$\lim_{n\to\infty} \mathbb{P}(\mathbb{G}_{n,p}$ *has a Hamilton cycle*$) = \lim_{n\to\infty} \mathbb{P}(\delta(\mathbb{G}_{n,n,p}) \geq 2)$.

Proof. The second statement will not be proved in this paper. For the proof of the first statement, we start with the assumption that $c_n = \omega \to \infty$, where $\omega = o(\log\log n)$. Now under this assumption we see that $\delta(\mathbb{G}_{n,p}) \geq 2$ w.h.p. due to 7.3. We see that for larger p, it will follow due to monotonicity.
Now, we will set up Pósa's lemma. We define P to be a path with end points a, b, as seen in figure 9.1. Now, suppose that b does not have a neighbor outside of P.



This is figure 9.1, the path of P.
Now observe that P' in figure 9.2 is a path of the same length as P, just rotated with vertex a fixed as the endpoint.
To be formal, we suppose that $P = (a, ..., x, y, y', ..., b', b)$ and $\{b, x\}$ is an edge where x is an interior vertex of P. The path $P' = (a, ..., x, b, b', ..., y', y)$ is said to be obtained from P by a rotation.

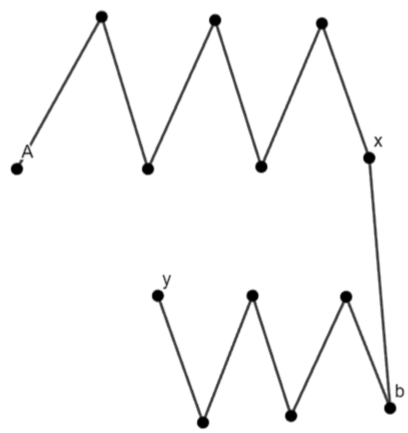


Figure 9.2, the path P' obtained after a single rotation

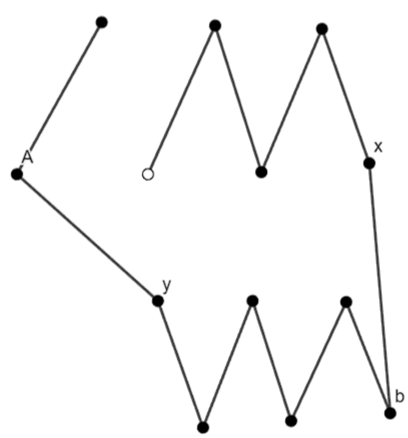


Figure 9.3
Now we let END = END(P) denote a set of vertices v such that there exists a path P_v from a to v such that P_v is obtained from P by a sequence of rotations with vertex a fixed as in Figure 9.3. We let the set END consist of all the white vertices on the path drawn below in figure 9.4.

This leads us to the following lemma

Lemma 9.2. *If $v \in P \backslash END$ and v is adjacent to $w \in END$, then there exists $x \in END$ such that the edge $\{v, x\} \in P$.*

Proof. We suppose the contrary that x, y are neighbors of v on P and that $v, x, y \notin END$ and that v is adjacent to $w \in END$. We now consider the path P_w. Let $\{r, t\}$ be the neighbors of v on P_w. Now $\{r, t\} = \{x, y\}$, as if the rotation deleted $\{v, y\}$ say then v or y becomes an endpoint. But then after a further rotation from P_w we see that $x \in END$ or $y \in END$.

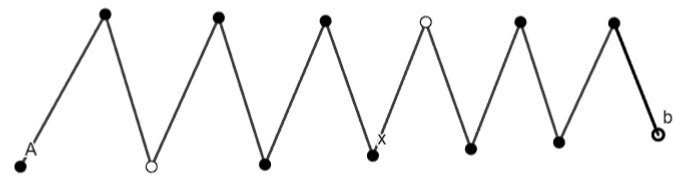


Figure 9.4 The set END

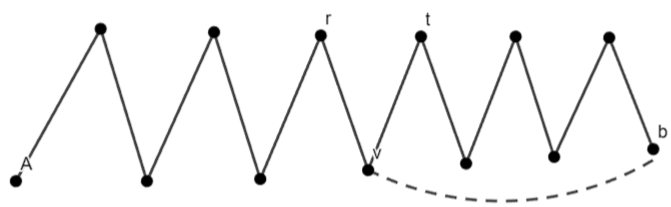


Figure 9.5: One of r, t will become an endpoint after a rotation ■

One surprising cool application of this helps us determine an algorithm which searches for a Hamilton Cycle in some connected graph G. For this case, the probability p_1 is above the connectivity threshold. We take some $\mathbb{G}_{n,p_1}$ and this algorithm will proceed in stages and grow it one by one.

9.1. **Algorithm Pósa:**

(1) Let P be our path at the start of stage K with endpoints x_0, y_0. If x_0 or y_0 have neighbors outside P, then we just extend P to include one of these neighbors and we proceed to the next stage.
(2) If this fails, we can perform rotations, where we fix x_0 until we get two possible scenarios.
   (a) We produce some path Q with an endpoint y that has a neighbor outside of Q. Now we are back at step 1.
   (b) If no rotation gives us the first case, let END denote the set of endpoints our paths produced. If $y \in END$ then P_y denotes a path with endpoints x_0, y that is obtained from P by a sequence of rotations. For each $y \in END$ we let $END(y)$ denote the set of vertices z such that there exists a longest path Q_z from y to z such that Q_z is obtained from P_y by a sequence of rotations where we fix y. Now repeat from the start of step 2. This either allows us to extend a path and move on to the next stage, or we proceed to the next step.

(3) Suppose that we are still unable to extend our path and that we have constructed all sets of END and $END(y)$. Suppose that G contains some edge (y,z) where $z \in END(y)$. This edge implies the existence of some cycle $C = (z, Q_y, z)$. If this is not a Hamilton Cycle, then as we know our graph is connected, we know there exists some $u \in C$ and $v \notin C$, such that u, v are joined by an edge. Let w be a neighbor of u on C and let P' be the path obtained from C by deleting the edge (u, w) . Now if we connect w, P', v, we get a cycle and we can move to the next stage.

A pair z, y where $z \in END(y)$ is called a booster in the sense that if we added this edge to $\mathbb{G}_{n,p_1}$, then it would either (1) make the graph Hamiltonian or (2) make the current path longer. We argue now that $\mathbb{G}_{n,p_2}$ can be used to "boost" P to a Hamilton cycle if needed.
Now we observe that $G = \mathbb{G}_{n,p_1}$, $|END| \geq \alpha n$ w.h.p. due to our previous lemma with $\theta = 1$. We also have $|END(y)| \geq \alpha n$ for all $y \in END$. This means we have $\omega(n^2)$ boosters.
For a graph G, we define $\lambda(G)$ as the length of the longest path in G, where G is not Hamiltonian and let $\lambda(G) = n$ when G is Hamiltonian. Now we let the edges of $\mathbb{G}_{n,p_2}$ be $\{f_1, f_2, ..., F_s\}$ in some random order with $s \approx \omega n/4$. Now let $\mathbb{G}_0 = \mathbb{G}_{n,p_1}$ and $\mathbb{G}_{\beth} = \mathbb{G}_{n,p_1} + \{f_1, f_2, ..., f_i\}$ for $i \geq 1$. Now due to the lemmas we had in perfect matching, if $\lambda(\mathbb{G}_i) < n$ then, we have

$$\mathbb{P}(\lambda(\mathbb{G}_{i+1})) \geq \lambda(\mathbb{G}_i) + 1 | f_1, f_2, ..., f_i) \geq \frac{\alpha^2}{2},$$

By replacing $A(v)$ by $END(v)$.
Now it follows that

$$\mathbb{P}(\mathbb{G}_{n,p} \text{ is not Hamiltonian}) \leq o(1) + \mathbb{P}(\text{Bin}(s, \alpha^2/2) < n) = o(1).$$

■

This concludes our proof for the phase transition of a Hamilton Cycle. The technique of rotations is pivotal to the proof and is certainly worth taking away from the proof.

References

[BB98] Béla Bollobás and Béla Bollobás. *Random graphs*. Springer, 1998.

[DE123] Jun 2023.

[FK16] Alan Frieze and Michał Karoński. *Introduction to random graphs*. Cambridge University Press, 2016.

[KS13] Michael Krivelevich and Benny Sudakov. The phase transition in random graphs: A simple proof. *Random Structures & Algorithms*, 43(2):131–138, 2013.

[W$^+$01] Douglas Brent West et al. *Introduction to graph theory*, volume 2. Prentice hall Upper Saddle River, 2001.

The Earth
Email address: `dieteryang@gmail.com`