On random graphs

Leo Davy

March 2019

Contents

1 An introduction to random graphs **2**
- 1.1 Graph theory . . . 2
- 1.2 Random graphs . . . 4
- 1.3 Cayley's formula . . . 4

2 The Erdos-Renyi Model **7**
- 2.1 Different approaches of the same space . . . 7
- 2.2 Connectivity . . . 8
- 2.3 Existence of thresholds . . . 10
- 2.4 The stability number . . . 12
- 2.5 The diameter . . . 13

3 Branching processes on random graphs **21**
- 3.1 Galton-Watson trees . . . 21
- 3.2 The exploration process, Karp's new approach . . . 29
- 3.3 The subcritical case : $\lambda < 1$ . . . 29
- 3.4 The supercritical case : $\lambda > 1$ . . . 29
- 3.5 Some words on the critical case . . . 29

4 The configuration model **30**
- 4.1 Random regular graphs . . . 30
- 4.2 An arbitrary degree sequence - the Newman-Watts-Strogatz model 30

A Some probabilistic tools **31**
- A.1 Common inequalities . . . 31
- A.2 Tail inequalities . . . 31
- A.3 Markov chains . . . 31
- A.4 Martingales . . . 31

Chapter 1

An introduction to random graphs

1.1 Graph theory

Intuitively, graphs are just about dots and lines, any phenomenon that can be represented as dots connected, or not, by lines can be thought of as a graph. Hence it is clear that graph theory, the study of the so called graphs that we will define in the following, will apply to a very wide variety of problems, such as, epidemiology, sociology, internet analysis, electric circuits, road traffic, ... and of course mathematics. Historically graphs first appeared in mathematics from Leonard Euler who gave, again, his name to a formula that would be the basis for the development of topology.
Formally a graph G will be defined as $G = (V(G), E(G))$, with $V(G)$ designing the vertex set (the points) of G and $E(G)$, disjoint from $V(G)$, the set of edges (the lines) of G. For now we will consider edges only as a pair of elements, without order, contained in $V(G)$. In the following of this report hypergraphs and directed graphs will be studied in which edges are composed of more than two elements or edges have a direction.
As an example of a graph we can consider the following graph with $V(G) = \{a, b, c, d, e\}, E(G) = \{e_1, e_2, e_3, e_4, e_5, e_6, e_7, e_8\}$ altogether with an incidence function $\psi_G : V(G) \to E(G)$ such as:

$$\psi_G(e_1) = ab \quad \psi_G(e_2) = ac \quad \psi_G(e_3) = bc \quad \psi_G(e_4) = ad$$
$$\psi_G(e_5) = cd \quad \psi_G(e_6) = cd \quad \psi_G(e_7) = ee \quad \psi_G(e_8) = ae$$

An equivalent way to define a graph would be through the incidence matrix $M_G = (m_{ve})$ with m_{ve} the number of times the vertice v and the edge e are incident. So m_{ve} can take the values 0 (not incident), 1 or 2 (e is a loop). It is also possible to define a graph in a third way, that is equivalent for the structure but doesn't take in account the labelling of the edges, it is through the adjacency matrix $A_G = (a_{uv})$. This is usually the most useful version as

typically a graph will have less vertices than edges, the adjacency matrix will be much smaller to write (hence to store in a computer) and it usually gives all the information needed to study a graph. It's interesting to note that an adjacency matrix is real and hermitian, thus all of it's eigenvalues are real and the study of it's distribution is a common topic in graph theory.
An interesting property of graphs is the degree of the vertices, so we will denote by $d_G(v)$ the number of edges incident with $v \in V(G)$. And we can also define the two following notations that will prove useful in the following of this report, $\delta(G)$ as the minimal degree of G and $\Delta(G)$ as the maximal degree of G. From these definition we can obtain the following lemma, with m the number of edges.

Theorem 1.1.1. For any graph finite graph G

$$\sum_{v \in V(G)} d_G(v) = 2m \tag{1.1}$$

Proof. The sum of the elements of each columns in the incidence matrix is equal to two. So the sum of the values in the colums over all the columns is equal to two times the number of columns, so $2m$. As the sum of the columns is equal to the sum of the rows, and the sum of each row is exactly the degree of a vertex, we have the result. □

This theorem will prove useful in the following as it connects the number of edges and the degrees of the vertices. In graph theory a graph can be represented in many different ways and then it can be really non trivial to know if two graphs with different labelling are the same. More formally, for a same graph, we will call the set of permutation of the labellings that doesn't change the structure of the graph, it's automorphism group, denoted $Aut(G)$ and $aut(G)$ it's cardinal. And for the anecdote, finding the problem of showing that two graphs are in the same automorphism group is a NP-hard problem.
One of the most fundamental properties of graphs is also the connexity. We will say that a graph is connected, if there is a path connecting any two edges. We define a path as a sequence of vertices connected by edges linking it's two ends. In fact we will consider simple paths, that are paths without loops, a simple path is always defined when there is a path. If v is a vertex, we will write $N(v)$ the set of vertices adjacent to v, from this definition we may observe that $d_G(v) = |N(v)|$. And we will call a component of a vertex the set of the vertices that can be reached from this vertex. Then a connected graph is a graph with only one component.
Some interesting graphs to which we might often refer are the complete graph on n vertices, denoted by $\mathbb{K}_n$, and the complete bipartite graph $\mathbb{K}_{n,m}$. A complete graph is a graph in which for any vertex, the set of neighbours is the rest of the graph. A graph is bipartite if it's set of vertices can be partitioned in two subsets X and Y such that every edge has one end in X and one in Y. The complete bipartite graph is a bipartite graph such that for all $x \in X$ we have $N(x) = Y$. This implies the same condition on the vertices in Y.
We call a simple graph a graph that doesn't contain any loop or multiple edge.

We will mainly study simple graphs as multigraphs or loopy graphs only add redundant information. We will see later that it is possible to mimic loopy graphs and multiple graphs by assigning weight to the edges.
It is also possible to define the union of two graphs simply as the union of each of the vertex sets and edge sets.
A very important kind of graphs are the directed graphs, it's a very intuitive notion, these are simply graphs on which there are arrows on the edges, equivalently, it is like defining the adjacency between two vertices as non symmetric, or as studying graph with non hermitian adjacency matrix
As there is usually no confusion possible we will denote $V = V(G), E = E(G), \psi = \psi_G, ...$

1.2 Random graphs

This section will try to give reasons behind the study of random graphs but it's purpose is not to go deep in the details and sophistication of their study. The study of random graphs is a flourishing area of mathematics since it's founding papers have been published by Erdos and Renyi between 1959 and 1963. Since then a lot of work has been done on random graphs, most of the questions on the Erdos-Renyi model have found satisfying answers, and the model being simplistic, many new models have been developed. So we will use the very vague definition by Janson.

Definition 1.2.1. A random graph is a graph where nodes, or edges, or both are selected by a random procedure.

Random graphs are interesting subject for pure mathematicians as they create a lot of open problems and offers many links with combinatorics. And for an applied mathematicians, random graphs are an entertaining tool as they may be used to simulate real world phenomenons (most famously in sociology, epidemiology or the study of internet). And accordingly to their level of matching with real life situations, they will be able to show the presence of complexity in the situation studied.

1.3 Cayley's formula

This section will first of all demonstrate an important result that will be used several times in crucial demonstrations in this report. Although it is not a demonstration that is specific to random graphs it may give an insight to the variety of techniques that may be used in the study of random graphs and how elegant are the results (at least quite often). This formula has been demonstrated in many different ways and we will use the demonstration by Joyal that is really elegant and also is a good place to introduce several notions that will be used in the rest of the report.

Theorem 1.3.1. Cayley's formula

$$t_n = n^{n-2} \tag{1.2}$$

with t_n the number of spanning trees on n vertices.

Before beginning the proof some definitions will be needed. A structure (graph or tree) is called spanning on the vertices (resp. edges) if it intersects all vertices (resp. edges). A tree is a special case of graph structure, that can be defined in several equivalent ways. For instance, a tree is a connected graph such that upon removal of any of it's edges it becomes disconnected, equivalently, it's a graph in which every two vertices are linked by exactly one path, equivalently, it's a connected and acyclic graph (doesn't contain any cycle).
The trees being a subset of the graphs, it is also possible to define directed trees in which you can follow an edge only in one direction (otherwise it would not be a tree anymore). We also define doubly rooted trees as trees with two special labels "Start" and "End" that can be attached to any vertices of the tree and which canonically maps on each edges the direction such that any vertice can reach the end. And we will call "SEL" the vertices that are in the "Start" to "End" line. We also denote by DRT_n the set of doubly rooted trees on n vertices.
As a consequence of this definition we have $|DRT_n| = n^2 t_n$. With $|\ |$ denoting the cardinal. To prove the theorem it would then be sufficient to prove that the number of elements in DRT_n is equal to n^n. So we will base our approach on Joyal's proof and show a bijection between the set of doubly rooted trees on n vertices and the the set of functions on n elements.

Proof. We will use the notation $[n] = \{1, 2, ..., n\}$ and $V = [n]$. Let's take $f : V \longrightarrow V$, and let's consider the graph of f. That is, $\forall v_1, v_2 \in V$ we have $v_1 \to v_2$ if and only if $f(v_1) = v_2$. Drawing such a graph for any function, and will appear two different kind of structures, first there will have directed line leading to cycles, and then cycles. And the whole graph will be a disjoint union of such components. It can be interesting to observe the case in which f is a permutation and then observe that the graph of f is a union of disjoint cycles as expected from the common group theory result.
We now take $C \subseteq V$ the set of vertices that are part of a cycle under the action of f. Equivalently,

$$C = \{x : \exists i \geq 1 s.t. f^i(x) = x\}$$

Let $k = |C|$ and write $C_<$ as $C_< = \{c_1 < c_2 < ... < c_k\}$ the ordered set and now we will construct a graph with the vertice set $D = f(C)$, and the edge set $E = \bigcup_{i=1}^{k-1} f(c_i)f(c_{i+1})$. We now have $G = (D, E)$ as a line of k vertices, and we will call $f(c_1)$ the "Start" and $f(c_k)$ the "End".
Now we will just append to this line the set of vertices that are not in G. So we construct $\tilde{E} = \bigcup_{x \in VC} xf(x)$ and $\tilde{G} = (V, E \cup \tilde{E})$ is a (directed doubly rooted) tree as it doesn't contain any cycle by construction and is clearly connected.

It's obviously directed and doubly rooted. We have now done the biggest part of the proof, that is, going from a function to a doubly rooted tree.
We will now take a doubly rooted tree and transform it in a function. From the definition of trees there is a unique "Start" to "End" (SEL) path.
For vertices on not on the SEL, for instance some vertice j, we define $f(j)$ as the first neighbour on the j to end line.
For vertices on the SEL,

$$SEL = \{x_1, x_2, ..., x_k\}, \text{ and } SEL_< = \{x_{\sigma_1}, x_{\sigma_2}, ..., x_{\sigma_k}\} \tag{1.3}$$

we define $f(x_{\sigma_i}) = x_i, \forall i \in [k]$.
Thus, we have two injective constructions, if composed give the identity, hence we have a bijection between the set of endomorphism of $[n]$ and the space of doubly rooted trees on n vertices. So the proof is complete. □

Chapter 2

The Erdos-Renyi Model

2.1 Different approaches of the same space

As said in the title of the section there are different ways to approach the Erdos-Renyi model that we may call paradigms as they will give us the same kind of results but depending on the context, one might be much more convenient to use than the others.
Historically the first paper published on random graphs was from Erdos and Renyi in 1959, in which they give the following construction :

Definition 2.1.1. We call a random graph $\mathcal{G}_{n,M}$ having n labelled vertices and M edges. That is we choose at random (with equal probability) one of the $\binom{\binom{n}{2}}{M}$ possible graphs.

One may observe that some changes in notations are made between this paraphrasing of the article of Erdos and Renyi, they are made in order to be more adapted with the modern study of random graphs. We will also adopt for the following $N = \binom{n}{2}$ to denote the total number of edges possible on n labelled vertices.
We then arrive to our main model that has been the most extensively studied in the literature of random graphs, that is $\mathcal{G}_{n,p}$ on which the coin tosses are no longer fair, but the probability of drawing an edge is now p. And the coin tosses are still independent. Now if we denote by e_G the number of edges of a graph G on the vertex set $[n]$. We have :

$$\mathbb{P}(G) = p^{e_G}(1-p)^{N-e_G} \tag{2.1}$$

This model is called the binomial model. It is easily seen that this model is asymptotically equivalent to the first one if Np is close to M on several aspects.
The third model that we will investigate is on the form of a Markov process, see in Annex for a discussion on properties used here from Markov chains. At time 0 there is no edge and an edge is selected at random among all of the possible edges. At time t, the edge is chosen among all the edges not already

present in the graph. We denote this process by $\{\mathcal{G}_{n,t}\}_t$, with t the number of edges added. It is clear that this model is perfectly equivalent to the first model presented in the case $t = M$. This model was also introduced in 1959 by Erdos and Renyi and is usually refered to as the random graph process. The advantage of this model is that it allows one to study properties on the verge of their realisations. For instance, using this model Bollobas proved that a graph is fully connected, when the last connection made is between an isolated vertice and the giant component. But we will study this in the following.

2.2 Connectivity

One of the most fundamental structure of a graph will be it's number of components. Hence, the first question we will try to ask is how often a random graph in the Erdos-Renyi model is connected. It is essential to answer this question as many other questions might not make sense on a graph that is not connected (for instance the diameter, the existence of hamiltonian paths or the stability number of the graph).
It is also an interesting first topic to have an insight of the kind of elegant results that arise from the study of random graphs. The main aim of this section will be to prove the following theorem in a didactic way as it is the first random graphs proof that we will study.

Theorem 2.2.1. Let $p = p(n) = \frac{log(n)+c}{n}$
Then $\lim_{n\to\infty} P(G \in \mathcal{G}_{n,p}$ is connected $) = e^{-e^{-c}}$

The proof of this theorem will be in two parts, first we will show that a graph will be connected if and only if there are no isolated vertices and then we will estimate the distribution that follows the number of isolated vertices.

Theorem 2.2.2. With $p = \frac{log(n)+c}{n}$ Almost every $G \in \mathcal{G}_{n,p}$ consists of a giant component and isolated vertices.

Proof. During this proof we will consider the random value X_k that counts the number of isolated vertices of order k. So, let's estimate the probability $P(X_2 > 0) = P(X_2 \geq 1)$. In order to do so we will use the method of first moment.

$$\mathbb{P}(X_2 \geq 1) \leq \mathbb{E}(X_2) = \binom{n}{2}\mathbb{P}(\text{"drawing an isolated edge"}) \tag{2.2}$$

$$= \binom{n}{2}p((1-p)^{n-1})^2 \tag{2.3}$$

$$\leq (\frac{ne}{2})^2 p(e^{-p})^{2(n-2)} \tag{2.4}$$

$$= \mathcal{O}\left(n^2 p\frac{e^2 e^{-2p(n+1)}}{4}\right) \tag{2.5}$$

$$= \mathcal{O}(n^2 p) \tag{2.6}$$

So it's sufficient that $p = o^(n-2)$ in order to have almost surely no edges in G. This is clearly satisfied by the p we use in the theorem.
However, this is not sufficient to prove that there is no isolated other than vertices. We will observe that there can't have any component of size larger than $\lceil \frac{n}{2} \rceil$ that is not the largest component in the graph. Hence, we will study, the probability that there is any component of intermediary size that is not connected to the greatest component.

$$\mathbb{P}(X_k \geq 1) \leq \mathbb{E}(X_k) \quad , \forall k \geq 3 \tag{2.7}$$
$$\leq \binom{n}{k} k^{k-2} q_k \tag{2.8}$$
$$\leq \binom{n}{k} k^{k-2} p^{k-1} ((1-p)^{n-k})^k \tag{2.9}$$

In the above, q_k represents the probability that a spanning tree on k vertices doesn't connect to the greatest connected components. A tree on k vertices having $k-1$ edges, this means in terms of probability that it must have $k-1$ "success" and on each of the k vertices $n-k$ failures. Which leads to the following line.
Now we will try to have an upper bound of the RHS such that the sum on k will converges to a $o(n^{-\delta})$ for some $\delta > 0$.

$$\mathbb{P}(X_k \geq 1) \leq k^{-2} p^{-1} (\frac{ne}{k})^k k^k p^k e^{-p\frac{n}{2}(n-k))} \tag{2.10}$$
$$\leq k^{-2} p^{-1} ((\frac{ne}{k}) kpe^{-p\frac{n}{2}})^k \tag{2.11}$$
$$\leq p^{-1} (nepe^{-p\frac{n}{2}})^k \tag{2.12}$$

If we denote the bracketed term by A, then

$$A = \mathcal{O}(\log(n) n^{-\frac{1}{2}}) \tag{2.13}$$

Hence, A goes to 0 without any condition on k, so we obtain

$$\sum_{k=3}^{\lceil n/2 \rceil} \mathbb{P}(X_k \geq 1) \leq p^{-1} \sum_{k=3}^{\lceil n/2 \rceil} A^k = o(1) \tag{2.14}$$

Which gives the fact that whether $\mathcal{G}_{n,p}$ is connected doesn't depend on the existence of isolated connected components of size 2 or more. So it only depends on the existence of isolated vertices. An intuitive point of view on that is some isolated component emerges it is really likely to be eaten by the giant component, then the only components that have a chance of not being absorbed by the giant component are the isolated vertex. So we have proved 2.2.2. □

The proof of 2.2.1 will be given in the part on branching processes.

2.3 Existence of thresholds

One of the most surprising features on random graphs, which seems to have motivated Erdos to publish results from 1959, is the existence of thresholds. That is, for many graph properties, with a small variation on the number of edges (in the ER model) or on $p(n)$, the limiting probability would jump between 0 and 1. This zone which produces great difference in limiting probability will be called a threshold. It has been shown by Bollobas and Thomason that this is in fact not exclusive to random graphs, but true for all monotone properties on random subsets.

Definition 2.3.1. We will call a graph property a family of graphs that is closed under isomorphism.

This means that a graph property is independent of the labelling and of the drawing of the graph. We can refine properties in the following definition.

Definition 2.3.2. A property is monotone increasing (resp. decreasing) if it's stable under the the addition (resp. removal) of an edge. A graph property $\mathcal{Q}$ is convex if when $A, C \in \mathcal{Q}$ and $A \subseteq B \subseteq C$ then $B \in \mathcal{Q}$.

For instance, being connected or containing a specific subgraph are monotone increasing properties where as being planar or containing an isolated vertice are monotone decreasing. As an example of property that is neither monotone increasing or decreasing, we can think of being k-regular for some k (this means that all vertices are of degree k). Having exactly k isolated vertices is an example of a convex not monotone property.
Here is a theorem showing that monotone increasing properties make probability distributions on these properties also monotone increasing.

Theorem 2.3.1. Suppose $\mathcal{Q}$ is a monotone increasing property and $0 \leq M_1 \leq M_2 \leq N$ and $0 \leq p_1 \leq p_2 \leq 1$.
Then

$$\mathbb{P}_{M_1}(\mathcal{Q}) \leq \mathbb{P}_{M_2}(\mathcal{Q}) \text{ and } \mathbb{P}_{p_1}(\mathcal{Q}) \leq \mathbb{P}_{p_2}(\mathcal{Q})$$

Proof. The first inequality is clear, as the only difference between the two spaces on which we evaluate the property $\mathcal{Q}$ is that on the RHS edges have been added, hence, the probability of realising a monotone increasing has been increased.
For the second inequality, let $p = \frac{p_2 - p_1}{1 - p_1}$. Let $G_1 \in \mathcal{G}_{n,p_1}, G_2 \in \mathcal{G}_{n,p}$
So if $G_2 = G_1 \cup G$ it's edges are chosen with probability $p_1 + p - p_1 p = p_2$. So G_1 is in G_2, the property being monotone increasing, we have $\mathbb{P}_{p_1}(\mathcal{Q}) \leq \mathbb{P}_{p_2}(\mathcal{Q})$ □

The following result follows from definition with $\mathcal{Q}$ a monotone increasing property

$$\mathbb{P}(\mathcal{Q}) = \sum_{A \in \mathcal{Q}} p^{|A|}(1-p)^{N-|A|} \tag{2.15}$$

However this result requires to know all of the elements in $\mathcal{Q}$ and as we are often interested with properties for very large n this result won't be magical...

However from the following lemmas it is very useful to obtain some results on the links between $\mathcal{G}_{n,p}$ mentionned in the introduction $\mathcal{G}_{n,M}$. Indeed the following theorem shows that if we know quite accurately $\mathbb{P}_M(\mathcal{Q})$ for every M close to pN then we know $\mathbb{P}_p(\mathcal{Q})$ with a comparable accuracy. The converse being clearly false, for instance the property of containing M edges.

Theorem 2.3.2. Suppose $\mathcal{Q}$ is any property and $0 < p = M/N < 1$
Then $\mathbb{P}_M(\mathcal{Q}) \leq 3\sqrt{M}\mathbb{P}_p(\mathcal{Q})$

Proof. Let $\mathcal{Q}$ be any property, then we will write $\mathcal{Q}$ as a partition based on the number of edges in each graph contained in $\mathcal{Q}$.
So we have

$$\mathcal{Q} = \bigsqcup_{m=0}^{N} \mathcal{Q}_m \quad , \text{ with } \forall G \in \mathcal{Q}_m, e(G) = m$$

$\mathbb{P}_m(\mathcal{Q}) = |\mathcal{Q}_m|\binom{N}{M}^{-1}$ From this we can obtain, with $q = 1 - p$

$$\begin{aligned}
\mathbb{P}_p(\mathcal{Q}) &= \sum_{A \in \mathcal{Q}} p^{|A|} q^{N-|A|} \\
&= \sum_{m=0}^{N} \sum_{A \in \mathcal{Q}_m} p^{|A|} q^{N-|A|} \\
&= \sum_{m=0}^{N} \sum_{A \in \mathcal{Q}_m} p^{m} q^{N-m} \\
&= \sum_{m=0}^{N} |\mathcal{Q}_m| p^{m} q^{N-m} \\
&= \sum_{m=0}^{N} p^{m} q^{N-m} \binom{N}{M} \mathbb{P}_m(\mathcal{Q}) \\
&\geq \binom{N}{M} p^{M} q^{N-M} \mathbb{P}_M(\mathcal{Q}) \\
&\geq \mathbb{P}_M(\mathcal{Q}) (e^{\frac{1}{6M}} \sqrt{2\pi pqN})^{-1}
\end{aligned}$$

So we have

$$\mathbb{P}_M(\mathcal{Q}) \leq \mathbb{P}_p(\mathcal{Q}) e^{\frac{1}{6M}} \sqrt{2\pi qM} \tag{2.16}$$

Observing that $q \leq 1$ and $\sqrt{2\pi} e^{\frac{1}{6}} \approx 2.961... < 3$ the proof is complete. □

The previous section was about connectivity in $\mathcal{G}_{n,p}$, in this section we have seen that connectivity can be characterized as a monotone increasing property. Also it was observed that the function p was somehow best possible, by that we mean that modifying it slightly would impply to only have a zero-one law. We call such a function p a threshold (in that case for the connectivity).

More formally, let $\mathcal{Q}$ a monotone increasing property, in $\mathcal{G}_{n,p}$, we call $\hat{p} = \hat{p}(n)$ a threshold if

$$\mathbb{P}(\mathcal{G}_{n,p} \in \mathcal{Q}) \to \begin{cases} 0 & \text{if } p \ll \hat{p}, \\ 1 & \text{if } p \gg \hat{p}. \end{cases} \tag{2.17}$$

Analogously, in $\mathcal{G}_{n,M}$, we call $\hat{M} = \hat{M}(n)$ a threshold if

$$\mathbb{P}(\mathcal{G}_{n,M} \in \mathcal{Q}) \to \begin{cases} 0 & \text{if } M \ll \hat{M}, \\ 1 & \text{if } M \gg \hat{M}. \end{cases} \tag{2.18}$$

In fact, thresholds are unique with respect to the multiplication by a scalar. So for the following, we should denote a threshold for a property as the threshold.

2.4 The stability number

Another property of graphs that one might be interested to study is the stability number. The stability number of a graph is the size of the largest set of vertices we can choose in a graph such that no two vertices are adjacent. One of the reasons that makes this an interesting property to study is that it is linked to one of the most fundamental property of graphs, $\chi(G)$, the chromatic number. Indeed, the chromatic number being the smallest number of colours that make a proper colouring of a graph. That means a colouring on which there are no two adjacent vertices of the same color. We are then led to a similar question than finding the stability number of a graph. If we denote by $\alpha(G)$ the stability number of a graph, we have the simple lower bound

$$\chi(G) \geq \frac{n}{\alpha(H)} \tag{2.19}$$

Before giving a lower bound on the stability number of a graph it might be interesting to denote that the notion of stable set is dual to the notion of clique and is analogous to the notion of perfect matching that concerns the edges. Although the following theorem will give a bound that is quite tight in $\mathcal{G}_{n,p}$ the problem of finding the actual maximum stable set of a graph is a NP-hard problem.

Theorem 2.4.1. The stability number of a graph in $\mathcal{G}_{n,p}$, is at most $\lceil 2p^{-1}\log(n)\rceil$.

Proof. Let $G \in \mathcal{G}_{n,p}$ and $S \subseteq V$ such that V contains $k+1$ vertices. Then we have

$$\mathcal{P}(\text{"S is a stable set""}) = (1-p)^{\binom{k+1}{2}} \tag{2.20}$$

as none of the $\binom{k+1}{2}$ possible edges must be selected.
Let's define our random values as follow, $X_S = \mathbb{1}(\text{" } S \text{ is a stable set "})$ and

$$X_{k+1} = \sum_{\substack{S \subseteq V \\ |S|=k+1}} X_S \tag{2.21}$$

the random variable counting the number of stable sets of size $k+1$, so what we are investigating here is the rank α such that $X_k = 0, \forall k > \alpha$. Such an α would then be maximal stability number.

$$\mathbb{E}X_{k+1} = \sum_{\substack{S\subseteq V \\ |S|=k+1}} \mathbb{E}X_S = \sum_{\substack{S\subseteq V \\ |S|=k+1}} \mathbb{P}(X_S = 1) \tag{2.22}$$

$$= \sum_{\substack{S\subseteq V \\ |S|=k+1}} (1-p)^{\binom{k+1}{2}} = (1-p)^{\binom{k+1}{2}} \sum_{\substack{S\subseteq V \\ |S|=k+1}} 1 \tag{2.23}$$

$$= \binom{n}{k+1}(1-p)^{\binom{k+1}{2}} \tag{2.24}$$

Now we will use the common inequalities $\binom{n}{k+1} \leq \frac{n^{k+1}}{(k+1)!}$ and $(1-p) \leq e^{-p}$. And we have

$$\mathbb{E}X_{k+1} \leq \frac{n^{k+1}}{(k+1)!} e^{-p\binom{k+1}{2}} = \frac{n^{k+1}}{(k+1)!} e^{-p\frac{k(k+1)}{2}} \tag{2.25}$$

$$\leq \frac{(ne^{-p\frac{k}{2}})^{k+1}}{(k+1)!} \tag{2.26}$$

So, if we consider $k = \lceil 2p^{-1}\log(n) \rceil \leq 2p^{-1}\log(n)$ we have that $ne^{-p\frac{k}{2}} \leq 1$. We finally obtain

$$\mathbb{E}X_{k+1} \longrightarrow_{n\to\infty} 0 \tag{2.27}$$

And then $X_{k+1} = 0$ almost surely which proves the theorem.

□

2.5 The diameter

Definition 2.5.1. The diameter of a graph is the greatest distance between any pair of vertice. We denote it by $\mathrm{diam}(G)$ and say it is equal to ∞ if the graph is not connected.

It is quite easy to understand that the diameter is a value that is a great importance particularly in applied systems. For instance, the small world phenomena is quite notorious (and will be discussed later in this report) but we can also think of optimization problems in which the fact that two points are far appart might be of great consequences. This section won't be focused on real world applications of the diameter because as we will see we are not studying the right model for this. Hence we will first discuss some graph theoretic problems and results on the diameter and after giving the main theorem on the diameter we will demonstrate it through several technical lemmas, some of which will be admitted. Finally some corollary will be obtained from the theorem.
One of the challenging questions in graph theory is estimating the following function

$$n(D,\Delta) = max\{|G|, \mathrm{diam}(G) \leq D, \Delta(G) \leq \Delta\} \tag{2.28}$$

n is the function that for a fixed diameter and a fixed maximal degree, gives the graph with a maximal number of verifies both conditions.
For instance if we take $\Delta = 2$ we obtain easily by construction that a graph that maximizes the number of vertices with a diameter D is a $(2D+1)$-cycle. Hence,

$$n(D, 2) = 2D + 1, \forall D \in \mathbb{N} \tag{2.29}$$

But it is in fact very hard to obtain such a formula for other values of D or Δ. If we take a graph of max degree Δ we observe that there are at most $\Delta(\Delta-1)^{k-1}$ vertices at distance k from a chosen vertex. It is easily to be convinced of this simply by drawing such a graph. From this very simple construction we can obtain the following upped bound

$$n(D, \Delta) \leq 1 + \Delta \sum_{j=1}^{D} (\Delta - 1)^{j-1} = \frac{\Delta(\Delta-1)^D - 2}{\Delta - 2} = n_0(D, \Delta) \tag{2.30}$$

n_0 is called the Moore bound and a graph for which the Moore bound is best possible is called a Moore graph. As Donald Knuth would have suspected, the Petersen's graph is such a graph with parameter $D = 2$ and $\Delta = 3$.
Now let's give some functions that will allow us to make a study of the diameter in random graphs. If we choose x a vertex in a graph, then we define $\Gamma_k(x)$ as the set of vertices at distance k from x. And from this we will define $N_k(x)$ as the set of vertices of distance less than or equal to k. Formally, we have

$$\Gamma_k(x) = \{v : d(x, v) = k\} \tag{2.31}$$

$$N_k(x) = \bigcup_{i=1}^{k} \Gamma_i(x) \tag{2.32}$$

And we can link those with the diameter using, $\text{diam}(G) \leq d$ if and only if $N_d(x) = V(G), \forall x$.
Similarly $\text{diam}(G) \geq d$ if and only if $\exists y, N_{d-1}(y) \neq V(G)$.
Now we will define the probabilty function that we will use in the following of this section as

$$p^d n^{d-1} = \log(\frac{n^2}{c}) \quad , \text{ for some } c > 0 \tag{2.33}$$

and as a teaser the following abbreviated version of the theorem we will properly give and prove later.

Theorem 2.5.1. $\mathbb{P}(d \leq \text{diam}(\mathcal{G}_{n,p}) \leq d + 1) \longrightarrow_{n\to\infty} 1$

This (inexact)theorem that is very strong states that all random graphs in $\mathcal{G}_{n,p}$ have (nearly exactly) the same diameter. In the actual theorem that we will prove we will be able to characterize exactly the probability of obtaining either d or $d + 1$ as a function of the parameter c in

Before proving such a theorem we will need some technical lemmas and assumptions. So we will give here some equations for reference later. So we will assume that $0 < p < 1$ and $d = d(n)$ is a natural number. Also

$$p\frac{\log(n)}{n}^{-1} \longrightarrow_{n\to\infty} \infty \tag{2.34}$$

that should not be too surprising as the threshold for connectivity is $\frac{\log(n)+c}{n}$ as shown before. We may also consider that $p = o(n^{-\frac{1}{2}+\epsilon}), \forall\epsilon > 0$ as it can be shown that otherwise the diameter of $\mathcal{G}_{n,p}$ is lower or equal than 2 (TODO : CHECK THIS) One might be interested in what $p(n, d, c)$ or $d(n, p, c)$ look like so here they are

$$p = n^{\frac{1}{d}-1}(\log(\frac{n^2}{c})^{\frac{1}{d}} \tag{2.35}$$

$$d = \frac{1}{\log(pn)}(\log(n) + \log\log(n) + log(2) + \mathcal{O}(\frac{1}{\log(n)})) \tag{2.36}$$

$$= \mathcal{O}((1 + o(1))\frac{\log(n)}{\log\log(n)}) \tag{2.37}$$

And finally

$$p(pn)^{d-2} = o(1) \tag{2.38}$$

The first lemma that we present here gives a tail inequality of $\Gamma_k(x)$ conditionnaly on some space that we will now define. $\Omega_k \subseteq \mathcal{G}_{n,p}$ is a set of graphs with $a = |\Gamma_{k-1}(x)|$ and $b = |N_{k-1}(x)|$ that satisfy

$$\begin{cases} \frac{1}{2}(pn)^{k-1} \leq a & \leq \frac{1}{2}(pn)^{k-1} \\ b & \leq 2(pn)^{k-1} \end{cases} \tag{2.39}$$

Lemma 2.5.2. Let x be a fixed vertex.
$1 \leq k = k(n) \leq d - 1$
And $K = K(n)$ that satisfy $6 \leq K \leq \frac{1}{12}\sqrt{pn\frac{1}{\log[n}}$ We also define

$$\alpha_k = K\sqrt{\frac{\log n}{(pn)^k}}, \quad \beta_k = p(pn)^{k-1}, \quad \gamma_k = 2\frac{(pn)^{k-1}}{n} = \frac{2\beta_k}{pn} \tag{2.40}$$

Then

$$\mathbb{P}(|\gamma_k(x) - apn| \geq (\alpha_k + \beta_k + \gamma_k)apn \quad | \quad \Omega_k) \leq n^{-\frac{K^2}{9}} \tag{2.41}$$

Proof. Knowing that we are in the space Ω_k, then the probability that some vertex y is at distance k, that is the probability that y is not in $N_{k-1}(x)$ but is connected to $\Gamma_{k-1}(x)$ is

$$p^a = 1 - (1 - p)^a \tag{2.42}$$

Hence, the random value $|\Gamma_k(x)|$ follows a binomial on $n - b$ elements, with a propability p_a. In the following inequalities we will assume that n is large

enough, this will allow us to do certain inequalities that are not really precise but asymptotically satisfying. Also to make it more easy for the reader we consider that all of these inequalities take place in Ω_k without writing it explicitly.

$$|\gamma_k(x) - apn| \geq (\alpha_k + \beta_k + \gamma_k)apn \tag{2.43}$$
$$\geq (\alpha_k + \beta_k)apn + \gamma_k apn \tag{2.44}$$
$$\geq (\alpha_k + \beta_k)apn + ap(n - n_k) \tag{2.45}$$
$$\geq (\alpha_k + \beta_k + 1 - \frac{n_k}{n})apn \tag{2.46}$$
$$\geq (\alpha_k + \beta_k)apn \geq (\alpha_k + \beta_k)apn_k \tag{2.47}$$
$$\tag{2.48}$$

From this first sequence of inequalities we managed to remove γ_k and we used some n_k that is less than n. The point in evaluating these inequalities is that we now have

$$\mathbb{P}(|\gamma_k(x) - apn| \geq (\alpha_k + \beta_k + \gamma_k)apn \quad | \quad \Omega_k) \tag{2.49}$$
$$\leq \mathbb{P}(|\gamma_k(x) - apn_k| \geq (\alpha_k + \beta_k)apn_k \quad | \quad \Omega_k) \tag{2.50}$$

Now using $ap - p_a \leq \beta_k ap$ and the triangular inequality (TODO : CHECK TRI INEQ) we have

$$\leq \mathbb{P}(|\gamma_k(x) - p_a n_k| \geq \alpha_k apn_k \quad | \quad \Omega_k) \tag{2.51}$$
$$\leq \mathbb{P}(|\gamma_k(x) - p_a n_k| \geq \alpha_k p_a n_k \quad | \quad \Omega_k) \tag{2.52}$$
$$\tag{2.53}$$

And using Theorem 1.7 from Bollobas we have

$$\leq \frac{1}{\sqrt{\alpha_k^2 p_a n_k}} \exp(-\frac{1}{3}\alpha_k^2 p_a n_k) \tag{2.54}$$
$$\leq \exp(-\frac{1}{3}\alpha_k^2 p_a n_k) \tag{2.55}$$

And using $p_a \geq pa(1 - \frac{pa}{2})$, $a \geq \frac{1}{2}(pn)^{k-1}$ and using $n_k = n - b$ we obtain $p_a n_k > \frac{(pn)^k}{3}$ that we insert in the previous inequality that will give us the result expected.

$$\leq \exp(-\alpha_k^2 \frac{(pn)^k}{9}) = n^{-\frac{K^2}{9}} \tag{2.56}$$

□

We will now prove another lemma

Lemma 2.5.3. Let $K > 12$ a constant and $\alpha_k, \beta_k, \gamma_k, k = 1, ..., d-1$ as before. Set

$$\delta_k = \exp(2\sum_{l=1}^{k}(\alpha_l + \beta_l + \gamma_l)) - 1 \tag{2.57}$$

Then if n is sufficiently large, with probability at least $1 - n^{-K-2}$ for every vertex x and every natural number $k, 1 \leq k \leq d-1$ we have

$$||\Gamma_k(x)| - (pn)^k| \leq \delta_k (pn)^k \tag{2.58}$$

Proof. As $\delta_{d-1} \longrightarrow_{n\to\infty} 0$ we may assume that $\delta_{d-1} < \frac{1}{4}$ For a fixed vertex x, we denote by Ω_k^* the set of graph for which

$$||\Gamma_l(x) - (pn)^l| \leq \delta_l (pn)^l \quad , 0 \leq l \leq k \tag{2.59}$$

And it is easy to verify that it is decreasing. Anf if one replaces $|\Gamma_l(x)|$ by a it is clear that we have

$$\Omega_k^* \subseteq \Omega_{k-1}^* \subseteq \Omega_k \tag{2.60}$$

Now, simply applying Bayes formula (for the complementary) we have

$$1 - \mathbb{P}(\Omega_k^*) = 1 - \mathbb{P}(\Omega_{k-1}^*) + \mathbb{P}(\Omega_{k-1}^*)\mathbb{P}(||\Gamma_k(x)| - (pn)^k| \leq \delta_k (pn)^k | \Omega_{k-1}^*) \tag{2.61}$$

If $G \in \Omega_{k-1}^*$ and $|a| = |\Gamma_{k-1}(x)|$ and by applying the definition of belonging to Ω_{k-1}^* and multiplying both sides by pn we have

$$|(pn)^k - apn| \leq \delta_{k-1}(pn)^k \tag{2.62}$$

And we obtain using the second triangular inequality

$$\mathbb{P}(\not\Omega_k^* | \Omega_{k-1}^*) \leq \mathbb{P}(\Omega_{k-1}^*)^{-1} \mathbb{P}(||\gamma_k(x)| - apn| \geq (\delta_k - \delta_{k-1})(pn)^k | \Omega_k) \tag{2.63}$$

$$\leq (1 - 2(k-1)n^{-\frac{K^2}{9}})^{-1} \mathbb{P}(||\Gamma_k(x)| - apn| \geq 2(\alpha_k + \beta_k + \gamma_k)(pn)^k | \Omega_k) \tag{2.64}$$

The last inequality being obtained from the hypothesis of induction and using $(1+x) \leq \exp(x)$. Now using the fact that $apn \leq \frac{3}{2}(pn)^k$ we have

$$\leq 2\mathbb{P}(||\Gamma_k(x)| - apn| \geq (\alpha_k + \beta_k + \gamma_k) apn | \Omega_k) \tag{2.65}$$

$$\leq 2n^{-\frac{K^2}{9}} \tag{2.66}$$

If we combine the last inequality that is obtained applying the previous lemma with 2.61 then we have the result. □

Theorem 2.5.4. Using all previous definitions on p and d, in $\mathcal{G}_{n,p}$ we have

$$\mathbb{P}(\text{diam}(G) = d) \longrightarrow e^{-\frac{c}{2}} \tag{2.67}$$

$$\mathbb{P}(\text{diam}(G) = d+1) \longrightarrow 1 - e^{-\frac{c}{2}} \tag{2.68}$$

Definition 2.5.2. For x and y two vertices of G, we say that x and y are remote if $y \notin N_d(x)$.

Proof. TODO : CHECK THE FOLLOWING EQ

$$\mathbb{P}(|N_{d-1}(x)| < \frac{5}{6}(pn)^{d-1}) < n^{-4} \tag{2.69}$$

Now we estimate the probability that a vertex y is joined to no vertex in a set W with $|W| \geq \frac{5}{6}(pn)^{d-1}$

$$(1-p)^{|W|} \leq \exp(-|W|) \leq \exp(-p|W|) = \exp(-\frac{5}{6}\log(\frac{n^2}{c})) = c^{\frac{5}{6}}n^{-\frac{5}{3}} \tag{2.70}$$

Hence, if x, y, z are vertices we have

$$\begin{aligned}
&\mathbb{P}(x \text{ is remote from } y \text{ and } z) && (2.71)\\
&\quad \leq \mathbb{P}(|N_{d-1}(x)| \leq \frac{5}{6}(pn)^{d-1}) && (2.72)\\
&\quad + \mathbb{P}(\{y,z\} \cap N_{d-1}(x) = \emptyset || N_{d-1}(x)| \geq \frac{5}{6}(pn)^{d-1}) && (2.73)\\
&\quad \leq \mathbb{P}(|N_{d-1}(x)| \leq \frac{5}{6}(pn)^{d-1}) && (2.74)\\
&\quad + (\mathbb{P}(y \text{ is not joined to } W = N_{d-1}(x) | W \geq \frac{5}{6}(pn)^{d-1}))^2 && (2.75)\\
&\quad \leq n^{-4} + c^{\frac{5}{3}}n^{-\frac{10}{3}} && (2.76)\\
&\quad \leq n^{-3}n^{-\frac{1}{4}} && (2.77)
\end{aligned}$$

So, we obtain

$$\begin{aligned}
&\mathbb{P}(\mathcal{G}_{n,p} \text{ contains two remote vertices pairs sharing a vertex }) && (2.78)\\
&\quad \leq \sum_x \sum_{(y,z)} \mathbb{P}(\ x \text{ is remote from } y \text{ and } z) && (2.79)\\
&\quad \leq \sum_x \binom{n}{2} n^{-3-\frac{1}{4}} = n\binom{n}{2} n^{-3-\frac{1}{4}} && (2.80)\\
&\quad \leq n^{-\frac{1}{4}} && (2.81)
\end{aligned}$$

The following lemma can be obtained quite easily simply by construction.

Lemma 2.5.5. From r disjoint pair of vertices, there are 2^r r-tuples of vertices meeting each pair.

TODO : CHECK FOLLOWING LEMMA

Lemma 2.5.6. The r-th factorial moment of a random value is within $o(1)$ of the expected number of ordered r-tuples of disjoint remote pairs.

If we denote by $\mathbb{F}_r$ the probability that a fixed r-tuple consists of vertices remote from each other. Then with X the random value that counts the number of remote pair of vertices.

$$\mathbb{E}_r(X) = \frac{n!}{(n-r)!}2\mathbb{F}_r(1+o(1)) + o(1) \tag{2.82}$$

Admitting a lemma from Bollobas TODO : ADD IT / PROVE IT ?
With probability at least $1 - n^{-K}$

$$(1 - n^{-K})Q_r \leq \mathbb{F}_r \leq (1 - n^{-K})Q_r + n^{-K} \tag{2.83}$$

With $Q_r = (\frac{c}{n})^r(1 + o(1))$, we obtain

$$\mathbb{F}_r = (\frac{c}{n})^r(1 + o(1)) \tag{2.84}$$

And we can obtain the asymptotical estimate

$$\mathbb{E}_r(X) = n^r 2^{-r}(\frac{c}{n})^r(1 + o(1)) + o(1) \tag{2.85}$$

Using the following theorem (TODO : PROVE IT SOMEWHERE)

Theorem 2.5.7. $\lim_{n\to\infty} \mathbb{E}_r(X_n) - \lambda^r = 0$
$\implies X \longrightarrow_d \mathcal{P}_\lambda$

Then we have that X converges in distribution to Poisson law of parameter $\frac{c}{2}$.
We can then obtain

$$\mathbb{P}(X = 0) = \mathbb{P}(\text{diam}(G) \leq d) \longrightarrow e^{-\frac{c}{2}} \tag{2.86}$$

that proves the first part of the theorem. We can also prove the case with $d = 1$ that gives

$$\mathbb{P}(\text{diam}(G) \leq 1) = \mathbb{P}(G = K_n) = p^{\binom{n}{2}} \longrightarrow 0 \tag{2.87}$$

And to finish the proof we need to consider the two following cases

$$p^{d-1}n^{d-2} = \log(\frac{n^2}{c_1(n)}) \tag{2.88}$$

And in that case we have $c_1 \longrightarrow \infty$.

$$\mathbb{P}(\text{diam}(G) \leq d + 1) = e^{-\frac{c_1}{2}} \longrightarrow 0 \tag{2.89}$$

If we now consider

$$p^{d+1}n^d = \log(\frac{n^2}{c_2(n)}) \tag{2.90}$$

we obtain $c_2 \longrightarrow 0$ which gives

$$\mathbb{P}(\text{diam}(G) \leq d + 1) = e^{-\frac{c_2}{2}} \longrightarrow 1 \tag{2.91}$$

Finally, if we combine the last result with 2.86 it is clear that the proof if complete. □

To end this section we present two theorems that are much easier to prove with a similar flavor to the previous one.

Theorem 2.5.8. Suppose

$$i) \quad p^2 n - 2\log(n) \longrightarrow \infty \tag{2.92}$$

$$ii) \quad n^2(1-p) \longrightarrow \infty \tag{2.93}$$

Then almost every graph in $\mathcal{G}_p$ has diameter 2.

Proof.

$$\mathbb{P}(\text{dist}(x,y) > 2) = (1-p^2)^{n-2} \tag{2.94}$$

as it is the probability that two vertices x and y are not connected by a path of length two.
Then using the the first moment method.

$$\mathbb{E}(\text{ pairs not connected by a path of length 2}) \tag{2.95}$$

$$= \binom{n}{2}(1-p^2)^{n-2} \geq \mathbb{P}(\text{diam}(G) > 2) \tag{2.96}$$

As from i) the LHS converges to zero, we are almost sure that no graph is of diameter more than 2 and then we just have to prove that there is almost no graph of diameter 1. A graph of diameter 1 being a complete graph we have

$$\mathbb{P}(\text{diam}(G) = 1) = \mathbb{P}(G = K_n) = p^{\binom{n}{2}} \longrightarrow 0 \tag{2.97}$$

$$\text{iff } \log p^{\binom{n}{2}} \longrightarrow -\infty \tag{2.98}$$

$$\text{iff } n^2 \log p \longrightarrow -\infty \tag{2.99}$$

$$\text{iff } n^2(1-p) \longrightarrow +\infty \tag{2.100}$$

□

And simply if we remember a previous corollary linking $\mathcal{G}_{n,p}$ and $\mathcal{G}_{n,M}$ we obtain the following corollary

Corollary 2.5.8.1. If $M = M(n) < \binom{n}{2}$ satisfies

$$\frac{2M^2}{n^3} - \log(n) \longrightarrow \infty \tag{2.101}$$

Then almost every graph in $\mathcal{G}_{n,M}$ is of diameter 2.

Proof. The first condition makes that we are not studying a complete graph then the diameter is not 1. And we have the proof simply using 2.3.2 with $pN = M$ and combining it with the previous theorem 2.5.8 we have the result. □

Chapter 3

Branching processes on random graphs

3.1 Galton-Watson trees

A branching process is the simplest model that can be used to describe the evolution of a population over time. Typically in a branching process we start with one individual, consider it will create a number of individuals through his lifetime. This number following a distribution, we call it the offspring distribution and denote it by $\{p_i\}_0^\infty$ such as

$$p_i = \mathbb{P}(\text{ having i child }) \tag{3.1}$$

When we say that a branching process is a "something" branching process, that means that the offspring distribution follows the "something" law (typically a Poisson branching process). And we also denote by Z_n the number of individuals in the n-th generation. Then if we consider that the offspring distribution doesn't depend on the generation of the individual considered we have

$$Z_n = \sum_{i=1}^{Z_{n-1}} X_{n,i} \quad , \text{ with } X = \{X_{n,i}\}_{n,i}, \text{ i.i.d.} \tag{3.2}$$

Observing this distribution, we observe that if for some generation k_0 we have $Z_{k_0} = 0$, then $Z_{k_0+k} = 0$ for any k. We would say that that the population dies out at k_0 and one might be interested to study under which condition a population will die out. It was in fact this question that was studied by Galton and Watson (TODO : CHECK AND ADD HISTORICAL DETAILS) that led to the study of branching processes. Hence, we might refer to these branching proces as Galton-Watson processes or trees (GW). We can obtain the following theorem

Theorem 3.1.1. If $\mathbb{E}X \leq 1$ then the population dies out almost surely.

Proof. TODO : ADD PROOF □

Theorem 3.1.2. If $\mathbb{E}X > 1$ then the population survives with a non-zero probability.

Proof. TODO : ADD PROOF □

We will now define an exploration process of such a branching process. We will use the model of Bollobas and Riordan (TODO : ADD REF). In this model we consider a graph with n vertices and the exploration will take n steps. For now we will consider the case where we are exploring a connected component. With this process we think of vertices in three different positions, a vertex can be active, that means the algorithm knows the existence of the vertex and is evaluating it. A vertex can be explored, in that case we can consider that the vertex has been completely evaluated and sorted, in some fashion we can forget about it. Otherwise a vertex can be unseen, meaning that we still have no idea of what it is. So in terms of set, we can consider that at time (step) t we have :

$$A_t : \text{ the set of active vertices at time } t \tag{3.3}$$
$$E_t : \text{ the set of explored vertices at time } t \tag{3.4}$$
$$U_t : \text{ the set of unseen vertices at time } t \tag{3.5}$$

The process starts as follow, at $t = 0$, no vertex has been seen yet, so $U_0 = V$ and the process has not started yet then $A_0 = \emptyset, E_0 = \emptyset$. In the case of a branching process we consider a rooted tree and we denote the root as r, so at $t = 1$, $A_1 = r, U_1 = V\backslash\{r\}, E_1 = r$. The two-steps initialisation might seem redundant but we will use it in the future when we will extend this process to study any kind of graph.
For the following steps, at time $t > 1$, the process is as follow, a vertex v_t is picked [1] at random in A_{t-1}. For convenience we will add the variable $\eta_t = |N_(v_t) \cap U_{t-1}|$ the number of vertices not yet seen that are neighbours from v_t. And then $A_t = (N(v_t) \cap U_{t-1}) \cup A_{t-1}\backslash\{v_t\}$ Finally, we move v_t to E_t and the process stops when all vertices are explored $|E_t| = \text{n}$, equivalently $t = n$, equivalently $|U_t| = 0$.
Connecting this to a Galton-Watson tree, η_t is the direct progeny of v_t so η_t follows the offspring distribution defined by $X_{1,1}$ and we consider that the population dies out only if the algorithm finishes.

$$\begin{cases} |A_0| &= 1 \\ |A_i| &= |A_{i-1}| + \eta_i - 1 = \eta_1 + ... + \eta_i - (i-1) \end{cases} \tag{3.6}$$

With all the η_i independent and identically distributed random variables. We can then define T as the instant the population dies out.

$$T := \min\{t : A_t = 0\} \tag{3.7}$$

[1] We use "picked" as follow, if an element x is picked in X_t then $x \notin X_{t+1}$.

If $T = \infty$ then we say that the population survives.
We have that $\{A_t\}_t$ is a random walk on the tree and we also get the following evolution equation

$$|S_t| = |S_{t-1}| + \eta_t - 1 \tag{3.8}$$

When we are considering the Erdos Renyi model, in the Markovian paradigm, we consider that in this random walk each vertex v has a probability p of turning active. Studying the random walk in that case we observe that the number of vertices to which we can connect the vertex we v is

$$|U_t| = n - |E_{t-1}| - |A_{t-1}| = n - (t-1) - |A_{t-1}| \tag{3.9}$$

So we have

$$\eta_t \text{ Bin}(n - (t-1) - |A_{t-1}|, p) \tag{3.10}$$

and we observe comparing it to 3.6 that our random values η_i are no longer independently distributed. However we notice that if A_{t-1} is "small enough" and n "large enough" the r.v. are almost independently distributed. We also denote that the random walk we defined explores only a connected component so intuitively we want to say that if the connected components are small enough and sparse enough, then, they follow a Galton Watson process, hence our random graph would be made of Galton Watson trees. This will be the point of this chapter, studying links between random graphs and branching process, so we will consider that our probability p has to be small enough such that there is not a single connected component. From 2.2.2 we will consider that $p = \frac{\lambda}{n}$.
As the Poisson law is more convenient to work with than the Bernoulli, and also being it's limit in distribution, let's observe through a few theorems the connecting between Poisson and Bernoulli branching process.

Theorem 3.1.3. For a branching process with a binomial offspring distribution with parameter n and p and a branching process with a Poisson offspring distribution with parameter $\lambda = np$

$$\mathbb{P}_{n,p}(T \geq k) = \mathbb{P}^*_\lambda(T^* \geq k) + e_k(n) \quad , \forall k \geq 1 \tag{3.11}$$

With T (resp. T^*) the size of the binomial (resp. Poisson) resulting branching process. And

$$|e_k(n)| \leq \frac{2\lambda^2}{n} \sum s = 1^{k-1} \mathbb{P}^*_{lambda}(T^* \geq s) \leq \frac{2\lambda^2 k}{n} \tag{3.12}$$

Proof. The proof makes use of couplings, see Annex (TODO : WRITE ON COUPLINGS IN THE ANNEX), they are a way to define distinct random values on a common probability space giving information on the intertwinning of both.

$$X_i \text{ Bin}(n, \frac{\lambda}{n}) \quad , \quad X_i^* \text{ Poi}(\lambda) \tag{3.13}$$

If we denote by $\mathbb{P}$ the joint probability distribution of X_i and X_i^*n then according to the theorem (TODO : ADD THE THEOREM IN ANNEX + PROVE IT) then

$$\mathbb{P}(X_i \neq X_i^*) \leq \frac{\lambda^2}{n} \tag{3.14}$$

Also,

$$\mathbb{P}_{n,p}(t \leq k) = \mathbb{P}(T \geq k, T^* \geq k) + \mathbb{P}(T \geq k, T^* < k) \tag{3.15}$$

$$\mathbb{P}_{\lambda}^*(t \leq k) = \mathbb{P}(T \geq k, T^* \geq k) + \mathbb{P}(T < k, T^* \geq k) \tag{3.16}$$

Which gives,

$$|\mathbb{P}_{n,p}(T \geq k) - \mathbb{P}_{\lambda}^*(T^* \geq k)| \tag{3.17}$$

$$\leq \mathbb{P}(T \geq k, T^* < k) + \mathbb{P}(T < k, T^* \geq k) \tag{3.18}$$

The following part, until the end of the proof, is valid if we exchange T by T^* and X by X^*.
By construction of T, the event $\{T \geq k\}$ is only defined by the events $X_1, ..., X_{k-1}$ Then we have $T \geq k$ and $T^* < k$ if there exists some s such as $X_s \neq X_s^*$. Hence,

$$\mathbb{P}(T \geq k, T^* < k) \leq \sum_{s=1}^{k-1} \mathbb{P}(T \geq K, X_i \neq X_i^*, \forall i \leq s-1, X_s \neq X_s^*) \tag{3.19}$$

If we are in the event,$T \geq K, X_i \neq X_i^*, \forall i \leq s-1$ then

$$X_1^* + ... + X_i^* \geq i, \forall i \leq s-1 \tag{3.20}$$

In particular,

$$X_1^* + ... + X_s^* = T^* - 1 \geq s-1 \tag{3.21}$$

Then the event $\{T^* \geq s\}$ depends only on the $X_i^*, i \leq s-1$ thus it is independent of the event $X_s \neq X_s^*$. Combining these elements we obtain,

$$\mathbb{P}(T \geq k, T^* < k) \leq \sum_{s=1}^{k-1} \mathbb{P}(T^* \geq s, X_s \neq X_s^*) \tag{3.22}$$

$$\leq \sum_{s=1}^{k-1} \mathbb{P}(T^* \geq s)(X_s \neq X_s^*) \tag{3.23}$$

$$\leq \frac{\lambda^2}{n} \sum_{s=1}^{k-1} \mathbb{P}(T^* \geq s) \tag{3.24}$$

The last inequality being obtained by the theorem on couplings (TODO : ADD REFERENCE). Using the remark on the fact that this portion of the proof is valid for both the binomial and the Poisson case we obtain the following inequality that finishes the proof.

$$e_k(n) = |\mathbb{P}_{n,p}(T \geq k) - \mathbb{P}_{\lambda}^*(T^* \geq k)| \leq \frac{2\lambda^2}{n} \sum_{s=1}^{k-1} \mathbb{P}(T^* \geq s) \tag{3.25}$$

□

The last theorem gives us some kind of "point wise" convergence between Poisson and binomial branching for trees to be larger than some fixed constant. Now we want to investigate the typical size of a connected component in $\mathcal{G}_{n,p}$. We denote the connected comported of a vertex v by $\mathcal{C}(v)$ and as a typical component we take $\mathcal{C}(1)$.

Theorem 3.1.4. For all $k \geq 1$,

$$\mathbb{P}_{n,p}(|\mathcal{C}(1)| \geq k) \leq \mathbb{P}_{n,p}(T \geq k) \tag{3.26}$$

Proof. We consider $|U_i| = n-i-|A_i|$ the number of vertice at step $t = i$ and the random variable X_i that is the offspring distribution on vertices that make the branching process stay a tree. And Y_i the converse of X_i that is the offspring distribution for elements that have already appeared in our branching process. We then have

$$X_i \text{ Bin}(|U_{i-1}|, p) \quad , \quad Y_i \text{ Bin}(n - |U_{i-1}|, p) \tag{3.27}$$

And we construct the random variable $X_i^{\geq} = X_i + Y_i$ Then we have that $X_i{\geq}$ Bin(n, p) if X_i and Y_i are independent, so it is true if we consider $X_i{\geq}$ conditionnally on $\{X_j\}_{j=1}^{i-1}$. Which gives that they are i.i.d. variables.
Also, as $Y_i \geq 0$, then $X_j{\geq} \geq X_i$. And if we denote by

$$A_i^{\geq} = X_1 + ... + X_i^{\geq} - (i-1) \tag{3.28}$$

Then,

$$\mathbb{P}_{n,p}(|\mathcal{C}(1)| \geq k) = \mathbb{P}(|A_t| > 0, \forall t \leq k-1) \leq \mathbb{P}(|A_t^{\geq}| > 0, \forall t \leq k-1) = \mathbb{P}(T \geq k) \tag{3.29}$$

□

Another similar theorem (TODO : PROVE IT CORRECTLY ??) is the following one that gives a lower bound on the size of a typical connected component although it is less useful than the previous because it is less fit for asymptotical evaluations.

Theorem 3.1.5.

$$\mathbb{P}_{n,p}(|\mathcal{C}(1)| \geq k) \geq \mathbb{P}_{n-k,p}(T \geq k) \tag{3.30}$$

Here T is the total progeny of a branching process with parameter $n-k$ and p.

Here is a sketch of the proof, the following changes shall just be plugged in the previous proof and the result will appear.

Proof. The proof is very similar to the previous one but the following random variables are used

$$Y_i \sim \text{Bin}(N_{i-1} - (n-k), p) \tag{3.31}$$

then

$$X_i = X_i^{\leq} + Y_i \tag{3.32}$$

□

Before finishing this section, here is a theorem that gives the probability law of $|A_t$ in a random graph branching process.

Theorem 3.1.6.

$$|A_t| + (t-1) \sim \text{Bin}(n-1, 1-(1-p)^t) \tag{3.33}$$

Proof. Let's first observe by symmetry that

$$X \sim \text{Bin}(m,p) \iff Y = m - X \sim \text{Bin}(m, 1-p) \tag{3.34}$$

So to prove the theorem we will prove the equivalent statement

$$n - t - |A_t| = |U_t| \sim \text{Bin}(n-1, (1-p)^t) \tag{3.35}$$

Indeed,

$$\begin{aligned} |U_t| &= n - t - |A_t| = n - t - |A_{t-1}| - |\eta_t| + 1 && (3.36)\\ &= n - (t-1) - A_{t-1} - \eta_t && (3.37)\\ &= n - (t-1) - A_{t-1} - \text{Bin}(n-(t-1) - |A_{t-1}|, p) && (3.38)\\ &= |U_{t-1}| - \text{Bin}(|U_{t-1}|, p) = \text{Bin}(|U_{t-1}|, 1-p) && (3.39) \end{aligned}$$

Simply applying a recursion on the last result gives the expected result. □

Now we will apply this result to prove the following theorem.

Theorem 3.1.7. $\mathbb{P}_\lambda(|\mathcal{C}(1)| > t) \leq e^{-I_\lambda t}$

In the previous theorem, I_λ stands for the large deviation rate function and is defined as follow.

$$I_\lambda = \lambda - 1 - \log(\lambda) \tag{3.40}$$

It is interesting to note that I_λ is positive if $\lambda \neq 0$

Proof. This proof uses the fact that $|A_t| = 0$ means that the whole connected component has been explored after t steps, so the connected component is of size less than t. Using 3.1.6 we obtain that $|A_t| \sim \text{Bin}(n-1, 1-(1-p)^t) - (t-1)$, so

$$\mathbb{P}_\lambda(|\mathcal{C}(1)| > t) \leq \mathbb{P}(|A_t| > 0) \leq \mathbb{P}_\lambda(\text{Bin}(n-1, 1-(1-p)^t) \leq t) \tag{3.41}$$

Using Bernoulli's inequality (TODO : PROVE IT SOMEWHERE) $1-(1-p)^t \leq tp$ and observing that for all s positive the following is true

$$\mathbb{P}_\lambda(e^{s\text{Bin}(n-1,tp)} \leq e^{st}) \tag{3.42}$$

Then we can apply Markov inequality which gives

$$\mathbb{P}_\lambda(|\mathcal{C}(1)| > t) \leq e^{-st}\mathbb{E}_\lambda(e^{s\text{Bin}(n, \frac{t\lambda}{n})}) \tag{3.43}$$

Replacing the moment generating function of the binomial with it's value (TODO : COMPUTE IT SOMEWHERE AND ADD REF), we obtain

$$\mathbb{P}_\lambda(|\mathcal{C}(1)| > t) \leq e^{-st}(1 - \frac{t\lambda}{n} + e^s(\frac{t\lambda}{n})^n) \leq e^{-t(s-\lambda e^s-\lambda)} \tag{3.44}$$

Using $s = \log(1/\lambda)$, which minimizes the bound, we obtain

$$\mathbb{P}_\lambda(|\mathcal{C}(1)| > t) \leq e^{-I_\lambda t} \tag{3.45}$$

□

From this theorem we can obtain that a typical component will be smaller than any (finite) composition of logaithm (TODO : CHECK THIS) on n almost surely.Now using this result we will obtain a logarithmic bound on the largest connected component.
For this we will use the random variable

$$\mathcal{Z}_{\leq k} = \sum_{v \in V} \infty |\mathcal{C}(v)| \leq k \tag{3.46}$$

Observing that $\mathcal{Z}_{\leq k}$ is equal to 0 if k is larger that the size of the greatest connected component [2] and we denote it by $\mathcal{C}_{\max}$. Hence we have

$$|\mathcal{C}_{\max}| = \max\{k : \mathbb{Z}_{\leq k} \leq k\} \tag{3.47}$$

And we obtain

$$\mathbb{E}_\lambda(\mathbb{Z}_{\leq k}) = n\mathbb{P}_\lambda(|\mathcal{C}(1)| \leq k) \tag{3.48}$$

Applying the theorem above 3.1.7 we immediately have the following result.

Lemma 3.1.8. $\mathbb{P}_\lambda(|\mathcal{C}_{\max}| \leq a \log n) \longrightarrow_{n\to\infty} 0 \quad , a > I_\lambda^{-1}$

We will now prove the next lemma that is similar to the previous one but gives an upper bound on the greatest connected component instead (in the subcritical regime).

Lemma 3.1.9. $\mathbb{P}_\lambda(|\mathcal{C}_{\max}| \leq a \log(n)) \leq \mathcal{O}(n^{-\delta})$

Proof. This prove will be a little bit more technical as it uses the second moment methods. First of all we will need an estimate of the variance on $Z_{\geq}$, for this purpose we will use the following function

$$\chi_k(\lambda) = \mathbb{E}_\lambda(|\mathcal{C}(v)|\mathbb{1}_{\{|\mathcal{C}(v)|\geq k\}}) \tag{3.49}$$

Lemma 3.1.10. $\mathbb{V}_\lambda(Z_{\geq k}) \leq n\chi_k(\lambda)$

[2] we assume unicity on the greatest connected component as asymptotically it is almost surely unique

Proof. By definition of the variance

$$\mathbb{V}_\lambda(Z_{\geq k}) = \sum_{i,j\in V} (\mathbb{P}_\lambda(|\mathcal{C}(i)| \geq k, |\mathcal{C}(j)| \geq k) - \mathbb{P}_\lambda(|\mathcal{C}(i)| \geq k)\mathbb{P}_\lambda(|\mathcal{C}(j)| \geq k) \tag{3.50}$$

And we can split those probabilities as components form an obvious partition of the vertex set as follow

$$\mathbb{P}_\lambda(|\mathcal{C}(i)| \geq k, |\mathcal{C}(j)| \geq k) = (\mathbb{P}_\lambda(|\mathcal{C}(i)| \geq k, i \leftrightarrow j) + (\mathbb{P}_\lambda(|\mathcal{C}(i)| \geq k, |\mathcal{C}(j)| \geq k, i \not\leftrightarrow j) \tag{3.51}$$

Furthermore,

$$\mathbb{P}_\lambda(|\mathcal{C}(i)| \geq k, |\mathcal{C}(j)| \geq k) = \sum_{l=k}^{n} \mathbb{P}_\lambda(|\mathcal{C}(i)| = k, |\mathcal{C}(j)| \geq k) \tag{3.52}$$

$$= \sum_{l=k}^{n} \mathbb{P}_\lambda(|\mathcal{C}(i)| = k)\mathbb{P}(|\mathcal{C}(j)| \geq k||\mathcal{C}(i)| = l) \tag{3.53}$$

And simply observing that in $\mathcal{G}_{n,p}$, the event $\{|\mathcal{C}(j)| \geq k\}$ naturally increases with n, thus decreases if l vertices are removed, we have in the event where i and j are not connected.

$$\mathbb{P}_\lambda(|\mathcal{C}(i)| \geq k, |\mathcal{C}(j)| \geq k) \leq \sum_{l=k}^{n} \mathbb{P}_\lambda(|\mathcal{C}(i)| = k)\mathbb{P}(|\mathcal{C}(j)| \geq k) \tag{3.54}$$

$$\leq \mathbb{P}(|\mathcal{C}(j)| \geq k) \sum_{l=k}^{n} \mathbb{P}_\lambda(|\mathcal{C}(i)| = k) \tag{3.55}$$

$$\leq \mathbb{P}(|\mathcal{C}(j)| \geq k)\mathbb{P}_\lambda(|\mathcal{C}(i)| \geq k) \tag{3.56}$$

And in the case where i and j are part of the same connected component then

$$\mathbb{P}_\lambda(|\mathcal{C}(i)| \geq k, |\mathcal{C}(j)| \geq k) = \mathbb{P}_\lambda(|\mathcal{C}(i)| \geq k) \tag{3.57}$$

So, if we denote by Γ the partition in connected components of V. Then we have the following

$$\mathbb{Z}_{\geq k} = \sum_{C\in\Gamma} \sum_{i\in C} (\sum_{j\in C} (\mathbb{P}_\lambda(|\mathcal{C}(i)| \geq k, |\mathcal{C}(j)| \geq k) - \mathbb{P}_\lambda(|\mathcal{C}(i)| \geq k)\mathbb{P}_\lambda(|\mathcal{C}(j)| \geq k)) \tag{3.58}$$

$$+ \sum_{j\notin C} (\mathbb{P}_\lambda(|\mathcal{C}(i)| \geq k, |\mathcal{C}(j)| \geq k))) - \mathbb{P}_\lambda(|\mathcal{C}(i)| \geq k)\mathbb{P}_\lambda(|\mathcal{C}(j)| \geq k)) \tag{3.59}$$

$$\leq \sum_{C\in\Gamma} \sum_{i\in C} (\sum_{j\in C} \mathbb{P}_\lambda(|\mathcal{C}(i)| \geq k) - \mathbb{P}_\lambda(|\mathcal{C}(i)| \geq k)^2 \tag{3.60}$$

$$\leq \sum_{C\in\Gamma} (\mathbb{P}_\lambda(|C| \geq k) - \mathbb{P}_\lambda(|C| \geq k)^2) \sum_{i\in C} \sum_{j\in C} 1 \tag{3.61}$$

$$\leq \sum_{C\in\Gamma} (\mathbb{P}_\lambda(|C| \geq k) - \mathbb{P}_\lambda(|C| \geq k)^2)|C|^2 \tag{3.62}$$

□

□

3.2 The exploration process, Karp's new approach

3.3 The subcritical case : $\lambda < 1$

3.4 The supercritical case : $\lambda > 1$

3.5 Some words on the critical case

Chapter 4

The configuration model

4.1 Random regular graphs

4.2 An arbitrary degree sequence - the Newman-Watts-Strogatz model

Appendix A

Some probabilistic tools

A.1 Common inequalities

A.2 Tail inequalities

A.3 Markov chains

A.4 Martingales