

Article

Use of Enumerative Combinatorics for Proving the Applicability of an Asymptotic Stability Result on Discrete-Time SIS Epidemics in Complex Networks

Carlos Rodríguez Lucatero ^{1,*}  and Luis Angel Alarcón Ramos ²

¹ Departamento de Tecnologías de la Información, Universidad Autónoma Metropolitana, Cuajimalpa, Mexico-City 05348, Mexico

² Departamento de Matemáticas Aplicadas y Sistemas, Universidad Autónoma Metropolitana, Cuajimalpa, Mexico-City 05348, Mexico; lalarcon@correo.cua.uam.mx

* Correspondence: crodriguez@correo.cua.uam.mx; Tel.: +52-55-5814-6500 (ext. 2953)

Received: 24 November 2018; Accepted: 24 December 2018; Published: 29 December 2018



Abstract: In this paper, we justify by the use of Enumerative Combinatorics, the applicability of an asymptotic stability result on Discrete-Time Epidemics in Complex Networks, where the complex dynamics of an epidemic model to identify the nodes that contribute the most to the propagation process are analyzed, and, because of that, are good candidates to be controlled in the network in order to stabilize the network to reach the extinction state. The epidemic model analyzed was proposed and published in 2011 by of Gómez et al. The asymptotic stability result obtained in the present article imply that it is not necessary to control all nodes, but only a minimal set of nodes if the topology of the network is not regular. This result could be important in the spirit of considering policies of isolation or quarantine of those nodes to be controlled. Simulation results using a refined version of the asymptotic stability result were presented in another paper of the second author for large free-scale and regular networks that corroborate the theoretical findings. In the present article, we justify the applicability of the controllability result obtained in the mentioned paper in almost all the cases by means of the use of Combinatorics.

Keywords: asymptotic graph enumeration problems; complex networks control; virus spreading on complex networks

1. Introduction

The main aim of the present article is to prove the applicability of an asymptotic stability result on discrete-time SIS epidemics in complex networks by the use of a combinatorial enumeration argument. Here SIS means that the nodes can be susceptible to contagion, can transit to a state of infection and once they are healthy they can become again susceptible to contagion.

The bifurcation analysis that will be presented in Section 4 as well as the criterion for selection of control nodes that will be presented in Section 5 of the present article were unpublished results prior to the more general and refined criterion obtained in the article [1]. During the development of the criterion of selection of nodes to be controlled, presented in Section 5 of the present article, we faced the problem that in the case that the behavior transition and interaction of the nodes were homogenous and the network had a topology of regular type, that is to say that the nodes had the same degree, we would be forced to have to control all the nodes since we could not distinguish them. Due to this drawback, we decided to show through an enumerative combinatorial argument that, even in the case of homogeneous behavior of the nodes, the criterion is applicable in most cases because the probability that a randomly generated graph will be regular tends to zero as the number of nodes

grows arbitrarily. Simultaneously, the second author, in collaboration with other colleagues, refined the selection criterion of nodes to be controlled obtained in Section 5 of the present article and tested by simulation that the criterion worked under the hypothesis of non-homogeneity, for scale-free type topologies in the article [1].

Combinatorics is the science of combinations. It is a very important subject in the field of Discrete [2] Among many other things, it helps us to conceive methods for enumerating a wide range of objects that accomplish a certain property of interest in a given domain. In particular, the subfield of analytic combinatorics has as a goal the precise prediction of large structured combinatorial configurations under the approach of analytic methods by the use of generating functions. Analytic combinatorics is the study of finite structures built according to a finite set of rules. Generating functions is an enumerative combinatorics tool that connects discrete mathematics and continuous analysis. One of the applications of the enumerative combinatorics is the probabilistic method. His rapid development is related with the important role played by randomness in Theoretical Computer Science. The probabilistic method is one of the most simple and beautiful nonconstructive proving methods that has been used in combinatorics in the last sixty years. This method was one of the most important contributions of the great Hungarian mathematician Paul Erdős [3,4]. This method is used for proving the existence of some kind of mathematical object by showing that if we chose some object randomly from a specified class, the probability that the mathematical object is of a prescribed type is more than zero. The method is applied in many fields of the mathematics as can be number theory, combinatorics, graph theory, linear algebra, information theory and computer science.

The second author of the present article, who is also the first author of [1], in a previous result that is less general than the model of the virus spreading on complex networks presented in [1], was interested in establishing the conditions allowing for detecting those nodes of a complex network that should be controlled such that the system can be steered to a stable virus extinction state. In fact, this previous model was presented as an application example of the results obtained in a section of [1]. The criteria obtained in [1] as well as the one obtained in the mentioned previous model, can be considered as a good option when the number of nodes to be controlled with respect to the total number of nodes is small. The number of nodes to be controlled depends on the topology of the complex network as well as on the values of the state transition parameters in each node. If all the nodes have the same degree and the transition parameters are the same for each node, then we can not distinguish the nodes to be controlled from the ones that don't have to be controlled. In such a situation, we have to control all the nodes. As a consequence, the applicability of the criteria for the detection of the nodes to be controlled obtained in [1] is no longer a good choice given that it becomes very expensive to control the totality of the nodes in the complex network. In [1], the authors avoid this problem giving as an argument that, in the literature of complex networks, the majority of the topologies that can be found in practice are not regular. Additionally, the authors of [1] generalise the previous model that was referred to, assigning to each node different internal transition probabilities, as well as different rates of interaction among nodes. The first author of the present paper thinks that the situation where each node has the same internal transition probabilities as well as the same rate of interaction among nodes can not be discarded because it is the most common case observed in the behavior of the agents in the social networks and the one which is based on the good predictability of the models used for simulating social phenomena [5–14]. Because of that, in the present article, we want to give an argument in favor of the applicability of the result obtained in [1], by means of using the mathematical tool of enumerative combinatorics. For that end, we are going to assume that the transitions' probabilities of each node as well as the rate of interaction among nodes are the same for all of the networks and use an enumerative combinatoric argument to show that, even in that case, the criteria of detection of nodes to be controlled is still applicable.

The paper is organized as follows: in Section 2, the SIS mathematical model is introduced. In Section 3, we introduce the previous model that was presented as an application example in [1]. In Section 4, we present a bifurcation analysis that allows for obtaining the epidemic transition

threshold of the complex network. In Section 5, we propose a method to determine the nodes to be controlled. In order to show the application of the criteria obtained in Section 5, we give some simulations in Section 6. In Section 7, we talk about the enumerative combinatoric argument that allow us to justify the applicability of the criteria for the selection of nodes to be controlled obtained in Section 5 as well as the one obtained in [1]. In Section 8, we make an overview of the graph enumeration problem and also about some results obtained for the enumeration of connected graphs and asymptotic analysis. In Section 9, we talk about some results obtained for the enumeration of regular graphs. Finally, the main results and conclusions are presented in Section 10.

2. Epidemics Spreading in Complex Networks

Many systems have been studied using structures called complex networks due to the number of its elements and their interaction. Each node in the network has associated some variables that represent the state of the node. The edges that connect the nodes are weighted and can be undirected or directed. The resulting graph is an undirected network or a directed network. The state in any node depends on the state and interaction intensity with the nodes in his neighbour. Examples of such systems are the social networks, the computer networks and the electrical systems.

Recently, many researchers have concentrated their efforts on studying the problem of controllability in complex networks [15–19], i.e., to drive the dynamical system from any initial state to some desired state in a finite time.

The controllability problem bring up many questions, and one of the most important is to determine which nodes must be controlled. In [15], the authors propose a method to find the minimum set of driver nodes that control the system. This minimum set, which is equal to the number of inputs, is determined by the maximum matching in the network. This approach of structured control theory makes use of tools of graph theory.

In [16], the authors propose to make use of the maximum matching graph theory concept, but, unlike the approach proposed in [15], where dynamics were considered at node level to control the system, they propose to make use of the dynamic at edges level and associate the state of the system to the state of the edges of the network. In the latter case, the authors of [16] determine the smallest set of control paths and therefore of driver nodes.

The concept of controllability of linear systems has been applied in the context of complex networks. In [18,19], the authors give the requirements that must be satisfied in order to achieve the Kalman's controllability rank condition [20]. Finally, in [17], the authors propose a control strategy which supposes that the selection of control nodes must be based on the partition of the nodes of the network, defined in terms of the energy to control the network and propose as well a distributed control law to drive the system to a desired state. However, in all the articles mentioned in the previous paragraphs, a node is controlled by an external input that modifies the state of the node, but, in the present paper, a node is controlled only if it meets some property. The purpose of controlling a node consists in stating conditions in order to determine which nodes are the best ones for monitoring and controlling to stabilize the system in the extinction state for stopping the propagation of a virus. In order to verify our theoretical results, we make simulations on a scale free network as well as on a regular network. In contrast with other authors, we consider an undirected network where each node follows the well-known Susceptible-Infected-Susceptible (SIS) model [10], so that the control mechanism brings the system to the extinction state. Of course, we are interested in determining the number of nodes that will be controlled.

3. Control Problem Statement

For us to control the virus spreading in the network will mean to modify in a set of selected nodes one or more of its parameters in such a way that the infection tends to the extinction. The criteria for selecting will be described in Section 5. The discrete time Markov process dynamical system with the SIS epidemiological model that we will use to illustrate the problem is the one proposed in [10] and was

mentioned in a section of [1] as a feedback control example. Consider a discrete time Markov process dynamical system with the SIS epidemiological model as is described in [10]. Transitions between states depend on μ_i that is the recovery probability of each node i , and $\zeta_i(t)$ that is the probability that a node i is not being infected by interaction with their neighbours, as is shown in Figure 1. In each time step, the node i try to get in contact with each of its neighbours, the probability that node i perform at least one contact with its neighbour j is given by r_{ij} , and the probability that the node i will be infected by contact with an infected node is β_i . In that case, we consider that β_i can be tuned in such a way that we can improve the nodes' health. By considering that the interaction between the node i with each one of its neighbours is independent, the probability that node i is not infected by its neighbours is

$$\zeta_i(t) = \prod_{j=1}^N [1 - a_{ij}\beta_i r_j p_j(t)], \tag{1}$$

where a_{ij} are the entries of the adjacency matrix $A \in \mathbb{R}^{N \times N}$ that represents the existing connections between the N nodes of the undirected network, and $p_i(t)$ is the probability that a node i is infected. According to the transitions diagram (see Figure 1), the following nonlinear dynamics is obtained:

$$p_i(t + 1) = (1 - \mu_i)p_i(t) + (1 - \zeta_i(t))(1 - p_i(t)). \tag{2}$$

The purpose of the present work is to identify the nodes to be controlled in the network, so a control mechanism will allow to bring the system to the extinction state from any initial state, i.e., $p_i(t) \rightarrow 0$ as $t \rightarrow \infty$, from any state $p_i(0) = p_{oi}, i = 1, 2, \dots N$. We will assume that $r_{ij} = r_{ji} = r_j$, that is, a node j tries to connect to a node i in the same way that a node i tries to connect to a node j . We will assume that, if a node i reduces the number of connection attempts with the node j , the node j will do the same. We have to mention that, in our work, we will not be considering in advance a particular complex network structure. Moreover, we will consider that not all the parameters in each node have the same value. This means that we will assume non-homogeneity both in the structure and in the properties in our model.

Where SIS in Figure 1 means that the nodes can be susceptible to contagion, can transit to a state of infection and once they are healthy they can become again susceptible to contagion.

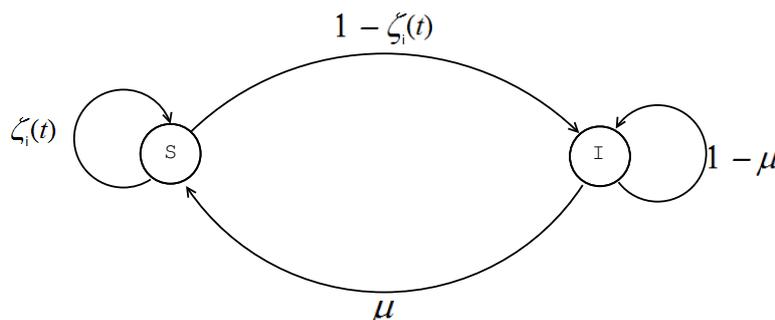


Figure 1. State transition diagram associated with the SIS model, where S represents the susceptible state and I the infected state.

4. Bifurcation Analysis

Recently, an important number of scientific works have been dedicated to the study of the dynamics of the spread of epidemics in complex networks [5,10,21], so we have followed the same research path in order to determine under what conditions the epidemic extinction state $p_i = 0$ becomes a global attractor for any node i . First, we consider the following bound

$$1 - \zeta_i(t) \leq \sum_{j=1}^N a_{ij}\beta_i r_j p_j(t) \tag{3}$$

proved in [5] and then determine a linear dynamics substituting (3) into (2)

$$p_i(t + 1) = (1 - \mu_i)p_i(t) + (1 - p_i(t)) \sum_{j=1}^N a_{ij}\beta_i r_j p_j(t) \tag{4}$$

$$\leq (1 - \mu_i)p_i(t) + \sum_{j=1}^N a_{ij}\beta_i r_j p_j(t). \tag{5}$$

Therefore, we propose the following bounded linear dynamics that is given by

$$x_i(t + 1) = (1 - \mu_i)x_i(t) + \beta_i \sum_{j=1}^N a_{ij}r_j x_j(t), \quad p_i(t) \leq x_i(t). \tag{6}$$

In matrix notation our system, $X(t + 1) = \mathbf{H}X(t)$ is summarized, where $\mathbf{H} = [I - \mu + \beta\mathbf{R}A]$, $\mathbf{R} = \text{diag}(r_i)$, $\mu = \text{diag}(\mu_i)$, $\beta = \text{diag}(\beta_i)$ and $X(t) = [x_1(t), x_2(t), \dots, x_N(t)]^T$. These linear dynamics given by the matrix \mathbf{H} are asymptotically stable if and only if the spectral radius ($\sigma(\mathbf{H})$) of the matrix \mathbf{H} meets the condition that $\sigma(\mathbf{H}) \in \{\lambda \in \mathbb{C} : |\lambda| < 1\}$.

If the contact probabilities μ_i and β_i belong to the closed interval $[\mu_{min}, \mu_{max}]$ and $[\beta_{min}, \beta_{max}]$, respectively, it is possible to bound the spectral radius as follows:

$$\sigma(\mathbf{H}) < 1. \tag{7}$$

This equation is similar to that reported in [5,10]. Moreover, from (7), we have several bifurcation parameters given by β_{max} , μ_{min} and r , and, for control purposes, it is appropriate to take β as a control parameter. In this case, β_i will be our control parameter in some nodes, in order to stabilize our system in the epidemic extinction state.

5. Selection of Nodes to Be Controlled

For the selection of the nodes to be controlled, we have to consider (6) as an uncoupled system due to the symmetry of the matrix \mathbf{R} . Consider Equation (6) as a sum over all nodes as follows.

The proposed model is

$$p_i(t + 1) = (1 - \mu_i)p_i(t) + (1 - \zeta_i(t))(1 - p_i(t)), \tag{8}$$

$$\zeta_i(t) = \prod_{j=1}^N (1 - a_{ij}\beta_i r_j p_j(t)). \tag{9}$$

From (8), we can deduce (3) and, by substitution of (3) on (8), we obtain the inequality of (4). This lends us to propose the following new dynamics

$$x_i(t + 1) = (1 - \mu_i)x_i(t) + \beta_i \sum_{j=1}^N a_{ij}r_j x_j(t), \tag{10}$$

where $p_i(t) \leq x_i(t)$. Representing (10) in matrix form, we obtain the following mathematical expression

$$\mathbf{X}(t + 1) = [I - \mu + \beta\mathbf{R}A]\mathbf{X}(t). \tag{11}$$

We can bound the Equation (11) in terms of the eigenvalues, in order to determine the system stability, as follows:

$$\sigma([I - \mu + \beta\mathbf{R}A]) < 1. \tag{12}$$

If the condition (12) is satisfied, the system will be asymptotically stable in the extinction state. In order to select those nodes that participate in the spreading process, we have the following result known as the Gerschgorin theorem [22,23]:

$$\rho(A) \leq \max_i \sum_{j=1}^N |a_{ij}|, \tag{13}$$

which implies

$$|\lambda - (1 - \mu_i)| \leq \sum_{j=1}^N |\beta_i a_{ij} r_j|, \tag{14}$$

whose consequence is

$$|\lambda| - |1 - \mu_i| \leq |\lambda - (1 - \mu - i)| \leq \sum_{j=1}^N \beta_i a_{ij} r_j \tag{15}$$

from where we get

$$|\lambda| \leq |\lambda - 1 + \mu_i| + |1 - \mu_i| \leq \sum_{j=1}^N \beta_i a_{ij} r_j + 1 - \mu_i \leq 1 \tag{16}$$

and we conclude that

$$\beta_i \sum_{j=1}^N a_{ij} r_j < \mu_i. \tag{17}$$

Note that the last equation does not tell us how to control the spread of the disease but instead tells us which nodes we need to control in order to reach the epidemic extinction state. In this case, the nodes that have to be controlled are those who do not satisfy the inequality (17).

6. Simulations

The purpose of the control design is to make the system reach the extinction state (i.e., $p_i \rightarrow 0$, for $t > 0$). In this work, we suppose that β_i can be manipulated in both cases: homogeneous (i.e., $\beta_i = \beta$, $\mu_i = \mu$ and $r_i = r$) and non-homogeneous for scale-free and regular networks.

Following the results obtained in [1], the selection of nodes to be controlled, for any network topology, is given by (see Equation (25) from [1])

$$\beta_i \sum_{j=1}^N r_j a_{ij} < \mu_i \tag{18}$$

for the non-homogeneous case, and

$$\beta \sum_{j=1}^N r a_{ij} < \mu \tag{19}$$

for the homogeneous case.

From Equations (18) and (19), it is evident that, in the homogeneous case, it will be necessary to control all nodes of the network because the necessary condition (19) is the same for every node in the network.

In order to verify the above results, we will apply the node selection criteria in the following cases:

- A scale-free network proposed according to the algorithm in [24], for homogeneous and non-homogeneous cases.
- A regular network for homogeneous and non-homogeneous cases.

6.1. Non-Homogeneous Scale-Free Network

Let μ_i , β_i and r_i be distributed uniformly in $[0.35, 0.65]$, $[0.40, 0.80]$ and $[0.25, 0.65]$, respectively, and $N = 100,000$ nodes. Simulation results predict that, on average, it will be necessary to control 67,610 nodes. In order to control the selected nodes, let there be

$$\bar{\beta}_i = \frac{\mu_i}{\sum_{j=1}^N r_j a_{ij}},$$

and the linear control law

$$\beta_i(t) = k\bar{\beta}_i(1 - p_i(t)), \tag{20}$$

where $k \in (0, 1)$ is the gain that depends on the resources available to the control. The control law is applied just to those nodes that do not satisfy the necessary condition (18). In this case, the gain is set to $k = 0.99$. The simulation results are shown in Figure 2.

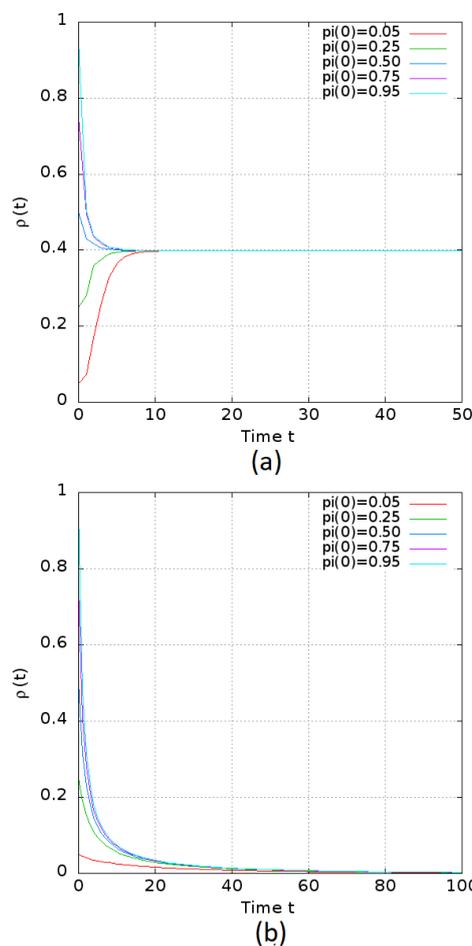


Figure 2. Non-homogeneous scale-free network. (a) results without control and (b) results with the control law given by (20).

6.2. Homogeneous Scale-Free Network

In this case, let $\mu = 0.35$, $\beta = 0.40$, $r = 0.25$, and $N = 100,000$ nodes. Simulations results predict that 33,387 must be controlled. In order to control the selected nodes, let be

$$\bar{\beta}_i = \frac{\mu}{\sum_{j=1}^N r a_{ij}},$$

and the linear control law

$$\beta_i(t) = k\bar{\beta}_i(1 - p_i(t)), \tag{21}$$

for those nodes i that do not satisfy (19), and $\beta_i = \beta$ for those nodes i that satisfy (19). The gain is set again to $k = 0.99$. The simulation results are shown in Figure 3.

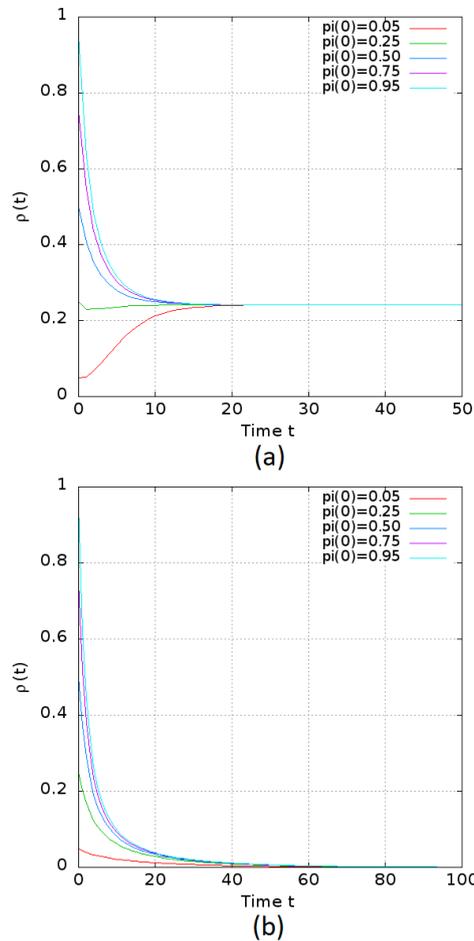


Figure 3. Homogeneous scale-free network. (a) results without control and (b) results with the control law given by (21).

6.3. Non-Homogeneous Regular Network

Let μ_i , β_i and r_i be distributed uniformly in $[0.35, 0.65]$, $[0.40, 0.80]$ and $[0.25, 0.65]$ respectively, and $N = 100,000$ nodes with degree 30. Simulation results predict that all nodes must necessarily be controlled.

In order to control the selected nodes, let there be

$$\bar{\beta}_i = \frac{\mu_i}{\sum_{j=1}^N r_j a_{ij}},$$

and the linear control law

$$\beta_i(t) = k\bar{\beta}_i(1 - p_i(t)), \tag{22}$$

for all nodes in the network. The gain is set to $k = 0.99$ and the simulation results are shown in Figure 4.

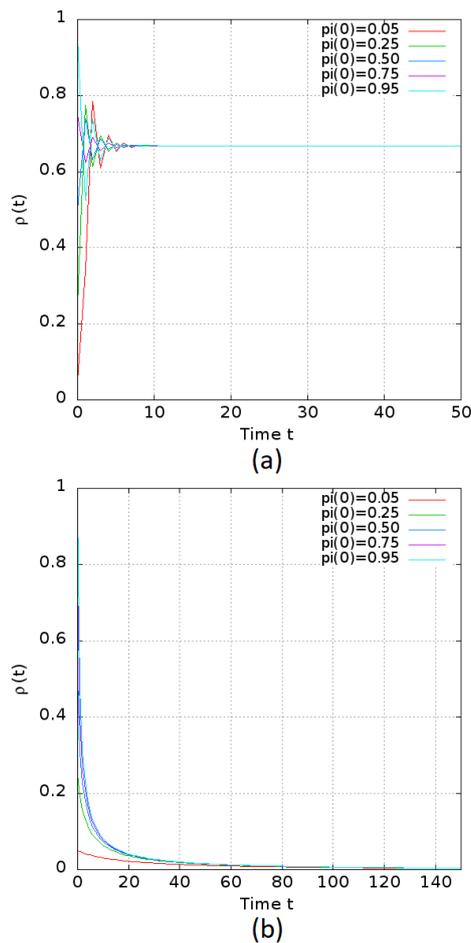


Figure 4. Non-homogeneous regular network. (a) results without control and (b) results with the control law given by (22).

6.4. Homogeneous Regular Network

In this case, let $\mu = 0.35$, $\beta = 0.40$, $r = 0.25$, and $N = 100,000$ nodes with degree 30. Simulations results predict that all the nodes must necessarily be controlled. In order to control them, let be

$$\bar{\beta}_i = \frac{\mu}{\sum_{j=1}^N ra_{ij}} = \frac{0.35}{0.40 \times 0.25 \times 30} = 0.11667,$$

and the linear control law

$$\beta_i(t) = k\bar{\beta}_i(1 - p_i(t)), \tag{23}$$

for all nodes. The gain is again set to $k = 0.99$. The simulation results are shown in Figure 5.

Note that, for regular homogeneous networks, it is necessary to control all the nodes in order to stabilize the network in the extinction state.

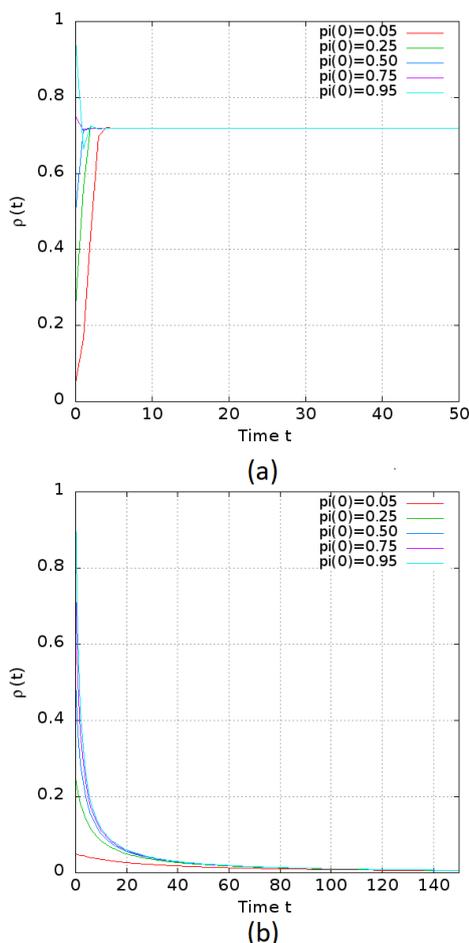


Figure 5. Homogeneous regular network. (a) results without control and (b) results with the control law given by (23).

7. The Applicability of the Result

The applicability of the result obtained in Section 5 of the present article has the drawback that, if the topology of the network is regular, which means that each node has the same degree, and the transition/interaction probabilities are the same for all the nodes in the network, and then we cannot detect what are the nodes that have to be controlled and then we have to control all the nodes. The authors of [1] tested the applicability of a similar result by simulation under the assumption of heterogeneity of the transition probabilities inside each node as well as in the interaction rates among the nodes. Another claim made in [1] is that the regular topology seldom arise in practice because the data available show that the real world computer networks are not homogeneous and follow a power law topology [24,25]. The point of view of the first author of the present article is that the homogeneous case cannot be discarded because in many mathematical models applied successfully to social phenomena [6–8,11,14] as well as to virus spreading phenomena [5,12,13,26] the models were homogeneous. The assumption of homogeneity in the aforementioned models is based on the behavior of individuals in social networks. Because of that, in the present article, we are assuming that the model is homogeneous and we will justify under such circumstances that the result of Section 5 is still applicable by proving with a combinatorial enumerative argument that the regular topology of a network is extremely unfrequent.

8. Enumeration and Generating Functions

Enumeration in Combinatorics is one of the most fertile and fascinating topics of Discrete Mathematics [2,27]. The purpose of this field has been to count the different ways of arranging

objects under given constraints. With the combinatorial enumerative methods, words, permutations, partitions, sequences and graphs can be counted. One of the mathematical tools that is frequently used for this end is the generating functions or formal power series. The generating functions represent a bridge between discrete mathematics and continuous analysis (particularly complex variable theory). When we face a problem whose answer is a sequence of numbers a_0, a_1, a_2, \dots and we want to know what the closed mathematical expression is that enables us to obtain the element a_n of that sequence. Sometimes, we can do it at first sight by inspection. For example, if the numerical sequence $1, 3, 5, 7, 9, \dots$ we recognize at first sight that it is a sequence of odd numbers and the n -th element is $a_n = 2n - 1$. In other cases, it is unreasonable to expect a simple formula as can be the case of the sequence $2, 3, 5, 7, 11, 13, 17, 19, \dots$, whose a_n is the n -th prime number. For some other sequences, it is very hard to obtain directly a simple formula, but it can be very helpful to transform it in a power series whose coefficients are the elements of that sequence as follows:

$$\sum_{i=0}^{\infty} a_i x^i. \quad (24)$$

The expression (24) is called ordinary generating function. This kind of series can be easily algebraically manipulated. In this paper, we are particularly interested in the application of the mentioned mathematical tools for the enumeration of graphs accomplishing some given properties. Many questions about the number of graphs that have some specified property can be answered by the use of generating functions. Some typical questions about the number of graphs that fulfill a given property are for example: How many different graphs can I build with n vertices? How many different connected graphs with n vertices exists? How many binary trees can be constructed with n vertices? [28,29] etc. For some of these questions, the application of generating functions allows us to easily obtain a simple formula. For some other questions, the answer is an asymptotic estimation formula. The most commonly used generating functions are the Ordinary Generating functions and the Exponential Generating functions. The generating functions are the tool used for enumerating graphs. From the point of view of the generating functions, they are to type of graph enumerating problems:

1. Labeled graphs problems,
2. Unlabeled graphs problems.

The labeled graphs problems can be easily solved with the direct application of the exponential generating functions. The case of the unlabeled enumeration problems can be solved by using ordinary generating functions but require the use of more combinatorial theory and the application of Pólya's theorem. For more details on generating functions subject, consult Appendix A.

9. Enumerating Regular Graphs

As we mentioned in Section 8, some enumerating problems can be solved easily using generating tools for obtaining a closed formula. Some other problems are more hard to deal with for obtaining a closed mathematical expression, but we can resort in such a case to the asymptotic approximation of the coefficients of the power series [30–33]. It was also mentioned in Section 8 that there are some graph enumerating problems where the nodes are labelled and in such a case the use of the exponential generating functions is well adapted for these kinds of problems. The other case of graph enumerating problems is when we are dealing with graphs whose nodes does not have an assigned label. Then, we can resort in such case to Polya's enumerating method [34,35] and the best choice is to use ordinary generating functions. It should also be mentioned that the edges of the graphs to be enumerated can be directed or undirected.

One of the seminal articles of enumerating graphs is [36], where a fundamental theorem was proved in the theory of random graphs on n unlabelled nodes and with a given number q of edges.

In [36], the authors obtained a necessary and sufficient condition for relating asymptotically the number of unlabelled graphs with n nodes and q edges with the number of labelled graphs with

n nodes and q edges. Let T_{nq} be the number of different graphs with n nodes and q edges, F_{nq} the corresponding number of labelled graphs, $N = \frac{n(n-1)}{2}$ the possible edges and $F_{nq} = \binom{N}{q} = \frac{N!}{q!(N-q)!}$. The result obtained in [36] can be stated as the following theorem:

Theorem 1. *The necessary and sufficient conditions that*

$$T_{nq} \sim \frac{F_{nq}}{n!} \tag{25}$$

as $n \rightarrow \infty$ is that

$$\min(q, N - q) / n - (\log n / 2) \rightarrow \infty. \tag{26}$$

The formal result expressed in Theorem 1 for unlabelled graphs is a starting point on the enumeration of regular graphs because it allows for enumerating those unlabelled graphs that have some number of edges. In fact, the author of [36] proved that, if a graph with $|E| = E(n)$ edges, where n is the number of vertices or order of such a graph, has no isolated vertices or two vertices of degree $n - 1$, then the number of unlabelled graphs of order n and number of edges $|E|$ divided by the number unlabelled graphs is asymptotic to $n!$.

Another interesting article on asymptotic enumeration of labelled graphs having a given degree sequence was [37]. The authors of [37] obtained their asymptotic result for $n \times n$ symmetric matrices subject to the following constraints:

- (i) each row sum is specified and bounded,
- (ii) the entries are bounded,
- (iii) a specified *sparse* set of entries must be zero.

The authors of [37] mentioned that their results can be interpreted in terms of incidence matrices for labelled graphs. The results of [37] can be stated as follows. Let $\mathcal{M}(n, z)$ be the set of all $n \times n$ symmetric $(0, 1)$ matrices with at most z zeroes in each row, \mathbf{r} a vector over $[d] = \{0, 1, \dots, d\}$ and $G(\mathbf{M}, \mathbf{r}, t)$ the number of $n \times n$ symmetric matrices (g_{ij}) over $[t] = \{0, 1, \dots, t\}$ such that

- (i) $g_{ij} = 0$ if $m_{ij} = 0$,
- (ii) $\sum_j g_{ij} = r_i$.

Theorem 2. *For given d, t and z ,*

$$G(\mathbf{M}, \mathbf{r}, t) \sim \frac{T(f, \delta) e^{\epsilon a - b}}{\prod r_i!}. \tag{27}$$

Uniformly, for $(\mathbf{M}, \mathbf{r}) \in \cup_{n=1}^{\infty} (\mathcal{M}(n, z) \times [0, d]^n)$ as $f \rightarrow \infty$, where $f = \sum_i r_i$, $\epsilon = -1$ if $t = 1$ and $+1$ if $t > 1$, for $a = \left(\frac{\sum_i \binom{r_i}{2}}{f}\right)^2 + \left(\frac{\sum_{m_{ij} \binom{r_i}{2}}}{f}\right)$, $b = \left(\sum_{m_{ij}=0, i < j} r_i r_j + \sum_i \frac{\binom{r_i}{2}}{f}\right)$, $\delta = \sum_{m_{ij}=0} r_i$ and $T(f, \delta)$ being the number of involutions on $[1, f]$ such that no element in some specified set of size δ is fixed.

Three years later, the article [38] appeared giving a different approach of [37] allowing for obtaining a more general asymptotic formula without reference to an exact formula. The asymptotic result obtained by Bela Bollobas in [38] for enumerating labelled regular graphs is proved by a probabilistic method. This result can be stated as follows. Let Δ and n be natural numbers such that $\Delta n = 2m$ is even and $\Delta \leq (2 \log n)^{\frac{1}{2}}$, where n is the number of vertices and m is the number of edges of the graph G . Then, as $n \rightarrow \infty$, the number of labelled Δ -regular graphs on n vertices is asymptotic to

$$e^{-\lambda - \lambda^2} \frac{(2m)!}{m! 2^m (\Delta!)^m}, \tag{28}$$

where $\lambda = \frac{(\Delta-1)}{2}$.

The authors of [38] affirm that the asymptotic Formula (28) holds not only for Δ constant but also for Δ growing slowly as $n \rightarrow \infty$ and summarized this in the following theorem.

Theorem 3. Let $d_1 \geq d_2 \geq \dots d_n$ be natural numbers with $\sum_{i=1}^n d_i = 2m$ even. Suppose $\Delta = d_1 \leq (2 \log n)^{\frac{1}{2}} - 1$ and $m \geq \max \{ \epsilon \Delta n, (1 + \epsilon)n \}$ for some $\epsilon > 0$. Then, the number $L(\mathbf{d})$ of labelled graphs with degree sequence $\mathbf{d} = (d_i)_1^n$ satisfies

$$L(\mathbf{d}) \sim e^{-\lambda - \lambda^2} \frac{(2m)_m}{\{2^m \prod_{i=1}^n d_i!\}}, \tag{29}$$

where $\lambda = \frac{1}{2m} \sum_{i=1}^n \binom{d_i}{2}$.

In the next year, the author on [38] extended this result to the case of unlabelled graphs in [39]. The result of Theorem 3 extended for the case of unlabelled graphs can be summarized in the following theorem.

Theorem 4. If $\Delta \geq 3$ and $L_\Delta = e^{-\lambda - \lambda^2} \frac{(2m)!}{m! 2^m (\Delta!)^m}$, then

$$U_\Delta \sim \frac{L_\Delta}{n!} \sim e^{-\frac{(\Delta^2-1)}{4}} \frac{(2m)! (\Delta!)^{-n}}{2^m m! n!}, \tag{30}$$

where $m = \frac{\Delta n}{2}$.

For the details of the proof of Theorems 3 and 4, see [38,39], respectively.

10. Combinatorial Proof of Applicability of the Result on Control Node Selection

As was mentioned in Section 1, the case of homogeneity in the behavior of the nodes and their interaction can not be discarded given what has been observed in the reaction of the agents in the context of social networks is that they try to minimize the conflict. Many successful models can, for example [5,6,11,12], base their predicting effectiveness on the homogeneity of the behaviour of the nodes and their interaction. By other side, in Section 5 of the present paper, we have obtained a criteria for selecting the nodes to be controlled, but such criteria fails if we have homogeneity in the behavior of the nodes and, at the same time, the topology of the network is regular. Then, what we want to do here is to justify the applicability of obtained in Section 5, keeping the homogeneity of the nodes and try to compare the number of regular graphs with n vertices with the total of graphs that can be constructed with n vertices. For this end, based on the results on combinatorial graph enumeration mentioned on the Theorems 3 and 4, we can state our main result as follows. First of all, we suppose that our graph is labelled, that $G = (V, E)$ is r -regular with $r \geq 3$ constant and $rn = 2m$, where $n = |V|$ corresponds to the number of vertices and $m = |E|$ corresponds to the number of edges. Let L_r the number of labelled regular graphs of degree r whose asymptotic value is ([38])

$$L_r \sim e^{-\frac{r^2-1}{4}} \frac{(2m)!}{2^m m!} (r!)^n. \tag{31}$$

Let G_n be the number of all possible graphs with n vertices whose value is

$$G_n = 2^{\binom{n}{2}}. \tag{32}$$

Theorem 5. If $r \geq 3$ and $nr = 2m$, then

$$\lim_{n \rightarrow \infty} \frac{L_r}{G_n} = 0. \tag{33}$$

Proof of Theorem 5. If $nr = 2m$ and r is constant of value c_1 , then we can deduce that $m = \frac{r}{2}n = \frac{c_1}{2}n$ and this implies that $m = O(n)$ so lets say that $m = c_2n$; then,

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{L_r}{G_n} &= \lim_{n \rightarrow \infty} \frac{e^{-\frac{c_1^2-1}{4}} \frac{(2c_2n)!(c_1!)^{-n}}{2^{c_2n} (c_2n)!}}{2^{\frac{n(n-1)}{2}}} \\ &= \lim_{n \rightarrow \infty} \frac{e^{-\frac{c_1^2-1}{4}} \frac{(2c_2n)!}{(c_1!)^n 2^{c_2n} (c_2n)!}}{2^{\frac{n(n-1)}{2}}}, \end{aligned}$$

applying the approximation Stirling formula $n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$

$$(c_1!)^n = \left(\sqrt{2\pi c_1} \left(\frac{c_1}{e}\right)^{c_1}\right)^n$$

then

$$= \lim_{n \rightarrow \infty} \frac{\frac{e^{-\frac{c_1^2-1}{4}} (2c_2n)!}{\left(\sqrt{2\pi c_1} \left(\frac{c_1}{e}\right)^{c_1}\right)^n 2^{c_2n} (c_2n)!}}{2^{\frac{n(n-1)}{2}}}$$

simplifying

$$= \lim_{n \rightarrow \infty} \frac{1}{(\sqrt{2\pi c_1})^n (c_1)^{c_1 n} e^{\frac{c_1^2+4c_1n+1}{4}} \frac{(2c_2n)!}{2^{c_2n} (c_2n)!}} 2^{\frac{n(n-1)}{2}},$$

applying the approximation Stirling formula

$$(2c_2n)! = \sqrt{2\pi(2c_2n)} \left(\frac{2c_2n}{e}\right)^{(2c_2n)}$$

and

$$(c_2n)! = \sqrt{2\pi(c_2n)} \left(\frac{c_2n}{e}\right)^{(c_2n)}$$

we get

$$= \lim_{n \rightarrow \infty} \frac{1}{(\sqrt{2\pi c_1})^n (c_1)^{c_1 n} e^{\frac{c_1^2+4c_1n+1}{4}} \frac{\sqrt{2\pi(2c_2n)} \left(\frac{2c_2n}{e}\right)^{(2c_2n)}}{2^{c_2n} \sqrt{2\pi(c_2n)} \left(\frac{c_2n}{e}\right)^{(c_2n)}}} 2^{\frac{n(n-1)}{2}},$$

simplifying

$$= \lim_{n \rightarrow \infty} \frac{\frac{\sqrt{2}}{(\sqrt{2\pi c_1})^n c_1^{c_1 n} e^{(c_1^2+4c_1n+1)/4}} \left(\frac{c_2n}{e}\right)^{c_2n}}{2^{\frac{n(n-1)}{2}}}, \tag{34}$$

$$\tag{35}$$

given that $r \geq 3$, for which we assumed that r is a constant c_1 and that $nr = 2m$, then we have that $c_1 = 2c_2$ and replacing that, in (34), we can express it in terms of c_1 , which is the regular degree r assumed as fixed; then, we get

$$\lim_{n \rightarrow \infty} \frac{\frac{\sqrt{2}}{(\sqrt{2\pi c_1})^n c_1^{c_1 n} e^{(c_1^2+4c_1n+1)/4}} \left(\frac{c_1n/2}{e}\right)^{c_1n/2}}{2^{\frac{n(n-1)}{2}}}, \tag{36}$$

and, replacing c_1 by r in (36), we get

$$\lim_{n \rightarrow \infty} \frac{\frac{\sqrt{2}}{(\sqrt{2\pi r})^n r^n e^{(r^2+4rn+1)/4}} \left(\frac{rn/2}{e}\right)^{rn/2}}{2^{\frac{n(n-1)}{2}}}. \tag{37}$$

Therefore, if the degree r is constant, the $\lim_{n \rightarrow \infty} \frac{L_r}{G_n} = 0$. \square

Now, our main result can be stated as a consequence of Theorem 5.

Theorem 6. *If we assume that all graphs are uniformly distributed and that the nodes have homogeneous behavior, then the criteria for selecting nodes to be controlled obtained in Section 5 is almost always applicable.*

Proof of Theorem 6. As a consequence of Theorem 5, we know that the probability that a regular graph appears tends to zero as $n \rightarrow \infty$. Then, the mentioned criteria is almost always applicable. \square

11. Conclusions

In the present article, we have applied the techniques of combinatorial enumeration to prove the applicability of a selection result of nodes to be controlled in a regular network modeled by a graph under the hypothesis of homogeneity in the behavior of the nodes. This allowed us to explore the application of combinatorial techniques to control problems in networks and thus verify the goodness of said methods for network analysis and the control of virus propagation in them. This still needs to be studied by applying the combinatorial methods discussed above, if there are any other types of topologies that prevent the application of the selection criteria of nodes to be controlled under the hypothesis of behavior of partially heterogeneous nodes, that is, if in the network we have subsets of nodes with the same behavioral parameters. The selecting node criteria described in Section 5 is based on the combination of the parameter values of the selected nodes as well as their degrees. In many recent publications [40–47], interesting and elaborated methods for detecting the influencer nodes in complex networks have been proposed that we will try to apply in combination with our criteria in the future in order to reduce the number of nodes to be controlled. We must mention as another contribution of the present article that the simulations of Section 6 allowed us to illustrate the application of the criterion obtained in Section 5 for scale-free as well as 30-regular topologies under conditions of both homogeneity and heterogeneity and obtain the number of nodes to be controlled in each case.

Author Contributions: All of the authors did the research, in particular the combinatorial enumeration part, enumeration and generating functions, combinatorial proof of applicability of the result on control node selection, the applicability result and respective proof as well as the MATLAB calculation was performed by the C.R.L.; L.A.A.R. performed the simulations, the Control Problem Statement, the bifurcation analysis and selection of nodes to be controlled; C.R.L. wrote the paper; all authors reviewed the manuscript; and all authors have read and approved the final manuscript.

Acknowledgments: This work was partially sponsored by Division de Ciencias de la Comunicación y Diseño Universidad Autónoma Metropolitana Unidad Cuajimalpa “Proyectos Divisionales de la Division de Ciencias de la Comunicación y Diseño de la Universidad Autónoma Metropolitana Unidad Cuajimalpa”.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

In [48], the following concepts are defined:

Definition A1. *A combinatorial class is a finite or denumerable set on which a size function is defined, satisfying the following conditions:*

- (i) *the size of an element is a nonnegative integer;*
- (ii) *the number of elements of any given size is finite.*

Definition A2. The combinatorial classes \mathcal{A} and \mathcal{B} are said to be (combinatorially) isomorphic which is written $\mathcal{A} \cong \mathcal{B}$ iff their counting sequences are identical. This condition is equivalent to the existence of a bijection from \mathcal{A} to \mathcal{B} that preserves size, and one also says that \mathcal{A} and \mathcal{B} are bijectively equivalent.

In [48], it is mentioned that the ordinary generating functions (OGF) as (24) of a sequence $\mathcal{A} = \{a_0, a_1, \dots\}$ or combinatorial class \mathcal{A} is the generating function of the numbers \mathcal{A}_n whose sizes $a_n = \text{card}(\mathcal{A}_n)$ such that the OGF of class \mathcal{A} admits the combinatorial form

$$A(x) = \sum_{\alpha \in \mathcal{A}} x^{|\alpha|}. \tag{A1}$$

It is also said that the variable x marks size in the generating function. The OGF form (A1) can be easily interpreted by observing that the term x^n occurs as many times as there are objects in \mathcal{A} of size n .

It can be defined the operation of coefficient extraction of the term x^n in the power series $A(x) = \sum a_n x^n$ as follows:

$$[x^n] \left(\sum_{n \geq 0} a_n x^n \right) = a_n. \tag{A2}$$

Such is the case of the sequence 0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, ... know as Fibonacci numbers and that satisfy the the following recurrence relation

$$F_{n+1} = F_n + F_{n-1} \text{ where } n \geq 1, F_0 = 0, F_1 = 1. \tag{A3}$$

The n -th element of this sequence F_n is the coefficient of x^n in the expansion of the function

$$\frac{x}{1 - x - x^2} \tag{A4}$$

as a power series about the origin. The very interesting book [49] talks about all the answers that can be obtained by the use of the generating functions and mentioned the following list:

1. Sometimes it can be found an exact formula for the members of the sequence in a pleasant way. If it is not the case, when the sequence is complicated, a good approximation can be obtained.
2. A recurrence formula can be obtained. Most often generating functions arise from recurrence formula. Sometimes, however, a new recurrence formula, from generating functions and new insights of the nature of the sequence can be obtained.
3. Averages and other statistical properties of a sequence can be obtained.
4. When the sequence is very difficult to deal with, asymptotic formulas can be obtained instead of an exact formula. For example, the n -th prime number is approximately $n \log n$ when n is very big.
5. Unimodality, convexity, etc. of a sequence can be proved.
6. Some identities can be proved easily by using generating functions. For instance,

$$\sum_{j=0}^n \binom{n}{j}^2 = \binom{2n}{n} \quad (n = 0, 1, 2, \dots). \tag{A5}$$

7. Relationship between problems can be discovered from the striking resemblance of the respective generating functions.

As was mentioned in the list of answers obtained by the use of generating functions, sometimes it is hard to obtain an exact formula and, in that case, we can resort to the use of asymptotic formulas. The mathematical tools used for this purpose belong to field of the Analytic Combinatorics and can be consulted in [48]. The purpose of the Analytic Combinatorics is to predict with accuracy the properties

of large structured combinatorial configurations with an approach based on mathematical analysis tools [48]. Under this approach, we can start with an exact enumerative description of the combinatorial structure by the use of a generating function. This description is taken as a formal algebraic object. After that, the generating function is used as an analytic object interpreting it as a mapping of the complex plane into itself. The singularities help us to determine the coefficients of the function in asymptotic form given as a result very good estimations of the counting of sequences. For this end, the authors of [48] organize the Analytic Combinatorics in the following three components:

1. Symbolic Methods that establish systematically relations discrete mathematics constructions and operations on generating functions that encode counting sequences.
2. Complex Asymptotics that allow for extracting asymptotic counting information from the generating functions by the mapping to the complex plane mentioned above.
3. Random structures concerning the probabilistic properties accomplished by large random structures.

Rich material relative to Complex Asymptotic Analysis can be found in [50]. A nice text that can be consulted for applications of the enumerative combinatorics tools to the analysis of algorithms is [51].

In this paper, we are particularly interested in the application of the mentioned mathematical tools for the enumeration of graphs accomplishing some given properties. Let us define what is a graph. In the present article, we will define a graph $G = \langle V, E \rangle$ as a structure with a set of vertices $V = \{v_1, v_2, \dots, v_p\}$ with cardinality $|V| = p$ called the order of G and a set of unordered pairs of adjacent vertices, called edges, $E = \{\{v_{i_1}, v_{j_1}\}, \{v_{i_2}, v_{j_2}\}, \dots, \{v_{i_q}, v_{j_q}\}\}$ if G is undirected or a set of ordered pairs of adjacent vertices $E = \{(v_{i_1}, v_{j_1}), (v_{i_2}, v_{j_2}), \dots, (v_{i_q}, v_{j_q})\}$, if G is a directed graph, having cardinality $|E| = q$, without loops and without multiple edges. A graph G with p vertices and q edges is called a (p, q) graph. In a labeled graph of order p , one integer from 1 to p is assigned to each vertex as its label.

Many questions about the number of graphs that have some specified property can be answered by the use of generating functions. Some typical questions about the number of graphs that fulfill a given property are for example: How many different graphs can I build with n vertices? How many different connected graphs with n vertices exists? How many binary trees can be constructed with n vertices? etc. For some of these questions, the application of generating functions allow us to obtain easily a simple formula. For some other questions, the answer is an asymptotic estimation formula.

The most commonly used generating functions are the Ordinary Generating functions and the Exponential Generating functions.

The Ordinary Generating functions can be defined as follows:

$$a(x) = \sum_{k=0}^{\infty} a_k x^k, \tag{A6}$$

where the coefficients are elements of the sequence of numbers a_0, a_1, a_2, \dots . This kind of generating functions are used in combinatorial selection problems where the order is not important. The Exponential Generating functions can be defined as follows:

$$b(x) = \sum_{k=0}^{\infty} b_k \frac{x^k}{k!}, \tag{A7}$$

where the coefficients are elements of the sequence of numbers b_0, b_1, b_2, \dots . These kinds of generating functions are used in combinatorial disposition problems where the order is crucial. The sequence of numbers are counting sequences and can be encoded exactly by generating functions. Combinatorics deals with discrete objects as, for example, graphs, words, trees and integer partitions. One interesting problem is the enumeration of such objects. In the present article, we are interested in the enumeration

of graphs. One of the main contributors of the combinatorial enumeration of graphs is the great mathematician George Pólya who counted graphs with a given number of vertices and edges and proposed closed formulas for many graph counting problems based on group theory [34]. From Polyá’s formulas, it became relatively easy to enumerate rooted graphs, connected graphs, etc. One of the first graph enumerating problems was the enumeration of number of triangulations of certain plane polygons by Leonard Euler [52] in the 18th century. After that, the electrical engineer Kirchhoff in [53] found the number of spanning trees in a connected graph, that is to say, the number of labelled trees. Some years later, the English mathematician Arthur Cayley was interested in the enumeration of trees (labeled trees, rooted trees and ordinary trees) and obtained in [54] the closed formulas for solving such enumerating problem. Another great mathematician for a long time unknown and who anticipated Polyá in his discoveries was John Howard Redfield [55]. Many objects and their configuration that are not graphs can be enumerated by clever transformations to graphs as, for example, automata, boolean functions or chemical isomers. The generating functions are the tool used for enumerating graphs. From the point of view of the generating functions, they are to type of graph enumerating problems:

1. Labeled graphs problems,
2. Unlabeled graphs problems.

The labeled graphs’ problems can be easily solved with the direct application of the exponential generating functions. The case of the unlabeled enumeration problems can be solved by using ordinary generating functions but require the use of more combinatorial theory and the application of Pólya’s theorem.

The first labeled graph enumeration that can be asked is: How many labeled graphs with p vertices and q edges are there? For answering that question, let $G_p(x)$ be the polynomial or ordinary generating function whose coefficient of the term x^k represents the number of labelled graphs with p vertices and k edges. If V is the set of vertices of cardinality p , there are $q = \binom{p}{2}$ pairs of these vertices. In every vertex set V , each pair is adjacent or not adjacent. The number of labeled graphs with k edges is therefore $\binom{q}{k} = \binom{\binom{p}{2}}{k}$. Thus, the ordinary generating function $G_p(x)$ for labeled graphs with p vertices is given by

$$G_p(x) = \sum_{k=0}^m \binom{m}{k} x^k = (1 + x)^m, \tag{A8}$$

where $m = \binom{p}{2}$. Then, the number of labeled graphs with p vertices is $G_p(1)$ so we have that

$$G_p = 2^m = 2^{\binom{p}{2}}. \tag{A9}$$

For example, if we want to know how many labeled graphs with $p = 3$ vertices can be obtained, we apply the formula (A9) and we get

$$G_3 = 2^{\binom{3}{2}} = 2^{\frac{3!}{2!1!}} = 2^3 = 8. \tag{A10}$$

If we want to know how many labeled graphs with $p = 4$ vertices and exactly $q = 5$ edges exist, before the expression (A8), we use the coefficient the term x^5

$$[x^5]G_p(x) = [x^5] \sum_{k=0}^6 \binom{6}{k} x^k = \binom{6}{5} = \frac{6!}{5!1!} = 6. \tag{A11}$$

When we are working with labeled graphs, it is very normal to ask in how many ways can be labeled a graph. To give an answer, the symmetries or automorphisms of a graph have to be considered. A graph isomorphism between a graph G and a graph G_1 is a one-to-one map $A : V(G) \mapsto V(G_1)$ that preserve adjacency. If $G_1 = G$, then A is called automorphism of G . The collection of all automorphisms of G , represented by $\Gamma(G)$, constitutes the group of G . The elements of $\Gamma(G)$ are permutations over V .

Let $s(G) = |\Gamma(G)|$ be the order of the group or number of symmetries of G . Therefore, the number of ways in which a graph G of order p can be labeled is

$$l(G) = \frac{p!}{s(G)}. \tag{A12}$$

Another illustrative example of graph enumeration problem is to enumerate labeled connected graphs. A path of length n can be defined as a sequence of vertices $\{v_0, v_1, \dots, v_n\}$ such that the edges involved $\{v_i, v_{i+1}\}$ for $i = 0, \dots, n$ are distinct. A connected graph is a graph in which any two vertices are joined by a path. If we want to obtain a formula for enumerating all the connected graphs C_p of order p , we will have to resort to the notion of subgraph. A graph H is a subgraph of a graph G if $V(H) \subset V(G)$ and $E(H) \subset E(G)$. A component is a maximal connected subgraph. A rooted graph is a graph that has a distinguished vertex called root. Two rooted graphs H_1 and H_2 are isomorphic if there exists a one-to-one function from $f : V(H_1) \mapsto V(H_2)$ that preserves the adjacency relation and the roots. A similar definition applies to the case of labelled graphs.

Let us assume that a_k for $k = 1, 2, 3, \dots$ is the number of ways of labeling all graphs of order that fulfil the property $P(a)$ and whose exponential generating function is

$$a(x) = \sum_{k=1}^{\infty} \frac{a_k x^k}{k!}. \tag{A13}$$

Similarly, let us assume that b_k for $k = 1, 2, 3, \dots$ is the number of ways of labeling all graphs of order that fulfil the property $P(b)$ and whose exponential generating function is

$$b(x) = \sum_{k=1}^{\infty} \frac{b_k x^k}{k!}. \tag{A14}$$

If we make the product of series (A13) and (A14), the coefficients of $\frac{x^k}{k!}$ in $a(x)b(x)$ is the number of ordered pairs (G_1, G_2) of two disjoint graphs, where G_1 meets the property $P(a)$, G_2 fulfils the property $P(b)$, k is the number of vertices in $G_1 \cup G_2$ and the labels from 1 to k have been distributed over $G_1 \cup G_2$. If $C(x)$ is the exponential generating function for labeled connected graphs

$$C(x) = \sum_{k=1}^{\infty} \frac{C_k x^k}{k!}, \tag{A15}$$

then $C(x)C(x)$ is the generating function for ordered pairs of labeled connected graphs. If we divide by 2 this generating function, we obtain the generating function for labeled graphs having exactly two components. If we make the same operation n times, the coefficient of $\frac{x^k}{k!}$ represents the number of labeled graphs of order k with exactly n components

$$G(x) = \sum_{n=1}^{\infty} \frac{C^n(x)}{n!}. \tag{A16}$$

From (A16), we obtain the following relation

$$1 + G(x) = e^{C(x)}. \tag{A17}$$

Riordan in [56] obtained the relation $C_p = J_p(2)$, where $J_p(x)$ enumerates the trees by of inversions and then deduced

$$C_p = \sum_{k=1}^{p-1} \binom{p-2}{k-1} (2^k - 1) C_k C_{p-k}. \tag{A18}$$

From (A18), it can be noticed that, if the exponential generating function for a class of graphs is known, then the exponential generating function for the class of graphs will be the logarithm of the first series, just as in (A17).

Another equivalent recurrence that can be obtained for enumerating the connected labeled graphs of order p (p. 7 in [35]) is:

$$C_p = 2^{\binom{p}{2}} - \frac{1}{p} \sum_{k=1}^{p-1} k \binom{p}{k} 2^{\binom{p-k}{2}} C_k. \quad (\text{A19})$$

Using (A19), the first author of the present article implemented the following Matlab (R2019a, MathWorks, Natick, Massachusetts, MA, USA) code that can help us to obtain the number of connected graphs C_p of orders going from $p = 1$ to $p = 20$.

```
function y = CuentaGrafConnEtiq( p )

% recurrence satisfied by the number of connected graphs
% Harary Graph Enumeration pag. 7
% C_p= 2^{combinations(p,2)}-1/p* \sum_{k=1}^{p-1} k*
% combinations(p,k)*2^{combinaciones(p-k,2)}*C_{k}

C(1:p)=0;
C(1,1)=1;

for k=2:p
    acum=0;
    for j=1:k
        acum = acum + j * combinaciones(k,j) * CuentaGrafEtiq(k-j) * C(1,j);
    end
    C(1,k) = CuentaGrafEtiq(k)-(1/k)*acum;
end
y=C(1,p);
sprintf('%u',y) %% displaying the value with unsigned integer format
end

function z=combinaciones(n,k)
z= factorial(n)/(factorial(k)*factorial(n-k));
end

%% calling the function from the matlab prompt for the calculation of the
%% evaluation from graphs of order 1 to 20

>> for i=1:20
R(1,i)=CuentaGrafConnEtiq( i );
end
```

We get the following table (Table A1).

Table A1. Order 1 to 20.

p	C_p
1	1
2	1
3	4
4	38
5	728
6	26,704
7	1,866,256
8	251,548,592
9	66,296,291,072
10	34,496,488,594,816
11	35,641,657,548,953,344
12	7.335460×10^{19}
13	3.012722×10^{23}
14	2.471649×10^{27}
15	4.052768×10^{31}
16	1.328579×10^{36}
17	8.708969×10^{40}
18	1.41641×10^{46}
19	2.992930×10^{51}
20	1.569216×10^{57}

As we can see in the table above, the number of possible connected graphs C_p grows very fast with respect to the number p of vertices. Concerning the expressions (A18) and (A19), it can be mentioned that they are recurrence relations instead of a closed formula. The recurrences (A18) or (A19) can be used for performing the calculation of C_p with a computer program. Some recurrences are very hard to be solved and others cannot be solved, but, for obtaining an approximate value for p very big, there are methods in analytic combinatorics that help us to calculate a very good approximation of the p -th coefficient. The generating functions a very important concept in combinatorial theory whose algebraic structure allow to reflect the structure of combinatorial classes. The approach taken by the analytic combinatorics is to examine the generating functions from the point of view of the mathematical analysis by assigning not only real value values to its variables but also values in the complex plane. The assignation of complex values to the variables of the generating functions gives as a consequence that the function becomes a geometric transformation of the complex plane. This type of geometrical mapping is regular (holomorphic) near the origin of the complex plane. Far from the origin some singularities appear that correspond to the absence of smoothness of the function and give a lot of information about the function coefficients and their asymptotic growth. Sometimes, elementary real analysis suffices for estimating asymptotically enumerative sequences. At the next level of difficulty, the generating functions are still explicit, but its form does not allow for obtaining the coefficients of the series easily. In such cases, the complex plane analysis is a good option for estimating asymptotically these coefficients.

With the purpose of illustrating the notion of singularities, let us take as an example the ordinary generating function of the Catalan numbers

$$f(x) = \frac{1}{2}(1 - \sqrt{1 - 4x}). \tag{A20}$$

The expression (A20) is a compact description of the power series of the form

$$(1 - y)^{1/2} = 1 - \frac{1}{2}y - \frac{1}{8}y^2 - \dots \tag{A21}$$

The coefficients of the generating function associated with (A20) have the following explicit form

$$f_n = [x^n]f(x) = \frac{1}{n} \binom{2n-2}{n-1}. \tag{A22}$$

If we use Stirling formula, we can obtain the following asymptotic approximation to (A22)

$$f_n \sim \lim_{n \rightarrow \infty} \frac{4^n}{\sqrt{\pi n^3}}. \tag{A23}$$

The approximation of the kind of (A23) can be derived by use of the generating functions as analytic objects. For this end, we can substitute in the power series expansion of the generating function $f(x)$ any real or complex value ρ_f whose modulus small enough for example $\rho_f = 4$. The graph associated with (A20) is smooth and differentiable in the real plane and tends to the limit $\frac{1}{2}$ as $x \rightarrow (\frac{1}{4})^-$, but, if we calculate its derivative, we obtain the following function

$$f(x) = \frac{1}{1 - \sqrt{1 - 4x}}, \tag{A24}$$

and it can be noticed that the derivative (A24) becomes infinite in $\rho_f = \frac{1}{4}$. The points where the smoothness stops are called singularities. It can also be observed that the region where function (A20) is still continuous can be extended; let us take, for instance, $x = -1$

$$f(-1) = \frac{1}{2}(1 - \sqrt{5}). \tag{A25}$$

We can proceed in a similar way and try to evaluate (A20) with values in the complex plane whose modulus is less than the radius of convergence of the series defined by (A20) and observe that the orthogonal and regular grid in that can be defined in the real plane get transformed in a grid on the complex plane that preserves the angles of the curves of the grid which correspond to the property of complex differentiability and which also is equivalent to the property of analyticity. As regards the asymptotic behavior of the coefficients f_n of the generating function, it can be observed that it has a general asymptotic pattern composed by an exponential growth factor A^n and a subexponential factor $\theta(n)$. In the case of (A23) $A = 4$ and $\theta(n) \sim \frac{1}{4}(\sqrt{\pi n^3})^{-1}$, we can relate the exponential growth factor with the radius of convergence of the series by $A = \frac{1}{\rho_f}$ that is the singularity found along the positive real axis of the complex plane that in general correspond to the pole of the generating function and the subexponential part $\theta(n) = O(n^{-\frac{3}{2}})$ arises from the singularity of the square root type. This asymptotic behavior can be summarized as follows:

$$[x^n]f(x) = A^n \theta(n). \tag{A26}$$

The exponential growth part of (A26) is known as first principle of coefficient asymptotics and the subexponential growth part as second principle of coefficient asymptotics. More general generating functions can be addressed with complex variable theory results as can be the Cauchy residue theorem that relates global properties of a meromorphic function (its integral along closed curves) to purely local characteristics at the residues poles. An important application of the Cauchy residue theorem concerns a coefficient of analytic functions. This is stated in the following theorem [48]

Theorem A3. (Cauchy's coefficient formula). Let $f(z)$ be analytic in a region containing 0 and let λ be a simple loop around 0 that is positively oriented. Then, the coefficient $[z^n]f(z)$ admits the integral representation

$$f_n \equiv [z^n]f(z) = \frac{1}{2i\pi} \int_{\lambda} f(z) \frac{dz}{z^{n+1}}. \tag{A27}$$

For more details about Analytic Combinatorics, we recommend to consult [48] as well as [50]

References

- Alarcón Ramos, L.A.; Bernal Jaquez, R.; Shaum, A. Output-Feedback Control for Discrete-Time Spreading Models in Complex Networks. *Entropy* **2018**, *20*, 204. [[CrossRef](#)]
- Graham, R.L.; Knuth, D.E.; Patashnik, O. *Concrete Mathematics*; 6th Printing with Corrections; Addison-Wesley: Boston, MA, USA, 1990
- Erdős, P. Graph theory and Probability. *Can. J. Math.* **1959**, *11*, 34–38. [[CrossRef](#)]
- Alon, N.; Spencer, J.H. *The Probabilistic Method*, 2nd ed.; New York Wiley-Interscience: New York, NY, USA, 2000.
- Chakrabarti, D.; Wang, Y.; Wang, C.; Leskovec, J.; Faloutsos, C. Epidemic thresholds in real networks. *ACM Trans. Inf. Syst. Secur.* **2008**, *10*, 1–26. [[CrossRef](#)]
- Galam, S. Rational Group Decision Making: A random field Ising model at $T = 0$. *arXiv* **1997**, arXiv:cond-mat/9702163v1.
- Galam S, From 2000 Bush? Gore to 2006 Italian elections: Voting at fifty-fifty and the contrarian effect. *Qual. Quant.* **2007**, *41*, 579–589. [[CrossRef](#)]
- Galam, S. Social Paradoxes of Majority Rule Voting and Renormalization Group. *J. Stat. Phys.* **1990**, *61*, 943–951. [[CrossRef](#)]
- Galam, S. Sociophysics: A review of Galam models. *arXiv* **2008**, arXiv:physics.soc-ph/0803.183v1.
- Gómez, S.; Gómez-Gardeñes, J.; Moreno, Y.; Arenas, A. Non-perturbative heterogeneous mean-field approach to epidemic spreading in complex networks. *Phys. Rev. E* **2011**, *84*, 036105. [[CrossRef](#)]
- Axelrod, R. The Dissemination of Culture: A model with local convergence and global polarization. *J. Confl. Resolut.* **1997**, *41*, 203–226. [[CrossRef](#)]
- Gonzalez-Avella, J.C.; Eguiluz, V.M.; Cosenza, M.G.; Klemm, K.; Herrera, J.L.; San Miguel, M. Nonequilibrium transition induced by mass media in a model for social influence. *Phys. Rev. E* **2005**, *72*, 065102. [[CrossRef](#)]
- Gonzalez-Avella, J.C.; Cosenza, M.G.; Tucci, K. Local versus global interactions in nonequilibrium transitions: A model of social dynamics. *Phys. Rev. E* **2006**, *73*, 046119. [[CrossRef](#)]
- Klemm, K.; Eguiluz, M.; Toral, R.; San Miguel, M. Nonequilibrium transitions in complex networks: A model of social interaction. *Phys. Rev. E* **2003**, *67*, 026120. [[CrossRef](#)]
- Liu, Y.-Y.; Slotine, J.-J.; Barabási, A.-L. Controllability of complex networks. *Nature* **2011**, *473*, 167–173. [[CrossRef](#)] [[PubMed](#)]
- Nepusz, T.; Vicsek, T. Controlling edge dynamics in complex networks. *Nat. Phys.* **2012**, *8*, 568–573. [[CrossRef](#)]
- Pasqualetti, F.; Zampieri, S.; Bullo, F. Controllability metrics, limitations and algorithms for complex networks. *IEEE Trans. Control Netw. Syst.* **2014**, *1*, 40–52. [[CrossRef](#)]
- Lombardi, A.; Hörnquist, M. Controllability analysis of networks. *Phys. Rev. E* **2007**, *75*, 056110. [[CrossRef](#)]
- Tanner, H.G. On the controllability of nearest neighbor interconnections. In Proceedings of the 2004 CDC—43rd IEEE Conference on Decision and Control, Nassau, Bahamas, 14–17 December 2004; Volume 3, p. 2467.
- Kalman, R.E. Mathematical description of linear dynamical systems. *J. Soc. Ind. Appl. Math. Ser. A* **1963**, *1*, 152–192. [[CrossRef](#)]
- Ahn, H.J.; Hassibi, B. Global dynamics of epidemic spread over complex networks. In Proceedings of the 2013 IEEE 52nd Annual Conference on Decision and Control, Florence, Italy, 10–13 December 2013; pp. 4579–4585.
- Cullen, C.G. *Matrices and Linear Transformations*, 2nd ed.; Addison-Wesley: Boston, MA, USA, 1972.
- Gerschgorin, S. Über die Abgrenzung der Eigenwerte einer Matrix. *Izv. Akad. Nauk. USSR Otd. Fiz.-Mat. Nauk.* **1931**, *6*, 749–754.
- Barabási, A.L.; Albert, R. Emergence of scaling in random networks. *Science* **1999**, *286*, 509–512.
- Wang, W.; Tang, M.; Stanley, H.E.; Braunstein, L.A. Unification of theoretical approaches for epidemic spreading on complex networks. *Rep. Prog. Phys.* **2017**, *80*, 036603. [[CrossRef](#)]

26. Prakash, B.A.; Chakrabarti, D.; Faloutsos, M.; Valler, N.; Faloutsos, C. Got the Flu (or Mumps)? Check the Eigenvalue! *arXiv* **2010**, arXiv:physics.soc-ph/1004.0060v1.
27. Courcelle, B.; Makowsky, J.A.; Rotics, U. On the fixed parameter complexity of graph enumeration problems definable in monadic second order logic. *Discret. Appl. Math.* **2001**, *108*, 23–52. [[CrossRef](#)]
28. Nijenhuis, A.; Wilf, H.S. The Enumeration of Connected Graphs and Linked Diagrams. *J. Comb. Theory* **1979**, *27*, 356–359. [[CrossRef](#)]
29. Ronald, C. Read, Some unusual enumeration problems. *Ann. N. Y. Acad. Sci.* **1970**, *175*, 314–326.
30. Mackay Brendan, D. Asymptotics for symmetric 0–1 matrices with prescribed row sums. *Ars Comb.* **1985**, *19A*, 15–25.
31. Mackay Brendan, D.; Wormald Nicholas, C. Uniform generation of random regular graphs of moderate degree. *J. Algorithms* **1990**, *11*, 52–67. [[CrossRef](#)]
32. Mackay Brendan, D.; Wormald Nicholas, C. Asymptotic enumeration by degree sequence of graphs of high degree. *Eur. J. Comb.* **1990**, *11*, 565–580. [[CrossRef](#)]
33. Mackay Brendan, D.; Wormald Nicholas, C. Asymptotic Enumeration by Degree Sequence with Degrees $O(n^{\frac{1}{2}})$. *Combinatorica* **1991**, *11*, 369–382. [[CrossRef](#)]
34. Pólya, G. Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen. *Acta Math.* **1937**, *68*, 145–254. [[CrossRef](#)]
35. Harary, F.; Palmer, E.M. *Graph Enumeration*; Academic Press: New York, NY, USA; London, UK, 1973.
36. Wright, E.M. Graphs on unlabelled nodes with a given number of edges. *Acta Math.* **1971**, *126*, 1–9. [[CrossRef](#)]
37. Bender, E.A.; Canfield, E.R. The asymptotic number of labeled graphs with given degree sequences. *J. Comb. Theory* **1978**, *24*, 296–307. [[CrossRef](#)]
38. Bollobás, B. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *Eur. J. Comb.* **1980**, *1*, 311–316. [[CrossRef](#)]
39. Bollobás, B. The asymptotic number of unlabelled regular graphs. *J. Lond. Math. Soc.* **1981**, *1*, 201–206. [[CrossRef](#)]
40. Pei, S.; Morone, F.; Makse, H.A. Theories for influencer identification in complex networks. *arXiv* **2018**, arXiv:physics.soc-ph/1707.01594v2.
41. Cha, M.; Haddadi, H.; Benevenuto, F.; Gummandi, P.K. Measuring user influence in Twitter: The million follower fallacy. In Proceedings of the 4th International AAAI Conference on Weblogs and Social Media, Washington, DC, USA, 23–26 May 2010; Volume 10, pp. 10–17.
42. Watts, D.J.; Dodds, P.S. Influential networks and public opinion formation. *J. Consum. Res.* **2007**, *34*, 441–458. [[CrossRef](#)]
43. Kitsak, M.; Gallos, L.K.; Havlin, S.; Liljeros, F.; Muchnik, L.; Stanley, H.E.; Makse, H.A. Identification of influential spreaders in complex networks. *Nat. Phys.* **2010**, *6*, 888–893. [[CrossRef](#)]
44. Pei, S.; Makse, H.A. Spreading dynamics in complex networks. *J. Stat. Mech. Theory Exp.* **2013**, *2013*, P12002. [[CrossRef](#)]
45. Min, B.; Morone, F.; Makse, H.A. Searching for influencers in big-data complex networks. In *Diffusive Spreading in Nature, Technology and Society*; Bunde, A., Caro, J., Karger, J., Vogl, G., Eds.; Springer: Berlin, Germany, 2016.
46. Leskovec, J.; Adamic, L.A.; Huberman, B.A. The dynamics of viral marketing. *ACM Trans. Web* **2007**, *1*, 5. [[CrossRef](#)]
47. Rogers, E.M. *Diffusion of Innovations*; Simon and Schuster: New York, NY, USA, 2010.
48. Sedgewick, R.; Flajolet, P. *Analytic Combinatorics*, Zeroth ed.; Fifth Printing; Cambridge University Press: Cambridge, UK, 2005.
49. Wilf, H.S. *Generatingfunctionology*, 3rd ed.; A K Peters Ltd.: Wellesley, MA, USA, 2006.
50. Louis, C. *Advanced Combinatorics: The Art of Finite and Infinite Expansions*; D. Reidel Publishing Company: Dordrecht-Holland, The Netherlands; Boston, MA, USA, 1974.
51. Sedgewick, R.; Flajolet, P. *An Introduction to the Analysis of Algorithms*; 2nd Printing; Addison-Wesley: Boston, MA, USA, 2001.
52. Euler, L. *Novi Commentarii Academiae Scientiarum Imperialis Petropolitanae*; Holding Institution, American Museum of Natural History Library: New York, NY, USA, 1750–1776; Volume 7, pp. 13–14. Available online: <https://www.biodiversitylibrary.org/bibliography/9527#/summary> (accessed on 21 March 2018).

53. Kirchhoff, G. Über die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Verteilung galvanischer Ströme geführt wird. *Ann. Phys. Chem.* **1847**, *72*, 497–508. [[CrossRef](#)]
54. Cayley, A. A Theorem on trees. *Q. J. Math. Oxf. Ser. Collect. Pap. Camb.* **1897**, *13*, 26–28.
55. Redfield, J.H. The theory of group-reduced distributions. *Am. J. Math* **1927**, *49*, 433–455. [[CrossRef](#)]
56. Mallows, C.L.; Riordan, J. The inversion enumerator for labeled trees. *Bull. Am. Math. Soc.* **1968**, *74*, 92–94. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).