30

83670

UMI

MICROFILMED 2003

A

Thresholds in Random Graphs

by

Ariel Halpert

A dissertation submitted to the Graduate Faculty in Mathematics in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2003

UMI Number: 3083670

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

This manuscript has been read and accepted for the Graduate Faculty in Mathematics in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

May 1st, 2003
Date

Pach János
Chair of Examining Committee
Prof. János Pach
City University of New York

May 1st 2003
Date

Alvany Rocha
Executive Officer
Prof. Alvany Rocha

Prof. Amotz Bar–Noy
City University of New York

Prof. Joseph Malkevitch
City University of New York

Prof. Richard Pollack
New York University

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

Thresholds in Random Graphs

By: Ariel Halpert

Advisor: Prof. János Pach

This thesis follows the work of Erdős and Rényi, on the evolution of random graphs. We consider several models of random graphs and present results relating to the asymptotic existence of certain invariant graph theoretic properties, such as connectedness and planarity.

In Chapter 1 we describe a dynamic load sharing algorithm for a cellular telephone network with n transceiver towers. We model this network via a hypergraph and show how a random sequence of calls within the network can be modelled by a random bipartite graph. We show that an admissible sequence, i.e. a sequence of calls, each of which is accepted by the network, is equivalent to the existence of a matching of the random bipartite graph and find thresholds on the number of calls for admissibility.

In Chapter 2 we study the $\mathcal{G}_{k\text{-out}}$ model for random graphs. In this model, each vertex randomly connects "out" to k other vertices. We find connectivity, planarity and maximal degree properties of the model. We also extend the model to non-integer values of k and find connectivity and planarity properties for this model as well.

In Chapter 3 we use combinatorial techniques for counting the number of connected random graphs in $\mathcal{G}_{k\text{-out}}$ to solve number theoretic problems related to determining the asymptotics of $\sum w_{t_1} w_{t_2} \cdots w_{t_k}$ where $w = (w_1, w_2, w_3, \cdots)$ is a sequence of numbers called weights and the sum is taken over all k-part compositions $t_1 + t_2 + \cdots + t_k$ of n. We note that although this problem has been recently solved using methods of complex analysis, our combinatorial approach is of interest due to its relative simplicity.

Acknowledgements

I would like to thank my advisor, Professor János Pach for allowing me the opportunity to work with him on the research presented here. His incisively clear thinking and ability to discern the core mathematical issues in a problem were invaluable in developing this thesis.

Many others were instrumental in my research. Most notably Noson Yanofsky, Florian Lengyel and Dan Ismailescu. Chapter One stems from work I did with Florian, with whom I have a coauthored paper. Noson, who has been and will continue to be a lifelong friend, has read through and commented on most of Chapter Two. His comments and our many discussions have led to a better understanding of the mathematics and to several results. Much of the form of Chapter Three is the result of many discussions and collaborative research with Dan.

I attribute the ongoing motivation for the completion of this effort to my parents, Yoram and Sara Halpert. Indeed my debt to them goes well beyond this paper. In addition to acting as a sounding board for several of the mathematical issues I had been thinking about, my father's unyielding encouragement and desire to see me through to completion served as a constant source of inspiration for the research. They have been for me, boundless sources of love, strength and support. An infinite search would not find more dedicated parents. I have been blessed and am forever grateful.

To my wife Devora Halpert, I offer heartfelt thanks and deepest appreciation, for bringing sunshine into my life and for bringing to this world a thing of absolute beauty and perfection greater than any mathematical proof, our daughter, Aliza Faigy Halpert. For the former I am constantly grateful, for the latter I am constantly awed. Without her by my side, my current and future accomplishments, indeed my life would be void of all meaning.

Contents

0 Introduction **1**

1 Cellular Telephone Networks and Random Bipartite Graphs **4**

1.1 Introduction 4

1.1.1 Non-Idealized Systems 5

1.1.2 Random Graphs 7

1.1.3 The Idealized Cellular Telephone Network Problem . . 8

1.2 Polygonal Models 9

1.2.1 The One Dimensional Model 10

1.2.2 The Two Dimensional Model 12

1.3 The Hypergraph Model 16

1.4 The Probabilistic Method 20

1.5 A Lower Threshold for Admissibility 21

1.5.1 An Upper Bound For $E[Y_j]$ 25

1.5.2 A Lower Bound for $E[Y_j]$ 27

1.5.3 Proof of the Expectation Threshold Function 28

1.6 Applying Martingale Techniques 30

1.6.1 An Upper Threshold for Rejection 30

1.6.2 Closing the Gap between $n^{\frac{r}{r+1}}$ and $n^{\frac{r}{r+1/2}}$ 34

1.7 A Limiting Probability Distribution 35

2 The $\mathcal{G}_{k\text{-out}}$ Model for Random Graphs **39**

2.1 Introduction 39

2.2 Defining the Model 41

2.3 On the Connectedness of $\mathcal{G}_{k\text{-out}}$ 42

2.4 An Extension of $\mathcal{G}_{k\text{-out}}$ and the Poisson Paradigm 49

2.5 On Higher Connectivity of $\mathcal{G}_{k\text{-out}}$ 62

2.6 On the Degree Sequence Distribution of $\mathcal{G}_{k\text{-out}}$ 65

2.7 On the Planarity of $\mathcal{G}_{k\text{-out}}$ 69

3 An Application of Graph Theory to Number Theory **78**

3.1 Introduction 79

3.2 Counting Connected Graphs 84

3.3 Proof for $p = 3/2$ 85

3.4 Proof for large enough powers 88

3.5 Proof for all powers $p \geq 3/2$ 88

Bibliography **90**

List of Figures

1.1 Section of a One Dimensional System 10

1.2 A Set of Switches on the System 11

1.3 Saturated Cells and a Rejected Call on the Linear System . . 11

1.4 Section of a Two Dimensional Regular Hexagonal System . . . 13

1.5 Saturated Cells and a Rejected Call on the Hexagonal System 14

1.6 Subsets of Polygonal Systems that do not have Property $\mathcal{P}$ . . 15

1.7 An Admissible Call Sequence on the Hypergraph Model 18

1.8 A Matched Bipartite Graph of the Sequence in Figure 1.7 . . . 19

1.9 A Rooted 6–Cluster with 2 of Several Associated Trees 23

2.1 A Section from a random graph $G \in \mathcal{G}_{3\text{-out}}$ 42

2.2 A connected graph from $\mathcal{G}_{1\text{-out}}$ 44

2.3 A component of order 4 from the $\mathcal{G}_{2\text{-out}}$ model 46

2.4 Isolated edges of the $\mathcal{G}_{1\text{-out}}$ model 52

2.5 A 3–set of isolated edges 54

2.6 A 3–set of edges with one non-isolated edge 55

2.7 Forbidden subgraphs 71

2.8 Pruning T at vertex v 74

3.1 A Connected Graph from $\mathcal{G}_{1\text{-out}}$ 84

Chapter 0

Introduction

This thesis follows the work of Erdős and Rényi of half a century ago, on the evolution of random graphs. In their work they considered among other models, the random graph model $\Gamma_{n,N}$. In this model each graph with n labelled vertices and N edges is equally likely. In the limit as $n \to \infty$, they found that the probabilities of a random graph having certain fundamental structural properties, properties that remain invariant with a graph isomorphism, suddenly changed from 0 to 1, as the number of edges reached a critical function of n. They termed such a function a threshold function. For example, the threshold for the number of edges required so that there is a.s. a triangle in a random $G \in \Gamma_{n,N}$, is n and the threshold for connectedness is $n \log n$. We consider several models of random graphs and present results relating to the asymptotic existence of certain invariant graph theoretic properties, such as connectedness, planarity and a matching within these models.

In Chapter 1 we describe a dynamic load sharing algorithm for a cellular telephone network with n transceiver towers. We model this network via a hypergraph and show how a random sequence of calls within the network can be modelled by a random bipartite graph. We show that an admissible sequence, i.e. a sequence of calls, each of which is accepted by the network, is equivalent to the existence of a matching of the random bipartite graph. We then prove that the expected number of rejected calls $E[\overline{X}_m]$, in a sequence of m calls has the threshold function

$$\lim_{n\to\infty} E[\overline{X}_m] = \begin{cases} 0 & \text{if } m \ll n^{\frac{r}{r+1}}, \\ \infty & \text{if } m \gg n^{\frac{r}{r+1}}. \end{cases}$$

We use martingale techniques to show that random call sequences of length $m \ll O(n^{\frac{r}{r+1}})$ will a.s. have a matching, i.e. be admissible and that sequences of length $m \gg O(n^{\frac{r}{r+1/2}})$ will a.s. not be admissible. Here r represents the uniform number of transceivers available within each calling region. In the case $r = 2$, we find a probability distribution for admissibility.

In Chapter 2 we study the $\mathcal{G}_{k\text{-out}}$ model for random graphs. In this model, each vertex randomly connects "out" to k other vertices. We find connectivity, planarity and maximal degree properties of the model. We also extend the model to non-integer values of k and find structural properties for this model as well. We use the "Poisson Paradigm" to show that the number of isolated vertices in $\mathcal{G}_{1\text{-out}}$ is Poisson with mean $e^{-2}/2$. From there we prove that the the model $\mathcal{G}_{2-\epsilon\text{-out}}$ is disconnected with positive probability for $\epsilon > 0$. In contrast, if $\epsilon = 0$ then the graph is a.s. connected. We also prove that for constant k the maximum degree Δ of $\mathcal{G}_{k\text{-out}}$ is a.s.

$$\Delta = \Theta\left(\frac{\log n}{\log\log n}\right).$$

From there we prove that $\mathcal{G}_{1+\epsilon\text{-out}}$ is a.s. non-planar for $\epsilon > 0$. In contrast if $\epsilon = 0$ then the graph is always planar.

In Chapter 3 we use combinatorial techniques for counting the number of connected random graphs in $\mathcal{G}_{k\text{-out}}$ to solve number theoretic problems related to determining the asymptotics of $\sum_{\gamma_k(n)} w_{t_1} w_{t_2} \cdots w_{t_k}$, where $w = (w_1, w_2, w_3, \cdots)$ is a sequence of numbers called weights and the sum is taken over all k-part compositions $t_1 + t_2 + \cdots + t_k$ of n. In particular, we prove that if $p \geq 3/2$ and n, k are positive integers so that $k \leq n-1$, then

$$\sum_{\gamma_k(n)} \frac{1}{(t_1 t_2 \cdots t_k)^p} = \sum_{\substack{t_1+t_2+\cdots+t_k=n \\ t_j \geq 1,\ 1 \leq j \leq k}} \frac{1}{(t_1 t_2 \cdots t_k)^p} \leq \frac{k\, a^k}{\sqrt{2\pi}} (n-k)^{-p},$$

where $a = e^{1/12}\sqrt{2\pi}$. We note that although this problem has been recently solved using sophisticated methods of complex analysis, our combinatorial approach is of interest due to its relative simplicity.

Chapter 1

Cellular Telephone Networks and Random Bipartite Graphs

1.1 Introduction

In the first part of this dissertation we study the call acceptance behavior of an idealized cellular telephone network. Given a large network with n calling regions and a sequence of m calls placed randomly to regions of the network we find thresholds for m so that all calls of the sequence are accepted. We do this by modelling the network via a hypergraph and then consider a bipartite graph with the modelled hypergraph as one component and the sequence of calls as the other. We then show that a perfect matching of this bipartite graph is equivalent to the acceptance of every call in the sequence and find conditions under which the random bipartite graph has a perfect matching.

1.1.1 Non-Idealized Systems

At the second International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications in 1998, Matula, Iridon and Yang presented a paper [29] which motivated the problems that we have analyzed. In our research we idealized the system in their paper and consider only the combinatorial elements of their work. We thus ignored any physical and engineering problems present in the design and implementation of cellular telephone networks. We briefly mention here some of the physical and engineering issues that are present in the real world system invented by Matula and Yang [40] in order to motivate our research.

The system of Matula and Yang allows for callers within a calling region to connect to a set of transceivers. A single transceiver can accommodate several calls by several resource sharing methods. Typically, each call will communicate with a transceiver within a narrow frequency band. Several calls can be directed to the same transceiver each at different frequencies. This method is known as *Frequency Division Multi Access* or FDM. *Time Division Multi Access* or TDM is a method by which calls can communicate with a transceiver at the same frequency but at discretized time intervals, each call communicating with the transceiver for a an integer number of milliseconds at a time after which a different call is allowed access to the transceiver. A combination of FDM and TDM known as *Code Division Multi Access* or CDM is now most often used in cellular telephone networks.

Each of the three methods FDM, TDM and CDM have relative advantages and disadvantages which in practice would affect the capacity of a system to accept a random sequence of calls. In our analysis we assume that each transceiver has a finite load capacity c of calls that it can handle at any point in time. We do not consider the method by which a call is transmitted to the network and assume the only issue in connecting a call to a transceiver is that of finding a matching of our modelled bipartite graph. In reality issues such as limited frequency bandwidth, transmission interference, topography, weather and transmission power, among others will affect the performance of the system[1].

The main result in the Matula and Yang paper is the Capacity-Demand Theorem which shows that an assignment of calls to transceivers exists under certain conditions similar to those of the König-Hall Lemma. Finding such a matching could be accomplished only via a centralized decision algorithm. Such an algorithm would have information on the location of every call and the load of every transceiver. It would then make the necessary switches of calls from one transceiver to another in order to accommodate all calls. In practice switching calls creates a load on the power and computing resources of the system and decreases reliability and thus is to be minimized. Additionally, a centralized algorithm would take too long to find a perfect matching and thus local algorithms are needed. Our results represent a theoretical maximum for the call acceptance performance of the system for a centralized

[1]My thanks to Prof. Amotz Bar-Noy for our discussion on the physical and engineering issues mentioned here.

algorithm.

1.1.2 Random Graphs

We now turn our attention to the study of random graphs, which we use as an idealized model of the cellular telephone network. The study of random graphs was pioneered by Paul Erdős and Alfred Rényi with an investigation of the connectivity of random graphs which resulted in a series of seminal papers published from 1959 - 1961 [12], [13] and [14]. In these papers they analyzed connectivity properties of a random graph process in which an edge is randomly added to a graph of large order n at each point in the process. They considered two related models, $\mathcal{G}_{n,p}$ and $\mathcal{G}_{n,M}$. In the $\mathcal{G}_{n,p}$ model each vertex pair is given an edge with probability $p = p(n)$ and in the $\mathcal{G}_{n,M}$ model the process randomly adds edges from the remaining vertex pairs without edges until there are M edges in the graph. Among other things, they found thresholds on M and p for *almost sure* (a.s.) connectivity. In the next chapter we will discuss connectivity thresholds for a third model of random graph.

From 1964 - 1968 Erdős and Rényi published a second set of papers finding among other things, thresholds for the a.s. existence of perfect matchings in random graphs [15], [16] and [17]. A perfect matching in a graph $G = (V, E)$ is a subgraph of G that covers V and consists of disjoint edges. They show that the threshold for minimum vertex degree $\delta = 1$ is the same threshold as that for a perfect matching. In particular if $G \in \mathcal{G}_{n,n,p}$ is random graph from a random bipartite graph space with n vertices in each class and probability

p for an edge between any two vertices in different classes, they showed that

$$\Pr[G \text{ has a perfect matching}] = \begin{cases} 1 & \text{if } np - \ln n \to \infty, \\ e^{-2e^c} & \text{if } np - \ln n \to c, \\ 0 & \text{if } np - \ln n \to -\infty. \end{cases}$$

This is an example of the "minimum degree phenomenon" [31], which states that if a necessary condition for a random graph to have Property $\mathcal{P}$ is a minimum degree of δ then the threshold for Property $\mathcal{P}$ is the same as the threshold for a minimum degree of δ.

A similar problem to the one considered in our research is contained in the paper of Gimbal and Palmer [24]. They consider a random bipartite graph model $\mathcal{B}_{n,r}$ with vertex sets U and V of size n, and in which each vertex u in U independently chooses a set of r vertices in V as neighbors. They find conditions on r so that a random graph from $\mathcal{B}_{n,r}$ has a.s. a perfect matching.

1.1.3 The Idealized Cellular Telephone Network Problem

We now turn to our problem, which is an idealization of the cellular telephone network that considers only the graph theoretic elements of the physical system. The network consists of n calling regions, each region being associated with a fixed number of transceivers s, and each transceiver being capable of servicing at most a constant number of calls c, at any one point in time.

We consider a random sequence of calls placed within the calling regions of the network. The network is equipped with the load sharing algorithm of Matula and Yang, which finds, if possible, a set of switches of calls. Each call may be switched from its current transceiver to a neighboring transceiver, with the intent of freeing a transceiver somewhere in the network to accept a subsequent call. As previously implied the algorithm finds a matching of calls within regions to neighboring transceivers, if in fact such a matching exists.

For the purposes of our analysis we assume that each call, once placed within a particular calling region remains immobile and additionally, does not terminate. Thus we consider essentially the state of the network at a single point in time.

Given such a network and algorithm or system, one may inquire as to the length of admissible sequences as a function of n, the number of calling regions within the system. In our analysis we find threshold functions for the lengths of admissible sequences. Additionally, Flórián Lengyel [26] has also found a limiting probability distribution for the linear system which is described in the next section.

1.2 Polygonal Models

Our first step in this analysis is to model the system in a most direct manner. We do this via a polygonal lattice-like tiling. By *lattice-like* we mean that vertices of the tiling polygons meet other polygons at vertices and not at edges. Subsequently we model the system in a graph theoretic setting, realizing through this process that the geometry plays no role in the system's call sequence acceptance behavior. Rather, it is the incidence relationships of regions and their associated transceivers that plays the pivotal role. In fact, our results are independent of the dimension of the system. Despite this statement we begin with a description of the simplest geometric model.

1.2.1 The One Dimensional Model

We model the linear system as a segment of length n over $\mathbb{R}$. More precisely, the system is represented by the interval $[0, n]$ with integer points representing transceiver towers and the segments between towers representing calling regions or *cells*. Figure 1.1 depicts a section of a linear system. Calls placed to cells of the system, represented by x's, may be routed through either the right or left transceiver tower of that cell, as indicated by the arrows from the x's. Cells are labelled $i - 1$ through $i + 4$.

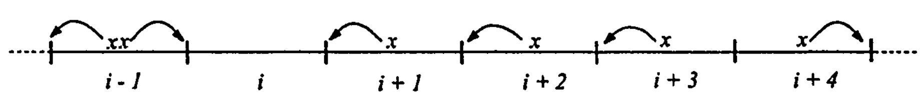


Figure 1.1: Section of a One Dimensional System

In our current example each tower may serve at most one call. Thus a set of switches is necessary to accommodate a new call, represented by a y, placed in cell i. Broken line arrows in Figure 1.2 indicate switched calls.

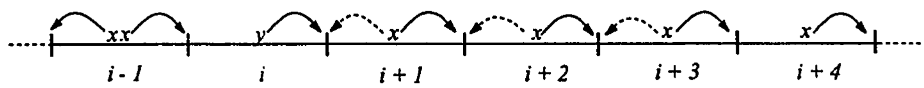


Figure 1.2: A Set of Switches on the System

Definition 1.1. *A cell is said to be* saturated *if no set of switches of calls to neighboring transceivers exists within the network to connect a new call placed within that cell to a neighboring transceiver.*

Definition 1.2. *A call placed within a saturated cell is said to be* rejected.

It is clear that cells $i-1, i, \ldots, i+4$ in Figure 1.2 are saturated. Figure 1.3 depicts a rejected call, represented by a z with no arrow, placed into cell i of the saturated set of Figure 1.2. There is no set of switches that will free a transceiver for call z.

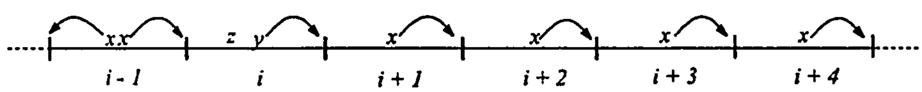


Figure 1.3: Saturated Cells and a Rejected Call on the Linear System

Definition 1.3. *A call that is not rejected it is said to be* accepted.

Definition 1.4. *A sequence of calls is said to be* admissible *if every call of the sequence is accepted.*

Thus, a sequence for which there exists a matching of calls to transceiver towers is an admissible sequence. After discussing the hypergraph model the following will become clear.

Remark 1.5. *Admissibility of a call sequence is invariant with respect to permutation of the sequence.*

1.2.2 The Two Dimensional Model

Matula and Yang's patent illustrates an implementation of their algorithm over an hexagonal lattice of the plane. Figure 1.4 depicts a section of an hexagonal system with a set of connected calls. Vertices represent transceiver towers, while the interiors of the hexagons represent cells. In the figure, calls are represented by x's and connections to vertices are represented by arrows from the x's. Cells are labelled A through G. A call placed within a cell may connect only to an adjacent vertex, i.e., a vertex of the hexagon in which it was placed.

In Figure 1.5 we add a new call y to cell B, this causes calls in cells E and F to be switched. These switches are again represented by broken line arrows. Addition of the call y also causes cells B, C and E to become saturated. We indicate this by shading those cells. Finally, a rejected call z with no associated arrow is shown placed to cell E.

In the hexagonal case each cell is adjacent to and thus served by $s = 6$ transceiver towers or vertices. We will need to impose a technical condition

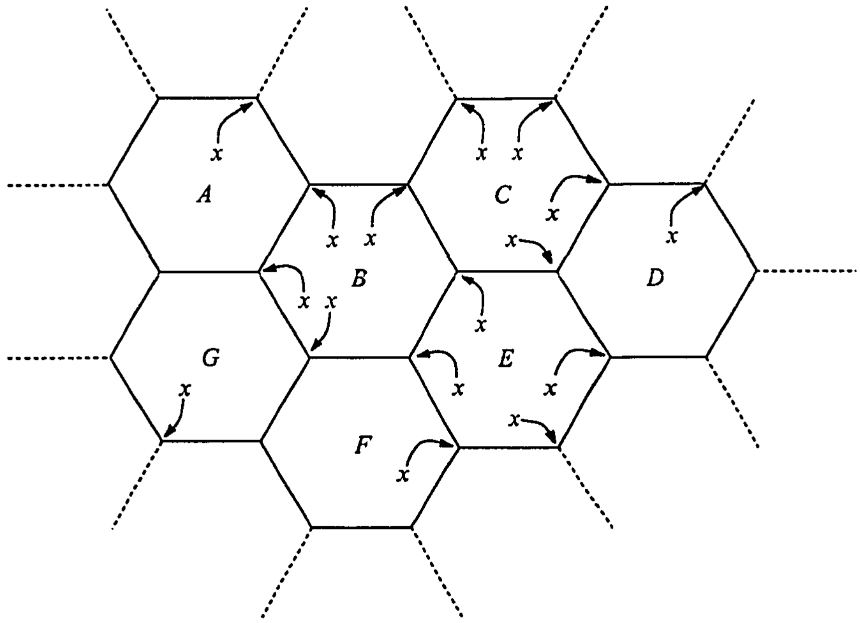


Figure 1.4: Section of a Two Dimensional Regular Hexagonal System

on the structure of the tiling or lattice of the system. We will require that all systems under consideration, have the Property $\mathcal{P}$ which we define below for polygonal models of a system.

Definition 1.6. *We will say that a polygonal system of s-gons has Property $\mathcal{P}$ if for every non-empty subset A of cells we have*

$$|A| + s - 1 \leq |\cup A|,$$

where $\cup A$ is the set of all vertices associated with cells in A.

A simple check will show that a one-dimensional system whose end points

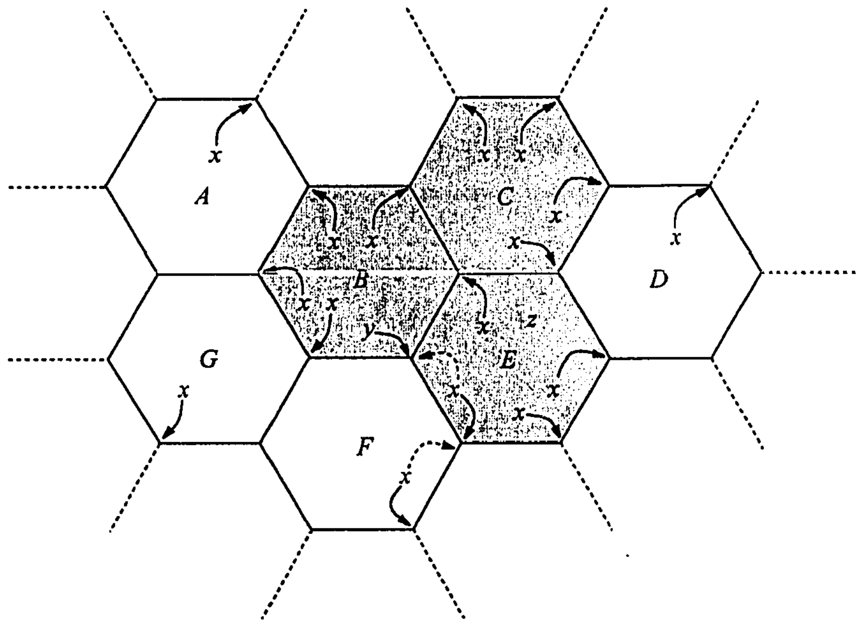


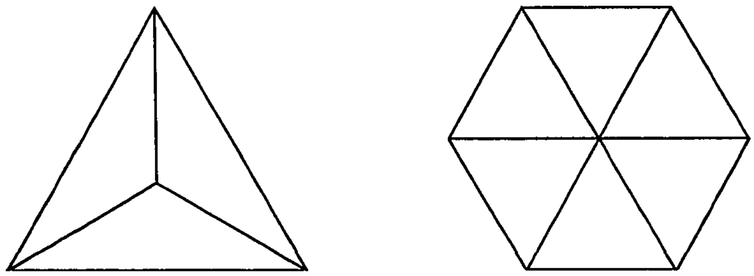
Figure 1.5: Saturated Cells and a Rejected Call on the Hexagonal System

are not identified has Property $\mathcal{P}$. In two dimensions however it is not true that every polygonal system has Property $\mathcal{P}$. Figure 1.6 shows subsets of two polygonal systems which do not have Property $\mathcal{P}$. For example, Figure 1.6a[2] has $|A| = 3$ cells, $s = 3$ and $|\cup A| = 4$ vertices in all. Thus, in this case $|A| + s - 1 \nleq |\cup A|$.

The following proposition gives a class of polygonal systems having Property $\mathcal{P}$.

Proposition 1.7. *Let $\mathcal{L}$ be a lattice-like tiling of a disk in the plane via congruent translates of a (convex) polygon Q, then $\mathcal{L}$ has Property $\mathcal{P}$.*

[2]This illustrative example was bought to my attention by Prof. Alex Heller.



a b

Figure 1.6: Subsets of Polygonal Systems that do not have Property $\mathcal{P}$

Proof. For $s \geq 3$ let Q be an s–gon. Furthermore, let $\emptyset \neq A \subset \mathcal{L}$. Fix a vertex of Q to be associated with each copy of Q in $\mathcal{L}$. Since the copies of Q are translates, each cell in A has a distinct associated vertex. Choose $Q' \in \partial A$ so that Q' has an unassociated vertex. Such a Q' must exist since all associated vertices have the same orientation.

We must show the existence of at least $s-1$ unassociated vertices of $\cup A$. This can be seen by induction on $|A|$. For $|A| = 1$ there is 1 associated vertex and exactly $s-1$ unassociated vertices. Now consider the set $A - Q'$, this has cardinality $|A-1|$ and thus by the inductive hypothesis has $s-1$ unassociated vertices. Addition of Q' to this set will add at least one unassociated vertex, by our choice of Q'. On the other hand it will associate only 1 previously unassociated vertex. There remain at least $s-1$ unassociated vertices after the addition of Q'. $\square$

1.3 The Hypergraph Model

We begin with the definition of a hypergraph and an r-uniform hypergraph.

Definition 1.8. *A* hypergraph $\mathcal{H}$ *is a pair* (V, E) *of disjoint sets where the elements of* E *consist of nonempty subsets of* V.

Definition 1.9. *An* r-uniform hypergraph *is a hypergraph where the each element of* E *has cardinality precisely* r.

Remark 1.10. *A* 2*-uniform hypergraph is a graph.*

The hypergraph model of the system of Matula and Yang is a *connected* r-uniform hypergraph $\mathcal{H} = (V, E)$ with $|E| = n$ and such that each vertex is adjacent to at most d hyperedges or cells. In the sequel we refer to the hyperedges of $\mathcal{H}$ as cells to emphasize the analogy with the polygonal model. Additionally, as in the polygonal model, we require that $\mathcal{H}$ have Property $\mathcal{P}$ given in Definition 1.7 with r replacing s.

In this model, each vertex is capable of accepting at most one call. We begin to establish equivalence with the polygonal model by setting $r = cs$ where c is the number of calls a transceiver tower in the polygonal model can accept and s in the uniform number of vertices in each polygon.

Denote by $[m]$ the set $\{1, 2, \ldots, m\}$. A sequence of calls to the cells of the system is represented by a function $f : [m] \to E$. The elements of $[m]$ represent calls. f assigns each call $j \in [m]$ to the cell $f(j) \in E$, where E is the edge set of $\mathcal{H}$.

Given $\mathcal{H}$ and f a bipartite graph $\Gamma_f = (V_f, E_f)$ can be assigned as follows. We define and bipartition the vertex set of Γ_f as

$$V_f = [m] \bigcup V,$$

where V is the vertex set of $\mathcal{H}$. The edge set of Γ_f is defined as

$$E_f = \{(j, v) \mid j \in [m] \text{ and } v \in f(j) \subseteq V\}.$$

In analogy with the polygonal model the edge set of Γ_f lists all possible connections between a call $j \in [m]$ and the vertices of the cell $f(j)$, in which it was placed.

Remark 1.11. *All vertices in the call set $[m]$ of Γ_f have uniform degree r.*

Figure 1.7 shows a bipartite graph Γ_f for the admissible function $f[5] \to E$ whose call sequence is given by $\{2, 1, 1, 3, 4\}$ on a system with 4 cells. Note that both vertices and cells of $\mathcal{H}$ are labelled in the figure. For clarity only $f(1)$ and $f(3)$ are drawn with connections to $\mathcal{H}$. Also for clarity, the cells of $\mathcal{H}$ are labelled with circled numbers. For example $f(3) = 1$ means that the third call is placed within cell number 1 and by Figure 1.7 has as neighboring vertices, those numbered $5, 6, 8$ and 9. Thus we draw 4 connections between vertex $3 \in [m]$ and vertices $\{5, 6, 8, 9\} \subseteq V$. In Figure 1.8 we eliminate the geometry and draw the same bipartite graph as in Figure 1.7. Notice the vertices of $[m]$ have uniform degree 4. We now give a definition and widely

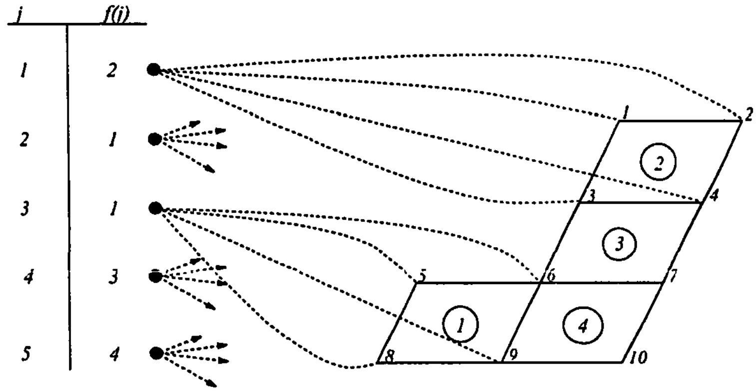


Figure 1.7: An Admissible Call Sequence on the Hypergraph Model

used theorem that we will use as well.

Definition 1.12. *Let $G = (V, E)$ be a bipartite graph with bipartition $\{A, B\}$. A* matching *in G from A to B is a set of independent edges of G that cover A.*

The bipartite graph in Figure 1.8 has a matching, the edges of which are in bold.

Theorem 1.13. *(König-Hall 1935) A bipartite graph G contains a matching of A iff*

$$|S| \leq |N(S)| \ \forall \ S \subseteq A,$$

where $N(S)$ is the set of vertices in B, each adjacent to some vertex in S.

We refer the reader to books on Graph Theory by Bollobás [5] and Diestel [10] for several proofs of this theorem.

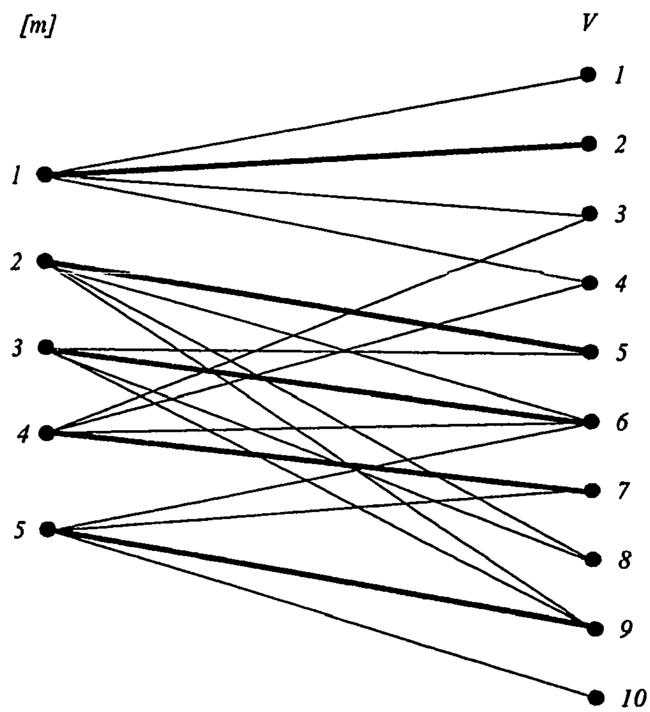


Figure 1.8: A Matched Bipartite Graph of the Sequence in Figure 1.7

We have previously defined a sequence to be admissible in the polygonal model if each of it's calls was accepted. That is, each call in the sequence was connected to a vertex to which no other call in the sequence was connected. This however translates into the existence of a matching in Γ_f between the sequence of calls $[m]$ and the set of vertices V, of $\mathcal{H}$. Figures 1.7 and 1.8 illustrate this by displaying a matching for the given admissible call sequence f.

By the König-Hall theorem we may say that a sequence f is *admissible* iff for every $S \subseteq [m]$ we have $|S| \leq |N(S)|$. This is precisely the statement that no subsequence of calls is placed into a set of cells with fewer vertices than calls. Such a subsequence will certainly have a rejected call under the polygonal model. Thus, we have equivalence of the two models in the sense that a call sequence is admissible in one model iff it is also admissible in the other. In Chapter 2 we will be concerned with finding limiting probabilities that a random uniformly distributed call sequence placed into the n cells of the hypergraph model will be admissible.

The definitions of saturated cell and rejected call carry over naturally to the hypergraph model. Given a finite function $f : [m] \to E$ we may define the *admissible domain* A_f of $[m]$ as the lexicographically smallest subset $A_f \subseteq [m]$ of maximal cardinality such that f restricted to A_f is admissible. A call in $[m]$ is *rejected* if it is not in A_f. A call $j \in [m]$ *saturates* the cell $e = f(j)$ iff $j \in A_f$ and for any call $k > j$ is rejected if $f(k) = e$.

1.4 The Probabilistic Method

In the sequel we state and prove our results related to the cellular telephone switching system of Matula and Yang. We use the probabilistic method in [1] to find upper and lower threshold functions for the asymptotic probability of admissibility of a random sequence of calls f, uniformly distributed to the cells of the hypergraph model $\mathcal{H}$. Additionally, we state a limiting probability distribution of Lengyel [26] for the one dimensional case where $r = 2$ and use it as a basis for a conjecture in the general case.

We use the notation $f \ll g$, $f \gg g$ and $f \sim g$ to indicate the limiting ratio $\lim_{n\to\infty} \frac{f(n)}{g(n)}$ tends to 0, ∞ and 1 respectively. We will also use the notation $f = \Theta(g)$ if $f = O(g)$ and $g = O(f)$. Lastly, we say that $s(n)$ is a *lower threshold function* for a Property $\mathcal{Q}$ of a model of the telephone switching system on n cells if $\Pr[\mathcal{Q}] \to 0$ when the number of calls $m(n) \ll t(n)$. Similarly $t(n)$ is an *upper threshold function* for $\mathcal{Q}$ if $\Pr[\mathcal{Q}] \to 1$ when $m(n) \gg t(n)$.

1.5 A Lower Threshold for Admissibility

Let $\mathcal{H} = (V, E)$ be an r-uniform hypergraph with $|E| = n$ cells such that each cell intersects at most d other cells, with r and d fixed. Furthermore, assume $\mathcal{H}$ has Property $\mathcal{P}$ of Definition 1.7. We consider the uniformly distributed probability space of all sequences $f : [m] \to E$, where $n \to \infty$ and $m = m(n) \to \infty$.

Definition 1.14. *Let R_k be the event that the k-th call of the sequence f is rejected. Furthermore let X_k be the indicator random variable for R_k and let*

$$\overline{X}_k = \sum_{j=1}^{k} X_j.$$

Then we have the following theorem, which gives a lower threshold function on the expected number of rejections in a random call sequence $f : [m] \to E$.

Theorem 1.15.

$$\lim_{n\to\infty} E[\overline{X}_m] = \begin{cases} 0 & \textit{if } m \ll n^{\frac{r}{r+1}}, \\ \infty & \textit{if } m \gg n^{\frac{r}{r+1}}. \end{cases} \tag{1.1}$$

We will need a few more definitions and lemmas before proceeding to the proof of Theorem 1.15.

Definition 1.16. *Let Y_j be the random variable that is the number of previously not saturated cells of $\mathcal{H}$ that become saturated precisely at the j-th call. Let*

$$\overline{Y}_k = \sum_{j=1}^{k} Y_j,$$

the number of saturated cells of $\mathcal{H}$ after k calls.

Remark 1.17. *If the j-th call is rejected then $Y_j = 0$. Additionally, $Y_j = 0$ for all $j < r$ since at least r calls are needed to saturate even a single cell.*

Definition 1.18. *A u–*cluster *is a connected set U of u cells of E. A* rooted *u–*cluster *is a pair (U, e) consisting of a u–cluster U and a root cell $e \in U$.*

Lemma 1.19. *For any $1 \le u \le n$, the number of u–clusters in $\mathcal{H}$ is bounded above by*

$$\frac{n}{u} \cdot \frac{(3d)^u}{(d-1)u+1}.$$

Proof. As depicted in Figure 1.9 for each rooted u–cluster there is at least one rooted d–ary tree with u vertices. Simply take the graph induced by the u–cluster, vertices of the graph are cells of the u–cluster and edges of the

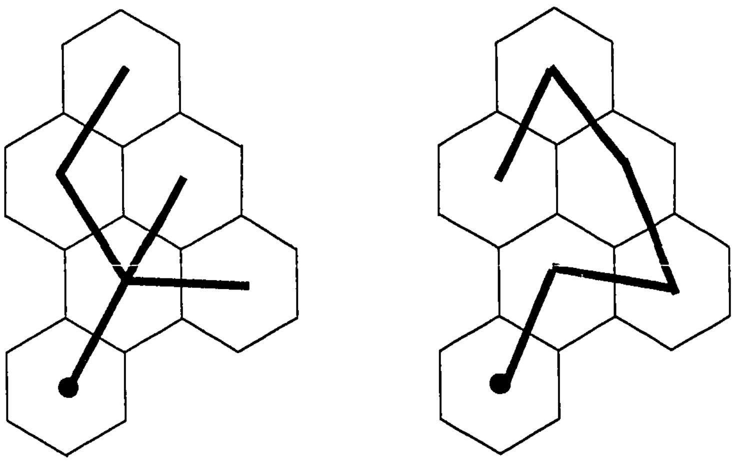


Figure 1.9: A Rooted 6–Cluster with 2 of Several Associated Trees

graph are adjacencies between cells. This graph has at least one spanning tree. Thus the number of rooted u–clusters is bounded above by the number of rooted d–ary trees with u vertices. But this number is known [36] to be

$$\binom{du}{u}\frac{1}{(d-1)u+1} \leq \frac{(3d)^u}{(d-1)u+1}.$$

Now there are n possible choices for a root and we identify the u rooted u–clusters that differ only in their root. Thus we have the factor of $\frac{n}{u}$ in the upper bound. □

Lemma 1.20. *For $r \leq j \leq m \ll n$,*

$$E[Y_j] = \Theta\left(\left(\frac{j}{n}\right)^{r-1}\right).$$

Proof. We will need more definitions in the proof of this lemma. As in the statement of the lemma for the proof we assume $r \leq j \leq m \ll n$.

Definition 1.21. *For $e \in U \subseteq E$ let (U, e) be a rooted u–cluster. Then by $S_{j,(U,e)}$ we denote the set of all sequences $f : [m] \to E$ whose j–th call saturates each of the $u = |U|$ cells of U and no others and $f(j) = e$. Furthermore let*

$$S_{j,U} = \bigcup_{e \in U} S_{j,(U,e)}.$$

Thus $S_{j,U}$ is simply the set of all sequences of length m that saturate the u-cluster U precisely at the j–th call. We now have the equation

$$E[Y_j] = \sum_{\substack{1 \leq u \leq j \\ |U|=u}} u \Pr[S_{j,U}] = \sum_{\substack{1 \leq u \leq j \\ |U|=u}} u \sum_{e \in U} \Pr[S_{j,(U,e)}] = \sum_{\substack{1 \leq u \leq j \\ |U|=u}} u \Pr[S_{j,U}], \quad (1.2)$$

where the first sum is over all u–clusters U and all $u \in \{1, 2, \ldots, j\}$. As the probability $\Pr[S_{j,U}]$ is simply the number of sequences of length m that saturate the u–cluster U precisely at the j–th call, we have

$$\Pr[S_{j,U}] = \frac{|S_{j,U}|}{n^m}.$$

We thus look for upper and lower bounds on $|S_{j,U}|$.

1.5.1 An Upper Bound For $E[Y_j]$

We consider first an upper bound on $|S_{j,(U,e)}|$. Let $v(U)$ be the number of vertices of U. There will be $v(U)$ calls connected to the $v(U)$ vertices of U when the j-th call of the sequence f arrives to saturate U. We may order the $v(U)$ vertices of (U, e) so that the last vertex is one of the r vertices of e, in $(v(U) - 1)! \cdot r$ ways.

The $(v(U) - 1)$ calls that will precede the j-th call in saturating U can appear in $\binom{j-1}{(v(U)-1)}$ ways within the first $j - 1$ calls. Each of these initial $v(U) - 1$ calls must appear in one of the, at most, d cells containing the vertex to which it is assigned. This can be done in at most, $d^{v(U)-1}$ ways.

The remaining calls prior to j are assigned to cells in $E - U$ in $n^{j-v(u)}$ ways and the $m - j$ calls after the j-th call can be assigned to any of the n cells of E in n^{m-j} ways. With this we have the following upper bound for $|S_{j,(U,e)}|$.

$$|S_{j,(U,e)}| \leq \Big(v(U) - 1\Big)! \cdot r \cdot \binom{j-1}{v(U)-1} \cdot d^{v(U)-1} \cdot n^{j-v(U)} \cdot n^{m-j} \qquad (1.3)$$

Combining 1.2 and 1.3 we obtain the upper bound for $E[Y_j]$,

$$\begin{aligned}
E[Y_j] &= \sum_{1\le u\le j} u \sum_{|U|=u} \frac{|S_{j,U}|}{n^m} \\
&\le \sum_{1\le u\le j} u \sum_{|U|=u} \sum_{e\in U} \frac{|S_{j,(U,e)}|}{n^m} \\
&\le \sum_{1\le u\le j} u \sum_{|U|=u} \sum_{e\in U} \frac{r}{n} \left(\frac{d}{n}\right)^{v(U)-1} \frac{(j-1)!}{(j-v(U))!} \\
&= O\left(\sum_{1\le u\le j} \frac{u^2}{n} \left(\frac{dj}{n}\right)^{u-1} \sum_{|U|=u} \left(\frac{dj}{n}\right)^{v(U)-u} \right).
\end{aligned}$$

Recall that requiring $\mathcal{H}$ to have Property $\mathcal{P}$ means that every non-empty subset $A \subset E(\mathcal{H})$ must have $|A| + r - 1 \le |\cup A|$. In our current notation this becomes $r - 1 \le v(U) - u$ for every u–cluster U. Using this inequality together with the last upper bound and then Lemma 1.19 we obtain,

$$E[Y_j] = O\left(\sum_{1\le u\le j} \frac{u^2}{n} \left(\frac{dj}{n}\right)^{u-1} \sum_{|U|=u} \left(\frac{dj}{n}\right)^{r-1} \right) \tag{1.4}$$

$$= O\left(d^{r-1} \left(\frac{j}{n}\right)^{r-1} \sum_{1\le u\le j} \frac{u^2}{n} \left(\frac{dj}{n}\right)^{u-1} \sum_{|U|=u} 1 \right)$$

$$= O\left(\left(\frac{j}{n}\right)^{r-1} \sum_{1\le u\le j} \frac{u^2}{n} \left(\frac{dj}{n}\right)^{u-1} \frac{n}{u} \frac{(3d)^u}{(d-1)u+1} \right). \tag{1.5}$$

The last sum in Equation 1.5 can be bounded above by the series

$$\sum_{u=0}^{\infty} \left(\frac{3d^2 j}{n}\right)^u, \tag{1.6}$$

and since $j \leq m \ll n$ we may choose n large enough so that $\frac{3d^2 j}{n} < 1$ and the sum in 1.6 becomes a convergent geometric series. We have thus shown

$$E[Y_j] = O\left(\left(\frac{j}{n}\right)^{r-1}\right). \tag{1.7}$$

It remains to find a lower bound for $E[Y_j]$.

1.5.2 A Lower Bound for $E[Y_j]$

To obtain a lower bound for $E[Y_j]$ we underestimate the probability that a 1–cluster is saturated by a sequence $f : [m] \to E$ on the j–th call, where again $j \leq m \ll n$. Using the same notation as above let $S_{j,e}$ be the set of all sequences that saturate the 1–cluster consisting of the cell e at the j–th call. Since $\Pr(S_{j,e}) \leq E[Y_j]$, this underestimate will also be valid for $E[Y_j]$. If $f \in S_{j,e}$, then necessarily $r \leq j$, and there must be a subsequence S of length $r-1$ of the first $j-1$ calls that connect to the at most d cells adjacent to $f(j) = e$. We underestimate $S_{j,e}$ by counting sequences $f : [m] \to E$ that send all of the calls of $S \cup \{j\}$ to e, and which send the remaining calls to cells disjoint from e. This is possible because there are at least $n-d-1$ cells disjoint from e. This is done in $(n-d-1)^{r-1}$ ways. We choose the subsequence S is $\binom{j-1}{r-1}$ ways. Finally we choose e in n ways. This gives the following lower bound,

$$\frac{n(n-d-1)^{j-r}}{n^j}\binom{j-1}{r-1} \leq \Pr(S_{j,1}) \leq E[Y_j].$$

By Stirling's formula [49], for some constant $C > 0$,

$$C\left(\frac{n-d-1}{n}\right)^{j-1}\left(\frac{j-1}{n-d-1}\right)^{r-1} \leq E[Y_j]. \tag{1.8}$$

Now for $j \ll n$, $((n-d-1)/n)^{j-1} \to 1$ and $(\frac{j-1}{n-d-1})^{r-1} \to (\frac{j}{n})^{r-1}$ as $n \to \infty$. This shows

$$\left(\frac{j}{n}\right)^{r-1} = O\left(E[Y_j]\right),$$

which completes the proof of the lemma. □

1.5.3 Proof of the Expectation Threshold Function

Proof. We now conclude our proof of Theorem 1.15. By linearity of expectation and the provided definitions we have the following identity:

$$E[\overline{X}_m] = \sum\nolimits_{k=r+1}^{m} E[X_k] = \sum\nolimits_{k=r+1}^{m} \frac{1}{n} E[\overline{Y}_{k-1}] = \sum\nolimits_{k=r+1}^{m} \frac{1}{n} \sum\nolimits_{j=r}^{k-1} E[Y_j].$$

The second equality holds since the probability of a rejection is equal to the expected number of saturated cells, divided by the number n, of cells. The index k starts at $r+1$ since all calls up to r are accepted. The index j starts at r since at least r calls are needed for the j-th to saturate a cluster. By Lemma 1.20,

$$E[\overline{X}_m] = \frac{1}{n}\sum\nolimits_{k=r+1}^{m}\sum\nolimits_{j=r}^{k-1}\Theta\left(\frac{j}{n}\right)^{r-1} = \Theta\left(\frac{m^{r+1}}{n^r}\right), \tag{1.9}$$

from which the statement of the theorem,

$$\lim_{n\to\infty} E[\overline{X}_m] = \begin{cases} 0 & \text{if } m \ll n^{\frac{r}{r+1}}, \\ \infty & \text{if } m \gg n^{\frac{r}{r+1}} \end{cases} \tag{1.10}$$

follows immediately. □

It should be noted that Property $\mathcal{P}$ is not a necessary condition for Theorem 1.15. This is because Property $\mathcal{P}$ was only used in Equation 1.4 where all u–clusters, or equivalently sets of hyperedges A had cardinality at most $m = m(n)$. Thus for example, cycles in 2–uniform graphs do not have Property $\mathcal{P}$ but Theorem 1.15 is true for large cycles since the inequality in question, $|A| + r - 1 \leq |\cup A|$ holds for all proper subsets of cycles.

In general however, to prove Theorem 1.15 for lattice-like tilings of the plane that do not have Property $\mathcal{P}$ a more subtle approach may be needed. In particular the lattice-like tiling by equilateral triangles given in Figure 1.6 violates the inequality in Property $\mathcal{P}$ for sets of constant cardinality as well as for sets of very large cardinality $c = c(n)$. For these types of systems it is unclear how one would arrive at the upper bound in Equation 1.4 by methods similar to those used here.

As a corollary to Theorem 1.15 we have a lower threshold function for admissibility of a call sequence $f : [m] \to E$.

Corollary 1.22. $\overline{X}_m = 0$ *almost surely if* $m \ll n^{\frac{r}{r+1}}$.

Proof. The random variable $\overline{X}_m$ takes on non-negative integer values as it is the sum of indicator random variables. We thus take the expectation over non-negative integers and by Theorem 1.15, for $m \ll n^{\frac{r}{r+1}}$ we have trivially,

$$0 = E[\overline{X}_m] = \sum_{k=0}^{m-r+1} k \Pr[\overline{X}_k = k] \geq \sum_{k=1}^{m-r+1} \Pr[\overline{X}_k = k] = \Pr[\overline{X}_m > 0]$$

and hence $\Pr[\overline{X}_m = 0] = 1$. $\square$

1.6 Applying Martingale Techniques

We will write $\overline{X}_m \sim E[\overline{X}_m]$ *almost surely* when the limiting ratio

$$\lim_{n\to\infty} \frac{\overline{X}_m}{E[\overline{X}_m]} = 1$$

for all sequences $f[m] \to E$ except for perhaps on a subset of sequences of zero probability measure.

One may ask the next natural question, what is the behavior, almost surely, of the random variable $\overline{X}_m$ for $m \gg n^{\frac{r}{r+1}}$. Unfortunately, we have not been able to show the analog to Corollary 1.22, that is, for $m \gg n^{\frac{r}{r+1}}$ we have been unable to show $\overline{X}_m \sim E[\overline{X}_m]$ almost surely.

1.6.1 An Upper Threshold for Rejection

Using martingale techniques and Azuma's inequality (also called Hoeffding's Inequality) we have been able to prove the weaker statement of the following theorem.

Theorem 1.23. $\overline{X}_m \sim E[\overline{X}_m]$ *almost surely, if* $m \gg n^{\frac{r}{r+1/2}}$

Before proceeding to the proof we will go through some of the machinery and notation of martingale theory much of which is adopted here from [1]. For a detailed discussion of Azuma's inequality and proofs, we also refer the reader to the original papers of Azuma [3] and Hoeffding [27] as well as two other papers of Maury [41] and Milman [43].

Definition 1.24. *A* martingale *is a sequence of random variables* $\{M_i\}_{i=0}^{m}$ *with the conditional expectation property that almost surely,*

$$E[M_{i+1}|M_i] = M_i, \ \forall \ i = 0, 1, 2, \ldots, m-1.$$

Denote by $E^{[m]}$ the uniformly distributed probability space of all sequences $f : [m] \to E$, then the random variable $\overline{X}_m$ is a functional on this space, as it maps a call sequence to a real number, which is the number of rejected calls in that sequence. Thus we write,

$$\overline{X}_m : E^{[m]} \to \mathbb{R}.$$

Consider now the *gradation*

$$\emptyset = B_0 \subset B_1 \subset \cdots \subset B_m = [m] \tag{1.11}$$

where $B_i = [i] = \{1, 2, \ldots, i\}$. We now define the martingale we will use, $\{M_i\}_{i=0}^m$ as follows. Let $f, g \in E^{[m]}$ then

$$M_i(f) = E[\overline{X}_m(f)|\ f(b) = g(b)\ \forall\ b \in B_i]. \tag{1.12}$$

M_i is the random variable that is the expected number of rejected calls in a randomly chosen sequence $f \in E^{[m]}$, after the first i values of this sequence are revealed, i.e. known to equal the first i values of a fixed $g \in E^{[m]}$. Clearly then, $M_0(f) = E[\overline{X}_m]$ and $M_m(f) = \overline{X}_m(f)$, which is the total number of rejections in the complete sequence f.

M_i as defined in Equation 1.12 is a martingale. For consider M_{i+1}, this is the expectation for the number of rejections in a random sequence g, the first $i+1$ calls of which are known. But conditioning $E[M_{i+1}|M_i](f)$ on M_i means that we take the expectation only with respect to random sequences f, from a subset $F^{[m]} \subseteq E^{[m]}$. But these sequences all have the same expectation for number of rejections, namely, M_i.

We will need two forms of Azuma's inequality. The first of which is quoted here.

Theorem 1.25 (Martingale Form of Azuma's Inequality). *Let* $\{M_i\}_{i=0}^m$ *be a martingale with* $M_0 = 0$ *and*

$$|M_{i+1} - M_i| \leq 1,\ \forall\ i = 0, 1, 2, \ldots, m.$$

Then for any $\lambda > 0$ we have

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

Definition 1.26. *A functional $Y : E^{[m]} \to \mathbb{R}$ satisfies the Lipschitz condition relative to the gradation in Equation 1.11 if for any two sequences $g, g' \in E^{[m]}$ that differ only over $B_{i+1} - B_i$ we have*

$$|Y(g) - Y(g')| \leq 1.$$

Clearly $\overline{X}_m(f)$ satisfies the Lipschitz condition as two sequences that differ in a single call can differ in at most a single rejected call. This is most easily seen via the hypergraph model as there it is clear that the order of the calls is unimportant in determining the total number of rejections and then we may assume g and g' differ only in their last call.

Remark 1.27. *It can be proven that the martingale defined in Equation 1.12 also satisfies the Lipschitz condition as a consequence of $\overline{X}_m(f)$ satisfying this condition, see [1] for a proof. This fact together with the martingale form of Azuma's inequality in Theorem 1.25 gives the generalized form of Azuma's inequality.*

Theorem 1.28 (Generalized Form of Azuma's Inequality). *Let $Y : E^{[m]} \to \mathbb{R}$ satisfy the Lipschitz condition relative to a gradation of length m*

and let $\mu = E[Y]$. *Then for all* $\lambda > 0$,

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

Since $\overline{X}_m(f)$ satisfies the conditions of Azuma's inequality as stated in Theorem 1.28 we have for all $\lambda > 0$

$$\Pr[|\overline{X}_m - E[\overline{X}_m]| > \lambda\sqrt{m}] < 2e^{-\lambda^2/2}. \tag{1.13}$$

To complete the proof of Theorem 1.23 choose $\epsilon > 0$ and set $\lambda = \epsilon E[\overline{X}_m]/\sqrt{m}$. Equation 1.13 then becomes

$$\Pr\left[\left|\frac{\overline{X}_m}{E[\overline{X}_m]} - 1\right| > \epsilon\right] < 2e^{-\epsilon^2\left(E[\overline{X}_m]\right)^2/2m}. \tag{1.14}$$

Taking $m \gg n^{\frac{r}{r+1/2}}$, Equation 1.9 in the proof of Theorem 1.15 implies that $E[\overline{X}_m]/\sqrt{m} \to \infty$ as $n \to \infty$, which with Inequality 1.14 implies that $\overline{X}_m \sim E[\overline{X}_m]$ almost surely. □

1.6.2 Closing the Gap between $n^{\frac{r}{r+1}}$ and $n^{\frac{r}{r+1/2}}$

The upper and lower thresholds for lengths of admissible sequences and those with infinitely many rejections are $n^{\frac{r}{r+1}}$ and $n^{\frac{r}{r+1/2}}$, respectively. The question of whether this gap can be narrowed or eliminated remains open.

One may attempt to close this gap via Chebyschev's inequality in the form,

$$\Pr[|\overline{X}_m - E[\overline{X}_m]| > \epsilon E[\overline{X}_m]] \leq \frac{\mathrm{Var}[\overline{X}_m]}{\epsilon^2 E[\overline{X}_m]^2}$$

where one chooses any $\epsilon > 0$ and $\mathrm{Var}[\overline{X}_m]$ is the variance of $\overline{X}_m$. With this inequality, showing $\mathrm{Var}[\overline{X}_m] = o(E[\overline{X}_m]^2)$ for $m \gg n^{\frac{r}{r+1}}$ would imply $\overline{X}_m \sim E[\overline{X}_m]$ almost surely. If this were indeed the case all sequences of length $m \gg n^{\frac{r}{r+1}}$ with infrequent rejections would themselves occur with such infrequency so as to have measure zero.

Now if the events of call rejection R_i, were independent then the covariances $\mathrm{Cov}[X_i X_j] = 0,\ \forall\ i \neq j$ and $\mathrm{Var}[\overline{X}_m] = o(E[\overline{X}_m]^2)$ would follow immediately. In reality however the R_i are not independent. Consider the pair of calls i, j, with $i < j$. Knowledge that call i was rejected, no matter how far back in the sequence is essentially the same as one less call in the system and hence call j actually now has the higher acceptance probability of call $j - 1$.

Conjecture 1.29. *Based on the results of the next section for hypergraphs $\mathcal{H}$ for which the parameter $r = 2$ we conjecture that $\overline{X}_m \sim E[\overline{X}_m]$ almost surely for $m \gg n^{\frac{r}{r+1}}$.*

1.7 A Limiting Probability Distribution

Theorem 1.30 is attributable to Florian Lengyel. It gives a limiting probability distribution for the admissibility of a random sequence $f : [m] \to [n]$ on the linear system, i.e., the hypergraph model $\mathcal{H}$ in the case where $r = 2$.

Theorem 1.30. *Let A be a positive constant. Then in the 2-uniform hypergraph model $\mathcal{H}$, we have*

$$\lim_{n\to\infty} \Pr[f : [m] \to [n] \text{ is admissible }] = \begin{cases} 1 & \text{if } m \ll n^{2/3}, \\ e^{-A^3/6} & \text{if } m \sim An^{2/3}, \\ 0 & \text{if } m \gg n^{2/3}. \end{cases}$$

We refer the reader to [26] for a detailed proof and offer here only a broad outline. The algebraic proof has several steps. Set w_f to be the word of length n on the alphabet $\{\mathbf{0}, \mathbf{1}, \mathbf{2}, \ldots, \mathbf{n}\}$, where the k-th symbol of w_f is the number of times that a call is placed to the k-th segment of the 2-dimensional system. Then the set of all admissible words w_f is given by $\mathcal{L} =(\mathbf{0}^*\mathbf{1}^*(\mathbf{21}^*\mathbf{0})^*)^* (\epsilon \cup \mathbf{21}^*)$, where ϵ is the empty word and $*$ is the Kleene $*$-operation, i.e., $\ell*$ represents any finite number of ℓ's.

We write $A(m,n) = \widehat{M}(m,n) + N(m,n)$, where $\widehat{M}(m,n)$ is the number of admissible functions $f : [m] \to [n]$ whose word w_f does not end in $\mathbf{21}^*$ and where $N(m,n)$ is the number of admissible functions f whose word w_f ends in $\mathbf{21}^*$.

Towards solving for $\widehat{M}(m,n)$ we define $M(m,z) = \widehat{M}(m, m+z)$, where $z \geq 0$. We write a formal sum of words of $\mathcal{L}$ that represents the admissible functions counted by $M(m,z)$ as follows

$$\binom{m+z}{z}\mathbf{1}^m + \sum_{k=1}^{\lfloor m/2 \rfloor} \sum_{s=0}^{m-2k} \sum_{\substack{\vec{d}\in D_{s,k} \\ \vec{e}\in E_{s,k}}} \binom{s+k+z}{z} \mathbf{1}^{d_0}\alpha_{e_1}\mathbf{1}^{d_1}\cdots\alpha_{e_k}\mathbf{1}^{d_k}, \quad (1.15)$$

where $D_{s,k} = \{(d_0, \ldots, d_k) : d_i \geq 0, \sum_{j=0}^{k} d_j = s\}$, $E_{s,k} = \{(e_1, \ldots, e_k) : e_i \geq 2, \sum_{j=1}^{k} e_j = m - s\}$, and where α_i denotes the word $\mathbf{2}\mathbf{1}^{i-2}\mathbf{0}$, where $i \geq 2$.

Now every admissible function x not ending in $\mathbf{21}^*$ comes from a unique word y of the form

$$\mathbf{1}^{d_0}\alpha_{e_1}\mathbf{1}^{d_1}\cdots\alpha_{e_k}\mathbf{1}^{d_k} \tag{1.16}$$

by inserting $z = |x| - |y|$ $\mathbf{0}$'s into y. It remains to compute the number of functions $f : [m] \to [m+z]$ that have an associated word of the form (1.16).

Let $\#w_f$ denote the number of functions with $w_g = w_f$. For j with $1 \leq j \leq n$ let $|f^{-1}(j)|$ denote the cardinality of the pre-image of j, i.e., the number of calls placed to the j-th segment of the system. Then

$$\#w_f = \binom{m}{|f^{-1}(1)|, \ldots, |f^{-1}(n)|}. \tag{1.17}$$

Furthermore, if $f : [j] \to [r]$ and $g : [k] \to [s]$ are functions with words $u = w_f$ and $v = w_g$ respectively, then

$$\#(uv) = \#u \cdot \#v\binom{j+k}{j}. \tag{1.18}$$

After some algebra we have

$$M(m,z) = \frac{(m+z)!}{z!} + m!\sum_{k=1}^{\lfloor m/2 \rfloor}\frac{1}{2^k}\sum_{s=0}^{m-2k}\binom{s+k+z}{s,k,z}\binom{m-s-k-1}{k-1}. \tag{1.19}$$

We compute $N(m,n)$, the number of admissible functions ending in $\mathbf{21}^*$

as follows. Suppose that $f : [m] \to [n]$ has word w_f ending in $\mathbf{21}^{k-1}$ where $1 \leq k \leq m-1$. Then $w_f = uv$, where u does not end with $\mathbf{21}^*$. Furthermore $|u| = n-k$, and $C(u) = m-k-1$. The number of functions $g : [m-k-1] \to [n-k]$ with word $w_g = u$ is by definition $M(m-k-1, n-m+1)$. By (1.17), the number of functions $h : [k+1] \to [k]$ with $w_h = v$ is $\frac{(k+1)!}{2}$. By (1.18), the number of functions f is

$$\sum_{k=1}^{m-1} M(m-k-1, n-m+1)\frac{(k+1)!}{2}\binom{m}{k+1}.$$

An application of Gosper's algorithm to the inner sum of Equation 1.19 gives $M(m,z)$ in terms of the generalized hypergeometric function

$${}_pF_q\left[\begin{matrix} a_1, \ldots, a_p \\ b_1, \ldots, b_q \end{matrix}; x\right] = \sum_{n=0}^{\infty} \frac{(a_1)_n \cdots (a_p)_n}{(b_1)_n \cdots (b_q)_n} \frac{x^n}{n!}.$$

Further asymptotic analysis concludes the proof.

Using Theorem 1.30 it is not difficult to show that the conjecture of the previous section is true for hypergraphs in which $r = 2$. We must show that for sequence lengths $m \gg n^{2/3}$ there are arbitrarily many rejections. In particular for any $M \in \mathbb{N}$ choose $\epsilon = 1/M$ then the sequence of length $\epsilon m \gg n^{2/3}$ and hence by Theorem 1.30 has at least on rejected call, almost surely. We may append M such sequences of length $\epsilon m \gg n^{2/3}$ to the first sequence to obtain at least one rejected call per appended sequence for a total of at least M rejected calls.

Chapter 2

The $\mathcal{G}_{k\text{-out}}$ Model for Random Graphs

2.1 Introduction

In the second part of this dissertation we study the $\mathcal{G}_{k\text{-out}}$ model for random graphs. After defining the model we study connectedness and vertex degree distribution properties. In particular we bring down two proofs [34], [37] showing that $\mathcal{G}_{k\text{-out}}$ for $k = 1$ is a.s. disconnected and offer a new proof to the known result that for $k = 2$ $\mathcal{G}_{k\text{-out}}$ is a.s. connected.

We extend the model from integer values of k to positive real values r and study the connectedness properties for this model as well. Lastly, we study the planarity properties of the $\mathcal{G}_{k\text{-out}}$ model and our extension of this model. We find bounds on r and c so that a.s. $G \in \mathcal{G}_{r\text{-out}}$ has a minor K_c. In the next chapter we apply some of the techniques developed here to a problem in analytic number theory.

As mentioned in the introduction to Chapter One, from 1959-1961 Paul Erdős and Alfréd Rényi published a series of seminal papers on the random graph models $\mathcal{G}_{n,p}$ and $\mathcal{G}_{n,M}$ [12], [13] and [14]. In the papers they studied some of the basic properties of these models, including degree distributions and connectivity thresholds. There has since been a plethora of research on these models, [31] and [4] are thorough accounts of the recent research. We will study a relatively new model for generating random graphs.

Our study of the $\mathcal{G}_{k\text{-out}}$ model is motivated by discussion of a graph theoretic problem stemming from the study of neural structure and development. The brain may be described on some level as a graph where neurons are vertices and connections between neurons are edges. In studying neural development we allow for an initial stage where there are no connections between neurons. Each neuron is then allowed to connect, at random, to an integer number k of other neurons. The resulting structure describes the $\mathcal{G}_{k\text{-out}}$ model.

An important property of this a structure is the average distance between any two vertices, where distance is defined as the least number of edges in a path connecting two vertices. When simulating the $\mathcal{G}_{k\text{-out}}$ model, we found that the average distance between any two vertices increased with the total number of vertices in the graph. Unexpectedly though, we also noted that the rate of increase was periodic when viewed on a logarithmic scale [51].

The explanation of this phenomena is at present incomplete and will not be addressed here.

2.2 Defining the Model

Bollobás describes the model $\mathcal{G}_{k\text{-out}}$ in [4] (page 37). For $k \geq 1$ a fixed integer, n a large integer and vertex set $V = \{1, 2, \ldots, n\}$, we construct the edge set E, of a random graph from $\mathcal{G}_{k\text{-out}}$, as follows. For each vertex $v \in V$ select k distinct vertices $\{u_1, u_2, \ldots, u_k\}$ uniformly and at random from $V - \{v\}$. This selection describes the set of directed edges $\{vu_1, vu_2, \ldots, vu_k\}$ of E. However, unless stated otherwise, we ignore the direction of an edge in subsequent analysis and consider $\mathcal{G}_{k\text{-out}}$ an undirected random graph space. In the event v connects "out" to u and u connects "out" to v we count the edge $vu \equiv uv$ only once. It is immediate that there are no parallel edges in the $\mathcal{G}_{k\text{-out}}$ model. It is also clear that vertices have minimum degree k and maximum degree $n-1$. Figure 2.1 shows a subset of three vertices from a random $G \in \mathcal{G}_{3\text{-out}}$ model. Note that the edge between vertices u and v is both an "in" and "out" edge. We will later show the maximum degree in $\mathcal{G}_{k\text{-out}}$ is almost surely, in a sense defined below, $O\left(\frac{\log n}{\log \log n}\right)$.

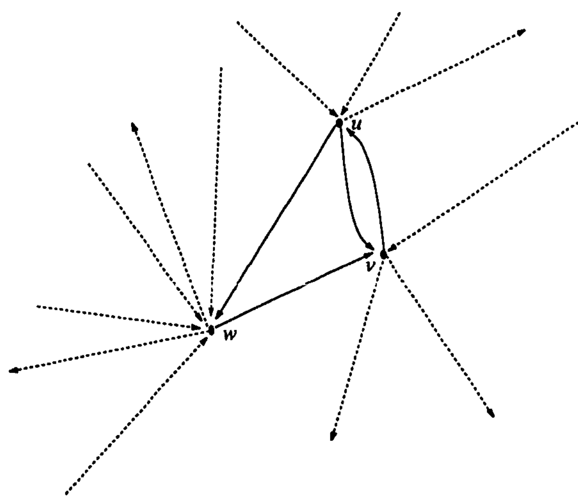


Figure 2.1: A Section from a random graph $G \in \mathcal{G}_{3\text{-out}}$

2.3 On the Connectedness of $\mathcal{G}_{k\text{-out}}$

Recall that a graph is said to be *simply connected* if any two vertices can be joined by a path, otherwise it is *disconnected.* It is natural to ask for a probability distribution for simple connectedness on $\mathcal{G}_{k\text{-out}}$. We will show at least asymptotically, that such probabilities are either 0 or 1 depending on the value of k. We say that a random graph model $\mathcal{G}$ *almost surely* or *a.s.* has Property $\mathcal{P}$ if $\Pr[G_{\mathcal{P}}] \to 1$ as $n \to \infty$, where $G_{\mathcal{P}}$ is the event that a randomly chosen graph $G \in \mathcal{G}$ has Property $\mathcal{P}$.

Theorem 2.1. *(Katz 1955) $\mathcal{G}_{1\text{-out}}$ is a.s. disconnected.*

In the next chapter we apply an alternative method of counting connected graphs $G \in \mathcal{G}_{1\text{-out}}$, which, together with Theorem 2.1, gives a graph theoretic technique for arriving at results in analytic number theory.

There are several proofs of Theorem 2.1 in the literature. We outline here two such proofs.

Proof 1. Kolchin [37] (Theorem 1.7.1) enumerates the set U_n of all labelled graphs on n vertices each of whose components is *unicyclic*, i.e., contains exactly one cycle. He gives this as

$$u_n = |U_n| = \frac{\sqrt{2\pi}\, e^{3/4}}{2^{1/4}\,\Gamma(1/4)}\, n^{n-1/4}\left(1 + o(1)\right), \tag{2.1}$$

where

$$\Gamma(x) = \int_0^\infty \xi^{x-1}\, e^{-\xi} d\xi$$

is the gamma function.

By the construction of a random $G \in \mathcal{G}_{1\text{-out}}$ it is clear that any connected component is unicyclic (see Figure 3.1). We show this as follows. Each cycle on q vertices has q edges and since each vertex has only a single "out" edge which must be in the cycle there can be no "out" edges emanating from this cycle. Assume that a single component has more than one cycle, C_1 and C_2 . Then there must be a vertex with an edge oriented "out" toward C_1 and an edge oriented "out" toward C_2. But this is a violation of the $\mathcal{G}_{1\text{-out}}$ construction, requiring only one edge "out" of any single vertex.

Let $\mathcal{U}_n$ be the event that a random graph $G \in \mathcal{G}_{1\text{-out}}$ is unicyclic with a single component. Since the set of unicyclic graphs with a single component

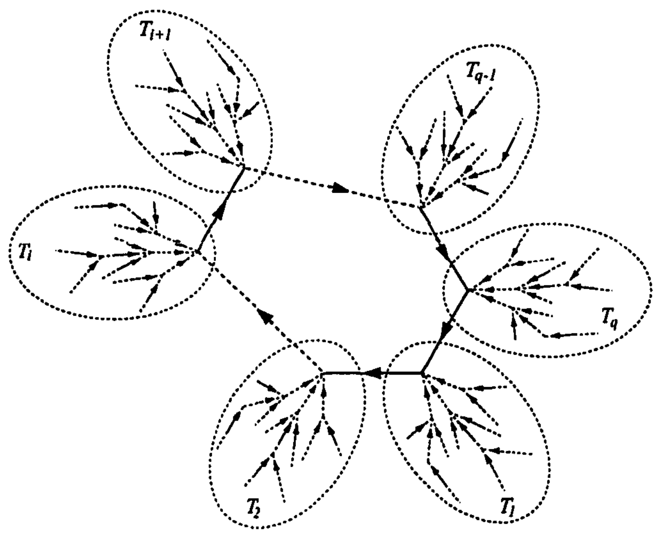


Figure 2.2: A connected graph from $\mathcal{G}_{1\text{-out}}$

is a subset of U_n, we have by equation 2.1

$$\Pr[\mathcal{U}_n] \leq \frac{|U_n|}{|\mathcal{G}_{1\text{-out}}|} = \frac{u_n}{(n-1)^n} = O\left(\frac{n^{n-1/4}}{n^n}\right) = O\left(n^{-1/4}\right) \to 0,$$

which shows $\mathcal{G}_{1\text{-out}}$ a.s. disconnected. □

Proof 2. Katz [34] proves the same result with faster convergence of $O\left(1/\sqrt{n}\right)$ by considering random mappings from a set onto itself, $f : \Omega \to \Omega$.

A map is *hollow* if $f(\omega) \neq \omega$ for all $\omega \in \Omega$ and a map is *indecomposable* if the minimal non-null subset $\Psi \subseteq \Omega$ for which $f(\Psi) = \Psi$ and $f^{-1}(\Psi) = \Psi$ is the whole set, that is $\Psi = \Omega$. There is a one-one correspondence between hollow mappings from the set $[n]$ to itself and the random graph space $\mathcal{G}_{1\text{-out}}$.

Furthermore, the set of indecomposable mappings corresponds to connected graphs $G \in \mathcal{G}_{1\text{-out}}$.

To complete our proof we note that Katz proves that the probability of a random hollow mapping $f : \Omega \to \Omega$ being indecomposable is $O(1/\sqrt{n})$. Thus we see the probability of a graph $G \in \mathcal{G}_{1\text{-out}}$ being connected tends to zero. $\square$

The result of Katz shows faster convergence than that of Kolchin since the former finds the probability of only single component unicyclic graphs while the latter finds the probability of all unicyclic graphs. These results, while tending to zero with $n \to \infty$ do so rather slowly, much more so than say $1/n$. We will return to this fact when considering connectivity thresholds for the model $\mathcal{G}_{r\text{-out}}$ for real $1 \leq r \leq 2$. We now consider the connectedness of $\mathcal{G}_{k\text{-out}}$ for $k \geq 2$.

Theorem 2.2. *(Fenner and Frieze 1982) For $k \geq 2$, $\mathcal{G}_{k\text{-out}}$ is a.s. connected.*

Proof. It is enough to consider the case $k = 2$. For assume the Theorem true for $k = 2$, then for $k > 2$, in the directed model, randomly color black 2 "out" edges from each vertex, color white the remaining $k - 2$ "out" edges. The subgraph consisting of only black edges is a random element of $\mathcal{G}_{2\text{-out}}$ and hence is a.s. connected by the assumption. But this subgraph spans all vertices so the entire graph is a.s. connected.

Let G_{disconn} be the event that a randomly chosen graph $G \in G_{2\text{-out}}$ is disconnected. We will prove $\Pr[G_{\text{disconn}}] \to 0$. For every disconnected graph

$G \in \mathcal{G}_{2\text{-out}}$ there is a bipartition of the vertex set $V(G) = U \cup W$ into disjoint sets U, W so that no edges cross between U and W. Furthermore each set of the bipartition contains at least 3 vertices since each vertex must connect to at least 2 distinct vertices. Denote by $G_{i\text{-disconn}}$ the event that there is an i element subset $I \subset V(G)$ such that there are no edges between I and its complement $\overline{I}$. Note that i need not be the order of the smallest component. We have the inequality,

$$\Pr[G_{\text{disconn}}] \leq \sum_{i=3}^{\lfloor \frac{n}{2} \rfloor} \Pr[G_{i\text{-disconn}}]. \tag{2.2}$$

We do not have strict equality since the terms on the right are not mutually exclusive, as a graph may be disconnected with several components of order $\leq \lfloor \frac{n}{2} \rfloor$.

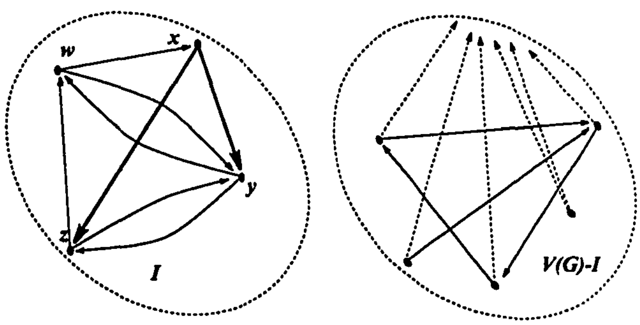


Figure 2.3: A component of order 4 from the $\mathcal{G}_{2\text{-out}}$ model

To find an upper bound on $\Pr[G_{i\text{-disconn}}]$ we consider a set $I \subset V(G)$ of order i and count the number of mappings from I into the set of distinct pairs of I, such that if $x \mapsto (y,z)$ then $x \neq y \wedge x \neq z$, where the last condition on the mapping stems from the restriction that $\mathcal{G}_{k\text{-out}}$ contain no loops. Figure 2.3 illustrates this with $|I| = 4$. The event that we refer to as $x \mapsto (y,z)$ is illustrated with the bold edges xz and xy. There are $\binom{i-1}{2}^i$ such mappings. We thus have an upper bound,

$$\Pr[G_{i\text{-disconn}}] \leq \binom{n}{i} \frac{\binom{i-1}{2}^i \binom{n-i-1}{2}^{n-i}}{\binom{n-1}{2}^n}, \tag{2.3}$$

where the first factor finds a specific choice for a bipartition of $V(G)$ into sets of order i and $n-i$. Using the version of Stirling's formula (see [49]),

$$m! = \sqrt{2\pi m}\left(\frac{m}{e}\right)^m e^{\alpha_m} \qquad \text{where} \quad \frac{1}{12m+1} \leq \alpha_m \leq \frac{1}{12m},$$

on the right side of Inequality 2.3 we obtain the upper bound

$$\Pr[G_{i\text{-disconn}}] \leq \frac{e^{\frac{1}{12n}-\frac{1}{12i}-\frac{1}{12(n-i)}}}{\sqrt{2\pi}} \cdot \frac{n^{n+1/2}}{i^{i+1/2}\,(n-i)^{n-i+1/2}} \cdot \frac{i^{2i}(n-i)^{2n-2i}}{(n-1)(n-2)}$$

$$\leq \frac{i^{i-1/2}\,(n-i)^{n-i-1/2}}{(n-2)^{n-1/2}}. \tag{2.4}$$

Together with Inequality 2.2 we have

$$\Pr[G_{\text{disconn}}] \le \sum_{i=3}^{\lfloor \frac{n}{2} \rfloor} (n-2)^{1/2-n}\, i^{i-1/2}\,(n-i)^{n-i-1/2} \tag{2.5}$$

$$\le (n-2)^{1/2-n} \sum_{i=3}^{\lfloor \frac{n}{2} \rfloor} i^{i}\,(n-i)^{n-i-1/2} \tag{2.6}$$

$$\le 3^3 n^{1/2-n}\,\frac{n}{2}\,n^{n-7/2} \tag{2.7}$$

$$= O\big(n^{-2}\big), \tag{2.8}$$

where the bound in (2.7) comes from the fact that the product $i^i \cdot (n - i)^{n-i}$ is maximized over the interval $[3, \lfloor \frac{n}{2} \rfloor]$ when $i = 3$. We have shown $\Pr[G_{\text{disconn}}] = O\big(n^{-2}\big) \to 0$ and thus $\mathcal{G}_{2\text{-out}}$ is a.s. connected. $\square$

It seems reasonable to expect the approximations in Inequalities 2.6 and 2.7 to be refined with further analysis and that one could show the bound in Inequality 2.8 to converge faster than $O\left(n^{-2}\right)$. Thus for some extension of our model there may be an a.s. connected random graph space between the a.s. connected $\mathcal{G}_{2\text{-out}}$ and the a.s. disconnected $\mathcal{G}_{1\text{-out}}$. In the next section we define an extension of $\mathcal{G}_{k\text{-out}}$ to study this idea.

2.4 An Extension of $\mathcal{G}_{k\text{-out}}$ and the Poisson Paradigm

We define a random graph space $\mathcal{G}_{r\text{-out}}$ on n vertices, for $r = r(n) \in \mathbb{R}^+$, which is an extension of the space $\mathcal{G}_{k\text{ out}}$ for $k \in \mathbb{N}$. Basically each vertex chooses $\lfloor r \rfloor$ vertices to draw edges out to and then with probability $r - \lfloor r \rfloor$ the $\lceil r \rceil$-th edge is also drawn. So that each vertex has either $\lfloor r \rfloor$ edges drawn out or $\lfloor r \rfloor + 1 = \lceil r \rceil$ drawn out. We find the threshold on r for a.s. connectedness.

For fixed $r \in \mathbb{R}^+$, n a large integer and vertex set $V = \{1, 2, \ldots, n\}$, we construct the edge set E, of a random graph from $\mathcal{G}_{r\text{-out}}$, in a manner similar to the construction of the integer based model $\mathcal{G}_{k\text{-out}}$. Let $p = r - \lfloor r \rfloor$. For each vertex $v \in V$ select $\lfloor r \rfloor$ distinct vertices $\{u_1, u_2, \ldots, u_{\lfloor r \rfloor}\}$ uniformly and at random from $V - \{v\}$. Add the edges $\{vu_1, vu_2, \ldots, vu_{\lfloor r \rfloor}\}$ to E. With probability p select an additional vertex w, uniformly and at random from $V - \{v, u_1, u_2, \ldots, u_{\lfloor r \rfloor}\}$ and add the edge vw to E.

In studying $\mathcal{G}_{r\text{-out}}$, we are primarily interested in the *phase transition*[1] interval $1 \leq r \leq 2$. Being that the probability of the $G \in \mathcal{G}_{1\text{-out}}$ model being connected converges to zero rather slowly, on the order of $1/n^{1/2}$ by a result of Katz [34] previously stated, it would seem intuitive to expect small values of p to be associated with the connectedness phase transition. In fact as the

[1]My thanks to Prof. Noson Yanofsky for our discussions on phase transitions.

next result shows, this is not the case at all. [2]

Jaworski and Smit [32] study some of the properties of similar graphs. They consider doubly random digraphs, where first a random number of "out" edges from a particular vertex is determined and next the terminal vertices to these edges are selected.

In what follows we will need to find a distribution for the number of isolated edges in the $\mathcal{G}_{1\text{-out}}$ model. We denote this number by the random variable X. Denote by ξ_i the event that the i–th vertex pair of $G \in \mathcal{G}_{1\text{-out}}$ is an isolated edge, $i = 1, \ldots, \binom{n}{2}$ and let X_i be the indicator random variable for ξ_i so that $X = \sum_i X_i$. There is an implicit assumption here that X is a function of n, the number of vertices in a random $G \in \mathcal{G}_{1\text{-out}}$.

If the events ξ_i were independent then X would be the sum of independent trials and hence binomial. This is not the case however, since knowledge that a particular edge is isolated, say ξ_1, increases the probability that some other edge is isolated, say ξ_2. This, since vertices of edge 1 are necessarily isolated from vertices of edge 2. To take the most extreme case, if we know that all but one edge is isolated, the last edge must also be isolated. However, the events ξ_i are "mostly" independent and thus we use the "Poisson Paradigm", which roughly states that if X is a random variable which is the sum of many "mostly" independent random variables, each with small probability, then X is asymptotically Poisson.

We use the Brun's sieve approach to the "Poisson Paradigm" to show

[2] I am grateful to Prof. Béla Bollobás for his insights in arriving at this result.

that X is asymptotically Poisson. The approach is outlined in [1] (Chapter 8), and in the methods in the classic paper of Erdős and Rényi [13] (Lemma 1 of §2). In the sequel we follow the approach in [13].

Lemma 2.3. *(Erdős and Rényi 1961) Let $X_{n,1}, X_{n,2}, \ldots, X_{n,\ell_n}$, $n = 1, 2, \ldots$ be indicator random variables on some probability space and let*

$$X_n = \sum_{i=1}^{i=\ell_n} X_{n,i}.$$

If

$$\lim_{n\to\infty} \sum_{1\le i_1<i_2<\cdots<i_\rho\le\ell_n} E[X_{n,i_1}X_{n,i_2}\cdots X_{n,i_\rho}] = \frac{\lambda^\rho}{\rho!}$$

uniformly in ρ for all $\rho = 1, 2, \ldots$, where $\lambda > 0$ and the summation is over all combinations $\{i_1, i_2, \ldots, i_\rho\} \subseteq \{1, 2, \ldots, \ell_n\}$ then $X_n \to X$ is Poisson distributed with mean λ.

Remark 2.4. *Although the statement of the lemma requires uniform convergence with respect to ρ, the actual proof of the lemma uses only a weaker form of convergence. In particular the weaker convergence is given by*

$$\lim_{n\to\infty} \sum_{\rho=0}^{\infty} \Gamma(n,\rho) z^\rho = e^{\lambda z}, \tag{2.9}$$

where

$$\Gamma(n,\rho) = \sum \Pr[\xi_{i_1} \wedge \ldots \wedge \xi_{i_\rho}],$$

the sum being over all ρ-sets and $|z| < 1$. By a ρ-set we mean a subset

$\{i_1, i_2, \ldots, i_\rho\} \subseteq \{1, 2, \ldots, \ell_n\}$. *We will use this weaker condition in applying the lemma.*

Theorem 2.5. *If $r(n) \leq 2 - \epsilon$ for some $\epsilon > 0$, then $\mathcal{G}_{r\text{-out}}$ is disconnected with positive probability $p(\epsilon) > 0$.*

On the other hand, if $r(n) > 2 - \frac{2}{n-2}$ then $\mathcal{G}_{r\text{-out}}$ is a.s. connected.

Proof. Let $q = q(n) = 2 - r(n)$ and $p = 1 - q$. We construct a random graph in $\mathcal{G}_{r\text{-out}}$ by first choosing a random $G \in \mathcal{G}_{1\text{-out}}$ and then adding an additional random edge with probability p from each vertex in G to generate a random $H \in \mathcal{G}_{r\text{-out}}$. By Theorem 2.1 G is a.s. disconnected. Let X be the number of isolated edges in G. Each such edge must be both an "in" and an "out" edge as illustrated in Figure 2.4. Then it is clear that

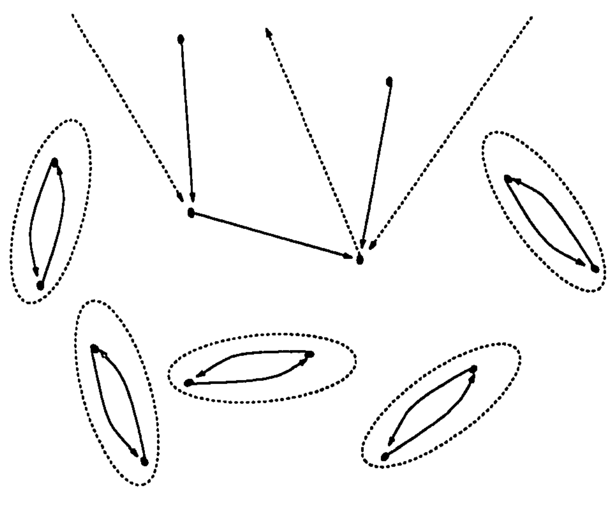


Figure 2.4: Isolated edges of the $\mathcal{G}_{1\text{-out}}$ model

$$\lim_{n\to\infty} E[X] = \lim_{n\to\infty} \binom{n}{2} \left(\frac{1}{n-1}\right)^2 \left(1 - \frac{2}{n-1}\right)^{n-2} = \frac{e^{-2}}{2}. \qquad (2.10)$$

The second term finds the probability that a particular edge is a double edge, while the last term finds the probability that no other vertex connects to that edge. A *double* edge connects the same vertex pair via two directed edges, an in-edge and an out-edge.

We use Lemma 2.3 to show that X is asymptotically Poisson. Denote by ξ_i be the event that the i–th vertex pair of $G \in \mathcal{G}_{1\text{-out}}$ is an isolated edge, $i = 1, \ldots, \binom{n}{2}$ and let X_i be the indicator random variable for ξ_i so that $X = \sum_i X_i$. Note that we have omitted the subscripts of n so that $X = \sum_i X_i$ should be understood in the context of Lemma 2.3 as $X_n = \sum_i X_{n,i}$, the implicit assumption being that every $G \in \mathcal{G}_{r\text{-out}}$ has n vertices and $n \to \infty$.

By Remark 2.4 we need show only that (2.9) holds instead of the stronger uniform convergence requirement in [13]. We will first show that

$$\begin{aligned} \lim_{n\to\infty} \Gamma(n,\rho) &= \lim_{n\to\infty} \sum \Pr[\xi_{i_1} \wedge \cdots \wedge \xi_{i_\rho}] \\ &= \lim_{n\to\infty} \sum E[X_{i_1} X_{i_2} \cdots X_{i_\rho}] = \frac{\left(\frac{e^{-2}}{2}\right)^\rho}{\rho!}, \end{aligned} \qquad (2.11)$$

for all $\rho = 1, 2, \ldots$, and then that this convergence is fast enough so that (2.9) also holds. Here again, the summations are over all ρ–sets of vertex pairs, i.e. potential edges of G, so that $\ell(n) = \binom{n}{2}$. The first equality in (2.11) holds by definition and the second holds since the random variables X_{i_j} are indicators and thus their joint expectations are the joint probabilities of the ξ_{i_j}.

For the third equality in (2.11) we need to compute the probabilities of ρ–sets of edges being isolated edges. To illustrate these events Figure 2.5 depicts a 3–set of isolated edges and Figure 2.6 depicts a 3–set of edges not all of which are isolated.

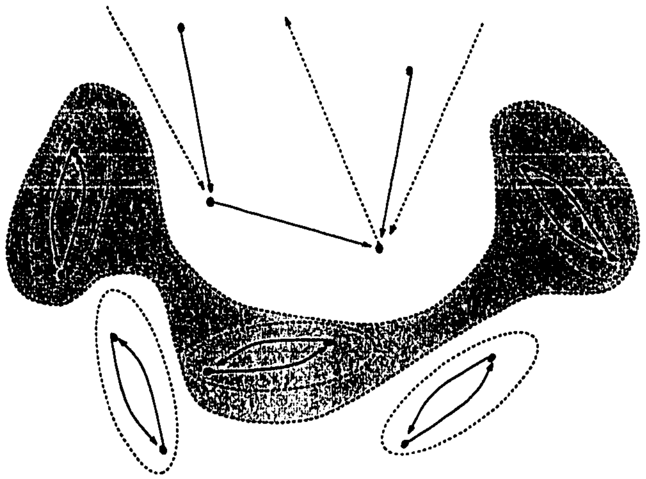


Figure 2.5: A 3–set of isolated edges

Now there are $\binom{\binom{n}{2}}{\rho}$ summands in the sum of (2.11) but not all events have equal probability. In particular we must exclude summands where not all of the edges in the ρ–set are mutually vertex disjoint, for these by definition can not all be isolated and thus have probability zero. The number of ρ–sets

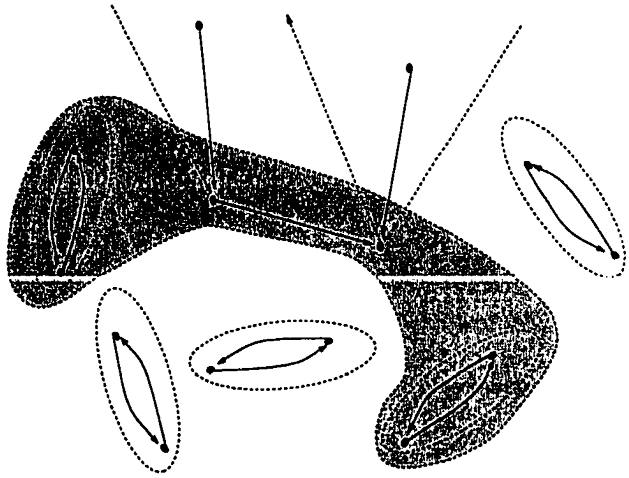


Figure 2.6: A 3–set of edges with one non-isolated edge

for which all edges are mutually vertex disjoint is given by

$$\frac{\binom{n}{2}\binom{n-2}{2}\cdots\binom{n-2\rho+2}{2}}{\rho!} = \frac{\frac{n!}{(n-2)!}\frac{(n-2)!}{(n-4)!}\cdots\frac{(n-2\rho+2)!}{(n-2\rho)!}}{2^\rho \rho!} = \frac{n!}{(n-2\rho)!\rho!2^\rho}. \tag{2.12}$$

Now each of the summands for these mutually vertex disjoint ρ–sets of edges have identical probability and given the previous argument we rewrite the second term in (2.11) as

$$\lim_{n\to\infty} \frac{n!}{(n-2\rho)!\rho!2^\rho} \Pr[\xi_{i_1} \wedge \cdots \wedge \xi_{i_\rho}] = \lim_{n\to\infty} \frac{(n^2/2)^\rho}{\rho!} \Pr[\xi_{i_1} \wedge \cdots \wedge \xi_{i_\rho}]$$

for fixed ρ. Multiplying the last term and (2.11) by $\rho!2^\rho$ we have that (2.11) is equivalent to

$$\lim_{n\to\infty} n^{2\rho}\Pr[\xi_{i_1}\wedge\cdots\wedge\xi_{i_\rho}] = \lim_{n\to\infty} n^{2\rho}\prod_{\sigma=1}^{\rho}\Pr\left[\xi_{i_\sigma} \mid \bigwedge_{s<\sigma}\xi_{i_s}\right] = e^{-2\rho} \tag{2.13}$$

for fixed ρ. The first equality follows from the definition of conditional probability. Expanding the conditional probability in the middle term of (2.13) and again noting that there are a fixed number of terms in the product, we bring the limit and the n^2 term into the product,

$$\begin{aligned}\lim_{n\to\infty} n^{2\rho}\prod_{\sigma=1}^{\rho}\Pr\left[\xi_{i_\sigma} \mid \bigwedge_{s<\sigma}\xi_{i_s}\right] &= \prod_{\sigma=1}^{\rho}\lim_{n\to\infty} n^{2}\Pr\left[\xi_{i_\sigma} \mid \bigwedge_{s<\sigma}\xi_{i_s}\right]\\ &= \prod_{\sigma=1}^{\rho}\lim_{n\to\infty} n^2\left(\frac{1}{n-1-2\sigma}\right)^2\left(1-\frac{2}{n-1-2\sigma}\right)^{n-2-2\sigma}\\ &= \prod_{\sigma=1}^{\rho} e^{-2} = e^{-2\rho}.\end{aligned} \tag{2.14}$$

We compute the conditional probability,

$$\Pr\left[\xi_{i_\sigma} \mid \bigwedge_{s<\sigma}\xi_{i_s}\right] = \left(\frac{1}{n-1-2\sigma}\right)^2\left(1-\frac{2}{n-1-2\sigma}\right)^{n-2-2\sigma}$$

in (2.14) as follows. We have the event ξ_{i_σ} for fixed $\sigma \leq \rho$ conditioned upon the events $\wedge_{s<\sigma}\xi_{i_s}$. The first term of

$$\left(\frac{1}{n-1-2\sigma}\right)^2$$

gives the probability of the i–th edge being a double edge. Since we know there are σ isolated edges there are only $n-1-2\sigma$ vertices that a vertex can randomly be drawn "out" to. The second term gives the probability that no other vertex draws an edge "in" to ξ_i. But again since we know there are σ isolated edges, there are only $n-2-2\sigma$ vertices for which an edge can be drawn "in" to one of the 2 vertices of ξ_i. The term

$$\left(1-\frac{2}{n-1-2\sigma}\right)^{n-2-2\sigma}$$

gives the probability that no such edge is drawn to ξ_i. Since σ is fixed the first two factors in the product of (2.14) are

$$\lim_{n\to\infty} n^2\left(\frac{1}{n-1-2\sigma}\right)^2 = 1,$$

while the last factor is

$$\lim_{n\to\infty}\left(1-\frac{2}{n-1-2\sigma}\right)^{n-2-2\sigma} = e^{-2}.$$

We have now shown

$$\lim_{n\to\infty}\Gamma(n,\rho) = \frac{\lambda^\rho}{\rho!},$$

for every fixed ρ, where $\lambda = \frac{e^{-2}}{2}$.

It remains to show that (2.9) also holds. Taking a Taylor series expansion of $e^{\lambda z}$ about $z = 0$ we have

$$e^{\lambda z} = \sum_{\rho=0}^{\infty} \frac{\lambda^\rho}{\rho!} z^\rho,$$

so that (2.9) becomes

$$\lim_{n\to\infty} \sum_{\rho=0}^{\lfloor \frac{n}{2} \rfloor} \Gamma(n,\rho) z^\rho - e^{\lambda z} = \lim_{n\to\infty} \sum_{\rho=0}^{\lfloor \frac{n}{2} \rfloor} \left(\Gamma(n,\rho) - \frac{\lambda^\rho}{\rho!} \right) z^\rho = 0.$$

Setting

$$\gamma(n,\rho) = \Gamma(n,\rho) - \frac{\lambda^\rho}{\rho!}$$

the previous equation becomes

$$\lim_{n\to\infty} \sum_{\rho=0}^{\lfloor \frac{n}{2} \rfloor} \gamma(n,\rho) z^\rho = 0.$$

We take the sum from $\rho = 0$ to $\rho = \lfloor \frac{n}{2} \rfloor$ and not to $\rho = \infty$ since there can be at most $\lfloor \frac{n}{2} \rfloor$ isolated edges. Write

$$\lim_{n\to\infty} \sum_{\rho=0}^{\lfloor \frac{n}{2} \rfloor} \gamma(n,\rho) z^\rho = \lim_{n\to\infty} \sum_{\rho=0}^{N} \gamma(n,\rho) z^\rho + \lim_{n\to\infty} \sum_{\rho=N+1}^{\lfloor \frac{n}{2} \rfloor} \gamma(n,\rho) z^\rho,$$

as $n \to \infty$ the first term on the right converges to zero, since for $\rho < N$ each of the $\gamma(n,\rho)$ can be made arbitrarily small. To show that the second term also converges to zero we note that $\gamma(n,\rho) \leq 1$ for all n and for $N \leq \rho \leq \lfloor \frac{n}{2} \rfloor - 1$, since under these conditions both $\lambda^\rho/\rho! \leq 1$ and $\Gamma(n,\rho) \leq 1$. In particular for $\rho > e\lambda$ we have $\lambda^\rho/\rho! \leq 1$. We now show that $\Gamma(n,\rho) \leq 1$ for all n and for $N \leq \rho \leq \lfloor \frac{n}{2} \rfloor - 1$. Using the same analysis as in (2.12) and (2.14) we have

$$\begin{aligned}
\Gamma(n,\rho) &= \frac{n!}{(n-2\rho)!\rho!2^\rho} \prod_{\sigma=1}^{\rho} \Pr\left[\xi_{i_\sigma} \mid \bigwedge_{s<\sigma} \xi_{i_s}\right] \\
&= \frac{n!}{(n-2\rho)!\rho!2^\rho} \prod_{\sigma=1}^{\rho} \left[\left(\frac{1}{n-1-2\sigma}\right)^2 \left(1-\frac{2}{n-1-2\sigma}\right)^{n-2-2\sigma}\right] \\
&< \frac{n!}{(n-2\rho)!\rho!2^\rho} \left(\frac{1}{n-1-2\rho}\right)^{2\rho} e^{-2\rho} \\
&< \frac{e^\rho}{\rho^\rho 2^\rho} \left(\frac{n}{e(n-2\rho)}\right)^{2\rho} \\
&= \left(\frac{n}{\sqrt{2e\rho}(n-2\rho)}\right)^{2\rho} \\
&= \left(\frac{n}{\sqrt{2e}\rho^{1/2}n - 2\sqrt{2e}\rho^{3/2}}\right)^{2\rho} < 1 \qquad (2.15)
\end{aligned}$$

if $n < \sqrt{2e\rho}n - 2\sqrt{2e}\rho^{3/2}$ or $\rho < \frac{n}{2}$. We thus rewrite the second term as

$$\lim_{n\to\infty} \sum_{\rho=N+1}^{\lfloor \frac{n}{2} \rfloor} \gamma(n,\rho) z^\rho \leq \lim_{n\to\infty} \sum_{\rho=N+1}^{\lfloor \frac{n}{2} \rfloor} z^\rho = \lim_{n\to\infty} \frac{z^{N+1}}{1-z}.$$

By Remark 2.4 we have $|z| < 1$, thus the last term can be made smaller than any $\epsilon > 0$ by choosing

$$N > \max\left[e\lambda, \ \frac{\log\left[\epsilon(1-z)\right]}{\log z}\right].$$

We have shown X to be asymptotically Poisson with mean $\frac{e^{-2}}{2}$. Thus, with probability $1-\exp(-e^{-2}/2)$ a random graph $G \in \mathcal{G}_{1\text{-out}}$ has an isolated edge. Now, if $r(n) < 2$ (i.e. $p = p(n) < 1$), then this edge remains isolated in $\mathcal{G}_{r\text{-out}}$ with probability

$$(1-p)^2\left(1-\frac{2p}{n-1}\right)^{n-2} \to (1-p)^2 e^{-2p} > 0,$$

which can be bounded from below by a positive number, unless $r(n) \to 2$ (i.e. $p(n) \to 1$). This proves the first statement of Theorem 2.5.

It remains to show if $q(n) < \frac{2}{n-2}$ then $\mathcal{G}_{r\text{-out}}$ is a.s. connected. As in the proof of Theorem 2.2, we find a bound to $\Pr[G_{i\text{-disconn}}]$, the probability that a random $G \in \mathcal{G}_{r\text{-out}}$ is disconnected with a component of order i. We count the probability of a subset of vertices of order i having "out" edges only to itself and having j vertices with out-degree 2 and $i-j$ vertices with out-degree 1, summed over all possible j. We proceed in the same way for the complement of order $n-i$. This gives

$$\Pr[G_{i\text{-disconn}}] \leq \binom{n}{i}\sum_{j,\ell}\binom{i}{j}\binom{n-i}{\ell}p^j q^{i-j} p^\ell q^{n-i-\ell}. \tag{2.16}$$

$$\frac{\binom{i-1}{2}^j\binom{i-1}{1}^{i-j}}{\binom{n-1}{2}^j\binom{n-1}{1}^{i-j}}\frac{\binom{n-i-1}{2}^\ell\binom{n-i-1}{1}^{n-i-\ell}}{\binom{n-1}{2}^\ell\binom{n-1}{1}^{n-i-\ell}} = \binom{n}{i}\frac{A_i B_i}{(n-1)^n},$$

where

$$A_i = \sum_{j=0}^{i}\frac{\binom{i}{j}p^j q^{i-j}\binom{i-1}{2}^j (i-1)^{i-j}}{\binom{n-2}{2}^j}$$

$$= (i-1)^i\left[q + \frac{p(i-2)}{n-2}\right]^i$$

and

$$B_i = \sum_{\ell=0}^{n-i} \frac{\binom{n-i}{\ell} p^\ell q^{n-i-\ell} \binom{n-i-1}{2}^\ell (i-1)^{n-i-\ell}}{\binom{n-2}{2}^\ell}$$

$$= (n-i-1)^{n-i}\left[q + \frac{p(n-i-2)}{n-2}\right]^{n-i}.$$

Now X, the number of isolated edges in $\mathcal{G}_{1\text{-out}}$ is Poisson with mean $\lambda = e^{-2}/2$. Thus for large s

$$\Pr\left[X > s\right] \leq 2e^{-\lambda}\frac{\lambda^s}{s!}.$$

Furthermore, the probability that at least one out of s fixed isolated edges remains isolated after the addition of extra edges to $\mathcal{G}_{1\text{-out}}$ is $\leq s(1-p)^2$. Thus, no matter how slowly $p = p(n)$ tends to 1, we can find $s = s(n) = o\big(1/(1-p)^2\big) \to \infty$ such that the last probability as well as $\Pr[X > s] \to 0$. In other words, a.s. $\mathcal{G}_{r\text{-out}}$ will have no isolated edges. Therefore, it is sufficient to consider the sum in (2.2), starting with $i = 3$.

$$\Pr[G_{\text{disconn}}] \leq \frac{1}{(n-1)^n}\sum_{i=3}^{\lfloor \frac{n}{2} \rfloor}\binom{n}{i}(i-1)^i\left[q + \frac{p(i-2)}{n-2}\right]^i.$$

$$(n-i-1)^{n-i}\left[q + \frac{p(n-i-2)}{n-2}\right]^{n-i} + o(1)$$

$$\leq \frac{1}{(n-1)^n} \sum_{i=3}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{i} (i-1)^i (n-i-1)^{n-i} \left[\frac{q(n-2)+i-2}{n-2} \right]^i + o(1).$$

For $q < \frac{2}{n-2}$ we have

$$\Pr[G_{\text{disconn}}] \leq \frac{1}{(n-1)^n} \sum_{i=3}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{i} (i-1)^i (n-i-1)^{n-i} \left[\frac{i}{n-2} \right]^i + o(1),$$

$$\leq \frac{1}{(n-1)^n} \sum_{i=3}^{\lfloor \frac{n}{2} \rfloor} (i-1)^i (n-i-1)^{n-i} = o(1),$$

by the same argument as in (2.5) to (2.8). We have shown $\Pr[G_{\text{disconn}}] \to 0$ if $q < \frac{2}{n-2}$. □

Remark 2.6. *Bollobás [6] has suggested methods for biasing the probability of a second edge so that we have a.s. connectedness in a model that has minimally more edges than $\mathcal{G}_{1\text{-out}}$. Some ideas for study are to determine the effects on a.s. connectedness of sending an edge from each component of $\mathcal{G}_{1\text{-out}}$ that is a cycle and from each vertex of degree one. A second idea is to send an edge from one of the vertices of every edge of total degree at least four.*

2.5 On Higher Connectivity of $\mathcal{G}_{k\text{-out}}$

A result of Erdős and Rényi [12] roughly states that a random graph on n vertices with $\frac{1}{2}n \log n + O(n)$ edges has a.s. connectivity equal to its minimum degree. Bollobás and Thomason [7] showed that for every function

$k = k(n) \leq n - 1$, the random graph models $\mathcal{G}_{n,p}$ and $\mathcal{G}_{n,M}$ have a.s. k–connectivity, where k is the minimum degree of the graph. Despite the fact that $\mathcal{G}_{k\text{-out}}$ has at most kn edges we can give similar connectivity results here as well.

A graph is k–vertex (edge)–connected if deletion of any $k - 1$ vertices (edges) will not disconnect the graph. As a corollary to Menger's theorem [42] one can show that a graph is k–vertex–connected if any two vertices can be joined by k independent paths, i.e. paths whose only vertices in common are the endpoints. One can also show as a corollary, that a graph is k–edge–connected if any two vertices can be joined by k edge-disjoint paths.

We say that a graph has vertex (edge)–connectivity k if k is the largest integer for which the graph is k–vertex (edge)–connected. It is clear that the minimum degree of a graph represents an upper bound on its edge and vertex connectivity. For deletion of all edges of a minimal degree vertex will render that vertex isolated and the graph disconnected in both the edge and vertex sense. A result of Fenner and Freize [20] which we quote here gives that this upper bound is also a lower bound.

Theorem 2.7. *(Fenner and Frieze 1982) Let $C_v(G)$ and $C_e(G)$ be the vertex and edge–connectivity (resp.) of a graph $G \in \mathcal{G}_{k\text{-out}}$ for $k \geq 2$. Then*

$$\lim_{n \to \infty} \Pr\left[C_v(G) = k\right] = 1$$

and

$$\lim_{n\to\infty} \Pr\left[C_e(G) = k\right] = 1.$$

We prove now a corollary to Theorem 2.7 relating to graph planarity. A graph G is said to be *planar* if there exists a representation of G in the plane, such that no two edges of G intersect. Kuratowski's Theorem [39] states that a graph G is not planar if and only if there is a K_5 or $K_{3,3}$ is imbedded G, i.e. G has a complete subgraph on 5 vertices or a complete bipartite graph with vertex sets of order 3 and 3. Wagner in an extension of Kuratowski's result [53] proves that a graph G is planar if and only if it is not contractible to K_5 or $K_{3,3}$. By *contractible* it is meant that a series of edge contractions exists, that is the identification of two vertices at the ends of an edge. The graph resulting from a series of edge contractions is said to be a *minor* of the original graph.

The existence of a set of 5 vertices of $V(G)$ such that all the vertices are connected to each other via mutually independent paths is equivalent to K_5 being a contraction of G. For first, contract all edges not on one of these independent paths and then contract each independent path to a single edge. The existence of such a contraction would render G non-planar.

Corollary 2.8. *For $k \geq 4$ $G \in \mathcal{G}_{k\text{-out}}$ is a.s. non-planar.*

Proof. By Theorem 2.7 $\mathcal{G}_{3\text{-out}}$ is a.s. 3–connected. By a well known theorem of Steinitz [2] the embedding of any 3-connected planar graph in the plane is essentially unique, i.e., the facets are uniquely determined. Add a random edge from each vertex to obtain $\mathcal{G}_{4\text{-out}}$. The probability that none of the additional edges connects two vertices in different cells is extremely small. The details are omitted, since in Theorem 2.13 we establish a stronger result.

□

Another related result of Fenner and Frieze [21] says that for $k \geq k_0$, $\mathcal{G}_{k\text{-out}}$ is a.s. Hamiltonian. Bollobás [4] shows that their theorem holds for $k_0 = 23$.

2.6 On the Degree Sequence Distribution of $\mathcal{G}_{k\text{-out}}$

Many results are known for random regular graphs [56] and for graphs whose degree sequence is known to have fixed upper and lower bounds. In this section we study some properties of the degree sequence of $\mathcal{G}_{k\text{-out}}$. We first show that the in-degree of any vertex in $\mathcal{G}_{k\text{-out}}$ converges to a Poisson distribution with mean k, as the number of vertices, $|V| = n \to \infty$.

Let Y_k be the random variable on $\mathcal{G}_{k\text{-out}}$ that gives the in-degree of a vertex $v \in V$. Of course by definition the out-degree of each vertex is k. Clearly Y_k is independent of v and so we do not include this in the notation of the random variable.

Proposition 2.9. *For fixed k, Y_k converges uniformly to the Poisson distri-*

bution with mean k as $n \to \infty$. That is, for $0 \leq j \leq n-1$

$$\lim_{n \to \infty} \Pr[Y_k = j] = e^{-k} \frac{k^j}{j!}.$$

Proof. Consider an oriented edge uv from $u \in V - \{v\}$ a "success" of a *Bernoulli* trial with probability of success $p = \frac{k}{n-1}$. For any $n > k$ the in-degree of v is the sum of independent successes and is thus binomially distributed. Since pn converges monotonically for $n \to \infty$ to k we have the following inequalities (see [19], p. 161)

$$e^{-k-\frac{j^2}{n-j}-\frac{k^2}{n-k}} \frac{k^j}{j!} < \Pr[Y_k = j] < e^{-k+\frac{k^2(n+1)}{n(n-1)}} \frac{k^j}{j!}. \tag{2.17}$$

□

Theorem 2.10. *For constant k and $G \in \mathcal{G}_{k\text{-}out}$ let $\Delta(G) = \max_{v \in G}$ in$-\deg(v)$. Then a.s.*

$$\Delta(G) = \Theta\left(\frac{\log n}{\log \log n}\right).$$

Proof. Denote by $\xi_{v,j}$, $\psi_{v,j}$ and $\phi_{v,j}$ the events of vertex $v \in V(\mathcal{G}_{k\text{-out}})$ having in-degree precisely j, $< j$, and $\geq j$ respectively. Now knowledge that some vertex $v \in V(\mathcal{G}_{k\text{-out}})$ has in-degree $< j$ can not increase the probability that some other vertex $u \in V(\mathcal{G}_{k\text{-out}})$ has in-degree $< j$. For each edge that does not connect "in" to v may potentially connect "in" to u thereby raising the total in-degree of u and thus lower the probability that u has in-degree less than some number j. Thus we have for $u \neq v$

$$\Pr\left[\psi_{u,j} \,\middle|\, \psi_{v,j}\right] \leq \Pr\left[\psi_{u,j}\right].$$

Let N_j be the number of vertices with in-degree $\geq j$. By the same reasoning this inequality holds for sets as well. Fix a set $S \subset V$ with $u \notin S$. Denote by T_S the event that every vertex in S has in-degree $< j$, then

$$\Pr\left[\psi_{u,j} \,\middle|\, T_S\right] \leq \Pr\left[\psi_{u,j}\right]. \tag{2.18}$$

By repeated application of (2.19) and (2.18) we have

$$\begin{aligned}
\Pr\left[N_j > 0\right] &= 1 - \Pr\left[N_j = 0\right] \\
&= 1 - \Pr\Big[\bigwedge_{v \in V} \psi_{v,j}\Big] \\
&\geq 1 - \prod_{v \in V} \Pr\left[\psi_{v,j}\right].
\end{aligned}$$

By Proposition 2.9, we have from the upper bound given by (2.17) that

$$\lim_{n \to \infty} \Pr[\psi_{v,j}] = \lim_{n \to \infty} \sum_{i=0}^{j-1} \Pr[\xi_{v,j}] < e^{\frac{k^2(n+1)}{n(n-1)}}\, e^{-k} \sum_{i=0}^{j-1} \frac{k^i}{i!}. \tag{2.19}$$

Thus,

$$\lim_{n\to\infty} \Pr\left[N_j > 0\right] > 1 - \lim_{n\to\infty} \left(e^{\frac{k^2(n+1)}{n(n-1)}} e^{-k} \sum_{i=0}^{j-1} \frac{k^i}{i!} \right)^n$$
$$\geq 1 - \lim_{n\to\infty} \left(1 - e^{-k} \frac{k^j}{j!}\right)^n, \tag{2.20}$$

But the last term in (2.20) tends to zero as $n \to \infty$ provided $j \ll \frac{\log n}{\log\log n}$. Using Stirling's approximation [49] on the last term in (2.20) we have

$$\left(1 - e^{-k} \frac{k^j}{j!}\right)^n \leq \left(1 - \frac{c_1^j}{j^{j+1/2}}\right)^n \leq \left(1 - \left(\frac{c_2}{j}\right)^j\right)^n,$$

where c_1 and c_2 are constants $\leq 2ek$. Now, let $s(n) \to \infty$ as $n \to \infty$, but arbitrarily slowly. Then

$$\left(\frac{c_2}{j}\right)^j \geq \frac{s(n)}{n} \Rightarrow \left(1 - \left(\frac{c_2}{j}\right)^j\right)^n \to 0. \tag{2.21}$$

But

$$\lim_{n\to\infty} n \frac{c^j}{j^j} = \lim_{n\to\infty} \frac{\log n + j\log(c)}{j\log(j)} = \infty$$

for $j \ll \frac{\log n}{\log\log n}$.

We now show that $\Pr[N_j > 0] = 0$ for $j \gg \frac{\log n}{\log\log n}$. Since N_j is the number of vertices with in-degree $\geq j$ we have

$$\lim_{n\to\infty} \Pr[N_j > 0] \leq \lim_{n\to\infty} n \Pr[\phi_{v,j}]$$

$$\leq \lim_{n\to\infty} n\, e^{-k+\frac{k^2(n+1)}{n(n-1)}} \sum_{i=j}^{\infty} \frac{k^i}{i!} \leq \lim_{n\to\infty} 2ne^{-k}\frac{k^j}{j!},$$

since the first summand is dominating. The right hand side tends to zero for $j \gg \frac{\log n}{\log\log n}$ since

$$\lim_{n\to\infty} n\frac{k^j}{j!} \leq \lim_{n\to\infty} n\frac{(2k)^j}{j^j} = \lim_{n\to\infty} \frac{\log n + j\log(2k)}{j\log(j)} = 0$$

for $j \gg \frac{\log n}{\log\log n}$. □

2.7 On the Planarity of $\mathcal{G}_{k\text{-out}}$

Recall that a graph is planar if it can be represented in the plane, such that no two edges intersect. In this section[3] we prove the threshold for a.s. non-planarity is $k = 2$, i.e. for $k \geq 2$ $G \in \mathcal{G}_{k\text{-out}}$ is a.s. non-planar. We also prove, in the $\mathcal{G}_{r\text{-out}}$ model that for $r = 1 + p$ and $p = p(n)$ very small, that $G \in \mathcal{G}_{k\text{-out}}$ is a.s. non-planar.

Proposition 2.11. *Every $G \in \mathcal{G}_{1\text{-out}}$ is planar.*

Proof. By construction every connected component of any $G \in \mathcal{G}_{1\text{-out}}$ contains at most one cycle, see the proof to Theorem 2.1. Since all unicyclic graphs are planar, each component of G is planar and thus G is planar. □

We have previously noted with Corollary 2.8 that for $k \geq 4$, $\mathcal{G}_{k\text{-out}}$ is a.s. non-planar. The well known result attributable to Euler (see for example [5]

[3]My thanks to Dr. Jeffrey Jaffe Head of Research and Advanced Technologies at Bell Labs for raising the question of planarity in these graphs.

Theorem 16) stating that a planar graph on $n > 3$ vertices with more than $3n - 6$ edges is non-planar by itself is not sufficient to establish that $\mathcal{G}_{k\text{–out}}$ is a.s. non-planar for $k = 3$. There are at most $3n$ edges in this model but in general there will be less than that due to the existence of double edges. Recall that uv is a *double edge* if u has an "out" edge and an "in" edge adjacent to v. The expected number of double edges is

$$\binom{n}{2}\left(\frac{3}{n-1}\right)^2 \sim 4.5$$

and there may be non-zero probability that the number of double edges exceeds 6.

Before proving that indeed we have a.s. non-planarity for $k \geq 3$ we state and prove a lemma relating to the planarity of multigraphs. A *multigraph* is a graph with more than one edge connecting any two vertices.

Lemma 2.12. *Let G be a multigraph that does not contain any of the forbidden (not necessarily induced) subgraphs of Figure 2.7. Then contraction of all multiple edges of G results in a graph H with no multiple edges.*

Proof. We characterize the forbidden subgraphs as follows: any double edge is not part of a triangle and any two double edges are separated by at least two other edges. Assume now that G does not possess any of the forbidden subgraphs, then any cycle containing 2 double edges must have length at least 6 and any cycle containing one double edge must have length at least 4. Contraction of the double edges on any cycle results in a cycle of length

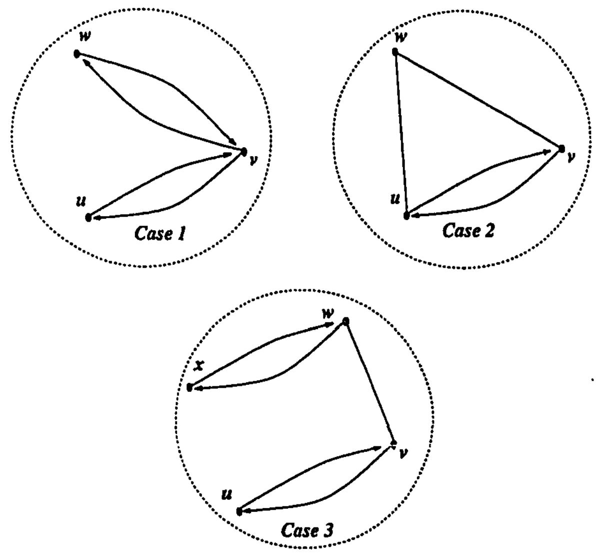


Figure 2.7: Forbidden subgraphs

at least 3, but a cycle of length 2 is needed to create a double edge. □

Theorem 2.13. *Let $G \in \mathcal{G}_{k\text{-out}}$ be a random graph with $k \geq 3$. Then G is a.s. non-planar.*

Proof. We address the case of $k = 3$ with higher k obvious from the definition of the $\mathcal{G}_{k\text{-out}}$ model. One can think of generating a random $G \in \mathcal{G}_{k\text{-out}}$ by first generating a random $G' \in \mathcal{G}_{k-1\text{-out}}$ and then adding random edges.

Let d be the number of double edges of G. We will show the existence of a graph minor H of G such that H has a.s. $m = n - d$ vertices and $3n - 2d > 3m$ edges, where here we count a double edge only once. From this we have that H is a.s. non-planar by the Euler condition mentioned above. Let H be the minor of G obtained by contraction of all double edges of G. In the event contraction of a double edge creates a new double edge, the new double edge must also be contracted.

If there are no forbidden subgraphs of Figure 2.7 in G, then H has $n-d$ vertices and $3n-2d$ edges (counting double edges once). Starting with n vertices and $n-d$ edges in G, since there are d contractions of edges in G and each contraction eliminates one vertex and one edge. We now show that a.s. there are no forbidden subgraphs. The probability of Case 1 is bounded from above by

$$\binom{n}{3}\left(\frac{6}{n-1}\right)^2\left(\frac{5}{n-1}\right)^2 = O\left[\frac{1}{n}\right] \to 0,$$

similarly the probability of Case 2 is bounded above by

$$\binom{n}{3}\left(\frac{6}{n-1}\right)^2\left(\frac{6}{n-1}\right)^2 = O\left[\frac{1}{n}\right] \to 0.$$

The probability of Case 3 is bounded above by

$$\binom{n}{4}\left(\frac{6}{n-1}\right)^2\left(\frac{5}{n-1}\right)^2\left(\frac{4}{n-1}\right) = O\left[\frac{1}{n}\right] \to 0.$$

□

Using the same model as in Theorem 2.5 we prove that fewer than 2 "out" edges per vertex are needed for a.s. non-planarity.

Lemma 2.14. *Let $s = s(n)$, $c = c(n)$ and $p = p(n)$ be functions such that*

$s \to \infty$, $sp < n$ *and* $c > e$. *If*

$$\frac{s^2 p}{\log c} \gg n$$

then

$$\lim_{n \to \infty} c^2 \left(1 - \frac{sp}{n}\right)^s = 0.$$

Proof.

$$\lim_{n \to \infty} c^2 \left(1 - \frac{sp}{n}\right)^s = \lim_{n \to \infty} c^2 \left[\left(1 - \frac{sp}{n}\right)^n\right]^{\frac{s}{n}}$$

$$= \lim_{n \to \infty} c^2 \exp\left(-sp\right)^{\frac{s}{n}} = \lim_{n \to \infty} c^2 \exp\left(\frac{s^2 p}{n}\right)$$

$$= \lim_{n \to \infty} \exp\left(2 \log c - \frac{s^2 p}{n}\right) = 0$$

when the given condition is satisfied. □

Definition 2.15. *Let T be a tree and let S be a subtree formed by deletion of an edge $e_S \in E(T)$. We say that T is pruned into the subtrees S and $T - S$ by deletion of edge e_S.*

Lemma 2.16. *Let T be a tree on m vertices with maximum degree Δ and let $c < m/\Delta$. Then T can be pruned into c subtrees each containing at least $\lfloor \frac{m}{\Delta c} \rfloor$ vertices.*

Proof. We describe an algorithm for pruning T. Assign a root vertex r to T and orient all edges toward r. For each vertex $v \neq r$, let T_v be the pruned subtree that results from deletion of the edge adjacent to v in the direction of r. Let $t(v) = |T_v|$. We prune T as follows.

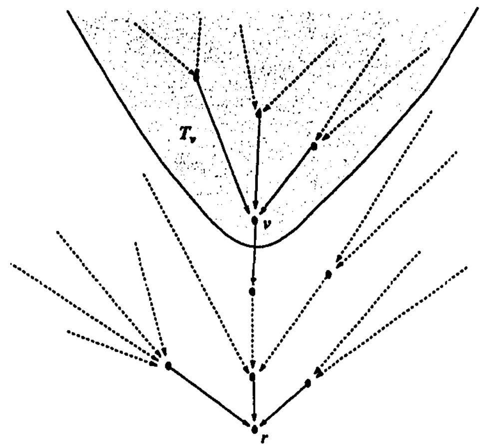


Figure 2.8: Pruning T at vertex v

1. Starting at a leaf, take successive adjacent vertices, proceeding toward r until we reach the first vertex v such that $t(v) \geq \frac{m}{\Delta c}$.

2. If $t(v) < \frac{m}{c} + 1$ then prune T_v and repeat Step 1 for $T - T_v$, unless $|T - T_v| < \frac{m}{\Delta c}$. See Figure 2.8.

3. If $t(v) \geq \frac{m}{c} + 1$, then, since $\deg(v) \leq \Delta$, there is an edge uv toward vertex u adjacent to v and oriented toward v such that $t(u) \geq \frac{m}{\Delta c}$. Substitute vertex u for vertex v and return to Step 2.

Note that,since $t(u) < t(v)$, the process will eventually end. Since each pruning removes at most $\frac{m}{c}$ vertices from T, the pruning step (Step 2) can be repeated at least $c - 1$ times, resulting in at least c trees meeting the requirement. □

Theorem 2.17. *Let $c = c(n)$ be an integer, $p = p(n)$ be a probability and*

$G \in \mathcal{G}_{r\text{-}out}$ be a random graph with $r(n) = 1 + p(n)$. If

$$\frac{p}{c^2 \log c} \gg \frac{\log^4 n}{n(\log\log n)^2}$$

then G a.s. has a minor K_c.

Proof. We generate a random $G \in \mathcal{G}_{r\text{–out}}$ by first choosing a random $H \in \mathcal{G}_{1\text{–out}}$ and then adding a random edge "out" of each vertex with probability $p(n)$. By a result of Kolchin [37] the number of connected components of H is asymptotically normally distributed with mean $\frac{1}{4}\log n$ and variance similarly $\frac{1}{4}\log n$. Thus, there are a.s. $O(\log n)$ connected components of H and at least one component C of order $m = \Omega(n/\log n)$, i.e. there exists a constant α such that $m \geq \alpha n/\log n$.

Since C is connected it has a spanning tree T with $m = \Omega(n/\log n)$ vertices. By Proposition 2.10 the maximum degree of T is $O(\frac{\log n}{\log\log n})$. Applying Lemma 2.16, there is a pruning of T into c subtrees, $T_1, T_2, \ldots, T_c$, each containing at least

$$s(n) = \Omega\left(\frac{n \log\log n}{c \log^2 n}\right)$$

vertices. Now add the random "out" edges with probability p. The probability of an "out" edge between some specific vertex $v \in T_i$ and any vertex in T_j is at least $ps/(n-1)$ and the probability of no such edge is at most $(1 - ps/(n-1))$. Thus, the probability of no "out" edges from any vertex of

T_i to any vertex of T_j is at most

$$\left(1 - \frac{ps}{n-1}\right)^s$$

and the probability of no edges between any of the trees $T_1, T_2, \ldots, T_c$ is at most

$$c^2\left(1 - \frac{ps}{n-1}\right)^s.$$

But for

$$\frac{p}{c^2 \log c} \gg \frac{\log^4 n}{n(\log\log n)^2}$$

the conditions of Lemma 2.14 are satisfied, and thus in the limit the probability that there is a pair (T_i, T_j) with no connecting edges is at most

$$\lim_{n\to\infty} c^2\left(1 - \frac{sp}{n}\right)^s = 0.$$

We have shown the a.s. existence of an edge between every pair of subtrees T_i and T_j. Contraction of each T_i to a point gives K_c. □

Corollary 2.18. *Let $G \in \mathcal{G}_{k\text{-}out}$ be a random graph with $k \geq 2$. If*

$$\frac{c}{\sqrt{\log c}} \ll \frac{\sqrt{n}\log\log n}{\log^2 n}$$

then G a.s. has a complete minor K_c.

Proof. Setting $p = 1$ in Theorem 2.17 we have that

$$1 \gg \frac{c^2 \log^4 n}{\log cn\left(\log\log n\right)^2},$$

which is implied by the condition. □

Corollary 2.19. *Let $G \in \mathcal{G}_{k\text{-}out}$ be a random graph with $k \geq 2$. Then G is a.s. non-planar.*

Proof. By the previous corollary there is a.s. a minor K_c for $c = 5$ (in fact much larger than $c = 5$), which by Kurotowski's Theorem [39] implies non planarity. □

Corollary 2.20. *Let c be a fixed integer and $G \in \mathcal{G}_{r-out}$ be a random graph with $r = 1 + p$. If*

$$p \gg \frac{\log^4 n}{n\left(\log\log n\right)^2}.$$

then G a.s. has a K_c and is thus a.s. non-planar.

Proof. Fix c in Theorem 2.17. □

Chapter 3

An Application of Graph Theory to Number Theory

This chapter gives an application of the $\mathcal{G}_{k\text{-out}}$ model of random graphs to a problem in analytic number theory. In attempting to find a new proof for Theorem 2.1 we generated a sum over the ordered partitions of an integer and reduced the proof of the theorem to finding the asymptotic nature of this sum.

Recent methods devised in analytic number theory [22] apply to the problem we present here. These however, require sophisticated techniques in complex analysis. We present a different line of proof using combinatorial methods, which we note is of interest due to its relative simplicity.

3.1 Introduction

Definition 3.1. *A* composition *of n is an ordered partition of n, i.e. a representation of n as the sum of positive integers with regard to order in the sum.*

For instance, 4 has the following eight compositions

$$4 = 3+1 = 1+3 = 2+2 = 2+1+1 = 1+2+1 = 1+1+2 = 1+1+1+1.$$

A *k-part composition* is an ordered partition consisting of exactly k terms. We denote by $\gamma_k(n)$ the set of all k-part compositions of n and by $\gamma(n)$ the set of all compositions of n.

Examples of weighted sums over compositions have appeared in the literature (see [44, 28, 35]). Let $w = (w_1, w_2, w_3, \cdots)$ be a sequence of numbers called weights. The general problem is to find exact formulas or at least good asymptotic estimates for sums of the following type:

$$c_{n,k}(w) = \sum_{\gamma_k(n)} w_{t_1} w_{t_2} \cdots w_{t_k} \tag{3.1}$$

and

$$c_n(w) = \sum_{k=1}^{n} c_{n,k}(w) = \sum_{\gamma(n)} w_{t_1} w_{t_2} \cdots , \tag{3.2}$$

where the first sum is taken over all k-part compositions $t_1 + t_2 + \cdots + t_k$ of n, while the second one is taken over all compositions of n.

So far, estimates for sums as above were obtained by using the power series generating function of w namely,

$$W(x) = \sum_{n=1}^{\infty} w_n \, x^n.$$

To obtain an explicit representation, notice that

$$1 + \sum_{k=1}^{\infty} [W(x)]^k \, y^k = \frac{1}{1 - y\,W(x)} = 1 + \sum_{k=1}^{\infty} \sum_{n=k}^{\infty} c_{n,k}(w) x^n \, y^k. \tag{3.3}$$

As an example, let $w_n = 1, \forall n \geq 1$, so that $W(x) = x + x^2 + \cdots = x/(1-x)$ and replacing in (3.3)

$$\begin{aligned} 1 + \frac{yW(x)}{1 - yW(x)} &= 1 + \frac{xy}{1 - x(1+y)} = 1 + xy \sum_{n=0}^{\infty} x^n (1+y)^n \\ &= 1 + \sum_{n=0}^{\infty} \sum_{k=0}^{n} \binom{n}{k} x^{n+1} y^{k+1} = 1 + \sum_{n=1}^{\infty} \sum_{k=1}^{n} \binom{n-1}{k-1} x^n y^k, \end{aligned} \tag{3.4}$$

so that

$$c_{n,k}(w) = \sum_{\gamma_k(n)} 1 = \binom{n-1}{k-1}$$

is the number of compositions of n into k parts. This appears in [48] and can also be proved by a simple combinatorial argument. It follows that in this case

$$c_n(w) = \sum_{\gamma(n)} 1 = \sum_{k=1}^{n} c_{n,k}(w) = \sum_{k=1}^{n} \binom{n-1}{k-1} = 2^{n-1}$$

is the total number of compositions of n. For some other similar examples the reader is referred to [28, 44].

However, it is not always the case that a closed formula can be obtained. Knopfmacher and Ridley [35] considered the case

$$w_n = \frac{1}{n}, \quad \forall n \geq 1.$$

It is known (see [33], p.146) that in this case

$$c_{n,k}(w) = \frac{k!}{n!} s(n,k) \tag{3.5}$$

where $s(n,k)$ is the signless Stirling number of the first kind. The asymptotic behavior of $s(n,k)$ has been studied in [45, 52, 54]; for k fixed and $n \to \infty$ it is known that

$$s(n,k) \sim \frac{(n-1)!}{(k-1)!} (\ln n)^{k-1}.$$

Using the above estimate in (3.5) we obtain that

$$c_{n,k}(w) = \sum_{\substack{t_1+t_2+\cdots+t_k=n \\ t_j \geq 1,\, 1 \leq j \leq k}} \frac{1}{t_1 t_2 \cdots t_k} \sim \frac{k}{n} (\ln n)^{k-1}. \tag{3.6}$$

In [35], the authors prove that in this case

$$c_n(w) = \sum_{k=1}^{n} c_{n,k}(w) = \frac{1}{e-1}\left(\frac{e}{e-1}\right)^n + O(1). \tag{3.7}$$

In attempting to find a new proof for a well known theorem from graph theory (Theorem 2.1) we were lead to a similar sum over the k-part compositions of an integer. More precisely, we had to estimate $c_{n,k}(w)$ when $w_n = n^{-p}, \quad \forall n \geq 1$, where p is some constant ≥ 1. Being unable to successfully estimate this sum, we proceeded to prove the theorem in an alternate manner and then show that the theorem itself may be used to describe the order of magnitude of this sum.

It is natural to consider the more general problem of estimating $c_{n,k}(w)$ when $w_n = n^{-p}, \quad \forall n \geq 1$, where p is some constant ≥ 1. Here is our main result:

Theorem 3.2. *Given $p \geq 3/2$ and n, k positive integers so that $k \leq n-1$, then*

$$C(n,k,p) := \sum_{\gamma_k(n)} \frac{1}{(t_1 t_2 \cdots t_k)^p} = \sum_{\substack{t_1+t_2+\cdots+t_k=n \\ t_j \geq 1,\ 1\leq j\leq k}} \frac{1}{(t_1 t_2 \cdots t_k)^p} \leq \frac{k\,a^k}{\sqrt{2\pi}}(n-k)^{-p},$$
$$\text{where} \quad a = e^{1/12}\sqrt{2\pi} = 2.7244\ldots. \tag{3.8}$$

Remark. It is obvious that

$$C(n,k,p) = \sum_{\substack{t_1+t_2+\cdots+t_k=n \\ t_j\geq 1,\ 1\leq j\leq k}} \frac{1}{(t_1 t_2 \cdots t_k)^p} \geq k n^{-p}, \quad \forall p. \tag{3.9}$$

For take the k-part compositions of n where $k-1$ of the terms are equal to 1 and the remaining one equals $n-k+1$. There are k such compositions which already gives that $C(n,k,p) \geq k\,(n-k+1)^{-p} \geq k\,n^{-p}$.

From this and the above theorem it follows that if we fix p and k (k positive integer and $p \geq 3/2$) and let $n \to \infty$ we have that

$$C(n,k,p) = \Theta(n^{-p}). \tag{3.10}$$

On the other hand, relation (3.6) shows that n^{-p} is not the correct order of magnitude when $p = 1$. We suspect that our result remains valid for all values of $p > 1$. Since the theorem is obvious for $k = 1$, in the sequel we will assume $k \geq 2$.

In the following section we mention a few well-known facts related to the graph enumeration problems of the previous chapter. These facts will be needed in the proof of the theorem.

3.2 Counting Connected Graphs

Denote by $f(n,e)$ the number of connected graphs having n labelled vertices and e edges. It is trivial that $f(n,e) = 0$ if $e < n-1$ and it has long been known [9] that $f(n, n-1) = n^{n-2}$.

Formulas for $f(n,n)$ have been found by Katz [34], Renyi [47] and Wright [57]. An important observation is that any connected labelled graph with n vertices and n edges contains just one cycle of k points where $k \geq 1$ (we allow loops and multiple edges), and each vertex in the cycle is a root of a rooted tree (possibly the degenerate tree consists of a single vertex). Figure 3.1 illustrates this characterization, with trees labelled $T_1, T_2, \cdots, T_q$.

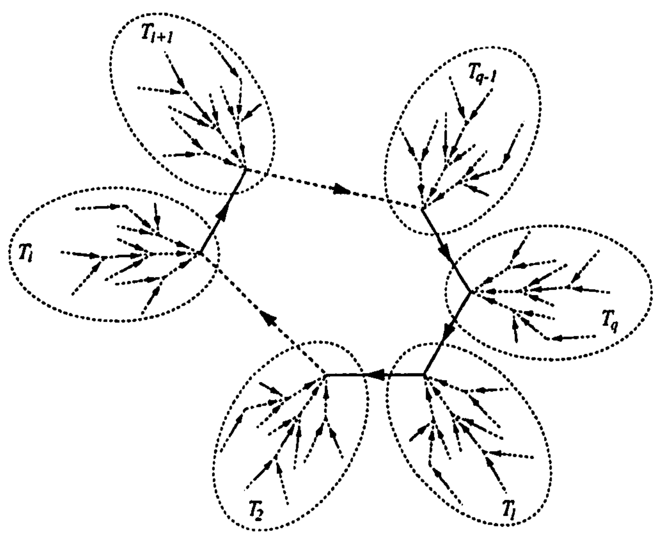


Figure 3.1: A Connected Graph from $\mathcal{G}_{1\text{-out}}$

If the edges of the cycle are deleted from such a graph there remains a graph consisting of k disjoint rooted trees, or in other words a planted forest with k components. The number $p_k(n)$ of planted forests has been determined by Cayley (see e.g. [50] p. 23-26 for a nice Prüfer type proof);

$$p_k(n) = k\, n^{n-k-1} \binom{n}{k}.$$

Let $f_k(n,n)$ be the number of connected graphs on n labelled vertices, having n edges and a k-cycle. Then

$$f_k(n,n) = p_k(n)\, \frac{(k-1)!}{2} = \frac{n!\, n^{n-k-1}}{2\,(n-k)!} \tag{3.11}$$

because the planted forest with k components can be chosen in $p_k(n)$ different ways and the roots can be arranged in a cycle in $(k-1)!/2$ different ways. This last relation will be useful in the sequel.

3.3 Proof for $p = 3/2$

We use an alternate way for counting the unicyclic labelled graphs on n vertices, containing a k-cycle, $2 \le k \le n-1$. Partition first $V(G)$ into k sets T_i, each of cardinality $|T_i| = t_i \ge 1$. The k-cycle will have one vertex in each of these subsets. Then cyclically order these sets; this can be done in

$$\frac{1}{2k} \binom{n}{t_1\ t_2\ \cdots\ t_q} \tag{3.12}$$

ways. Each set T_i must be the vertex set of a tree. By Cayley's theorem [9] we form these trees in

$$\prod_{i=1}^{k} t_i^{t_i-2} \tag{3.13}$$

ways. For each tree we choose a root vertex which is to become part of the k-cycle. This is done in

$$\prod_{i=1}^{k} t_i \tag{3.14}$$

ways. Summing over the k-part compositions the product of the factors in (3.12), (3.13) and (3.14) we have the following equation for $f_k(n,n)$:

$$f_k(n,n) = \sum_{\gamma_k(n)} \frac{1}{2k} \binom{n}{t_1\; t_2\; \cdots\; t_k} \prod_{i=1}^{k} t_i \prod_{i=1}^{k} t_i^{t_i-2} = \frac{n!}{2k} \sum_{\gamma_k(n)} \frac{\prod_{i=1}^{k} t_i^{t_i-1}}{t_1!\, t_2! \cdots t_k!}.$$

Using now (3.11) we obtain that

$$\sum_{\gamma_k(n)} \frac{t_1^{t_1-1}\, t_2^{t_2-1} \cdots t_k^{t_k-1}}{t_1!\, t_2! \cdots t_k!} = \frac{k\, n^{n-k-1}}{(n-k)!}. \tag{3.15}$$

We will need the following version of Stirling's formula (see [49]):

$$m! = \sqrt{2\pi m}\, \left(\frac{m}{e}\right)^m e^{\alpha_m} \qquad \text{where} \quad \frac{1}{12m+1} \le \alpha_m \le \frac{1}{12m}. \tag{3.16}$$

This implies that

$$\begin{aligned} t_1!\,t_2!\cdots t_k! &\leq \frac{\left(\sqrt{2\pi}\right)^k}{e^n} \prod_{i=1}^{k} t_i^{t_i+1/2}\, exp\left(\frac{1}{12t_1}+\cdots+\frac{1}{12t_k}\right) \\ &\leq \frac{\left(e^{1/12}\sqrt{2\pi}\right)^k}{e^n}\, t_1^{t_1+1/2}\, t_2^{t_2+1/2}\cdots t_k^{t_k+1/2}. \end{aligned} \tag{3.17}$$

From (3.15) and (3.17) it follows that

$$\frac{k\,n^{n-k-1}}{(n-k)!} \geq \frac{e^n}{\left(e^{1/12}\sqrt{2\pi}\right)^k} \sum_{\gamma_k(n)} \frac{1}{(t_1t_2\cdots t_k)^{3/2}}. \tag{3.18}$$

Stirling's formula gives that

$$\begin{aligned} \frac{n^{n-k-1}}{(n-k)!} &\leq \frac{n^{n-k-1}\,e^{n-k}}{\sqrt{2\pi}\,(n-k)^{n-k+1/2}} = \frac{e^{n-k}}{\sqrt{2\pi}\,(n-k)^{3/2}}\left(\frac{n}{n-k}\right)^{n-k-1} \\ &\leq \frac{e^n}{\sqrt{2\pi}\,(n-k)^{3/2}} \end{aligned}$$

which used in (3.18) implies

$$C(n,k,3/2) = \sum_{\gamma_k(n)} \frac{1}{(t_1t_2\cdots t_k)^{3/2}} \leq \frac{k\,\left(e^{1/12}\sqrt{2\pi}\right)^k}{\sqrt{2\pi}}\,(n-k)^{-3/2}. \tag{3.19}$$

This proves the theorem when $p = 3/2$.

3.4 Proof for large enough powers

Fix n and $2 \le k \le n-1$ and choose $N_0(n,k) = \lceil (k-1)\log_2 n \rceil$. We will now prove the theorem for all powers $N \ge N_0(n,k)$. Indeed, we have

$$C(n,k,N) = \sum_{\gamma_k(n)} \frac{1}{(t_1 t_2 \cdots t_k)^N} \le \frac{k}{(n-k+1)^N} + \binom{n-1}{k-1} \frac{1}{2^N (n-k)^N} \tag{3.20}$$

The first term in the right hand member represents the sum over all k-part compositions of n in which $k-1$ of the terms are equal to 1 and the remaining one equals $n-k+1$. The second term represents an upper bound for the sum over the remaining k-part compositions. There are $\binom{n-1}{k-1} - k$ such compositions and it is easy to see that the product $t_1 t_2 \cdots t_k$ attains its minimum when $t_1 = t_2 = \cdots = t_{k-2} = 1$, $t_{k-1} = 2$ and $t_k = n-k$.

From (3.20) and our choice for N it immediately follows that for every $k \ge 2$

$$C(n,k,N) \le \frac{k}{(n-k)^N} + \frac{n^{k-1}}{2^N (n-k)^N} \le \frac{k+1}{(n-k)^N} \le \frac{k\, a^k}{\sqrt{2\pi}} (n-k)^{-N}. \tag{3.21}$$

3.5 Proof for all powers $p \ge 3/2$

Let $p \ge 3/2$ be an arbitrary constant. Given n and $k \le n-1$, choose N large enough so that inequality (3.21) is satisfied. Obviously, we can restrict ourselves to the case $p < N$. By Hölder's inequality we have

$$
\begin{aligned}
C(n,k,p) &= \sum_{\gamma_k(n)} (t_1 t_2 \cdots t_k)^{-p} \\
&= \sum_{\gamma_k(n)} \left[(t_1 t_2 \cdots t_k)^{-\frac{3(N-p)}{2N-3}} \cdot (t_1 t_2 \cdots t_k)^{-\frac{N(2p-3)}{2N-3}} \right] \\
&\le \left[\sum_{\gamma_k(n)} (t_1 t_2 \cdots t_k)^{-\frac{3(N-p)}{2N-3} \cdot \frac{2N-3}{2(N-p)}} \right]^{\frac{2(N-p)}{2N-3}} \cdot \\
&\quad \left[\sum_{\gamma_k(n)} (t_1 t_2 \cdots t_k)^{-\frac{N(2p-3)}{2N-3} \cdot \frac{2N-3}{2p-3}} \right]^{\frac{2p-3}{2N-3}} \\
&= [C(n,k,3/2)]^{\frac{2(N-p)}{2N-3}} \cdot [C(n,k,N)]^{\frac{2p-3}{2N-3}} \\
&\le \left(\frac{k\, a^k}{\sqrt{2\pi}} \right)^{\frac{2(N-p)}{2N-3}} \cdot (n-k)^{-\frac{3}{2} \cdot \frac{2(N-p)}{2N-3}} \cdot \\
&\quad \left(\frac{k\, a^k}{\sqrt{2\pi}} \right)^{\frac{2p-3}{2N-3}} \cdot (n-k)^{-N \cdot \frac{2p-3}{2N-3}} \\
&= \frac{k\, a^k}{\sqrt{2\pi}} (n-k)^{-p}.
\end{aligned}
$$

□

Bibliography

[1] N. Alon, J. Spencer and P. Erdős, *The Probabilistic Method*, John Wiley & Sons, New York, (1992).

[2] P. K. Agarwal, J. Pach, *Combinatorial Geometry*, Wiley Interscience, New York, (1995).

[3] K. Azuma, Weighted sums of certain dependent variables *Tôku Math Journal*, **3**, 357-367, (1967).

[4] B. Bollobás, *Random Graphs*, Academic Press, London (1985).

[5] B. Bollobás, *Modern Graph Theory*, Springer, New York (1998).

[6] B. Bollobás, *Private Communication*, July, 2000.

[7] B. Bollobás and A.G. Thomason, Random graphs of small order, *Annals of Discrete Mathematics*, 47-97, (1985).

[8] B. Bollobás and W. F. de la Vega, The diamter of random regualar graphs, *Combinatorica* **2**, 125-134, (1982).

[9] A. Cayley, A theorem on trees. *Quart. J. Math.* **23**, 376-378, (1889).

[10] R. Diestel, *Graph Theory*, Springer, New York, (1997).

[11] E.W. Dijkstra, A note on two problems in connexion with graphs, *Numerische Math.*, **1**, 269-271 (1959).

[12] P. Erdős and A. Rényi, On the strength of connectedness of a random graph, *Acta Mathematica Academiae Scientiarum Hungaricae*, **12**, 261-267 (1961).

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

[14] P. Erdős and A. Rényi, On random graphs, *I. Publ. Math. Debrecen*, **6**, 290-297 (1959).

[15] P. Erdős and A. Rényi, On random matrices, *Acta Mathematica Academiae Scientiarum Hungaricae*, **8**, 455-461 (1964).

[16] P. Erdős and A. Rényi, On the existemce of a factor of degree one of a connected random graph, *Acta Mathematica Academiae Scientiarum Hungaricae*, **17**, 359-368 (1966).

[17] P. Erdős and A. Rényi, On random matrices II, *Acta Mathematica Academiae Scientiarum Hungaricae*, **3**, 459-464 (1968).

[18] S. Even, *Graph Algorithms*, Computer Science Press, 27, (1979).

[19] W. Feller, *An Introduction to Probability Theory and Its Applications, Volume I*, Second Edition, Wiley Publications in Statistics, New York, (1957).

[20] T. Fenner and A. M. Frieze, On the connectivity of random m–orientable graphs and digraphs, *Combinatorica*, **2**(4), 347-359, (1982).

[21] T. Fenner and A. M. Frieze, On the existence of Hamiltonian cycles in a class of random graphs, *Discrete Mathematics*, **45**, 301-305, (1983).

[22] P. Flajolet, Singularity analysis and asymptotics of Bernoulli sums, *Theoretical Computer Science* **215**, 371-381, (1999).

[23] A. Frieze and T. Luczak (editors), Sparse random graphs with a given degree sequence, *Random Graphs, Volume 2* Wiley Interscience Series in Discrete Mathematics, New York, 165-182, (1992).

[24] J. Gimbal and E. M. Palmer, Matchings and cycle covers in random digraphs, *Discrete Applied Mathematics* **34**/1-3, 121-128, (1991).

[25] I.P. Goulden and D. M. Jackson, *Combinatorial Enumeration*, Wiley Interscience Series in Discrete Mathematics, §3.3.10, (1983).

[26] A. Halpert, F. Lengyel and J. Pach, Cellular telephone networks and random maps in hypergraphs, *Jouranal of Discrete Applied Mathematics*, **103**/1-3, 111-126, (2000).

[27] W. Hoeffding, Probability inequalities for sums of random variables, *Journal of the American Statistical Association*, **58**, 13-30, (1963).

[28] V. E. Hoggatt and D. A. Lind, Compositions and Fibonacci numbers. *Fibonacci Quart.* **7**, 253-266,(1969).

[29] M. Iridon, D. W. Matula and C. Yang, A graph theoretic approach for channel assignment in cellular networks, *Wireless Networks*, **7(6)**, 567-574, (2001).

[30] S. Janson, Poisson approximation for large deviations, *Random Structures and Algorithms*, **1**, 221-230, (1990).

[31] S. Janson, T. Luczak and A. Rucinski, *Random Graphs*, Wiley Interscience Series in Discrete Mathematics, New York, 81, (2000).

[32] J. Jaworski and I. H. Smit, On a random digraph, *Annals of Discrete Mathematics* **33**, 111-127, (1987).

[33] C. Jordan, *Calculus of finite differences*, Third Edition, Chelsea Publishing Co., New York, (1965).

[34] L. Katz, Probability of indecomposability of a random mapping function, *Annals of Mathematical Statistics*, **25**, 512-517, (1955).

[35] A. Knopfmacher and J. N. Ridley, Reciprocal sums over partitions and compositions. *SIAM J. Discrete Math.* **6**, no. 3, 388-399, (1993).

[36] D. E. Knuth, *The Art of Computer Programming, Volume One*, Second Edition, Addison-Wesley Publishing Company, New York, (1973).

[37] V. F. Kolchin, *Random Graphs, Encyclopedia of Mathematics*, Cambridge University Press, (1999).

[38] D. C. Kozen, *The Design and Analysis of Algorithms*, Texts and Monographs in Computer Science, Springer, New York, (1991).

[39] K. Kuratowski, *Sur le probl'eme des courbes gauches en topologie*, Fundamenta Mathematicae 15, 271–283, (1930).

[40] D. W. Matula and C. Yang, A multilayered arrangement for load sharing in cellular communication systems, *United States patent number 5,633,915.*

[41] B. Maury, Construction de suites symétrques, *Comptes Rendu Academie des Sciences Paris*, 288, 679-681 (1979).

[42] K. Menger, Zur allgemeinen Kurventheorie, *Fund. Math.* **10**, 96-115, (1927).

[43] D. Milman and G. Schechtman, Asymptotic theory of finite dimensional normed spaces, *Lecture Notes in Mathematics*, **1200**, Springer, New York (1986).

[44] L. Moser and E. L. Whitney, Weighted compositions. *Canad. Math. Bull.* 4, 39-43, (1961).

[45] L. Moser and M. Wyman, Asymptotic development of the Stirling numbers of the first kind. *J. London Math. Soc.*, **33**, 133-146, (1958).

[46] F. Riesz and B. Sz.-Nagy, *Functional Analysis*, Dover Publications, New York, 40, (1990).

[47] A. Rényi, On connected graphs I, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* 4, 385-388, (1959).

[48] J. Riordan, *An introduction to combinatorial analysis*, Wiley Publications in Mathematical Statistics, New York, (1958).

[49] H. Robbins, A remark on Stirling's formula, *Americam Mathematical Monthly*, **62**, 26-29, (1955).

[50] R. Stanley, *Enumerative Combinatorics, Vol. 2*, Cambridge Studies in Advanced Mathematics, **62**, Cambridge University Press, Cambridge, (1999).

[51] M. Stead, Downtate Medical Center, Brooklyn, NY, *Private Communtication with Janos Pach* (1995).

[52] N. M. Temme, Asymptotic estimates of Stirling numbers. *Stud. Appl. Math.*, **89**, no. 3, 233-243, (1993).

[53] K. Wagner, Uber eine Eigenschaft der ebenen Komplexe, Math. Ann. 114 ,570-590, (1937).

[54] Herbert S. Wilf, *Generatingfunctionology*, Academic Press, New York, (1994).

[55] N. C. Wormald, The asymototic connectivity of labelled regular graphs, *Journal of Combinatorial Theory, Series B*, **31**, 156-167, (1981).

[56] N. C. Wormald, Models of random regular graphs, *London Mathematical Society Lecture Note Series Surveys in Combinatorics*, (2000).

[57] E. M. Wright, The number of connected sparsely edged graphs. *J. Graph Theory*, **1**, no. 4, 317-330, (1977).