Alma Mater Studiorum · Università di Bologna

Scuola di Scienze
Dipartimento di Fisica e Astronomia
Corso di Laurea Magistrale in Fisica

Random Matrix Theory and Renormalization Group: Spectral Theory for Network Ensembles

Relatore:
Prof. Armando Bazzani

Presentata da:
Alessio Catanzaro

Anno Accademico 2018/2019

Two is company, three is complexity.

Neil F. Johnson

Abstract

In questa tesi si affronta il tema delle proprietà spettrali dei network. La teoria dei grafi ha sviluppato metodi per studiare le proprietà spettrali di matrici stocastiche e di adiacenza dei network basandosi su tecniche algebriche, mentre le applicazioni alla teoria dei sistemi complessi si basano essenzialmente su metodi di fisica statistica. Adottiamo un approccio al problema sfruttando i risultati della Random Matrix Theory, collegandoli ad alcune tecniche di meccanica statistica. La RMT permette di chiarire il significato fisico della legge di Wigner nel caso di random network nel limite termodinamico e di calcolarne alcune correzioni. In seguito si propone un collegamento tra i risultati della RMT e le tecniche del Gruppo di Rinormalizzazione per analizzare le proprietà spettrali delle classi di network Scale-Free e Small-World. In particolare si propone un meccanismo che potrebbe contribuire a comprendere il comportamento di strutture di network autosimili a basso diametro.

In this thesis we tackle the issue of spectral properties of networks introduced by Complex Systems Physics. Graph theory has developed methods to study the spectral properties of adjacency or stochastic matrices associated to networks based on algebraic techniques, whereas the applications to Complex Systems Theory have been essentially based on the methods of Statistical Mechanics. We use an approach to the problem using the results of the RMT in connection with some statistical mechanics techniques. RMT is of use in clarifying the physical meaning of Wigner law for large, random networks, and to compute its corrections. Then, we try to build a bridge between the sound results of RMT and Renormalisation Group methods in order to investigate the spectral properties of the Scale Free and Small World class of networks. In particular, we propose a mechanism who could help understanding the behaviour of self similar network structures with low diameter.

Contents

Introduction

Complex systems are not complicated ones. There must be more to them than that[1]. In complicated systems the overall behaviour is the all-embracing demeanor of individual agents. Complexity is a different thing [24]. A complex system requires an approach that focuses on the interactions among the agents that compose it; an interwoven nexus rather than a folded map. Finding a description, if not a proper definition, for alike yet vastly heterogeneous systems of this kind is not trivial. Wide and yet to be refined definitions are those that refer to complex systems as "those that display emergent features", or "non-trivially regular systems" or even philosophically as "those where the reductionist paradigm fails". However none of these definitions are complete nor, in fact, actually distinct from one another. Emergence, strong (nonlinear) interactions and the failure of reductionist approaches are all features that require a different approach from the ones customarily used in almost all fields of science. In the present work we tackle some of the issues of Complex Systems Physics by the approach of Network Science. Complex networks have been revealed to be powerful tools to understand and portray some of the emergence and non linear phenomena that characterise complexity [38], [37]. Classic graph theory has been a cradle for the study of networks, nonetheless the usual approach of exactly solving for specific graph attributes lacks some of the universality features we look for when speaking about complexity. The approach based on spectra of networks may reveal to be of greater generality. The need to seek for spectral properties of network classes leads to the study of Random Matrix Theory [3][27], which reveals to be a powerful tool to understand some key features of the issue. Random Matrices Theory is an edge-cutting topic in contem-

[1] According to its etymology, complicated means "repeated folding" (*cum + plecare*) of diverse parts in a single item. Conversely, complex comes from *cum + plectere*, meaning the repeated weaving of the parts.

porary mathematical research as it finds applications in data analytics [43], financial markets modeling [2], control theory and so on. These fields are closely related to Complex Systems Science, as they share common points like long range dependence between agents, extreme events appearance, and the overall need to use statistical methods. Moreover the usage of Stochastic matrices - describing random walk dynamics over a graph - make the study of properties of random matrices even more relevant. We will show how spectral densities reflect prominent general features of graph ensembles, namely symmetry in the spectrum, indicating a tree-like underlying structure and bipartiteness; bounds for the eigenvalues, setting the scales for the relaxation time of the random walk; self-similarity, conveying the fractal structure of certain classes of networks and so on. This analysis is complemented by the implementation of the renormalisation group technique to the field of networks, showing the presence of different "phases" of networks which we try to identify. Finally, in the framework of renormalisation group, we try to explain how some self-similar structures may also display small world features. This last issue is of interest because it is commonly thought [40] [5] that Complex Networks found in nature should display self similarity structures with navigability properties that purely scale-free models seldom replicate.

A more detailed outline of the thesis is the following:

- In the first chapter we outline the issue of the dynamics of a Random Walk on a network, which is equivalent to a Markov process. We give examples of applications trying to emphasise the natural link with network science;

- In the second chapter we give some methods of work in the field of spectral distribution of random matrices, introducing random matrix theory. Some relevant results are obtained, among them the correction to the finite size and non zero average distribution of entries;

- In the third chapter we expose the basics of Network Science. Then we present a section where spectra of some known graphs are given. We also try to model the behaviour of the spectra for the ensemble of Scale-Free networks. Finally a physical interpretation for the results of the spectra of a random walk on a network is proposed;

- In the fourth chapter the analysis over the classes of different network is performed with Renormalisation Group techniques. The result are

commented highlighting the relation with the previous results and giving way to an overall picture which is finalized in the last chapter. We also propose a way of classification of network "phases" or ensembles according to spectral properties defining some distinctive dynamical traits.

To speak the truth, this thesis should start with the analysis performed in the last chapter, because it was the question "What if we applied renormalisation to networks?" that gave the spark that ignited everything up. However, trying to explain the different phases behaviour needed a sound background of both network science concepts and random matrix spectral properties. Hence, the topics were rearranged in this order to provide for a more linear flow of thought.

Chapter 1

Physical Motivations

The physics of complex systems is a vast and heterogeneous topic [24]. In this introductory chapter we give the gist of the issue in a twofold way. First we analyse the prominent features of a random walk performed on a network, focusing on the spectral decomposition of the stochastic matrix that describes the walk. Then we present two (among many others) simple applications in which we focus on how the network features are relevant to the issue of complexity and why they are a suitable instrument to models such systems. The first example is the spread of epidemics [20] on network structures. We put the emphasis on the structural properties of the network and how different topologies give rise to a different outbreak probabilities. The second example is to show how networks are ubiquitous and how they are an inherent feature of society [19]. We make the case of the "networking" issue when it comes to employment.

1.1 Elements of random walk dynamics

A random walk on a network may be easily understood as a particle, or a set of non interacting[1] particles, moving from a node of the network to another one, the probability of the transition depending on how the network is connected. Needless to say that this portrays a Markov process, *viz.* a so-called "process with no memory". If we indicate with the probability of being in the state i at time t with $p_i(t)$, then the evolution equation reads

[1] By non interacting we mean also that the presence of a particle on a certain node does not alter the chances of the others to fall in that very node.

$$p_i(t+1) = \sum_j \pi_{ij} p_j(t)$$

where π_{ij} were introduced, that is the probability of passing from node j to node i. There are some constraints for π_{ij}, namely

$$\sum_i \pi_{ij} = 1 \quad \text{and} \quad 0 \leq \pi_{ij} \leq 1$$

to make sure that probability (flow of particles) is conserved. We arrange those coefficients in a matrix Π and we call it the transition probability matrix, or *stochastic matrix*. Indeed, the stochastic matrix wholly represents the process, as it describes every possible shift from one state to another.

It has the following properties:

- It always has the maximal eigenvalue $\lambda = 1$, following from Perron-Frobenius theorem (see [17], pag 53.).
- Since the possible physical states are in the first quadrant, which is invariant, there exists an eigenvector with eigenvalue $\lambda = 1$ in the first quadrant $p_i^s \geq 0$. We call this the "stationary distribution".
- If the multiplicity of the eigenvalue $\lambda = 1$ is one, the stationary state is unique and attractive. If the multiplicity is greater than one, the network is disconnected and each component has a stationary distribution, the multiplicity being the number of separate components.
- The relaxation time towards the stationary distribution depends on the "Spectral Gap", that is the difference between the two greatest eigenvalues $|1 - \lambda_2|$ of the network.

If we interpret the Markov chain as a random walk on a network, then the stochastic matrix stems from the adjacency matrix of the underlying network as $\Pi = \Delta^{-1} A$, where A is the adjacency matrix and Δ the diagonal degree matrix ($\Delta_{ij} = \delta_{ij} d(i)$). They are related in the following way. We perform the similarity operation

$$\Delta^{\frac{1}{2}} \Pi \Delta^{-\frac{1}{2}} = \Delta^{-\frac{1}{2}} A \Delta^{-\frac{1}{2}} = \Sigma$$

being Σ a symmetric matrix, we know that it has real eigenvalues, and so does Π as similarity transformations do not alter the spectrum. Still, the

relation among the stochastic matrix, the adjacency one and the diagonal degree one is deeper. Indeed, we may define the scalar product between states as $\langle \phi, \psi \rangle = \phi \cdot \Delta^{-1} \psi$. Then we get that the stochastic matrix Π is symmetric with respect to this product

$$\langle \phi, \Pi\psi \rangle = \phi \cdot \Delta^{-1} A \Delta^{-1} \psi = A\Delta^{-1}\phi \cdot \Delta^{-1}\psi = \langle \Pi\phi, \psi \rangle$$

which means that the degree matrix induces a metrics to whom the stochastic matrix is symmetric. In this metrics, we also see that the eigenvalue problem of the stochastic matrix reduces to the eigenvalue problem of the adjacency matrix with respect to this newly defined metrics, *i.e.*

$$\det(\Pi - \lambda I) = \det\left(\Delta^{-1}\right) \det(A - \lambda\Delta) = \prod_{i=1}^{N} \frac{1}{d(i)} \det(A - \lambda\Delta)$$

This means that the distribution of the degree of the network play a key role in understanding the dynamics which is taking place on the graph. Indeed, as we will see in the followings, one of the most used classifications of networks is the one based on degree distributions.

Finally, we show how the dynamics depends on the spectrum of the stochastic transition matrix. As Σ is symmetric, it may be decomposed in terms of orthogonal matrices O as

$$\Sigma = O\Lambda O^T = \sum_{i=1}^{N} \lambda_i v_i v_i^T$$

where the representation in terms of the orthonormal eigenvectors that compose the matrix O was made explicit. Then the stochastic matrix may be decomposed as

$$\Pi = \sum_{j=1}^{N} \lambda_j \Delta^{-\frac{1}{2}} v_j v_j^T \Delta^{\frac{1}{2}}$$

Taking into consideration the so-called "basic law of the degree"

$$\frac{d}{2L} \cdot u^T = 1$$

where d is the degree vector, u the unitary one, L the total number of edges in the network[2], we may write the stationary distribution $p^* = \frac{d}{2L}$ or, equivalently, as

$$p_j^* = \frac{d_j}{2L}$$

Moreover, the similarity transform $\Delta^{\frac{1}{2}}$ maps the stationary state onto $v_i = \Delta^{-\frac{1}{2}} p^*$ so we may write the first of the eigenvectors of Σ as

$$v_{1j} = \frac{\frac{\sqrt{d_j}}{2L}}{\sqrt{\sum_{j=1}^{N} \left(\frac{\sqrt{d_j}}{2L}\right)^2}} = \sqrt{\frac{d_j}{2L}} = \sqrt{p_j^*}$$

and using this expression in the decomposition of Π we have

$$\Pi = up^* + \sum_{i=2}^{N} \lambda_j \Delta^{-\frac{1}{2}} v_j v_j^T \Delta^{\frac{1}{2}}$$

and for the l^th step transition probability from state i to j

$$\Pi_{ij}^l = \frac{d_j}{2L} + \sqrt{\frac{d_j}{d_i}} \sum_{k=2}^{N} \lambda_k^l (v_k v_k^T)_{ij}$$

This expression allow to express the convergence rate of any state to the stationary one. We may estimate the distance between a state and the stationary one as

$$|\Pi_{ij}^l - p_j^*| \leq \sqrt{\frac{d_j}{d_i}} \sum_{k=2}^{N} |\lambda_k^l| |(v_k v_k^T)_{ij}| \leq \sqrt{\frac{d_j}{d_i}} \sum_{k=2}^{N} |\lambda_k^l|$$

In particular, if we take the largest of the eigenvalues apart from $\lambda_1 = 1$, which is the one coupled with the stationary state and call it λ_M, we may estimate the distance from stationary state as

$$|\Pi_{ij}^l - p_j^*| < \sqrt{\frac{d_j}{d_i}} \lambda_M^l + O(\lambda_{k \neq 1,2})$$

hence, the definition of the spectral gap.

[2]This is immediate from the fact that the sum of all the degrees in a graph must be twice the number of links.

1.2 Some examples

The vastness of the applications of network science has already been mentioned, however, some examples may be useful in order to get deeper into the topic. We show here two cases in which modelling of particular systems requires the knowledge of the network structure on which it is based. In the first, we focus on the relevance of how different structural properties may give rise to quite different thresholds for the same transition phenomenon. In the second we highlight how the network features inherently modify the long-term state characteristics.

1.2.1 Disease spread

Epidemic models have been among the first cases of application of network science. In some cases, they even date back the topic [21], given the importance of the issue. Indeed, outbreaks of epidemics seem to be classic threshold phenomena often encountered in complex systems. The vast majority of epidemic models derive from the SIS model or its refinement, the SIR model. The SIS (Susceptible - Infectious - Susceptible) model is based on the division of the population in two states: Susceptible to the infection and infected, (S/I) and the fraction of the population at time t that are respectively sane (but susceptible) and ill are $s(t)$ and $i(t)$. The model reads

$$\begin{cases} \frac{ds(t)}{dt} = gi(t) - vs(t) \\ \frac{di(t)}{dt} = vs(t) - gi(t) \end{cases}$$

which is nothing more than a revisited Lotka-Volterra population dynamic model in which g is the rate of healing, and v is the expansion rate of the infection. This kind of model may describe well sexually transmitted diseases for two reasons: the latency of the disease, which causes the infected individual not to know about his situation, thus causing repeated infections to occur; and the fact that in most cases immunity or healing cannot be reached. Moreover the latency of the disease may let us neglect the death rate of the disease. The SIR model is suitable to portray diseases, like flu, which confer instead life-long immunity through healing. Its acronym stands for Susceptible - Infected - Removed[3], which are the states that the individual

[3]The term may refer to the removal of the disease, through healing, but also to the removal of the individual because of death.

can be in. In this case we get Lotka-Volterra like equations

$$\begin{cases} \frac{ds(t)}{dt} = -\alpha i(t)s(t) \\ \frac{di(t)}{dt} = \alpha i(t)s(t) - \gamma i(t) \\ \frac{di(t)}{dt} = \gamma i(t) \end{cases}$$

in which we set α to be the rate at which the infected individuals spread the disease, while γ is the removal rate, that is the healing and dying rate of the disease. We remark that in this case "removed" individuals do not partake to the spread of the disease because they are either dead or they have acquired immunity through healing. When solving such system, the relation network is fundamental. Let us take two connected nodes, i, and j, each representing an individual. We take i to be ill, and j to be susceptible. The rate that sets the probability that i infects j is r_{ij}, and is a random variable depending on how frequently these people meet, how close their bond is, and so on. We also have the infectious time τ_i which is the time for which i remains infectious. The probability of non transmission between i and j is $1 - T_{ij} = (1 - r_{ij}\Delta t)^{\frac{\tau_i}{\Delta t}}$ where Δt is the interaction time. If we make the continuous limit we readily have that the probability of disease transmission between i and j is

$$T_{ij} = 1 - e^{-r_{ij}\tau_i}$$

and

$$T_{ij} = 1 - (1 - r_{ij})^{\tau_i}$$

in the discrete time step $dt = 1$. A reasonable assumption is that the variables r_{ij}, τ_i are drawn form a random distribution and are independent and identically distributed variables. If so, the mean transmission T probability is simply the average of T_{ij} over the two distributions $P(\tau)$ and $P(r)$

$$T = 1 - \int dr d\tau P(r)P(\tau)e^{-r\tau}$$

This allows us to greatly simplify the problem, because in the overall population the disease propagates as if the transmission probabilites were all T. The problem thus reduces to a Percolation problem. Percolation is one of the most studied and analysed problems in network theory, and statistical physics as a whole. It may be synthesised as the answer to the question:

"Given a source on a network, how does the flux flow on the grid?". Few analytical results are available in this field, as they often depends on the network structure[4].

Here we list some of the available results and thresholds for percolation theory to happen.

- Random Networks. The mixing that defines the individual who compose the populations is randomly performed. This means each node has on average pN neighbours, randomly chosen among the network population. In this case, the threshold for which the outbreak of the disease becomes epidemic is

$$T = \frac{1}{d}$$

  where d is the average degree of the network.

- Scale Free networks. These self-similar networks follow a power law distribution for the degree. In this case the "hub" phenomenon occurs, leading to a supposedly more rapid disease spread. If we take the degree distribution to be of the form

$$p(k) = Ck^{-a}e^{-\frac{k}{\kappa}}$$

  then the result in [31] states that the transition rate of the disease T leads to a critical epidemic behaviour when

$$T = \frac{Li_{a-1}(e^{-\frac{1}{\kappa}})}{Li_{a-2}(e^{-\frac{1}{\kappa}}) - Li_{a-1}(e^{-\frac{1}{\kappa}})}$$

  where $Li_n(x)$ is the n^{th} polylogarithm of x.

- Small world [30]. In this case, the scale free behaviour is kept but shortcuts between random, possibly even distant nodes are added. If we set the average number of shortcuts to be ϕ, then we have a threshold for the transition probability

$$T = \frac{\sqrt{4\phi^2 + 12\phi + 1} - 2\phi - 1}{4\phi} \simeq 1 - 4\phi + O(\phi^2)$$

[4] A sound definition of the different kind of structures a network possesses has not been reached yet. We may refer to concepts like "scale free", "random", "small world", which are better explained in chapters 3 and 4.

- Regular networks. In square regular lattices, the percolation threshold, *i.e.* the critical transmission rate is $t = \frac{1}{2}$ [6]. In k-regular trees, the percolation threshold is $T = \frac{1}{1-k}$.

The diversity of behaviours when taking into account for different structures is one of the crucial features of complex systems. Hence the need to better formalize the categories networks populate emerges.

1.2.2 Finding a job with NetSci

Networks seem to be naturally suited to model markets. The behaviour of individuals who exchange goods and services over a net of relations should be easily transposed over a graph with an overlying dynamics. Here we introduce some of the ideas used when modelling a seminal sector of economics: the job market. It is well known that finding a job is often an issue of being offered or knowing about the vacancy of a position[5]. Let us try to sketch the prominent feature of this complex market along with [7]. We begin by an undirected network in which nodes are workers, which have two statuses: employed, or jobless. There is a random chance that a worker may loose his position, becoming unemployed. Then, there is an information about a vacant position travelling over the nodes. If a node is employed, it passes the information over one neighbour with uniform probability. If the node is unemployed, it takes the vacant position[6]. It is immediately clear that the model favours the workers with more connections[7], which can find a job with more ease than a colleague who has fewer access to information. The matrix A is the adjacency matrix of such network, and the vector $O(t)$ is the vector that the describe the employment situation at time t, taking value $O_i(t) = 1$ if i is employed at time t and zero otherwise. At time $t + 1$ the following things happen:

- With probability a each node independently get to know about a job opportunity.
- If unemployed, a node takes the position, if already employed passes the information to an unemployed neighbour of his. The choice of the unemployed neighbours is random.

[5]Hence the very definition "networking" in the Merriam Webster dictionary: "the cultivation of productive relationships for employment or business".

[6]No choosy workers in this model.

[7]In socio-economical sciences, this is called "social capital".

- If an employed node knows about a position but has no unemployed neighbours, the position information is lost.
- At the end of the step, each node randomly looses hos job with probability b.

The probability that i get to know about a vacant job and this job is taken by j is

$$p_{ij}(O(t)) = \begin{cases} a & \text{if } O_i(t) = 0 \text{ and } i = j, \\ \frac{a}{\sum_{k:O_k(t)=0} A_{ik}} & \text{if } O_i(t) = 1,\ O_j(t) = 0 \text{ and } A_{ij} = 1, \\ 0 & \text{otherwise.} \end{cases}$$

The focus is over the long run expectation value of the employment vector, which we call μ and it is nothing but the stationary distribution of the Markov chain described by this model.

We begin by solving this model in the trivial case: a single worker. The stationary distribution is nothing but a number which must satisfy

$$\mu = \frac{1-b}{\mu + a(1-\mu)}$$

which is solved for μ to give the long term occupation chance of the worker

$$\mu = \frac{1}{1 + \frac{b}{(1-b)a}}$$

If we take the case for a two worker network, then the result for the chance of one of them to be employed is

$$\mu = \frac{a}{a + b - \frac{ab}{2a+b}}$$

which is greater than in the case of the single worker, showing that "networking" works. Another analytically solvable model is the one of the complete graph of N nodes. When the chance of being fired/offered a job is respectively $\frac{b}{T}$ and $\frac{a}{T}$, as T grows the probability of having k employed workers is

$$\mu_k = \frac{1}{\sum_{j=0}^{N} \frac{k!}{j!} \left(\frac{b}{Na}\right)^{k-j}}$$

The thing to point out is that as the network grows in size, the correlation among workers shrinks. Nonetheless the "network effect" does not. An interesting feature is that quite the contrary happens. As pointed out in [19] we may compute the variance of the probability of total employment of the stationary distribution in the network case and then normalize it by the variance of total employment in the case of a population of the same size who does not adopt networking as a way to get a job. We can see that it increases as the population grows in size, as shown in [19], figure 10.2.2, page 442. Even in a simple case like this one, we clearly see that networks play a prominent role in shaping the emerging behaviour of the systems.

Chapter 2

Elements of Random Matrix Theory

2.1 Introduction

This chapter is devoted to the exposure of the main results that Random Matrix Theory (RMT) develops which are relevant to the main goal of the present work. RMT is is a very active research field with numerous applications in Theoretical and Statistical Physics, Number Theory and Combinatorics. The recent development of Complex Systems Physics has offered the possibility of new applications of RMT. In the followings we present the main techniques that are relevant to provide insights in the study of complex networks. Besides, RMT is a field of knowledge worth studying *per se*, not only for the wide range of applications, but also as a generalization of many classic results from probability theory. The focus will be twofold, concerning about both the methods that allow to prove the cornerstone and birth act of RMT, *viz.* Wigner Semicircular Law, and their actual meaning in terms of network theory, trying to build a sound bridge between the fields. Indeed, the different approaches used to tackle the problem of determining the Spectral distribution of a certain random matrix will be of great utility in order to understand the development of the following chapters. However I have no presumption of exhaustiveness or to display such branches of mathematics thoroughly. A classic reference for those who are interested to be introduced to the field of RMT is [27], while for a more modern approach I recommend [3] and the very useful and updated site [35].

The study of the spectral properties of a random matrix is one of the main problems of RMT, the word "random" meaning that the entries of such a matrix are drawn from a distribution. This field of studies originated from the investigation of heavy nuclei Hamiltonian eigenvalue problem, which was brilliantly solved by E.P. Wigner in his seminal papers [41], [42], replacing the true Hamiltonian with a random one and calculating its spectral distribution. Wigner law for the distribution of eigenvalues of a random matrix has been generalised and expanded since, focusing about its domain of validity and its universality. In this chapter a four-fold view of such theorem will be presented, beginning with the classic moments method, showing that the ESD moments of a random matrix and those of the semicircular law coincide in the limit of large N; then by the Stieltjes transform it is shown how the semicircular law is the most probable distribution, again in the large N limit; then a brief *excursus* in the free probability theory is taken to show how Wigner law is recovered in the central limit theorem for non-commuting variables: finally the Replica approach is introduced, employed to get Wigner law and its first corrections.

2.2 RMT basic framework

The first and foremost task to comply with in order to tackle the problem of calculating the eigenvalues and the spectra of a random matrix is the definition of a framework in which usual probabilistic concepts are translated to. This section is devoted to introducing the basic tools and their meaning. A thorough exposure of such techniques lies beyond the purpose of this work. However, for a sound and comprehensive introduction to methods of RMT, we recommend the reference [3].

2.2.1 Empirical Spectral Distribution

The Empirical Spectral Distribution (ESD) of a random matrix is defined as follows

Definition 2.2.1 (Empirical Spectral Distribution)**.** The ESD of a $N \times N$ matrix A with real eigenvalues λ_i is

$$\rho_A(x) = \frac{1}{N} \sum_i \delta(x - \lambda_i) \tag{2.1}$$

2.2.2 Moments

We define the moments of a distribution in the conventional way as

Definition 2.2.2 (Moments of distribution)**.** The k-moment of a ESD ρ_A is

$$m_k = \int x^k \rho_A(x) dx$$

Moments are of use since they are the first method by which the central result of RMT is proven. Indeed, a canon result in classic probability theory is the following

Theorem 2.2.1 (Carleman, [3] page 509)**.** Let m_k be the moments of a distribution ρ, if the Carleman Condition

$$\sum_k m_{2k}^{-\frac{1}{2k}} = \infty$$

holds, then ρ is uniquely determined by the sequence of moments m_k.

A weaker condition is given by a corollary due to Riesz

Theorem 2.2.2 (Riesz, [3] page 508)**.** Let m_k be the moments of a distribution ρ, if

$$\liminf_{k\to\infty} \frac{1}{k} m_{2k}^{\frac{1}{2k}} < \infty$$

then ρ is uniquely determined by the sequence of moments m_k.

For the purposes and the scope of this work, the aforementioned conditions are more than sufficient. When dealing with the continuous type of ESD (Wigner, McKay, *etc.*), they reduce to the condition of the moments to be finite.

2.2.3 Stieltjes Transform

Another tool of great use in RMT is the resolvent of the matrix, which also develops a whole new theory to treat more handily ESDs.

Definition 2.2.3 (Stieltjes Transform)**.** The Stieltjes transform, or Green function, or Resolvent of a distribution ρ is

$$G(z) = \int dx \frac{\rho(x)}{z - x}$$

The usefulness of such transform is twofold. First we may readily recover the moments of the distribution

$$G(z) = \int dx \frac{\rho(x)}{z-x} = \frac{1}{z}\int dx \frac{\rho(x)}{1-\frac{x}{z}} = \frac{1}{z}\sum_{k=0}^{\infty}\int dx \rho(x)\left(\frac{x}{z}\right)^k = \sum_{k=0}^{\infty}\frac{m_k}{z^{k+1}}$$

and secondly, computing the transform at $z = y - i\epsilon$

$$G(y-i\epsilon) = \int dx \frac{\rho(x)}{y-i\epsilon-x} = \int dx \frac{\rho(x)(y-x)}{(y-x)^2+\epsilon^2} + i\int dx \frac{\rho(x)\epsilon}{(y-x)^2+\epsilon^2}$$

and taking into account the representation of the delta function

$$\delta(x) = \frac{1}{\pi}\lim_{\epsilon\to 0}\frac{\epsilon}{x^2+\epsilon^2}$$

we have that the knowledge of the resolvent leads to the distribution

$$\rho(y) = \frac{1}{\pi}\lim_{\epsilon\to 0}\operatorname{Im} G(y-i\epsilon)$$

2.3 Moments Method

2.3.1 Wigner Law by moments method

The moments method was the first and most brute-force way to prove Wigner law, but it is worth to sum up its proof as it is of interest for purposes that will be clear in the following chapters. The core of the theorem is to show that the moments of the Semicircular distribution and those of the ESD of a symmetric matrix whose random elements are (to a certain extent) independent coincide, thus also the distributions do.

Theorem 2.3.1 (Semicircle Law)**.** Let A_N be a symmetric N x N matrix whose non-diagonal elements are independent and identically distributed variables drawn from a Gaussian distribution with unitary variance, let $W_N = \frac{1}{\sqrt{N}}A_N$, then as $N \to \infty$, $\rho_W(x)$ converges to $\rho_{sc}(x)$, where

$$\rho_{sc}(x) = \begin{cases} \frac{1}{2\pi}\sqrt{4-x^2}, & \text{if } x \in [-2,2] \\ 0, & \text{otherwise} \end{cases}$$

Proof. We first calculate the moments of the semicircular distribution

$$m_{2k} = \frac{1}{2\pi}\int_{-2}^{2} x^{2k}\sqrt{4-x^2}dx = \frac{1}{\pi}\int_{0}^{2} x^{2k}\sqrt{4-x^2}dx.$$

and by substituting $x = 2\sqrt{y}$

$$m_{2k} = \frac{1}{\pi}\int_0^1 2^{2k+1}y^{k-\frac{1}{2}}\sqrt{1-y}dy = \frac{2^{2k+1}}{\pi}\frac{\Gamma(k+\frac{1}{2})\Gamma(\frac{3}{2})}{\Gamma(k+2)} = \frac{1}{k+1}\binom{2k}{k} = C_k$$

we get that the moments of the semicircular law are equivalent to the so-called Catalan numbers C_k. These numbers are of importance in the field of combinatorics [22]. Among many useful properties, we cite the recurrence relation they satisfy which will be also used later

$$C_0 = 1 \qquad C_n = \sum_{i=0}^{n} C_i C_{n-1}$$

The second step of the proof is to compute the averaged quantities

$$\langle m_k(W_N)\rangle = \frac{1}{N}\langle \mathrm{Tr}\left(\frac{A_N}{\sqrt{N}}\right)^k\rangle = \frac{1}{N^{1+\frac{k}{2}}}\sum_{i_1\cdots i_k=1}^{N}\langle A_{i_1i_2}A_{i_2i_3}\cdots A_{i_{k-1}i_k}A_{i_ki_1}\rangle$$

It is immediate to gather that, since the variables taken into account are independent with null mean, the summands in the series are non-zero only if a factor A_{jl} appears only twice in it. Hence k must be even, and so up to $\frac{k}{2}+1$ different indeces are allowed to appear.

The computation of the remaining terms is best performed using graph theory. Each summand in the RHS of the previous equation represents a path on the multigraph built upon the set of indices $\mathbf{i} = \{i_1, i_2, \cdots i_k\}$, where each of the i_ν is a node, and two nodes i, j are linked by l edges where l is the number of times that the term A_{ij} appears in the computation. Let $r = |\{i_1 \cdots i_k\}|$ be the size of the walk from i_1 to i_k.

Rearranging the summatory as

$$\sum_{i_1\cdots i_k=1}^{N}\langle A_{i_1i_2}A_{i_2i_3}\cdots A_{i_{k-1}i_k}A_{i_ki_1}\rangle = \sum_{r=1}^{N}\sum_{|\{i_1\cdots i_k\}|=r}\langle A_{i_1i_2}A_{i_2i_3}\cdots A_{i_{k-1}i_k}A_{i_ki_1}\rangle$$

An estimation of the order of such quantity may be done counting the possible realization of **i**. We have an upper bound for choosing the r possible different numbers in $\{1, \cdots N\}$ of N^r, and an upper bound for choosing among the k position of each of the r elements that is r^k, hence an upper bound for the term in the summatory is $r^k N^r$. Accordingly, the term $\sum_{\{i_1 \cdots i_k\}=r} \langle A_{i_1 i_2} A_{i_2 i_3} \cdots A_{i_{k-1} i_k} A_{i_k i_1} \rangle$ may be neglected, as it is of order N^r and we have a prefactor of $\frac{1}{N^{1+\frac{k}{2}}}$, when the length of **i** $r < \frac{k}{2} + 1$. Then we are left with the computation of

$$\sum_{|\{i_1 \cdots i_k\}|=1+\frac{k}{2}} \langle A_{i_1 i_2} A_{i_2 i_3} \cdots A_{i_{k-1} i_k} A_{i_k i_1} \rangle$$

For the sake of notation we call

$$I_k^{(N)} = \{\mathbf{i} \text{ such that } |\{i_1 \cdots i_k\}| = 1 + \frac{k}{2} \text{ and each edge appears only twice}\}$$

When k is even, as the variance is unitary by hypothesis, the previous term reduces to

$$\langle A_{i_1 i_2} A_{i_2 i_3} \cdots A_{i_{k-1} i_k} A_{i_k i_1} \rangle = 1$$

hence the sum to be computed is simply

$$\langle m_k(W_N) \rangle = \frac{1}{N^{1+\frac{k}{2}}} |I_k^{(N)}|$$

Then the graphs that have exactly 2 edges for each node, have k edges and $1 + \frac{k}{2}$ nodes are equivalent to the simple connected graphs that have $\frac{k}{2}$ edges and $1+\frac{k}{2}$ node, *i.e.* trees. The number of trees with h edges is precisely $\frac{1}{h+1}\binom{2h}{h}$, *viz.* the Catalan numbers. All the possible Eulerian paths on a tree are readily computed by assigning $1 + \frac{k}{2}$ indices to the $1, \cdots N$ vertices of the tree, *i.e.* $\frac{N!}{(N-(1+\frac{k}{2}))!} \approx N^{1+\frac{k}{2}}$ ways. Hence the averaged moments of W are

$$\lim_{N \to \infty} \langle m_k(W_N) \rangle = C_k$$

□

Wigner theorem is one of the classic results in RMT and in the theory of spectral properties of random matrices. However, I reckon the impact it has on network theory has not been stressed enough when taking into account

the actual meaning of this method of proof. Indeed when the moments are computed through the representative multigraph we should be aware that, if the matrix A is an adjacency matrix of a network, the counting of such moments is equivalent to counting the loops that the network possesses. Then, thanks to the hypothesis of independence of the adjacency matrix elements we may discard the odd moments[1], and thanks to the normalizing term in the front we get rid of the short cycles that are negligible in the limit of large N. So we are left with counting the tree cycles, which from a dynamical perspective are the most inefficient way of moving on the network. What is crucial in such proof is not only the independence of nodes of the underlying graph but the fact that the cycles of finite length are few with respect to trees of the same length when the thermodynamic limit is taken. However a remark can be made about the fact that the rate by which the number of non-tree cycles becomes negligible may be a good estimate of the true rate of convergence to the limiting case of Wigner law. Therefore, this quantity could be a viable candidate to be the order parameter of the phase transition that will be presented in the followings. Another remark to be kept in mind is that a network which does not show any odd cycle is completely bipartite, *i.e.* its vertices may be split into two sets where there are no internal edges. Also, bipartiteness of a network implies that its spectral density is an even function. These considerations may suggest that the number of odd cycles and the independence of the nodes should somehow be deeply related to the behaviour of the spectral density function, showing more than a feature that may lead to think to a phase transition phenomenon.

2.4 Stieltjes transform

We now consider a different approach to the whole problem of determining the limiting ESD of a certain ensemble of matrices. The Wigner law proven by the moments method is quite intuitive in its development, although being rather tedious in some calculation. However, besides giving hints about the

[1]Odd moments correspond to odd tours. Even in the most simple case, that is a tour of length 3 from a node i to j, k and back to i, it is immediate to see that the independence request, *viz.* the vanishing of odd cycles, is equivalent to imposing that structure of the network should display no immediate way of going back to the starting point. Hence, the network should be expected to have a diameter that grows according to a power law of the number of nodes. Such topic will be resumed in chapter 4.

reason by which the convergence to such a limit can be performed, the motives by which the limit is precisely Wigner law are not so explicit. A clearer way to answer this question is the derivation of Wigner law with the method of the resolvent, or Stieltjes transform, as introduced before. Such method is also handy for trying to set bounds to the convergence rate to the limit (see for instance [1]).

The framework in which we set such discussion is focused on the eigenvalues from the beginning. We may wonder how do the eigenvalue of a certain matrix distributes according to the pdf of the entries. What is needed is a transformation for the joint PDF of the entries into the joint PDF of the eigenvalues for a real symmetric matrix.

It is known that the jPDF of Gaussian variables[2] is the product of each singular PDF, thanks to the very definition of independence. For a $N \times N$ matrix H with iid Gaussian entries the probability density reads

$$p(H)dH = \prod_{i,j=1}^{N} \frac{1}{\sqrt{2\pi}} e^{-\frac{H_{ij}^2}{2}} dH$$

where the element of the matrix H means $dH = \prod_{ij=1}^{N} dH_{ij}$.

Now, the symmetry of the matrix H entails two facts: the former is that there are at most $\frac{N(N+1)}{2}$ distinct entries and the latter that such matrix is diagonalizable via an orthogonal transformation O of the kind $H = O\Lambda O^{-1}$. Hence the element dH may be written as

$$dH = \prod_{i \leq j} dH_{ij} = J_{(H \to \Lambda)} d\lambda_1 d\lambda_2 \cdots d\lambda_N d\alpha_1 d\alpha_2 \cdots d\alpha_{N(N-1)/2}$$

where the α_i are the angular-like parameters of the orthogonal transformation O, λ_i are the eigenvalues and $J_{H \to \Lambda}$ is the Jacobian of the transformation

$$J_{H \to \Lambda} = \det \begin{vmatrix} \frac{\partial H_{11}}{\partial \lambda_1} & \frac{\partial H_{11}}{\partial \lambda_2} & \cdots & \frac{\partial H_{11}}{\partial \lambda_N} & \frac{\partial H_{11}}{\partial \alpha_1} & \frac{\partial H_{11}}{\partial \alpha_2} & \cdots & \frac{\partial H_{11}}{\partial \alpha_{N(N-1)/2}} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{vmatrix}$$

[2] it is easy to generalize to the case in which the variance is non unitary.

We may make some crucial consideration about this Jacobian. As the transformation is linear in the λ_i, the first N columns of the Jacobian matrix will be independent in the eigenvalues, while the remaining ones will be linear in the λ_i as well. Hence in the determinant each factor will contain $N(N-1)/2$ terms, each linear in the λ_i so that the Jacobian is a polynomial of degree $N(N-1)/2$ in the eigenvalues. Now, for the transformation to be invertible the Jacobian must be non-zero and this implies it must be a function of the eigenvalues differences $|\lambda_i - \lambda_j|$. Hence

$$J_{H\to\Lambda} \propto \left[\prod_{i<j=1}^{N} |\lambda_i - \lambda_j|\right]^s$$

and the requirement for the polynomial to be at most of grade $N(N-1)/2$ in the eigenvalues sets the exponent $s = 1$[3].

Equating marginals and integrating over the α_i distributions in which we are not interested we get

$$p(\lambda_1 \cdots \lambda_N) = \frac{1}{Z_N} \prod_{i,j=1}^{N} e^{-\frac{\lambda_i^2}{2}} \prod_{i<j=1}^{N} |\lambda_i - \lambda_j|$$

where Z_N is an appropriate normalizing constant.

If we rescale the eigenvalues $\lambda_i \to \lambda_i/\sqrt{N} = x_i$, in order to have control over the distribution, we have the final result

$$p(\lambda_1 \cdots \lambda_N) = \frac{1}{Z_N} e^{N\left(\frac{1}{2}\sum_i x_i^2 - \frac{1}{2N}\sum_{i\neq j} \ln|x_i - x_j|\right)}$$

2.4.1 An heuristic derivation

Having the jPDF of the eigenvalues it is natural to wonder which is their most probable eigenvalue distribution for i.i.d. entries of a symmetric matrix. We

[3]The value of the exponent of this determinant identifies various classes of matrices whose jPDF is known. When $s = 1$ it is the case of Gaussian Orthogonal Ensembles, (GOE). The other two main ensembles are the Gaussian Unitary Ensemble (GUE) and the Gaussian Simplectic Ensemble (GSE). These terms refers to the requirement of the matrices to have real eigenvalues, *i.e.* symmetricity for real entries, unitarity for complex entries ($s = 2$) and self-duality for quaternionic entries ($s = 4$).

introduce a potential

$$U(x) = \frac{1}{2}\sum_i x_i^2 - \frac{1}{2N}\sum_{i\neq j} \ln|x_i - x_j| \tag{2.2}$$

and we impose an equilibrium condition $\frac{\partial U(x)}{\partial x_j} = 0$, getting

$$x_j = \frac{1}{2N}\sum_{i\neq j}\frac{1}{x_j - x_i}$$

then we multiply both sides by $\frac{1}{N(z-x_i)}$ and sum over j, yielding

$$\frac{1}{N}\sum_j \frac{x_j}{z-x_i} = \frac{1}{2N^2}\sum_j\sum_{i\neq j}\frac{1}{x_j - x_i}\frac{1}{z-x_i}$$

The LHS of the equation can be written in terms of discrete resolvents as

$$\frac{1}{N}\sum_j \frac{x_j - z + z}{z - x_j} = -1 + \frac{z}{N}\sum_j \frac{1}{z-x_j} = -1 + zG_N(z)$$

While the RHS term may be recast as

$$\frac{1}{N^2}\sum_j\sum_{i\neq j}\frac{1}{z-x_i}\left(\frac{1}{z-x_j} + \frac{1}{x_j - x_i}\right) = \frac{G_N^2(z)}{2} + \frac{G_N'(z)}{2N}$$

thus obtaining the relation

$$G_N^2(z) + 2zG_N(z) + \frac{G_N'(z)}{N} + 2 = 0$$

This is a differential relation for the discrete counterpart of the resolvent. When going into the continuum limit over large N we simply replace the sum with an averaging integral over the distribution, having the Stieltjes transform as its natural outcome

$$\langle G_N(z)\rangle = \langle \frac{1}{N}\sum_j \frac{1}{z-x_j}\rangle = \int dx \frac{\rho(x)}{z-x} = G(z)$$

If we perform such limit, we immediately see that the derivative term disappears, as it is of the same order of the others but divided by N. Hence we have to solve the equation

$$G(z)^2 - 2zG(z) + 2 = 0$$

which has the solutions

$$G(z) = z \pm \sqrt{z^2 - 2}$$

Then we may employ the inversion formula for the transform to find the distribution.

$$G(x - i\epsilon) = (x - i\epsilon) \pm \sqrt{(x^2 - \epsilon^2 - 2) - 2ix\epsilon}$$

and as we are interested only in the imaginary part of $G(z)$ we have[4]

$$\operatorname{Im} G(z) = -\epsilon \pm \frac{sgn(-2x\epsilon)}{\sqrt{2}} \sqrt{\sqrt{(x^2 - \epsilon^2 - 2)^2 + 4x^2\epsilon^2} - x^2 + \epsilon^2 + 2}$$

taking the limit of $\epsilon \to 0$ and dividing by π we recover the distribution

$$\rho(x) = \pm \frac{sgn(-x)}{\sqrt{2}\pi} \sqrt{|x^2 - 2| - x^2 + 2}$$

which is exactly the semicircular distribution when the right sign is chosen

$$\rho_{sc}(x) = \begin{cases} \frac{1}{\pi}\sqrt{2 - x^2}, & \text{if } x \in [-\sqrt{2}, +\sqrt{2}] \\ 0, & \text{otherwise} \end{cases}$$

This heuristic derivation for the semicircular law gives immediate hints to its physical meaning. Indeed, it is the distribution that maximizes the jPDF of the eigenvalues. Also, working with resolvents points out the physical meaning of the problem. Indeed, the reader may have noticed the analogies with an electrostatic problem, having to find the equilibrium solution for charges on a line with a logarithmic repulsive potential. To conclude this section we highlight that this is rescaled semicircular law with respect to the previous result. Substituting $x \to \sqrt{2}x$ gives immediately the "usual" Wigner law, for which we compute the canonical resolvent Stieltjes transform

$$G(z) = \frac{1}{2\pi} \int dx \frac{\sqrt{4 - x^2}}{z - x} = \frac{-z + \sqrt{z^2 - 4}}{2}$$

which will be the form henceforth used when referring to the resolvent of the semicircular distribution.

[4]Following from $\sqrt{x + iy} = \frac{1}{\sqrt{2}} \sqrt{\sqrt{x^2 + y^2} + x} + i \frac{sgn(y)}{\sqrt{2}} \sqrt{\sqrt{x^2 + y^2} - x}$

2.5 Free Central Limit Theorem

In the previous section we have seen that the "most probable" distribution of eigenvalues is the one described by the semicircular law. One may speculate if this is somehow related to a central limit theorem. The answer is yes. In this section I give the gist of this result and then introduce the framework in which it can be proven soundly, that is Free Probability Theory.

2.5.1 Free Central Limit Theorem

When performing the limit we cannot pass through the characteristic function of the distribution as in the classic case since the requirement for them to work is the commutativity of the variables when computing the exponent. We need to find a transformation of the distribution laws that takes into account the non commutativity properties of random matrices. As pointed out in [44], the right candidate is the functional inverse of the resolvent, also known as "Blue function", that is

$$B(G(z)) = z$$

This function however may display a singular behaviour for small z, so we subtract z^{-1} to regularize, defining the so-called R-transform of the resolvent of the matrix X

$$R_X(G) = B_X(z) - \frac{1}{z}$$

The R-transform is the fondamental tool to perform the CLT procedure as we see that it is additive

$$R_{X+Y}(G) = R_X(G) + R_Y(G)$$

and may be expanded in terms as

$$R_X(G) = \sum_{i=1}^{k} C_i(X) G^{i-1}$$

where the C_i coefficients are given by

$$C_i(X) = \tau(X^i) - \sum_{j=1}^{i-1} C_j(X) \sum_{a_i + a_j + \cdots a_h = i-j} \tau(X^{a_i}) \cdots \tau(X^{a_h})$$

where, following [16], we introduced the expected trace operator of the matrix $\tau(X) = \frac{1}{N}\langle \mathrm{Tr}\, X\rangle$.

Now, using the additivity of the R-transform we may compute it for a normalized sequence of random matrices $S_n = \frac{X_1+X_2+\cdots X_n}{n}$, that is

$$R_{S_n}(G) = nR_{X/\sqrt{n}}(G)$$

The coefficients in the expansion of the R-transform are given by

$$C_i\left(\frac{X}{\sqrt{n}}\right) = \tau\left(\frac{X^i}{n^{i/2}}\right) - \sum_{j=1}^{i-1} C_j\left(\frac{X}{\sqrt{n}}\right) \sum_{a_i+a_j+\cdots a_h=i-j} \tau\left(\frac{X^{a_i}}{n^{a_i/2}}\right)\cdots\tau\left(\frac{X^{a_h}}{n^{a_h/2}}\right)$$

and as $n \to \infty$ we see that only the terms with $k \leq 3$ do not vanish. To proceed further, we assume that the matrices X_i have null mean and unitary average. Then the computation of the coefficient reduces to

$$C_1\left(\frac{X}{\sqrt{n}}\right) = \tau\left(\frac{X}{n^{1/2}}\right) = 0$$

$$C_2\left(\frac{X}{\sqrt{n}}\right) = \tau\left(\frac{X^2}{n}\right) = \frac{1}{n}$$

$$C_3\left(\frac{X}{\sqrt{n}}\right) = \tau\left(\frac{X^3}{n^{3/2}}\right)$$

which means

$$R_{S_n}(G) = nR_{X/\sqrt{n}}(G) = G + \frac{\tau(X^3)}{\sqrt{n}}G^2 \tag{2.3}$$

which reduces to G in the limit of $n \to \infty$. Now, if we compute the R-transform of the resolvent of the semicircular law we have

$$R\left(G_\sigma(z)\right) = G_\sigma(z)$$

This proves that the distribution of sequence of random matrices X_i whose mean is null and variance is unitary converges in the limit of large n to the semicircular law, establishing a free counterpart of the Central limit theorem for non-commuting variables, in which the role of the Gaussian distribution is played by Wigner law for the spectral density.

2.5.2 Elements of Free Probability Theory

Free Probability Theory develops a framework in which non-commutative random variables - as the matrices we deal with are - may be studied. The core of FPT is to replace the notion of statistical independence with that of "Freeness"[5]. Why so? When trying to find the PDF of a sum of independent variables the standard technique is via the cumulant generating function, which under the assumption of Independence is additive. Then we use the characteristic function and eventually to the PDF of the sum with Fourier inverse transform. When dealing with matrices this for the spectral density, as matrix properties (orthogonality, symmetry, row/column sum) may impose weak correlations among entries that impede to treat them as independent. It is required the notion of Freeness, described as follows

Definition 2.5.1 (Freeness of matrices)**.** Let X_1, X_2 be two $N \times N$ matrices and τ be the expected trace operator $\tau = \frac{1}{N}\langle X \rangle$, then the two matrices are said free if

$$\tau((P_1(X_1) - \tau(P_1(X_1)))(P_1(X_2) - \tau(P_1(X_2)))(P_2(X_1) - \tau(P_2(X_1)))\times \\ (P_2(X_2) - \tau(P_2(X_2)))(P_3(X_1) - \tau(P_3(X_1)))\cdots) = 0$$

where $P_i(X)$ is a i^{th} grade polynomial of X.

Also, this definition generalises to that of asymptotic freedom as follows.

Definition 2.5.2 (Asymptotic freeness of matrices)**.** Let X_1, X_2 be two $N \times N$ matrices and τ be the expected trace operator $\tau = \frac{1}{N}\langle X \rangle$, then the two matrices are said asymptotically free if

$$\lim_{N\to\infty} \tau((P_1(X_1) - \tau(P_1(X_1)))(P_1(X_2) - \tau(P_1(X_2)))(P_2(X_1) - \tau(P_2(X_1)))\times \\ (P_2(X_2) - \tau(P_2(X_2)))(P_3(X_1) - \tau(P_3(X_1)))\cdots) = 0$$

To give the gist of these definitions let us compute via the freeness formula

$$\tau(X_1^2 X_2^2) = \tau(X_1^2)\tau(X_2^2)$$

[5] I am aware that English word formation rules would require the property of being free to be called "Freedom", however this is the standard nomenclature in the literature.

which is analogous to the classical case, but when the non commutativity takes place we have

$$\tau(X_1X_2X_1X_2) = \tau^2(X_1)\tau(X_2^2) + \tau(X_1^2)\tau^2(X_2) - \tau(X_1^2)\tau(X_2^2)$$

which has no counterpart in the classical case.

2.6 Replica Trick

We now introduce a new formalism in which to study the problem of determining the limiting spectral densities of a random matrix. This approach was developed in close analogy with statistical mechanics, thus being addressed to the problems of statistical mechanics. Besides, the main method of solution, *viz.* Replica Trick, is not rigorous nor mathematically sound. A great deal of debate has been spent about the validity of this trick[6] to solve many models in statistical mechanics. A review of the achievements of this trick and a good try to soundly formalize it may be found in [29], while for a critical approach one may refer to [39] and [45]. However, almost all the results that will be proven in this section are may be proven in some other way. The purpose of this section is therefore to provide a framework which may be related with ease to a statistical mechanics one, and to give physical interpretation hints about the results. Moreover, replica trick will be useful to compute some of the corrections we must add to Wigner law in certain cases, namely when the average of the entries distribution does not vanish and the finite size corrections.

2.6.1 Derivation of general formula

We take advantage from the representation of the Dirac δ function to rewrite the density function for the eigenvalues as

$$\rho(x) = \frac{1}{N}\sum_i \delta(z-\lambda_i) = \frac{1}{N\pi}\lim_{\epsilon\to 0^+}\operatorname{Im}\sum_i \frac{1}{(z-i\epsilon-\lambda_i)}$$

where the parameter ϵ is needed to regularize the function later on. As the determinant is the product of eigenvalues of the matrix A, we have

[6] Also, the very fact that it is still called a "trick" shows the recalcitrance to fully accept it.

$$det(Ix - A) = \prod_i (x - \lambda_i)$$

hence we get the general formula

$$\rho_\epsilon(x) = \frac{1}{N\pi} \operatorname{Im} \frac{\partial}{\partial x} \ln(\det(I(x - i\epsilon) - A))$$

Now the core of the replica trick is introduced. The idea is to make use of the identity of the logarithm[7]

$$\ln x = \lim_{n \to 0} \left(\frac{x^n - 1}{n} \right)$$

to create n copies (replicas) of the determinant to calculate, which is expressed in terms of a Fresnel integral

$$\frac{1}{\sqrt{det(Ix - A)}} = \left(\frac{e^{i\frac{\pi}{4}}}{\sqrt{\pi}} \right)^N \int \prod_{i=1}^{N} dw_i e^{-i \sum_{ij} w_i (\delta_{ij} x - A_{ij}) w_j} \tag{2.4}$$

so that we may rewrite the distribution as

$$\rho(x, A_{ij}) = \frac{-2}{N\pi} \operatorname{Im} \frac{\partial}{\partial x} \lim_{n \to 0} \frac{1}{n} \left(\left(\frac{e^{i\frac{\pi}{4}}}{\sqrt{\pi}} \right)^{Nn} \int \prod_{i=1}^{Nn} dw_i e^{-i \sum_{ij\alpha} w_i^\alpha (\delta_{ij} x - A_{ij}) w_j^\alpha} \right)$$

The integration being Nn-fold as i runs from 1 to N and the auxiliary index α from 1 to n, the limit of $\epsilon \to 0$ being implicit. The handiness of the replica trick is that when we need to calculate the ESD we just have to average the formula with the distribution of the entries $p(A_{ij})$, so that we have a general formula

$$\langle \rho(x) \rangle = \int dA_{ij} \rho(x, A_{ij}) \prod_{ij} p(A_{ij})$$

to evaluate, in principle, the eigenvalue distribution of any ensemble of matrices. Of course, some assumptions are taken into account about the fact that replicas are to be treated as independent variables, *i.e.* their coupling is negligible.

[7] This comes from the expansion $\lim_{n\to 0} \left(\frac{x^n - 1}{n}\right) = \lim_{n\to 0} \left(\frac{e^{n \ln x} - 1}{n}\right) = \lim_{n\to 0} \left(\frac{n \ln x + \frac{1}{2}(n \ln x)^2 + \cdots}{n}\right) = \ln x$.

2.6.2 ESD for complete graph

The first and most *naïve* application of the replica trick formula is the case in which all the entries are equal. At first glance this may look of little use in the context of RMT for network theory, however, the result which will be found will be a basic block of some of the further developments. Indeed, the case in which the matrix elements are all the same is (almost) the case of a wholly connected graph, also called a complete graph. Namely, each node will have the same degree, which is also the dimension of the network N, and the stochastic matrix Π will have only $\pi_{ij} = \frac{1}{N}$ entries for all possible i, j.

Introducing the auxiliary identity

$$e^{i\frac{1}{N}\left(\sum_i w_i^\alpha\right)^2} = \frac{e^{-i\frac{\pi}{4}}}{\sqrt{2\pi}} \int dq e^{\frac{iq^2}{2}} e^{-\sqrt{\frac{2}{N}} q \sum_i w_i^\alpha}$$

the integral in 2.6.1 may be recast as

$$I = \prod_\alpha \frac{e^{-\frac{i\pi}{4}}}{\sqrt{2\pi}} \int dq \prod dw_i^\alpha e^{-ix\sum_i (w_i^\alpha)^2 - \sqrt{\frac{2}{N}} q \sum_i w_i^\alpha + \frac{iq^2}{2}}$$

by completing the square and computing the integrals in w_i we have

$$I = \prod_\alpha \frac{e^{-\frac{i\pi}{4}}}{\sqrt{2\pi}} e^{-\frac{iN\pi}{4}} \left(\frac{\pi}{x}\right)^{\frac{N}{2}} \int dq e^{\frac{iq^2}{2} - \frac{iq^2}{2x}}$$

i.e.

$$I = \prod_\alpha e^{-\frac{iN\pi}{4}} \pi^{\frac{N}{2}} \frac{1}{x^{\frac{N-1}{2}}} \frac{1}{\sqrt{x-1}}$$

Inserting it back into the general formula 2.6.1 (there is no need of averaging over a constant distribution) we get

$$\begin{aligned} \rho(x) = \frac{-2}{N\pi} \operatorname{Im} \frac{\partial}{\partial x} \ln \frac{1}{x^{\frac{N-1}{2}}} \frac{1}{\sqrt{x-1}} = \\ = \frac{2}{N\pi} \operatorname{Im} \left(\frac{N-1}{2x} + \frac{1}{2(x-1)} \right) = \\ = \frac{N-1}{N} \delta(x) + \frac{1}{N} \delta(x-1) \end{aligned}$$

where we used again the representation of the delta function. This result is indeed quite intuitive when we think of it. The distribution is a peak centered in 0, *viz.* all the eigenvalues are null, except from a peak of order $\frac{1}{N}$ in 1 that represents the stationarity, its eigenvector being the stationary state of the random walk.

We may readily see that this is the correct result by taking into account the adjacency matrix for a complete graph, $A_N = 1_{N\times N} - I$, where $1_{N\times N}$ is the all-one matrix made by $1_{N\times N} = u \cdot u^T$ and u is the all-one vector. Then a direct computation shows

$$\begin{aligned}\det(1_{N\times N} - I - xI) &= \det\big(u\cdot u^T - (x+1)I\big) = \\ &= (-(x+1))^N \det\left(I - \frac{u\cdot u^T}{x+1}\right) = (-(x+1))^N\left(1 - \frac{N}{x+1}\right) = \\ &= (-1)^N(x+1)^{N-1}(x+1-N)\end{aligned}$$

Then, in the limit of large N by rescaling the eigenvalues by their degree $N-1 \simeq N$ we immediately glean that they all concentrate in 0 but one localized in 1[8].

A few remark may arise as this simple result is proven. The first is that in terms of network dynamics the relaxation time τ is maximal for this kind of graph, as all the eigenvalues have the maximal possible spectral gap. Another observation is that a sizeable peak in the eigenvalue distribution corresponds to a certain extent to the fact that the matrix entries have no variability. This latter fact will be insightful later on.

2.6.3 ESD for random graph

Replica trick may be as well used to prove Wigner law. We begin by postulating the same Gaussian distribution for the elements A_{ij}

$$p(A_{ij}) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{A_{ij}^2}{2\sigma^2}}$$

and rescale the variance $J^2 = N\sigma^2$. Inserting it in the averaging equation we previously obtained we have

[8]Such rescaling corresponds, *en passant*, to the Stochastic matrix of such network.

$$\langle\rho(x)\rangle = \int \rho(x, A_{ij}) \prod p(A_{ij}) dA_{ij} =$$
$$= -\frac{2}{N\pi} \int \mathrm{Im}\, \frac{\partial}{\partial x} \lim_{n\to 0} \frac{1}{n} \left[\left(\frac{e^{i\frac{\pi}{4}}}{\sqrt{\pi}} \right)^{Nn} \int \prod_{i=1}^{N} dw_i e^{-i\sum_{ij\alpha} w_i^\alpha (\delta_{ij} x - A_{ij}) w_j^\alpha} \right] \times$$
$$\times \prod \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{A_{ij}^2}{2\sigma^2}} dA_{ij}$$

Such computation is best performed in some steps. We start by solving the Gaussian integrations, yielding

$$\langle\rho(x)\rangle = -\frac{2}{N\pi}\, \mathrm{Im}\, \frac{\partial}{\partial x} \lim_{n\to 0} \frac{1}{n} \left[\left(\frac{e^{i\frac{\pi}{4}}}{\sqrt{\pi}} \right)^{Nn} \int \prod_{i=1}^{Nn} dw_i e^{t(x,w_i)} \right]$$

where

$$t(x, w_i) = -ix \sum_{i,\alpha} (w_i^\alpha)^2 - \frac{J^2}{N} \sum_{i,j} (\sum_\alpha w_i^\alpha w_j^\alpha)^2 + \frac{J^2}{N} \sum_i (\sum_\alpha (w_i^\alpha)^2)^2$$

Let us take into consideration the second term in the t expression.

$$\frac{J^2}{N} \sum_{i,j} (\sum_\alpha w_i^\alpha w_j^\alpha)^2 = \frac{J^2}{N} \sum_{i,j,\alpha,\beta} w_i^\alpha w_j^\alpha w_i^\beta w j^\beta =$$
$$= \frac{J^2}{N} \sum_\alpha (\sum_i (w_i^\alpha)^2)^2 + \frac{J^2}{N} \sum_{i,j,\alpha\neq\beta} w_i^\alpha w_j^\alpha w_i^\beta w_j^\beta$$

We make the following considerations about the terms in the RHS of such equation. The former is of order Nn, and must be retained, while the latter has null mean and its square is of order n. The remaining term is of order n^2, so we retain only the terms where $\alpha = \beta$, as we did before. In the limit of large N thus

$$\langle\rho(x)\rangle = -\frac{2}{N\pi} Im \frac{\partial}{\partial x} \lim_{n\to 0} \frac{1}{n} \left[\left(\frac{e^{i\frac{\pi}{4}}}{\sqrt{\pi}} \right)^{Nn} \int \prod_{i=1}^{Nn} dw_i e^{-ix\sum_i w_i^2 - \frac{J^2}{N}(\sum_i w_i^2)^2} \right]$$

Introducing the identity to parametrize the exponential in terms of an auxiliary variable s

$$e^{-\frac{J^2}{N}(\sum_i w_i^2)^2} = \left(\frac{N}{2\pi}\right)^{\frac{1}{2}} \frac{x}{\sqrt{2J^2}} \int ds e^{-\frac{x^2}{4J^2}Ns^2 - ixs\sum_i (w_i)^2}$$

we may recast the integral as

$$\begin{aligned}
&\int \prod_{i=1}^{Nn} dw_i e^{-ix\sum_i w_i^2 - \frac{J^2}{N}(\sum_i w_i^2)^2} = \\
&\quad = \left[\int ds \prod_i dw_i \left(\frac{N}{2\pi}\right)^{\frac{1}{2}} \frac{x}{\sqrt{2J^2}} e^{-ix(1+s)\sum_i (w_i)^2 - \frac{Nx^2 s^2}{4J^2}}\right]^n = \\
&\quad\quad = \left[\left(\frac{N}{2\pi}\right)^{\frac{1}{2}} \left(\frac{\pi^N}{2J^2}\right)^{\frac{1}{2}} x e^{-\frac{N}{2} lnx} \int ds e^{-Ng(s)}\right]^n \qquad (2.5)
\end{aligned}$$

where in the last step we computed Fresnel integrals over the w_is. We introduced the auxiliary function $g(s)$ as

$$g(s) = \frac{x^2 s^2}{4J^2} + \frac{\ln(i(1+s))}{2}$$

Then we take the integral to be very peaked in the limit on large N, so that we may evaluate it using the saddle point method. Such approximation is consistent as we see that having x a small imaginary part $x - i\epsilon$ let us locate the branching point of the integral in $s = -1$. Then the cut on the complex plane is a line from -1 to $-\infty$, lying just above the real axis, that defines the right integration domain. The equation for the saddle point is $g'(s) = 0$, which gives 4 solutions, 2 for each case of x.

$$\begin{cases} s_0^{\pm} = -\frac{1}{2} \pm \frac{i}{2}\left(\frac{4J^2}{x^2} - 1\right)^{\frac{1}{2}} & \text{per } |x| < 2J \\ s^{\pm} = -\frac{1}{2} \pm \frac{1}{2}\left(1 - \frac{4J^2}{x^2}\right)^{\frac{1}{2}} & \text{per } |x| > 2J \end{cases}$$

We start by taking into account the case in which $|x| < 2J$. Following [13] we choose the contour passing trough s_0^-, and the saddle point evaluation yields

$$\int ds e^{-Ng(s)} \simeq \frac{\sqrt{2\pi}}{\sqrt{|Ng''(s_0^-)|}} e^{-Ng(s_0^-)} e^{i\phi^-}$$

where ϕ^- is given by

$$\phi^- = -\frac{1}{2}\arctan\left(-\frac{1}{\sqrt{\frac{4J^2}{x^2}-1}}\right)$$

By substituting this result in 2.6.3 for the replica ensemble, it is easy to see that the semicircular law is recovered

$$\rho(z) = -\frac{2}{N\pi} Im \frac{\partial}{\partial x}\left(Ng(s_0^-) - \frac{N}{2}\ln x\right) = \ = \frac{1}{2\pi J^2}\sqrt{4J^2 - x^2}$$

When $|x| > 2J$ it is immediate to see that $g(s_0^-)$ is real, hence the density is identically zero.

Two observations may now arise, focusing about the validity of the result and its applicability to the purposes of the present work. For what concerns the effective domain of validity we discarded the terms that did not grow as Nn. The n-replica limit is fine with this, as long as the replica symmetry is not violated the limit is soundly performable. However the limit for N large is not so straightforward, and we may wonder what are the effect of a finite size matrix on the eigenvalue distribution. The second main remark is that if we are to investigate the spectra of graph, the assumption that the stochastic matrix elements are drawn from a zero-mean Gaussian distribution is not correct, as the entries will surely have a finite mean. This latter question is the one which will be addressed in the next subsection.

2.6.4 A first correction to Wigner Law

Let us now take into consideration the case where the distribution by which the matrix entries are drawn is a Gaussian with non-zero mean. Such problem is not unavailing, as for our aims it is precisely the case in which the ESD that will be found will be that of the actual adjacency matrix of the underlying graph. Indeed, we shall take into account a real symmetric matrix A_N whose elements are drawn from the PDF with mean A_0/N[9]

[9]We note that $A_0 = \langle A_{ij}\rangle$ is the connectivity of the network.

$$p(A_{ij}) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\left(A_{ij}-\frac{A_0}{N}\right)^2}{2\sigma^2}}$$

rescaling as before the variance $J^2 = N\sigma^2$. By computing the Gaussian integrations and discarding the negligible terms in Nn we have

$$\langle \rho(x) \rangle = -\frac{2}{N\pi} \operatorname{Im} \frac{\partial}{\partial x} \lim_{n\to 0} \frac{1}{n} \left[\left(\frac{e^{\frac{i\pi}{4}}}{\sqrt{\pi}} \right)^{Nn} \int \prod^{Nn} dw_i^\alpha e^{t(x,w_i)} \right]$$

where now the $t(x, w_i)$ function is

$$t(x, w_i) - ix \sum_{i\alpha} (w_i^\alpha)^2 - \frac{J^2}{N} \sum_\alpha \left(\sum_i (w_i^\alpha)^2 \right)^2 + \frac{iA_0}{N} \sum_\alpha \sum_i (w_i^\alpha)^2$$

Then we make use of the Hubbard-Stratonovich transform and parametrize the integral as before introducing the auxiliary fields s and q

$$\begin{aligned} I_1 &= \int \prod^{Nn} dw_i^\alpha e^{t(x,w_i)} = \\ &= \left[\frac{e^{\frac{-i\pi}{4}}}{2\pi} \frac{xN}{\sqrt{4A_0 J^2}} \right]^n \int ds dq \prod_i^{Nn} dw_i^\alpha e^{-ix(1+s)\sum_i (w_i^\alpha)^2 - q\sum_i w_i^\alpha - \frac{x^2 s^2 N}{4J^2} + \frac{iq^2 N}{4A_0}} \end{aligned}$$

then by completing the square integration over the w_i^α is possible

$$I_1 = \left[\frac{e^{\frac{-i\pi}{4}}}{2\pi} \frac{xN\pi^{\frac{N}{2}}}{\sqrt{4A_0 J^2}} e^{-\frac{N}{2}\ln x} \int ds dq e^{-Ng(s) - \frac{iNq^2}{4}\left(\frac{1}{x(1+s)} - \frac{1}{A_0}\right)} \right]^n$$

and by integrating over q

$$I_1 = \left[\frac{e^{\frac{-i\pi}{4}}}{2\pi} \frac{xN\pi^{\frac{N}{2}}}{J} e^{-\frac{N}{2}\ln x} \sqrt{\frac{\pi}{N}} \int ds \sqrt{1+s} \frac{e^{-Ng(s)}}{\sqrt{i(s-s_1)}} \right]^n$$

where

$$s_1 = -1 + \frac{A_0}{x}$$

Now, to analyse this integral we make some considerations. If we are far from the pole s_1 we may perform the same saddle point evaluation that previously gave the semicircular law. However, when near s_1 we may expand in Taylor series

$$g(s) = g(s_1) + \frac{x}{2J^2}(x_m - x)(s - s_1) + \frac{x^2}{4}\left(\frac{1}{J^2} - \frac{1}{A_0^2}\right)\frac{(s - s_1)}{2} + \cdots$$

where $x_m = A_0 + \frac{J^2}{A_0}$. We observe that when $x = x_m$ $g(s)$ has an extremal point in s_1. If $A_0 < J$ the point is a maximum, whilst being a minimum for $A_0 > J$.

In the former case we may discard the contribution, as it implies a minimum in the integrand. On the other side, when $A_0 > J$ the mean value of the matrix entries lies outside the semicircle bounds, so a correction is mandatory. We may evaluate the integral by introducing a dummy variable $ih^2 = s - s_1$ and taking into account the first two terms of the Taylor expansion, yielding a Gaussian integral that gives

$$I_1 = \left[\sqrt{\frac{2A_0\pi^N}{J^2}} e^{-\frac{N}{2}\ln x} e^{-Ng(s_1) - \frac{1}{2}\ln(x_m - x)}\right]^n$$

where

$$g(s_1) = -\frac{x^2}{4J^2}\left(\frac{A_0}{x} - 1\right)^2 + \frac{\ln x}{2} - \frac{\ln A_0}{2} - \frac{i\pi}{4}$$

By substituting in the espression for I_1 we have

$$I_1 = \left[\sqrt{\frac{2A_0\pi^N}{J^2}} e^{-\frac{i\pi}{4}} e^{-\frac{N}{2}\ln A_0} e^{-\frac{x^2 N}{4J^2}\left(\frac{A_0}{x} - 1\right)^2 - \frac{1}{2}\ln(x_m - x)}\right]^n$$

Making use of the formula for the averaged density, we have a contribution

$$\langle \rho(x)_1 \rangle = \frac{1}{N\pi} \operatorname{Im} \frac{\partial}{\partial x} \ln(x_m - x) = \frac{1}{N}\delta(x - x_m)$$

with no counterpart in the case of the variance J being larger than the average A_0. Hence, the general result for a real symmetric random matrix

with a positive mean value of the entries is

$$\langle \rho(x) \rangle = \begin{cases} \frac{1}{2\pi}\sqrt{4-x^2} + \frac{1}{N}\delta\left(x - \left(A_0 + \frac{J^2}{A_0}\right)\right), & \text{if } A_0 > J \\ \frac{1}{2\pi}\sqrt{4-x^2} & \text{if } A_0 < J \end{cases} \tag{2.6}$$

From this result it is immediate to see that, if the variance of the matrix distribution falls below the threshold set by the mean value A_0, the semicircular law gains a peaked contribution due to an extensive population of eigenvalues no more following the semicircular law and due to the fact that the variance of the elements distribution shrinks. The possible implications of such fact, *i.e.* the variance of the entries distribution having a threshold which reflects on spectral properties, will be discussed more extensively in a following chapter.

2.6.5 Finite size effects

Replica trick is also handy for answering the following question: what are the effect of finite N on the spectral distribution?

We rewrite the ESD for Gaussian distribution in the replica form

$$\langle \rho(x) \rangle = -\frac{2}{N\pi} \operatorname{Im} \frac{\partial}{\partial x} \lim_{n\to 0} \frac{1}{n} \left[\left(\frac{e^{i\frac{\pi}{4}}}{\sqrt{\pi}} \right)^{Nn} \times \right.$$
$$\left. \times \int \prod_{i,\alpha=1}^{Nn} dw_i^\alpha e^{-ix\sum_{i,\alpha}(w_i^\alpha)^2 - \frac{J^2}{N}\sum_{i,j}(\sum_\alpha w_i^\alpha w_j^\alpha)^2} \right.$$

and express the exponentiated term as

$$\sum_{i,j}(\sum_\alpha w_i^\alpha w_j^\alpha)^2 = \sum_\alpha \left(\sum_i (w_i^\alpha)^2 \right)^2 + \sum_{\alpha\neq\beta,i,j} w_i^\alpha w_j^\alpha w_i^\beta w_j^\beta$$

so that the previous formula yields

$$\langle\rho(x)\rangle = -\frac{2}{N\pi}\,\mathrm{Im}\,\frac{\partial}{\partial x}\lim_{n\to 0}\frac{1}{n}\Bigg[\left(\frac{e^{i\frac{\pi}{4}}}{\sqrt{\pi}}\right)^{Nn}\times$$
$$\times\int\prod_{i,\alpha=1}^{Nn}dw_i^\alpha e^{-ix\sum_{i,\alpha}(w_i^\alpha)^2-\frac{J^2}{N}\sum_\alpha\left(\sum_i(w_i^\alpha)^2\right)^2-\frac{J^2}{N}\sum_{\alpha\neq\beta,i,j}w_i^\alpha w_j^\alpha w_i^\beta w_j^\beta}\Bigg]$$

Now, if the term $\sum_{\alpha\neq\beta,i,j}w_i^\alpha w_j^\alpha w_i^\beta w_j^\beta$ is to be neglected, as it is the case for the replica symmetry when $N\to\infty$ and $n\to 0$, the semicircular law is readily recovered. However, we want to compute the first correction ($O(1/N)$) to that limit. In order to do so we perform a Hubbard Stratonovich transformation of the kind

$$e^{-\frac{J^2}{N}\left(\sum_i(w_i^\alpha)^2\right)^2}=\sqrt{\frac{N}{2\pi}}\frac{x}{\sqrt{2J^2}}\int ds^\alpha e^{-\frac{x^2}{4J^2}N(s^\alpha)^2-ixs^\alpha\sum_i(w_i^\alpha)^2}$$

so that the density may be expressed in terms of the auxiliary variable s^α (under the assumption of replica symmetry) mediated over the ensemble at the desired order

$$\langle\rho(x)\rangle = -\frac{2}{N\pi}\,\mathrm{Im}\,\frac{\partial}{\partial x}\lim_{n\to 0}\frac{1}{n}\Bigg[\left(\frac{e^{i\frac{\pi}{4}}}{\sqrt{\pi}}\right)^{Nn}\sqrt{\frac{N}{2\pi}}\frac{x}{\sqrt{2J^2}}\times$$
$$\times\int ds^\alpha e^{-\frac{Nx^2}{4J^2}}\sum_\alpha(s^\alpha)^2\int\prod_{i,\alpha=1}^{Nn}dw_i^\alpha e^{-ix\sum_{i,\alpha}(1+s^\alpha)(x_i^\alpha)^2-\frac{J^2}{N}\sum_{\alpha\neq\beta,i,j}w_i^\alpha w_j^\alpha w_i^\beta w_j^\beta}\Bigg]$$

and the integral over the ws may be written perturbatively as

$$I=\int\prod_{i,\alpha=1}^{Nn}dw_i^\alpha e^{-ix\sum_{i,\alpha}(1+s^\alpha)(x_i^\alpha)^2}\times$$
$$\times\left[1-\frac{J^2}{N}\sum_{\alpha\neq\beta,i,j}w_i^\alpha w_j^\alpha w_i^\beta w_j^\beta+\frac{1}{2}\frac{J^4}{N^2}\sum_{\alpha\neq\beta,\delta\neq\gamma,i,j}\cdots\right]$$

Then, expanding this integral in a sum for each contribution we may have a perturbative result for the desired order $I=\sum_k I^{(k)}$, $I^{(k)}$ being the

contribution of order $\left(-\frac{J^2}{N}\right)^k$. A diagrammatic approach may be followed henceforth, nonetheless the method by which the result is obtained will be discussed from another point of view in the following chapter, so we refer to the paper [10] for the rather tedious calculations and give only the result for the $O(1/N)$ correction to Wigner law, that is

$$\rho_1(x) = \frac{1}{N}\left[\frac{1}{2}\left(\delta(x+2J)+\delta(x-2J)\right) - \frac{1}{2\pi\sqrt{4J^2-x^2}}\right]$$

so that the final result is

$$\langle\rho(x)\rangle = \frac{1}{2\pi J^2}\sqrt{4J^2-x^2}\left[1+\frac{J^2}{N(x^2-4J^2)}\right] + \\ + \frac{1}{4N}\left(\delta(x+2J)+\delta(x-2J)\right) \quad (2.7)$$

Equation 2.6.5 shows that the correction is made up of two terms: a border effect one (the peaks) and a more interesting "line-graph" effect. The spectrum of the line graph indeed, as it will be shown in the following chapter, is of the form $\frac{1}{\pi\sqrt{4-x^2}}$. The effect on the semicircular distribution is that of fattening a bit the tails. From a dynamical point of view, this could lead to a change in the relaxation time. Being a consequence of the finiteness of the number of sites, it might be related to the fact that the structure of the network is not only tree-like (which give rise to the Semicircular law), and a contribution of closed path to the moments counting is not negligible at the order $O(N^{-1})$.

Chapter 3

Network Theory and Spectral Density

This chapter is dedicated to the exposure of some results in the Spectral Analysis of networks [38]. First, some basic nomenclature and definitions are given to clarify the relevance of the analysis performed on the spectra and their "physical" meaning. Then we briefly expose the main classes of networks which are commonly considered in the literature [4] [40] [15]. The core of this chapter is the section about the spectral properties, in which we make clear their relevance and expose some results about the emerging properties that should be related to dynamical one. Both exact and limiting results are shown, as well as non analytical results of simulations. Finally, we perform an analogy between the spectral theory of networks and Quantum Field Theories via the representation of the ESD which we discuss. This latter section introduces the renormalization issue which is addressed in the fourth chapter.

3.1 Basic definitions

We introduce the basic tools to perform and understand the analysis we are going to develop in the followings. All the results and the definitions are restricted to the case of undirected networks.

Definition 3.1.1 (degree)**.** The degree d_i of a node i is the number of links (edges) the node has.

Definition 3.1.2 (Size)**.** The size of a network N is the number of nodes (even disconnected) in the network.

Definition 3.1.3 (Adjacency matrix)**.** The adjacency matrix A, also known as connectivity matrix, is the matrix whose elements A_{ij} are null if there is no connection between node i and node j, while taking the value 1 if they are connected.

The following properties hold for the adjacency matrix

- $\operatorname{Tr} A = 0$;
- $\sum_j A_{ij} = d_i$;
- It is symmetric, hence it has real eigenvalues;
- The maximal eigenvalue is $\lambda \leq N - 1$[1].

From the adjacency matrix it is possible to define stochastic matrix that represents the random walk performed on the network. If we introduce the degree matrix Δ whose entries are defined as $\Delta_{ij} = \delta_{ij}/d_j$, we are able to define the stochastic matrix

$$\Pi = \Delta^{-1} A$$

that is the one which characterizes the random walk on the network.

Definition 3.1.4 (Walk)**.** A walk is a sequence of connected nodes, which may contain a single node multiple times.

A specific term is devised for walk with no repetitions.

Definition 3.1.5 (Path)**.** A path is sequence of connected nodes, in which each node can appear only once.

The number of walks of a network is encoded in the adjacency matrix, indeed

Theorem 3.1.1. The number of walks of k hops between node i and node j is $(A^k)_{ij}$.

[1] A tighter bound is the maximal degree, reached only in the case of a regular graph.

Proof. If $k = 1$ is obvious, as it is the single link (if present) between the two nodes. Let us assume it works for $k-1$, so that the number of walks between node i and node l is $(A^{k-1})_{il}$. Then the number of walks of length 1 between node l and node j is simply A_{lj}. Hence the total number of walks between node i and j is $\sum_{l=1}^{N}(A^{k-1})_{il}A_{lj} = (A^k)_{ij}$. □

Definition 3.1.6 (Tour in a network)**.** A tour is circular sequence of connected nodes, *i.e.* the starting node is also the ending node, which may contain a single vertex multiple times.

Again, tours with no vertex repetition have a specific name

Definition 3.1.7 (Cycle)**.** A cycle, or loop, is a tour in which the only repeated vertex is the starting and ending one.

Definition 3.1.8 (Diameter)**.** The diameter D of a network is the longest of the shortest path between two nodes in the network.

Mathematically speaking, it is the lowest k for which $(A^k)_{ij} \neq 0$ for every i, j nodes.

Definition 3.1.9 (Bipartite network)**.** A network is said bipartite if the set of nodes can be divided into two sets of nodes A and B such that there are no links among the nodes in set A and among those in set B.

3.2 Classes of Networks

Networks are complex objects. A plethora of metrics, parameters, and measures may be extracted from a graph, still not describing it uniquely. Therefore the need of categorization, as regrouping them into classes describing a shared average demeanour allows to be much more general when speaking about the properties of certain graphs. This categorisation is mainly done according to the generative models that build the network, hence their growth. However when talking about general properties these networks are displaying, such as "Randomness", "Structure", "Regularity", and so on we must be aware that we refer to their intuitive meaning, as there is no commonly agreed or precise definition for them.

3.2.1 Random networks

The model for the generation of a random network by Paul Erdös and Alfred Renyi is widely known [15]. It was the act of birth of Network science, for which still nowadays the most of the results is achieved. It is the most intuitive model for building a graph and its spectrum is known to a certain extent. However complex networks often show rather different behaviours and features, hence it is not the most suited to mimic real-world counterparts. This model prescribes two nodes i , j to be linked (or not) according to a given probability p (or $1-p$). Indeed, having N nodes, we expect to have $p\frac{N(N-1)}{2}$ links in the whole graph. The average link number $\bar{N}_l$ will be $\bar{N}_l = p\binom{N}{2}$, and its variance can be computed according to a binomial distribution $Var(N_l) = p(1-p)\binom{N}{2}$. Moreover when studying properties of random graphs, the limit of large N is performed, as it provides a nice expression for the probability of having a certain average degree:

$$p(d) = \binom{N-1}{d} p^d (1-p)^{N-1-d} \simeq \frac{\bar{d}^d e^{-\bar{d}}}{N_l!}$$

where $\bar{d} = \frac{2}{N}\binom{N}{2} = p(N-1)$ is the average degree. According to various values p, graphs may look rather different one to another. As a matter of facts, if $p < \frac{1}{N}$ the whole graph will not be connected, but various detached components will be present, increasing p to be of order $\sim \frac{1}{N}$ ($\bar{d} = 1$) a phase transition-like phenomenon occurs, and a maximal connected component of order $O(N)$ appears.

3.2.2 Scale-free networks

Scale-free (SF) networks are sometimes called also "fractals" or "self-similar", although such properties are not really equivalent. The idea behind this kind of networks is that when varying the scale at which the network is observed the structure should not change drastically. Their most prominent feature is the rising of hubs, that is nodes whose degree is significantly larger than the others. This leads to a power-law distribution of the kind

$$p(d) = Cd^{-\alpha}$$

in the large N limit. It is a widespread conjecture that this kind of network should imitate more "natural" network models, for which a degree

distribution of the power law kind was observed *e.g.* World-Wide-Web, aerial transportation, scientific paper citations, *etc.* The most widespread and simple generative method for SF networks is the one proposed by Albert Barabasi and Rèka Albert in [4] (BA model), in which the proposal for a preferential attachment type of growth rather than a random one appeared to explain the SF behaviour. The power law distribution, or at least the fat tailed appearance, is related to long-range dependence, that is correlations over distant sites.

3.2.3 Small-World networks

Scale-Free networks are a step towards modeling of real world networks, but they are not the final one. Indeed, when it comes to portray the behaviour of graphs generated by natural processes, some issues arise. The first is the robustness, as purely SF networks are not robust with respect to perturbation performed on hubs: secondly the question of the diameter that grows with a power law in the case of BA model, while it has been observed that the emergence of shortcuts in real world networks is a prominent feature. Then the need to explain such overall (although not exact) self similarity also displaying small scale properties when it comes to navigability. This is what lead to the "Small-World" (SW) class of network, which are defined mainly by the fact that their diameter grows logarithmically with respect to the size of the network. The most prominent model for their generation is the one proposed by Duncan Watts and Steven Strogatz [40], in which a regular graph undergoes a random "rewiring" process that, in the case of low rewiring probability, achieves the small world properties. This compromise between the regular structure offered by the regular network (in which all the degrees are equal) and a random, though small, perturbation, might be the key to explain the mixed behaviour of this ensemble.

3.3 Spectra

Before going deeper into the problem of graph spectra, it is to be remarked that the question whether a network is uniquely determined by its spectrum is not a lazy one. Indeed, there are co-spectral graphs that are different up to relabelling nodes when their size is $N \geq 5$[2] so the question should be

[2]the so-called "Saltire pair in [37]".

rephrased to investigate if classes of networks are uniquely determined by their spectrum, and how. In the words of Piet van Miegheim [38]: "I believe that we still do not understand networks sufficiently. For example, if the data (e.g., the adjacency matrix) of a large graph is given, and you are not allowed to visualize the network, it seems quite complex to tell, by computing graph metrics only, what the properties of the network are. [...] We as humans see a pile of numbers, but often miss the overall picture and understanding. I believe that the spectrum, that is for a sufficiently large graph a unique fingerprint [...], may reveal much more."

As a matter of facts, all the metrics of a network may be derived from its spectrum, which often also gives immediate information such as bipartiteness[3], local resemblance to known network patterns[4], dynamical properties of the walk that may be performed, *etc.*

Keeping these considerations in mind, we begin to expose some of the results that involve spectra of certain classes of networks, highlighting the physical reasons that underlie those patterns.

3.3.1 The spectral-structural relation

In order to build a sound bridge between some of the RMT results and structural properties of networks, we give the following theorem.

Theorem 3.3.1 (Equivalence of tours and moments)**.** Let $G(N)$ be a graph with N nodes, A_N its adjacency matrix with spectral distribution ρ_A. Then the number C_k of tours of length k in the graph is equivalent to the k-th moment of the distribution.

Proof. It is a straightforward generalization of what has been done in the proof of the Wigner law by the moment method. We know that the moments of the distribution ρ_A are

$$m_k^A = \int x^k \rho_A(x) dx = \mathrm{Tr}(A^k) = \sum_{i_1,\cdots i_k=1}^{N} A_{i_1 i_2} A_{i_2 i_3} \cdots A_{i_{k-1} i_k} A_{i_k i_1}$$

Now, being A the adjacency matrix of the network, each term A_{ij} is non-zero only if the two nodes i and j are linked, otherwise the term is unitary.

[3] When the spectrum is symmetric a graph is bipartite, if it is symmetric over a certain value it is "locally bipartite".

[4] See [38], section 7.1

We may represent the resulting graph as a multigraph (*i.e.* a graph where more than one edge between two nodes is allowed), where for each term A_{ij} an edge between node i and node j is added. The sum is therefore the number of walks that start from the node i_1, goes into the node i_2 and so on until it jumps back from the node i_k to the starting node. This is precisely the number of possible tours of length k of the network. □

A remark that may be necessary is that tours are directional, that is the tour that start from node i_1 and steps to i_2, all the way to i_k before jumping back to i_1 and the tour that proceeds backwards are counted as 2 distinct tours. Hence, the multigraph representing the moments is not only distinct from the underlying network as it has multiple edges between nodes, but these edges are also directional if we are to make the identification of moments and tours.

Also, for the first moments it is quite easy to count the number of tours of length k that a network possesses, indeed

$$\begin{aligned} C_2 &= 2e \\ C_3 &= 6t \\ C_4 &= 8s + 4c + 2e \\ &\dots \end{aligned}$$

where e is the number of edges of the network (the 2 is for the directionality issue), t is the number of triangles (which may be travelled in $3 \cdot 2$ possible ways), s is the number of squares (which may be travelled in $4 \cdot 2$ possible ways), c is the number of cherries (which may be travelled in $2 \cdot 2$ possible ways), and so on. Devising a method to generalise this kind of approach to real-world graphs is not an easy task. However, this line of thought may be useful to resolve perturbatively the spectra of a given network. A reference where this approach is more thoroughly treated is [9].

3.3.2 Symmetries

Before analysing some of the known spectra of networks, we expose some results about general spectral properties and their structural meaning. The first and most important is that a network has a symmetric spectrum if it is (almost) bipartite.

Theorem 3.3.2 (Bipartiteness of a network)**.** If the spectrum of a network is symmetric over $\lambda = 0$ then the network is bipartite.

Proof. For a block matrix, we have the formal relation

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & 0 \\ CA^{-1} & I \end{bmatrix} \times \begin{bmatrix} A & B \\ 0 & D - CA^{-1}B \end{bmatrix}$$

given that A is invertible. Then for determinants it is

$$\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(A) \det\big(D - CA^{-1}B\big)$$

Now we take the adjacency matrix A_N of a network G_N

$$A_N = \begin{bmatrix} O & K \\ K^T & H \end{bmatrix}$$

where O is a $m \times m$ zero matrix, K is a $m \times n$ matrix and H is $n \times n$. From the previous relation, the characteristic polynomial is

$$\det(A_N - \lambda I) = (-\lambda)^{m-n} \det\big(\lambda H - \lambda^2 I_{n\times n} - K^T K\big)$$

which is symmetric only if also H is a zero-matrix. Then, the matrix

$$A_(N) = \begin{bmatrix} O & K \\ K^T & O \end{bmatrix}$$

is that of a bipartite network. □

The vice versa may be easily proven, as it is done in [38], pages $131 - 134$. From this result, it is also immediate that a bipartite graph must show a symmetric ESD, hence the odd moments vanish.

We should remark however that bipartiteness is a feature of networks that have a thorough regularity in their structure. This is not the case for many of the cases will be analyzing.

3.3.3 Random networks: Wigner law

The first result about limiting ESD of networks is that of the random graphs built according to the Erdös Renyi model. It is known that the empirical spectral density of such network converges to the semicircular law [8]. The

reason for which this holds is clear in the very proof of the Wigner law for random matrices. Indeed, when building the "multigraph" to count the moments of the matrix A, we simply take A to be the adjacency matrix of the ER network and thanks to the equivalence of moments and tours we may readily recover the Wigner Law with rescaled variance J

$$\langle \rho(x) \rangle = \begin{cases} \frac{1}{2\pi J^2}\sqrt{4J^2 - x^2}, & \text{if } x \in [-2J, 2J] \\ 0, & \text{otherwise} \end{cases}$$

The results of simulations are in accordance with this result, as shown in figure 3.1. The corrections developed for this law in the previous chapter are now relevant as they allow us to see some of them with the eyes of the spectral analysis of networks. In particular, the line graph effect emerging in 2.6.5 will be computed in 3.3.5; while the correction for a positive mean value 2.6.4 may be now linked to variance shrinking and peak appearance, a feature common to scale-free networks as shown in 3.6 and 3.5.

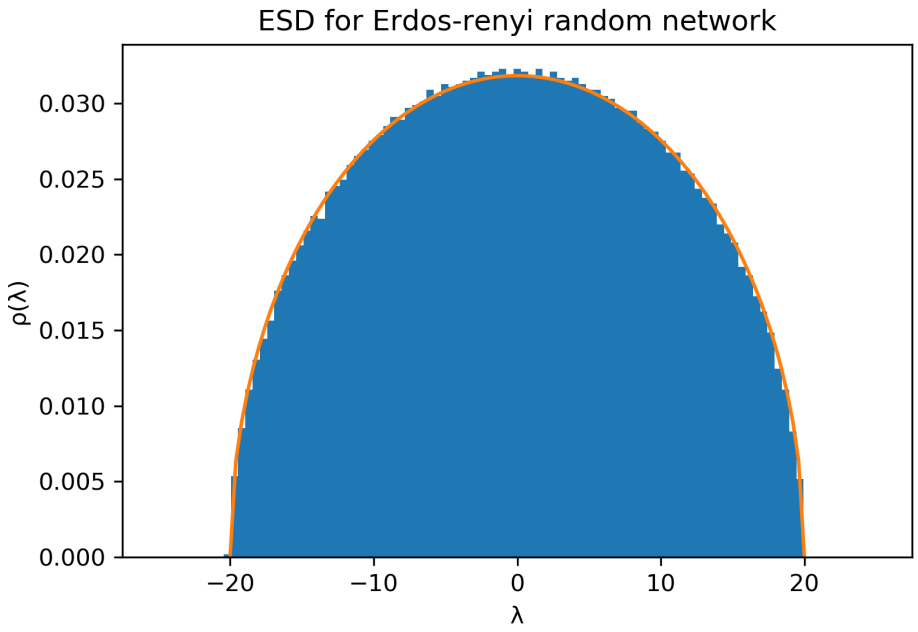


Figure 3.1: Semicircular law for a Erdös Renyi random graph for a network 10000 nodes and probability of linkage $p = 0.01$. The normalised histogram density is shown in blue while the Wigner law is plotted in orange.

3.3.4 Regular graphs: McKay law

The McKay law is the limiting spectral distribution for random regular graphs. It was developed independently from proper RMT and achieved in the context of random regular graphs. A regular graph is a graph whose nodes all have the same degree d. The theorem proven by McKay in [26] is the following

Theorem 3.3.3. Let G_N be a random regular graph with degree $d \geq 2$. Then, when $N \to \infty$ the ESD of A_N, that is the adjacency matrix of G_N, tends to

$$\langle \rho(x) \rangle = \begin{cases} \frac{d}{2\pi} \frac{\sqrt{4(d-1)-x^2}}{d^2-x^2}, & \text{if } x \in [-2\sqrt{d-1}, 2\sqrt{d-1}] \\ 0, & \text{otherwise} \end{cases}$$

Proof. A sketch of the proof is the following. Firstly, in the limit of $N \to \infty$ the structure is a pure d-tree for each node. Counting the moments is then counting the possible walks on such trees. As such, we may readily discard the odd moments, as closed paths on trees are always even. Then, we need to count the number of paths of length $2k$. Starting from the "root" node we have $d-1$ possible "down"-steps and, after one of those, 1 "up"-step. Let us suppose the walk returns to the root after $2(i+1)$ steps, with $0 \geq i \geq k-1$. For each down-step these walks are equivalent to the so-called "Dyck" paths[5], whose number is precisely the i^{th} Catalan number. Then the number of possible walks between the first step and the $2(i+1)$ one is $C_i(d-1)^i$. We may write then a recursion relation for the m_{2k} number of paths as

$$m_{2k} = d \sum_{i=0}^{k-1} C_i (d-1)^i m_{2(k-i-1)}$$

and by multiplying by the variable y^k and summing over it we have

$$M(y) = \sum_{k=0}^{\infty} m_{2k} y^k = 1 + d \sum_{k=1}^{\infty} \sum_{i=0}^{k-1} C_i (d-1)^i m_{2(k-i-1)} y^k = \\ = 1 + dy \left(\sum_{k=0}^{\infty} C_k (d-1)^k y^k \right) \left(\sum_{k=0}^{\infty} m_{2k} y^k \right)$$

[5]Dyck paths are those staircase walks that, in a square $n \times n$ go from the origin to the opposite vertex without going above the diagonal.

in which we have a kind-of generating function of the moments. Now, the following relation holds for the Catalan numbers

$$\sum_{k=0}^{\infty} C_k (d-1)^k y^k = \frac{1-\sqrt{1-4(d-1)y}}{2(d-1)y}$$

so that we may get to an equation for the generating function

$$M(y) = 1 + \frac{d(1-\sqrt{1-4(d-1)y}}{2(d-1)} M(y)$$

which means

$$M(y) = \left(1 - \frac{d(1-\sqrt{1-4(d-1)y})}{2(d-1)}\right)^{-1}$$

Now, we compute the moments for McKay law, which are

$$m_k = \int_{-2\sqrt{d-1}}^{2\sqrt{d-1}} dx \frac{d}{2\pi} \frac{\sqrt{4(d-1)-x^2}}{d^2-x^2} x^k$$

in which only even moments $(2k)$ are non-vanishing. Their generating function is

$$\begin{aligned}
\sum_{k=0}^{\infty} \int_{-2\sqrt{d-1}}^{2\sqrt{d-1}} dx \frac{d}{2\pi} \frac{\sqrt{4(d-1)-x^2}}{d^2-x^2} x^{2k} y^k = \\
= \int_{-2\sqrt{d-1}}^{2\sqrt{d-1}} dx \frac{d}{2\pi} \frac{\sqrt{4(d-1)-x^2}}{d^2-x^2} \frac{1}{1-x^2 y} = \\
= -\int_{-\pi}^{\pi} d\alpha \frac{d}{2\pi} \frac{4(d-1)\sqrt{1-\cos^2\alpha}}{d^2-4(d-1)\cos^2\alpha} \frac{\sin\alpha}{1-4(d-1)\cos^2\alpha y}
\end{aligned}$$

where the substitution $x = 2\sqrt{d-1}\cos\alpha$ was made. Passing to a complex variable

$$\frac{d}{2\pi i}\oint_{|z|=1}\frac{dz}{z}\frac{(d-1)\left(z-\frac{1}{z}\right)^2}{d^2-(d-1)\left(z+\frac{1}{z}\right)^2}\frac{1}{1-(d-1)\left(z+\frac{1}{z}\right)^2 y}=$$
$$=\frac{d}{2\pi i}\oint_{|z|=1}\frac{dz}{z}\frac{(d-1)(z^4-2z^2+1)}{dz^4-(d-1)(dyz^2-z^2+y)(z^4+2z^2+1)}=$$
$$=\frac{1}{1-\frac{d(1-\sqrt{1-4(d-1)y})}{2(d-1)}}=M(y)$$

solving through residues. The moment generating functions coincide, so do the moments and we may conclude that the distributions also do. □

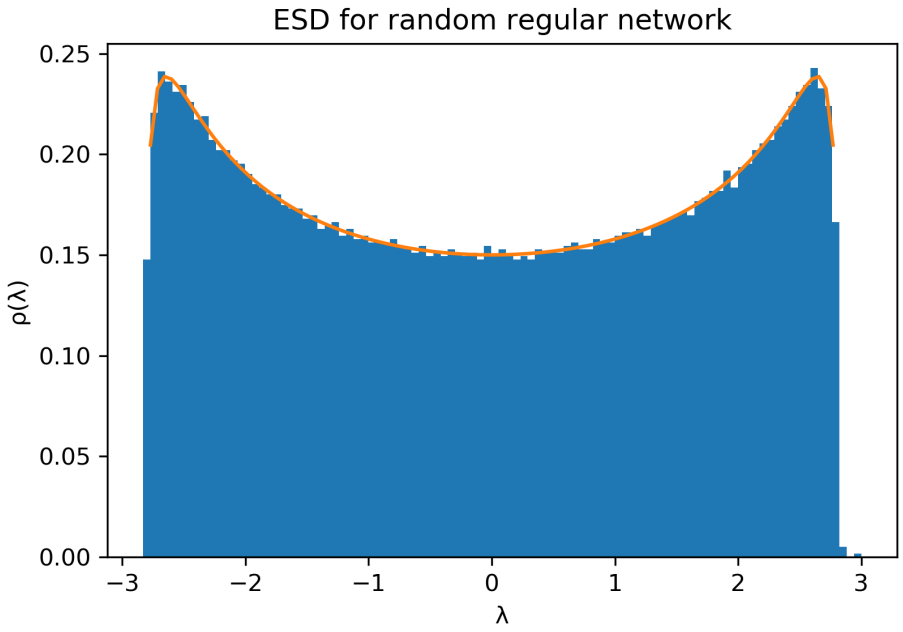
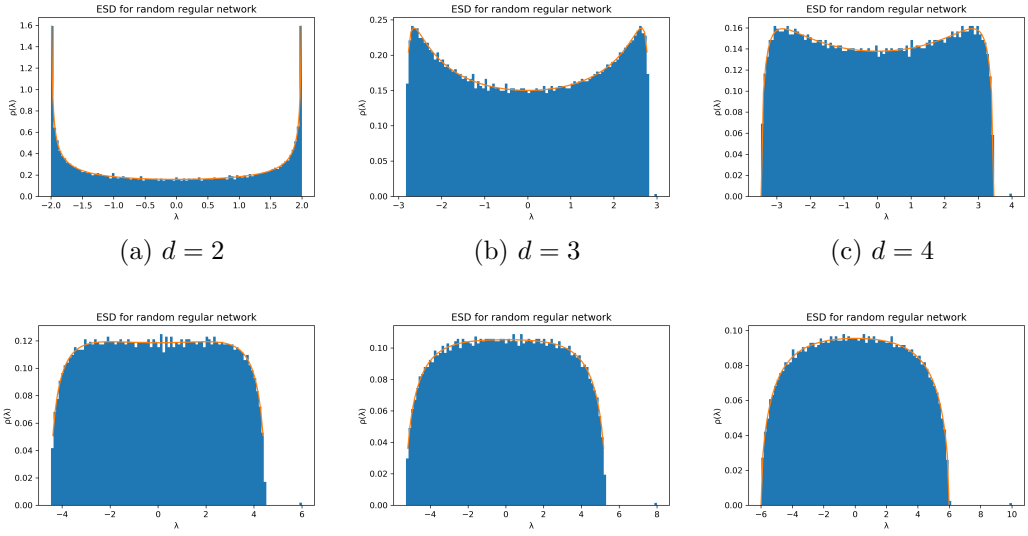


Figure 3.2: McKay law for a regular graph of 10000 nodes with degree $d = 3$. The normalised histogram distribution is shown in blue, while McKay law is plotted in orange.

This result may be interpreted as a constrained thermodynamical limit in the case of random regular graphs, where the unconstrained limit is the

semicircular law. Indeed, if the independence of matrix entries was seminal in the proof of that general result, in this case the entries, *i.e.* the nodes, are strongly correlated by the regularity condition imposition. As such this law behaves in a mixed way: the ESD is continuous as in the semicircular case, but a repulsion effect for the eigenvalues appear[6], "fattening" the tails of the distribution. Nonetheless, a startling result is that when the connectivity grows, the limiting spectral distribution is the semicircular law [36], showing that Wigner distribution has quite universal features. The transition is shown in figure 3.3, by which we see that already in the case of $d = 8$ the semicircular law holds quite well.



(a) $d = 2$ (b) $d = 3$ (c) $d = 4$

(d) $d = 6$ (e) $d = 8$ (f) $d = 10$

Figure 3.3: 6 realisations of d-regular networks with 5000 nodes. The degree is specified in the subcaptions. The normalised histogram distributions are shown in blue, while the respective McKay laws are plotted in orange.

[6]The so-called repulsion effect of eigenvalues may be understood from the electro-like potential 2.4.1. *Ça va sans dire*, in this case the potential is not logarithmic and should be determined.

3.3.5 Some exact spectra

In rare occasions exact spectra are known. In these cases, it is easy to reconstruct the ESD. Let us suppose to have the spectrum λ_k. Then, expressing the ESD as in 2.2.1, we make use of the delta representation

$$\delta(x) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} dz e^{zx}$$

so that the ESD may be recast as

$$\rho(x) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} dz e^{zx} h_\lambda(z)$$

where the function $h_\lambda(z)$ is

$$h_\lambda(z) = \frac{1}{N}\sum_{k=1}^{N} e^{-z\lambda_k} \tag{3.1}$$

Then, it is sufficient to insert the exact spectrum in this expression, which plays the role of a generating function for the ESD.

The line graph

The first fully known spectrum is one of the simplest network possible, that is the line. Each node is linked to other two except two nodes, the start one and the end one. The adjacency matrix of such a network is of the form[7]

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & & \\ 1 & 0 & 1 & 0 & \cdots & \\ 0 & 1 & 0 & 1 & \cdots & \\ & \cdots & \cdots & \cdots & & \\ & & \cdots & 1 & 0 & 1 \\ & & \cdots & 0 & 1 & 0 \end{pmatrix}$$

for which we know all the eigenvalues, as it is a special case for a Toepliz tridiagonal matrix [18] whose spectrum specializes to

$$\lambda_k = -2cos\left(\frac{\pi k}{N+1}\right)$$

[7]With a suitable labelling of indices, an operation that is always possible.

Then the ESD is simply given through the $h_\lambda(z)$ function

$$h_\lambda(z) = \frac{1}{N}\sum_{k=1}^{N} e^{2zcos\frac{\pi k}{N+1}}$$

whose limit when $N \to \infty$ is taken making k continuous so that

$$h_\lambda(z) = \frac{1}{N}\int_0^N dk e^{2zcos\frac{\pi k}{N+1}} = \frac{N+1}{N\pi}\int_0^{\frac{N\pi}{N+1}} dt e^{2z\cos t} \to \frac{1}{\pi}\int_0^{\pi} dt e^{2z\cos t} = I_0(2z)$$

$I_0(2z)$ is the first modified Bessel function of the first kind. If we put it in the ESD formula we have

$$\rho(x) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} dz e^{zx} I_0(2z) = \frac{1}{\pi}\frac{1}{\sqrt{4-x^2}} \tag{3.2}$$

A worthwhile remark is that this term is of the same kind of the one which appears in the finite size correction to the semicircular law. Then, the natural interpretation of that result may be that of the fact that for finite size there are paths which are simply lines, *i.e.* path that starts and do not come back "tree-like", simply ending the possible hops performable $(N-1)$. Note also that McKay law reduces to this result when the degree is 2.

Regular graph

Regular graphs have a known spectrum too. It can be found as follows. Being the degree d constant for every node, the graph may be visualized on a circle, which is invariant with respect to rotations. Travelling on this circle we have that each node has precisely k previous and k subsequent neighbours, where of course $2k = d$. Then we introduce the rotating operator P^n, that rotates the network by n vertices (anticlockwise). Being a symmetry operation, P^n commutes with the adjacency matrix, so they may be used to create a basis of common eigenvectors, that are $\vec{e_i}$ with

$$(e_i)_j = e^{2\pi i\frac{jl}{N}}$$

where $l = 0, 2, ...N-1$. The eigenvalues of P^n will hence be

$$s_l^{(n)} = e^{2\pi i\frac{nl}{N}}$$

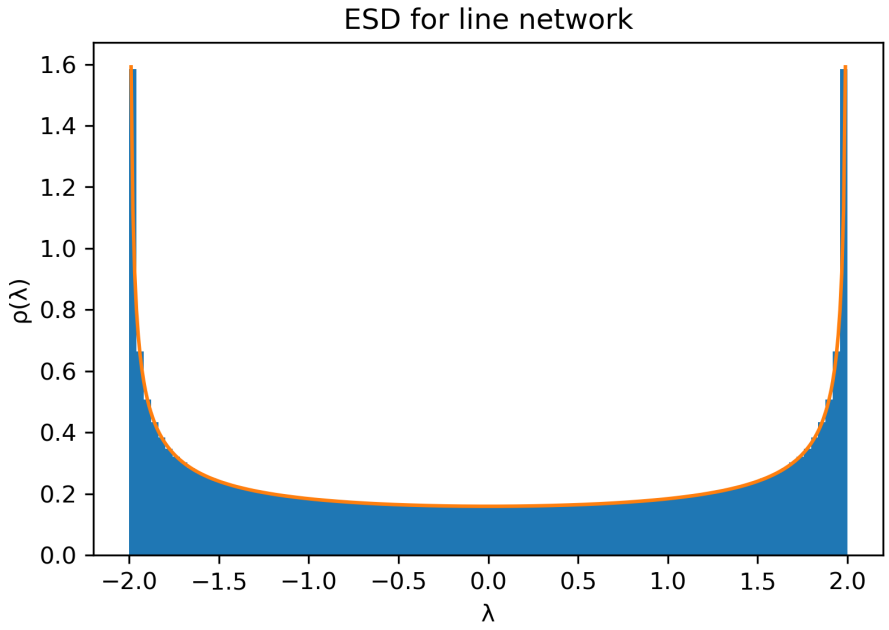


Figure 3.4: Line graph spectrum for a network of 10000 nodes. In blue the normalised histogram distribution, while in orange the analytic ESD.

which are then summed pairwise to find the eigenvalues of the adjacency matrix

$$\lambda_l = 2\sum_{j=1}^{k} \cos\frac{2\pi jl}{N} = \frac{\sin[(2k+1)l\pi/N]}{\sin(l\pi/N)} - 1$$

which in the continuum limit for $N \to \infty$ reads

$$\lambda(x) = \frac{[\sin(2k+1)x]}{\sin x} - 1 = U_{2k}(\cos x) - 1$$

where $U_{2k}(\cos x)$ are the Chebyshev polynomials of the second kind. Then the h function is

$$h_\lambda(z) = \frac{1}{\pi}\int_0^\pi dx e^{-2z\sum_{j=1}^{k}\cos(2jx)}$$

which for $k = 1$ is equivalent (again) to the line graph[8].

It is quite hard to find an exact form to express the ESD of regular graphs, beside from the cycle one. However, when $k \geq 2$ we may find a recursion law for the h_λ functions, namely

$$h_\lambda^{(k)}(z) = h_\lambda^{(k-1)}(z) \int_0^\pi dx e^{-2z\cos(kx)} = \\ = h_\lambda^{(k-1)}(z) \left[I_0(2z) + \sum_{j=1}^{\infty} I_j(2z)\cos(2kjx) \right]$$

from the generating function of the I_j. Knowing an exact form for the ESD for a regular graph would be of great use. Indeed of the three great classes of networks, that is Random, Scale Free and Small World, we have insights of the general ESDs of the former two, but not a real hint about the ESD of a Small World network. However the most widely used generative process for a SW network is that due to Watts and Strogatz [40], and it starts from considering a regular graph whose links are rewired according to a probability parameter. A regular graph ESD then should be close to the one that a Small world network possesses when the rewiring is small. We will deal with this more extensively in the last chapter.

3.3.6 Scale-free spectra

A conjecture about the SF type of networks is that they should display a scale-freeness in the spectrum as a consequence of the scale invariance of the structure itself. Indeed, the BA model gives a peaked spectrum, in which each peak seems to be somehow related to the scale at which the dynamics described by the eigenvalue happens. It is shown in figure 3.5. In the followings we try to give a hint of a proof by a generative algorithm for a scale-free network.

The cone of a network

Let G_N be a network of N nodes, with adjacency matrix A_N. The operation of "cone" of G_N is the addition of one node which is linked to every other

[8]The line graph is actually disconnected at the extrema, whilst the regular graph is always closed for $k \geq 1$ so they are not the same graph unless in the large N limit.

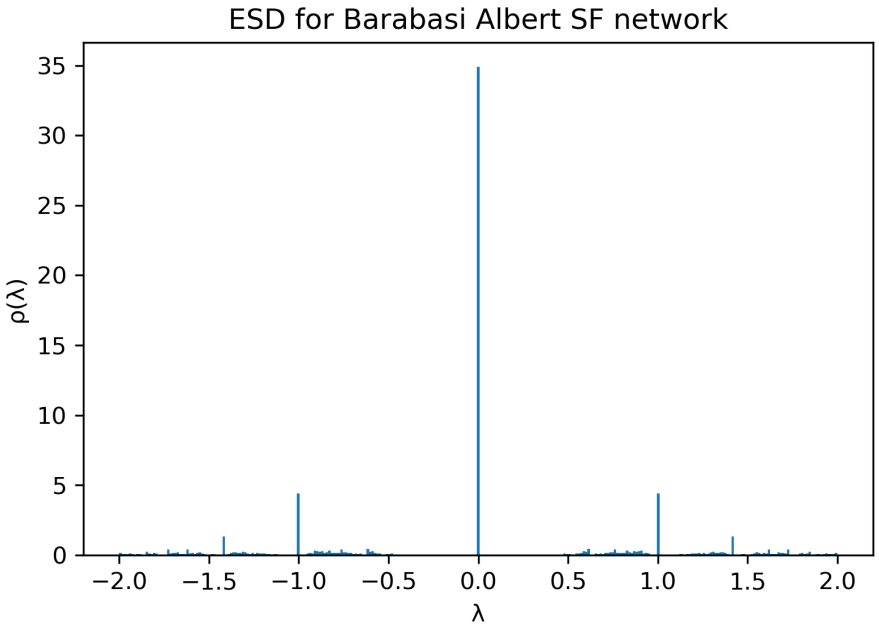


Figure 3.5: The spectrum of a Barabasi-Albert SF network [4] of 10000 nodes with parameter $k = 1$.

node. The resulting adjacency matrix is A_{N+1} and is built as

$$A_{N+1} = \begin{bmatrix} A_N & u \\ u^T & 0 \end{bmatrix}$$

from the relation about the determinant of a block matrix

$$\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(A) \det\left(D - CA^{-1}B\right)$$

we have that the spectrum of the cone may be expressed in terms of the spectrum of the graph G_N with the upcoming of a new eigenvalue. Applying it to this case, where $B = u$ and $C = u^T$, and $D = \lambda$, and $A = A_N - \lambda I$

$$\begin{aligned}\det(A_{N+1}-\lambda I) &= \det(A_N-\lambda I)\det\big(-\lambda - u^T(A_N-\lambda I)^{-1}u\big) = \\ &= -(\lambda + u^T(A_N-\lambda I)^{-1}u)\det(A_N-\lambda I) = \\ &= -\left[\lambda - \frac{1}{\lambda}\sum_{k=0}^{\infty}\frac{u^TA^ku}{\lambda^k}\right]\det(A_N-\lambda I)\end{aligned}$$

where the relation for the resolvent of a matrix was used

$$(A-\lambda I)^{-1} = -\frac{1}{\lambda}(I-\lambda^{-1}A)^{-1} = -\frac{1}{\lambda}\sum_{k=0}^{\infty}\frac{A^k}{\lambda^k}$$

Starring the network

Now, perform a slightly different operation: take a network G_N with adjacency matrix A_N and make l copies of the network, resulting in the matrix A_{N^l}. Then perform a cone on that ensemble of disconnected copies. This operation is called the "star" of the network G_N. Applying the previous formula for the characteristic polynomial yields

$$\det(A_{N^l+1}-\lambda I) = -\left[\lambda - \frac{1}{\lambda}\sum_{k=0}^{\infty}\frac{u^TA_{N^l}^ku}{\lambda^k}\right]\det(A_N-\lambda I)^l$$

where the relation $\det(A_{N^l}-\lambda I) = \det(A_N-\lambda I)^l$ comes readily from the fact that A_{N^l} is a block matrix.

The scheme

The core of the idea is to make stars of stars to implement a kind of renormalization transformation. Let G_N be the "seed" network, with adjacency matrix $A_N = A_0$. Let l_1 be the number of copies that are made of G_N. Let $A_{N^{l_1}} = A_1$ be the adjacency matrix of the yet-to-be-starred network and $A_{N^{l_1}+1} = A_{1^*}$ be the adjacency matrix for the starred one, so that it is immediate from the previous formula that

$$\det(A_{1^*}-\lambda I) = -\left[\lambda - \frac{1}{\lambda}\sum_{k=0}^{\infty}\frac{u^TA_1^ku}{\lambda^k}\right]\det(A_0-\lambda I)^{l_1}$$

Then make l_2 copies of the starred network G_{1^*}. The resulting adjacency matrix A_2 will be a block matrix formed blocks of A_{1^*}. Then cone the copies of G_{1^*}. By the previous relation this starring gives

$$\det(A_{2^*} - \lambda I) = -\left[\lambda - \frac{1}{\lambda}\sum_{k=0}^{\infty}\frac{u^T A_2^k u}{\lambda^k}\right]\det(A_{1^*} - \lambda I)^{l_2} =$$

$$= (-1)^{l_2+1}\left[\lambda - \frac{1}{\lambda}\sum_{k=0}^{\infty}\frac{u^T A_2^k u}{\lambda^k}\right]\cdot\left[\lambda - \frac{1}{\lambda}\sum_{k=0}^{\infty}\frac{u^T A_1^k u}{\lambda^k}\right]^{l_2}\det(A_0 - \lambda I)^{l_1 l_2}$$

For the third starring step, that is making l_3 disconnected copies of G_{2^*} and connecting them, the same relation yields

$$\det(A_{3^*} - \lambda I) = -\left[\lambda - \frac{1}{\lambda}\sum_{k=0}^{\infty}\frac{u^T A_3^k u}{\lambda^k}\right]\cdot\det(A_{2^*} - \lambda I)^{l_3} =$$

$$= (-1)^{(l_2+1)l_3+1}\left[\lambda - \frac{1}{\lambda}\sum_{k=0}^{\infty}\frac{u^T A_3^k u}{\lambda^k}\right]\cdot\left[\lambda - \frac{1}{\lambda}\sum_{k=0}^{\infty}\frac{u^T A_2^k u}{\lambda^k}\right]^{l_3}\cdot$$

$$\cdot\left[\lambda - \frac{1}{\lambda}\sum_{k=0}^{\infty}\frac{u^T A_1^k u}{\lambda^k}\right]^{l_2 l_3}\det(A_0 - \lambda I)^{l_1 l_2 l_3}$$

The scheme is now clear, for a sequence of h starring operations, each involving a number l_i of copies at each step i, the spectrum will be given by the characteristic polynomial

$$\det(A_{h^*} - \lambda I) = (-1)^{\left(\left((l_2+1)^{l_3}+1\right)^{\cdots}\cdots+1\right)}\cdot\det(A_0 - \lambda I)^{\prod_{i=1}^{h} l_i}\cdot$$

$$\cdot\left[\lambda - \frac{1}{\lambda}\sum_{k=0}^{\infty}\frac{u^T A_h^k u}{\lambda^k}\right]\cdot\prod_{i=1}^{h-1}\left[\lambda - \frac{1}{\lambda}\sum_{k=0}^{\infty}\frac{u^T A_i^k u}{\lambda^k}\right]^{\prod_{j=i+1}^{h} l_j}$$

A clarification should be made about the vectors u. They are the unity vectors of the dimension required by the scale l_i, that is the unity vector of the dimension required by the step i, so we indicate also them by u_i. This may look a pedantry but it may be useful later on.

Comments on the result

If the vector u_i is in fact the unity vector the quadratic form in u_i reduces to

$$u_i^T A_i^k u_i = W_k^{(i)}$$

that is the number of walk of length k in the network described by the matrix A_i. Then we may discard the seed graph (collapse the copies of the seed in a single node) to see what happens to the spectrum when this purely fractal structure is created. Ignoring the (-1) prefactor, the spectrum is given by the solution of the equation

$$\prod_{i=1}^{h} \left[\lambda - \frac{1}{\lambda} \sum_{k=0}^{\infty} \frac{u_i^T A_i^k u_i}{\lambda^k} \right]^{\prod_{j=i+1}^{h} l_j} = 0$$

so that for a process involving always the same number of copies l at each step, the equation reduces to

$$\prod_{i=1}^{h} \left[\lambda - \frac{1}{\lambda} \sum_{k=0}^{\infty} \frac{W_k^{(i)}}{\lambda^k} \right]^{l(h-i)} = 0$$

so that for each scale we have with multiplicity $l(h-i)$[9] the solutions of

$$\lambda^2 = \sum_{k=0}^{\infty} \frac{W_k^{(i)}}{\lambda^k}$$

Solving this equation is not easy. It should be solvable however for special classes of network. Indeed, the number of walks $W_k^{(i)}$ of the network G_i may be expressed, in terms of of the angle α_j between eigenvector v_j and the unity vector u as

$$W_k^{(i)} = N_{G_i} \sum_{j=1}^{N_{G_i}} \lambda_j^k \cos \alpha_j$$

so that the equation for the eigenvalues becomes

$$\lambda^2 = N_{G_i} \sum_{k=0}^{\infty} \frac{\sum_{j=1}^{N_{G_i}} \lambda_j^k \cos \alpha_j}{\lambda^k}$$

[9] Only 1 for the h last step.

For instance, in the case of a first step of the algorithm applied to single nodes we have

$$W_k^{(1)} = \begin{cases} 2l_1^k, & \text{if } k \text{ odd} \\ l_1^{k/2}(l_1+1), & \text{if } k \text{ even} \end{cases}$$

Finally, when the vector u_i is not all-ones and drawn from a distribution, the peaks centered in the eigenvalues that solve the self-similar equations are smoothed. Instead of having a λ_i characteristic of the i^{th} scale with multiplicity l_i we will have a distribution over $\langle\lambda_i\rangle$, whose weight in the total distribution is linked with the exponent l_i.

Properties of the generated network

A simulation was run for the network generated by this algorithm, whose spectrum is shown in figure 3.6.

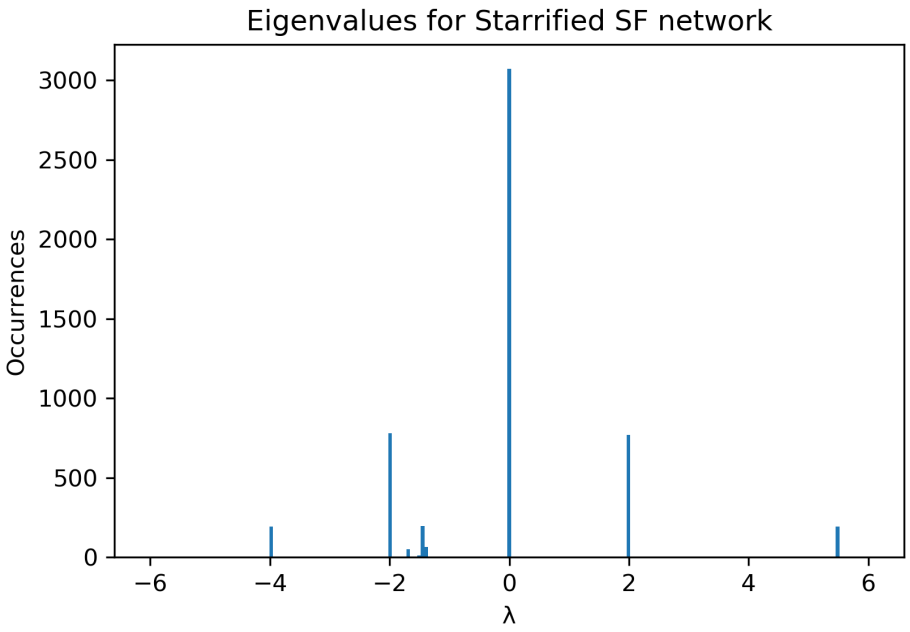
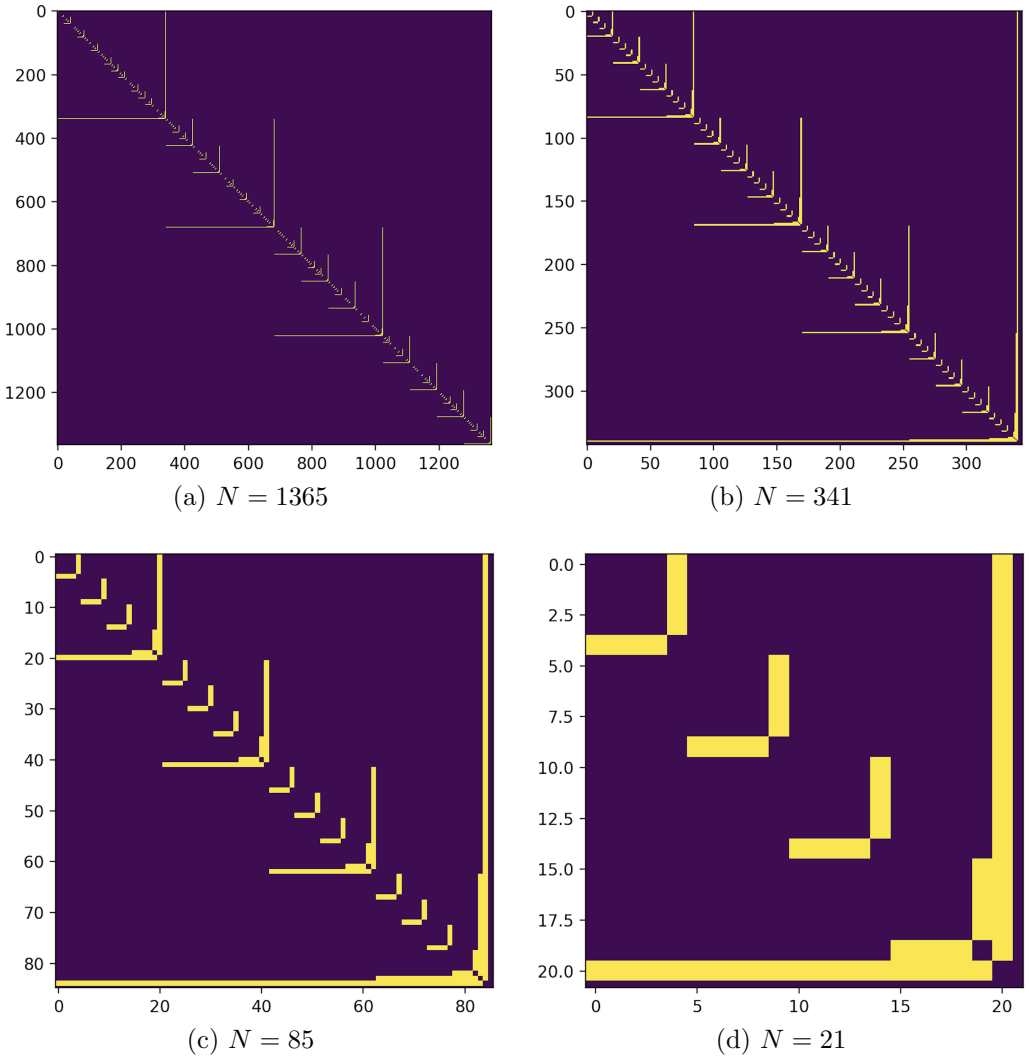


Figure 3.6: A network generated by the starrification algorithm of 5460 nodes, which corresponds to a six iterations procedure with parameter $l_i = l = 4$.

We see the self replicating peaks behaviour to appear. The fractal structure in the adjacency matrix for this network is evident from figure 3.7.

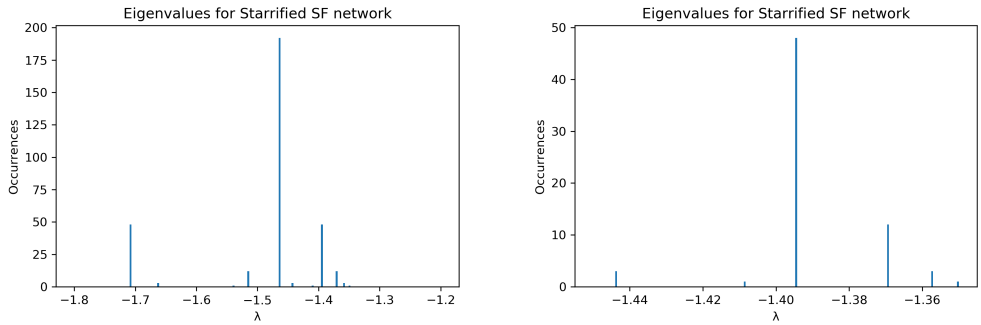


(a) $N = 1365$ (b) $N = 341$

(c) $N = 85$ (d) $N = 21$

Figure 3.7: Adjacency matrices for the starred network, the purple pixels are zero value cells while the yellow ones take the value one. In the top left corner, the adjacency matrix of the network generated by the starring algorithm with 5 iterations of 4 copies at each step, the size being 1365. In the top right, bottom left and bottom right corners, we zoomed for the structure of the subnetworks of respectively 341, 85, 21 nodes.

Moreover, the peaks representing the eigenvalues multiplicity are in the same proportion as the copies parameter l, as expected. However, an unexpected feature emerges as some noise seems to appear as the third iteration is taken into account. We see that the eigenvalues having the third greater "weight" are not symmetric, and the noise seems to have the same weight. If we zoom in, we see that the self replicating structure is in fact shared by this noise, as in figure 3.9. Then, we may think that this is not a true noise, but rather a feature of the network we generated.



(a) (b)

Figure 3.8: Zoom of the perturbation in figure 3.6. We see that the self-similarity structure is kept by the noise, as the ratio of eigenvalues multiplicity is in both cases $1:4$.

This breaking of symmetry should be understood as strong non-bipartiteness. As a matter of fact, it can be linked to the very algorithm by which we build the network. Why then this breaking of symmetry appears only at the third iteration? The reason may be due to the fact that three is the minimal number of edges for minimal loops (triangles) to appear. And as triangles appear, the starrification algorithm replicates that structure over and over, preserving their relative "weight" in the moment counting.

Let us try to compute the moments for a the case in which $l_i = l$.

$$m_1 = \mathrm{Tr}(A) = \sum_{i=1}^{N} \lambda_i = 0$$

by definition of adjacency matrix.

$$m_2 = \mathrm{Tr}(A^2) = \sum_{i=1}^{N} \lambda_i^2 = J^2$$

which defines the variance of the distribution. Then for the third moment we have

$$m_3 = \mathrm{Tr}(A^3) = \sum_{i=1}^{N} \lambda_i^3 \neq 0$$

If we expand this sum we have

$$\sum_{i=1}^{N} \lambda_i^3 = \sum_{i_0} \lambda_{i_0}^3 + \sum_{i_1} \lambda_{i_1}^3 + \sum_{i_2} \lambda_{i_2}^3 + \cdots = \sum_{j}^{h} \sum_{i_j} \lambda_{i_j}^3$$

where we rearranged the sum according to each eigenvalue scale multiplicity. An estimate (large) for each term $\sum_{i_j} \lambda_{i_j}^3$ is $(\lambda_j^M)^3$ by which we mean the largest (in absolute value) eigenvalue relative to the scale j, multiplied by its multiplicity $l^{(h-j)}$.

Then let us go into the thermodynamic limit $N \to \infty$ and rescale this moment. N grows as l^h. The scaled third moment is

$$\frac{m_3}{N} \sim \sum_{j=0}^{h} \frac{(\lambda_j^M)^3}{l^j} \tag{3.3}$$

If the thermodynamic limit is performed by a fixed cloning parameter and an infinite number of iterations then the third moment does not vanish. This implies that the number of triangles grows as the number of nodes. If, however, the limit is performed by taking a very large number of copies and fixed steps, the third moments, hence the number of triangles vanishes. As a matter of fact, the part of the algorithm that generate connections are the starring steps (parameter h), while the cloning ones are devoted to making the structure scale-free. If this reasoning is correct, the symmetry should break only when starrification has been performed at least three times, because three is the minimal number of link addition steps that requires for a ensemble of disconnected nodes to form triangles among them. Indeed, from the figure we see that the symmetry breaking appears for the third

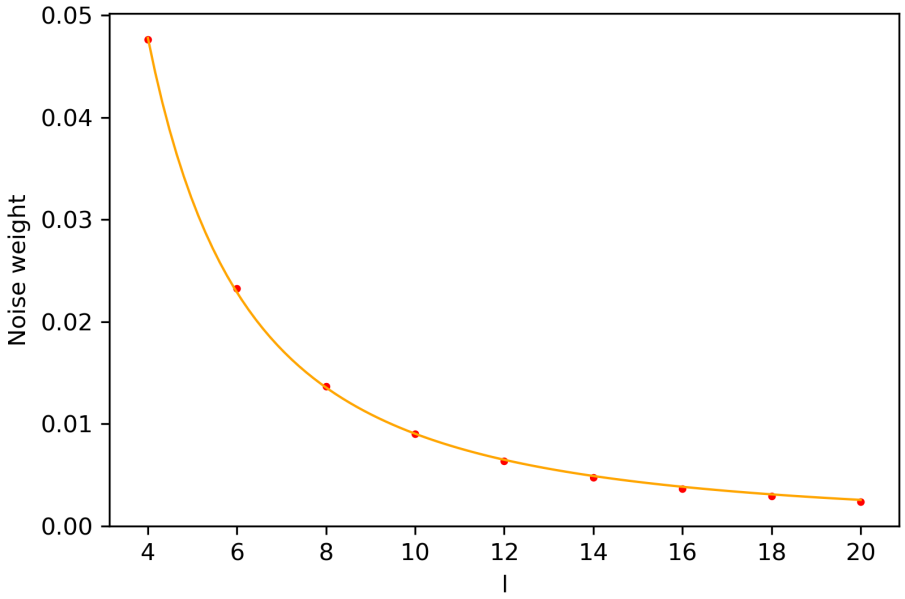


Figure 3.9: SB versus cloning parameter l. The networks were realised by fixing the iteration number to 3 but varying the number of copies l performed at each iteration. The power law decay fits the points almost perfectly.

most abundant eigenvalue, which is the one relative to the third step of the process.

In order to verify this conjecture, we may measure the magnitude of this symmetry breaking in two ways: we could count all the eigenvalues which do not have a symmetrical partner, however this method may overestimate the noise by taking into account also the eigenvalues from the third scale on that are only "perturbed" by the noise presence. Another method comes from the fact that we expect each eigenvalue at the scale j to be greater in magnitude with respect to the ones of the scale $j-1$. Then the noise estimate is performed by counting all the λ_j which are smaller than λ_{j-1}. In figure 3.9 we plot the estimate of the symmetry breaking (SB) versus the cloning parameter l according to this method. The simulation seems to confirm the conjecture about the third moment vanishing.

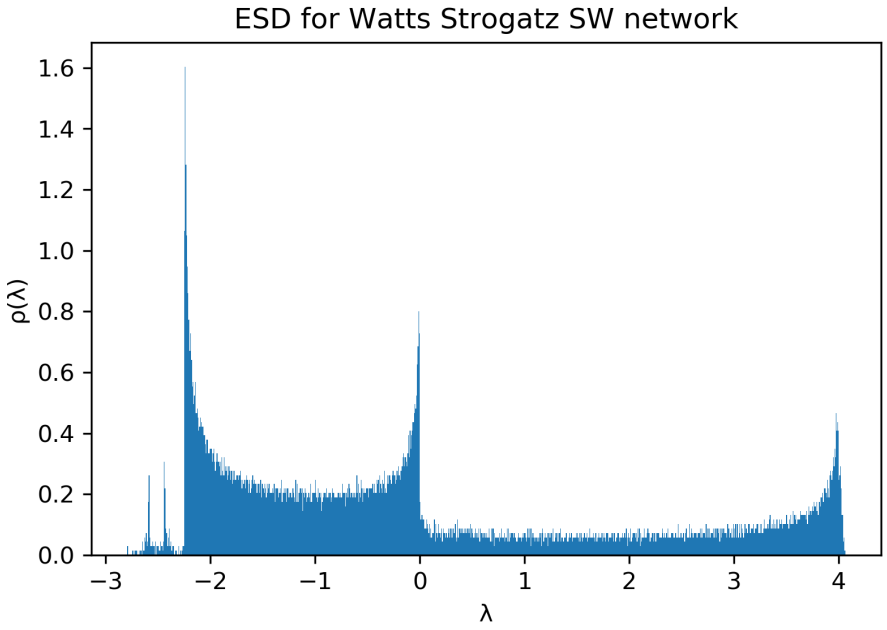


Figure 3.10: A Watts-Strogatz network with 10000 nodes, initial degree $d = 4$ and rewiring probability $p = 0.01$

3.3.7 The Small World network

There are no analytical results about the spectra of SW networks which are not empirical or results of simulations. In figure 3.10 we show the result of the spectrum of a Small World network obtained via the Watts-Strogatz algorithm. The most prominent feature is the lack of symmetry in the spectrum, meaning that the network is not bipartite. This breaking of symmetry may be explained by the non vanishing of the odd moments in the large limit. We will talk more extensively about this in the final remarks. Another feature is the mixed behaviour between a peaked and continuous distribution, which is linked both to hub structure (peaks meaning less variability of degree) and to navigability (continuous transition between states rather than discrete hops).

3.4 An analogy with QFT

We have seen in section 2.6 how the replica trick may be of use to give correction to the ESD of ideal infinite random networks. The reader may have noticed that those correction were actually perturbative expansions, according to the modified PDF in the case of finite mean value and a higher order one in the case of finite size effects. In this subsection we try to formalize that scheme of thought prior to the use of replica variables. The first thing that we notice is that when we have to compute the non-averaged SD we may recast the expression as

$$\rho(x) = -\frac{2}{N\pi}\,\mathrm{Im}\,\frac{\partial}{\partial x}\ln Z_N(x)$$

where we introduced the expression for a "partition-like" function for the adjacency matrix entries' ensemble, being

$$Z_N(x) = \int \prod_{i=1}^{N} d\phi_i e^{-\frac{i}{2}\sum_{i,j}\phi_i H_{ij}(x)\phi_j} = Z_N^0(x)\int \prod_{i=1}^{N} d\phi_i e^{-\frac{i}{2}\sum_{i,j}\phi_i A_{ij}\phi_j}$$

where $H_{ij}(x) = \delta_{ij}x - A_{ij}$ and $Z_N^0(x) = \int \prod_{i=1}^{N} d\phi_i e^{-\frac{ix}{2}\sum_i \phi_i^2}$ and we may see that a potential term appears

$$\begin{aligned}\langle Z_N(x)\rangle &= \int \prod_{i}^{N} d\phi_i e^{-\frac{ix}{2}\sum_i \phi_i^2}\left\langle e^{-\frac{i}{2}\sum_{ij}\phi_i A_{ij}\phi_j}\right\rangle = \\ &= \int \prod_{i}^{N} d\phi_i e^{-\frac{ix}{2}\sum_i \phi_i^2}\int \prod_{ij} dA_{ij}p(A_{ij})\int \prod_{i}^{N} d\phi_i e^{-\frac{i}{2}\sum_{ij}\phi_i A_{ij}\phi_j} = \\ &= \int \prod_{i}^{N} d\phi_i \prod_{ij} dA_{ij} e^{-\frac{i}{2}x\sum_i \phi_i^2 - \frac{i}{2}\sum_{ij}\phi_i A_{ij}\phi_j + \sum_{ij}\ln p(A_{ij})}\end{aligned}$$

Now, we see that a meaning arises from this formula. Let us assume to run a single-particle random walk on the network described by the adjacency matrix A. A vector ϕ_i represents a possible state of the particle on the graph, *i.e.* the particle being in the i^{th} node, and the product $A_{ij}\phi_i$ is the one-step evolution from node i to node j. Then, the product $\phi_i A_{ij}\phi_j$ represents the

one-step evolution from node i to node j, if possible. Let us pass to the Einstein convention for summation over repeated indices. Then, if we expand the exponential, we have the series

$$e^{-\frac{i}{2}\phi_i A_{ij}\phi_j} = 1 - \frac{i}{2}\phi_i A_{ij}\phi_j - \frac{1}{8}(\phi_i A_{ij}\phi_j)(\phi_i A_{ij}\phi_j) - \\ - \frac{i}{48}(\phi_i A_{ij}\phi_j)(\phi_i A_{ij}\phi_j)(\phi_i A_{ij}\phi_j) + \cdots$$

Now, let us take the second order term and rearrange it

$$\phi_i A_{ij}\phi_j\phi_i A_{ij}\phi_j = \phi_i A_{ij}\phi_j\phi_j A_{ji}\phi_i = \phi_i A_{ij}A_{ji}\phi_j$$

as of course $\phi_j\phi_j = 1$. This is the evolution of a state from node i to node j and back to i.

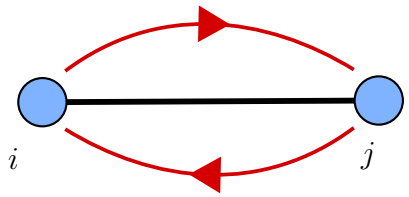


Figure 3.11: A one-loop correction to the state ϕ_i.

Then the third order term, recalling that $A_{ij} = A_{ji}$ and expanding $A_{ij} = \delta_{ik}A_{kj}$ and $\phi_i = \delta_{ik}\phi_k$, is readily

$$\phi_i A_{ij}\phi_j\phi_i A_{ij}\phi_j\phi_i A_{ij}\phi_j = \delta^2_{ik}\delta^2_{kj}\phi_i A_{ij}\phi_j\phi_j A_{jk}\phi_k\phi_k A_{ki}\phi_i = \delta_{ijk}\phi_i A_{ij}A_{jk}A_{ki}\phi_i$$

We introduced the δ_{ijk} to resume the deltas in a single one having the same effect.

The meaning of this is more clear when we consider the fourth order term

$$\begin{aligned}
&\phi_i A_{ij}\phi_j\phi_i A_{ij}\phi_j\phi_i A_{ij}\phi_j\phi_i A_{ij}\phi_j = \\
&\quad = \frac{1}{2}\left((\phi_i A_{ij}\phi_j\phi_i A_{ij}\phi_j)^2 + \phi_i A_{ij}\phi_j\phi_i A_{ij}\phi_j\phi_i A_{ij}\phi_j\phi_i A_{ij}\phi_j\right) = \\
&\quad = \frac{1}{2}\left((\phi_i A_{ij}A_{ji}\phi_i)^2 + \delta^2_{ik}\delta^2_{jl}\phi_i A_{ij}\phi_j\phi_j A_{jk}\phi_k\phi_k A_{kl}\phi_l\phi_l A_{li}\phi_i\right) = \\
&\quad = \frac{1}{2}\left((\phi_i A_{ij}A_{ji}\phi_i)^2 + \delta_{ijkl}\phi_i A_{ij}A_{jk}A_{kl}A_{li}\phi_i\right)
\end{aligned}$$

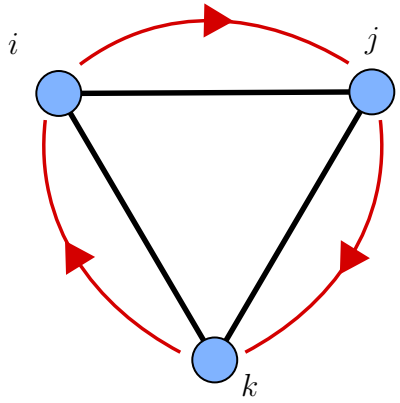


Figure 3.12: A two-loop correction to the state ϕ_i.

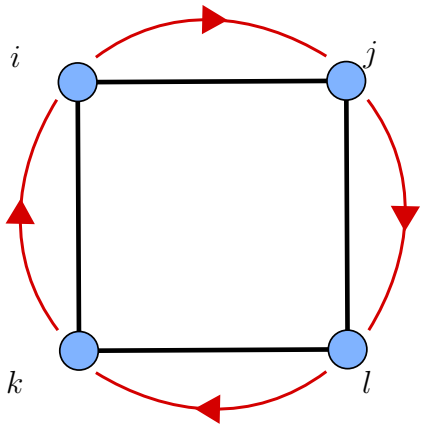


Figure 3.13: A three-loop correction to the state ϕ_i.

A clear path emerges. The expansion of the exponential is precisely counting all the possible loops that a state may perform on the network, the n^{th} term corresponding to a n-length path to follow. A weight $\delta_{i_1 \cdots i_k}$ is introduced to see that for every term, the more a loop is "further-reaching", the less its contribution to the overall term. For more than three steps paths we also see that they are possible by summing prior blocks and adding a new "species" of paths when increasing the order, *i.e.* the length.

According to this interpretation the formula for the partition function becomes

$$\langle Z_N(x)\rangle = \int \prod_i^N d\phi_i \prod_{ij} dA_{ij} \prod_{ij} e^{-\frac{i}{2}x\phi_i\delta_{ij}\phi_j - \frac{i}{2}\phi_i A_{ij}\phi_j + \ln p(A_{ij})} =$$

$$= \int \prod_i^N d\phi_i \prod_{ij} dA_{ij} \prod_{ij} e^{-\frac{i}{2}x\phi_i\phi_i} \Bigg[1 - \frac{i}{2}\phi_i A_{ij}\phi_j -$$

$$- \frac{1}{8}\phi_i A_{ij}A_{ji}\phi_i - \frac{i}{48}\sum_k \delta_{ijk}\phi_i A_{ij}A_{jk}A_{ki}\phi_i -$$

$$- \frac{1}{768}\Big[(\phi_i A_{ij}A_{ji}\phi_i)^2 + \sum_{kl}\delta_{ijkl}\phi_i A_{ij}A_{jk}A_{kl}A_{li}\phi_i\Big] - \cdots\Bigg] p(A_{ij})$$

Now we may try giving a meaning to this formula. The Z_N^0 part of the function is a soliton wave[10]. Then, when averaging over the ensemble of the adjacency matrix entries what we are doing is propagating such soliton wave for each possible (weighted) self-loop of each state ϕ_i, the term $p(A_{ij})$ telling if such expansion is possible. This represents a random walk from the state ϕ_i that ends in the same state, taking into considerations all the possible ways of exploring the states network. It is to be remarked that we are taking into consideration tours, hence directional paths, so that the RW performed may be reversible as well as it may be not. Finally, the integration over the states represents a mean of the random walk, which is a random walk of N particles, each starting from one of the N different states. The loop expansion here performed is, to the best of my knowledge, new in the field of spectral analysis of network. However, it looks like the paper [28] is implicitly taking this framework to develop a continuous and more QFT-like theory. Still, I believe the continuum limit under the replica assumptions, while helping greatly in the actual computation of the spectral density, may lead astray from the interpretation just given. Before moving to the next subsection, I believe it is worthwhile to highlight that what we just found could be a nice candidate to give an interpretation to the partition function formalism for networks spectra in a dynamic context. Also, if such interpretation were to be expanded alongside field theory lines of thought, we may wonder what an interaction term would represent to our eyes.

[10] We can see that by solving for ρ when only that term appears.

3.4.1 Renormalization equation

A proper QFT analogy would not be complete without an insight over renormalization. Indeed the next chapter will be entirely dedicated to the implementation of the Renormalization Group over classes of networks, so I deem trying to establish a connection to be useful. We had highligths in the previous section that contemporary theoretical physics methods are of use in the field of RMT, and on the same level of QFT perturbative approaches there should be mentioned the supersymmetric approach of [12]. The bridge between the QFT analogy and RG tecniques may come from the work [23]. They redefine an effective potential from the one that emerges from the averaging of the partition function

$$V(t,\phi) = \ln \int D\psi e^{\frac{1}{t}\sum_i \psi_i^2 + V_0(\phi+\psi)} - N \ln t$$

where $V_0(\phi)$ is the potential that appears in the non-free equation

$$V_0(\phi) = \ln\langle e^{-\frac{i}{2}\sum_i \phi_i A_{ij} \phi_j}\rangle$$

And $t = 1/z$ was introduced for convenience. Now, such potential obeys a Gaussian convolution rule

$$V(t+s,\phi) = \ln \int D\psi e^{-\sum_i \frac{\psi_i^2}{s} + V_0(t,\phi+\psi)} - N \ln s$$

and for a small $s = dt$ we have this kind of equation

$$\frac{\partial V(t,\phi)}{\partial t} = \sum_i \frac{\partial^2 V(t,\phi)}{\partial \phi_i^2} + \frac{\partial V(t,\phi)}{\partial \phi_i}\frac{\partial V(t,\phi)}{\partial \phi_i} \tag{3.4}$$

or, in the equivalent integral form,

$$V(t,\phi) = V_0(t,\phi) + \int_0^t ds \sum_i \frac{\partial^2 V(t,\phi)}{\partial^2 \phi_i} + \frac{\partial V(t,\phi)}{\partial \phi_i}\frac{\partial V(t,\phi)}{\partial \phi_i}$$

As it is said in the article, these equations correspond to a Renormalisation equation for a quantum field theory in 0 dimensions. On the LHS of equation 3.4.1 we recognise a term describing how the potential that generates the new theory, *i.e.* a "renormalised" network, scales according to the

dimension parameter t[11].

Let us take a closer look to the terms in the right hand side. The action of the first term, the second derivative one, on the loop expansion in the potential is the following

$$\begin{aligned}\phi_i A_{ij}\phi_j &\longrightarrow 0\\ \phi_i A_{ij}A_{ji}\phi_i &\longrightarrow A_{ij}A_{ji}\\ \phi_i A_{ij}A_{jk}A_{ki}\phi_i &\longrightarrow A_{ij}A_{jk}A_{ki}\end{aligned}$$

that is to truly count the loops which are present in the network generated by the potential V. The double derivative term instead acts as

$$\begin{aligned}\phi_i A_{ij}\phi_j &\longrightarrow \phi_j A_{ji}A_{ij}\phi_j\\ \phi_i A_{ij}A_{ji}\phi_i &\longrightarrow \phi_i A_{ij}A_{ji}A_{ij}A_{ji}\phi_i\\ \phi_i A_{ij}A_{jk}A_{ki}\phi_i &\longrightarrow \phi_i A_{ij}A_{jk}A_{ki}A_{ik}A_{kj}A_{ji}\phi_i\end{aligned}$$

that is a "counting trees" term. Indeed, the second derivative term "detaches" the possible loop expansions of a state and the summation in the equation 3.4.1 is precisely the counting of such loops, *i.e.* tours of the network; while the double derivative one has the effect of opening two loops and attaching them, making them to be travelled like trees from roots to leaves and back on. Then the invariance with respect to scaling for a network whose potential has Gaussian cumulants[12] requires that a loop-counting term equates a tree-counting term. I believe this is related to the proof we gave for the semicircular law with the moments method; as a matter of facts we showed that in the case of Wigner matrices the only moments (loops) that do not vanish in the large dimension limit are those tours performed on trees.

[11] When computing the eigenvalues z of an adjacency matrix they scale linearly with the dimension of the matrix, *viz.* the dimension of the network.

[12] This request is made in order to achieve the convolution formula.

Chapter 4

Renormalization Group Approach

4.1 Introduction

A fractal framework such as the one provided by the scale-free class of network should be ideal for the application of Renormalization Group (RG) techniques. Indeed, the machinery provided by such method will be insightful to discover the "phases" networks may show. First, we need to define the environment in which the RG is employed. A fractal (scale-free) network is a network that shows the properties of self-similarity at different scales. The first problem one may encounter is that real world networks are not self similar at all scales, being self similar only when changing certain scopes, this fact due to finiteness of real world examples. To overcome this fact we may assume to work with an infinitely extended network, whose number of nodes and edges diverges, ensuring that the scope-changing (yet to be properly defined) may be performed indefinitely. The first issue arising in the development of such framework is to properly define the method by which the RG may be implemented. Some works [34], [33] suggest that the "box-covering" algorithm is the one suited to perform such task. Indeed, the procedure is to regroup nodes in the underlying network in boxes whose capacity b means that nodes in each box may be at most at distance b. Then, each box collapses into a "supernode", and the edges between such nodes exist if any of the "undernodes" in two different boxes were connected. Such procedure allows also to define quite naturally a fractal dimension of the network, *i.e.*

a measure of the scaling feature of the graph. Indeed, letting N_0 to be the number of nodes of the starting network and N_b the number of nodes in the b-renormalized network, the ratio between them shall be

$$\frac{N_0}{N_b} \sim b^{d_F}$$

where d_F is the now defined fractal dimension. We also remark that for a thoroughly scale-free (SF) network the above relation is exact. The action of such renormalization that maps a network G_0 into G_b is shown in figure 4.1

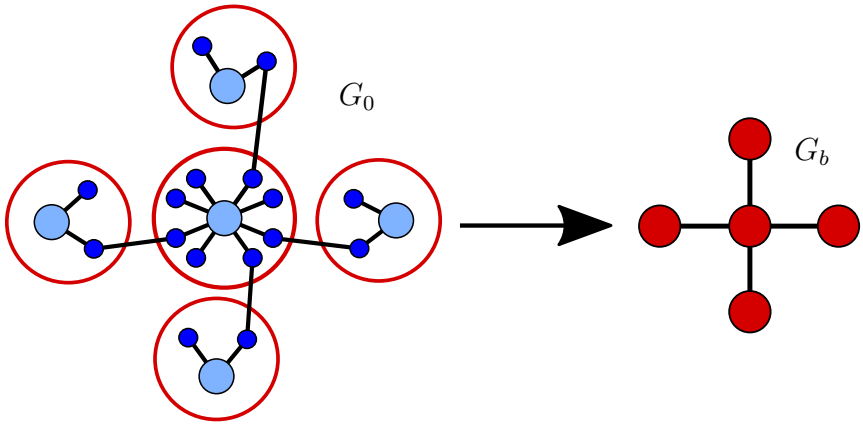


Figure 4.1: Visualization of the action of the renormalization transformation via the box-covering implementation.

4.2 Implementation of perturbations

In the framework of thermodynamic limit of ideally SF networks one has that the action of the RG transformation R_b is to left unchanged the starting network G_0 so that

$$G_b = R_b[G_0] = G_0$$

We may wonder then what happens when small perturbations W occur to interfere with the purely SF structure in the renormalized case. A perturbation to the self-similar graph G_0 is performed through the random

addition of edges between nodes, that act as shortcuts when they link two distant nodes. We introduce those shortcuts with probability depending on the distance between two nodes r

$$p(r) = Ar^{-\alpha}$$

where A is an appropriate renormalizing constant, thus the perturbed graph G_0' is obtained. Then, we may wonder what is the effect of the renormalization on the perturbation, *i.e.* how to realize the W_b that correctly gives the perturbed and renormalized graph G_b'. The scheme is the following

$$\begin{array}{ccc} G_0 & \xrightarrow{W} & G_0' \\ {\scriptstyle R_b}\downarrow & & \\ G_b & \xrightarrow{W_b} & G_b' \end{array}$$

and the visualization is shown in figure 4.2.

We do so by computing the probability of *not having* shortcuts between two nodes at distance br in the renormalized network G_b'. This is

$$P_{NS} = [1 - p(br)]^{b^{2d_F}}$$

Then the probability that *there are* shortcuts at distance r in G_b' is $1-P_{NS}$

$$p(r) = 1 - [1 - p(br)]^{b^{2d_F}} = 1 - [1 - A(br)^{-\alpha}]^{b^{2d_F}}$$

The addition of shortcuts obviously increase the connectivity of the network, but we may wonder if there are any fixed points when the renormalization is applied indefinitely. Indeed, one may take the limit for $b \to \infty$ in order to represent the infinite repetition of R_b. With the substitutions $x = \frac{(br)^\alpha}{A}$ and $B(r) = A^{\frac{2d_F}{\alpha}} r^{-2d_F}$ the limit is performed

$$p^*(r) = \lim_{b\to\infty} p_b(r) = 1 - \lim_{x\to\infty} \left(1 - \frac{1}{x}\right)^{B(r)x^{\left(\frac{2d_F}{\alpha}\right)}} = 1 - \lim_{x\to\infty} e^{-B(r)} e^{-x^{\left(\frac{2d_F}{\alpha}-1\right)}}$$

Now, according to the value of the exponent $s = \frac{\alpha}{d_F}$ we may divide the space of configurations in phases that corresponds to the fixed points of the flow. Indeed, the transformation is self-similar in three cases. The first one is when $s < 2$, *viz.* $p^*(r) = 1$, so that there is always a shortcut between two

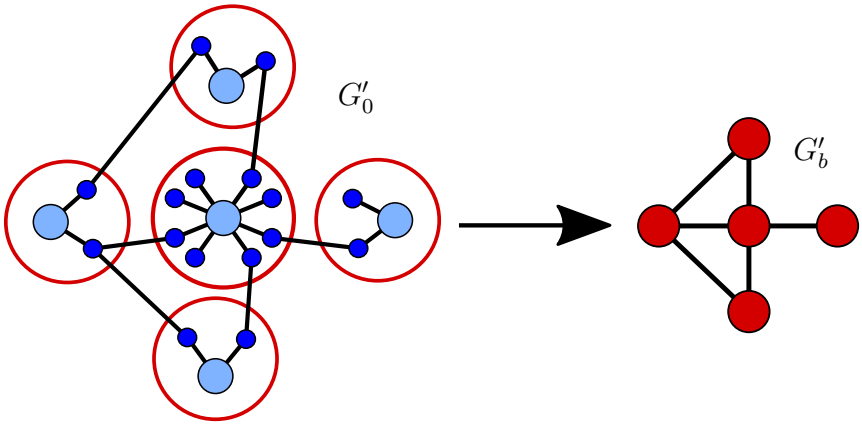


Figure 4.2: The renormalization action on the perturbed network when shortcuts are added.

nodes at any distance. This is the case of a totally connected network, *i.e.* a complete graph. It is not surprising that a complete graph is self-similar, indeed, the wholly connected network is nothing but a single node to the eyes of the RG transformation. Then, the case in which $s > 2$ is the one of a purely fractal network, as the probability of finding any shortcut between two nodes shrinks quite rapidly to $p^*(r) = 0$. This too does not come unexpected, as the addition of shortcuts is insufficient to give a significant contribution to the renormalized network. To visualize it, it would be as if the shortcut are added so scarcely that they are only inside the boxes, so they have no effect when widening the scope. However there is also a third fixed point, which may seem startling at a first glance. Indeed, when the critical exponent $s = 2$ there is a fixed numbers of shortcuts at any scale of length. This is quite remarkable, as it identifies a phase in which the self-similarity features does not hinder the navigability properties like in the pure SF case. The analysis of this latter fixed points will be best performed in the next section.

4.3 Analysis of the phase space

The first thing to point out about the fixed points is their stability. With respect to the perturbation, it is clear that the fixed point of the complete graph, which we call G_c, is attractive and stable. The purely SF graph G_0 shows a mixed behaviour; any G' is thrown apart from it by the RG flow

when $s \leq 2$, while being attractive only in the case $s > 2$, however it is clearly not stable. The third and most intriguing fixed point, which we call G_m, is attractive, but only for a unique value of the critical exponent.

Nonetheless complete and SF graphs are not the only network models we know, hence it is worthwhile trying to locate in the space of configuration random and regular graphs too, as they are the models we know the most about. Indeed, the random network model should locate in a region between G_0 and G_c. The reason for this fact, which should be anyway intuitive, will be clearer in the followings when trying to link the RG analysis with network SD. Also, the fact that when randomized regular graphs tend to the Erdos-Renyi type allows us to say that the RG flow thrusts G_r in the same region of "randomness transition". In figure 4.3 the action of the RG is represented in the space of configurations.

4.3.1 Identification of G_m

The third fixed point that was found is not immediately identifiable. In order to further investigate it we may look closer at its structural properties. When the perturbation maps G_0 into G' we should be careful of what happens to the average degree. Indeed, let d_0 and d' be the average degree of the unperturbed and perturbed network respectively. Then, their difference is

$$d' - d_0 = \frac{2M(D)}{N_0}$$

where $M(l)$ is the number of shortcuts at distance l, D is the diameter of the network and the factor 2 is due to each edge increasing the degree of 2 nodes. Now, the number of shortcuts is given by

$$M(l) \simeq \int_1^l p(r) d_F r^{d_F - 1} dr = d_B \int_1^l A r^{-\alpha} r^{d_F - 1} dr = \frac{A}{1-s} [l^{d_F(1-s)} - 1]$$

hence the average degree difference is

$$d' - d_0 = \frac{2A}{1-s} \left[\frac{D^{d_F(1-s)} - 1}{N_0} \right]$$

when we apply the RG the total number of shortcuts is simply the number of shortcuts linking different boxes, that is

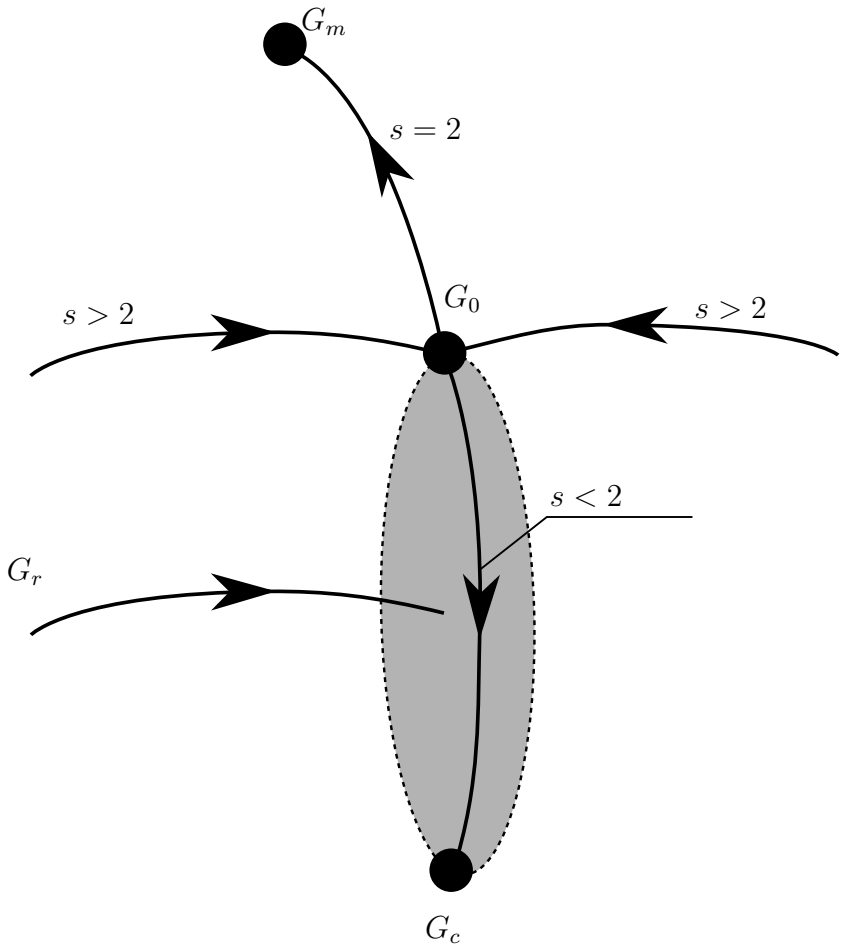


Figure 4.3: A representation of the RG flow on the space of configurations. The three fixed points are G_0, G_c and G_m. The flow lines are drawn and the value of the critical exponent for which they have such direction is specified. The grey region between G_0 and G_c is the one where we locate the random network model. The line describing the behaviour of G_r is only indicative that the flow should push it into the random grey region.

$$d_b - d_0 = \frac{2(M(D) - M(b))}{N_b} = (d' - d_0) f_N(b)$$

where the function $f_N(b)$ is given by

$$f_N(b) = \frac{D^{d_F(1-s)} - b^{d_F(1-s)}}{D^{d_F(1-s)} - 1} b^{d_F}$$

To perform the limit of infinite renormalization we define $x_b = \frac{N_0}{N_b} = b^{d_F}$, and we have the scaling

$$f_N(x_b) \sim x_b^{\lambda}$$

where λ is a new critical exponent given by

$$\lambda = \begin{cases} 1, & \text{if } s \leq 1 \\ 2 - s, & \text{if } s > 1 \end{cases}$$

This exponent allows us to better understand the fixed point G_m. When the critical exponent $s = 2$ we see that λ introduces a new separation of phases. There is phase where $\lambda < 0$ and $s > 2$, where the SF structure is conserved, as previously stated, and wee see that this is consistent with the self similarity requirement that nullifies the difference between the degrees of the renormalized and starting networks. Conversely in the case when $\lambda > 0$ and $s < 2$ the result is a growth of the degree difference, which is consistent with the fact that the flow thrusts toward the complete graph point. Then, in the very condition of $s = 2$ we have $\lambda = 0$, meaning there is no change in the degree difference, which implies that the effect of the RG flow on the average degree is the same as that of the perturbation. We are in a transition region where the diameter of the graph grows logarithmically, whilst in the SF region the diameter growth is a power law. This fact may encourage us to think to the phase $s = 2$ as a region where "small-world" phenomena occur.

4.3.2 Space partition

According to the three possible values of the critical exponent s we may part the configuration space into three main regions. Although being a preliminary analysis, we may identify each region with one of the main classes of networks, trying to link these phase partition to the results we had in previous chapters.

Scale-Free region

This region is the easiest to identify. The intuitive reason was made clear before when explaining that in this case the addition of shortcuts is overwhelmed by the boxing of nodes. Also, SF networks are classical cases of rich-get-richer processes, also called in a fancier way "Indian buffet processes[1]" or "Matthew effects[2]". These kind of processes are based on preferential attachment strategies, that is nodes that have a higher degree tend to be chosen to be linked with more than others when new nodes are added. The preferential attachment scheme leads to the shrinking of the variance of the adjacency matrix entries distribution, which results in peaks in the ESD outcome. Quite intuitively too, less variability implies the accumulation of entries on certain sites rather than a smoother and more sparse distribution.

Small World region

The identification of the G_m point with a Small World network was established before. Here we remark that the region is in fact just a line, for which the critical exponent s has the precise value of 2. In this case the growth process is peculiar, as it must show fractality, hence some form of preferential attachment, and some "intelligent randomness" that breaks the pure scale-free structure but not thrusting towards the randomness region. Such balance between preferential attachment and "guided" or "intelligent" choices of linking nodes is precarious, and should be investigated further.

Random region

This region is the most populated and explored as far as we know by now. Networks that are Random in the proper sense lie in the grey region of figure 4.3, but it is easy to see that many models (regular, non purely SF, *etc.*) have the RG flow to thrust them into this area. According to the interpretation and analogy with QFT in section 3.4.1, this is the region where "loops are trees", in the sense that all the self-states are mainly performed by tree paths on the network. It is the same result of the classic Wigner law, that is the

[1] The name is due to the fact that, when having to choose among many different unknown possibilities (a table full of different Indian plates), choices often fall on already sampled items (already tasted plates).

[2] 'For everyone who has will be given more, and he will have an abundance', Matthew, 25:29.

one saying that in the large limit the only contributions to the spectra are the path performed on trees (the Catalan numbers).

4.4 Phase transitions

Having established the three main phases, it is natural to ask which are their prominent feature and how the transition from one phase to another is achieved. In other words, are there any order parameters that describe these phase transitions? According to some of the results we found in the previous sections, we may try to answer to this question.

4.4.1 Scale Free to Wigner

When we analysed the behaviour of the Wigner law for non-zero average Random Matrices we noticed that there was a threshold in equation 2.6.4. When taking into consideration random networks, we have to take into consideration this ESD, as the average entry of the adjacency matrix is surely positive. The threshold was determined by the ratio between the average of the entries and their variance

$$\omega = \frac{\langle A_{ij} \rangle}{J} = \begin{cases} > 1 \text{ peak behaviour} \\ < 1 \text{ Wigner behaviour} \end{cases}$$

Following the conjecture that the peak behaviour in the spectrum is to be linked with the emergence of scale free properties, then this parameter sets the threshold for a first fractal scale to emerge. In other words, there is not a scale free overall behaviour, but there is a scale at which the behaviour of nodes is really similar. The emergence of other scales of fractality is more complicated, as it needs the separation of the spectrum to take place. This change between continuity of the support and discretisation may be another candidate to set the threshold for a more proper fractal transition, however it is quite hard to understand how to describe it.

4.4.2 Scale Free to Small World

This transition, as we have hinted at in the course of the previous chapters, is linked to the breaking of the spectral symmetry. Such hypothesis is corroborated by the empirical spectral densities we have shown before. Also, there

is a physical meaning to that. As we know by now the spectral symmetry breaks if the odd moments of the adjacency matrix do not vanish. Such odd moments are the number of odd tours in the network, in particular triangles for the third moment. Then this breaking of symmetry corresponds to the presence of "faster ways" to travel the network than by even tours, leading to the shrinkage of the diameter, hence a Small World behaviour.

The model we introduced in section 3.3.6 to portray the behaviour of Scale Free network is of use to better understand this transition. The Scale Free behaviour is implementing quite naturally the action of the Renormalisation Group, but it has two different ways of reaching the termodynamical limit. The first is by fixing the number of iterations steps h (starring operations) but making a large amount of copies l. The second one is by fixing the number of copies and performing a very large number of iterations. As we have seen in figure 3.9, if we measure the symmetry breaking magnitude it vanishes when the former way of implementing the termodynamic limit is chosen. This would lead to a perfect Scale Free network, very close to G_0 in terms of the phase it represents.

If we take a finite number of copies however the situation changes. At the third iteration the symmetry breaking appears, and this is due to the fact that we must certainly have created at least a triangle in the second step, this triangle being replicated and scaling as the number of copies itself. So after the second step if the number of copies is not large, the number of triangles Δ begins to grow as the number of copies $\Delta \sim l^{h-2}$ and their ratio to the size of the network N is non vanishing and almost constant

$$\frac{\Delta}{N} \sim \frac{1}{l^2}$$

We remark that is consistent with the estimate we did for the third moment in 3.3.6, as for the first term in that sum $j = 0$ the value is null while for the second term $j = 1$ they are symmetric and their sums vanish pairwise.

This behaviour for the relative presence of triangles in a network and the fact that it is the first odd moment which does not always vanish may lead us to think that it is a suitable order parameter for the phase transition from a purely Scale Free network to another that exhibits Small World features.

Another hint that the parameter Δ is the right one to describe the transition between the two phases is given by the Free Central Limit Theorem. In particular in equation 2.5.1 we see that the vanishing of the normalised third moment is the requirement for the sequence of matrices to follow the

semicircular law. In that very equation if the matrix X is not renormalised by the factor $\sqrt{n}$ and we choose it to be an adjacency matrix A, then the quantity to vanish in order for the semicircular law to hold is precisely the number of triangles in the network described by A.

Chapter 5

Final Remarks

Complex systems physics is still a developing field of knowledge. Its soundest results are the ones that rely on the study of complex networks, still in most cases there are no exact results other than those on toy models. In this thesis we explore a new approach that focuses on perturbed, non ideal networks rather than exact graphs. To do so, statistical mechanics and other physics-inspired methods are borrowed to investigate the effect of finite size, fluctuations and correlations effects on ideal random networks. We chose to focus on spectra of networks, mainly because eigenvalue distributions encode many features both about dynamical and structural properties.

Algebraic graph theory is the most powerful tool for obtaining exact result for the spectra of a precise network (a line, a wheel, a complete one, *etc*). However, most of these graphs show regular features which fails to suit the so-called real world networks. Thus the need to provide with models that mimic the traits of certain "classes" or "ensembles" of networks emerges. When facing such task, the exact results of algebraic graph theory are no longer holding, and we should go for new methods to solve the issue.

In this thesis we adopted two different approaches: Random Matrix Theory and Renormalisation techniques. When dealing with the former we tried to see how the rigorous results of this field may be transposed over networks.

In sections 2.3, 3.3.1 we managed to show that the Wigner law for moments for a semicircular distribution, when computed with the moments method is giving the physical insight about a structure that the network displays, *i.e.* the tree skeleton.

Moreover, in section 2.4.1 we saw that the semicircular distribution for

the spectrum is the "most probable" one when we consider a random process of distribution of links. This was complemented by showing that Wigner law is also the counterpart of the Central Limit Theorem in the case of non commuting random variables 2.5.1.

Replica theory 2.6 allowed us to compute the corrections to the limiting spectral distribution we had found, to better fit the problem we were studying and giving a possible insight to a dynamical interpretation of random walk performed on a network which was sketched in section 3.4.

In the successive chapter we tried to set a framework in which algebraic graph theory, Random Matrix Theory and other physics-inspired techniques could work together in order to investigate properties of the main ensembles of networks that literature is interested in.

Some exact law about spectra of regular graphs 3.3.4 and line ones 3.3.5 were obtained to better interpret the results of the first chapter. This gave way to the analysis of a new generative model for a Scale Free network which could describe the behaviour of the spectrum of this class of graphs in section 3.3.6. The unexpected result about the symmetry breaking of this model, while preserving the fractal structure, led us to conjecture that it might be a phase transition phenomenon to the class of so called Small World networks. We also introduced a possible dynamical interpretation of what is a structural feature, *viz.* the spectrum of the adjacency matrix. This analogy, if properly expanded, may lead to the understanding of different processes other than simple random walks on networks in which agents may interact among them.

Eventually we introduced the Renormalisation group technique to analyse the configuration space of the different ensembles of networks. In this language, the classes we previously established became the phases in which the space was partitoned.

We point out that three are the fixed points, *i.e.* the ensembles which are stable with respect to the flux of the RG. The purely scalefree one, which was attributed to the Scale Free ensemble of Barabasi-Albert-like networks, was easy to identify. The complete network, a trivial point in terms of interest, was still helpful because it allowed to identify in which region the Random network ensemble should be located, being the complete graph the limiting case for a Erdös graph with connection probability $p = 1$. The third fixed point was the most startling feature of the analysis. If the case we made along

with [32] holds, then this point is a special case of fractality that hinders the growth of the diameter, making it a good candidate to portray the behaviour of Small World networks.

In order to analyse this kind of transition, the model we developed to mimic the Scale Free behaviour in section 3.3.6 could be suitable. Indeed, one of the few information we have about the spectrum of SW networks is that it is not symmetric, making the contribution of third moments, *viz.* triangles loops, extensive even in the thermodynamic limit.

The breaking of symmetry in the model we propose show some promising features that may make it suitable to represent the behaviour of self-similar networks when near to undergo a phase transition mechanism of the kind SF $\rightarrow$ SW. This could help understanding how the typical self-replicating structure of many systems encountered in nature is balanced by the small world phenomenon in order to achieve better control of processes.

Bibliography

[1] ZD Bai, Baiqi Miao, and Jhishen Tsay. "Remarks on the convergence rate of the spectral distributions of Wigner matrices". In: *Journal of Theoretical Probability* 12.2 (1999), pp. 301–311.

[2] Zhidong Bai, Zhaoben Fang, and Ying-Chang Liang. *Spectral theory of large dimensional random matrices and its applications to wireless communications and finance statistics: random matrix theory and its applications*. World Scientific, 2014.

[3] Zhidong Bai and Jack W Silverstein. *Spectral analysis of large dimensional random matrices*. Vol. 20. Springer, 2010.

[4] Albert-László Barabási and Réka Albert. "Emergence of scaling in random networks". In: *science* 286.5439 (1999), pp. 509–512.

[5] Marian Boguna, Dmitri Krioukov, and Kimberly C Claffy. "Navigability of complex networks". In: *Nature Physics* 5.1 (2009), pp. 74–80.

[6] Béla Bollobás and Oliver Riordan. "Sharp thresholds and percolation in the plane". In: *Random Structures & Algorithms* 29.4 (2006), pp. 524–548.

[7] Antoni Calvo-Armengol and Matthew O Jackson. "The effects of social networks on employment and inequality". In: *American economic review* 94.3 (2004), pp. 426–454.

[8] Alessio Catanzaro. "Spectral properties of stochastic matrices: an application to random walks". URL: `http://amslaurea.unibo.it/14519/`.

[9] Giovanni M Cicuta. "Real symmetric random matrices and path counting". In: *Physical Review E* 72.2 (2005), p. 026122.

[10] Gurjeet S Dhesi and RC Jones. "Asymptotic corrections to the Wigner semicircular eigenvalue spectrum of a large real symmetric random matrix using the replica method". In: *Journal of Physics A: Mathematical and General* 23.23 (1990), p. 5577.

[11] Pierfrancesco Dionigi. "A random matrix theory approach to complex networks". URL: `http://amslaurea.unibo.it/18513/`.

[12] Margherita Disertori and Vincent Rivasseau. "Random matrices and the Anderson model". In: *arXiv preprint math-ph/0310021* (2003).

[13] Samuel F Edwards and Raymund C Jones. "The eigenvalue spectrum of a large symmetric random matrix". In: *Journal of Physics A: Mathematical and General* 9.10 (1976), p. 1595.

[14] George FR Ellis. "Physics, complexity and causality". In: *Nature* 435.7043 (2005), p. 743.

[15] Paul Erdős and Alfréd Rényi. "On the evolution of random graphs". In: *Publ. Math. Inst. Hung. Acad. Sci* 5.1 (1960), pp. 17–60.

[16] Adina Roxana Feier. "Methods of proof in random matrix theory". PhD thesis. Harvard University, 2012.

[17] FR Gantmacher. "Theory of Matrices, Vol. I & II. Chelsea". In: *New York* (1959).

[18] MJC Gover. "The eigenproblem of a tridiagonal 2-Toeplitz matrix". In: *Linear Algebra and its Applications* 197 (1994), pp. 63–78.

[19] Matthew O Jackson. *Social and economic networks*. Princeton university press, 2010.

[20] Matt J Keeling and Ken TD Eames. "Networks and epidemic models". In: *Journal of the Royal Society Interface* 2.4 (2005), pp. 295–307.

[21] William Ogilvy Kermack and Anderson G McKendrick. "A contribution to the mathematical theory of epidemics". In: *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character* 115.772 (1927), pp. 700–721.

[22] Thomas Koshy. *Catalan numbers with applications*. Oxford University Press, 2008.

[23] Thomas Krajewski. "A Renormalisation Group Approach to the Universality of Wigner's Semicircle Law for Random Matrices with Dependent Entries". In: *Advances in High Energy Physics* 2017 (2017).

[24] Jarosław Kwapień and Stanisław Drożdż. "Physical approach to complex systems". In: *Physics Reports* 515.3-4 (2012), pp. 115–226.

[25] Giacomo Livan, Marcel Novaes, and Pierpaolo Vivo. *Introduction to random matrices: theory and practice*. Springer, 2018.

[26] Brendan D McKay. "The expected eigenvalue distribution of a large regular graph". In: *Linear Algebra and its Applications* 40 (1981), pp. 203–216.

[27] Madan Lal Mehta. *Random matrices*. Vol. 142. Elsevier, 2004.

[28] Fernando L Metz, Giorgio Parisi, and Luca Leuzzi. "Finite-size corrections to the spectrum of regular random graphs: An analytical solution". In: *Physical Review E* 90.5 (2014), p. 052109.

[29] Marc Mézard, Giorgio Parisi, and Miguel Virasoro. *Spin glass theory and beyond: An Introduction to the Replica Method and Its Applications*. Vol. 9. World Scientific Publishing Company, 1987.

[30] Cristopher Moore and Mark EJ Newman. "Epidemics and percolation in small-world networks". In: *Physical Review E* 61.5 (2000), p. 5678.

[31] Mark EJ Newman. "Spread of epidemic disease on networks". In: *Physical review E* 66.1 (2002), p. 016128.

[32] Hernán D Rozenfeld, Chaoming Song, and Hernán A Makse. "Small-world to fractal transition in complex networks: a renormalization group approach". In: *Physical review letters* 104.2 (2010), p. 025701.

[33] Chaoming Song, Shlomo Havlin, and Hernan A Makse. "Self-similarity of complex networks". In: *Nature* 433.7024 (2005), p. 392.

[34] Chaoming Song et al. "How to calculate the fractal dimension of a complex network: the box covering algorithm". In: *Journal of Statistical Mechanics: Theory and Experiment* 2007.03 (2007), P03006.

[35] Terence Tao. *Terry Tao's Site*. URL: `https://terrytao.wordpress.com`.

[36] Linh V Tran, Van H Vu, and Ke Wang. "Sparse random graphs: Eigenvalues and eigenvectors". In: *Random Structures & Algorithms* 42.1 (2013), pp. 110–134.

[37] Edwin R Van Dam and Willem H Haemers. "Which graphs are determined by their spectrum?" In: *Linear Algebra and its applications* 373 (2003), pp. 241–272.

[38] Piet Van Mieghem. *Graph spectra for complex networks*. Cambridge University Press, 2010.

[39] JJM Verbaarschot and MR Zirnbauer. "Critique of the replica trick". In: *Journal of Physics A: Mathematical and General* 18.7 (1985), p. 1093.

[40] Duncan J Watts and Steven H Strogatz. "Collective dynamics of 'small-world'networks". In: *nature* 393.6684 (1998), p. 440.

[41] Eugene P Wigner. "Characteristic vectors of bordered matrices with infinite dimensions I". In: *The Collected Works of Eugene Paul Wigner*. Springer, 1993, pp. 524–540.

[42] Eugene P Wigner. "Characteristic vectors of bordered matrices with infinite dimensions II". In: *The Collected Works of Eugene Paul Wigner*. Springer, 1993, pp. 541–545.

[43] Jianfeng Yao, Shurong Zheng, and ZD Bai. *Sample covariance matrices and high-dimensional data analysis*. Cambridge University Press Cambridge, 2015.

[44] Anthony Zee. "Law of addition in random matrix theory". In: *Nuclear Physics B* 474.3 (1996), pp. 726–744.

[45] Martin R Zirnbauer. "Another critique of the replica trick". In: *arXiv preprint cond-mat/9903338* (1999).