# On the Stability of Random Matrices 

## Diffusion on Economic Networks and Dynamical Properties of Random Ensembles

## HELLENIC REPUBLIC National and Kapodistrian University of Athens

EST. 1837

## Nikolaos Poulios

Department of Economics
National and Kapodistrian University of Athens

This dissertation is submitted for the degree of
Doctor of Philosophy
© Copyright by Nikolaos Poulios Candidate 2022
All Rights Reserved

I certify that I have read this dissertation and that in my opinion it is fully adequate in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

## (Prof. Ioannis Leventides) Principal Adviser

I certify that I have read this dissertation and that in my opinion it is fully adequate in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
(Prof. Dimitrios Cheliotis)

I certify that I have read this dissertation and that in my opinion it is fully adequate in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
(Ass. Prof. Michael T. Rassias)

## Members

I certify that I have read this dissertation and that in my opinion it is fully adequate in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
(Prof. Apostolos Christopoulos)

I certify that I have read this dissertation and that in my opinion it is fully adequate in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
(Prof. Stylianos Kotsios)

I certify that I have read this dissertation and that in my opinion it is fully adequate in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
(Ass. Prof. Nikolaos Michelakakis)

I certify that I have read this dissertation and that in my opinion it is fully adequate in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
(Ass. Prof. Theodoros Nicoleris)

I would like to dedicate this thesis to my brother Nikolaos Poulios who is watching me from heaven, to my parents for their support and to my beloved Diamanto Fotiou.
'Be a free thinker and don't accept everything you hear as truth. Be critical and evaluate what you believe in.‘

- Aristotle, 384-322 BC, Greek Philosopher


## Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains fewer than 55,000 words including appendices, bibliography, footnotes, tables and equations and has fewer than 30 figures.

Nikolaos Poulios
February 2022

## Acknowledgements

The PhD thesis that you keep in your hands is a three-year quintessence. This time period has been full of tension, stress and very tough effort which is finally completed with great success. This endeavour it could not have happened mainly without the fruitful, important and valuable contribution of my commitment, and secondly without those people who stood by me all those years. I would like to express my deepest appreciation to my supervisor and co-author of three papers Prof. Ioannis Leventides who since my graduate studies till now has been the professor who initiated me into fabulous paths of mathematical science and formed me a necessary and sufficient character in order to stand on the highest academic level. Furthermore, I can not forget the endless hours of discussions not only on the topic of this PhD thesis but also on new ideas to open problems in mathematics. I am deeply indebted to him for his trust in me even from the first year of the PhD studies not only with my participation in education of undergraduate program of our Department as a Teacher Assistant but also with my involvement into examinations per semester. I am convinced that our research collaboration will be continued when this thesis will be defensed.

I would also like to express my deepest gratitude to Ass. Prof. Dimitrios Cheliotis of the Department of Mathematics at the National and Kapodistrian University of Athens as a member of the committee who was very supportive in all stages of this thesis. Moreover, his contribution was more than vital. His very important and valuable suggestions regarding the part of the second research paper related to randomness led this paper to be completed and published in a famous international mathematical journal.

I am extremely grateful to Ass. Prof. Michael Rassias for the very valuable contribution to my PhD studies. It is my great honour that my first research work was published in his book which was later published by one of the best editors in the world. I hope that after my PhD graduation to participate and learn more about number theory which is the queen of Mathematical science.

I would like to extend my sincere thanks to my parents for their endless support and encouragement from the first day of this PhD journey up to now. In addition, when I was young, they taught me to believe in my strengths and to chase incessantly my dreams. I am also grateful to my siblings for their continuous support and encouragement. Many thanks
to my friend and co-author of a paper Costas for the amazing lectures in Set Theory in the Master program of Algorithms, Logic and Discrete Mathematics (ALMA) at the NKUA. Also, many thanks to the PhD team Aleka, Panagiota, Kalliopi (now she is a Dr.) for the beautiful and very interesting conversations we had in the office. I also wish to thank my friend Aggeliki who taught me English and we spent many hours talking about philosophy.

Special thanks to my girlfriend Diamantoula for her continuous support and the passion she gave me not only in the PhD but also during the master studies, where we met at the School of Applied Mathematical and Physical Sciences at the National Technical University of Athens. I will never forget how many hours we were talking about mathematics, computers, algorithms and tried to understand how the human brain works. Last but not least, I would like to thank the secretary staff of the PhD program for their support.


#### Abstract

This thesis considers a multidimensional family of problems in mathematics. Moreover, one of the three topics which we cover here deals with the diffusion process on dynamical networks in finance. Those graphs are connected, directed and weighted networks of banks, (nodes), where their initial capital and the amount of transfer loans from one bank, (node), to another are given. In other words, we try to treat problems in the financial sector where initially we have dynamical banking networks with different initial leverages and we conclude to have a stable banking network thanks to the diffusion process. Moreover, we present enough toy examples which our theory validates and are confirmed by practical examples. Last but not least, we speculate that this work can be extended by introducing tools from mathematical control theory in order to make the network more controllable.

In addition, another problem that we present in this thesis is the notion of almost zeros of polynomial vectors and matrices. Our motivation in this topic was to understand and to present how the notion of almost zero depends of randomness. To make it more clear we study and present the statistical characteristics of the almost zero and its value in a given random ensemble of polynomials. Furthermore, toy examples are presented in order to make our study more clear as well.

Finally, the last problem that we study is the linear, time-invariant multivariable system described by a system of equations. Moreover, we focus on random systems with the parameters having been replaced by random matrices. In addition, we define the concept of $\varepsilon$-controllability of a random system for a given positive number $\varepsilon$. Furthermore, we consider the fundamental Gaussian orthogonal ensemble of random matrices and we calculate that the $\varepsilon$-uncontrollability of such a random system depends on the distribution of the entries. Finally, in this case $\varepsilon$-uncontrollability is calculated.


## Table of contents

Abbreviations ..... xvii
List of figures ..... xix
List of tables ..... xxi
I First part ..... 1
1 Introduction ..... 3
1.1 Motivation and structure of the thesis ..... 3
2 Literature and Historical Review ..... 5
2.1 Mathematical control theory ..... 5
2.2 Random matrix theory ..... 8
2.3 Graph theory ..... 12
2.4 Diffusion ..... 14
II Second part ..... 17
3 Methodological Tools ..... 19
3.1 An introduction to the Laplace Transform method ..... 19
3.2 How to solve an Initial Value Problem ..... 22
3.2.1 Fundamental theory of ODE Linear Systems ..... 25
3.2.2 The role of Diagonalizable Matrices ..... 26
3.2.3 The notion of Matrix Exponential and its Laplace Transform ..... 27
3.2.4 Homogeneous Linear Systems with Constants Coefficients ..... 29
3.2.5 Nonhomogeneous Linear Systems with Constant Coefficients ..... 32
3.3 Basic Mathematical Theory of Linear Systems ..... 34
3.3.1 The Behavior of Input, State, Output Models ..... 35
3.3.2 Controllability, Observability, Stability, Poles and Zeros ..... 35
3.3.3 Reachability and Controllability ..... 36
3.3.4 Observability and Constructibility ..... 41
3.3.5 Fundamental of Stability Theory ..... 44
3.3.6 Poles and Zeros ..... 48
3.4 Random matrix theory ..... 50
3.4.1 What is random matrix theory ..... 51
3.4.2 Importance of random matrix theory ..... 51
3.4.3 Distribution of eigenvalues of random matrices ..... 52
3.4.4 Terminology ..... 54
3.4.5 Gaussian Ensembles ..... 54
3.4.6 Random matrices in Finance ..... 55
3.4.7 Matrix cleaning and RMT ..... 56
3.5 Introduction to Graph Theory ..... 57
III Third part ..... 61
4 Diffusion on Dynamical Interbank Loan Networks ..... 63
4.1 Introduction ..... 63
4.2 Diffusion equations and equilibrium points ..... 65
4.2.1 Connectivity and equilibria ..... 69
4.3 The basic two structural cases by toy-examples - calculations of equilibrium points ..... 72
4.3.1 Case 1: At least one non-zero element per line to the adjacency operator ..... 72
4.3.2 Case 2: Adjacency operator with two lines equals to zero ..... 74
4.4 Differential equation and its solution ..... 76
4.5 Solution by diffusion three structural examples ..... 77
4.5.1 Case 1: Two real negatives eigenvalues and one zero ..... 77
4.5.2 Case 2: Two complex eigenvalues with negative real part and one zero ..... 80
4.5.3 Case 3: Three real negative eigenvalues ..... 83
4.5.4 Solutions to the former three cases without solving differential equa- tions ..... 86
4.6 Case study in a banking network ..... 88
5 The Notion of Almost Zeros and Randomness ..... 93
5.1 Introduction ..... 93
5.2 Almost zeros of polynomial matrix ..... 94
5.3 Example of prime almost zero calculation ..... 97
5.4 Statistical results ..... 99
5.5 Conclusions ..... 102
6 Random Matrices and Controllability of Dynamical Systems ..... 103
6.1 Introduction ..... 103
$6.2 \quad \varepsilon$-uncontrollability of random systems ..... 104
6.3 Random matrix ensemble ..... 105
6.4 The case $n=2$ ..... 106
6.5 The general case ..... 109
6.6 Conclusions ..... 114
7 Conclusions ..... 115
7.1 Epilogue ..... 115
References ..... 119
IV Appendix ..... 123
. 1 Chapter 3 ..... 125
. 2 Chapter 4 ..... 126
. 3 Chapter 5 ..... 129
. 4 Chapter 6 ..... 132

## List of Abbreviations

BIBO Bounded-Input Bounded-Output<br>GDP Gross Domestic Product<br>GCD Greatest Common Divisor<br>GOE Gaussian Orthogonal Ensemble<br>GSE Gaussian Symplectic Ensemble<br>GUE Gaussian Unitary Ensemble<br>IPR Inverse Participation Ratio<br>LHS Left Hand Side<br>ODE Ordinary Differential Equations<br>PCA Principal Component Analysis<br>RHS Right Hand Side<br>RMT Random Matrix Theory

## List of figures

1 Plot of the vector field for the system in Example 4. ..... 125
2 Sort list of leverages in descending order. ..... 126
3 The mean initial leverage $\left(l_{M}\right)$ is equal to $7 / 9$. At the point of equilibrium $l_{i}=l_{j}=l \in \mathbb{R}, i \neq j$ and $l \geq l_{M}$ ..... 127
4 From dark blue to red we are moving from the highest level of leverage per bank to the lowest and furthermore from the heighest level of weights to the lowest, for the nodes and edges, respectively. ..... 128
5 Mean of p-index ${ }^{2}$ vs $k, k=1,2, \ldots, 25$ ..... 129
6 Standard deviation of p-index ${ }^{2}$ vs $k, k=1,2, \ldots, 25$ ..... 129
7 Mean of the size of almost zero vs $k, k=1,2, \ldots, 25$ ..... 129
8 Standard deviation of the size of almost zero vs $k, k=1,2, \ldots, 25$ ..... 130
9 Probability distribution functions of the p-index ${ }^{2}$ for degrees ranging from 1 up to 4 ..... 130
10 Descriptive statistics of polynomials degree 1 ..... 130
11 The distribution of the p-index ${ }^{2}$ of $s_{0}, P_{M}\left(\overline{s_{0}}, s_{0}\right)$ follows Gamma distribution and flattens with respect to the size of $M$. ..... 131
12 The distribution of almost zero $s_{0}$ of $\left\|s_{0}\right\|^{2}$ follows Gamma distribution and sharpens with respect to the size of $M$ ..... 131
13 Fit of various distributions for the size or p -index ${ }^{2}$ for the ensemble of degree 113
14 Probability distribution functions of sum of squares of polynomials degrees 1 up to 4 ..... 132
15 Moving from the lower to the upper value of $\varepsilon$ the probability bound increases with n . ..... 132
16 The growth of probability bound increases quadraticaly with $n$. ..... 133

## List of tables

3.1 Useful Laplace Transform properties. ..... 22
3.2 Basic Laplace Transforms where $n$ is a nonnegative integer and $\mathrm{a}, \mathrm{b}$ are real. ..... 25
3.3 Dayson's classification of Gaussian ensembles. ..... 55
4.1 Initial and final leverages $\left(l_{i}\right)$ per bank. ..... 80
4.2 Initial and final leverages $\left(l_{i}\right)$ per bank. ..... 83
4.3 Initial and final leverages $\left(l_{i}\right)$ per bank. ..... 85
4.4 Final leverages $\left(l_{i}\right)$ in equilibrium, taking account the new $\phi_{i j}$ per bank. ..... 86
4.5 Final leverages $\left(l_{i}\right)$ in equilibrium, taking account the new $\phi_{i j}$ per bank. ..... 87
4.6 Initial leverages $\left(l_{i}\right)$ per bank. ..... 88
4.7 Final leverages $\left(l_{i}\right)$ in equilibrium, taking account the new $\phi_{i j}(t)$ per bank. ..... 92

## Part I

## First part

## Chapter 1

## Introduction

'As you set out for Ithaka hope your road is a long one, full of adventure, full of discovery.' —K. P. Cavafis, 1863-1933, Greek Poet

### 1.1 Motivation and structure of the thesis

The aim of this thesis is multiple. Initially we study the diffusion process on dynamical financial networks. To be more precise, we study the effect of diffusion method to interbank networks in relation to connected, directed and weighted networks. We consider networks of $n$ different banks which exchange funds (loans) and the main feature is how the leverages of banks can be chosen to improve the financial stability of the network. This is done by considering differential equations of diffusion type.

Furthermore, we investigate the problem of almost zeros of polynomial matrices as used in the system theory. It is related to the controllability and observability notion of systems as well as the determination of Macmillan degree and complexity of systems. We also present some new results in this important invariant in the light of randomness and we prove an uncertainty type relation appearing in such ensembles of operators.

In addition, we introduce the concept of $\varepsilon$-uncontrollability for random linear systems, i.e. a linear system in which the usual matrices have been replaced by random matrices. We also estimate the $\varepsilon$-uncontrollability in the case where the matrices come from the Gaussian orthogonal ensemble. Our proof utilizes tools from systems theory, probability theory and convex geometry.

This thesis is divided into three parts: Introduction and literature review material in Part I, methodological tools which were used in this thesis in Part II, and the rest consists of three research papers among which the first one has been published in a collected volume under

Spinger publications and the other two have been published in Institute of Mathematics and its Applications journal (IMA), in Part III.

In Part I and especially in Chapter 1 we present an introduction and we provide information related to the motivation and the structure of this thesis. In addition in Chapter 2 we have a literature review about what other scientists have done so far on topics like mathematical control theory, diffusion process, graphs/networks and random matrix theory.

The methodological tools needed for the derivation of the main research results presented in Part III are given in Part II, Chapter 3. Furthermore, in this part, topics and notions from mathematical control theory, like controllability/observability, are presented. Moreover, we describe not only the meaning of RMT but also the the given examples of what a random matrix is and what makes it so special. Another topic that is covered in this Chapter is graph theory and examples about connected, directed and weighted graphs/networks and algebraic properties of graphs are given. Finally, the topic of diffusion process is presented as well.

Part III constitutes the most important part of this research study which is presented in this thesis. Thus, in Chapter 4, we describe the diffusion process on connected, directed and weighted interbank loan networks. We prove that diffusion can drive the bank network to its steady state where the leverages of all the banks of the network are equal. Diffusion in the bank network is modelled with a system of coupled ordinary linear differential equations. The last argument, we ended up with two different methods to the same results and a great success. In Chapter 5, we present the notion of almost zero of polynomial vectors and matrices in terms of a minimization problem. The purpose of this chapter is twofold: (a) Firstly, to explore further the algebraic properties and computations of almost zeroes of polynomial vectors or matrices and present new results of higher order almost zeros; (b) Secondly, to relate the notion of almost zero to that of randomness i.e. what are the statistical characteristics of the almost zero and its norm polynomial value in a given random ensemble of polynomials. In Chapter 6, we consider random systems where the parameters which are matrices have been replaced with random matrices. Moreover, we define the $\varepsilon$-uncontrollability for random systems and we also describe the GOE. In addition, we give the detailed calculation of the $\varepsilon$-uncontrollability of a random system where the matrix $\mathbf{A}$ comes from the GOE.

## Chapter 2

## Literature and Historical Review

'Imagination is more important than knowledge‘."
— Albert Einstein, 1879-1955, German-born Theoretical Physicist.

### 2.1 Mathematical control theory

Mathematical control theory is the area of mathematical science which deals with the analysis and control of complex systems. The main idea behind control is to act to on an object in order to manipulate its behavior so as to achieve the desired goal. There are two main key points in mathematical control theory that someone should take into account. The first one is that we have such a proper model of the object that we want it to be controlled e.g. a satellite and the goal is the optimization of its behavior, and the second is the existence of uncertainty about the model or the environment under which the object operates. Furthermore, a very important role in the whole procedure of mathematical control is the notion of feedback which gives the user the opportunity to check out the output of the model if it is the desirable goal and if not, to try to make some changes in the input in order to get the behavior of the object.

In 1868 Maxwell in his unique paper [45] brought great attention to engineers and mathematicians by the exhibition of the notion of governors to the theory of dynamics. The author in his work avoiding many details described four different cases which correspond to the motion of a machine under its governors and disturbance. According to him these cases about disturbance term are:

1. More and more the disturbance may be increasing.
2. It may constantly be reduced.
3. An oscillation of continually increasing amplitude.
4. An oscillation of continually decreasing amplitude.

Where the first and third option are signs of unstable motion while the remaining cases show the other side of a coin viz the stability of the motion.

A very special paper about the roots with negative real part of an equation [30] was written by A. Hurwitz in 1895. The author tried and gave answers related to the design of mechanical and electrical systems. His main goal was to determine the location of the roots a polynomial $n$ degree and got the coefficients without solving an equation numerically. Furthermore, Hurwitz used the techniques of Cauchy and Hermite in order to determine criteria which could be applied.

In 1932 H. Nyquist published a superb paper [56] of great importance. He realized the semantic role of the roots and poles of meromorphic function in the stability of linear systems. Furthermore, the aim of his paper was to describe not only the notion of stability of linear feedback or "regeneration" systems but also the meaning of steady-state and transient behavior. He proposed as a dominant result of his research the Nyquist criterion for stability which is a very helpful tool and it is used by control engineers.

Approximately 30 years later (1959) another important paper was published by LaSalle [42]. In his days there was an intuitive assumption that a control system could be moved between two states in the shortest time by consuming all available energy. A name to this hypothesis was given which was called "bang-bang principle". LaSalle has chosen linear time to control systems and a "steering" function in order to control the system. Moreover, in case of a more general control system he has shown that anything that can be done by a proper steering function is possible to be done by using bang-bang steering. In addition, he has clarified that it was not essential that all optimal steering functions are bang-bang and one of the key results of his research work was to establish the bang-bang principle for all control systems where the controlled elements are linear.

One year after LaSalle publication (1960) a very pioneering research work was published by R. Bellman and R. Kabala [9] where the first one is considered as the applied mathematician who introduced dynamic programming in 1953. The aim of their paper was to establish a keystone in the area of mathematical theory about an important class of decision processes which have not been studied in general. These processes have a name and they are called adaptive. Furthermore, these processes focus on and play a significant role in areas like statistics, operations research and stochastic control. In addition, the authors brought to our attention how the theory of dynamic programming could be used in these areas and gave answers to complex and galling questions.

In 1976 Sontag and Rouchaleau[61] in their paper worked on a more general procedure to systems which have algebraic constraints on their state set and the state transitions are given by the polynomial functions of the inputs and state variables. Furthermore, one of the key points of their research was to obtain finiteness conditions due to finite computational procedures which are very important in the area of applied algebra. Moreover, in their work they gave answers to interesting questions like reachability in bounded time and they worked under the problem of recognition of external behavior between two systems by applying finitely many inputs.

In 1979 Sontag in his paper [60] mentioned not only the important meaning of observability but also the special role of it in systems defined by polynomial difference equations. An intuitive meaning of observability was given by him according to which a system is observable if for any two states it is possible to be distinguished by some input/output experiment. According to him observability plays a vital system property role for a few reasons some of which are the following ones:

- the state-variable plays the role of feeding back a function of the state which should get through "observers".
- In stochastic case the beneficial well known solution of the Kalman filter contains with high accuracy an efficient observer construction with parameters being optimized according to the disposable statistical data.
- Observability is one of the best key ideas in realization theory.

Another novel paper which is noteworthy to be mentioned was written by Baillieul [7] in 1980. The author in his work presented a class of controlled dynamical systems which are related to finitely many differential equations having specific polynomial nonlinearities. His main aim was to study in depth such systems in the light of algebraic and differential geometry. Furthermore, one of the most important results which the author presented in his article about the study of nonlinear systems is that they provide alternatives to conventional approximation such as the Lie theory does. In the same year the former author published a very interesting paper [6] about the controllability and observability of polynomial dynamical system. The key idea in his work was the Hilbert basis theorem and elimination theory in order to bring a remarkable attention and give answers like integrability, stability, etc for polynomial systems. In other words, he applied the notions of Hilbert basis theorem to a wide class of systems in order to describe controllability and observability through some finitely conditions.

In 1986 a remarkable article about local feedback stabilization and bifurcation control was presented by E.H. Abed and J-H Fu [3]. The writers of this work demonstrated two main
issues. The first one was the demonstration of the link between local stabilization in critical cases and the problem to (local) bifurcation control. In addition, they pointed out that it is not rare the fact that the results on stabilization in critical cases are possible to be applied in cases of control of bifurcations. The other goal of the authors was to get the results on simple and pure imaginary eigenvalues from a state dynamical matrix of a linearised system. Finally, the approach they followed in their research was to produce more general and valid analytical criteria not only for stabilizability but also stabilizing feedback controls. This can be achieved by using the bifurcation formula which contains Taylor series of the vector field and calculation of eigenvectors. In the field of control theory there are areas which are mature such as

1. algebraic system theory
2. optimal control
3. $H_{\infty}$ control
4. adaptive control etc.
not only theoretic but also applied areas which make mathematical control theory, after almost 50 years, to be a complete field of mechanics and mathematical science.

### 2.2 Random matrix theory

The first time we met random matrices was in 1930 but in that moment, because it was not such a popular method, it was not given the necessary importance until the next decade when the method was studied by many scientists. The big question imposed by scientists at that time was what we can say about the eigenvalues of a very large matrix if we suppose that the elements of a matrix are random variables with known distribution. To give an answer to this question we have to go back in 1950 when physicists studied the motion of neutron in nuclear physics.

According to quantum mechanics, the energy levels of a system can be described by the properties of a Hermitian operator $\mathbf{H}$ which is called Hamiltonian. This Hamiltonian system should have the same structure, regarding to eigenvalues, thus it should be worked as an infinite Hilbert space and according to the $\mathbf{H}$ operator the aim was to collect information about the eigenvalues.

The first scientist who worked in this sector was the Austrian-Hungarian physicist Eugene Wigner who supposed that the statistical behavior of energy levels is identical to the
eigenvalues of a random matrix. Some decades later in 1960 when random matrices were recognised by many famous physicists and through Monte Carlo simulation they concluded that the hypothesis that Wigner first made about random matrices was correct.

It was in 1953 when F. Dyson [17] worked on theoretical physics and his research focused on dynamical behavior of a disorder linear chain system. By a disordered chain he meant that chain which consists of harmonic oscillators in one dimension. The main goal of his paper was to present a method in order to determine the exact spectrum or the distribution function of characteristic frequency of the chain as the size of the masses becomes very large. In his paper Dyson described his analysis and presented an exact solution of models in one dimension clarifying that those results which were produced by him could not say that it could be applied on a corresponding 3 dimensional problem. Furthermore, Dyson gave a more theoretical approach, in other words, from a group theoretical analysis concept and he concluded that an irreducible ensemble of matrices which are invariant under a symmetry group $G$ should belong to one of three classes, which he then named orthogonal, unitary and symplectic.

There was a great necessity to find a new analysis method due to failure of classical methods on prediction of a macroeconomic-model by economists. Ormerod and Mounfield [58] in 2000 tried to describe how the economists failed to forecast a system of macroeconomic model with classical methods and they proposed the approach of RMT. By classical method they meant that the model which was constructed by economists is a combination of economic theory and statistical regression in order to specify the relationships and to place values on the parameters in the relationships. The importance of such models was the short-term forecasting factors for the economy like inflation, unemployment, consumer spending and the overall level of output in the economy. According to them, the results of the forecasting were very helpful because they played a very important role in policy makers like setting interest rates and monetary policy.

In the last 30 years, even though the economists had disagreements about how the economy works, this situation drove them to study the relationships of macroeconomicmodels, no single method was better than any other about the forecasts and their accuracy. Indeed, the forecasting record of macroeconomic-models was very poor and as other scientists studied and they concluded that no evidence exists where the accuracy was improving over the time. Another step to their research was to compare the eigenvalues of the empirical correlation matrix with the range of eigenvalues of the correlation matrix of a purely random matrix. The reason why they did this step to their research was that if the computed eigenvalues are between the theoretical limits of a random matrix they come to the conclusion
that any correlations which exist between the rate of growth of GDP at time $t$ and any lagged values of itself are essential random and have no real significance.

Finally, they concluded that the reasons that the economists got poor forecasting record of GDP growth was due to the inherent characteristics of data and furthermore they said that the forecasting cannot be improved in a better way independently what economic theory or statistical technique used to generate them.

On the other side Plerou, Gropikrishnan and Rosenow [59] on their paper point out the importance of quantification of correlations between the returns of different stocks which is very important due to quantifying the risk of portfolios of stocks, pricing of options and forecasting. They mention that the idea of the problem to identify stocks that are correlated are not far away from a more complex problem like connectivity of gas molecules. On the one hand they calculate the correlation matrix with elements the correlations between any two stocks and on the other hand, they meet a few difficulties which they tried to overcome, such as the following:

- there is no "algorithm" to show us what the "interaction strength" between two firms $i, j$ is.
- correlations should involve clusters of stocks.
- the correlations between any pair of firm $i, j$ of stocks change in time.
- for any stock $i$ they estimate an average correlation from given finite records.

Moreover, the authors wondered for what reasons should someone manipulate with matrices with random elements. To answer that query, we need to look back in time, in 1950 when the famous Houngarian-American theoretical physicist Eugene Wigner and later on the British-American theoretical and mathematical physicist Freeman Dyson and of Indian origin theoretical physicist Madan Lal Mehta worked on nuclear physics and tried to understand the energy levels of complex nuclei when experimental data failed to be described by theoretical models.

The treatment to this problem was done by E. Wigner [62] who had the superb idea to think of the behaviour of nucleus components as random and furthermore, he assumed the elements of the Hamiltonian matrix $\mathbf{H}, H_{i j}$, as mutually independent numbers. According to the last assumption he derived all the statistical properties of the eigenvalues of the Hamiltonian (H) matrix which were very close to the results from the experimental model. In addition, E. Wigner worked on real symmetric matrices whose elements follow Gaussian distribution and he called such a class of matrices as GOE. Moreover, Plerou et al. in order to quantify and understand the correlations between different stocks started by the calculation
of the eigenvalues and then they compared the spacing distribution to the Wigner distribution. The authors came with very interesting results and confirmed that the cross-correlation matrix is perfectly aligned with the predictions of random matrix theory. Furthermore, what came out by the authors from the last results according to the comparison of eigenvalues between the cross-correlated matrix and the RMT, was that the elements of cross-correlated matrix have a considerable degree of randomness. A hypothesis made by authors about the randomness is due to either nonstationary correlations or the worked with the finite time series.

Ormerod [57] in 2008 worked on RMT with application to Macro-Economic Time Series. He studied the application of RMT by analyzing the synchronization of international business cycles from 1880 up to 2008. Furthermore, Ormerod worked onto correlation matrix of annual real GDP growth rates in a discrete set of countries. According to his research, he found that there is a trend towards great synchronization over time. Moreover, he illustrated that the results under RMT technique are not so differently compared to results taken with other, more close to economists, techniques. In addition, what he did was to compare the eigenvalues of the correlation matrix versus the theoretical maximum and minimum values of a random matrix of similar dimension.

Ormerod has investigated through comparison of the eigenspectra properties of the empirical matrix versus the theoretical eigenspectra properties of a random matrix in order to determine the degree to which an empirical correlation matrix is dominated by noise. After that, what he proved, was to identify those eigenstate of the empirical matrix which contain authentic information content and the remaining eigenstates will be noise dominated and thus unstable.

In addition, after the comparison of the eigenvalues distribution of the correlation matrices he found that the empirical formed matrix will not have random elements but there will be a structure in correlation matrix. Finally, for Ormerod so as to study and analyze this structure of eigenvectors, the IPR should have been calculated.

Next we not only have an application of RMT on Finance but also on cleaning correlation matrices through the approach of empirical eigenvalue distribution. Bouchard and Potters, in 2009, relied on the work of Marchenko and Pastur [44] (1967) in order to describe the application of RMT in finance. The main aim of their research as a new statistical tool was the analysis of multidimensional data sets which had a very good impact on almost all scientific research fields like physics, image analysis, mechanics economics, finance, etc. Moreover, the authors thought that they could investigate and interpret the dynamic of N return stocks. Thus, the easiest way to interpret the correlations among those quantities was to calculate the Pearson estimator of the correlation matrix.

Furthermore, they linked the calculation of eigenvalues of a correlation matrix to the risk of investment of a portfolio and they have observed that large eigenvalues correspond to more risky mix of assets in contrast with the lower eigenvalues. In addition, they mention that in financial markets the large eigenvalues correspond to the trend of investing in almost all stocks by the same manner and the last is the so called "market mode" which is correlated to the market index.

In addition, the authors not only present their results about the comparison between the empirical spectrum and RMT but also the comparison of different approximation schemes for the risk control. Thus, they compare the empirical spectrum with the equivalent as presented to the paper of Marchenko-Pastur and they observed that several eigenvalues take values which extend the proposed bound by Marchenko-Pastur. Moreover, they show that eigenvectors with less structure follow Gaussian distribution. Furthermore, they conclude that not only eigenvectors with large values but also those with small values are not random and they mention that such large eigenvalues correspond to economic sectors in contrast with the small values which correspond to long-short portfolios that correspond to investments with low volatility.

### 2.3 Graph theory

In 1956 an innovative paper was published by L. R. Ford [22] where he presented appropriate algorithms for the solution of the maximal network flow problem and its application to a numerous of different type of problems. In his article, he briefly described the definition and the notion of a network and its characteristics, like nodes, arcs and gave some special features to arcs like the length and the capacity, respectively. In addition, in his paper he stated and gave solution to three problems. The first one was on what the maximal steady state flow is in case of sending an item, say a good, from the origin to the terminal, say a rail network, by avoiding to take account the cost of this procedure viz the length of this flow. The second problem was starting from an origin to terminal node which is the cheapest route without considering any capacity constraints. Finally, the last problem which was considered by L. R. Ford was to give a solution on what the maximal amount of items is, say goods, that could be transferred from the initial, origin node to the final in a given time. Moreover, according to him, he stated the Minimal Cut Theorem which says that "In any network the maximal steady state flow is equal to the minimal cut value", where cut value according to the definition of a cut in a network is the sum of capacities of its arcs. The answer the author gave to the three previous problems was to present them as a linear programming and he concluded that the problems could be tackled by the simplex method.

In the same year a very interesting paper about network theory was published by L. R. Ford and D. R. Fulkerson [41]. The authors in their paper tried to give an answer to a connected network between two cities and their intermediates where they assigned a number to each link of the network which construed as the capacity. According to them, such a problem could be set up as a linear programming problem and could be solved by the simplex method but they observed that in the case of the network being planar, then a more efficient and simpler computing procedure should be followed. Furthermore, in their paper they gave a non-constructive proof of the minimal cut theorem and they concluded to the ability of giving an upper bound for flows over an arbitrary network. Finally, they pointed out a network where by giving any two points, they noticed the duality between the capacity problem and that of finding the shortest path.

Three years after the superb research work by L. R. Ford and D. R. Fulkerson we have another great paper by Dijkstra which was published in 1959 [15]. In his paper the author worked on two problems which are related to the connection of graphs. More precisely his initial idea was to consider $n$ points (nodes) which are connected by a branch where the length of each one is given. Furthermore, he focused on the case where between any two nodes there exists at least one path and according to the last one he studied those two problems. Thus the first problem he studied was a given set of $n$ nodes which make a tree with the characteristic of a minimum total length among those $n$ nodes. He argued that his solution to this problem was better than the solution given by J. B. Kruskal [40] and those given by H. Loberman and A. Weinberger [43]. Dijkstra claimed that the solution which was given by the last authors was much more complex than his own approach because his method produced the right answer to the problem taking account only the requirement of simultaneous storing of the data for the majority of $n$ branches, while the other authors not only sorted all the branches according to their length but also they stored simultaneously all data for all branches. Moreover, the second problem that had been studied by Dijkstra within the same article was to find a path of minimum total length between two nodes. In addition, the solution to the second problem which was given by Dijkstra again as in the first problem, was much less complex than the given answer by L. R. Ford [22] as described by C. Berge [10]. The reason was that Dijkstra's approach didn't store all the data for all branches simultaneously but he stored a sufficient number of them which was much less than $n$, where $n$ was the cardinality of nodes.

For more than a hundred years, one of the most popular ideas in graph theory literature is the idea that every map requires at most four colors or to draw a map five colors are required. A solution to this problem was proposed in 1977 by a joint superb work of Kenneth Appel and Wolfgang Haken [5]. Travelling back in time and more specifically in 1852, Francis

Guthrie after his graduation from the University College of London sent a letter to his brother who was a student of the very famous mathematician Augustus De Morgan. Francis was curious about an idea that he had and wanted to share it with his brother. His concept was that in order to draw a map on a sheet of paper, you should do it by filling up those countries in four only different colors in such a way that none of those countries share the same color with their neighbours. Although his brother Frederick was a mathematician student he couldn't prove his brother proposal mathematically and in turn, he asked for help by his professor De Morgan but without a success either. So, Francis Guthrie was the mathematician who posed the four color problem which remained unsolved though for 124 years approximately with many curious problem solvers from high school students to professor mathematicians trying to solve it unsuccessfully. Moreover, in 1976 Kenneth Appel and Wolfgang Haken solved the four-color problem and Guthrie's conjecture was proved in a different mathematical way than the ones former researchers tried all those years without any success. The authors pointed out that they successfully proved this difficult problem thanks to the power of computers and that was the main difference that they used as a method against the other researchers who tried to prove the conjecture through more classical methods of the mathematical science. Also, they mentioned that even though there are some theorems of great interest from a mathematical point of view, they believed that it could only be proved by computer methods. The last is true because according to their results they concluded that despite of their proof of the four-color theorem, there are limits to those researchers who try to treat difficult problems in mathematical science by theoretical methods only.

### 2.4 Diffusion

The word diffusion derived from the Latin word 'diffundere' which means 'to spread out'. If we want to think about it, we should bring in our minds a glass of water and deposit a droplet of ink without stirring it and then we get a demonstration of the process. The how it works part is due to Brownian motion of atoms or molecules that lead to full mixing.

This motion of diffusion process occurs in many states like gases, liquids and more difficult to a solid state. The scientific community had the belief that the diffusion process could occur only in liquids and gases and this idea was considered truthful until the end of the 19th century. Due to the paramount work of William Roberts-Austen and Georg von Hevesy the previous belief was rejected. Furthermore, scientists realized that due to the density of such materials, like solid, a different approach needed to be followed in order to understand the motion of atoms in such materials, i.e. the diffusion process. In other words, scientists
should focus on two key points which are the position and the motion of atoms in solids in order to get a better understanding in a diffusion process.

The first scientific work on diffusion in gases was studied in detail by Thomas Graham in 1828 [24]. In addition, he studied in depth the diffusion in gases from a quantitative point of view. Graham, in his research, realized that when different types of gases came into contact the result was that they diffused through each other and did not separate themselves according to their density, as it was initially believed.

In addition, a very interesting paper was published in 1855 by A. Fick [18] and the main topic of his research was a mathematical accomplishment about salt movement in liquids as a diffusion process, something very closed to heat diffusion. Fick was so genius that he was the scientist who published the first book on medical physics. In addition, in his book, he mentioned the applications of optics, solid mechanics, gas and heat diffusion to biological systems. Moreover, his ideas about molecular diffusion could be described by mathematical formulas in a more difficult way than he had imagined as a concept. Furthermore, his paper was divided into two main parts. The first one was a description of his experimental results after the applicability of Fourier's equation to liquid diffusion and the second one was a full description with every detail of the diffusion through a membrane.

Closer to today and more specifically in 2015 we have a very promising paper by N . Meghanathan. The author in his research work [46] introduced a probabilistic type of diffusion process in random network graphs. In his work the author has as the main idea a network with nodes and links which he called the initial set of node as early adopters and along the links of the network he tried to understand how the information from the starting point of a set of nodes diffuse to the remaining nodes of the network. Furthermore, he gave probabilities to links and diffusion process or in other words, he assigned a probability ( $p_{\text {link }}$ ) if there exists a link between any two nodes and a certain probability $\left(p_{\text {diff }}\right)$ assigned to information where it spreads along the links. Furthermore, according to him, we can think the process of diffusion in rounds where the initial set of nodes, i.e. early adapters, is the first round of this procedure. In addition, those nodes which get the information are said to be covered, and after that, those nodes are candidates which diffuse the information to other nodes which have not been covered yet and so on. This procedure continues until all the remaining uncovered nodes get the information or in other words there is no uncovered node in the network. Thus the main aim of his paper was to analyze the speed of information from one or more nodes, referred to as early adopters, to spread out to other nodes of the network. Finally, the author was interested in studying such random networks from the point of view that if the number of that initial set of nodes increases what the impact on the success of diffusion process is.

## Part II

## Second part

## Chapter 3

## Methodological Tools

'If I have seen further it is by standing on the shoulders of giants. To R. Hook', February 1676.
—Sir Isaac Newton, 1643-1727, English Mathematician.

### 3.1 An introduction to the Laplace Transform method

In this chapter we present some methodological tools that are necessary in order to get a better perception on the chapters that are following. Tools like the Laplace transform, fundamental matrix and how they are involved in order to solve systems of linear ordinary differential equations are given. In addition, we introduce useful tools and basic theory concerning linear systems like controllability, observability, stability, poles and zeros. Last but not least tools from graph theory are presented as well.

Suppose that we have a first order differential equation which is given by the following equation

$$
\begin{equation*}
\frac{d y}{d t}+p(t) y=f(t) \tag{3.1}
\end{equation*}
$$

where $p(t), f(t)$ are given functions. In order to solve Eq.(3.1) we should multiply it by an integrating factor $\mu$ so that the resulting equation is integrable. More precisely we have

$$
\begin{gather*}
\mu(t) \frac{d y}{d t}+\mu(t) p(t) y=\mu(t) f(t)  \tag{3.2}\\
(\mu(t) y)^{\prime}=\mu(t) y^{\prime}+\mu(t)^{\prime} y=\mu(t) \frac{d y}{d t}+\frac{d \mu(t)}{d t} y=\mu(t) f(t) \tag{3.3}
\end{gather*}
$$

Comparing the last two equations we conclude that the next relation must be satisfied

$$
\begin{equation*}
\frac{d \mu(t)}{d t} y=\mu(t) p(t) \tag{3.4}
\end{equation*}
$$

From the last equation we get

$$
\begin{equation*}
\ln \mu(t)=\int p(t) d t+\rho \tag{3.5}
\end{equation*}
$$

By choosing $\rho$, which is an arbitrary constant, as zero we obtain the following equation which is the simplest possible function for $\mu(t)$, and it is called integrating factori.e.

$$
\begin{equation*}
\mu(t)=\exp \int p(t) d t \tag{3.6}
\end{equation*}
$$

Furthermore in order to give a general solution to (3.1) we start from (3.3) and we have

$$
\begin{equation*}
(\mu(t) y)^{\prime}=\mu(t) f(t) \tag{3.7}
\end{equation*}
$$

and we get

$$
y=\frac{1}{\mu(t)} \int_{t_{0}}^{t} \mu(t) f(t) d t+c
$$

where $\mu(t)$ is given by (3.6).
Theorem 3.1.1. [12] If $f$ is piecewise continuous for $t \geq \alpha$, if $|f(t)| \leq g(t)$ when $t \geq M$ for some positive constant $M$, and if $\int_{M}^{\infty} g(t) d t$ converges, then $\int_{\alpha}^{\infty} f(t) d t$ also converges. On the other hand, if $f(t) \geq g(t) \geq 0$ for $t \geq M$, and if $\int_{M}^{\infty} g(t) d t$ diverges, then $\int_{\alpha}^{\infty} f(t) d t$ also diverges.

Unfortunately we can't find an appropriate integrating factor for ordinary linear equations of higher degree than Eq.(3.1). There is another tool that we can use to tackle this problem and solve nth order constant coefficient linear differential equation. The procedure is to multiply by $e^{-s t}$ the following equation

$$
\begin{equation*}
y^{(n)}+a_{n-1} y^{(n-1)}+a_{n-2} y^{(n-2)}+\cdots+a_{1} y^{\prime}+a_{0} y=f(t) \tag{3.9}
\end{equation*}
$$

where $a_{i}, i=0,1, \ldots, n-1$ are constants and then taking the integral from 0 to $\infty$ we get:

$$
\begin{equation*}
\int_{0}^{\infty} e^{-s t}\left(y^{(n)}+a_{n-1} y^{(n-1)}+a_{n-2} y^{(n-2)}+\cdots+a_{1} y^{\prime}+a_{0} y\right) d t=\int_{0}^{\infty} e^{-s t} f(t) d t \tag{3.10}
\end{equation*}
$$

We can rewrite (3.10) as

$$
\begin{equation*}
\mathscr{L}\{f(t)\}=F(s)=\int_{\alpha}^{\beta} L(s, t) f(t) d t \tag{3.11}
\end{equation*}
$$

where $L(s, t)=e^{-s t}$ is a given function and it's called kernel of the transformation and we call this procedure integral transforms. The key idea behind this procedure, is because there is a case that we don't know $f$ thus we use (3.11) in order to make the problem more simple and we work the problem with $F$ and then we go back from $F$ to $f$. This is known as "inverting the transform".

## Example 1: Integrating factor

- Solution Let's find the Laplace transform of the function $f(t)=t e^{a t}$, where $a$ is a real constant.

In order to find the Laplace transform we use the Eq.(3.11):

$$
\begin{aligned}
\mathscr{L}\{f(t)\}=F(s) & =\int_{0}^{\infty} e^{-s t} t e^{a t} d t=\int_{0}^{\infty} t e^{-(s-a) t} d t \\
& =\lim _{x \rightarrow \infty}\left(\left.\frac{t e^{(a-s) t}}{a-s}\right|_{0} ^{x}-\int_{0}^{x} \frac{e^{(a-s) t}}{a-s} d t\right) \\
& =\lim _{x \rightarrow \infty}\left[\frac{t e^{(a-s) t}}{a-s}-\frac{e^{(a-s) t}}{(a-s)^{2}}\right]_{0}^{x} \\
& =\frac{1}{(s-a)^{2}}
\end{aligned}
$$

where $s>a$.
A very useful property of the Laplace transform is that suppose we have two functions $g$ and $h$ and the Laplace transform exist for $s>a_{1}$ and $s>a_{2}$, respectively. Then we say that the Laplace transform is a linear operator and we write

$$
\begin{aligned}
\mathscr{L}\left\{c_{1} g(t)+c_{2} h(t)\right\} & =\int_{0}^{\infty} e^{-s t}\left[c_{1} g(t)+c_{2} h(t)\right] d t \\
& =c_{1} \int_{0}^{\infty} e^{-s t} g(t) d t+c_{2} \int_{0}^{\infty} e^{-s t} h(t) d t
\end{aligned}
$$

Thus, we get the following equation-property

$$
\begin{equation*}
\mathscr{L}\left\{c_{1} g(t)+c_{2} h(t)\right\}=c_{1} \mathscr{L}\{g(t)\}+c_{2} \mathscr{L}\{h(t)\} \tag{3.12}
\end{equation*}
$$

Table 3.1 Useful Laplace Transform properties.

|  | $f(t)=\mathscr{L}^{-1}\{F(s)\}$ | $F(s)=\mathscr{L}\{f(t)\}$ |
| :--- | :---: | :---: |
| 1. | Linearity | $\mathscr{L}\left\{k_{1} f(t)+k_{2} g(t)\right\}=k_{1} \mathscr{L}\{f\}+k_{2} \mathscr{L}\{g\}$ |
| 2. | First derivative principle | $\mathscr{L}\left\{f^{\prime}(t)\right\}(s)=s \mathscr{L}\{f(t)\}-f(0)$ |
| 3. | Second derivative principle | $\mathscr{L}\left\{f^{\prime \prime}(t)\right\}(s)=s^{2} \mathscr{L}\{f(t)\}-s f(0)-f^{\prime}(0)$ |
| 4. | Transform derivative principle | $\mathscr{L}\{-t f(t)\}(s)=\frac{d}{d s} F(s)$ |

### 3.2 How to solve an Initial Value Problem

In this new section we describe the method of Laplace transform in order to show how to solve an initial value problem of linear differential equations with constant coefficients. We mention the following theorem without a proof wich says

Theorem 3.2.1. [12] Suppose that $f$ is continuous and $f^{\prime}$ is piecewise continuous on any interval $0 \leq t \leq A$. Suppose further that there exist constants $K, a, M$ such that $|f(t)| \leq K e^{a t}$ for $t \geq M$. Then $\mathscr{L}\left\{f^{\prime}(t)\right\}$ exists for $s>a$ and moreover

$$
\mathscr{L}\left\{f^{\prime}(t)\right\}=s \mathscr{L}\{f(t)\}-f(0)
$$

Furthermore, from theorem 3.2.1 we have a useful corollary which says the following
Corollary 3.2.1.1. [12] Let $f, f^{\prime}, \ldots, f^{(n-1)}$ are continuous and that $f^{(n)}$ is piecwise continuous on any interval $0 \leq t \leq A$. Suppose that there exist constants $K, a, M$ such that $|f(t)| \leq K e^{a t},\left|f^{\prime}(t)\right| \leq K e^{a t}, \ldots,\left|f^{(n-1)}(t)\right| \leq K e^{a t}$ for $t \geq M$. Then $\mathscr{L}\left\{f^{(n)}(t)\right\}$ exists for $s>a$ and is given by

$$
\mathscr{L}\{f(t)\}=s^{n} \mathscr{L}\{f(t)\}-s^{(n-1)} f(0)-\cdots-s f^{(n-2)}(0)-f^{(n-1)}(0)
$$

Example 2: Use the Laplace transform method to solve the following differential equation. Give both $Y(s)$ and $y(t)$.

$$
\text { 1. } y^{\prime \prime}-3 y^{\prime}+2 y=4, y(0)=2, y^{\prime}(0)=3
$$

- Solution We suppose that there exists a solution $y=\phi(t)$ with its first two derivatives satisfying the conditions of Corollary 3.2.1.1. Taking the Laplace transform of equation we obtain

$$
\begin{equation*}
\mathscr{L}\left\{y^{\prime \prime}\right\}-3 \mathscr{L}\left\{y^{\prime}\right\}+2 \mathscr{L}\{y\}=0 \tag{3.13}
\end{equation*}
$$

What we did up to now is that we used the linearity property of Laplace transform as it was given by Eq.(3.12). Using the Corollary 3.2.1.1 we find that Eq.(3.13) becomes

$$
\begin{align*}
s^{2} \mathscr{L}\{y\}-s y(0)-y^{\prime}(0)-3[s \mathscr{L}\{y\}-y(0)]+2 \mathscr{L}\{y\} & =0  \tag{3.14}\\
\left(s^{2}-3 s+2\right) Y(s)-(s-3) y(0)-y^{\prime}(0) & =0 \tag{3.15}
\end{align*}
$$

where $Y(s)=\mathscr{L}\{y\}$. Next, substituting initial conditions in the last equation and in addition we substitute the Laplacian transform which corresponds to the nonhomogeneous part of initial equation i.e. $f(t)=4$ corresponds to $\mathscr{L}\{f(t)\}=F(s)=\frac{4}{s}$ as mentioned in the Table 3.2 and then solving for $Y(s)$, we obtain

$$
\begin{equation*}
Y(s)=\frac{2 s-3}{(s-1)(s-2)}+\frac{4}{s(s-1)(s-2)} \tag{3.16}
\end{equation*}
$$

Next we find the partial fraction decomposition of Eq.(3.16). More analytically we have

- $\frac{2 s-3}{(s-1)(s-2)}$

$$
\begin{aligned}
& \frac{A_{1}}{s-1}, A_{1}=\left.\frac{2 s-3}{s-2}\right|_{s=1} \Rightarrow A_{1}=1 \\
& \frac{A_{2}}{s-2}, A_{2}=\left.\frac{2 s-3}{s-1}\right|_{s=2} \Rightarrow A_{2}=1
\end{aligned}
$$

Thus, the first part of Eq.(3.16) can be rewritten as

$$
\begin{equation*}
\frac{2 s-3}{(s-1)(s-2)}=\frac{A_{1}}{s-1}+\frac{A_{2}}{s-2}=\frac{1}{s-1}+\frac{1}{s-2} \tag{3.17}
\end{equation*}
$$

We are working in the same way as before in order to decompose the second part of Eq.(3.16) and more specifically we obtain

- $\frac{4}{s(s-1)(s-2)}$

$$
\begin{aligned}
& \frac{A_{1}}{s-0}, A_{1}=\left.\frac{4}{(s-1)(s-2)}\right|_{s=0} \Rightarrow A_{1}=2 \\
& \frac{A_{2}}{s-1}, A_{2}=\left.\frac{4}{(s-0)(s-2)}\right|_{s=1} \Rightarrow A_{2}=-4 \\
& \frac{A_{3}}{s-2}, A_{3}=\left.\frac{4}{(s-0)(s-1)}\right|_{s=2} \Rightarrow A_{3}=2
\end{aligned}
$$

Thus, the second part of Eq.(3.16) can be rewritten as

$$
\begin{equation*}
\frac{4}{s(s-1)(s-2)}=\frac{A_{1}}{s-0}+\frac{A_{2}}{s-1}+\frac{A_{3}}{s-2}=\frac{2}{s}-\frac{4}{s-1}+\frac{2}{s-2} \tag{3.18}
\end{equation*}
$$

Finally, if we substitute the results of Eq.(3.18) and Eq.(3.17) in Eq.(3.16) we obtain

$$
\begin{equation*}
Y(s)=\frac{2}{s}-\frac{3}{s-1}+\frac{3}{s-2} \tag{3.19}
\end{equation*}
$$

Furthermore, consulting the Table 3.2 we obtain

$$
\begin{equation*}
y(t)=\phi(t)=2-3 e^{t}+3^{2 t} \tag{3.20}
\end{equation*}
$$

In this point we can mention various features of Laplace transform method. Some of them are
(i) a linear ODE, can be reduced from a differential to an algebraic equation
(ii) first we solve an algebraic equation and then by inverse Laplacian transform we find the solution in the time domain
(iii) it is not necessary to first solve for the homogeneous part
(iv) the same method can be applied to higher order equations taking account the conditions of Corollary 3.2.1.1.
(v) if $f$ is a continuous function then there is a one-to-one correspondence to Laplace function, viz the Laplace transform corresponds to $f$ is unique.

Table 3.2 Basic Laplace Transforms where $n$ is a nonnegative integer and $\mathrm{a}, \mathrm{b}$ are real.

|  | $f(t)=\mathscr{L}^{-1}\{F(s)\}$ | $F(s)=\mathscr{L}\{f(t)\}$ |
| :---: | :---: | :---: |
| 1. | 1 | $\frac{1}{s}$ |
| 2. | $e^{a t}$ | $\frac{1}{s-a}, s>a$ |
| 3. | $t^{n}$ | $\frac{n!}{s^{n+1}}, s>0$ |
| 4. | $\sin a t$ | $\frac{a}{s^{2}+a^{2}}, s>0$ |
| 5. | $\cos a t$ | $\frac{s}{s^{2}+a^{2}}, s>0$ |
| 6. | $e^{a t} \sin b t$ | $\frac{b}{(s-a)^{2}+b^{2}}, s>a$ |
| 7. | $e^{a t} \cos b t$ | $\frac{s-a}{(s-a)^{2}+b^{2}}, s>a$ |
| 8. | $u_{c}(t)$ | $\frac{e^{-c s}}{s}, s>0$ |
| 9. | $u_{c}(t) f(t-c)$ | $e^{-c s} F(s)$ |
| 10. | $e^{c t} f(t)$ | $F(s-c)$ |

### 3.2.1 Fundamental theory of ODE Linear Systems

In this section we will present basic knowledge on homogeneous systems with constant coefficients. More specifically, we call a system of equations of the form

$$
\begin{align*}
y_{1}^{\prime}(t) & =k_{11}(t) y_{1}(t)+\cdots+k_{1 n}(t) y_{n}(t)+h_{1}(t) \\
y_{2}^{\prime}(t) & =k_{21}(t) y_{1}(t)+\cdots+k_{2 n}(t) y_{n}(t)+h_{2}(t) \\
& \vdots  \tag{3.21}\\
y_{n}^{\prime}(t) & =k_{n 1}(t) y_{1}(t)+\cdots+k_{n n}(t) y_{n}(t)+h_{n}(t)
\end{align*}
$$

a first order linear system of ordinary differential equations or a vector differential equation. From Eq.(3.21) we can extract matrices which are

$$
y(t)=\left[\begin{array}{c}
y_{1}(t) \\
\vdots \\
y_{n}(t)
\end{array}\right], \mathbf{A}(t)=\left[\begin{array}{ccc}
k_{11}(t) & \ldots & k_{1 n} \\
\vdots & \ddots & \vdots \\
k_{n 1}(t) & \ldots & k_{n n}(t)
\end{array}\right] \text {, and } h(t)=\left[\begin{array}{c}
h_{1}(t) \\
\vdots \\
h_{n}(t)
\end{array}\right]
$$

and then Eq.(3.21) can be rewritten in a more compact way as

$$
\begin{equation*}
\mathbf{y}^{\prime}(t)=\mathbf{A}(t) \mathbf{y}(t)+\mathbf{h}(t) \tag{3.22}
\end{equation*}
$$

where the matrix $\mathbf{A}$ contains the constants and such a linear differential system is called constant coefficient. Furthermore, if $\mathbf{h}(t)$ is equal to zero then we say that we have a homogeneous linear differential system otherwise we call it nonhomogeneous. Thus substituting $\mathbf{h}(t)=0$ in Eq.(3.22) we get the homogeneous equation in matrix form for the system which is given by Eq.(3.21). In addition if we assign to Eq.(3.22) an initial value $\mathbf{y}\left(t_{0}\right)=\mathbf{y}_{0}$ then we say that we have an initial value problem.

Theorem 3.2.2. Suppose that we have a linear differential system which is given by Eq.(3.22) then

1. Let $y_{1}$ and $y_{2}$ are solutions to the homogeneous part of linear differential system i.e. for $\mathbf{h}(t)=0$ in Eq.(3.22) then $c_{1} y_{1}+c_{2} y_{2}$ is also a solution where $c_{1}, c_{2}$ are constants.
2. Let $y_{p s}$ is a particular solution to Eq.(3.22) viz $\mathbf{h}(t) \neq 0$ and $y_{h s}$ is solution to homogeneous case then $y_{g s}=y_{p s}+y_{h s}$ is also a solution to Eq.(3.22).

### 3.2.2 The role of Diagonalizable Matrices

When we work with systems of ordinary differential equations we face the difficulty of more than one unknown variables which are involved in some or all of the equations. Thus a good idea to treat this problem is to transform the coefficient matrix $\mathbf{A}$ into a diagonal matrix. Let $\mathbf{A}_{n \times n}$ be a matrix of $n$ linearly independent eigenvectors or in other words the matrix $\mathbf{A}$ has all the eigenvalues different or $\mathbf{A}$ is Hermitian. Moreover, suppose $\left(e^{(1)}, \ldots, e^{(n)}\right)$ and $\left(\lambda_{1}, \lambda_{2}, \ldots, \lambda_{n}\right)$ are the eigenvectors and eigenvalues of matrix $\mathbf{A}$, respectively. Next we form the matrix $\mathbf{P}$ whose columns are the eigenvectors of $\mathbf{A}$, that is

$$
\mathbf{P}=\left[\begin{array}{ccc}
e_{1}^{(1)} & \ldots & e_{1}^{(n)} \\
\vdots & \ddots & \vdots \\
e_{n}^{(1)} & \ldots & e_{n}^{(n)}
\end{array}\right]
$$

The columns of $\mathbf{P}$ are linearly independent vectors that is the matrix $\mathbf{P}$ is non-singular. In addition, if we calculate the product of $\mathbf{A P}$ and take into account that $\mathbf{A} e^{(r)}=\lambda_{r} e^{(r)}$ we can write the following

$$
\mathbf{A P}=\left[\begin{array}{ccc}
\lambda_{1} e_{1}^{(1)} & \ldots & \lambda_{n} e_{1}^{(n)} \\
\vdots & \ddots & \vdots \\
\lambda_{1} e_{n}^{(1)} & \ldots & \lambda_{n} e_{n}^{(n)}
\end{array}\right]=\mathbf{P D}
$$

where

$$
\mathbf{D}=\left[\begin{array}{cccc}
\lambda_{1} & 0 & \ldots & 0 \\
0 & \lambda_{2} & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & \lambda_{n}
\end{array}\right]
$$

is a diagonal matrix whose elements in the main diagonal are the eigenvalues of matrix $\mathbf{A}$. Thus we have the following equation

$$
\begin{equation*}
\mathbf{D}=\mathbf{P}^{-1} \mathbf{A} \mathbf{P} \tag{3.23}
\end{equation*}
$$

We call this procedure a similarity transformation and we say about matrix $\mathbf{A}$ that it is diagonalizable. If a transformation $\mathbf{P}$ applied on a differential equation then the diagonalization means that we get $n$ independent differential equations of one variable which can be solved one by one, separately.

### 3.2.3 The notion of Matrix Exponential and its Laplace Transform

A very powerful tool for solving linear differential equations with constant coefficients of order $n$ is the Laplace transform formula as it was mentioned in the second row of the Table 3.2 is

$$
\begin{equation*}
\mathscr{L}\left\{e^{a t}\right\}=\frac{1}{s-a} \tag{3.24}
\end{equation*}
$$

In this subsection we present a case where the Eq.(3.24) can be used as a superb tool in order to solve a system of differential equations with constant coefficients with the only exception now that we substitute the constant $a$ by a matrix $\mathbf{A}_{n \times n}$. Furthermore, the scalar exponential function $\exp (a t)$ can be represented by the power series

$$
\begin{equation*}
\exp (a t)=1+\sum_{n=1}^{\infty} \frac{a^{n} t^{n}}{n!} \tag{3.25}
\end{equation*}
$$

which converges on $\mathbb{R}$ for all $t$. If we substitute the scalar $a$ by the matrix $\mathbf{A}_{n \times n}$ the former equation can be written as

$$
\begin{equation*}
e^{\mathbf{A} t}=\mathbf{I}+\sum_{n=1}^{\infty} \frac{\mathbf{A}^{n} t^{n}}{n!}=\mathbf{I}+\mathbf{A} t+\frac{\mathbf{A}^{2} t^{2}}{2!}+\frac{\mathbf{A}^{3} t^{3}}{3!}+\cdots+\frac{\mathbf{A}^{n} t^{n}}{n!}+\ldots \tag{3.26}
\end{equation*}
$$

where $\mathbf{I}$ is a $n \times n$ matrix. By differentiating Eq.(3.26) we obtain

$$
\begin{align*}
\frac{d}{d t}\left(e^{\mathbf{A} t}\right) & =\frac{d}{d t}\left(\mathbf{I}+\mathbf{A} t+\frac{\mathbf{A}^{2} t^{2}}{2!}+\frac{\mathbf{A}^{3} t^{3}}{3!}+\cdots+\frac{\mathbf{A}^{n} t^{n}}{n!}+\ldots\right) \\
& =\mathbf{A}+\frac{\mathbf{A}^{2} t}{1!}+\frac{\mathbf{A}^{3} t^{2}}{2!}+\cdots+\frac{\mathbf{A}^{n} t^{n-1}}{(n-1)!}+\cdots \\
& =\mathbf{A}\left(\mathbf{I}+\mathbf{A} t+\frac{\mathbf{A}^{2} t^{2}}{2!}+\frac{\mathbf{A}^{3} t^{3}}{3!}+\cdots+\frac{\mathbf{A}^{n} t^{n}}{n!}+\ldots\right) \\
& =\mathbf{A} e^{\mathbf{A} t} \tag{3.27}
\end{align*}
$$

Now in the following lines we derive an extension of the Eq.(3.24) so that it can be applied in a linear system of differential equations. More precisely, supposing $f(t)=\exp (\mathbf{A} t)$ and $(d / d t) f(t)=\mathbf{A} \exp (\mathbf{A} t)$ and taking into account the first derivative principle as given in the second line of Table 3.1 we obtain

$$
\begin{align*}
L\left\{\frac{d}{d t} e^{\mathbf{A} t}\right\} & =\mathbf{A} \mathscr{L}\left\{e^{\mathbf{A} t}\right\} \\
s \mathscr{L}\{f(t)\}-f(0) & =\mathbf{A} \mathscr{L}\left\{e^{\mathbf{A} t}\right\} \\
-f(0) & =\mathbf{A} \mathscr{L}\left\{e^{\mathbf{A} t}\right\}-s \mathscr{L}\{f(t)\} \\
\mathbf{I} & =(s \mathbf{I}-\mathbf{A}) \mathscr{L}\left\{e^{\mathbf{A} t}\right\} \\
\mathscr{L}\left\{e^{\mathbf{A} t}\right\} & =\frac{\mathbf{I}}{s \mathbf{I}-\mathbf{A}}=(s \mathbf{I}-\mathbf{A})^{-1} \tag{3.28}
\end{align*}
$$

The matrix $(s \mathbf{I}-\mathbf{A})$ is called the characteristic matrix and $(s \mathbf{I}-\mathbf{A})^{-1}$ is called the resolvent matrix of $\mathbf{A}$, respectively. Furthermore, we write $c_{\mathbf{A}}=\operatorname{det}(s \mathbf{I}-\mathbf{A})=s^{n}+a_{1} s^{n-1}+$ $\cdots+a_{n-1} s+a_{n}=\left(s-\lambda_{1}\right)\left(s-\lambda_{2}\right) \ldots\left(s-\lambda_{n}\right)$ and we call it the characteristic polynomial of $\mathbf{A}$ and the roots of $c_{\mathbf{A}}$ are called characteristic values or eigenvalues of $\mathbf{A}$.
Example 3: For the matrix A compute the resolvent matrix $(s I-\mathbf{A})^{-1}$ and the matrix exponential $e^{\mathbf{A} t}$.

$$
\mathbf{A}=\left[\begin{array}{cc}
1 & -1 \\
-2 & 2
\end{array}\right]
$$

- Solution We calculate the characteristic polynomial and we obtain

$$
\begin{align*}
c_{\mathbf{A}}(s) & =\operatorname{det}(s \mathbf{I}-\mathbf{A}) \\
& =s(s-3) \tag{3.29}
\end{align*}
$$

Next we compute the resolvent matrix and we get

$$
\begin{align*}
(s \mathbf{I}-\mathbf{A})^{-1} & =\frac{1}{s(s-3)}\left[\begin{array}{cc}
s-1 & -1 \\
-2 & s-1
\end{array}\right] \\
& =\left[\begin{array}{cc}
\frac{s-2}{s(s-3)} & \frac{-1}{s(s-3)} \\
\frac{-2}{s(s-3)} & \frac{s-1}{s(s-3)}
\end{array}\right] \\
& =\left[\begin{array}{ll}
\frac{2}{3 s}+\frac{1}{3(s-3)} & \frac{1}{3 s}-\frac{1}{3(s-3)} \\
\frac{2}{3 s}-\frac{2}{3(s-3)} & \frac{1}{3 s}+\frac{2}{3(s-3)}
\end{array}\right] \\
& =\frac{1}{s}\left[\begin{array}{cc}
\frac{2}{3} & \frac{1}{3} \\
\frac{2}{3} & \frac{1}{3}
\end{array}\right]+\frac{1}{s-3}\left[\begin{array}{cc}
\frac{1}{3} & \frac{-1}{3} \\
\frac{-2}{3} & \frac{2}{3}
\end{array}\right] \tag{3.30}
\end{align*}
$$

The matrix exponential can be computed by using the Laplace inversion equation that is

$$
\begin{align*}
e^{\mathbf{A} t} & =\mathscr{L}^{-1}\left\{(s \mathbf{I}-\mathbf{A})^{-1}\right\} \\
& =1\left[\begin{array}{cc}
\frac{2}{3} & \frac{1}{3} \\
\frac{2}{3} & \frac{1}{3}
\end{array}\right]+e^{3 t}\left[\begin{array}{cc}
\frac{1}{3} & \frac{-1}{3} \\
\frac{-2}{3} & \frac{2}{3}
\end{array}\right] \\
& =\left[\begin{array}{ll}
\frac{2}{3}+\frac{e^{3 t}}{3} & \frac{1}{3}-\frac{e^{3 t}}{3} \\
\frac{2}{3}-\frac{2 e^{3 t}}{3} & \frac{1}{3}+\frac{2 e^{3 t}}{3}
\end{array}\right] \tag{3.31}
\end{align*}
$$

### 3.2.4 Homogeneous Linear Systems with Constants Coefficients

In this subsection we give a description of the solution of a homogeneous linear system with constant coefficients. Such a system has the form

$$
\begin{equation*}
\mathbf{y}^{\prime}=\mathbf{A} y \tag{3.32}
\end{equation*}
$$

where $\mathbf{A}$ is a constant $n \times n$ matrix, or in other words if we set $\mathbf{h}(t)=0$ in Eq.(3.22) we get the same system. Furthermore, we suppose that $\mathbf{A}$ is different from zero thus $\mathbf{y}=0$ is
the only equilibrium solution. In addition we are interested in finding if there are any other solutions that are very close to this equilibrium or depart from it as $t$ increases. This interest translated in more technical notion that whether $\mathbf{y}=0$ is stable or unstable solution. Systems which are described by Eq.(3.32) we can get a qualitative information by drawing a direction field which shows us the behavior of solutions or we can draw a sample of drajectories for a given system which is called a phase portrait.

Theorem 3.2.3. Let $\mathbf{A}_{n \times n}$ be a constant matrix and $\mathbf{h}(t)$ an $\mathbb{R}^{n}$-valued continuous function defined on an interval I. Suppose $t_{0} \in I$ and $\mathbf{y}(0) \in \mathbb{R}^{n}$. Then the unique solution to the initial value problem

$$
\begin{equation*}
\mathbf{y}^{\prime}(t)=\mathbf{A y}(t)+\mathbf{h}(t), \mathbf{y}\left(t_{0}\right)=\mathbf{y}_{0} \tag{3.33}
\end{equation*}
$$

is the function $\mathbf{y}(t)$ defined for $t \in I$ by

$$
\begin{equation*}
\mathbf{y}(t)=e^{\mathbf{A}\left(t-t_{0}\right)} \mathbf{y}_{0}+\int_{t_{0}}^{t} e^{\mathbf{A}(t-u)} \mathbf{h}(u) d u \tag{3.34}
\end{equation*}
$$

Corollary 3.2.3.1. Let $\mathbf{A}_{n \times n}$ be a constant matrix and $\mathbf{h}(t)$ an $\mathbb{R}^{n}$-valued continuous function defined on an interval I containing the origin. Suppose $\mathbf{y}_{0} \in \mathbb{R}^{n}$. Then the unique solution to

$$
\begin{equation*}
\mathbf{y}^{\prime}(t)=\mathbf{A} \mathbf{y}(t)+\mathbf{h}(t), \mathbf{y}(0)=\mathbf{y}_{0} \tag{3.35}
\end{equation*}
$$

is the function defined for $t \in I$ by

$$
\begin{equation*}
\mathbf{y}(t)=e^{\mathbf{A} t} \mathbf{y}_{0}+e^{\mathbf{A} t} * \mathbf{h}(t) \tag{3.36}
\end{equation*}
$$

Example 4: Find the general solution of the following system of equations, draw a direction field and find the fundamental matrix $\Phi(t)$ of the system.

$$
\mathbf{y}^{\prime}=\left[\begin{array}{ll}
3 & -2 \\
2 & -2
\end{array}\right] \mathbf{y}
$$

Solution Let $\mathbf{A}=\left[\begin{array}{ll}3 & -2 \\ 2 & -2\end{array}\right]$. We assume that $\mathbf{y}=\mathbf{k} e^{\lambda t}$ and substitute for $\mathbf{y}$ in the given system. Then we have a system of algebraic equations

$$
\left[\begin{array}{cc}
3-\lambda_{1} & -2  \tag{3.37}\\
2 & -2-\lambda_{2}
\end{array}\right]\left[\begin{array}{l}
k_{1} \\
k_{2}
\end{array}\right]=\left[\begin{array}{l}
0 \\
0
\end{array}\right]
$$

In order to get a nontrivial solution for the determinant of coefficients in the former equation should be zero that is $\operatorname{det}(\mathbf{A}-\lambda \mathbf{I})=0$. Solving the determinant we obtain the
characteristic polynomial which is equal to

$$
\begin{equation*}
c_{\mathbf{A}}=\lambda^{2}-\lambda-2 \tag{3.38}
\end{equation*}
$$

Eq.(3.38) has the roots $\lambda_{1}=-1, \lambda_{2}=2$ and $k_{1}=(1,2), k_{2}=(2,1)$ are the eigenvalues and eigenvectors, respectively. The corresponding solutions are

$$
\mathbf{y}^{(1)}(t)=\left[\begin{array}{l}
1  \tag{3.39}\\
2
\end{array}\right] e^{-t}, \mathbf{y}^{(2)}(t)=\left[\begin{array}{l}
2 \\
1
\end{array}\right] e^{(2 t)}
$$

We compute the Wronskian determinant of this solution

$$
W\left(\mathbf{y}^{(1)}, \mathbf{y}^{(2)}\right)(t)=\left|\begin{array}{cc}
e^{-t} & 2 e^{2 t}  \tag{3.40}\\
2 e^{-t} & e^{2 t}
\end{array}\right|=-e^{t} \neq 0
$$

Since $W\left(\mathbf{y}^{(1)}, \mathbf{y}^{(2)}\right)(t) \neq 0$ for $t>0$ we conclude that $\mathbf{y}^{(1)}$ and $\mathbf{y}^{(2)}$ form a fundamental set and a general solution of the given initial system of Example 4 is

$$
\mathbf{y}=c_{1}\left[\begin{array}{l}
1  \tag{3.41}\\
2
\end{array}\right] e^{-t}+c_{2}\left[\begin{array}{l}
2 \\
1
\end{array}\right] e^{2 t}
$$

where $c_{1}, c_{2}$ are arbitrary constants.
In order to compute the fundamental matrix $\Phi(t)$ we concentrate in Eq.(3.41) where the columns of $\Phi$ are solutions that satisfy the initial conditions $\mathbf{y}^{(1)}(0)=\left[\begin{array}{l}1 \\ 0\end{array}\right], \mathbf{y}^{(2)}(0)=\left[\begin{array}{l}0 \\ 1\end{array}\right]$. Next we substitute the initial conditions in Eq.(3.41) and solve for $c_{1}, c 2$. Finally, taking into account the first initial condition we obtain $\left(c_{1}, c_{2}\right)=(-1 / 3,2 / 3)$ and the second initial condition we get $\left(c_{1}, c_{2}\right)=(3 / 2,-3 / 4)$. Hence the fundamental matrix is given by

$$
\Phi(t)=\left[\begin{array}{ll}
\frac{-1}{3} e^{-t}+\frac{4}{3} e^{2 t} & \frac{2}{3} e^{-t}-\frac{2}{3} e^{2 t}  \tag{3.42}\\
\frac{-2}{3} e^{-t}+\frac{2}{3} e^{2 t} & \frac{4}{3} e^{-t}-\frac{1}{3} e^{2 t}
\end{array}\right]
$$

The direction field for the system in Figure 1 (see Appendix) shows clearly that all solutions depart from the origin or in other words $\mathbf{y}=\mathbf{0}$ is a saddle point due to the fact that the eigenvalues of the given system are real and have opposite signs.

### 3.2.5 Nonhomogeneous Linear Systems with Constant Coefficients

Example 5: Solve $y^{\prime}=\mathbf{A y}+\mathrm{h}$ for the given matrix $\mathbf{A}$, forcing function $h$ and initial conditions $\mathbf{y}(0)$.

$$
\mathbf{A}=\left[\begin{array}{cc}
5 & 2 \\
-8 & -3
\end{array}\right], \mathbf{h}(t)=\left[\begin{array}{c}
t \\
-2 t
\end{array}\right], \mathbf{y}(0)=\left[\begin{array}{l}
0 \\
1
\end{array}\right]
$$

- Solution We solve this problem with Laplace transform method.

$$
\begin{align*}
\mathscr{L}\left\{\mathbf{y}^{\prime}\right\} & =\mathbf{A} \mathscr{L}\{\mathbf{y}\}+\mathscr{L}\{\mathbf{h}(t)\} \\
s \mathbf{Y}(s)-\mathbf{A Y}(s) & =\mathbf{y}(0)+\mathbf{H}(s) \\
\therefore \mathbf{Y}(s) & =\left((s \mathbf{I}-\mathbf{A})^{-1}\right) \mathbf{y}(0)+\mathbf{H}(s) \tag{3.43}
\end{align*}
$$

By Table 3.2 we obtain

$$
\mathbf{H}(s)=\left[\begin{array}{c}
1 / s^{2} \\
-2 / s^{2}
\end{array}\right]
$$

Moreover we compute $(s \mathbf{I}-\mathbf{A})$ and its determinant

$$
\begin{align*}
(s \mathbf{I}-\mathbf{A}) & =\left[\begin{array}{cc}
s-5 & -2 \\
8 & s+3
\end{array}\right] \\
\operatorname{det}(s \mathbf{I}-\mathbf{A}) & =s^{2}-2 s+1=(s-1)^{2} \\
\therefore(s \mathbf{I}-\mathbf{A})^{-1} & =\left[\begin{array}{cc}
\frac{s+3}{(s-1)^{2}} & \frac{2}{(s-1)^{2}} \\
\frac{-8}{(s-1)^{2}} & \frac{s-5}{(s-1)^{2}}
\end{array}\right] \tag{3.44}
\end{align*}
$$

Next we find the partial fraction decomposition of the matrix above and we obtain

$$
\begin{align*}
(s \mathbf{I}-\mathbf{A})^{-1} & =\left[\begin{array}{cc}
\frac{4}{(s-1)^{2}}+\frac{1}{(s-1)} & \frac{2}{(s-1)^{2}} \\
\frac{-8}{(s-1)^{2}} & \frac{-4}{(s-1)^{2}}+\frac{1}{(s-1)}
\end{array}\right] \\
& =\frac{1}{(s-1)^{2}}\left[\begin{array}{cc}
4 & 2 \\
-8 & -4
\end{array}\right]+\frac{1}{(s-1)}\left[\begin{array}{ll}
1 & 0 \\
0 & 1
\end{array}\right] \\
& =t e^{t}\left[\begin{array}{cc}
4 & 2 \\
-8 & -4
\end{array}\right]+e^{t}\left[\begin{array}{ll}
1 & 0 \\
0 & 1
\end{array}\right] \tag{3.45}
\end{align*}
$$

By Corollary 3.2.3.1 the solution to homogeneous system is given by

- Homogeneous solution

$$
\begin{align*}
\mathbf{y}_{\mathbf{h}}(t)=e^{\mathbf{A} t} \mathbf{y}_{0} & =\left(t e^{t}\left[\begin{array}{cc}
4 & 2 \\
-8 & -4
\end{array}\right]+e^{t}\left[\begin{array}{ll}
1 & 0 \\
0 & 1
\end{array}\right]\right)\left[\begin{array}{l}
0 \\
1
\end{array}\right] \\
& =\left[\begin{array}{c}
2 t e^{t} \\
-4 t e^{t}+e^{t}
\end{array}\right] \tag{3.46}
\end{align*}
$$

- Partial Solution

$$
\begin{align*}
\mathbf{y}_{\mathbf{p}}=e^{\mathbf{A} t} \mathbf{h}(t) & =\left(t e^{t}\left[\begin{array}{cc}
4 & 2 \\
-8 & -4
\end{array}\right]+e^{t}\left[\begin{array}{ll}
1 & 0 \\
0 & 1
\end{array}\right]\right)\left[\begin{array}{c}
t \\
-2 t
\end{array}\right] \\
& =t e^{t}\left[\begin{array}{c}
1 \\
-2
\end{array}\right] \tag{3.47}
\end{align*}
$$

Finally, as Corollary 3.2.3.1 shows, the general solution of the system is given by

- General solution

$$
\mathbf{y}_{\mathbf{g s}}(t)=\mathbf{y}_{\mathbf{h}}+\mathbf{y}_{\mathbf{p}}=\left[\begin{array}{c}
2 t e^{t}+t e^{t}  \tag{3.48}\\
-4 t e^{t}+e^{t}-2 t e^{t}
\end{array}\right]
$$

### 3.3 Basic Mathematical Theory of Linear Systems

Suppose that we have a phenomenon that we want to study and to model. In addition we can give to a phenomenon two special features. More specifically we observe a phenomenon and concentrate on the outcomes that it produces, which is a set $\mathbb{U}$ and we call it the universum. In such a set $\mathbb{U}$ there exists a subset $\mathbb{B}$ which contains all the possible outcomes produced by a phenomenon and we call it behavior of the model.

Definition 3.3.1. We call a mathematical model a pair $(\mathbb{U}, \mathbb{B}), \mathbb{U}$ is the universum set with the elements that contains are called outcomes and B is a subset of $\mathbb{U}$ which we call it the behavior.

Furthermore, when we focus on a mathematical model we try to establish the outcomes and the behavior of the phenomenon. However, in most cases in order to extract more information from our model and get a better conception of the phenomenon and how it works except for the initial declared variables we need some other, auxiliary, variables which are called latent variables. In other words an auxiliary variable of particular importance is the state which is a combination of the past information that is relevant for the future progress of the system.

Definition 3.3.2. A dynamical system $\Sigma$ is defined as a triple

$$
\Sigma=(\mathbb{T}, \mathbb{W}, \mathbb{B})
$$

with $\mathbb{T}$ a subset of $\mathbb{R}$ called the time axis, $\mathbb{W}$ a set called the signal space and $\mathbb{B}$ a subset of $\mathbb{W}^{T}$ called the behavior $\left(\mathbb{W}^{T}\right.$ is a standard mathematical notation for the collection of all maps from $\mathbb{T}$ to $\mathbb{W}$ ).

Furthermore, a well-known, special and important class of latent variables are the state variables. The property of state variables is that they divide the behavior in two parts which are the past and the future. A very popular method to describe linear systems is input/state/output systems. Such systems are presented as

$$
\begin{align*}
\frac{d}{d t} x & =\mathbf{A} x+\mathbf{B} u \\
y & =\mathbf{C} x+\mathbf{D} u \tag{3.49}
\end{align*}
$$

where $u, x, y$ are the input, the state and the output, respectively. $\mathbf{A}_{n \times n}$, is sometimes called the state-feedback matrix and if $\mathbf{A}$ is zero then we can not talk about feedback loops in the realization, $\mathbf{B}_{n \times m}$, is called the input matrix, $\mathbf{C}_{p \times n}$ is called the output matrix and $\mathbf{D}_{p \times m}$,
is called the feedthrough term and the representation given by Eq.(3.49) is a state space representation. Moreover, according to the first line of Eq.(3.49) $x$ is the state and $u$ is an auxiliary variable which drives the evolution of the first one and $y$ is the trajectory of observed system and it is a static equation. Finally, Eq.(3.49) plays a central role in system theory as it forms the backbone in linear control theory.

### 3.3.1 The Behavior of Input, State, Output Models

In the next lines we describe the trajectories of the Eq.(3.49) and how they look like. In order to study such a system we pay attention to the input/state equation which is given by

$$
\begin{equation*}
\frac{d}{d t} x=\mathbf{A} x+\mathbf{B} u \tag{3.50}
\end{equation*}
$$

and the main reason is that due to the presentation of the derivative in the left hand side of Eq.(3.50). We split and focus on behavior of Eq.(3.50) in two main steps

1. By setting $u=0$. Then Eq.(3.50) reduces to the autonomous differential equation viz $\frac{d}{d t} x=\mathbf{A} x$ and can be solved as described in Eq.(3.32)
2. The more general case, i.e. $u \neq 0$.

Definition 3.3.3. The unique function $\Phi: \mathbb{R} \rightarrow \mathbb{R}^{n \times n}$ with the properties

1. $\forall t_{0}, t \in \mathbb{R}: \frac{d}{d t} \boldsymbol{\Phi}\left(t, t_{0}\right)=\mathbf{A} \boldsymbol{\Phi}\left(t, t_{0}\right)$
2. $\forall t_{0} \in \mathbb{R}: \Phi\left(t_{0}, t_{0}\right)=\mathbf{I}$
is called the state transition matrix or transition matrix or fundamental matrix (see Example 4) of the autonomous system, i.e. for $u=0$ in Eq.(3.50).

### 3.3.2 Controllability, Observability, Stability, Poles and Zeros

It is very common in some scientific areas where the main research activity involves dynamical systems to estimate an input that causes the states of a system in finite time. For instance we can think of an aircraft where a pilot is trying to control it and must change its orientation. Another useful property in linear dynamical systems that scientists need to know is to determine the state by taking into account measurements that come from the output of the system. All those legitimate concerns of scientists lead to the concepts not only of state reachability (or controllability-from-the-origin) and controllability (or controllability-to-the-origin) but also of state observability and constructibillity. The last notions are of
particular importance in order for someone to study the relationships between state-space and input-output realizations of a linear system. Furthermore, it is worth mentioning that notions like controllability and observability are extremely important in the case of state feedback controllers and state observers, respectively. On the one hand by state controllability we mean the capability to manipulate the state by changing the inputs in an appropriate manner. In addition the last notion is a necessary and sufficient condition for the eigenvalues of a matrix $\mathbf{A}$ of a given system. On the other hand by state observability we call the capability to define the initial state vector of the system taking into account the knowledge of the input and the corresponding information from outputs over time. In other words suppose that we have a set of measurements and we want to know how well can we determine the initial state of a system. Both reachability and observability play a central role in order for us to understand feedback control of systems.

### 3.3.3 Reachability and Controllability

Suppose we have a time-invariant system then we say that a state $x_{1}$ is reachable if there exists an input that transfer the state $x(t)$ from zero to $x_{1}$ in a finite time $T$. In addition, by reachability we mean that suppose we have a system starting from the origin we can reach $x_{1}$ in some finite time. At this point we have to say that reachability doesn't share any information about the specific time that is needed to reach $x_{1}$ and in addition it doesn't show us the trajectory to be followed. Furthermore, similarly we can say that by controllability we mean that a system is capable of transfering the state from $x_{0}$ to the zero state in finite time. As in the case of reachability again here we don't have any information about time or the trajectory to be followed. It is worth pointing out that reachability always implies controllability but the vice versa is true only when the state transition matrix (or fundamental matrix) $\Phi$ of the system has an inverse. Moreover, in the case of continuous time systems, it is always true but in the case of discrete-time systems the same holds true but with the particularity the matrix $\mathbf{A}$ of the system is nonsingular.
I. Continuous-Time Time-Varying Systems Let the following state equation

$$
\begin{equation*}
\dot{x}=\mathbf{A}(t) x+\mathbf{B}(t) u \tag{3.51}
\end{equation*}
$$

where $\mathbf{A}(t) \in \mathbb{R}^{n \times n}, \mathbf{B}(t) \in \mathbb{R}^{n \times m}$ and $u(t) \in \mathbb{R}^{m}$ are defined and continuous on some open interval $I=(a, b)$ with $a, b \in \mathbb{R}$. The following equation gives the state at time $t$

$$
\begin{equation*}
x\left(t, t_{0}, x_{0}\right) \triangleq x(t)=\Phi\left(t, t_{0}\right) x\left(t_{0}\right)+\int_{t_{0}}^{t} \Phi(t, \tau) \mathbf{B}(\tau) u(\tau) d \tau \tag{3.52}
\end{equation*}
$$

with $\Phi(t, \tau)$ being the state transition or fundamental matrix of the systems and $t_{0}, t \in I$ and $x\left(t_{0}\right)=x_{0}$ being the initial state at initial time. We want to transfer the state from $x_{0}$ to some other value, say $x_{1}$, in a finite time $t_{1}>t_{0}$, hence we can rewrite Eq.(3.52) and obtain

$$
\begin{align*}
x_{1} & =\Phi\left(t_{1}, t_{0}\right) x_{0}+\int_{t_{0}}^{t_{1}} \Phi\left(t_{1}, \tau\right) \mathbf{B}(\tau) u(\tau) d \tau \\
x_{1}-\Phi\left(t_{1}, t_{0}\right) x_{0} & =\int_{t_{0}}^{t_{1}} \Phi\left(t_{1}, \tau\right) \mathbf{B}(\tau) u(\tau) d \tau \tag{3.53}
\end{align*}
$$

We have the following definition
Definition 3.3.4. A state $x_{1}$ is reachable at time $t_{1}$ if for some time $t_{0}<t_{1}$ there exists an input $u_{t}, t \in\left[t_{0}, t_{1}\right]$ that transfers the state $x(t)$ from the origin at $t_{0}$ to $x_{1}$ at time $t_{1}$. Hence by the first relation of Eq.(3.53) there exists an input $u$ such that

$$
\begin{equation*}
x_{1}=\int_{t_{0}}^{t_{1}} \Phi\left(t_{1}, \tau\right) \mathbf{B}(\tau) u(\tau) d \tau \tag{3.54}
\end{equation*}
$$

Definition 3.3.5. A state $x_{0}$ is controllable at time $t_{0}$ iffor some finite $t_{1}>t_{0}$ there exists an input $u(t), t \in\left[t_{0}, t_{1}\right]$ that transfers the state $x(t)$ from $x_{0}$ at $t_{0}$ to the origin at time $t_{1}$

By the second relation and by the last definition we substitute $x\left(t_{1}\right)=0$ into Eq.(3.53) and we obtain

$$
\begin{align*}
-\Phi\left(t_{1}, t_{0}\right) x_{0} & =\int_{t_{0}}^{t_{1}} \Phi\left(t_{1}, \tau\right) \mathbf{B}(\tau) u(\tau) d \tau \\
-x_{0} & =\int_{t_{0}}^{t_{1}} \Phi\left(t_{1}, \tau\right) \mathbf{B}(\tau) u(\tau) d \tau \tag{3.55}
\end{align*}
$$

where we obtained the last equation by multiplying the first relation of Eq.(3.55) by $\Phi^{-1}\left(t_{1}, t_{0}\right)=\Phi\left(t_{0}, t_{1}\right)$ and taking into account the property $\Phi\left(t_{0}, t_{1}\right) \Phi\left(t_{1}, \tau\right)=\Phi\left(t_{0}, \tau\right)$. Suppose that there exists a system as described by Eq.(3.54) then the reachability space $R_{r}$ is given by

$$
\begin{equation*}
L \triangleq L\left(u, t_{0}, t_{1}\right) \triangleq \int_{t_{0}}^{t_{1}} \Phi\left(t_{1}, \tau\right) \mathbf{B}(t) u(\tau) d \tau \tag{3.56}
\end{equation*}
$$

where $u(t), t \in\left[t_{0}, t_{1}\right]$ and $t_{0} \leq t_{1}$. Eq.(3.56) holds true since $x_{1}$ is reachable if there exists a $t_{0}$ and $u$ such $x_{1} \in \mathscr{R}(L)$. In addition if the system (3.51) holds then we can say that the system is reachable if and only if it is controllable. The last holds true only for continuous-time systems and not for discrete-time. For discrete-time systems, if a system is reachable then it is controllable but we can't say with full of confidence the same for the opposite case. To put in other words controllability implies reachability if the state transition matrix $\Phi\left(q, q_{0}\right)$
has full column rank which does not always hold true in discrete-time systems. Hence, under other conditions viz in continuous-time systems the transition matrix $\Phi(t, \tau)$ is always nonsingular. Next we give the Gramian matrix definition
Definition 3.3.6. By $\mathbf{W}_{r}\left(t_{0}, t_{1}\right)$ we mean the $n \times n$ reachability Gramian of the system given by Eq.(3.51) which is given by

$$
\begin{equation*}
\mathbf{W}_{r}\left(t_{0}, t_{1}\right) \triangleq \int_{t_{0}}^{t_{1}} \Phi\left(t_{1}, \tau\right) \mathbf{B}(\tau) \mathbf{B}^{\top}(\tau) \Phi^{\top}\left(t_{1}, \tau\right) d \tau \tag{3.57}
\end{equation*}
$$

with $\Phi(t, \tau)$ is the state transition matrix.
At this point we should mention that $\mathbf{W}_{r}\left(t_{0}, t_{1}\right)$ is symmetric and positive semidefinite i.e. the property $\mathbf{W}_{r}\left(t_{0}, t_{1}\right)=\mathbf{W}_{r}\left(t_{0}, t_{1}\right)^{\top}$ it holds true.
Corollary 3.3.0.1. We say a system which is given by Eq.(3.51) is reachable at $t_{1}$ or the pair $(\mathbf{A}(\mathbf{t}), \mathbf{B}(\mathbf{t}))$ is reachable at $t_{1}$ if and only if there exists finite $t_{0}<t_{1}$ such that

$$
\begin{equation*}
\operatorname{rank} \mathbf{W}_{r}\left(t_{0}, t_{1}\right)=n \tag{3.58}
\end{equation*}
$$

Theorem 3.3.1. Let a system that is given by Eq.(3.51) then there exists an input $u$ that transfers the state of the system from $x_{0}$ at time $t_{0}$ to $x_{1}$ at time $t_{1}>t_{0}$ if and only if

$$
\begin{align*}
& x_{1}-\Phi\left(t_{1}, t_{0}\right) x_{0} \in \mathscr{R}\left(\mathbf{W}_{\mathbf{r}}\left(\mathbf{t}_{\mathbf{0}}, \mathbf{t}_{\mathbf{1}}\right)\right)  \tag{3.59}\\
& u(t)=\mathbf{B}^{\top}(t) \Phi^{\top}\left(t_{1}, t\right) \eta_{1}, \text { where } \eta_{1} \text { is the solution of }  \tag{3.60}\\
& \mathbf{W}_{r}\left(t_{0}, t_{1}\right) \eta_{1}=x_{1}-\Phi\left(t_{1}, t_{0}\right) x_{0} \tag{3.61}
\end{align*}
$$

Example 6: Consider $\dot{x}=\mathbf{A}(t) x+\mathbf{B}(t) u$. Find the state transition matrix of the system where

$$
\mathbf{A}(t)=\left[\begin{array}{cc}
t & -1 \\
1 & t
\end{array}\right]
$$

Solution We start by decomposing the matrix $\mathbf{A}$ as follows

$$
\mathbf{A}(t)=t \underbrace{\left[\begin{array}{ll}
1 & 0  \tag{3.62}\\
0 & 1
\end{array}\right]}_{\mathbf{M}_{\mathbf{1}}}+\underbrace{\left[\begin{array}{cc}
0 & -1 \\
1 & 0
\end{array}\right]}_{\mathbf{M}_{\mathbf{2}}}
$$

with $f_{1}(t)=t$ and $f_{2}(t)=1$. We use the next formula in order to obtain the transition matrix $\Phi$

$$
\begin{equation*}
\Phi(t, \tau)=e^{\mathbf{M}_{\mathbf{1}} \int_{\tau}^{t} \theta d \theta} e^{\mathbf{M}_{2} \int_{\tau}^{t} d \theta} \tag{3.63}
\end{equation*}
$$

Substituting $\mathbf{M}_{\mathbf{1}}, \mathbf{M}_{\mathbf{2}}$ we obtain

$$
\begin{align*}
\Phi(t, \tau) & =e^{\left[\begin{array}{ll}
1 & 0 \\
0 & 1
\end{array}\right]} \frac{\frac{1}{2} \int_{\tau}^{t}\left(\theta^{2}\right)^{\prime} d \theta}{} e^{\left[\begin{array}{cc}
0 & -1 \\
1 & 0
\end{array}\right] \int_{\tau}^{t}(\theta)^{\prime} d \theta} \\
& =e^{\left[\begin{array}{ll}
1 & 0 \\
0 & 1
\end{array}\right] \frac{1}{2}\left(t^{2}-\tau^{2}\right)} e^{\left[\begin{array}{cc}
0 & -1 \\
1 & 0
\end{array}\right](t-\tau)} \\
& =e^{\frac{1}{2}\left[\begin{array}{cc}
t^{2}-\tau^{2} & 0 \\
0 & t^{2}-\tau^{2}
\end{array}\right]} e^{\left[\begin{array}{cc}
0 & -(t-\tau) \\
t-\tau & 0
\end{array}\right]} \tag{3.64}
\end{align*}
$$

We are working with the second term on the RHS of Eq.(3.64) and we rewriting in the following form

$$
\underbrace{\left[\begin{array}{cc}
0 & -1  \tag{3.65}\\
1 & 0
\end{array}\right]}_{\mathrm{T}}(t-\tau)
$$

In the sequel as we have done in the previous examples in this chapter we continue in the same way, i.e. to calculate $(s \mathbf{I}-\mathbf{T})$ and $(s \mathbf{I}-\mathbf{T})^{-1}$, respectively. Hence

$$
\begin{align*}
(s \mathbf{I}-\mathbf{T}) & =\left[\begin{array}{cc}
s & 1 \\
-1 & s
\end{array}\right] \\
\operatorname{det}(s \mathbf{I}-\mathbf{T}) & =s^{2}+1 \\
(s \mathbf{I}-\mathbf{T})^{-1} & =\frac{1}{s^{2}+1}\left[\begin{array}{cc}
s & -1 \\
1 & s
\end{array}\right] \\
e^{\mathbf{T} t}=\mathscr{L}^{-1}\left((s \mathbf{I}-\mathbf{T})^{-1}\right) & =\left[\begin{array}{cc}
\cos [t] & -\sin [t] \\
\sin [t] & \cos [t]
\end{array}\right] \tag{3.66}
\end{align*}
$$

Hence, by Eq.(3.66) and Eq.(3.65) we obtain the final solution

$$
\Phi(t, \tau)=\left[\begin{array}{cc}
e^{\frac{1}{2}\left(t^{2}-\tau^{2}\right)} \cos (t-\tau) & -e^{\frac{1}{2}\left(t^{2}-\tau^{2}\right)} \sin (t-\tau)  \tag{3.67}\\
e^{\frac{1}{2}\left(t^{2}-\tau^{2}\right)} \sin (t-\tau) & e^{\frac{1}{2}\left(t^{2}-\tau^{2}\right)} \cos (t-\tau)
\end{array}\right]
$$

II. Continuous-Time Time-Invariant Systems In this case fortunately we can explicitly determine the transition matrix against the previous case where we couldn't define it as a simple exponential. Hence the transition matrix is given by $\Phi(t, \tau)=e^{\mathbf{A}(t-T)}$. Moreover, the definitions that were mentioned before in the case of time-variant systems are remaining true in this case as well and we should have in our mind that always $t_{0}=0$ and $t_{1}=T$. In addition we write again the reachability Gramian for this case which says

Definition 3.3.7. The $n \times n$ reachability Gramian of the time-invariant system $\dot{x}=\mathbf{A} x+\mathbf{B} u$ is given by

$$
\begin{equation*}
\mathbf{W}_{r}(0, T) \triangleq \int_{0}^{T} e^{(T-\tau) \mathbf{A}} \mathbf{B B}^{\top} e^{(T-\tau) \mathbf{A}^{\top}} d \tau \tag{3.68}
\end{equation*}
$$

Definition 3.3.8. We call the $n \times n m$ controllability (from-the-origin) matrix

$$
\begin{equation*}
\mathscr{L} \triangleq\left[\mathbf{B}, \mathbf{A B}, \ldots, \mathbf{A}^{n-1} \mathbf{B}\right] \tag{3.69}
\end{equation*}
$$

Theorem 3.3.2. The system $\dot{x}=\mathbf{A} x+\mathbf{B u}$ is reachable (controllable-from-the-origin)

1. if and only if

$$
\begin{equation*}
\operatorname{rank} \mathbf{\mathbf { W } _ { r }}(0, T)=n, \text { for some finite } T>0 \tag{3.70}
\end{equation*}
$$

where $\mathbf{W}_{r}(0, T)$ is the reachability Gramian as given by Eq.(3.68) or
2. if and only if the $n$ rows of $e^{\mathbf{A} t} \mathbf{B}$ are linearly independent on $[0, \infty]$ over $\mathbb{C}$ or $(s \mathbf{I}-\mathbf{A})^{-1} \mathbf{B}$ are linearly independent over $\mathbb{C}$
3. if and only if rank $\mathscr{L}=n$, where $\mathscr{L}$ is the controllability matrix as given by Eq.(3.69) or
4. if and only if $\operatorname{rank}\left[s_{i} \mathbf{I}-\mathbf{A}, \mathbf{B}\right]=n$ for all complex numbers $s_{i}$ or the eigenvalues of $\mathbf{A}$.

Example 7: For the system $\dot{x}=\mathbf{A} x+\mathbf{B} u$

1. check for reachability where

$$
\mathbf{A}=\left[\begin{array}{cc}
1 & -1 \\
-2 & 2
\end{array}\right], \mathbf{B}=\left[\begin{array}{l}
0 \\
1
\end{array}\right]
$$

Solution According to the theorem 3.3.2 we compute the rank of $\mathscr{L}=[\mathbf{B}, \mathbf{A B}]$. Thus,

$$
\mathscr{L}=\left[\begin{array}{cc}
0 & -1 \\
1 & 2
\end{array}\right]
$$

where the $\operatorname{rank} \mathscr{L}=2=n$. Hence, the system is reachable.

### 3.3.4 Observability and Constructibility

In this subsection we are briefly describing the notion of observability and constructibility but this time we focus on continuous-time time-invariant systems. Now, we are interested in knowing if it is possible to determine the state by observing the response of a system over a period of time and the system is equipped by some input. By observability we express the efficacy to determine the present state $x\left(t_{0}\right)$ from knowledge taken by the system output $y(t)$ and system inputs $u_{t}, t \geq t_{0}$. On the other hand by constructibility we mean the ability of determining the present state $x\left(t_{0}\right)$ taking into account the knowledge of past system outputs $y(t)$ and system inputs $u(t), t \leq t_{0}$.

Let the following continuous-time time-invariant system of the form

$$
\begin{align*}
& \dot{x}=\mathbf{A} x+\mathbf{B u} \\
& y=\mathbf{C} x+\mathbf{D u} \tag{3.71}
\end{align*}
$$

where $\mathbf{A} \in \mathbb{R}^{n \times n}, \mathbf{B} \in \mathbb{R}^{n \times m}, \mathbf{C} \in \mathbb{R}^{p \times n}, \mathbf{D} \in \mathbb{R}^{p \times m}, \mathbf{u}(t) \in \mathbb{R}^{m}$ is continuous. In addition the output of the system is given by

$$
\begin{equation*}
y(t)=\mathbf{C} e^{\mathbf{A} t} x(0)+\int_{0}^{t} \mathbf{C} \mathbf{e}^{\mathbf{A}(\mathbf{t}-\tau)} \mathbf{B u}(\tau) \mathbf{d} \tau+\mathbf{D u}(\mathbf{t}) \tag{3.72}
\end{equation*}
$$

In this type of systems it holds that in the present case $\Phi(t, \tau)=\Phi(t-\tau, 0)=e^{\mathbf{A}(t-\tau)}$ with initial time being $t_{0}=0$. Moreover, we can deduce state $x$ from $\mathbf{u}$ and $y$. That is, we can compute the derivative of $y$

$$
\begin{align*}
& y=\mathbf{C} x+\mathbf{D} u \\
& \dot{y}=\mathbf{C} \dot{x}+\mathbf{D} \dot{u}=\mathbf{C}(\mathbf{A} x+\mathbf{B u})+\mathbf{D} \dot{u}=\mathbf{C A} x+\mathbf{C B u}+\mathbf{D} \dot{\mathbf{u}} \\
& \ddot{y}=\mathbf{C A}^{2} x+\mathbf{C A B u}+(\mathbf{C B}+\mathbf{D}) \dot{\mathbf{u}} \\
& \vdots \tag{3.73}
\end{align*}
$$

Thus we can write the last system in matrix form as

$$
\underbrace{\left[\begin{array}{c}
y  \tag{3.74}\\
\dot{y} \\
\ddot{y} \\
\vdots \\
y^{(n-1)}
\end{array}\right]}_{\mathbf{y}}=\underbrace{\left[\begin{array}{c}
\mathbf{C} \\
\mathbf{C A} \\
\mathbf{C A}^{2} \\
\vdots \\
\mathbf{C A}^{(n-1)}
\end{array}\right]}_{\mathscr{O}(\mathbf{C}, \mathbf{A})} x+\underbrace{\left[\begin{array}{ccccc}
D & 0 & \ldots & 0 & 0 \\
\mathbf{C B} & \mathbf{D} & \ldots & 0 & 0 \\
\mathbf{C A B} & \mathbf{C B}+\mathbf{D} & 0 & \ldots & 0 \\
\vdots & & & & \\
\mathbf{C A}^{(n-2)} \mathbf{B} & \mathbf{C A}^{(n-3)} \mathbf{B} & \ldots & \mathbf{C B} & \mathbf{D}
\end{array}\right]}_{\mathbf{T}}+\underbrace{\left[\begin{array}{c}
\mathbf{u} \\
\mathbf{u} \\
\ddot{\mathbf{u}} \\
\vdots \\
\mathbf{u}^{(n-1)}
\end{array}\right]}_{\mathbf{u}}
$$

where $\mathbf{T}$ is a Toeplitz matrix and if $\mathscr{O}(\mathbf{C}, \mathbf{A})$ is invertible, then we obtain

$$
x(0)=\mathscr{O}^{-1}\left[\begin{array}{c}
y(0)  \tag{3.75}\\
\dot{y}(0) \\
\ddot{y}(0) \\
\vdots \\
y^{(n-1)}
\end{array}\right]-\mathbf{T}\left[\begin{array}{c}
\mathbf{u}(0) \\
\dot{\mathbf{u}}(0) \\
\ddot{u}(0) \\
\vdots \\
\mathbf{u}^{(n-1)}
\end{array}\right]
$$

and the observability matrix $p n \times n$ is given by

$$
\mathscr{O}=\left[\begin{array}{c}
\mathbf{C}  \tag{3.76}\\
\mathbf{C A} \\
\mathbf{C A}^{2} \\
\vdots \\
\mathbf{C A}^{n-1}
\end{array}\right]
$$

Finally we say that $\{\mathbf{C}, \mathbf{A}\}$ is an observable pair for the system given by Eq.(3.71) if $\mathscr{O}$ is nonsingular, viz is invertible.

Definition 3.3.9. We call a state $x$ as unobservable if the zero-input response of the system given by Eq.(3.71) is zero for every $t \geq 0$

$$
\begin{equation*}
\mathbf{C} e^{\mathbf{A} t} x=0, \forall t \geq 0 \tag{3.77}
\end{equation*}
$$

Definition 3.3.10. Suppose we have a system given by Eq.(3.71), then the observability Gramian is the $n \times n$ matrix

$$
\begin{equation*}
W_{0}(0, T) \triangleq \int_{0}^{T} e^{\mathbf{A}^{\top} \tau} \mathbf{C}^{\top} \mathbf{C} e^{\mathbf{A} \tau} d \tau \tag{3.78}
\end{equation*}
$$

which is nonsingular for any $T>0$.

Corollary 3.3.2.1. We say that a system like (3.71) is observable or the pair $(\mathbf{A}, \mathbf{C})$ is observable if and only if

$$
\begin{align*}
\operatorname{rank} \mathscr{O} & =n, \text { or }  \tag{3.79}\\
\operatorname{rank} \mathbf{W}_{0}(0, T) & =n \tag{3.80}
\end{align*}
$$

for all $T>0$. If the system is observable then the state at $x_{0}$ at $t=0$ is given by

$$
\begin{equation*}
x_{0}=\mathbf{W}_{0}^{-1}(0, T)\left[\int_{0}^{T} e^{\mathbf{A}^{\top} \tau} \mathbf{C}^{\top} \mathbf{C} e^{\mathbf{A} t} d \tau\right] \tag{3.81}
\end{equation*}
$$

Theorem 3.3.3. Suppose a system as given by Eq.(3.71). Then the system is observable

1. if and only if

$$
\begin{equation*}
\operatorname{rank} \mathbf{\mathbf { W } _ { 0 }}(0, T)=n, \text { for some finite } T>0 \tag{3.82}
\end{equation*}
$$

where $\mathbf{W}_{0}(0, T)$ is the observability Gramian as given by Eq.(3.78) or
2. if and only if the $n$ columns of $\mathbf{C} e^{\mathbf{A} t}$ are linearly independent on $[0, \infty]$ over $\mathbb{C}$ or $\mathbf{C}(s \mathbf{I}-\mathbf{A})^{-1}$ are linearly independent over $\mathbb{C}$
3. if and only if rank $\mathscr{O}=n$, where $\mathscr{O}$ is the observability matrix as given by Eq.(3.76) or
4. if and only if rank $\left[\begin{array}{c}s_{i} \mathbf{I}-\mathbf{A} \\ \mathbf{C}\end{array}\right]=n$, for all complex numbers $s_{i}$ or the eigenvalues of $\mathbf{A}$.

Example 8: Let a system of the form $\dot{x}=\mathbf{A} x, y=\mathbf{C} x$. Check for the observability where

$$
\mathbf{A}=\left[\begin{array}{ll}
1 & 0 \\
0 & 0
\end{array}\right], \mathbf{C}=\left[\begin{array}{ll}
1 & 0
\end{array}\right]
$$

Solution We can check the observability of the system taking into account the third option of the theorem 3.3.3. More specifically $\mathbf{C A}=\left[\begin{array}{ll}0 & 1\end{array}\right]$. Hence the observability matrix is given by

$$
\mathscr{O}=\left[\begin{array}{c}
\mathbf{C}  \tag{3.83}\\
\mathbf{C A}
\end{array}\right]=\left[\begin{array}{ll}
1 & 0 \\
1 & 0
\end{array}\right]
$$

The $\operatorname{rank} \mathscr{O}=1<2$. Thus, we conclude that the above system is not observable.

### 3.3.5 Fundamental of Stability Theory

A very important part of linear systems study is to investigate the stability. We give the notion of external or bounded-input, bounded-output stability for linear time-invariant systems and in addition we describe what we call internal stability. Furthermore, we briefly present a few stability criteria like

1. algebraic criteria
(a) Routh criterion
(b) Hurwitz criterion
2. Lyapunov criterion
I. External Stability We call a causal system that is externally stable if a bounded input $u(t)<K_{1},-\infty<-T \leq t<\infty$ gives a bounded output $y(t)<K_{2},-T \leq t<\infty$. In addition, for BIBO stability a necessary and sufficient condition is

$$
\begin{equation*}
\int_{0}^{\infty}\|h(t)\| d t<K<\infty \tag{3.84}
\end{equation*}
$$

where $h(t)$ is the impulse response.
II. Internal Stability Suppose that we have a realization

$$
\begin{align*}
\dot{x}(t) & =\mathbf{A} x(t)+\mathbf{B u}(t) \\
y(t) & =\mathbf{C} x(t) \tag{3.85}
\end{align*}
$$

it is internal stable or stable in the sense of Lyapunov if the solution of

$$
\begin{equation*}
\dot{x}(t)=\mathbf{A} x(t), x\left(t_{0}\right)=x_{0}, t \geq t_{0} \tag{3.86}
\end{equation*}
$$

goes to zero as $t \rightarrow \infty$ for any arbitrary $x_{0}$. In other words internal stability has the requirement that all the roots $\lambda_{i}$ of the characteristic polynomial which is produced by $\mathbf{A}$ must be negative if $\lambda_{i} \in \mathbb{R}$ or if $\lambda_{i} \in \mathbb{C}$ then $\mathfrak{R}\left(\lambda_{i}\right)<0$.

## Algebraic Crieteria

1. Routh criterion

Some of the very well known and very useful algebraic criteria for stability are the Routh and Hurwitz criteria, respectively. Their main feature is that they can define the stability of a system by calculating the roots of a characteristic polynomial. Hence, suppose we have a polynomial which is given by

$$
\begin{equation*}
p(s)=a_{n} s^{n}+a_{n-1} s^{n-1}+\cdots+a_{1} s+a_{0} \tag{3.87}
\end{equation*}
$$

where $a_{0}, a_{1}, \ldots, a_{n} \in \mathbb{R}$, constant coeficients.
Theorem 3.3.4. The polynomial given by Eq.(3.87) has on or more roots on the right complex plane when at least one coefficient is zero or (and) the coefficients have different signs.

## Routh Matrix

$$
\begin{array}{c|cccc}
s^{n} & a_{n} & a_{n-2} & a_{n-4} & \ldots \\
s^{n-1} & a_{n-1} & a_{n-3} & a_{n-5} & \ldots \\
s^{n-2} & b_{1} & b_{2} & b_{3} & \ldots \\
s^{n-3} & c_{1} & c_{2} & c_{3} & \ldots \\
\vdots & \vdots & \vdots & \vdots & \vdots
\end{array}
$$

where $a_{n}, a_{n-1}, \ldots, a_{0}$ are the polynomial coefficients while $b_{1}, b_{2}, \ldots, c_{1}, c_{2}, \ldots$ can be computed by

$$
\left.\begin{align*}
& b_{1}= \frac{-\left|\begin{array}{cc}
a_{n} & a_{n-2} \\
a_{n-1} & a_{n-3}
\end{array}\right|}{a_{n-1}}, b_{2}= \\
& c_{1}=\frac{-\left|\begin{array}{cc}
a_{n} & a_{n-4} \\
a_{n-1} & a_{n-5}
\end{array}\right|}{a_{n-1}}, \ldots  \tag{3.88}\\
& b_{1}-\frac{a_{n-1}}{a_{n-3}} b_{1} \\
& b_{2}
\end{align*} \right\rvert\,, c_{2}=\frac{-\left|\begin{array}{cc}
a_{n-1} & a_{n-5} \\
b_{1} & b_{3}
\end{array}\right|}{b_{1}}, \cdots .
$$

Theorem 3.3.5. Let $\lambda_{i}$ are the roots of the characteristic polynomial $p(s)$. The necessary and sufficient conditions in order to holds true that $\Re\left(\lambda_{1}<0\right)$ is that the first
column of Routh array must have the same sign. In other way we say that the system is unstable and the number of roots that have positive real part is equal to the number of sign changes.

Example 9: Let the polynomial characteristic of a system is given by $p(s)=$ $s^{3}+10 s^{2}+11 s+6$. Is it stable or unstable?

Solution We write down the Routh array as

| $s^{3}$ | 1 | 11 |
| :---: | :---: | :---: |
| $s^{2}$ | 10 | 6 |
| $s^{1}$ | $52 / 5$ | 0 |
| $s^{0}$ | 6 | 0 |

where $b_{1}=\frac{52}{5}, b_{2}=0, c_{1}=6, c_{2}=0$ are calculated by (3.88). We observe that the first column of the Routh array doesn't change sign hence the system is stable.

## 2. Hurwitz Criterion

Hurwitz criterion defines if the characteristic polynomial has the roots on the right complex plane or on the imaginary axis. Comparable with the Routh criterion, this one has a drawback, that is the amount of such roots can not be determined. The criterion is based on the Hurwitz determinants which are

$$
\begin{align*}
D_{o} & =a_{n} \\
D_{1} & =a_{n-1} \\
D_{2} & =\left|\begin{array}{ccc}
a_{n-1} & a_{n-3} \\
a_{n} & a_{n-2}
\end{array}\right| \\
D_{3} & =\left|\begin{array}{ccc}
a_{n-1} & a_{n-3} & a_{n-5} \\
a_{n} & a_{n-2} & a_{n-4} \\
0 & a_{n-1} & a_{n-3}
\end{array}\right| \\
& \vdots \\
D_{n} & =\left|\begin{array}{cccccc}
a_{n-1} & a_{n-3} & \ldots & a_{0}, \text { if } \mathrm{n} \text { is odd, } a_{1}, \text { if } \mathrm{n} \text { is even } & 0 & \ldots \\
a_{n} & a_{n-2} & \ldots & a_{1}, \text { if } \mathrm{n} \text { is odd }, a_{0}, \text { if } \mathrm{n} \text { is even } & 0 & \ldots \\
0 & a_{n-1} & a_{n-3} & \ldots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \ldots & 0 & a_{n}
\end{array}\right| \tag{3.89}
\end{align*}
$$

Theorem 3.3.6. The necessary and sufficient conditions in order $\mathfrak{R}\left(\lambda_{i}\right)<0$ holds true are $D_{i}>0, i=0,1, \ldots, n$, where $\lambda_{i}, i=0,1, \ldots, n$ are the eigenvalues of the characteristic polynomial.

Example 10: Let the polynomial characteristic of a system is given by $p(s)=$ $s^{3}+10 s^{2}+11 s+6$. Is it stable or unstable?

Solution According to Hurwitz criterion we compute the determinants as given by (3.89). More specifically

$$
\begin{align*}
& D_{o}=a_{n}=1 \\
& D_{1}=a_{n-1}=10 \\
& D_{2}=\left|\begin{array}{cc}
a_{n-1}=10 & a_{n-3}=6 \\
a_{n}=1 & a_{n-2}=11
\end{array}\right|=104 \\
& D_{3}=\left|\begin{array}{ccc}
a_{n-1}=10 & a_{n-3}=6 & a_{n-5}=0 \\
a_{n}=1 & a_{n-2}=11 & a_{n-4}=0 \\
0 & a_{n-1}=10 & a_{n-3}=6
\end{array}\right|=564 \tag{3.90}
\end{align*}
$$

As all the determinants $D_{i}, i=0,1,2,3$ are positive hence by Hurwitz criterion we have that the system is stable.

## The Lyapunov Criterion for Stability

Theorem 3.3.7 (Lyapunov). A matrix $\mathbf{A}$ is a stability matrix, i.e. $\mathfrak{R}\left(\lambda_{i}\right)<0$ for all eigenvalues $\lambda_{i}$ of $\mathbf{A}$, if and only iffor any given positive-definite symmetric matrix $\boldsymbol{Q}$ there exists a positive-definite (symmetric) matrix $\mathbf{P}$ that satisfies

$$
\begin{equation*}
\mathbf{A}^{\prime} \mathbf{P}+\mathbf{P A}=-\mathbf{Q} \tag{3.92}
\end{equation*}
$$

Corollary 3.3.7.1. If $\mathbf{A}$ is a stability matrix, then the Lyapunov equation

$$
\begin{equation*}
\mathbf{A}^{\prime} \mathbf{P}+\mathbf{P A}+\mathbf{Q}=\mathbf{0} \tag{3.93}
\end{equation*}
$$

has a unique solution for every $\mathbf{Q}$.

### 3.3.6 Poles and Zeros

The concept of poles and zeros of a time-invariant system is very closed to the eigenvalues of $\mathbf{A}$. Quite enough we call the eigenvalues of $\mathbf{A}$ as poles of the system $\{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\}$. First we briefly describe the Smith and Smith-McMillan forms.

Smith and Smith-McMillan forms Let first introduce the Smith and the Smith-McMillan form of polynomial $\mathbf{P}(s)$ and rational matrix $\mathbf{H}(s)$, respectively. We define the Smith form of $p \times m$ polynomial matrix as

$$
\mathbf{S}_{p}(s)=\left[\begin{array}{cc}
\Pi(s) & 0  \tag{3.94}\\
0 & 0
\end{array}\right]
$$

where $\Pi(s)=\operatorname{diag}\left[k_{1}(s), \ldots, k_{z}(s)\right]$, where $z=\operatorname{rank} \mathbf{P}(s)$ and $k_{i}(s)$ are the invariant factors of $\mathbf{P}(s)$. In addition $k_{i}(s)$ is given by $k_{i}(s)=G_{i}(s) / G_{i-1}(s), i=1, \ldots, z$, where $G_{i}(s)$ is the monic greatest divisor of all the nonzero ith-order minors of $\mathbf{P}(s)$ with $G_{0}(s)=1$ and $G_{i}(s)$ are the determinant divisors of $\mathbf{P}(s)$. Furthermore, every $k_{i+1}(s)$ can be divided by $k_{i}, i=1,2, \ldots, z-1$.

Now lets turn to the Smith-McMillan form and suppose that there exists a $p \times m$ rational matrix $\mathbf{H}(s)$. We write

$$
\begin{equation*}
\mathbf{H}(s)=\frac{\mathbf{N}(s)}{d(s)} \tag{3.95}
\end{equation*}
$$

where $\mathbf{N}(\mathrm{s})$ is a polynomial matrix. Let $\left.\mathbf{S}_{\mathbf{N}}(s)=\operatorname{diag}\left[n_{1}(s), n_{2}(s), \ldots, n_{z}(s), 0_{p-z, m-z}\right)\right]$. In order to obtain Smith-McMillan form we have to divide each $n_{i}(s)$ by $d(s)$ and exclude all those common factors of a transfer function matrix $\mathbf{H}(s)$, hence we obtain

$$
S M_{\mathbf{H}}(s)=\left[\begin{array}{cc}
\bar{\Pi}(s) & 0  \tag{3.96}\\
0 & 0
\end{array}\right]
$$

where $\bar{\Pi}(s)=\operatorname{diag}\left[k_{1}(s) / \psi_{1}(s), k_{2}(s) / \psi_{2}(s), \ldots, k_{z}(s) / \psi_{z}(s)\right]$ and $z=\operatorname{rank} H(s)$. As in the Smith form the same holds true here that is every $k_{i+1}(s)$ can be divided by $k_{i}, i=1,2, \ldots, z-$ 1 and $\psi_{i}(s)$ divided by $\psi_{i+1}$.

## Poles

Definition 3.3.11. Let $p_{\mathbf{H}}(s)$ be the characteristic polynomial or pole polynomial of a given $p \times m$ rational matrix $\mathbf{H}(s)$. We call poles of $\mathbf{H}(s)$ the roots of the pole polynomial $p_{\mathbf{H}(s)}$.

By $m_{\mathbf{H}(s)}$ we call the minimal polynomial of all nonzero first order minors of $\mathbf{H}(s)$. At this point it is important to highlight that the minimal polynomial of $\mathbf{H}(s)$, i.e. $m_{\mathbf{H}(s)}$ divides the pole or characteristic polynomial $p_{\mathbf{H}(s)}$ and if the roots of pole polynomial are distinct then it holds that $m_{\mathbf{H}(s)}=p_{\mathbf{H}(s)}$. It should be noted that, we observe the poles in the transfer function $\mathbf{H}(s)$ only to those eigenvalues of $\mathbf{A}$ that are controllable and observable.

Zeros Suppose that we have a scalar transfer function $\mathbf{H}(s)$ then the roots of its numerator polynomial are the zeros of $\mathbf{H}(s)$. To put it another way some or all of the characteristic values of the system matrix $\mathbf{P}(s)$, are called the zeros of the system. Suppose $\operatorname{rank} \mathbf{P}(s)=r$.

Then computing all those $r$ th order nonzero minors of the system matrix $\mathbf{P}(s)$ the zero polynomial of the system $z \mathbf{p}(s)$ is the monic greatest common divisor of all these minors.

Definition 3.3.12. We call the zeros of the system $\{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\}$ or the system zeros the roots of the zero polynomial of the system $z_{\mathbf{P}(s)}$. Moreover, by invariant zeros we call the roots of the invariant polynomials $\mathbf{P}(s)$.

More analytically suppose that we have a system matrix $\mathbf{P}(s)_{(p+n) \times(m+n)}$ and its Smith form is given by

$$
S_{\mathbf{P}}=\left[\begin{array}{cc}
\Pi(s) & 0  \tag{3.97}\\
0 & 0
\end{array}\right] \Pi(s)=\operatorname{diag}\left[k_{1}(s), k_{2}(s), \ldots, k_{r}(s)\right]
$$

Then the invariant zero polynomial of the system $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ is given by

$$
\begin{equation*}
z_{\mathbf{P}(s)}^{I}=k_{1}(s), k_{2}(s), \ldots, k_{r}(s) \tag{3.98}
\end{equation*}
$$

and its roots are the invariant zeros of the system. It should be mentioned that it holds true the fact that $z_{\mathbf{P}(s)}^{I} \subset z_{\mathbf{P}(s)}$. Moreover there exists a relation between the zeros of the system and the zeros of $\mathbf{H}(s)$ which is given by

$$
\mathbf{P}(s)=\left[\begin{array}{lll}
s \mathbf{I} & -\mathbf{A} & \mathbf{B}  \tag{3.99}\\
& -\mathbf{C} & \mathbf{D}
\end{array}\right]=\left[\begin{array}{lll}
s \mathbf{I} & -\mathbf{A} & 0 \\
& -\mathbf{C} & \mathbf{I}
\end{array}\right]\left[\begin{array}{cc}
\mathbf{I} & (s \mathbf{I}-\mathbf{A})^{-1} \mathbf{B} \\
0 & \mathbf{H}(s)
\end{array}\right]
$$

where $\mathbf{P}(s)$ is a square and nonsingular matrix.

### 3.4 Random matrix theory

Nowadays in societies people exchange a very big amount of goods and services under different type of markets. All those exchanges are numbers which are translated to data and can arise to useful and rational decisions by public or private authorities. The difficulty of this type of data is that they have a stochastic structure and high complexity. Those two reasons are enough to point out that not only economic but also finance science are difficult to make predictions about an economy or the capitals in general, respectively.

The answer and understanding of randomness and complexity can be given only by the mathematical science. It is common knowledge for many experts and scientists that if there is no evolution in mathematical science then there is no possibility for any other science to make a progress.

The branch of mathematics that studies the randomness is called stochastic analysis. One of the mathematical tools that is very useful, valuable, efficient and very promising to study such problems mentioned above is the RMT.

### 3.4.1 What is random matrix theory

In random matrix theory the keypoint is the study of matrices whose entries are random variables. Thus, probability theory is the backbone in this subject. What follows is a few examples of RMT.

One of those cases is random operator. There are several cases where someone has been interested in spectral properties of an operator, like Hamilton or Dirac or more generally in a matrix describing correlation of some kind but unfortunately such type of operators are too complicated and thus are not always possible to be analyzed explicitly. On the other hand, RMT approaches the former operators by finite-dimensional matrices whose elements are random variables. By doing this we are driven not only to a more straightforward solution but also we keep some of its important stimulus. Furthermore, another point of view is that the entries of random matrix itself can be considered as data. In this case in multivariate statistics, a random matrix represents a sample covariant matrix. In this case the study of eigenvalues and eigenvectors of an arbitrary size becomes important.

Another aspect where we can view RMT and its applications as counting device which is a very powerful tool in order to perform enumerative combinatorics either the matrices are of finite or infinite size. For instance the GUE generates Catalan numbers by simply taking moments of the semi-circular density.

### 3.4.2 Importance of random matrix theory

There are some reasons why RMT is so important to mathematicians, physicists and other scientists. Such a few reasons are flexibility, universality and predictivity.

Flexibility means that it gives us the ability to build in extra global (anti-unitary) symmetries, such as time reversal, spin, chiral or particle-hole symmetry while maintaining its exact solvability for all correlation functions of eigenvalues.

In addition by universality the precise choice of probability distribution is not so important because the correlation functions depend on this choice only through a few universal paramaeters.

Moreover predictivity means that the correlation functions typically depend on few paramaters and RMT is not able to know the physical scale of the problem in question. The last can be answered by being compared to the effective theory.

### 3.4.3 Distribution of eigenvalues of random matrices

The first study about the distribution of random matrices go back to the work of I. M. Lifsic and F. Dyson in 1947 and 1953, respectively. In his paper Dyson presented a method on how to determine accurately the spectrum or distribution function of the characteristic frequencies of a chain N masses which are all connected by elastic spring obeying the Hooke's law as the number of masses N becomes very large.

One of the aspects of his methods was the case of a disordered chain i.e. an infinite chain whose elements are distributed in a random way according to some known probability law. Furthermore, he considered three types to rely on the property of time-reversal invariance. Those types of random matrices ensembles have elements that are complex, real or self-dual quaternion. Moreover, he introduced two types of disordered chain and he called them type I and Type II with the first one being more simple from mathematical point of view where the second type he worked in a real chain of randomly arranged atoms. At the Type I, each of the parameters is an independent random variable with probability distribution function $G(\lambda)$ and at the Type II he assigned to each mass $m_{j}$ an independent random variable with distribution function $G(m)$, the spring-constants $K_{j}$ being fixed and equal. Additionaly Dayson said that it could be possible to go beyond the two Types and suppose a more general type of chain that the masses $m_{j}$ are independent random variables while the spring power $K_{j}$ is a known function of the two adjoining masses $m_{j}$ and $m_{j+1}$. Under this consideration it would be possible to have a model for a chain composed by different kinds of atoms characterizing by randomness and the only thing that matters in strength of the bond between two atoms is the chemical substance of those atoms. Thus, the systems that he considered were models of one dimensional glass, a disordered array of atoms in one dimension. He pointed out that the results of his paper couldn't be generalized in 3-dimensions but it was very important that a 1-dimensional problem could be solved.

In 1950 RMT gained attention due to the work of E. Wegner in mathematical physics. E. Wigner in 1957 had labored to study the distribution of the latent roots (characteristic values) of certain sets of real symmetric matrices of very high dimensionality. More specifically Wigner wished to describe a general concept of properties about the properties of nucleui as measured in nuclear reaction. He tried to describe a nuclear system with high complexity by an unknown Hermitian operator $\mathbf{H}$ in an infinite-dimensional Hilbert space governed by physical laws.

On the other hand even if the Hermetian operator $\mathbf{H}$ was known, due to high complexity it was not possible to solve eigenequation $\mathbf{H v}=\lambda \mathbf{v}$ (the so called Schrödinger equation of the physical system, where $\mathbf{v}$ and $\lambda$ are the eigenvectors and eigenvalues of $\mathbf{H}$ ). Furthermore, Wigner claimed as an alternative to Hamiltonian $\mathbf{H}$ to choose a large random matrix that is a
member of large class (ensemble) of Hamiltonians but all they have approximately the same properties in general. In such a case the energy levels of the physical system will be able to be approximated by the eigenvaleus of a large random matrix.

In addition Marchenko and Pastur in 1967 studied the distribution of eigenvalues for two sets of random Hermitian matrices and one set of random unitary matrices, where unitary is a complex square matrix if its conjugate transpose $\mathbf{U}^{*}$ is also its inverse, that is, it satisfies $\mathbf{U}^{*} \mathbf{U}=\mathbf{U U}^{*}=\mathbf{I}$, where $\mathbf{I}$ is the identity matrix. He proposed a unitary operator $H_{N}$ and a selfadjoint operator $\mathbf{B}_{N}(n)$ which is given by the next formula:

$$
\begin{equation*}
\mathbf{B}_{N}(n)=\mathbf{A}_{N}+\sum_{i=1}^{n} \tau_{i} q^{(i)}\left(\cdot, q^{(i)}\right) \tag{3.100}
\end{equation*}
$$

where $\mathbf{A}_{N}$ is a nonrandom selfadjoint operator, n is nonrandom number, $\tau_{i}$ is i.i.d. random variable $\in \mathbb{R}$ and $q^{i}$ are mutually independent random vectors in $\mathbf{H}_{N}$. He also supposed that the operator $\mathbf{B}_{N}(n)$ is a sum of not only a nonrandom operator but also of a number of independent random 1-dimensional operators. $\tau_{i}, q^{(i)}$ give a realization of the random operator $\mathbf{B}_{N}$. They are interested in the function $v\left(\lambda ; \mathbf{B}_{N}(n)\right)$ which is showing the ratio of those eigenvalues of $\mathbf{B}_{N}(n)$ which are lying to the left of $\lambda$.

They found that for fixed $\lambda$ the function $v\left(\lambda ; \mathbf{B}_{N}(n)\right)$ is giving a random quantity determined by the random numbers $\tau_{i}$ and random vectors $q^{(i)}$. Moreover, the calculation of such probability of this random quantity is one of the most groundwork problems in the spectral analysis of random operators.

In the recent past there had been existed a general interest in RMT. One of the best and most important research areas in RMT was its connection to quantum chaos (Bohigas, Giannoni and Schmit, 1984) which lead to a random matrix theory of quantum transport (Beenakker, 1997). Then RMT has become a major tool in many scientific areas like number theory and combinatorics, wireless communications (Tulino and Verdu, 2004), in multivariate statistical analysis and PCA (Johnstone, 2001) and economics/finance (Potters, Bouchard, Laloux, 2005) and biology (Wolte, Kohane, Butte, 2005).

In addition the recent work on random matrices has lead to a conjecture which is called univerality conjecture. The conjecture states that if we have large random matrices then its eigenvalues have limits that are independent of the probability distribution of the matrix ensembles.

Yet, this conjecture is not true in general and this is the reason why the researchers have focused on showing that the conjecture is true under specific families of probability distributions.

### 3.4.4 Terminology

We call ensemble of random matrices a set of random matrices assigned with a probability density p and show us the probability of realization of any member belonging to the set.

Wigner and Dyson were interested in how to approximate Hermitian matrix $\mathbf{H}$ by an ensemble of finite but very large dimension Hermitian matrices $\mathbf{H}_{n}$ where the probability density function is given by:

$$
\begin{equation*}
p\left(\mathbf{H}_{n}\right) \propto e^{-\beta \operatorname{tr}\left[V\left(\mathbf{H}_{n}\right)\right]} \tag{3.101}
\end{equation*}
$$

where V is some function or in other words a finite polynomial function, for instance: $V\left(\mathbf{H}_{n}\right)=x_{1} \mathbf{H}_{n}^{2}+x_{2} \mathbf{H}_{n}+x_{3}$ where $x_{i}, i=\{1,2,3\}, \in \mathbb{R}$ and $x_{1}>0$. Furthermore, the elements of $\mathbf{H}_{n}$ belonging to real or complex or real-quaternion and values of $\beta$ are equal to $\beta=1, \beta=$ $2, \beta=4$, respectively. In addition and a special case that we can get is that if $V\left(\mathbf{H}_{n}\right) \propto \mathbf{H}_{n}^{2}$ then Eq.(4.2) reduces and we have a Gaussian ensemble.

### 3.4.5 Gaussian Ensembles

As it was mentioned above F. Dyson introduced three types of Gaussian ensembles i.e Gaussian orthogonal ensemble (GOE), Gaussian unitary ensemble (GUE) and Gaussian symplectic ensemble (GSE).

## 1. Gaussian Orthogonal Ensemble (GOE)

Suppose $\mathbf{P}$ is a square $n \times n$ matrix with each elements being iid as $\mathrm{N}(0,1)$. Moreover, a symmetric Hermitian square matrix $n \times n$ is given by: $\mathbf{H}_{n}=\left(\mathbf{P}+\mathbf{P}^{\mathbf{T}}\right) / \sqrt{2}$ with elements $\left[H_{i j}\right] \sim N(0,1)$ and the off-diagonal being iid as $\mathrm{N}(0,1 / 2)$.

## 2. Gaussian Unitary Ensemble (GUE)

Suppose $\mathbf{P}$ is a asquare $n \times n$ matrix with each element being complex iii as $N^{\mathbb{C}}(0,1)$. Furthermore, $\mathbf{H}_{n}$ is a Hermitian matrix which is given by: $\mathbf{H}_{n}=\left(\mathbf{P}+\mathbf{P}^{*}\right) / 2$ where $\mathbf{P}^{*}$ is the Hermitian transpose of the complex matrix $\mathbf{P}$. In addition the diagonal elements of Hermitian matrix are distributed by iid $\mathrm{N}(0,1)$ and the off-diagonal elements are distributed as $N_{2}(0,1 / 2)$.

## 3. Gaussian Symplectic Ensemble (GSE)

Let $\mathbf{P}$ is a $n \times n$ matrix whose elements one by one are real - quaternian and distributed as iid $N^{Q}(0,1)$ and a self - dual $(n \times n)$ matrix is given by $\mathbf{H}_{n}=\left(\mathbf{P}+\mathbf{P}^{D}\right) / 2$ where $\mathbf{P}^{D}$ describes the dual transpose of the quaternian matrix $\mathbf{P}$. Moreover, the diagonal elements of $\mathbf{H}_{n}$ are distributed as iid $N_{4}(0,1 / 2)$.

Furthermore, F. Dyson and as presented by Table 3.3 proved that GOE is more suitable for systems that are time-reversal invariant in contradiction with GUE which is not. GSE are appropriate for those systems that are time-reversal invariant but do not have spin-rotational symmetry. Moreover, Porter and Rosenzweig in 1960 proved that if all the above ensembles are time-reversal invariant and all the entries of $\mathbf{H}_{n}$ are statistically independent then the probability of $\mathbf{H}_{n}$ is given by:

$$
\begin{equation*}
V\left(\mathbf{H}_{n}\right)=x_{1} \mathbf{H}^{2}+x_{2} \mathbf{H}_{n}+x_{3} \tag{3.102}
\end{equation*}
$$

where $x_{1}, x_{2}, x_{3} \in \mathbb{R}$ and $x_{1}>0$
Table 3.3 Dayson's classification of Gaussian ensembles.

| $\beta$ | Ensemble | TRI | SRS | $H_{i j}$ | $\mathbf{U}$ |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | GOE | Yes | Yes | real | orthogonal |
| 2 | GUE | No | NA | complex | unitary |
| 4 | GSE | Yes | No | real-quaternion | symplectic |

The three types of random matrices as given by Dyson above have some special characteristics. One of them and the most important is the repulsion, in other words if we have any two correlated eigenvalues from GOE or GUE or GSE, they are unlikely to be closer together, i.e. the probability that adjacent eigenvalues are close together is small and the probability goes to zero more quickly, as a power of the distance between them.

### 3.4.6 Random matrices in Finance

Let a portfolio of N assets with a weight $w_{i}$ where i index is the ith asset. According to Potters, Bouchard and Laloux (2005) the variance of the portfolio return is given by:

$$
\begin{equation*}
R^{2}=\sum_{i j} w_{i} \sigma_{i} C_{i j} \sigma_{j} w_{j} \tag{3.103}
\end{equation*}
$$

where $\sigma_{i}^{2}$ is the daily variance of asset i and $C_{i j}$ is the correlation matrix. Suppose that $g_{i}$ is the expected gains then the expected gain of portfolio is given by $G=\sum w_{i} g_{i}$. The measure and optimization of this portfolio can be done by estimating the correlation matrix $C_{i j}$. But the last is very difficult as Bouchard, Laloux, Cizeau et al have shown due to the determination of order $N^{2} / 2$ coefficients out of N time series of length T , where T is very close to N . An accurate determination of true correlation matrix is given by $q=N / T$, where
$q \ll 1$. Furthermore the empirical variance of each stock is given by:

$$
\begin{equation*}
\sigma_{i}^{2}=\frac{1}{T} \sum_{t}^{T}\left(r_{t}^{i}\right)^{2} \tag{3.104}
\end{equation*}
$$

They define an empirical correlation matrix as:

$$
\begin{equation*}
E_{i j}=\frac{1}{T} \sum_{t}^{T} x_{t}^{i} x_{t}^{j} \tag{3.105}
\end{equation*}
$$

where $x_{t}^{i} \equiv r_{t}^{i} / \sigma_{i}$ If $T<N$ then E has rank $T<N$ and $\mathrm{N}-\mathrm{T}$ zero eigenvalues. Moreover they assumed a "true" correlation matrix $\mathbf{C}$ from which past and future $x_{t}^{i}$ are drawn. Thus the risk of a portfolio is given by:

$$
\begin{equation*}
<R_{E}^{2}>=\frac{1}{T} \sum_{i j t} w_{i} \sigma_{i}<x_{t}^{i} x_{t}^{j}>w_{j} \sigma_{j} \approx \sum_{i j} w_{i} \sigma_{i} \mathbf{C}_{i j} \sigma_{j} w_{j} \tag{3.106}
\end{equation*}
$$

In addition Bouchard et al mentioned that due to the fact that the portfolio is not constructed by $\mathbf{E}$, this estimate is unbiased and the relative mean square-error is small approximately to 1 / T.

### 3.4.7 Matrix cleaning and RMT

Bouchard et al in 2005 thought that how could someone "clean" the empirical correlation matrix in order to avoid any biased estimation of future risk? In order to answer this question they had to refer to to the Markowitz solution of the eigenvalues $\lambda_{k}$ and eigenvectors $V_{i}^{k}$ of the correlation matrix which is given by:

$$
\begin{equation*}
w_{i} \propto \sum_{k j} \lambda_{k}^{-1} V_{i}^{k} V_{j}^{k} g_{j} \equiv g_{i}+\sum_{k j}\left(\lambda_{k}^{-1}-1\right) V_{i}^{k} V_{j}^{k} g_{j} \tag{3.107}
\end{equation*}
$$

By correlation term meaning that if the weights of eigenvectors with $\lambda>1$ are suppressed and if the weights of eigenvectors has $\lambda<1$ are enhanced. The optimal solution by Markowitz is to allocate a very large weight to small eigenvalues but this kind of solution may contain noise which leads to be unstable. To avoid this situation they proposed to go for the naive solution i.e.

$$
\begin{equation*}
w_{i} \propto g_{i}-\sum_{k \leq k^{*} ; j} V_{i}^{k} V_{j}^{k} g_{j} \tag{3.108}
\end{equation*}
$$

this result leads that the portfolio is market neutral.

Moreover one way to clean all eigenvalues and eigenvectors is the shrinkage estimator which is a Bayesian method that assigns a prior empirical matrix to identity matrix. Analytically:

$$
\begin{equation*}
\mathbf{E}_{c}=a \mathbf{E}+(1-a) \mathbf{1} \tag{3.109}
\end{equation*}
$$

so $\lambda_{c}^{k}=1+a\left(\lambda^{k}-1\right)$ where index c is the cleaned objects. Another way to eigenvalue cleaning was introduced again by Laloux, Cizeau, Bouchaud, Potters in 1999. What they did was to replace all low lying eigenvalues with a unique value and keep the rest of high eigenvalues corresponding to economical information:

$$
\begin{equation*}
\lambda_{c}^{k}=1-\delta \quad \text { if } k>k^{*}, \quad \lambda_{c}^{k}=\lambda_{E}^{k} \quad \text { if } k \leq k^{*} \tag{3.110}
\end{equation*}
$$

where $k^{*}$ is the important number of sectors. Furthermore, they introduced a way to choose $k^{*}$ and they proposed to use RMT in order to determine the randomness part of the eigenvalue distribution.

### 3.5 Introduction to Graph Theory

Fundamentals A basic introduction to graph theory is presented in this subsection. We call graph an order pair $G=(V, E)$, where $V(G)=\left\{v_{1}, v_{2}, \ldots, v_{n}\right\}$ it is a finite and a non empty set of veritices or nodes and $E(G)=\left\{e_{1}, e_{2}, \ldots, e_{n}\right\}$ is a finite and non-empty set of edges. For the empty set which is the trivial $G(\emptyset, \emptyset)$ graph we write just $\emptyset$. Moreover some graphs are finite but some others are infinite depending on their order, but in this thesis we focus only to finite graphs. Each element of $E$ contains two elements of $V$ that is $e_{i}=\left\{v_{i}, v_{j}\right\} \in E, i \neq j, i, j=1, \ldots, n$. In addition we define the cardinality of those sets as $n(G)=|V(G)|$ and $m(G)=|E(G)|$, correspondingly. Moreover, we call a vertex $v_{i}$ as incident with an edge $e$ if $v_{i} \in e$, then $e$ is an edge at $v_{i}$. Let $G$ be a graph; we say that $v_{i}, v_{j}, i \neq j$ are adjacent or neighbours, if $v_{i} v_{j}$ is an edge of $G$ and denoted by $N_{G}\left(V_{i}\right)$. On top of that we say two edges $e_{i}, e_{j}, i \neq j$ are adjacent if they have an end point in common. Additionally we call a graph $G$ complete if all the vertices are pairwise adjacent. Hence, we define $K^{n}$ a complete graph on $n$ vertices and for instance if $n=3$ we call it triangle.

Definition 3.5.1. Let $G=(V, E)$ and $G^{\prime}=\left(V^{\prime}, E^{\prime}\right)$ be two graphs. We call $G$ and $G^{\prime}$ isomorphic and write $G \simeq G^{\prime}$ if there exists a bijection $\phi: V \rightarrow V^{\prime}$ with $v_{i} v_{j}, \in E \Leftrightarrow \phi\left(v_{i}\right) \phi\left(v_{j}\right) \in E^{\prime}$ for all $v_{i}, v_{j} \in V$. Such a map $\phi$ is called an isomorphism; if $G=G^{\prime}$, it is called an automorphism.

Between two graphs $G$ and $G^{\prime}$ we say that they are disjoint if $G \cap G^{\prime}=\emptyset$, where $G \cap G^{\prime}:=$ $\left(V \cap V^{\prime}, E \cap E^{\prime}\right)$. On top of that if $V^{\prime} \subseteq V$ and $E^{\prime} \subseteq E$ then $G^{\prime}$ is a subgraph of G, or $G$ is a supergraph of $G^{\prime}$, viz $G^{\prime} \subseteq G$. A graph $G^{\prime}$ is a spanning subgraph of $G$, i.e. $G^{\prime} \subseteq G$ if $V^{\prime}$ spans all of $G$ that is if $V^{\prime}=V$. A very important meaning of graph theory is the notion of the degree of a graph $G$. Let $G$ be a graph, we call the degree (or valency) of a vertex $v$ the cardinality of edges at $v$ which is equal to the number of neighbours of $v$ and denoted by $d_{G}(v)=d(v)$. On top of that if a vertex is isoleted then the degree of that vertex is equal to zero. Furthermore, by $\delta(G):=\min \{d(v) \mid v \in V\}$ is the minimum degree of $G$ and by $\Delta(G):=\max \{d(v) \mid v \in V\}$ we denote the maximum degree, as well. In case that all the vertices of graph $G$ have the same degree $k$ then we call it k-regular; for instance a 3 - regular graph is called cubic.

## Theorem 3.5.1. Let $G$ be a graph. Then, the following holds

$$
\begin{align*}
\sum_{v \in V(G)} \operatorname{deg}_{G}(v) & =2|E|=2 m(G) \\
d(G) & :=\frac{1}{|V|} \sum_{v \in V} d(v) \\
\delta(G) & \leq d(G) \leq \Delta(G) \\
\varepsilon(G) & =\frac{d(G)}{2} \\
\varepsilon(G) & =\frac{|E|}{|V|} \tag{3.111}
\end{align*}
$$

where $d(G)$ and $\varepsilon(G)$ are the average degree of $G$ and the number of edges of $G$ per vertex, correspondingly. The last equation can be obtained by substituting the first and the second line of Eq.(3.111) into the third equation, respectively. Let $G(V, E)$ a non-empty graph, then a path is a finite or infinite sequence of distinct vertices $V=\left(v_{0}, v_{1}, \ldots, v_{k}\right)$ where the vertices $v_{0}, v_{k}$ are called end points and $v_{1}, v_{2}, \ldots, v_{k-1}$ are called inner vertices. Moreover, we call the length of a graph $G$ the total number of edges of a path and denoted by $p^{k}$. In addition we call distance $d_{G}\left(v_{i}, v_{j}\right)$ of graph $G$ the length of a shortest $v_{i} v_{j}$ path in $G$ and if we take the greatest distance between two vertices in $G$ we obtain the diameter of $G$. A very important characteristic of a graph is the notion of connectivity. Let $G$ be a non-empty graph, we say that $G$ is connected if any two of its vertices are linked by a path in $G$. In addition to undirected graphs we have another category as well those of directed. We call a graph $G$ as directed graph if the set of vertices are connected by directed edges
often called arcs. Between any two vertices we say a vertex $v_{i}$ as initial vertex and $v_{j}$ the terminal vertex, if an arc starts from $v_{i}$ and ends at $v_{j}$.

Adjacency and Incidence Matrix Linear algebra plays an important role in graph theory. Let $G(V, E)$ be a graph. Then we can extract the adjacency and incident matrix and we can obtain very important features of the graph.

Definition 3.5.2. The adjacency matrix $\mathbf{A}_{n \times n}$ of $G(V, E)$ is defined by

$$
\mathbf{A} \triangleq \begin{cases}1 & \text { if } v_{i} v_{j} \in E \\ 0 & \text { otherwise }\end{cases}
$$

Remark 3.5.1. Every adjacency matrix is symmetric, that is $\left(a_{i j}\right)=\left(a_{j i}\right)$, for al $i, j$. The adjacency of a simple graph has 0 s on the main diagonal. On top of that, the degree of $v_{i}$ is the row sum of adjacency matrix $\mathbf{A}_{\text {adj }}$.

Definition 3.5.3 (Undirected graph). The incidence matrix of a graph $G(V, E)$ defined as $\mathbf{B}_{n \times m}$, where $n, m$ are the number of vertices and edges, respectively is given by

$$
\mathbf{B} \triangleq \begin{cases}1 & \text { if } v_{i} \in e_{j} \\ 0 & \text { otherwise }\end{cases}
$$

Definition 3.5.4 (Directed graph). The incidence matrix of a loopless directed graph (or digraph) $G$, we set $\left(b_{i j}\right)=+1$ if $v_{i}$ is the tail of $e_{j}$ and $\left(b_{i j}\right)=-1$ if $v_{i}$ is the head of $e_{j}$ and 0 , otherwise. That is

$$
\mathbf{B}_{n \times m}=\left(b_{i j}\right) \triangleq\left\{\begin{array}{cc}
+1 & \text { if } v_{i} \text { tail of } e_{j}  \tag{3.112}\\
-1 & \text { if } v_{i} \text { head of } e_{j} \\
0 & \text { otherwise }
\end{array}\right.
$$

Proposition 3.5.1. Let $\mathbf{D}_{n \times n}$ be a diagonal matrix $\in \mathbb{R}^{n \times n}$

$$
\mathbf{D}=\left(d_{i j}\right)=\left\{\begin{array}{c}
d_{i i}=d\left(v_{i}\right) \\
d_{i j}=0
\end{array}\right.
$$

Then it holds true

$$
\begin{equation*}
\mathbf{B B}^{\top}=\mathbf{A}+\mathbf{D} \tag{3.113}
\end{equation*}
$$

Definition 3.5.5 (In-degree and Out-degree). Let v be a vertex in a directed graph (or digraph). Then by $d^{+}(v)$ we define the outdegree as the number of edges with tail $v$. In the same way, by $d^{-}(v)$ we define the indegree as the number of edges with head of $v$.

Example 11: Extract the adjacency and incidence matrix of $G(4,6)$.


Graph 1, G(4, 6)
Solution The adjacency matrix of Graph 1 according to the Definition 3.5 .2 is given by

$$
\mathbf{A}_{a d j}=\left[\begin{array}{llll}
0 & 1 & 1 & 0 \\
0 & 0 & 0 & 1 \\
0 & 1 & 0 & 1 \\
1 & 0 & 0 & 0
\end{array}\right]
$$

The incidence matrix of Graph 1 according to the Definition 3.5.4 is given by

$$
\mathbf{B}_{\text {inc }}=\left[\begin{array}{ccccccc}
v_{i} & \phi_{12} & \phi_{13} & \phi_{24} & \phi_{32} & \phi_{34} & \phi_{41} \\
1 & -1 & -1 & 0 & 0 & 0 & 1 \\
2 & 1 & 0 & -1 & 1 & 0 & 0 \\
3 & 0 & 1 & 0 & -1 & -1 & 0 \\
4 & 0 & 0 & 1 & 0 & 1 & -1
\end{array}\right]
$$

Finally, we can compute the outdegree and indegree of the graph taking in account the Definition 3.5.5

$$
\left[\begin{array}{ccc}
v_{i} & d^{-} & d^{+} \\
1 & 1 & 2 \\
2 & 2 & 1 \\
3 & 1 & 2 \\
4 & 2 & 1
\end{array}\right]
$$

## Part III

## Third part

## Chapter 4

## Diffusion on Dynamical Interbank Loan Networks

'Wise men speak because they have something to say; fools because they have to say something. ${ }^{\text {' }}$
— Plato, 427-347 B.C., Greek Philosopher

### 4.1 Introduction

Diffusion is the study of motion of atom, molecules, etc. from a region of higher concentration to a region of lower concentration. The idea behind the diffusion has roots primarily in physics (particle diffusion) but we can find it in chemistry, biology, geology, economics/finance and others scientific areas. A few examples are molecular diffusion, Brownian motion and heat diffusion.

The word diffusion derives from the latin word diffundere, which means "to spread out". The first systematic study of diffusion (in gases) as we looking back to the past was given by Thomas Graham from 1831 to 1833. Moreover, a mathematical description of diffusion and Brownian motion was elaborated by Joseph Fourier in 1822, Robert Brown in 1827, Adolf Fick in 1855 and by Albert Einstein in 1905.

In this study we consider a connected, directed and weighted dynamic network $G(V, E)$ with $V=\{1,2, \ldots, n\}$ and $E=\left\{e_{i j}\right\}, i \neq j$, where V denotes the nodes (banks) with $|V|=n$ and $E$ are the edges $|E|=e$, respectively. The edges represent the fund flow from bank ito bank j, characterized by a positive weight $\phi_{i j}$ and a direction. Furthermore, to each bank we assign a leverage index $\left(l_{i}\right)$, which is the sum of all outgoing fund flows $\phi_{i j}$ over the own capital of bank $\mathrm{i},\left(k_{i}\right)$, as described by (1) with $i \neq j$ i.e. self loops are not allowed (no capital
increase).

$$
\begin{equation*}
l_{i}=\frac{\sum_{j} \phi_{i j}}{k_{i}} \tag{4.1}
\end{equation*}
$$

Moreover, we extract from the graph $G(V, E)$ the adjacency operator $\mathbf{A} \in \mathbb{R}^{n \times n}$, where the main diagonal has all the elements equal to zero and off diagonal are zero or one, zero if there is no connection between two nodes and 1 otherwise, respectively. $\Phi(0)$ is an $n \times n$ graph matrix of the initial loans of the network and $\Phi(t)$ is the evolution of this matrix in time. We assume that $\Phi(t)$ is governed by a diffusion law transfering funds from banks of heigher leverages to those of lower. Thus is described by the following set of matrix differential equations:

$$
\begin{equation*}
\dot{\Phi}=[\mathbf{A}, \mathbf{L}]=\mathbf{A} \operatorname{diag}\left(l_{i}\right)-\operatorname{diag}\left(l_{i}\right) \mathbf{A} \tag{4.2}
\end{equation*}
$$

where $l_{i}$ is the leverage of bank i. In equilibrium we have $\dot{\Phi}=0$ and in a conected network all the leverages of banks will be equal i.e. $l_{i}=l_{j} \Rightarrow l_{i}-l_{j}=0$, for any $(i, j) \in V$. In addition if $\dot{\Phi}=0$ then (4.2) gives a homogeneous system with $e$ the number of unknowns and $n-1$ equations. Thus the dimension of the solution space of the system $(\Sigma)$ is $\operatorname{dim}(\Sigma)=$ $e-\operatorname{rank}\left(\mathbf{I}_{k}\right)=e-(n-1)=e-n+1$ where $\mathbf{I}_{k}$ is a non rectangular operator and in order to have a unique solution should be: $e-n+1=0 \Rightarrow e=n-1$. Therefore if the dimension of $e$ is $n-1$ then there exist a unique solution and $\dot{\Phi}=0$.

Equation (4.2) may be rewritten as a vector differential equation of all loans $\phi(t) \in \mathbb{R}^{e}$ as:

$$
\begin{equation*}
\underset{\sim}{\dot{\phi}}(t)=\Delta \underset{\sim}{\phi}(t) \tag{4.3}
\end{equation*}
$$

where $\phi(t)$ is the vector of loans of network with the solution, described by:

$$
\begin{equation*}
\underset{\sim}{\phi}(t)=e^{\Delta t} \underset{\sim}{\phi}(0) \tag{4.4}
\end{equation*}
$$

where $\underset{\sim}{\phi}(0)$ is the vector of initial loans $(t=0)$. Moreover the operator $\Delta$ contains information from the interconnections and capital of banks and can be written in the form $\Delta=\mathbf{P L} \mathbf{P}^{-\mathbf{1}}$ where $\mathbf{P}, \mathbf{L} \in \mathbb{R}^{n \times n}$ are the operators of eigenvectors and eigenvalues of $\Delta$, respectively. In order to have a stable equilibrium on our system the eigenvalues $z_{i}$ which are derived from $\Delta$ operator should have negative real parts and at least one equal to zero. The eigenvector of 0 corresponds to an atractor of the system or equilibrium and leads to a minimum dispersion of levereges.

In equilibrium the final solution of differential equation as described by (4.3) and given by (4.4), will be a vector $\underset{\sim}{\phi}(+\infty) \in \mathbb{R}^{e \times 1}$ describing all those directed flows $\phi_{i j}$ which all banks in our network should have at $t \rightarrow+\infty$. The leverage of any bank in equilibrium will be given again by (4.1) but now we replace in the numerator the steady state $\phi_{i j}$ which corresponding
to the vector $\underset{\sim}{\phi}(+\infty)$. Both the solution of (4.3) and the equilibrium solution depends on the properties of operator $\Delta$ and the final equilibrium flows calculated by applying an appropriate projection operator on the initial values $\underset{\sim}{\phi}(0)$.

### 4.2 Diffusion equations and equilibrium points

Let a network of n-banks exchanging funds (loans), forming a graph (V, E) V are the nodes of graph i.e. banks and E are the edges denoted by $\phi_{i j}$ (i bank lends to j bank amount $\phi_{i j}$ ). On every node (bank) the leverage is defined as in (4.1) and the aim of this paper is to provide means to minimize the dispersion of leverages in the network making it more stable. We propose the diffusion method i.e. funds are transfered from those banks of high leverage to those of low this is written as:

$$
\dot{\phi}_{i j}=l_{j}-l_{i}
$$

where $\phi_{i j}$ is a function of time t and with initial condition $\phi_{i j}(0)$. These equations can be written in matrix form as (4.2). We may rewrite them in matrix form by introducing the following notation:

1. $\Phi$ is the matrix of loans.
2. $\operatorname{Vec}(\Phi)$ is the column vector of all columns of $\Phi$ superimposed.
3. $D_{n^{2}-e}^{(0)}(\operatorname{Vec}(\Phi))$ is the vector taken from $\operatorname{Vec}(\Phi)$ by deleting the fixed $n^{2}-e$ zeros appearing in $\Phi$ and $D_{n^{2}-e}^{(0)}$ is the corresponding selection matrix.
4. $D_{n^{2}-n}^{(0)}\left(\operatorname{Vec}\left(\operatorname{diag}\left(l_{i}\right)\right)\right)$ is the $n \times 1$ vector $\left[l_{1}, l_{2}, \ldots, l_{n}\right]^{T}$ and $D_{n^{2}-n}^{(0)}$ is the corresponding selection matrix.
5. The above two operators will be also applied to matrices as induced by matrix multiplication on $\operatorname{Vec}(\Phi)$ and we will loosely apply the same notation to them.
6. $\otimes$ is the matrix tensor product.

Theorem 4.2.1. The diffusion equation (4.2) may be written in vector form as

$$
\begin{equation*}
\dot{\phi}(t)=\Delta \phi(t) \tag{4.5}
\end{equation*}
$$

which is a system of homogeneous linear differential equations
Theorem 4.2.2 (General approach). For any connected, directed and weighted network in the equillibrium all the leverages $l_{i}, i=\{1,2, \ldots, n\}$ are equal, i.e. $l_{i}=l_{j}, i \neq j$.

Proof. Let the following equation:

$$
\begin{equation*}
\dot{\Phi}=\mathbf{A} \operatorname{diag}\left(l_{i}\right)-\operatorname{diag}\left(l_{i}\right) \mathbf{A} \tag{4.6}
\end{equation*}
$$

Taking the transpose of (4.6) we get:

$$
\begin{equation*}
\dot{\Phi}^{T}=\operatorname{diag}\left(l_{i}\right) \mathbf{A}^{T}-\mathbf{A}^{T} \operatorname{diag}\left(l_{i}\right) \Rightarrow \dot{\Phi}^{T}=\mathbf{I}_{n} \operatorname{diag}\left(l_{i}\right) \mathbf{A}^{T}-\mathbf{A}^{T} \operatorname{diag}\left(l_{i}\right) \mathbf{I}_{n} \tag{4.7}
\end{equation*}
$$

From linear algebra and more precisely from Kronecker product we know a property which says that:

$$
\begin{equation*}
\operatorname{Vec}(\mathbf{A X B})=\left(\mathbf{B}^{T} \otimes \mathbf{A}\right) \operatorname{Vec}(\mathbf{X}) \tag{4.8}
\end{equation*}
$$

Apply the $\operatorname{Vec}(\cdot)$ operator to (4.7) and taking in account (4.8) we get:

$$
\begin{align*}
\operatorname{Vec}\left(\dot{\Phi}^{T}\right) & =\left(\mathbf{A} \otimes \mathbf{I}_{n}\right) \operatorname{Vec}\left(\operatorname{diag}\left(l_{i}\right)\right)-\left(\mathbf{I}_{n} \otimes \mathbf{A}^{T}\right) \operatorname{Vec}\left(\operatorname{diag}\left(l_{i}\right)\right) \\
\operatorname{Vec}\left(\dot{\Phi}^{T}\right) & =\left[\left(\mathbf{A} \otimes \mathbf{I}_{n}\right)-\left(\mathbf{I}_{n} \otimes \mathbf{A}^{T}\right)\right] \operatorname{Vec}\left(\operatorname{diag}\left(l_{i}\right)\right) \\
D_{n^{2}-e}^{(0)}\left(\operatorname{Vec}\left(\dot{\Phi}^{T}\right)\right) & =D_{n^{2}-e}^{(0)}\left[\left[\left(\mathbf{A} \otimes \mathbf{I}_{n}\right)-\left(\mathbf{I}_{n} \otimes \mathbf{A}^{T}\right)\right] \operatorname{Vec}\left(\operatorname{diag}\left(l_{i}\right)\right)\right\} \\
\dot{\phi}(t) & =D_{n^{2}-e}^{(0)}\left[\left(\mathbf{A} \otimes \mathbf{I}_{n}\right)-\left(\mathbf{I}_{n} \otimes \mathbf{A}^{T}\right)\right]\left[D_{n^{2}-n}^{(0)}\right]^{T} \underbrace{D_{n^{2}-n}^{(0)}\left[\operatorname{Vec}\left(\operatorname{diag}\left(l_{i}\right)\right)\right]}_{A} \tag{4.9}
\end{align*}
$$

The part (A) of (4.9) is just the vector of levereges $l_{i}$ after canceling out all the zeros. This can be further expressed as:

$$
\left[\begin{array}{c}
l_{1} \\
l_{2} \\
\vdots \\
l_{n}
\end{array}\right]=\operatorname{diag}\left(1 / k_{i}\right)\left(\mathbf{I}_{n} \otimes 1_{n}\right) \operatorname{Vec}\left(\Phi^{T}\right)=\operatorname{diag}\left(1 / k_{i}\right)\left(\mathbf{I}_{n} \otimes 1_{n}\right) D_{n^{2}-e}^{(0)} \phi(t)
$$

where $k_{i}$ is the initial capital per bank i. Plug in the last equation to (4.9) we get:

$$
\begin{align*}
& \dot{\phi}(t)=\underbrace{D_{f, n^{2}-e}^{(0)}\left\{\left[\left(\mathbf{A} \otimes \mathbf{I}_{n}\right)-\left(\mathbf{I}_{n} \otimes \mathbf{A}^{T}\right)\right]\right\} D_{f, n^{2}-n}^{(0)}}_{\Delta_{1} \in \mathbb{R}^{e \times n}} \underbrace{\operatorname{diag}\left(1 / k_{i}\right)}_{\in \mathbb{R}^{n \times n}} \underbrace{\left(I_{n} \otimes 1_{n}\right) D_{f, n^{2}-e}^{(0)}}_{\Delta_{2} \in \mathbb{R}^{n \times e}} \phi(t) \\
& \dot{\phi}(t)=\underbrace{\Delta_{1} \operatorname{diag}\left(1 / k_{i}\right) \Delta_{2}}_{\Delta \in \mathbb{R}^{e \times e}} \phi(t) \\
& \dot{\phi}(t)=\Delta \phi(t) \tag{4.10}
\end{align*}
$$

Equation (4.10) define a set of diffusion differential equations describing the evolution of interbank loans. The state space of this dynamical system is $\mathbb{R}^{e}$ with dimension equal to the number of loans, and the state matrix $\Delta \in \mathbb{R}^{e x e}$. The matrix $\Delta$ contains information of the structure of the network and the capitals of banks. Thus defines an operator semigroup $e_{t>0}^{\Delta t}$ which acts on the initial vector of loans $\phi(0)$ and the equilibrium is reached at $t=+\infty$ provided $\Delta$ is such that $\lim _{t \rightarrow+\infty} e^{\Delta t}$ exists. Similarly we may define evolution equations for the leverages $\underset{\sim}{l}=\left(l_{i}\right)_{i=1}^{n}$ as follows:

Theorem 4.2.3. The vector $l$ of leverages is governed by a set of linear differential equations $\dot{\sim}=\mathscr{L} \underline{\sim}$ where $\mathscr{L}$ is an appropriate Laplacian.

Proof. Consider the set of matrix diffusion equations

$$
\begin{equation*}
\dot{\Phi}=\mathbf{A} \operatorname{diag}\left(l_{i}\right)-\operatorname{diag}\left(l_{i}\right) \mathbf{A} \tag{4.11}
\end{equation*}
$$

then as $\underset{\sim}{l}=\operatorname{diag}\left(1 / k_{i}\right) \Phi 1_{n}$ we get

$$
\begin{equation*}
\underset{\sim}{\dot{i}}=\operatorname{diag}\left(1 / k_{i}\right) \mathbf{A} \underline{\sim}-\operatorname{diag}\left(1 / k_{i}\right) \operatorname{diag}\left(a_{i}\right) \underline{l} \tag{4.12}
\end{equation*}
$$

where

$$
\mathbf{A} 1_{n}=\left[\begin{array}{c}
a_{1} \\
a_{2} \\
\vdots \\
a_{n}
\end{array}\right]
$$

and $a_{i}, i \in\{1,2, \ldots, n\}$ are the column degrees of the graph. Thus

$$
\begin{equation*}
\underset{\sim}{i}=\mathscr{L} \underline{\underline{l}} \tag{4.13}
\end{equation*}
$$

where $\mathscr{L}=\operatorname{diag}\left(1 / k_{i}\right)\left(\mathbf{A}-\operatorname{diag}\left(a_{i}\right)\right)$

Thus (4.13) describes the evolution of leverages, $\mathscr{L} \in \mathbb{R}^{n \times n}$ is the Laplacian for the network and the equilibrium leverages are given by:

$$
\begin{equation*}
\underset{\sim}{l(+\infty)=\lim _{t \rightarrow+\infty} e^{\mathscr{L} t} \underset{\sim}{l}(0), ~(0)} \tag{4.14}
\end{equation*}
$$

The equilibria for both (4.13) and (4.10) depends on the spectral structure of the operatrors $\Delta$ and $\mathscr{L}$. We seak for both there are systems to lead to an equilibrium in which case the eqilibrium loans and leverages are the ones we require.

Theorem 4.2.4. When a network $G(V, E)$ is connected then in the equilibrium the leverages $l_{i}, i=\{1,2, \ldots, n\}$ are all equal, i.e. $l_{1}=l_{2}=\cdots=l_{n}$ or in other words $l_{i}=l_{j}$ with $i \neq j$.

Proof. Let the following equation:

$$
\begin{equation*}
\dot{\Phi}=\mathbf{A} \operatorname{diag}\left(l_{i}\right)-\operatorname{diag}\left(l_{i}\right) \mathbf{A} \tag{4.15}
\end{equation*}
$$

At the point of steady-state we have $\dot{\Phi}=0$ and (4.15) gives:

$$
\begin{equation*}
\mathbf{A} \operatorname{diag}\left(l_{i}\right)=\operatorname{diag}\left(l_{i}\right) \mathbf{A} \tag{4.16}
\end{equation*}
$$

Therefore at equilibrium we have that $l_{i}=l_{j} \forall(i, j) \in E$. Consider now a connected network with adjacency matrix $\mathbf{A}$. Then for every pair $i, j \in V$ there is a path which connect $i$ and $j$, this means that there are pairs in the graph, i.e. $\left(i_{1}, j_{1}\right),\left(i_{2}, j_{2}\right),\left(i_{3}, j_{3}\right), \ldots,\left(i_{k}, j_{k}\right)$ that connect $i$ and $j$ such that satisfy:

$$
\left\{i_{k}, j_{k}\right\} \bigcap\left\{i_{k+1}, j_{k+1}\right\}=\left\{r_{k}\right\} .
$$

Hence $l_{i}=l_{r_{1}}=l_{r_{2}}=\cdots=l_{r_{k}}=l_{j}$ and thus for every pair $i, j \in V$ at equilibrium we have that $l_{i}=l_{j}$. Therefore when the network is connected at equilibrium we must have all leverages equal.

### 4.2.1 Connectivity and equilibria

It is worth noting that, from an operator we can understand if our network is connected. Suppose that we have an adjacency operator $\mathbf{A}$ with all the elements $a_{i j}=0$, for all $i=j$.

Definition 4.2.1. [45] A directed graph is called weakly connected if replacing all of its directed edges with undirected edges produces a connected (undirected) graph. It is connected if it contains a directed path from $u$ to $v$ or a directed path from $v$ to $u$ for every pair of vertices $u$, $v$. It is strongly connected or simply strong if it contains a directed path from $u$ to $v$ and a directed path from $v$ to $u$ for every pair of vertices $u, v$. The strong components are the maximal strongly connected subgraphs.

## Connected and unconnected network

Suppose we have a connected, directed and weighted network. In order to characterize it as connected it must satisfy the following four cases-theorems.

Theorem 4.2.5. [32] Suppose that a network $G(V, E)$ is directed. Let $\mathbf{A}$ be the adjacency operator of $G(V, E)$. We convert the number of ones from $\mathbf{A}$ as a new nodes and connecting them in such a way in order to get triangles. Moreover, let $\mathscr{X}_{E}(G)$ be the Euler characteristic of graph G. If the graph is (weakly) connected then the Euler characteristic is equal to 1 or $\mathscr{X}_{E}(G)=1$, i.e. $:$

$$
\begin{equation*}
\mathscr{X}_{E}(G)=|V|-|E|+|F|=1 \tag{4.17}
\end{equation*}
$$

where $|V|,|E|$ and $|F|$, are the number of vertices, edges and faces, respectively.
Theorem 4.2.6. Let $\mathbf{A}$ be the adjacency operator of a graph $G(V, E)$ and $f_{2}:\{0,1,2\} \rightarrow$ $\{0,1\}$. We call that the graph $G$ is unconnected if and only if:

$$
\mathbf{S}\left(\mathbf{A}+\mathbf{A}^{T}\right)_{i j}=f_{2}\left(a_{i j}+a_{j i}\right) \text { where } f_{2}=\left\{\begin{array}{lll}
x & \text { if } x \in\{0,1\} \\
1 & \text { if } x=2
\end{array}\right.
$$

has at least $n-1$ elements equal to one at symmetric positions based on the main diagonal and more than $n-2$ secondary diagonal with zeros.

Theorem 4.2.7. Let A be the adjacency operator of a network. The network is connected if there are $n-1$ ones in such a way that the pairs of the network compose a connection.

Theorem 4.2.8. Let $\mathbf{A}$ be an adjacency operator of a connected, directed and weighted network. If there is no way to make $\mathbf{A}$ block diagonal by permutating rows and columns by the same permutation matrix $\mathbf{Q}$, then the graph is connected.

Theorem 4.2.9. Let $G(V, E)$ be a graph with $n$ vertices and at least $n-1$ edges. Then we can say that:

If there are connected components then the leverages are all equals only in the connected components of the network and not among them.

Proof. We define a connected component network G with n nodes and at least $n-1$ edges. By (4.2) we have

$$
\begin{equation*}
\dot{\Phi}=[\mathbf{A}, \mathbf{L}]=\mathbf{A} \operatorname{diag}\left(l_{i}\right)-\operatorname{diag}\left(l_{i}\right) \mathbf{A} \tag{4.18}
\end{equation*}
$$

In the equilibrium $\dot{\Phi}=0$. Thus (4.18) gives:

$$
\begin{equation*}
[\mathbf{A}, \mathbf{L}]=\mathbf{A} \operatorname{diag}\left(l_{i}\right)-\operatorname{diag}\left(l_{i}\right) \mathbf{A}=0 \tag{4.19}
\end{equation*}
$$

We have proved from Theorem 2.2 that when we have a connected network G the leverages are all equal, i.e. $l_{i}=l_{j}$ or $l_{i}-l_{j}=0, \forall(i, j) \in V$ with $i \neq j$. Thus we have the following system ( $\Sigma$ )

$$
\underbrace{\left[\begin{array}{c|c|c|c|c|c|c}
\frac{1}{k_{1}} 1_{a_{1}} & \frac{-1}{k_{2}} 1_{a_{2}} & 0_{a_{3}} & \ldots & 0 & 0 & 0 \\
0 & \frac{1}{k_{2}} 1_{a_{2}} & \frac{-1}{k_{3}} 1_{a_{3}} & \ldots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ldots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \ldots & 0 & \frac{1}{k_{n-1}} 1_{a_{n-1}} & \frac{-1}{k_{n}} 1_{a_{n}}
\end{array}\right]}_{\mathbf{I}_{k}} \underset{\sim}{\phi}=0_{n-1,1}
$$

which is equivalent to $\Delta \phi=\underset{\sim}{0}$.
This is a homogeneous system of $n-1$ equations in $n$ unknowns. For $n$ banks to form a connected graph, with no self loops, we must have $n-1 \leq e \leq n^{2}-n$. The solution space of the homogeneous system $(\Sigma)$ is as linear space of dimension $\operatorname{dim}(\Sigma)$ which is equal to the dimension $\operatorname{dim}\left(\operatorname{Ker} \mathbf{I}_{k}\right)$. Hence:

$$
\begin{equation*}
\operatorname{dim}(\Sigma)=e-\operatorname{rank}\left(\mathbf{I}_{k}\right)=e-(n-1)=e-n+1 \tag{4.20}
\end{equation*}
$$

If we further impose conservation of the total volume of loans in the system i.e. $\sum_{i} \phi=c$ and $\mathbf{V}$ is a positive $(e-n+1) e$ bases matrix for $(\Sigma)$ then a solution for $\phi_{e q}$ for equilibrium is given by:

$$
\begin{equation*}
\phi_{e q}=\frac{1^{T} \Sigma}{1^{T} \Sigma 1} c \tag{4.21}
\end{equation*}
$$

Thus, from one hand if the operator $\mathbf{I}_{k}$ has at least one element per line not equal to zero then all the leverages in equilibrium are equal, i.e. $l_{1}=l_{2}=\cdots=l_{n}=l$ but on the other hand if there exists one line with no outgoing flows, then it's true that the leverages are all equal and more precisely equals to zero, i.e. $l_{1}=l_{2}=\cdots=l_{n}=0$.

Suppose that we have a connected, directed and weighted network of banks where they exchange loans at some unit of time. By (4.10) we have:

$$
\begin{equation*}
\underset{\sim}{\dot{\phi}}(t)=\Delta \underset{\sim}{\phi}(t) \tag{4.22}
\end{equation*}
$$

with solution:

$$
\begin{equation*}
\underset{\sim}{\phi}(t)=\Delta{\underset{\sim}{x}}_{\phi}^{\phi}(0) \tag{4.23}
\end{equation*}
$$

where $\underset{\sim}{\phi}(0)$ is the $n \times 1$ vector of initial values of loans $(\mathrm{t}=0)$, and at the same time the leverages $\left(l_{i}\right)$ of banks, are not all equal or in other words the banking system is not in balance. Then through the proposed diffusion process the network will approach balance, as $t \rightarrow+\infty$. Furthermore, at equilibrium all the loan flows, $\phi$, are describing a vector space, as the next formula says:

$$
\begin{equation*}
\mathbf{V}_{b}=\{\phi: \Delta \phi=0\} \tag{4.24}
\end{equation*}
$$

In addition this imbalance in the network can be described by the norm of $\|\Delta \phi(t)\|$ and if this measure increases then we can say that the leverages among banks diverge and on the other hand if this measure is equal to zero then we have an equilibrium in the network. Our main target is to make a projection of the initial flows $\phi(0)$ on $\mathbf{V}_{b}$ produces a row required vector of flow $\phi_{b}$. This is done by taking an orthonormal basis of $\mathbf{V}_{b}$ and $\phi_{b}$ is given by:

$$
\begin{equation*}
\phi_{b}=\mathbf{B}_{b} \mathbf{B}_{b}^{T} \phi(0) \tag{4.25}
\end{equation*}
$$

where $\mathbf{B}_{b} \mathbf{B}_{b}^{T}=\mathbf{P}_{\mathbf{V}_{b}} \in \mathbb{R}^{e \times e}$ is a projection operator to that vector space $\mathbf{V}_{b}$ so that $\left\|\Delta \phi_{b}\right\|$ is minimal and $\mathbf{B}_{b}$ is an orthonormal basis of $\mathbf{V}_{b} . \mathbf{P}_{b}$ also can be alternatively calculated as:

1. by difussion:

$$
\lim _{t \rightarrow+\infty} e^{\Delta t}=\mathbf{P}_{b}
$$

2. by operational calculus:

$$
\lim _{\varepsilon \rightarrow 0} \varepsilon\left(\varepsilon \mathbf{I}_{e}-\Delta\right)^{-1}=\mathbf{P}_{b}
$$

### 4.3 The basic two structural cases by toy-examples - calculations of equilibrium points

### 4.3.1 Case 1: At least one non-zero element per line to the adjacency operator

The reason that we are study next two cases is that we want to mention that those graphs they have something special according to their connections. As we can observe that both graphs 1 and 2 are connected in a such a way that in spite of diffusion process don't works properly for every node finally we get a solution that all the vertices in equilibrium will have the same leverage.

Let the following graph $\mathrm{G}(\mathrm{V}, \mathrm{E})$, with the set of vertices as $V=\{1,2,3,4,5\},|V|=5$ and edges, $E=\left\{e_{12}, e_{25}, e_{31}, e_{35}, e_{41}, e_{54}\right\},|E|=6$.


Graph 1, G(5, 6)
Every vertex on Graph 1 represented by a number which is a bank. The arrows shows us the direction of the loans which means which bank gives a loan to another and we denoted by $\phi_{i j}$. Furthermore, every bank is associated with a capital $k_{i}$ and we calculate the initial leverages for any bank-vertex i. Thus, let's say that the initial per bank capitals are: $k_{1}=81, k_{2}=9, k_{3}=27, k_{4}=4, k_{5}=3$. Thus, the leverages are:

$$
\begin{equation*}
l_{1}=\frac{\phi_{12}}{k_{1}}, l_{2}=\frac{\phi_{25}}{k_{2}}, l_{3}=\frac{\phi_{31}+\phi_{35}}{k_{3}}, l_{4}=\frac{\phi_{41}}{k_{4}}, l_{5}=\frac{\phi_{54}}{k_{5}} \tag{4.26}
\end{equation*}
$$

The adjacency operator which coresponds to Graph 1 is:

$$
\mathbf{A}_{a d j}=\left[\begin{array}{ccccc}
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0
\end{array}\right]
$$

According to the Theorem 2.2, in the equilibrium we have that: $l_{1}=l_{2}=l_{3}=l_{4}=l_{5}$ and from the last expression we get:

$$
l_{1}-l_{2}=l_{2}-l_{3}=l_{3}-l_{4}=l_{4}-l_{5}=0
$$

Due to (4.26) if we replace $l_{i}$ to the above equations and then write them in matrix form we get an $m \times n, n>m$ operator, $\mathbf{A}_{\mathbf{1}}$ which is a special type of incident operator. Thus we get:

$$
\underbrace{\left[\begin{array}{cccccc}
1 / k_{1} & -1 / k_{2} & 0 & 0 & 0 & 0 \\
0 & 1 / k_{2} & -1 / k_{3} & -1 / k_{3} & 0 & 0 \\
0 & 0 & 1 / k_{3} & 1 / k_{3} & -1 / k_{4} & 0 \\
0 & 0 & 0 & 0 & 1 / k_{4} & -1 / k_{5}
\end{array}\right]}_{\mathbf{A}_{\mathbf{1}}}\left[\begin{array}{l}
\phi_{12} \\
\phi_{25} \\
\phi_{31} \\
\phi_{35} \\
\phi_{41} \\
\phi_{54}
\end{array}\right]=\left[\begin{array}{l}
0 \\
0 \\
0 \\
0
\end{array}\right](\Sigma)
$$

If we replace $k_{i}$ to the operator $\mathbf{A}_{\mathbf{1}}$ and do some algebraic operations, finaly we get:

$$
\mathbf{A}_{\mathbf{1}}=\left[\begin{array}{cccccc}
1 / 81 & -1 / 9 & 0 & 0 & 0 & 0 \\
0 & 1 / 9 & 0 & 0 & -1 / 4 & 0 \\
0 & 0 & 1 / 27 & 1 / 27 & -1 / 4 & 0 \\
0 & 0 & 0 & 0 & 1 / 4 & -1 / 3
\end{array}\right]
$$

For those values of operator $\mathbf{A}_{\mathbf{1}}$ the system $(\Sigma)$ gives: $\phi_{12}=(81 / 4) \phi_{41}, \phi_{25}=(9 / 4) \phi_{41}, \phi_{31}=$ $(27 / 4) \phi_{41}-\phi_{35}, \phi_{54}=(3 / 4) \phi_{41}$ and the right kernel corresponding to equilibrium is:

$$
\left[\begin{array}{l}
\phi_{12} \\
\phi_{25} \\
\phi_{31} \\
\phi_{35} \\
\phi_{41} \\
\phi_{54}
\end{array}\right]=\phi_{41} \underbrace{\left[\begin{array}{c}
81 / 4 \\
9 / 4 \\
27 / 4 \\
0 \\
1 \\
3 / 4
\end{array}\right]}_{\mathbf{a}_{1}}+\phi_{35}^{\left[\begin{array}{c}
0 \\
0 \\
-1 \\
1 \\
0 \\
0
\end{array}\right]}
$$

Consequently we have found a set of two elements - vectors, $\mathbf{a}_{1}, \mathbf{a}_{2}$ which are compose the basis $\mathbf{V}$ and $\phi_{45}, \phi_{35}, \in \mathbb{R}$. Suppose that $\lambda \equiv \phi_{41}, \mu \equiv \phi_{35}$ and if we want the total, current number of loans into the interbank system to be for instance 93 monetary units then we have: $(81 / 4) \lambda+(9 / 4) \lambda+(27 / 4) \lambda+\lambda+(3 / 4) \lambda-\mu+\mu=93 \Rightarrow \lambda=3$. This gives rise to equilibrium flows as:

$$
\begin{align*}
\phi_{12} & =(81 / 4) \times 3=243 / 4 \\
\phi_{25} & =(9 / 4) \times 3=27 / 4 \\
\phi_{31} & =(27 / 4) \times 3=81 / 4 \\
\phi_{35} & =0 \\
\phi_{41} & =1 \times 3=3 \\
\phi_{54} & =(3 / 4) \times 3=9 / 4 \tag{4.27}
\end{align*}
$$

By (4.26) the initial leverages was: $l_{1}=1 / 4, l_{2}=27 / 16, l_{3}=27 / 16, l_{4}=1 / 4, l_{5}=3 / 16$. But now, if we replace in (4.26) the new flows which are given by (27) we get the final leverages which are: $l_{1}=l_{2}=l_{3}=l_{4}=l_{5}=3 / 4$.

## Conclusion

Equilibrium is achieved by transferring funds from banks of higher leverage to those of lower.

### 4.3.2 Case 2: Adjacency operator with two lines equals to zero

Let the following network $\mathrm{G}(\mathrm{V}, \mathrm{E})$ with a set of vertices $\mathrm{V}=\{1,2,3,4,5\},|V|=5$ and edges $\mathrm{E}=\left\{e_{12}, e_{32}, e_{34}, e_{45}\right\},|E|=4$.


Graph 2, G(5, 4)

Suppose that the initial per bank capitals are $k_{1}=81, k_{2}=9, k_{3}=27, k_{4}=3, k_{5}=3$. Moreover, the initial leverages for those banks which are given from Graph 2 are:

$$
\begin{equation*}
l_{1}=\frac{\phi_{12}}{k_{1}}, l_{2}=0, l_{3}=\frac{\phi_{32}+\phi_{34}}{k_{3}}, l_{4}=\frac{\phi_{45}}{k_{4}}, l_{5}=0 \tag{4.28}
\end{equation*}
$$

The adjacency operator which can be extracted from the Graph 2 is the following:

$$
\mathbf{A}_{a d j}=\left[\begin{array}{ccccc}
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0
\end{array}\right]
$$

In the equilibrium as we already mentioned at Theorem 2.2, all the leverages are equall, thus we have: $l_{1}=l_{2}=l_{3}=l_{4}=l_{5}$ or equivalently:

$$
l_{1}=0, l_{3}=0, l_{3}-l_{4}=0, l_{4}=0
$$

By (4.28) if we replace $l_{i}$ to the above equations and write them in matrix form we get an $m \times n, n>m$ operator $\mathbf{A}_{\mathbf{2}}$, which is a special type of incident operator:

$$
\underbrace{\left[\begin{array}{cccc}
1 / 81 & 0 & 0 & 0 \\
0 & 1 / 27 & 1 / 27 & 0 \\
0 & 1 / 27 & 1 / 27 & -1 / 3 \\
0 & 0 & 0 & 1 / 3
\end{array}\right]}_{\mathbf{A}_{\mathbf{2}}}\left[\begin{array}{l}
\phi_{12} \\
\phi_{32} \\
\phi_{34} \\
\phi_{45}
\end{array}\right]=\left[\begin{array}{l}
0 \\
0 \\
0 \\
0
\end{array}\right](\Sigma)
$$

We can observe that $\mathbf{A}_{\mathbf{2}}$ is an operator with rank $=3$, i.e. as the number of non-zero rows. Thus, from the system $(\Sigma)$ above we get: $\phi_{12}=0, \phi_{32}=-\phi_{34}, \phi_{45}=0$ or

$$
\left[\begin{array}{c}
\phi_{12} \\
\phi_{32} \\
\phi_{34} \\
\phi_{45}
\end{array}\right]=\left[\begin{array}{c}
0 \\
-\phi_{34} \\
\phi_{34} \\
0
\end{array}\right]=\phi_{34}^{\left[\begin{array}{c}
0 \\
-1 \\
1 \\
0
\end{array}\right]}(\Sigma)
$$

Any vector from $\Lambda(\Sigma)$ always gives rise to zero equilibrium leverages for all banks.

## Conclusion

If we have a connected, directed and weighted Graph $G(V, E)$ and the adjacency operator of the graph has at least one row with all the elements equals to zero, then the vertices (banks) of the graph at equilibrium will have all the leverages equal to zero.

### 4.4 Differential equation and its solution

In what it follows we do an approach of a more general solution in matrix form of the equation:

$$
\begin{equation*}
\dot{\Phi}=[\mathbf{A}, \mathbf{L}] \Leftrightarrow \dot{\phi_{i j}}=l_{j}-l_{i} \tag{4.29}
\end{equation*}
$$

The incident operator should has at least one non-zero element in every row. Furthermore, we are study connected, directed and weight networks. Thus we have:

$$
\dot{\phi}=\Delta \phi
$$

with solution is given by:

$$
\phi(t)=e^{\Delta t} \phi(0)
$$

Next we are working on the rectangular operator $\mathbf{A}$ and we diagonalize it in order to get as a resault the eigenvalues and eigenvectors, respectively. In our analysis we are interested to those cases where on the one hand the operator $\mathbf{A}$ has all the eigenvalues $\lambda_{i}$ negative and at least one equal to zero, i.e. $\left\{\lambda_{i} \in \mathbb{R}_{-}: \lambda_{i} \leq 0\right\}$ and on the other hand, if it has complex eigenvalues the real part should be negative, i.e. $\left\{\lambda_{i} \in \mathbb{C}: \mathfrak{R e}\left(\lambda_{i}\right) \leq 0\right\}$. Thus let, $\lambda_{1}, \lambda_{2}, \ldots, \lambda_{n}$ and $e_{1}, e_{2}, \ldots, e_{n}$ are the eigenvalues and eigenvectors, respectively. Furthermore, the solution of the system will be of the form:

$$
\begin{equation*}
\underset{\sim}{\dot{\phi}}(t)=e^{\Delta t} \underset{\sim}{\phi}(0) \tag{4.30}
\end{equation*}
$$

where $\underset{\sim}{\phi} i j(0)$ are the initial conditions. As we know the rectangular operator can be written in the form as:

$$
\begin{equation*}
\Delta=\mathbf{P} \Lambda \mathbf{P}^{-1} \tag{4.31}
\end{equation*}
$$

where $\mathbf{P}, \Lambda$ are the eigenvectors and eigenvalues matrices, respectively. Then we have:

$$
\begin{equation*}
\underset{\sim}{\dot{\phi}}(t)=\mathbf{P} e^{\Lambda t} \mathbf{P}^{-1} \underset{\sim}{\phi}(0) \tag{4.32}
\end{equation*}
$$

and provided $\Lambda$ has simple structure with only zero eigenvalues or eigenvalues with negative real part. Thus $\lim _{t \rightarrow \infty} e^{\Lambda t}$ can be calculated as:

$$
\lim _{t \rightarrow+\infty} e^{\Lambda t}=\left[\begin{array}{cccc}
\lim _{t \rightarrow+\infty} e^{\lambda_{1} t} & 0 & \cdots & 0 \\
0 & \lim _{t \rightarrow+\infty} e^{\lambda_{2} t} & \ddots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \lim _{t \rightarrow+\infty} e^{\lambda_{e} t}
\end{array}\right]
$$


and hence $\phi(+\infty)=\mathbf{P}_{b} \phi(0)$ where $\mathbf{P}_{b}$ is the required projection operator.

### 4.5 Solution by diffusion three structural examples

### 4.5.1 Case 1: Two real negatives eigenvalues and one zero

The operator $\Delta$ of differential equation as well as the Laplacian of this problem are not symmetric hence their eigenvalues may be complex and also it is possible that $\Delta$ has no zero eigenvalues. This is demonstrated in the following examples.

Suppose next graph, which describes an interbank loan system among three banks as V $=\{1,2,3\}$ and the exchange transfer loans among them as $\mathrm{E}=\left\{e_{12}, e_{23}, e_{32}\right\}$, respectively. Furthermore, each bank has an initial capital $k_{i}, i=\{1,2,3\}$ and more precisely $k_{1}=2, k_{2}=4$ and $k_{3}=3$. Moreover, suppose the initial leverage, i.e. $\phi_{i j}(0)=10$ for all those three banks.


## Graph 1, G(3, 3)

The adjacency operator from Graph 1 is given by:

$$
\mathbf{A}_{a d j}=\left[\begin{array}{ccc}
0 & 1 & 0 \\
0 & 0 & 1 \\
0 & 1 & 0
\end{array}\right]
$$

According to (4.29) we can write the differential equations for Graph 1 as follows:

$$
\begin{aligned}
& \dot{\phi}_{12}=\frac{\phi_{23}}{k_{2}}-\frac{\phi_{12}}{k_{1}} \\
& \dot{\phi}_{23}=\frac{\phi_{32}}{k_{3}}-\frac{\phi_{23}}{k_{2}} \\
& \dot{\phi}_{32}=\frac{\phi_{23}}{k_{2}}-\frac{\phi_{32}}{k_{3}}
\end{aligned}
$$

or in matrix form:

$$
\left[\begin{array}{c}
\dot{\phi}_{12}  \tag{4.33}\\
\dot{\phi}_{23} \\
\dot{\phi}_{32}
\end{array}\right]=\underbrace{\left[\begin{array}{ccc}
(-1) / k_{1} & 1 / k_{2} & 0 \\
0 & (-1) / k_{2} & 1 / k_{3} \\
0 & 1 / k_{2} & (-1) / k_{3}
\end{array}\right]}_{\Delta}\left[\begin{array}{l}
\phi_{12} \\
\phi_{23} \\
\phi_{32}
\end{array}\right]
$$

and the solution as given by (4.30). Replacing all $k_{i}$ as given into the above operator $\Delta$ we get:

$$
\Delta=\left[\begin{array}{ccc}
-1 / 2 & 1 / 4 & 0 \\
0 & -1 / 4 & 1 / 3 \\
0 & 1 / 4 & -1 / 3
\end{array}\right]
$$

The eigenvalues of operator $\Delta$ are:

$$
\lambda_{1}=\frac{-7}{12}, \lambda_{2}=\frac{-1}{12}, \lambda_{3}=0
$$

we rewrite the operator $\Delta$ in the form of (4.31) as:

$$
\Delta=\underbrace{\left[\begin{array}{ccc}
3 & 1 & 2 / 3  \tag{4.34}\\
-1 & 0 & 4 / 3 \\
1 & 0 & 1
\end{array}\right]}_{\mathbf{P}} \underbrace{\left[\begin{array}{ccc}
-7 / 12 & 0 & 0 \\
0 & -1 / 12 & 0 \\
0 & 0 & 0
\end{array}\right]}_{\Lambda} \underbrace{\left[\begin{array}{ccc}
0 & -3 / 7 & 4 / 7 \\
1 & 1 & -2 \\
0 & 3 / 7 & 3 / 7
\end{array}\right]}_{\mathbf{P}^{-1}}
$$

and we are studying the system of differential equations given by (4.32) in the equilibrium, i.e. for time $t \rightarrow+\infty$. Thus we calculate the $\lim _{t \rightarrow+\infty} e^{\Delta t}=\lim _{t \rightarrow+\infty} \mathbf{P} e^{\Lambda t} \mathbf{P}^{-1}$, where $\Lambda$ is the eigenvalues operator of the system and $\mathbf{P}, \mathbf{P}^{-1}$ are the operator and inverse operator of eigenvectors, respectively.

Taking the limit to infinity of $e^{\Lambda t}$ we get:

$$
\lim _{t \rightarrow+\infty} e^{\Lambda t}=\left[\begin{array}{lll}
0 & 0 & 0  \tag{4.35}\\
0 & 0 & 0 \\
0 & 0 & 1
\end{array}\right]
$$

Finally to find the solution we replace (35) to $\underset{\sim}{\dot{\phi}}(t)=e^{\Delta t} \underset{\sim}{\phi}(0)$ which is the solution to the system as was mentioned in (4.30) to get:

$$
\underset{\sim}{\phi}(+\infty)=\lim _{t \rightarrow+\infty} e^{\Delta t} \underset{\sim}{\phi}(0)=\left[\begin{array}{lll}
0 & 2 / 7 & 2 / 7  \tag{4.36}\\
0 & 4 / 7 & 4 / 7 \\
0 & 3 / 7 & 3 / 7
\end{array}\right]\left[\begin{array}{l}
10 \\
10 \\
10
\end{array}\right]=\left[\begin{array}{l}
40 / 7 \\
80 / 7 \\
60 / 7
\end{array}\right]
$$

The initial leverages $l_{i}, i=\{1,2,3\}$ for any of those three vertices-banks $V_{i}$ and the final leverages after replacing the final $\underset{\sim}{\dot{\phi}}(t)$ by (36) are presented in Table 4.1.

Table 4.1 Initial and final leverages $\left(l_{i}\right)$ per bank.

| $l_{i}$ | $\sum_{j} \phi_{i j} / k_{i}$ | initial values | equilibrium values |
| :---: | :---: | :---: | :---: |
| $l_{1}$ | $\frac{\phi_{12}}{k_{1}}$ | $\frac{10}{2}$ | $\frac{20}{7}$ |
| $l_{2}$ | $\frac{\phi_{23}}{k_{2}}$ | $\frac{10}{4}$ | $\frac{20}{7}$ |
| $l_{3}$ | $\frac{\phi_{32}}{k_{3}}$ | $\frac{10}{3}$ | $\frac{20}{7}$ |

To sum up we can say that the diffusion acting up at our system and the result is that in equilibrium all leverages from the three banks are equal. In other words in our system as $t \rightarrow+\infty$ all banks will have equal leverages.

### 4.5.2 Case 2: Two complex eigenvalues with negative real part and one zero

Suppose next graph which depicts 3 vertices-banks and three edges i.e. loans flows from one bank to another. Furthermore, all those banks in the graph they have an initial capital per bank $k_{i}, i=\{1,2,3\}$, which is given by $k_{1}=2, k_{2}=4, k_{3}=3$, respectively.


## Graph 1, G(3, 3)

The adjacency operator of the Graph 1 above is:

$$
\mathbf{A}_{a d j}=\left[\begin{array}{lll}
0 & 1 & 0 \\
0 & 0 & 1 \\
1 & 0 & 0
\end{array}\right]
$$

We apply (4.29) to write down the differential equations of the Graph 1 which are:

$$
\begin{aligned}
& \dot{\phi}_{12}=\frac{\phi_{23}}{k_{2}}-\frac{\phi_{12}}{k_{1}} \\
& \dot{\phi}_{23}=\frac{\phi_{31}}{k_{3}}-\frac{\phi_{23}}{k_{2}} \\
& \dot{\phi}_{31}=\frac{\phi_{12}}{k_{1}}-\frac{\phi_{31}}{k_{3}}
\end{aligned}
$$

or we can rewrite in matrix form as $\underset{\sim}{\dot{\phi}}=\Delta \underset{\sim}{\phi}$ where $\Delta \in \mathbb{R}^{3 \times 3}$ :

$$
\left[\begin{array}{c}
\dot{\phi}_{12}  \tag{4.37}\\
\dot{\phi}_{23} \\
\dot{\phi}_{31}
\end{array}\right]=\underbrace{\left[\begin{array}{ccc}
-1 / k_{1} & 1 / k_{2} & 0 \\
0 & -1 / k_{2} & 1 / k_{3} \\
1 / k_{1} & 0 & -1 / k_{3}
\end{array}\right]}_{\Delta}\left[\begin{array}{l}
\phi_{12} \\
\phi_{23} \\
\phi_{31}
\end{array}\right]
$$

If we replace the $k_{i}$ into the operator $\Delta$ we get:

$$
\Delta=\left[\begin{array}{ccc}
-1 / 2 & 1 / 4 & 0 \\
0 & -1 / 4 & 1 / 3 \\
1 / 2 & 0 & -1 / 3
\end{array}\right]
$$

The eigenvalues of operator $\Delta$ are:

$$
\begin{equation*}
\lambda_{1}=\frac{1}{24}(-13+i \sqrt{47}), \lambda_{2}=\frac{1}{24}(-13-i \sqrt{47}), \lambda_{3}=0 \tag{4.38}
\end{equation*}
$$

We write down the operator $\mathbf{P}, \in \mathbb{R}^{3 \times 3}$ with elements the eigenvectors of above operator $\Delta$.

$$
\mathbf{P}=\left[\begin{array}{ccc}
1 / 12(-5+i \sqrt{47}) & 1 / 12(-5-i \sqrt{47}) & 2 / 3 \\
1 / 12(-7-i \sqrt{47}) & 1 / 12(-7+i \sqrt{47}) & 4 / 3 \\
1 & 1 & 1
\end{array}\right]
$$

Thus as we already know by (4.31), we can rewrite the operator $\Delta$ in the form of:

$$
\Delta=\mathbf{P} \underbrace{\left[\begin{array}{ccc}
1 / 24(-13+i \sqrt{47}) & 0 & 0 \\
0 & 1 / 24(-13-i \sqrt{47}) & 0 \\
0 & 0 & 0
\end{array}\right]}_{\Lambda=\operatorname{diag}\left(\lambda_{i}\right)} \mathbf{P}^{-1}
$$

As in the previous case 1 , we calculate $e^{\Lambda t}$ as $t \rightarrow+\infty$ :

$$
e^{\Lambda t}=\left[\begin{array}{ccc}
\lim _{t \rightarrow+\infty} e^{1 / 24(-13+i \sqrt{47}) t} & 0 & 0  \tag{4.39}\\
0 & \lim _{t \rightarrow+\infty} e^{1 / 24(-13-i \sqrt{47}) t} & 0 \\
0 & 0 & 1
\end{array}\right]
$$

Thus, we have:

$$
\lim _{t B+\infty} e^{\Delta t}=\mathbf{P} \lim _{t \beta+\infty} e^{\Lambda t} \mathbf{P}^{-1}=\left[\begin{array}{ccc}
188 / 3 & 188 / 3 & 188 / 3  \tag{4.40}\\
376 / 3 & 376 / 3 & 376 / 3 \\
94 & 94 & 94
\end{array}\right]
$$

The solution to the system is given by (4.30) and in our case we have:

$$
\underset{\sim}{\dot{\phi}}(+\infty)=\left[\begin{array}{ccc}
94 / 423 & 94 / 423 & 94 / 423  \tag{4.41}\\
188 / 423 & 188 / 423 & 188 / 423 \\
1 / 3 & 1 / 3 & 1 / 3
\end{array}\right]\left[\begin{array}{l}
10 \\
10 \\
10
\end{array}\right]=\left[\begin{array}{c}
20 / 3 \\
40 / 3 \\
10
\end{array}\right]
$$

To sum up the initial leverages $l_{i}, i=\{1,2,3\}$ for any of those vertices-banks $V_{i}$ and the final leverages after replacing the final $\underset{\sim}{\dot{\phi}}(t)$ by (41) are presented in Table 4.2.

Again as in the previous case, the diffusion works in a sense of stabilizer. In other words comparing the results in Table 4.2, we can observe that at $t=0$ the leverage was different for every $V_{i}, i=\{1,2,3\}$ bank and as $t \rightarrow+\infty$, i.e. in the equilibrium, all banks will have equal leverages. The main idea is that a bank which has a small amount of loans it can absorbs

Table 4.2 Initial and final leverages $\left(l_{i}\right)$ per bank.

| $l_{i}$ | $\sum_{j} \phi_{i j} / k_{i}$ | initial values | equilibrium values |
| :---: | :---: | :---: | :---: |
| $l_{1}$ | $\frac{\phi_{12}}{k_{1}}$ | $\frac{10}{2}$ | $\frac{10}{3}$ |
| $l_{2}$ | $\frac{\phi_{23}}{k_{2}}$ | $\frac{10}{4}$ | $\frac{10}{3}$ |
| $l_{3}$ | $\frac{\phi_{31}}{k_{3}}$ | $\frac{10}{3}$ | $\frac{10}{3}$ |

that quantity of the over loaned banks in order the banking network-system to converge in an equilibrium and all of them to have the same leverage ratio.

### 4.5.3 Case 3: Three real negative eigenvalues

Suppose next network where we have four nodes (banks) $V_{i}, i=\{1,2,3,4\}$ and three edges $e_{i j}$ with $i \neq j$ as the loans flows, $\mathrm{E}=\left\{e_{12}, e_{32}, e_{24}\right\}$ as shown in the following Graph 1 . Furthermore, all those four banks they have an initial capital i.e. $k_{1}=2, k_{2}=4, k_{3}=3, k_{4}=5$, respectively.


$$
\text { Graph } 1, G(4,3)
$$

The adjacency operator of the Graph 1 above is:

$$
\mathbf{A}_{a d j}=\left[\begin{array}{cccc}
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0
\end{array}\right]
$$

By (4.29) we can get the system of differential equations which is:

$$
\begin{align*}
& \dot{\phi_{12}}=\frac{\phi_{24}}{k_{2}}-\frac{\phi_{12}}{k_{1}} \\
& \dot{\phi_{32}}=\frac{\phi_{24}}{k_{2}}-\frac{\phi_{32}}{k_{3}} \\
& \dot{\phi_{24}}=\frac{-\phi_{24}}{k_{2}} \tag{4.42}
\end{align*}
$$

or equivalence the former system of differential equations (42) we can rewrite in matrix form as:

$$
\begin{gather*}
\stackrel{\dot{\mathbf{z}}}{\sim}=\Delta \underset{\sim}{\mathbf{z}}  \tag{4.43}\\
{\left[\begin{array}{c}
\dot{\phi}_{12} \\
\dot{\phi}_{32} \\
\dot{\phi}_{24}
\end{array}\right]=\underbrace{\left[\begin{array}{ccc}
-1 / k_{1} & 0 & 1 / k_{2} \\
0 & -1 / k_{3} & 1 / k_{2} \\
0 & 0 & -1 / k_{2}
\end{array}\right]}_{\Delta}\left[\begin{array}{l}
\phi_{12} \\
\phi_{32} \\
\phi_{24}
\end{array}\right]} \tag{4.44}
\end{gather*}
$$

If we replace the initial capital to the operator $\Delta$ we get:

$$
\Delta=\left[\begin{array}{ccc}
-1 / 2 & 0 & 1 / 4 \\
0 & -1 / 3 & 1 / 4 \\
0 & 0 & -1 / 4
\end{array}\right]
$$

The operator $\Delta$ has the following eigenvalues:

$$
\begin{equation*}
\lambda_{1}=\frac{-1}{2}, \lambda_{2}=\frac{-1}{3}, \lambda_{3}=\frac{-1}{4} \tag{4.45}
\end{equation*}
$$

Taking the $\lim _{t \rightarrow+\infty}\left(e^{\Delta t}\right)$ then we get the zero operator, i.e.:

$$
\lim _{t \rightarrow+\infty}\left(e^{\Delta t}\right)=\left[\begin{array}{lll}
0 & 0 & 0  \tag{4.46}\\
0 & 0 & 0 \\
0 & 0 & 0
\end{array}\right]
$$

Thus taking in account (4.30) we get:

$$
\underset{\sim}{\dot{\phi}}(+\infty)=\left[\begin{array}{lll}
0 & 0 & 0  \tag{4.47}\\
0 & 0 & 0 \\
0 & 0 & 0
\end{array}\right]\left[\begin{array}{l}
10 \\
10 \\
10
\end{array}\right]=\left[\begin{array}{l}
0 \\
0 \\
0
\end{array}\right]
$$

Table 4.3 Initial and final leverages $\left(l_{i}\right)$ per bank.

| $l_{i}$ | $\sum_{j} \phi_{i j} / k_{i}$ | initial values | equilibrium values |
| :---: | :---: | :---: | :---: |
| $l_{1}$ | $\frac{\phi_{12}}{k_{1}}$ | $\frac{10}{2}$ | 0 |
| $l_{2}$ | $\frac{\phi_{23}}{k_{2}}$ | $\frac{10}{4}$ | 0 |
| $l_{3}$ | $\frac{\phi_{32}}{k_{3}}$ | $\frac{10}{3}$ | 0 |

To sum up the observation we can do as we compare the results in Table 4.3 is that the new leverages are all equal to zero due to the fact that the eigenvalues of the related operator are all negative (none of them is zero) and the final projection operator at infinity is identically zero i.e. $\Lambda=\operatorname{diag}\left(\lambda_{i}\right)$ goes to zero.

## Conclusion

In all of these three cases study presented at section 5 we can point out a general observation. First of all in this framework we consider a connected, directed and weighted network which is not necessarily symmetric and hence general theorems on operators or Laplacians on symmetric graphs do not apply. To achieve the goal for equilibrium in banking leverages of the network we must have the real parts of all the operator eigenvalues to be negative and at least one equal to zero. Furthermore, this observation can be true only within connected components in the network i.e. the steady state values may be different in every connected component.

### 4.5.4 Solutions to the former three cases without solving differential equations

## Review case 5.1

We are trying to find the leverages in equilibrium without solve any differential equation as we did in 5.1 up to 5.3 sections. Suppose we have the same graph (Graph 1) as was described in section 5.1. According to the Theorem 2.4 in equilibrium all the leverages are equal or in other words: $l_{1}=l_{2}=l_{3}$ or we get: $l_{1}-l_{2}=0, l_{2}-l_{3}=0$ and we have the following system:

$$
\begin{aligned}
\phi_{23} & =2 \phi_{12} \\
\phi_{23} & =4 / 3 \phi_{32} \\
\phi_{12} & =2 / 3 \phi_{32}
\end{aligned}
$$

or in matrix form we get:

$$
\left[\begin{array}{l}
\phi_{12} \\
\phi_{23} \\
\phi_{32}
\end{array}\right]=\phi_{32} \underbrace{\left[\begin{array}{c}
2 / 3 \\
4 / 3 \\
1
\end{array}\right]}_{\underset{\sim}{\mathbf{v}}}
$$

Thus, the solution to the system is: $\Lambda(\Sigma)=\left\{(2 / 3,4 / 3,1)^{T} \mid \phi_{32} \in \mathbb{R}\right\}$. Next, let $\lambda \equiv \phi_{32}$ and the total amount of transactions into bank network suppose that is equal to 90 monetary units. Then we have: $\lambda\left[\begin{array}{c}2 / 3 \\ 4 / 3 \\ 1\end{array}\right]=\left[\begin{array}{c}90 \\ 0 \\ 0\end{array}\right] \Rightarrow \lambda=30$. The new loan flows now are: $\phi_{32}=30, \phi_{23}=40, \phi_{12}=20$. We wrap them up all the results into next table. In Table 4.4

Table 4.4 Final leverages $\left(l_{i}\right)$ in equilibrium, taking account the new $\phi_{i j}$ per bank.

| $l_{i}$ | $\sum_{j} \phi_{i j} / k_{i}$ | initial values | equilibrium values |
| :---: | :---: | :---: | :---: |
| $l_{1}$ | $\phi_{12} / k_{1}$ | 5 | $20 / 2=10$ |
| $l_{2}$ | $\phi_{23} / k_{2}$ | $10 / 4$ | $40 / 4=10$ |
| $l_{3}$ | $\phi_{32} / k_{3}$ | $10 / 3$ | $30 / 3=10$ |

we conclude that in equilibrium all banks will have the same leverage and equals to 10 .

## Review case 5.2

Here we are try to resolve the case 5.3 without any use of the tools of differential equations. As was described in last subsection again here we have a particularity about the diffusion
process. If we look closely on the Graph 1 in case 5.3 we can see that the node 1,3 and 4 are nodes that only gives and takes from and to the node 2 , respectively. This is somehow paradox because if we look more closely on the graph we observe that the nodes 1,3 provide with loans to the node 2 but on the other hand they can't absorb any amount of loans from this node. More analytically and taking account the Theorem 2.4 suppose that: $l_{1}=l_{2}=l_{3}=l_{4} \Rightarrow l_{1}-l_{2}=0, l_{2}-l_{3}=0, l_{3}-l_{4}=0$.

$$
\begin{array}{rll}
\phi_{12} / k_{1}-\phi_{24} / k_{2} & =0 \Rightarrow \phi_{12} & =0 \\
\phi_{24} / k_{2}-\phi_{32} / k_{3} & =0 \Rightarrow \phi_{24} & =0 \\
\phi_{32} / k_{3} & =0 \Rightarrow \phi_{32} & =0
\end{array}
$$

Table 4.5 Final leverages $\left(l_{i}\right)$ in equilibrium, taking account the new $\phi_{i j}$ per bank.

| $l_{i}$ | $\sum_{j} \vec{\phi}_{i j} / k_{i}$ | initial values | equilibrium values |
| :---: | :---: | :---: | :---: |
| $l_{1}$ | $\phi_{12} / k_{1}$ | $1 / 2$ | 0 |
| $l_{2}$ | $\phi_{23} / k_{2}$ | $10 / 4$ | 0 |
| $l_{3}$ | $\phi_{32} / k_{3}$ | $10 / 3$ | 0 |
| $l_{4}$ | 0 | 0 | 0 |

According to Table 4.5 we conclude that in equilibrium all the banks will have the same leverage and equals to 0 .

### 4.6 Case study in a banking network

In this last section we study $\mathrm{G}(10,13)$, which has the characteristics of directed, connected and weighted network. Any node, $v_{i}$ from the set $V=\{1,2, \ldots, 10\}$ represents a bank. Furthermore we have a set E of directed edges, with elements $e_{i j}$, i.e. $E=\left\{e_{12}, e_{23}, \ldots, e_{109}\right\}$ , and weights $\phi_{i j}$ which are showing to us the loan(s) transfer of bank ito bank j , (with $i \neq j$ ) as depicted in the following network. Finally, the initial capital per bank are: $k_{1}=20, k_{2}=$ $30, k_{3}=20, k_{4}=20, k_{5}=50, k_{6}=100, k_{7}=30, k_{8}=40, k_{9}=30, k_{10}=100$.


Graph 1, $\mathbf{G}(10,13)$
So from the Graph 1, we can find the initial leverages, $l_{i}, i=\{1,2, \ldots, 10\}$ per bank (Table 4.6) and sort them in a descending order as Figure 2 shows.

Table 4.6 Initial leverages $\left(l_{i}\right)$ per bank.

| $l_{i}$ | $\sum_{j} \phi_{i j} / k_{i}$ | Results |
| :---: | :---: | :---: |
| $l_{1}$ | $\varphi_{12} / k_{1}$ | $1 / 2$ |
| $l_{2}$ | $\left(\varphi_{28}+\phi_{23}\right) / k_{2}$ | $4 / 3$ |
| $l_{3}$ | $\varphi_{34} / k_{3}$ | $3 / 2$ |
| $l_{4}$ | $\varphi_{45} / k_{4}$ | $1 / 2$ |
| $l_{5}$ | $\left(\varphi_{510}+\phi_{56}\right) / k_{5}$ | $3 / 5$ |
| $l_{6}$ | $\left(\varphi_{69}+\phi_{67}\right) / k_{7}$ | $3 / 5$ |
| $l_{7}$ | $\varphi_{71} / k_{7}$ | $1 / 3$ |
| $l_{8}$ | $\varphi_{87} / k_{8}$ | 1 |
| $l_{9}$ | $\varphi_{91} / k_{9}$ | 1 |
| $l_{10}$ | $\varphi_{109} / k_{10}$ | $1 / 2$ |

Next step is to write down the differential equations as the differences amongst leverages. In other words, we take the difference between two nodes which are connected by a directed edge as the final point minus the initial point as given by (48).

$$
\begin{align*}
\dot{\phi}_{12} & =l_{2}-l_{1}=\left(\phi_{28}+\phi_{23}\right) / k_{2}-\phi_{12} / k_{1} \\
\dot{\phi}_{23} & =l_{3}-l_{2}=\phi_{34} / k_{3}-\left(\phi_{28}+\phi_{23}\right) / k_{2} \\
\dot{\phi}_{28} & =l_{8}-l_{2}=\phi_{87} / k_{8}-\left(\phi_{28}+\phi_{23}\right) / k_{2} \\
\dot{\phi}_{34} & =l_{4}-l_{3}=\phi_{45} / k_{4}-\phi_{34} / k_{3} \\
\dot{\phi}_{45} & =l_{5}-l_{4}=\left(\phi_{510}+\phi_{56}\right) / k_{5}-\phi_{45} / k_{4} \\
\dot{\phi}_{510} & =l_{10}-l_{5}=\phi_{109} / k_{10}-\left(\phi_{510}+\phi_{56}\right) / k_{5} \\
\dot{\phi}_{56} & =l_{6}-l_{5}=\left(\phi_{69}+\phi_{67}\right) / k_{6}-\left(\phi_{510}+\phi_{56}\right) / k_{5} \\
\dot{\phi}_{69} & =l_{9}-l_{6}=\phi_{91} / k_{9}-\left(\phi_{69}+\phi_{67}\right) / k_{6} \\
\dot{\phi}_{67} & =l_{7}-l_{6}=\phi_{71} / k_{7}-\left(\phi_{69}+\phi_{67}\right) / k_{6} \\
\dot{\phi}_{71} & =l_{1}-l_{7}=\phi_{12} / k_{1}-\phi_{71} / k_{7} \\
\dot{\phi}_{87} & =l_{7}-l_{8}=\phi_{71} / k_{7}-\phi_{87} / k_{8} \\
\dot{\phi}_{91} & =l_{1}-l_{9}=\phi_{12} / k_{1}-\phi_{91} / k_{9} \\
\dot{\phi}_{109} & =l_{9}-l_{10}=\phi_{91} / k_{9}-\phi_{109} / k_{10} \tag{4.48}
\end{align*}
$$

The adjacency operator of Graph 1 is given by $\mathbf{A}_{a d j}$ :

$$
\mathbf{A}_{a d j}=\left[\begin{array}{cccccccccc}
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0
\end{array}\right]
$$

In the equilibrium, as the Theorem 2.2 says, we have all the leverages $l_{i}$ equals, i.e. $l_{1}=l_{2}=\cdots=l_{10}$ and we can write:

$$
\begin{equation*}
l_{1}-l_{2}=l_{2}-l_{3}=l_{3}-l_{4}=l_{4}-l_{5}=l_{5}-l_{6}=l_{6}-l_{7}=l_{7}-l_{8}=l_{8}-l_{9}=l_{9}-l_{10}=0 \tag{4.49}
\end{equation*}
$$

If we replace every leverage into (49), as described by (1) and rewrite them as operators we have:

$$
\underbrace{\left[\begin{array}{cccccccc}
1 / k_{1} & 0 & 0 & -1 / k_{3} & 0 & \cdots & 0 & 0 \\
0 & 1 / k_{2} & 1 / k_{2} & -1 / k_{3} & 0 & \cdots & 0 & 0 \\
0 & 0 & 0 & 1 / k_{3} & -1 / k_{4} & \cdots & 0 & 0 \\
0 & 0 & 0 & 0 & 1 / k_{4} & \cdots & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \ddots & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \cdots & -1 / k_{9} & 0 \\
0 & 0 & 0 & 0 & 0 & \cdots & 1 / k_{9} & -1 / k_{10}
\end{array}\right]}_{\mathbf{A}_{\mathbf{1}} \in \mathbb{R}^{9 \times 13}} \underbrace{\left[\begin{array}{c}
\phi_{12} \\
\phi_{23} \\
\phi_{28} \\
\phi_{34} \\
\phi_{45} \\
\phi_{56} \\
\phi_{510} \\
\vdots \\
\phi_{109}
\end{array}\right]}_{\Phi \in \mathbb{R}^{1 \times 13}}=\mathbf{0}
$$

By the above system we have the following equations-solutions for unknowns $\underset{i j}{ }{ }_{i j}$

$$
\begin{align*}
\phi_{12} & =\phi_{34} \\
\phi_{23} & =\phi_{91}-\phi_{28} \\
\phi_{34} & =\phi_{45} \\
\phi_{45} & =(2 / 3) \phi_{91} \\
\phi_{67} & =(10 / 3) \phi_{91}-\phi_{69} \\
\phi_{71} & =\phi_{91} \\
\phi_{109} & =(10 / 3) \phi_{91} \tag{4.50}
\end{align*}
$$

Moreover the base $\mathbf{V}$ with elements the vectors which are taken from the solution of the system (50) is:

$$
\mathbf{V}_{(13 \times 4)}=\left[\begin{array}{cccc}
2 / 3 & 0 & 0 & 0  \tag{4.51}\\
1 & -1 & 0 & 0 \\
0 & 1 & 0 & 0 \\
2 / 3 & 0 & 0 & 0 \\
2 / 3 & 0 & 0 & 0 \\
5 / 3 & 0 & -1 & 0 \\
0 & 0 & 1 & 0 \\
10 / 3 & 0 & 0 & -1 \\
0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 \\
4 / 3 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0
\end{array}\right]
$$

Suppose that we have $\phi_{91} \equiv k, \phi_{28} \equiv \lambda, \phi_{510} \equiv \mu$ and $\phi_{69} \equiv v$. Furthermore, we make a hypothesis that the total amount of courent loans in the interbank system must not exceed the value of 296 monetary units. Thus we found that $\mathrm{k}=24$. Therefore, now we are ready to calculate the $n e w{\underset{\sim}{i j}}^{\phi_{i j}}$ i.e. in equilibrium.

$$
\left[\begin{array}{c}
\phi_{12}  \tag{4.52}\\
\phi_{23} \\
\phi_{28} \\
\phi_{34} \\
\phi_{45} \\
\phi_{56} \\
\phi_{510} \\
\phi_{67} \\
\phi_{69} \\
\phi_{71} \\
\phi_{87} \\
\phi_{91} \\
\phi_{109}
\end{array}\right]=24\left[\begin{array}{c}
2 / 3 \\
1 \\
0 \\
2 / 3 \\
2 / 3 \\
5 / 3 \\
0 \\
10 / 3 \\
0 \\
1 \\
4 / 3 \\
1 \\
10 / 3
\end{array}\right]=\left[\begin{array}{c}
16 \\
24 \\
0 \\
16 \\
16 \\
40 \\
0 \\
80 \\
0 \\
24 \\
32 \\
24 \\
80
\end{array}\right]
$$

Therfore, the new, (final), leverages in the equilibrium to our interbank loan network are given by Table 4.7.

Table 4.7 Final leverages $\left(l_{i}\right)$ in equilibrium, taking account the new $\phi_{i j}(t)$ per bank.

| $l_{i}$ | $\sum_{j} \phi_{i j} / k_{i}$ | equilibrium results |
| :---: | :---: | :---: |
|  |  |  |
| $l_{1}$ | $\phi_{12} / k_{1}$ | $4 / 5$ |
| $l_{2}$ | $\left(\phi_{28}+\phi_{23}\right) / k_{2}$ | $4 / 5$ |
| $l_{3}$ | $\phi_{34} / k_{3}$ | $4 / 5$ |
| $l_{4}$ | $\phi_{45} / k_{4}$ | $4 / 5$ |
| $l_{5}$ | $\left(\phi_{510}+\phi_{56}\right) / k_{5}$ | $4 / 5$ |
| $l_{6}$ | $\left(\phi_{69}+\phi_{67}\right) / k_{7}$ | $4 / 5$ |
| $l_{7}$ | $\phi_{71} / k_{7}$ | $4 / 5$ |
| $l_{8}$ | $\phi_{87} / k_{8}$ | $4 / 5$ |
| $l_{9}$ | $\phi_{91} / k_{9}$ | $4 / 5$ |
| $l_{10}$ | $\phi_{109} / k_{10}$ | $4 / 5$ |

We can present all the previous analysis of leverages in next Figure 3, in order to compare the initial, mean and final value of leverage of our interbank network.

In Figure 4 our goal is twofold. The first one is to order the nodes taking account the initial leverage per bank and the second one is to order the edges of the interbank network taking acount the weights of any edge.

## Chapter 5

## The Notion of Almost Zeros and Randomness

'There is no royal road to geometry'.
— Euclides, 4th-3rd Centrury B.C., Greek Mathematician.

### 5.1 Introduction

We examine the notion of almost zero of polynomial vectors and matrices in terms of a minimization problem. The present work originates in the seminal paper [37] where the almost zeros were first defined as minima of a norm function of the polynomial vector $\mathbf{p}(s)$. In this paper basic properties such as the region of the complex plane which contains the prime almost zero (global minimum) were defined, and necessary and sufficient conditions characterizing the set of almost zeros were given. The role of almost zeros of $\mathbf{p}(s)$ in the zero assignment problem of the combinants of $\mathbf{p}(s)$ were also investigated. Finally location properties of almost zeros were given in terms of discs in the complex plane as well as how some of the properties of a zero extend naturally to the almost zero case.

From then on there is extended literature that relates almost zeros to coprimness and GCD of polynomials [34], [28] and other applications in system theory and control.

The purpose of the present chapter is twofold: (a) Firstly, to explore further the algebraic properties and computations of almost zeroes of polynomial vectors or matrices and present new results of higher order almost zeros; (b) Secondly, to relate the notion of almost zero to that of randomness i.e. what are the statistical characteristics of the almost zero and its norm polynomial value in a given random ensemble of polynomials. To this end we will consider
random ensembles of polynomial vectors and calculate various statistics related to the prime almost zeroes and present the findings.

### 5.2 Almost zeros of polynomial matrix

The notion of almost zero arises from the study of coprimeness properties of polynomial vectors and matrices. The zeros of polynomial matrices of composite MFD's of transfer functions as well as of the composite penciles of the type $[\mathbf{s I}-\mathbf{A}, \mathbf{B}]$ are related to controllability observability and feedback pole assignability or stabilization. The notion of almost zero is a generalization of the notion of zero of polynomial matrices. To define this concept we need to construct an indicator that captures the ordinary zero when this exists or returns a value that shows how close the polynomial matrix is to having such a zero. This can be done in many ways i.e. this indicator is not uniquely defined.

To define this indicator in a natural way we consider a static feedback scheme designed to change the dynamics of a system by pole placement. The closed loop polynomial can then be written as:

$$
P_{\mathbf{K}}(s)=<\mathbf{K}, \mathbf{p}(s)>
$$

where $\mathbf{K}$ is the controller vector and $\mathbf{p}(s)$ is a polynomial vector arising from an MFD. Then the almost zero may ne defined by the following min max problem:

The almost zero can then be defined as: $s_{0}=\arg \min _{s \in \mathbb{C}}\|\mathbf{p}(s)\|$ and by definition we have that the closed loop polynomial $\mathbf{p}_{K}(s)$ satisfies:

$$
\left\|\mathbf{p}_{K}\left(s_{0}\right)\right\| \leq\left\|\mathbf{p}\left(s_{0}\right)\right\|
$$

In other words the almost zero imposes an obstacle in pole placement by any feeback controller $K$. In case that $s_{0}$ is an ordinary zero then this frequence is fixed and cannot be removed by feedback. On account of this fact, we call $\left\|\mathbf{p}\left(s_{0}\right)\right\|$ the power-index or $\mathbf{p}$-index of the almost zero $s_{0}$. We also refer to $\left\|\mathbf{p}\left(s_{0}\right)\right\|^{2}$ as the $\mathbf{p}$-index ${ }^{2}$. Finally, we define $\left\|s_{0}\right\|^{2}$ as the size of the almost zero $s_{0}$. Consider a polynomial matrix $\mathbf{M}(s) \in \mathbb{R}[s]^{m \times p}, m>$ $p$ of rank p , assume also that the column degrees of $\mathbf{M}(s)$ are $n_{1} \leq n_{2} \leq \cdots \leq n_{p}$ with $n_{1}+n_{2}+\cdots+n_{p}=n$ where n is the total degree of this matrix denoted by $\operatorname{deg}(\mathbf{M}(\mathrm{s}))$. A
zero $s_{0} \in \mathbb{C}$ is defined to be a complex number such that:

$$
\operatorname{rank}\left(\mathbf{M}\left(s_{0}\right)\right)<p
$$

Furthermore, if we denote by $\mathbf{M}(\infty)$ the coefficient matrix corresponded to the highest degree monomials of the columns if

$$
\operatorname{rank}(\mathbf{M}(\infty))<p
$$

then we have zeros at infinity.
If $s_{1}, s_{2}, \ldots, s_{r}$ are all the (finite) zeros of $\mathbf{M}(s)$ then $\mathbf{M}(s)$ can be written as:

$$
\begin{aligned}
\mathbf{M}(s) & =\mathbf{M}^{\prime}(s) \mathbf{Q}(s), \mathbf{Q}(s) \in \mathbb{R}^{p \times p}[s] \\
\operatorname{det}(\mathbf{Q}(s)) & =\left(s-s_{1}\right)^{\lambda_{1}}\left(s-s_{2}\right)^{\lambda_{2}} \ldots\left(s-s_{r}\right)^{\lambda_{r}}=p_{0}(s)
\end{aligned}
$$

and $\lambda_{1}, \lambda_{2}, \ldots, \lambda_{r}$ are the multiplicities of the corresponding zeros $s_{1}, s_{2}, \ldots, s_{r}$. Furthermore, if $\mathbf{M}(s), \mathbf{M}^{\prime}(s)$ have no zeros at infinity we have that:

$$
\operatorname{deg}(\mathbf{M}(s))=\operatorname{deg}\left(\mathbf{M}^{\prime}(s)\right)+\sum_{i} \lambda_{i} .
$$

Additionally we have that:

$$
\left\|p_{0}(s)\right\|^{2} \text { divides the } \operatorname{det}\left(\mathbf{M}^{T}(s) \mathbf{M}(s)\right)
$$

where $\left\|p_{0}(s)\right\|^{2}=p_{0}(\bar{s}) p_{o}(s)$, (i.e. $\|\cdot\|$ is the Euclidean norm in $\mathbb{C}$ ) or

$$
\begin{aligned}
& \operatorname{det}\left(\mathbf{M}^{T}\left(\bar{s}_{i}\right) \mathbf{M}\left(s_{i}\right)\right)=0 \text { or } \\
& \left\|\operatorname{det}\left(\mathbf{M}^{T}\left(\overline{s_{i}}\right) \mathbf{M}\left(s_{i}\right)\right)\right\|=0
\end{aligned}
$$

If $s_{i}$ is a zero of $\mathbf{M}(s)$. In the case that there are no finite zeros then $\left\|\operatorname{det}\left(\mathbf{M}^{T}(\bar{s}) \mathbf{M}(s)\right)\right\| \neq$ $0, \forall s \in \mathbb{C}$ and in this case we define the notion of almost zero.

Definition 5.2.1. The prime almost zero $s_{0}$ is a complex number such that $\left|P_{\mathbf{M}}\left(\bar{s}_{0}, s_{0}\right)\right|=$ $\min _{s \in \mathbb{C}}\left|P_{\mathbf{M}}(\bar{s}, s)\right|$ where $P_{\mathbf{M}}(\bar{s}, s)=\operatorname{det}\left(\mathbf{M}^{T}(\bar{s}) \mathbf{M}(s)\right)$.

The existence of an exact zero i.e. a complex number such that $P_{\mathbf{M}}(\bar{s}, s)=0$ significe the fact that the polynomial matrix $\mathbf{M}(s)$ is not coprime. We define as GCD variety, the variety of all $\mathbf{M}(s)$ that are not coprime. If $s_{0}$ is the prime almost zero then $\left|P_{\mathbf{M}}\left(\bar{s}_{0}, s_{0}\right)\right|$ is a characteristic of $\mathbf{M}(s)$ that defines the distance from the GCD variety i.e. the smaller the $P_{\mathbf{M}}\left(\bar{s}_{0}, s_{0}\right)$ the closer to GCD variety is.

For a fixed $\mathbf{M}(s)$ the function $P_{\mathbf{M}}(\bar{s}, s)$ is a positive function from $\mathbb{C} \rightarrow \mathbb{R}^{+}$which acquires a global minimum. We can also write equivalently:

$$
P_{\mathbf{M}}(\bar{s}, s)=\prod_{i=1}^{p} \sigma_{i}^{2}(\mathbf{M}(s))
$$

where $\sigma_{i}(\mathbf{M}(s))$ are the singular values of the matrix $\mathbf{M}(s)$. If we also define by $C_{p}(\mathbf{M}(s))=$ $\left(p_{1}^{\mathbf{M}}(s), p_{2}^{\mathbf{M}}(s), \ldots, p_{k}^{\mathbf{M}}(s)\right)$ where $k=\binom{m}{p}$ and $C_{p}$ denotes the p-th compound of $\mathbf{M}(s)$ then:

$$
P_{\mathbf{M}}(\bar{s}, s)=\sum_{i}\left\|p_{i}^{\mathbf{M}}(s)\right\|^{2}
$$

by the Cauchy-Binet expansion of $\operatorname{det}\left(\mathbf{M}^{T}(s) \mathbf{M}(s)\right)$.
Theorem 5.2.1. If $\operatorname{rank}(\mathbf{M}(\infty))=p$ then $\operatorname{deg}\left(C_{p}(M)(s)\right)=n_{1}+n_{2}+\cdots+n_{p}$ and also

$$
P_{\mathbf{M}}(\bar{s}, s)=\sum_{i}\left[1, \bar{s}, \cdots, \bar{s}^{n}\right] a_{i} a_{i}^{t}\left[\begin{array}{c}
1 \\
s \\
\vdots \\
s^{n}
\end{array}\right]=\left[1, \bar{s}, \cdots, \bar{s}^{n}\right] \mathbf{T}\left[\begin{array}{c}
1 \\
s \\
\vdots \\
s^{n}
\end{array}\right]
$$

where $a_{i}^{t}$ is the coefficient vector of $p_{i}^{\mathbf{M}}(s)$ and $\mathbf{T}$ is an induced $(n+1) \times(n+1)$ self-adjoint matrix.

Theorem 5.2.2. If $\mathbf{T}=\mathbf{B}^{t} \Lambda \mathbf{B}$ is the Jordan decomposition of $\mathbf{T}$ where $\mathbf{B}$ is orthogonal and $\Lambda$ real diagonal $\Lambda=\operatorname{diag}\left(\lambda_{i}\right), \lambda_{i} \geq 0$ and $b_{i}$ are rows of $\mathbf{B}$. Then $P_{\mathbf{M}}(\bar{s}, s)=\sum_{i=1}^{n} \lambda_{i}\left|b_{i}(s)\right|^{2}$ where $b_{i}(s)=b_{i}\left(1, s, \cdots, s^{n}\right)^{t}$

Hence the coprimenes of a polynomial matrix $\mathbf{M}(s) \in \mathbb{R}[s]^{n \times p}$ of degree n is equivalent to the coprimeness of at most n -polynomials of degree n . Thus if we consider the set $\mathbf{M a t}=\left\{\mathbf{M}(s) \in \mathbb{R}[s]^{, m \times p}, m>p, \operatorname{deg}(\mathbf{M}(s))=n\right\}$ we have a function:

$$
\boldsymbol{\operatorname { M a t }}(n) \rightarrow\left\{\mathbf{T}: \mathbf{T} \in \mathbb{R}^{(n+1) \times(n+1)}, \mathbf{T}=\mathbf{T}^{t}, \mathbf{T} \geq 0\right\}
$$

and for a $\mathbf{M}(s) \in \operatorname{Mat}(n)$ its prime almost zero may be calculated by minimizing:

$$
\min _{s} \sum_{i=1}^{n} \lambda_{i}\left|\mathbf{b}_{i}(s)\right|^{2}
$$

So far we investigated the existence and calculation of almost zeros of a polynomial matrix $\mathbf{M}(s)$ i.e. complex number that almost drop the rank of $\mathbf{M}(s)$. We can generalize this notion to examine the size of the rank of $\mathbf{M}(s)$. To this end we utilize the identity:

$$
\operatorname{tr}\left(C_{k}\left(\mathbf{M}^{t}(\bar{s}) \mathbf{M}(s)\right)\right)=\operatorname{sym}_{k}\left(\sigma_{i}^{2}(\mathbf{M}(s))\right)
$$

where $\operatorname{sym}_{k}$ is the k -th symmetric function of p objects for $1 \leq k \leq p$. In the previous exposition we utilize $k=p$ where in this case $\operatorname{sym}_{p}$ was just the product of the squares of the singular values of $\mathbf{M}(s)$. For a given $\mathrm{k}: 1 \leq k \leq p$ we may define as:

$$
P_{\mathbf{M}}^{k}(\bar{s}, s)=\operatorname{tr}\left(C_{k}\left(\mathbf{M}^{t}(\bar{s}) \mathbf{M}(s)\right)\right)
$$

and as prime almost zero of degree k the complex number that minimizes $P_{\mathbf{M}}^{k}(\bar{s}, s)$. The polynomial $P_{\mathbf{M}}^{k}(\bar{s}, s)$ is the sum of squares of the norms of all $k \times k$ minors of $\mathbf{M}(s)$. If a complex number $s_{0}^{k}$ is a root of $P_{\mathbf{M}}^{k}(\bar{s}, s)$ then $\operatorname{rank}\left(\mathbf{M}\left(s_{0}^{k}\right)\right)<k$. This way for a given matrix $\mathbf{M}(s)$ we may define a set of almost zeros $s_{0}^{k}$ according to the examination of how much a complex number drops the rank of the matrix $\mathbf{M}(s)$. The calculation then of $s_{0}^{k}$ follows exactly the same steps as in ch. 2 as $P_{\mathbf{M}}^{k}(\bar{s}, s)$ has the same structure as $P_{\mathbf{M}}(\bar{s}, s)$.

In this respect, we have to minimize $P_{M}^{k}(\bar{s}, s)$ with respect to $s$. In all the above cases $P_{M}^{k}(\bar{s}, s)$ is written as:

$$
P_{M}^{k}(\bar{s}, s)=\|\mathbf{p}(s)\|^{2}
$$

where $\mathbf{p}(s)$ is an appropriately defined vector. To minimize $\|\mathbf{p}(s)\|^{2}$ we write (in the case where the coefficients of $\mathbf{p}(s)$ are real numbers) the F.O.C. of $\min _{s}\langle\mathbf{p}(\bar{s}), \mathbf{p}(s)\rangle$ with $s=x+i y$. Hence,

$$
\begin{aligned}
& \frac{\partial}{\partial x}\langle\mathbf{p}(\bar{s}), \mathbf{p}(s)\rangle=\left\langle\mathbf{p}(\bar{s})^{\prime}, \mathbf{p}(s)\right\rangle+\left\langle\mathbf{p}(\bar{s}), \mathbf{p}^{\prime}(s)\right\rangle=0 \\
& \frac{\partial}{\partial y}\langle\mathbf{p}(\bar{s}), \mathbf{p}(s)\rangle=-i\left\langle\mathbf{p}(\bar{s})^{\prime}, \mathbf{p}(s)\right\rangle+i\left\langle\mathbf{p}(\bar{s}), \mathbf{p}^{\prime}(s)\right\rangle=0
\end{aligned}
$$

where $\mathbf{p}^{\prime}(s)$ denotes the complex derivative with respect to $s$. Combining the equations, we take

$$
\left\langle\mathbf{p}^{\prime}(\bar{s}), \mathbf{p}(s)\right\rangle=0 .
$$

The last is the condition an almost zero must satisfy. (If $\mathbf{p}(s)$ has complex coefficients, then the above calculation remains valid by replacing $\mathbf{p}(\bar{s})$ with $\overline{\mathbf{p}^{\prime}(s)}$.

### 5.3 Example of prime almost zero calculation

We next present some examples of prime almost zero calculation in terms of minimization of $P_{M}(s, \bar{s})$, which can be done either directly or using numerical methods.

## Example 12:

Let $\overline{\mathbf{p}}(s)=\left(s-s_{1}, s-s_{2}, \ldots, s-s_{n}\right)^{t}$ then

$$
\begin{aligned}
\mathbf{p}(s) & =\left(s-s_{1}, s-s_{2}, \ldots, s-s_{n}\right) \\
\left\langle{\overline{\mathbf{p}(s)^{\prime}}}^{\prime}, \mathbf{p}(s)\right\rangle & =\sum_{i=1}^{n}\left(s-s_{i}\right)=0 \Rightarrow s=\frac{\sum_{i=1}^{n} s_{i}}{n} .
\end{aligned}
$$

Hence $s_{0}=\operatorname{Avg}\left(\left(s_{j}\right)_{j=1}^{n}\right)$

$$
s_{0}=\operatorname{Avg}\left(\left(a_{j}\right)_{j=1}^{n}\right)+i \operatorname{Avg}\left(\left(b_{j}\right)_{j=1}^{n}\right)=\operatorname{Avg}\left(\left(s_{j}\right)_{j=1}^{n}\right)
$$

where $\operatorname{Avg}\left(\left(c_{j}\right)_{j=1}^{n}\right)=\frac{\sum_{j=1}^{n} c_{j}}{n}$ denotes the average of the numbers $\left(c_{j}\right)_{j=1}^{n}$. Hence, the p -index ${ }^{2}$ of $p(s)$ is given by:

$$
\sum_{i}\left|s_{i}-\operatorname{Avg}\left(s_{i}\right)\right|^{2}
$$

which is the standard measure of dispersion of the roots $\left(s_{j}\right)_{j=1}^{n}$.

## Example 13:

To calculate the almost zeros of the polynomial matrix:

$$
\mathbf{M}(s)=\left[\begin{array}{cc}
s+1 & 0 \\
1 & s^{2} \\
0 & s \\
0 & 1
\end{array}\right]
$$

we calculate the 2 - compound of $\mathbf{M}(s)$ i.e. $\mathbf{C}_{2}(\mathbf{M}(s))^{t}=\left[\begin{array}{llllll}s^{3}+s^{2} & s^{2}+s & s+1 & s & 1 & 0\end{array}\right]$ and hence $\mathbf{T}$ is given by:

$$
\begin{aligned}
\mathbf{T}=\left[\begin{array}{llll}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 \\
0 & 0 & 1 & 1
\end{array}\right] & +\left[\begin{array}{llll}
0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 \\
0 & 1 & 1 & 0 \\
0 & 0 & 0 & 0
\end{array}\right]+\left[\begin{array}{llll}
1 & 1 & 0 & 0 \\
1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{array}\right]+\left[\begin{array}{llll}
0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{array}\right]+ \\
& {\left[\begin{array}{llll}
1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{array}\right]+\mathbf{0}=\left[\begin{array}{llll}
2 & 1 & 0 & 0 \\
1 & 3 & 1 & 0 \\
0 & 1 & 2 & 1 \\
0 & 0 & 1 & 1
\end{array}\right] }
\end{aligned}
$$

and the quality to be minimized is:

$$
\min _{s \in \mathbb{C}}\left(|s|^{6}+2|s|^{4}+3|s|^{2}+2+2 \mathfrak{R}(s)\left(1+|s|^{2}+|s|^{4}\right)\right)
$$

which is minimized for $s_{0}=-0.48$ and $P_{\mathbf{M}}\left(\bar{s}_{0}, s_{0}\right)=1.58$.

If we alternatively decompose $\mathbf{T}$ as $\mathbf{B}^{t} \Lambda \mathbf{B}$ then the eigenvalues in $\Lambda$ are:

$$
\{4.06,2.4,1.3,0.236\}
$$

and the quantity to be minimized is:

$$
\begin{aligned}
& p_{\mathbf{M}}(\bar{s}, s)=4.06\left\|0.38+0.79 s+0.46 s^{2}+0.149 s^{3}\right\|^{2}+ \\
& \quad+2.4\left\|-0.54-0.212 s+0.66 s^{2}+0.476 s^{3}\right\|^{2} \\
& \quad+1.3\left\|0.74-0.51 s+0.13 s^{2}+0.422 s^{3}\right\|^{2}+ \\
& \quad+0.236\left\|-0.15+0.26 s-0.58 s^{2}+0.76 s^{3}\right\|^{2}
\end{aligned}
$$

which has minimum as above.

## Example 14:

Suppose we have the following polynomial vector where $a_{i} \in \mathbb{R}$ :

$$
\mathbf{p}(s)=\left(s^{2}+a_{1}, s^{2}+a_{2}, \ldots, s^{2}+a_{n}\right) .
$$

Suppose we have the polynomial vector then:

$$
<\mathbf{p}^{\prime}(\bar{s}), \mathbf{p}(s)>=2\left(n|s|^{2} s+\bar{s}\left(\sum_{i} a_{i}\right)\right)=0
$$

Hence, $s_{0}^{2}=-\left(a_{i}\right)$ which implies that the absolute almost zero is:

$$
s_{0}= \pm i \sqrt{E\left(a_{j}\right)}
$$

### 5.4 Statistical results

In this study we consider a polynomial vector random ensembles of the form:

$$
\mathbf{P}_{k, n}(s)=\left[\begin{array}{c}
a_{1 n} s^{n}+a_{1 n-1} s^{n-1}+\cdots+a_{10} \\
a_{2 n} s^{n}+a_{2 n-1} s^{n-1}+\cdots+a_{20} \\
a_{3 n} s^{n}+a_{3 n-1} s^{n-1}+\cdots+a_{30} \\
\vdots \\
a_{k n} s^{n}+a_{k n-1} s^{n-1}+\cdots+a_{k 0}
\end{array}\right]
$$

where $\mathbf{P}_{k, n}$ is a polynomial rectangular matrix $k \times 1$ of degree $n$ and the coefficients $a_{i j}$ are i.i.d. $\sim \mathscr{U}(-1,1)$. It is worth noting that the polynomials are not monic as one expects
and if they are normalized the resulting ratio $a_{i j} / a_{i n}$ follows a distribution with probability distribution function:

$$
f(x)= \begin{cases}1 / 4 & |x| \leq 1 \\ 1 /\left(4 x^{2}\right) & |x|>1\end{cases}
$$

We carried out simulations for the ensembles $\mathbf{P}_{k, n}$ for $n=1,2,3,4$ and $k=1,2, \ldots, 25$ and created 2000 scenarios for every pair $(k, n)$. For every scenario $\mathbf{P}_{k, n}$ we calculated via a minimization process the almost zero $s_{0}^{j}$ and the minimum value $P_{M}\left(s_{0}, \overline{s_{0}}\right)$ of the objective function $P_{M}(s, \bar{s})$.

Since $s$ and $a_{i j}$ are random variables, it follows that the almost zero and the p-index ${ }^{2}$ are random variables. Therefore, our purpose is to calculate the mean and the variance of the p-index ${ }^{2}, P_{M}\left(s_{0}, \overline{s_{0}}\right)$ and of the size $\left\|s_{0}\right\|^{2}$ of the almost zero $s_{0}$. The results are depicted in the Figures: 5, 6, 7 and 8.

We observe that the p-index increases linearly with the size of the polynomial vector and independently to the polynomial degree as we can see in Figure 5. Next in Figure 6 we observe that the standard deviation of the p-index is strictly increasing as k increases from 1 to 25 and it is independent to the degrees of polynomials. In addition in Figure 7 the mean of the size of almost zero of polynomials is strictly decreasing and with a rapid decrease in the range of polynomials with size 1-5 with the graph to flatten thereafter, this phenomenon is more evident for the degree $n=1$. Furthermore, the standard deviation of the size of the almost zero is strictly concave for all polynomials degree as the size the polynomial vector k increases from 1 to 25 and only the graph corresponding to the polynomial of degree 1 is decreasing more rapidly than the graphs of the others and this is true up to $\mathrm{k}=11$ as depicted in Figure 8.

Moreover the distribution function of the p-index ${ }^{2} P_{M}\left(\overline{s_{0}}, s_{0}\right)$ for ensembles of polynomials of degree 1 up to 4 belongs to log-normal family distribution. The last can be observed by Figures 9 a up to 9 d . In addition Figure 13 shows us the distribution that best fits is the Pearson 6 distribution and on the other hand according to Anderson Darling and Chi-Squared criteria the distribution that best fits is the Johnson SB and Beta distributions, respectively. Furthermore the Figures from 14a up to 14 c depict the distributions of the p-index ${ }^{2}$ which for polynomial degree 1-3 have the same hyperbolic shape with the exception of Figure 14d. This is a Pearson Type 6 distribution and according to Kolmogorov Smirnov, Pearson 6 states is the best fit among the others.

In addition to this statistical analysis of polynomials as described before there exist a remarkable fact. According to Figures (5) - (8) both the mean and the variance of the p-index ${ }^{2}$ $P_{M}\left(\overline{s_{0}}, s_{0}\right)$ increase as the size of the ensemble $M$ increases and therefore the distribution of $P_{M}\left(\overline{s_{0}}, s_{0}\right)$ flattens with the size of $M$ (see Figure 11). On the contrary, the mean and the
variance of the size $\left\|s_{0}\right\|^{2}$ of almost zero $s_{0}\left\|s_{0}\right\|^{2}$ decrease with the size of $M$. Hence, the corresponding distribution sharpens with the size of $M$ (Figure 12).

This observation reveals that there is some kind of uncertainty in the determination of the size and p-index ${ }^{2}$ of almost zeros in the following sense: while the mean and the variance of the size $\left\|s_{0}\right\|^{2}$ of $s_{0}$ converge to 0 , as the size of $M$ increases, the corresponding magnitudes of $p$-index ${ }^{2}$ converge to infinity. Roughly speaking, although we can say that the value of $s_{0}$ is close to 0 , it is not possible to determine in any way the value of the p -index ${ }^{2}, P_{M}\left(\overline{s_{0}}, s_{0}\right)$. We use the term uncertainty principle to describe this phenomenon. However, we should point out that this term usually means uncertainty in two ways: when one determines the value of one quantity, then the value of the other becomes vague. In our framework, this only applies in one direction.

If we assume (as it is observed) that this uncertainty relation depends only on size and not on degree of the polynomials vectors, we may formulate the uncertainty principle in terms of polynomial vectors of degree 1, i.e. let

$$
\mathbf{p}(s)=\left(a_{i 1} s+a_{i 2}\right)_{i=1}^{k} .
$$

Then, $\left\langle\mathbf{p}^{\prime}\left(s_{0}\right), \mathbf{p}\left(s_{0}\right)\right\rangle=0$ implies that $\left\langle\mathbf{a}_{1}, \mathbf{a}_{1} s_{0}+\mathbf{a}_{2}\right\rangle=0$ and therefore

$$
s_{0}=-\frac{\left\langle\mathbf{a}_{1}, \mathbf{a}_{2}\right\rangle}{\left\|\mathbf{a}_{1}\right\|^{2}}
$$

Hence,

$$
\begin{aligned}
\left\|\mathbf{p}\left(s_{0}\right)\right\|^{2} & =\left\|\mathbf{a}_{1}\right\|^{2}\left\|s_{0}\right\|^{2}+\left\|\mathbf{a}_{2}\right\|^{2}+2\left\langle\mathbf{a}_{1}, \mathbf{a}_{2}\right\rangle s_{0} \\
& =\frac{\left\langle\mathbf{a}_{1}, \mathbf{a}_{2}\right\rangle^{2}}{\left\|\mathbf{a}_{1}\right\|^{2}}+\left\|\mathbf{a}_{2}\right\|^{2}-2 \frac{\left\langle\mathbf{a}_{1}, \mathbf{a}_{2}\right\rangle^{2}}{\left\|\mathbf{a}_{1}\right\|^{2}} \\
& =\left\|\mathbf{a}_{2}\right\|^{2}-\frac{\left\langle\mathbf{a}_{1}, \mathbf{a}_{2}\right\rangle}{\left\|\mathbf{a}_{1}\right\|^{2}}=\|\mathbf{p}(0)\|^{2}+s_{0} \\
\left\|\mathbf{p}\left(s_{0}\right)\right\|^{2} & =\|\mathbf{p}(0)\|^{2}+s_{0}
\end{aligned}
$$

where, $\left\|\mathbf{a}_{2}\right\|^{2}=\|\mathbf{p}(0)\|^{2}$ and $-\frac{\left\langle\mathbf{a}_{1}, \mathbf{a}_{2}\right\rangle}{\left\|\mathbf{a}_{1}\right\|^{2}}=+s_{0}$. Moreover, the numerator of the LHS of the last equality is approach to zero, due to inner product between of two random vectors $\mathbf{a}_{1}, \mathbf{a}_{2}$ and the denominator $\left\|\mathbf{a}_{1}\right\|^{2}$ goes to infinity.

This last relation highlights the fact that, as $k \rightarrow \infty$, then the mean value and the variance of $s_{0}$ converge to 0 , while the mean value and variance of the p-index, $\left\|\mathbf{p}\left(s_{0}\right)\right\|$ converge to $\infty$. This is proven next.

Theorem 5.4.1. For fixed $n$, assume that for each $k \geq 1, s_{k}$ denotes the almost zero of the multivariable random vector $\mathbf{P}_{k, n}(s)$. Then with probability 1 it holds $\lim _{k \rightarrow \infty} s_{k}=0$.

Proof. Let $\mathbf{p}_{k}(s)=\sum_{j=0}^{n} \mathbf{a}_{j} s^{j}$ where $\mathbf{a}_{j}$ is the $j$-th column of $P_{k, n}$. Since $\left\|\mathbf{p}_{k}(s)\right\|^{2}$ is a non negative continuously differentiable function of $x$ and $y(s=x+i y)$ with $\lim _{s \rightarrow \infty}\left\|\mathbf{p}_{k}(s)\right\|^{2}=$ $\infty$, it attains a minimum value at a finite point. If we define

$$
f_{k}(s):=\left\langle\mathbf{p}_{k}^{\prime}(s), \mathbf{p}_{k}(s)\right\rangle=\sum_{j=1}^{n} j| | \mathbf{a}_{j} \|^{2}|s|^{2 j-2} s+\sum_{\substack{1 \leq j, r \leq n \\ j \neq r}} j\left\langle\mathbf{a}_{j}, \mathbf{a}_{r}\right\rangle \bar{s}^{j-1} s^{r},
$$

then any point $\hat{s}$ where this minimum is attained satisfies $f_{k}(\hat{s})=0$. Note that $f_{k}$ is a polynomial in $s, \bar{s}$ of degree $2 n-1$ with leading term $n\left\|a_{n}\right\|^{2}|s|^{2 n-2} s$. The law of large numbers implies that with probability one, for each $j, r=1,2, \ldots, n$ with $r \neq j$, it holds

$$
\begin{align*}
\lim _{k \rightarrow \infty} \frac{1}{k}\left\|\mathbf{a}_{j}\right\|^{2} & =\mathbf{E}\left(a_{11}^{2}\right)=\frac{1}{3},  \tag{5.1}\\
\lim _{k \rightarrow \infty} \frac{1}{k}\left\langle\mathbf{a}_{j}, \mathbf{a}_{r}\right\rangle & =0 \tag{5.2}
\end{align*}
$$

It is then easy to show that, because of the leading term, for almost all $\omega$ in the probability space, there are $k_{0}(\omega) \in \mathbb{N}$ and $M(\omega)>0$ such that $\left|f_{k}(s)\right|>1$ for all $s$ with $|s|>M(\omega)$ and $k \geq k_{0}(\omega)$.

Now for each $k \geq 1$, let $s_{k}$ be any point satisfying $f_{k}\left(s_{k}\right)=0$. We will show that $\lim _{k \rightarrow \infty} s_{k}=0$ with probability 1 . For almost all $\omega$ in the probability space and for $k \geq k_{0}(\omega)$ we have $\left|s_{k}\right| \leq M(\omega)$. It is enough then to show that all convergent subsequences of $\left(s_{k}\right)_{k \geq 1}$ converge to 0 . Let $\left(s_{k_{v}}\right)_{v \geq 1}$ be such a subsequence that converges to a number $c$. Using (5.1) and (5.2) we get that

$$
0=\lim _{v \rightarrow \infty} \frac{1}{k_{v}} f_{k_{v}}\left(s_{k_{v}}\right)=\frac{1}{3} c \sum_{j=1}^{n} j|c|^{2 j-2} .
$$

Consequently, $c=0$.

### 5.5 Conclusions

We defined the notion of almost zeros of polynomial matrices which was generalized in various degrees. This is related to notions such as almost controllability and observability as well as almost McMillan degree and pole zero assignability. Furthermore we investigate the statistical properties of almost zeros in the light of randomness of elements of matrices and we prove an uncertainty type result for the determination of almost zeros.

## Chapter 6

## Random Matrices and Controllability of Dynamical Systems

'Let no-one ignorant of geometry enter'.

- Plato, 427-347 B.C., Greek Philosopher.


### 6.1 Introduction

Controllability is one of the most fundamental concepts in systems theory and control theory. Roughly speaking, a system is controllable if one can switch from one trajectory to the other provided that the laws governing the system are obeyed and some delay is allowed. In the present work, we focus on the linear, time-invariant multivariable system described by the following equations (see [39], [63]):

$$
\begin{equation*}
\frac{d \mathbf{x}}{d t}=\mathbf{A x}+\mathbf{b u}, \quad \mathbf{y}=\mathbf{C x}+\mathbf{d u} \tag{6.1}
\end{equation*}
$$

where $\mathbf{x}=x(t)=\left(x_{1}(t), x_{2}(t), \ldots, x_{n}(t)\right)^{t} \in \mathbb{R}^{n}$ is a vector describing the state of the system at time $t, \mathbf{u}=u(t) \in \mathbb{R}$ is the input and $\mathbf{y}=y(t)$ is the $m$-vector of outputs. A, C are respectively $n \times n, m \times n$ matrices, and $\mathbf{b} \in \mathbb{R}^{n}, \mathbf{d} \in \mathbb{R}^{m}$ are vectors. In this case the definition of controllability goes as follows. The system (6.1) is said to be state controllable or simply controllable, if there exists a finite time $T>0$, such that for any initial state $x(0) \in \mathbb{R}^{n}$ and any $x_{1} \in \mathbb{R}^{n}$, there is an input $\mathbf{u}=u(t)$ defined on $[0, T]$ that will transfer $x(0)$ to $x_{1}$ at time $T$ (i.e. $x(t)$ obeys the first equation of (6.1) and $x(T)=x_{1}$ ). Otherwise, the system (6.1) is called uncontrollable.

In this chapter, we consider random systems of the form (6.1), that are systems where the parameters $\mathbf{A}$ and $\mathbf{b}$ have been replaced with random matrices. Given a positive number
$\varepsilon$, we define the concept of $\varepsilon$-uncontrollability of a random system. It is natural that the $\varepsilon$-uncontrollability of a random system depends on the distribution of the entries of $\mathbf{A}$ and b. Consequently, we consider the fundamental Gaussian orthogonal ensemble of random matrices and we calculate the $\varepsilon$-uncontrollability in this particular case.

The rest of the chapter is organised as follows. In Section 6.2, we define the $\varepsilon$ uncontrollability for random systems. In Section 6.3, we describe the Gaussian orthogonal ensemble, which is going to be used in this work, and we state some known results about this ensemble that will be used in the sequel. In Section 6.4, we consider the case $n=2$, i.e. when the state space is $\mathbb{R}^{2}$, and we give the detailed calculation of the $\varepsilon$-uncontrollability of a random system where the matrix A comes from the Gaussian orthogonal ensemble. Finally, in Section 6.5, we deal with the general case of $\mathbb{R}^{n}$ for $n>2$. In this case, the situation is more complicated and we provide an upper bound for the $\varepsilon$-uncontrollability of a random system.

We close this introductory section with some comments concerning the motivation of the present work. Random matrix theory applies in problems of algebraic control in the presence of error or noise due to many factors. Here, we ignore systematic errors and we consider the effect of randomness to algebraic relations (equalities). For example, suppose that a set of equations $\mathbf{f}(\mathbf{x})=0$ is given. In order to study the effect of a random perturbation $\mathbf{x}+\varepsilon \mathbf{p}$, where $\mathbf{x}$ is a given solution and $\mathbf{p}$ is a random variable, we expand $f(\mathbf{x}+\mathbf{p})$ into

$$
\begin{equation*}
\mathbf{f}(\mathbf{x})+\varepsilon \mathbf{f}_{1}+\varepsilon^{2} \mathbf{f}_{2}+\ldots \tag{6.2}
\end{equation*}
$$

and we examine the statistical properties of $\mathbf{f}_{1}, \mathbf{f}_{2}$ as random variables.
In the case of controllability, we have the equations

$$
\begin{equation*}
\left[\left(s+\varepsilon s_{r}\right) \mathbf{I}-\mathbf{A}-\varepsilon \mathbf{A}_{r},-\mathbf{b}-\varepsilon \mathbf{b}_{r}\right]\left(\mathbf{x}+\varepsilon \mathbf{x}_{r}\right)=0 \tag{6.3}
\end{equation*}
$$

and we examine the statistical properties of the second-order term $\left[s_{r} \mathbf{I}-\mathbf{A}_{r}, \mathbf{b}_{r}\right] \mathbf{x}_{r}$, assuming that the properties of the first-order perturbation have already been resolved.

This study is to provide a statistical effect of such random perturbations on a system of a given model description as far as the controllability is concerned.

## 6.2 $\varepsilon$-uncontrollability of random systems

In order to formulate a suitable concept of uncontrollability for random systems, we utilize a characterisation of the controllability of systems of the form (6.1). This is provided in the next theorem (for more details see, for example, [39] Theorem 2.2).

Theorem 6.2.1. The following are equivalent.

1. The system (6.1) is controllable.
2. The matrix $[s \mathbf{I}-\mathbf{A},-\mathbf{b}]$ has full rank (i.e. rank $n$ ), for every $s \in \mathbb{C}$.

Additionally, we need the next lemma whose proof is based on elementary linear algebra and thus it is omitted.

Lemma 6.2.2. Let $\mathbf{A}$ be an $n \times n$-matrix and $\mathbf{b} \in \mathbb{R}^{n}$. Then, the following are equivalent.

1. The matrix $[s \mathbf{I}-\mathbf{A},-\mathbf{b}]$ has full rank (i.e. rank $n$ ), for every $s \in \mathbb{C}$.
2. There is no eigenvector $\mathbf{v} \in \mathbb{R}^{n}$ of the matrix $\mathbf{A}$ such that $\langle\mathbf{v}, \mathbf{b}\rangle=0$.

Motivated by the above results, we now define the $\varepsilon$-uncontrollability for random systems.
Definition 6.2.1. Assume that a random system is given:

$$
\begin{equation*}
\frac{d \mathbf{x}}{d t}=\mathbf{A x}+\mathbf{b u} \tag{6.4}
\end{equation*}
$$

where $\mathbf{A}$ is an $n \times n$ random matrix and $\mathbf{b}$ is an n-dimensional random vector. Given a positive number $\varepsilon$, the $\varepsilon$-uncontrollability of the above system is defined to be the probability:

$$
P_{\varepsilon}=\mathbb{P}\left(|\langle\mathbf{v}, \mathbf{b}\rangle|<\varepsilon: \text { for some eigenvector } \mathbf{v} \text { of } \mathbf{A} \text { with }\|\mathbf{v}\|_{2}=1\right) .
$$

### 6.3 Random matrix ensemble

It is quite evident that the measure of $\varepsilon$-uncontrollability $P_{\varepsilon}$ of the random system (6.4) depends on the distribution of the matrix $\mathbf{A}$ and the vector $\mathbf{b}$. In this chapter, we consider one important ensemble of real symmetric random matrices, namely the so-called GOE. On account of its applications, GOE is one of the most studied random matrix ensembles. It is placed in the more general framework of Wigner matrices, which are defined as follows. We consider $\xi, \zeta$ real-valued random matrices with zero mean. Let $\mathbf{W}=\left(w_{i j}\right)_{i, j=1}^{n}$ be a random symmetric matrix. We call $\mathbf{W}$ a Wigner matrix if its entries satisfy the next conditions:

- $\left\{w_{i j} \mid 1 \leq i \leq j \leq n\right\}$ are independent random variables;
- $\left\{w_{i j} \mid 1 \leq i<j \leq n\right\}$ are i.i.d. (independent, identically distributed) copies of $\xi$;
- $\left\{w_{i i} \mid i=1, \ldots, n\right\}$ are i.i.d. copies of $\zeta$.

The case of Wigner matrices in which $\xi$ and $\zeta$ are Gaussian with $\mathbb{E}\left[\xi^{2}\right]=1$ and $\mathbb{E}\left[\zeta^{2}\right]=2$ gives the GOE. Hence, if the symmetric matrix $\mathbf{W}$ belongs to GOE, then $w_{i i} \sim N(0,2)$ (for all $i=1, \ldots, n$ ), $w_{i j} \sim N(0,1)$ (for all $1 \leq i<j \leq n$ ) and the entries on and above the diagonal are independent random variables. (We write $\operatorname{GOE}(n)$ when an emphasis on the dimension is necessary. However, in majority of cases the dimension will be clear from the context.)

Additionally, as far as the random vector $\mathbf{b}$ is concerned, we have to choose some ensemble. More specifically, we consider the ensemble $\mathbf{S}_{b}$ containing all random vectors $\mathbf{b}=$ $\left(b_{1}, b_{2}, \ldots, b_{n}\right)$ such that $\left(b_{i}\right)_{i=1}^{n}$ are independent Gaussian random variables with zero mean and $\mathbb{E}\left[b_{i}^{2}\right]=1$. Furthermore, we assume that $\left(b_{i}\right)_{1 \leq i \leq n}$ and $\left(w_{i j}\right)_{1 \leq i \leq j \leq n}$ are all independent random variables.

For more information concerning Wigner matrices and the GOE we refer to [4]. For our purpose, we need a couple of result for the eigenstructure of the GOE, which are stated here without proof. First of all, it is known that a.s., the eigenvalues of a matrix $\mathbf{A}$ from GOE are all distinct (see [4], Theorem 2.5.2). Let now $v_{1}, \ldots, v_{n}$ denote the eigenvectors corresponding to the (real) eigenvalues of $\mathbf{A}$, with their first non zero entry positive real. Then the following proposition holds (see [4], Corollary 2.5.4).

Proposition 6.3.1. The collection $\left(v_{1}, v_{2}, \ldots, v_{n}\right)$ is independent of the eigenvalues. Each of the eigenvectors $v_{1}, v_{2}, \ldots, v_{n}$ is distributed uniformly on

$$
\mathbf{S}_{+}^{n-1}=\left\{\mathbf{x}=\left(x_{1}, x_{2}, \ldots, x_{n}\right) \mid x_{i} \in \mathbb{R},\|\mathbf{x}\|_{2}=1, x_{1}>0\right\}
$$

### 6.4 The case $n=2$

This section is entirely devoted to the calculation of the $\varepsilon$-uncontrollability of a random system of the form (6.4) when the state space is $\mathbb{R}^{2}$ and $\mathbf{A}, \mathbf{b}$ belong to GOE and $\mathbf{S}_{b}$ respectively. In order to achieve this goal, we firstly fix a vector $\mathbf{b} \in \mathbb{R}^{2}$ and we set

$$
P_{\varepsilon, \mathbf{b}}=\mathbb{P}\left(|\langle\mathbf{v}, \mathbf{b}\rangle|<\varepsilon: \text { for some eigenvector } \mathbf{v} \text { of } \mathbf{A},\|\mathbf{v}\|_{2}=1\right) .
$$

Then, the following result holds.
Theorem 6.4.1. Let the random system (6.4) be given, where the state space is $\mathbb{R}^{2}$ and $\mathbf{A}$ belongs to GOE(2). Then, for every non zero vector $\mathbf{b} \in \mathbb{R}^{2}$, we have

$$
P_{\varepsilon, \mathbf{b}}= \begin{cases}\frac{4}{\pi} \cdot \arcsin \left(\frac{\varepsilon}{\|\mathbf{b}\|_{2}}\right), & \text { if } \frac{\varepsilon}{\|\mathbf{b}\|_{2}}<\frac{\sqrt{2}}{2} \\ 1, & \text { if } \frac{\varepsilon}{\|\mathbf{b}\|_{2}} \geq \frac{\sqrt{2}}{2}\end{cases}
$$

Proof. Since the matrix $\mathbf{A}$ from $\operatorname{GOE}(2)$ is symmetric, there is an orthonormal basis $\left\{\mathbf{v}_{\mathbf{1}}, \mathbf{v}_{\mathbf{2}}\right\}$ of $\mathbb{R}^{2}$ consisting of eigenvectors of $\mathbf{A}$. Without loss of generality (replacing $\mathbf{v}_{1}$ with $-\mathbf{v}_{1}$ or changing the order of $\left\{\mathbf{v}_{\mathbf{1}}, \mathbf{v}_{\mathbf{2}}\right\}$, if necessary), we may assume that the first coordinate of $\mathbf{v}_{1}$ is positive. Hence, we can write $\mathbf{v}_{1}=(\cos \theta, \sin \theta)$ and $\mathbf{v}_{2}=(-\sin \theta, \cos \theta)$, for some $\theta \in\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. Now, we have:

$$
\begin{aligned}
P_{\varepsilon, \mathbf{b}} & =\mathbb{P}\left(|\langle\mathbf{v}, \mathbf{b}\rangle|<\varepsilon: \text { for some eigenvector } \mathbf{v} \text { of } \mathbf{A} \text { with }\|v\|_{2}=1\right) \\
& =\mathbb{P}\left(\left|\left\langle\mathbf{v}_{i}, \mathbf{b}\right\rangle\right|<\varepsilon: \text { for some eigenvector } \mathbf{v}_{i} \text { of } \mathbf{A}, i=1,2\right) .
\end{aligned}
$$

The non-zero vector $\mathbf{b}$ is written in the form $\mathbf{b}=\|\mathbf{b}\|_{2}\left(\frac{b_{1}}{\|\mathbf{b}\|_{2}}, \frac{b_{2}}{\|\mathbf{b}\|_{2}}\right)$. Let $T$ be the rotation through a suitable angle $\varphi$ such that $T\left(\frac{b_{1}}{\|\mathbf{b}\|_{2}}, \frac{b_{2}}{\|\mathbf{b}\|_{2}}\right)=e_{2}=(0,1)$. Then, $T$ is in orthogonal group $O(2)$ and hence,

$$
|\langle\mathbf{v}, \mathbf{b}\rangle|=|(\mathbf{v}), T(\mathbf{b})\rangle\left|=\|\mathbf{b}\|_{2} \cdot\right|\left\langle T(\mathbf{v}), e_{2}\right\rangle \mid
$$

for any eigenvector $\mathbf{v}=\mathbf{v}_{i}, i=1,2$ of $\mathbf{A}$. Thus,

$$
\begin{aligned}
P_{\varepsilon, \mathbf{b}} & =\mathbb{P}\left(\|\mathbf{b}\|_{2} \cdot\left|\left\langle e_{2}, T\left(\mathbf{v}_{i}\right)\right\rangle\right|<\varepsilon: \text { for } i=1 \text { or } i=2\right) \\
& =\mathbb{P}\left(\left|\left\langle e_{2}, T\left(\mathbf{v}_{i}\right)\right\rangle\right|<\frac{\varepsilon}{\|\mathbf{b}\|_{2}}: \text { for } i=1 \text { or } i=2\right) .
\end{aligned}
$$

The rotation $T$ is given by the following matrix representation

$$
T=\left[\begin{array}{cc}
\cos \varphi & -\sin \varphi \\
\sin \varphi & \cos \varphi
\end{array}\right] .
$$

Therefore,

$$
T\left(\mathbf{v}_{1}\right)=(\cos (\varphi+\theta), \sin (\varphi+\theta)) \quad \text { and } \quad T\left(\mathbf{v}_{2}\right)=(-\sin (\varphi+\theta), \cos (\varphi+\theta)),
$$

and, consequently,

$$
P_{\varepsilon, \mathbf{b}}=\mathbb{P}\left(|\sin (\varphi+\theta)|<\frac{\varepsilon}{\|\mathbf{b}\|_{2}} \text { or }|\cos (\varphi+\theta)|<\frac{\varepsilon}{\|\mathbf{b}\|_{2}}\right) .
$$

We have to distinguish two cases.
Case I If $\frac{\varepsilon}{\|\mathbf{b}\|_{2}} \geq \frac{\sqrt{2}}{2}$, then, clearly, one has $P_{\varepsilon, \mathbf{b}}=1$.

Case II If $\frac{\varepsilon}{\|\mathbf{b}\|_{2}}<\frac{\sqrt{2}}{2}$, then

$$
P_{\varepsilon, \mathbf{b}}=\mathbb{P}\left(|\sin (\varphi+\theta)|<\frac{\varepsilon}{\|\mathbf{b}\|_{2}}\right)+\mathbb{P}\left(|\cos (\varphi+\theta)|<\frac{\varepsilon}{\|\mathbf{b}\|_{2}}\right)
$$

where the equality follows from the disjointness of the two sets. Now, for the first summand, we observe that $\varphi+\theta$ belongs to a semicircle. Hence, the values of $\varphi+\theta$ for which we have $|\sin (\varphi+\theta)|<\frac{\varepsilon}{\|\boldsymbol{b}\|_{2}}$ belong to an arc or to the unions of two disjoint arcs whose length is $2 \cdot \arcsin \left(\frac{\varepsilon}{\|\mathbf{b}\|_{2}}\right)$. It follows that $\theta$ belongs either to an arc or to the union of two disjoint arcs with total length $2 \cdot \arcsin \left(\frac{\varepsilon}{\|\mathbf{b}\|_{2}}\right)$. By Proposition 6.3.1, $\theta$ is a random variable with the uniform distribution on the interval $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. Therefore,

$$
\mathbb{P}\left(\left.|\sin (\varphi+\theta)|<\frac{\varepsilon}{\|\mathbf{b}\|_{2}} \right\rvert\, \theta \in\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)\right)=\frac{2}{\pi} \cdot \arcsin \left(\frac{\varepsilon}{\|\mathbf{b}\|_{2}}\right) .
$$

Using similar argumentation for the second summand, we finally obtain:

$$
P_{\varepsilon, \mathbf{b}}=\frac{4}{\pi} \cdot \arcsin \left(\frac{\varepsilon}{\|\mathbf{b}\|_{2}}\right) .
$$

We are now ready to prove the main result of this section.
Theorem 6.4.2. Assume that $n=2$ and that $\mathbf{A}, \mathbf{b}$ belong to $G O E$ and $\mathbf{S}_{\mathbf{1}}, \mathbf{v}_{\mathbf{b}}$ respectively. For any positive number $\varepsilon$, the $\varepsilon$-uncontrollability of the random system (6.4) is given by

$$
P_{\varepsilon}=1-e^{2 \varepsilon^{2}}+\frac{4}{\sqrt{2} \pi} \int_{2 \varepsilon^{2}}^{\infty} \arcsin \left(\frac{\varepsilon}{\sqrt{x}}\right) e^{-x / 2} d x
$$

Proof. Let $\left\{\mathbf{v}_{1}, \mathbf{v}_{2}\right\}$ be an orthonormal basis of $\mathbb{R}^{2}$ consisting of eigenvectors of A. We set

$$
Z=\min \left\{\left|\left\langle\mathbf{v}_{1}, \mathbf{b}\right\rangle\right|,\left|\left\langle\mathbf{v}_{2}, \mathbf{b}\right\rangle\right|\right\}
$$

Then, $Z$ is a non-negative random variable, which follows from the coordinates of the random vectors $\mathbf{v}_{1}, \mathbf{v}_{2}, \mathbf{b}$ after multiplication, summation and absolute values. It is not hard to see that

$$
P_{\varepsilon}=\mathbb{P}(Z<\boldsymbol{\varepsilon})=\mathbb{E}\left[Z \cdot \mathbf{1}_{[0, \varepsilon]}\right],
$$

where $\mathbf{1}_{[0, \varepsilon]}$ is the characteristic (or indicator) function of the interval.
Recall (from Section 6.3) our assumption that the entries $\left(w_{i j}\right)_{1 \leq i \leq j \leq n}$ of $\mathbf{A}$ and $\left(b_{i}\right)_{1 \leq i \leq n}$ of $\mathbf{b}$ are independent. It follows that $\mathbf{v}_{1}, \mathbf{b}$ are independent random vectors and clearly this is
also true for the pair $\mathbf{v}_{2}, \mathbf{b}$. Using conditional expectation, we obtain that

$$
P_{\varepsilon}=\mathbb{E}\left[Z \cdot \mathbf{1}_{[0, \varepsilon]}\right]=\int_{\mathbb{R}^{2}} \mathbb{E}\left[Z \cdot \mathbf{1}_{[0, \varepsilon]} \mid \mathbf{b}\right] \cdot f(\mathbf{b}) d \mathbf{b}=\int_{\mathbb{R}^{2}} P_{\varepsilon, b} \cdot f(\mathbf{b}) d \mathbf{b},
$$

where $f(\mathbf{b})$ is the probability density function of the random vector $\mathbf{b}$. Since the coordinates $b_{1}, b_{2}$ of $\mathbf{b}$ are independent Gaussian random variables with zero mean and variance equal to 1 , it follows that

$$
P_{\varepsilon}=\int_{\mathbb{R}^{2}} P_{\varepsilon, \mathbf{b}} \cdot \frac{1}{2 \pi} \cdot \exp \left(-\frac{b_{1}^{2}+b_{2}^{2}}{2}\right) d b_{1} d b_{2}=\int_{\mathbb{R}^{2}} P_{\varepsilon, \mathbf{b}} \cdot \frac{1}{2 \pi} \cdot \exp \left(-\frac{\|\mathbf{b}\|_{2}^{2}}{2}\right) d b_{1} d b_{2}
$$

We observe now that $P_{\varepsilon, \mathbf{b}}$ depends only on $\|b\|_{2}=\sqrt{b_{1}^{2}+b_{2}^{2}}$. Hence, by changing in polar coordinates (or, equivalently using the fact that $b_{1}^{2}+b_{2}^{2}$ has the $\chi^{2}$-distribution with 2 degrees of freedom), we get that

$$
\begin{aligned}
P_{\varepsilon} & =\int_{0}^{2 \pi} \int_{0}^{\infty} P_{\varepsilon, \mathbf{b}}(r) \frac{1}{2 \pi} e^{-r^{2} / 2} r d r d \theta=\int_{0}^{\infty} P_{\varepsilon, \mathbf{b}}(r) e^{-r^{2} / 2} r d r=\int_{0}^{\infty} P_{\varepsilon, \mathbf{b}}(\sqrt{r}) \frac{1}{2} e^{-r / 2} d r \\
& =\int_{0}^{2 \varepsilon^{2}} \frac{1}{2} e^{-r / 2} d r+\int_{2 \varepsilon^{2}}^{\infty} \frac{4}{\pi} \cdot \arcsin \left(\frac{\varepsilon}{\sqrt{r}}\right) \frac{1}{2} e^{-r / 2} d r \\
& =1-e^{2 \varepsilon^{2}}+\frac{2}{\pi} \int_{2 \varepsilon^{2}}^{\infty} \arcsin \left(\frac{\varepsilon}{\sqrt{r}}\right) \cdot e^{-r / 2} d r
\end{aligned}
$$

and we have proved the desired result.

### 6.5 The general case

In this section, we consider the more general case where the state space of the random system (6.4) is $\mathbb{R}^{n}, n \geq 3$. This case is more complicated and we only give an upper bound for the $\varepsilon$-uncontrollability of the system.

Firstly, we need some estimates from the elementary convex geometry. Assume that $\mathbf{v} \in \mathbf{S}^{n-1}$ is a unit vector and $\varepsilon \in[0,1)$. The $\varepsilon$-spherical cap about $\mathbf{v}$ is the following subset of $\mathbf{S}^{n-1}$ :

$$
C(\varepsilon, \mathbf{v})=\left\{\theta \in \mathbf{S}^{n-1}:\langle\theta, \mathbf{v}\rangle \geq \varepsilon\right\}
$$

Observe that the number $\varepsilon$ does not refer to the radius of the cap. An easy calculation shows that the radius is $r=2(1-\varepsilon)$. In general, the cap of radius $r$ about $\mathbf{v}$ is:

$$
\left\{\theta \in \mathbf{S}^{n-1}:|\theta-\mathbf{v}| \leq r\right\}
$$

Let $A_{n}$ denote the surface area of the unit ball $S^{n-1}$, i.e. $A_{n}=\frac{2 \pi^{n / 2}}{\Gamma(n / 2)}$. Convex geometry provides the following upper and lower bounds for the surface area of a spherical cap (see, for example, [8]).

Lemma 6.5.1. For $0 \leq \varepsilon<1$, the cap $C(\varepsilon, v)$ on $S^{n-1}$ has surface area at most $e^{-n \varepsilon^{2} / 2} \cdot A_{n}$. Lemma 6.5.2. For $0 \leq r \leq 2$, a cap of radius $r$ on $S^{n-1}$ has surface area at least $\frac{1}{2} \cdot\left(\frac{r}{2}\right)^{n-1}$. $A_{n}$.

Following the lines of Theorem 6.4.1, we now prove the next result. Assume that we have a random system of the form (6.4), where the matrix $\mathbf{A}$ belongs to $\operatorname{GOE}(n)$.

Theorem 6.5.3. Let $\mathbf{A}$ be in the $G O E(n)$ and let $\mathbf{b} \in \mathbb{R}^{n}$ be any non-zero vector. Then, for the random system (6.4), we have the estimate:
$P_{\varepsilon, \mathbf{b}}=\mathbb{P}\left(|\langle\mathbf{v}, \mathbf{b}\rangle|<\varepsilon:\right.$ for some eigenvector $\mathbf{v}$ of $\left.\mathbf{A},\|\mathbf{v}\|_{2}=1\right) \leq n\left(1-\left(1-\frac{\varepsilon}{\|\mathbf{b}\|_{2}}\right)^{n-1}\right)$.
Proof. Let $\left\{\mathbf{v}_{i}\right\}_{i=1}^{n}$ be an orthonormal basis of $\mathbb{R}^{n}$ consisting of eigenvectors of the matrix $\mathbf{A}$. Without loss of generality (replacing $\mathbf{v}_{i}$ with $-\mathbf{v}_{i}$ if necessary), we may assume that the first non-zero coordinate of each $\mathbf{v}_{i}$ is positive. We now obtain:

$$
P_{\varepsilon, \mathbf{b}}=\mathbb{P}\left(\left|\left\langle\mathbf{v}_{i}, \mathbf{b}\right\rangle\right|<\varepsilon: \text { for some } i=1,2, \ldots, n\right) .
$$

We write $\mathbf{b}=\|\mathbf{b}\|_{2}\left(\frac{b_{1}}{\|\mathbf{b}\|_{2}}, \frac{b_{2}}{\|\mathbf{b}\|_{2}}, \ldots, \frac{b_{n}}{\|\mathbf{b}\|_{2}}\right)$ and we consider an orthogonal transformation $T \in O(n)$ that assigns $\left(\frac{b_{1}}{\|\mathbf{b}\|_{2}}, \frac{b_{2}}{\|\mathbf{b}\|_{2}}, \ldots, \frac{b_{n}}{\|\mathbf{b}\|_{2}}\right)$ to the vector $\mathbf{e}_{1}=(1,0, \ldots, 0)$. Since $T$ is orthogonal, it follows that:

$$
|\langle\mathbf{b}, \mathbf{v}\rangle|=|\langle T(\mathbf{b}), T(\mathbf{v})\rangle|=\|\mathbf{b}\|_{2} \cdot\left|\left\langle\mathbf{e}_{1}, T(\mathbf{v})\right\rangle\right|
$$

for any eigenvector $\mathbf{v}=\mathbf{v}_{i}, i=1,2, \ldots, n$ of $\mathbf{A}$. Hence,

$$
\begin{aligned}
& \mathbb{P}\left(\|\mathbf{b}\|_{2}\left|\left\langle T\left(\mathbf{v}_{i}\right), \mathbf{e}_{1}\right\rangle\right|<\varepsilon: \text { for some } i=1,2, \ldots, n\right) \\
= & \mathbb{P}\left(\left|\left\langle T\left(\mathbf{v}_{i}\right), \mathbf{e}_{1}\right\rangle\right|<\frac{\varepsilon}{\|\mathbf{b}\|_{2}}: \text { for some } i=1,2, \ldots, n\right) .
\end{aligned}
$$

Note that in $\mathbb{R}^{n}$ for $n \geq 3$, the sets $\left(\left|\left\langle T\left(\mathbf{v}_{i}\right), \mathbf{e}_{1}\right\rangle\right|<\frac{\varepsilon}{\|\mathbf{b}\|_{2}}\right), i=1,2, \ldots, n$ are not pairwise disjoint, even for small values of $\varepsilon$. Therefore, we cannot repeat the argumentation of the
case $n=2$. However, we may proceed as follows

$$
P_{\varepsilon, \mathbf{b}} \leq \sum_{i=1}^{n} \mathbb{P}\left(\left|\left\langle T\left(\mathbf{v}_{i}\right), \mathbf{e}_{1}\right\rangle\right|<\frac{\varepsilon}{\|\mathbf{b}\|_{2}}\right) .
$$

Since $\mathbf{v}_{i}$ is uniformly distributed in $S_{+}^{n-1}$ (see Proposition 6.3.1), we have that $T\left(\mathbf{v}_{i}\right)$ is uniformly distributed to some hemisphere. Therefore, if $\mathbf{A}$ denotes the surface area measure in the sphere $S^{n-1}$, then,

$$
\begin{aligned}
\mathbb{P}\left(\left|\left\langle T\left(\mathbf{v}_{i}\right), \mathbf{e}_{1}\right\rangle\right|<\frac{\varepsilon}{\|\mathbf{b}\|_{2}}\right) & \leq \frac{\mathbf{A}\left(\theta \in S^{n-1}: 0<\theta_{1}<\frac{\varepsilon}{\|\mathbf{b}\|_{2}}\right)}{A_{n} / 2} \\
& =\frac{A_{n} / 2-\mathbf{A}\left(\theta \in S^{n-1}: \frac{\varepsilon}{\|\mathbf{b}\|_{2}} \leq \theta_{1}\right)}{A_{n} / 2} .
\end{aligned}
$$

The set $\left\{\theta \in S^{n-1}: \frac{\varepsilon}{\|\mathbf{b}\|_{2}} \leq \theta_{1}\right\}$ is a spherical cap of radius $r=2\left(1-\frac{\varepsilon}{\|\mathbf{b}\|_{2}}\right)$. Hence, by Lemma 6.5.2, its surface area is at least $\frac{1}{2}\left(\frac{r}{2}\right)^{n-1} A_{n}$. Therefore,

$$
P_{\varepsilon, \mathbf{b}} \leq \frac{A_{n} / 2-\frac{1}{2}\left(\frac{r}{2}\right)^{n-1} A_{n}}{A_{n} / 2}=1-\left(\frac{r}{2}\right)^{n-1}=1-\left(1-\frac{\varepsilon}{\|\mathbf{b}\|_{2}}\right)^{n-1} .
$$

Hence,

$$
P_{\varepsilon, \mathbf{b}} \leq \sum_{i=1}^{n}\left(1-\left(1-\frac{\varepsilon}{\|\mathbf{b}\|_{2}}\right)^{n-1}\right)=n\left(1-\left(1-\frac{\varepsilon}{\|\mathbf{b}\|_{2}}\right)^{n-1}\right) .
$$

Theorem 6.5.4. Assume that $\mathbf{A}, \mathbf{b}$ belong to $G O E(n)$ and $S_{\mathbf{b}}$ respectively and let $\varepsilon$ be any positive number. For the $\varepsilon$-uncontrollability of the random system (6.4), the following inequality holds

$$
\begin{equation*}
P_{\varepsilon} \leq \frac{n}{2^{n / 2} \Gamma(n / 2)} \int_{0}^{\infty}\left(1-\left(1-\frac{\varepsilon}{\sqrt{r}}\right)^{n-1}\right) \cdot e^{-r / 2} \cdot r^{(n / 2)-1} d r . \tag{6.5}
\end{equation*}
$$

which can also be written as

$$
\begin{equation*}
P_{\varepsilon} \leq \sum_{k=1}^{n-1}(-1)^{k+1}\binom{n-1}{k} \frac{n \Gamma\left(\frac{n-k}{2}\right)}{2^{k / 2} \Gamma\left(\frac{n}{2}\right)} \varepsilon^{k} \tag{6.6}
\end{equation*}
$$

Consequently, the growth of $P_{\varepsilon}$ is at most polynomial of degree $n-1$ with respect to $\varepsilon$.

Proof. As in the proof of Theorem 6.4.2, it follows that

$$
P_{\varepsilon}=\int_{\mathbb{R}^{n}} P_{\varepsilon, \mathbf{b}} \cdot f(\mathbf{b}) d \mathbf{b},
$$

where $f$ is the probability density function of the random vector $\mathbf{b}$. Since the entries of $\mathbf{b}$ are independent Gaussian random variables with zero mean and variance equal to 1 , we have

$$
P_{\varepsilon}=\int_{\mathbb{R}^{n}} P_{\varepsilon, \mathbf{b}} \cdot \frac{1}{\sqrt{(2 \pi)^{n}}} \cdot \exp \left(-\frac{\|\mathbf{b}\|_{2}^{2}}{2}\right) d \mathbf{b}
$$

By Theorem 6.5.3, we obtain

$$
P_{\varepsilon} \leq \int_{\mathbb{R}^{n}} n\left(1-\left(1-\frac{\varepsilon}{\|\mathbf{b}\|_{2}}\right)^{n-1}\right) \cdot \frac{1}{\sqrt{(2 \pi)^{n}}} \cdot \exp \left(-\frac{\|\mathbf{b}\|_{2}^{2}}{2}\right) d \mathbf{b}
$$

We observe that, in the last integral, only the norm $\|\mathbf{b}\|_{2}$ of the vector $\mathbf{b}$ appears. Hence, using polar coordinates (see, for example, [21] Corollary 2.51), or equivalently, the fact that $\mathbf{b}_{1}^{2}+\ldots+\mathbf{b}_{n}^{2}$ has the $\chi^{2}$-distribution with $n$ degrees of freedom, we obtain

$$
\begin{aligned}
P_{\varepsilon} & \leq \frac{2 \sqrt{\pi^{n}}}{\Gamma(n / 2)} \int_{0}^{\infty} n\left(1-\left(1-\frac{\varepsilon}{r}\right)^{n-1}\right) \cdot \frac{1}{\sqrt{(2 \pi)^{n}}} \cdot \exp \left(-\frac{r^{2}}{2}\right) r^{n-1} d r \\
& =\frac{n}{2^{n / 2} \Gamma(n / 2)} \int_{0}^{\infty}\left(1-\left(1-\frac{\varepsilon}{\sqrt{r}}\right)^{n-1}\right) \cdot e^{-r / 2} \cdot r^{(n / 2)-1} d r
\end{aligned}
$$

and the proof of inequality (6.5) is complete.
As far as (6.6) is concerned, we use the binomial expansion formula to obtain

$$
\begin{equation*}
1-\left(1-\frac{\varepsilon}{\sqrt{ } r}\right)^{n-1}=\sum_{k=1}^{n-1}\binom{n-1}{k}(-1)^{k+1} \frac{\varepsilon^{k}}{r^{k / 2}} \tag{6.7}
\end{equation*}
$$

By inequality (6.5), it follows that

$$
\begin{aligned}
P_{\varepsilon} & \leq \sum_{k=1}^{n-1}(-1)^{k+1}\binom{n-1}{k} \frac{n}{2^{n / 2} \Gamma\left(\frac{n}{2}\right)} \cdot \varepsilon^{k} \int_{0}^{\infty} e^{-r / 2} r^{\frac{n-k}{2}-1} d r \\
& =\sum_{k=1}^{n-1}(-1)^{k+1}\binom{n-1}{k} \frac{n \Gamma\left(\frac{n-k}{2}\right)}{2^{k / 2} \Gamma\left(\frac{n}{2}\right)} \varepsilon^{k}
\end{aligned}
$$

and we have the desired result.
The next natural corollary is now straightforward.
Corollary 6.5.4.1. For any integer $n \geq 2$, we have that $\lim _{\varepsilon \rightarrow 0} P_{\varepsilon}=0$.
Furthermore, the following corollary shows that the upper estimate of Theorem 6.5.4 is sharp in the case $\mathrm{n}=2$.

Corollary 6.5.4.2. Under the assumptions of Theorem 6.4.1, we have

$$
\begin{equation*}
\frac{2}{\sqrt{\pi}} \varepsilon \leq P_{\varepsilon}<\sqrt{2 \pi \varepsilon} \tag{6.8}
\end{equation*}
$$

for $\varepsilon>0$ small enough.
Proof. The right-hand inequality follows immediately by Theorem 6.5.4 and holds for any positive $\varepsilon$. As far as the left-hand inequality is concerned, we utilize Theorem 6.4.2. Firstly, we observe that the inequality $e^{x}-1<(e-1) x$, for $0<x<1$, implies that $1-e^{2 \varepsilon^{2}}>$ $-2(e-1) \varepsilon^{2}$ for small values of $\varepsilon$ (namely, when $2 \varepsilon^{2}<1$ ). Furthermore, $\operatorname{since} \arcsin (x)>x$ for $0<x<1$, we obtain

$$
\begin{align*}
\int_{2 \varepsilon^{2}}^{\infty} \arcsin \left(\frac{\varepsilon}{\sqrt{x}}\right) e^{-x / 2} d x & \geq \\
& =\int_{0}^{\infty}\left(\frac{\varepsilon}{\sqrt{x}}\right) e^{-x / 2} d x-\int_{0}^{2 \varepsilon^{2}}\left(\frac{\varepsilon}{\sqrt{x}}\right) e^{-x / 2} d x \\
& \geq \sqrt{2} \Gamma\left(\frac{1}{2}\right) \varepsilon-\varepsilon \int_{0}^{2 \varepsilon^{2}} \frac{1}{\sqrt{x}} d x \\
& =\sqrt{2} \sqrt{\pi} \varepsilon-2 \sqrt{2} \varepsilon^{2} \tag{6.9}
\end{align*}
$$

Therefore, the formula of Theorem 6.4.2 gives

$$
\begin{align*}
P_{\varepsilon} & \geq-2(e-1) \varepsilon^{2}+\frac{4}{\sqrt{\pi}} \varepsilon-\frac{8}{\pi} \varepsilon^{2} \\
& =\left(\frac{4}{\sqrt{\pi}}-\left(\frac{8}{\pi}+2(e-1)\right) \varepsilon\right) \varepsilon \\
& \geq \frac{2}{\sqrt{\pi}} \varepsilon \tag{6.10}
\end{align*}
$$

(for small values of $\varepsilon$ ). Consequently, we have proved the next inequality:

$$
\begin{equation*}
\frac{2}{\sqrt{\pi}} \varepsilon \leq P_{\varepsilon} \leq \sqrt{2 \pi} \varepsilon \tag{6.11}
\end{equation*}
$$

(for small values of $\varepsilon$ ) showing that the upper bound of Theorem 6.5.4 is sharp in the case $n=2$.

Finally, an upper estimate for the growth rate of $P_{\varepsilon}$ at 0 can also be obtained.
Corollary 6.5.4.3. Assume that $P_{\varepsilon}$ is differentiable at 0 . Then,

$$
\left.\frac{d P_{\varepsilon}}{d \varepsilon}\right|_{\varepsilon=0} \leq \frac{n(n-1) \Gamma\left(\frac{n-1}{2}\right)}{\sqrt{2} \Gamma\left(\frac{n}{2}\right)} .
$$

Proof. It follows immediately by inequality (6.6) of Theorem 6.5.4.

### 6.6 Conclusions

We defined a measure of $\varepsilon$ uncontrollability in a Gaussian Random Ensemble of linear systems. We calculated tight bounds for this probability in terms of $\varepsilon$ and the number of states $n$. This is also depicted in the graphs included in the appendix (Figures 15, 16).

## Chapter 7

## Conclusions

'And if you find her poor, Ithaka won't have fooled you. Wise as you will have become, so full of experience, you'll have understood by then what these Ithakas mean'.
— K . P. Kavafis, 1863-1933, Greek Poet.

### 7.1 Epilogue

In this section we summarize the most important findings of this thesis and we outline prospects for future research.

The aim of this PhD Thesis is trifold. Initially, we looked at networks of banks which are modelled by connected, directed and weighted graphs and are sustained by interbank lending. We assumed that the interbank lending is governed by a diffusion-like law. We proved that the diffusion-like interbank lending drives the network of banks towards its steady state where the leverages of all the banks of the network are equal. This research can be of interest in the banking community as this method determines on systematic manner flows of funds between banks that make the overall system more stable. That said, it could be possible for the central bank which is the main authority and has under control all the banks to choose those leverages in order to have a stable banking network. This possibility is something that we are going to study in the future as another research paper to which we will focus on the regulation mechanism of a central bank. In other words, we want to study how the central banker could play the role of a stabilizer, a regulator in an unstable banking network. Thus, we will approach this problem by introducing the mathematical control theory.

The second topic we examined was not only to investigate the algebraic and computations of almost zeros of polynomial vectors or matrices but also to describe new results that we found and which rely on a high order of almost zeros. A connection between randomness and almost zeros was shown to exist through the study of problems of random ensembles of
polynomial vectors. Moreover, several statistical results in connection with the prime almost zeros is presented as well. More specifically, we studied problems of random ensembles of polynomial vectors and several statistical results in connection with the prime almost zeros are presented as well. As it is well known, the importance of zeros and poles of a linear dynamical system is shown via a Laplace transformation to the system of coupled linear ordinary differential equations which describe the dynamical system. The position of poles and zeros is related with the stability of the system as well as with the stabilization of the system with feedback [33][38][34][28]. Also, if we know the position of zeros and poles, we can obtain very useful information and conclusions not only about the behavior of systems but also about their controllability. It is well known that if zeros lay on the right of the complex plane, then the stabilization by feedback of this system is difficult. There are some cases where a system does not have strict zeros but has almost zeros. That being said not merely the existence of such almost zeros but also their position can efficiently be characterized equivalently by the dynamic of a system as it could have happened in the case of ordinary zeros. These are some cases where, we are facing the problem of not existence of zeros by the strict meaning. Thus, alternatively we can use the existence of almost zeros in order to study besides the dynamic behavior of the system its stabilization ability as well.

Focusing on an economic system poles and zeros could be computed and we can obtain very useful conclusions about the dynamic behavior of the system along with the availability to design a controller responsible for its stabilization. For instance let an economic system where, the almost zeros are laying on the right complex plain. Then, this system is characterized by a stabilization problem and probably it's very difficult to stabilize it by feedback. In addition, a system with almost zero that belongs to the right half plane and which is close to a pole of a system then it is very difficult to stabilize this system by feedback. For more information on those dynamical characteristics which have been studied in depth, we recommend to have a glimpse on the relative bibliography [35] [36] [37].

The third aim of this thesis was to study in depth a linear time invariant multivariable system from a controllability point of view. More specifically, we focused on systems with a special characteristic that given matrices were random. Moreover, we worked on the case of Gaussian Random Ensemble and we defined $\varepsilon$ controllability which is a measure for the previous category of random matrices. Finally, we found bounds in terms of $\varepsilon$ and the number of states $n$ of a given linear system. The notion of controllability relies on the ability of changing the behavior of the system. More specifically, suppose we have a system either economic, mechanic or electric which can be described by a system of differential equations with input, output and states. Then, the controllability gives an answer on how likely it is to transfer a vector state by changing the input values or the policy variables. In such a
case the system is called controllable; conversely, we say that the system is uncontrollable. Moreover, the properties of controllability, stability are of great importance which are used in the theory of systems and control very often. Uncontrollable systems are of zero measure in the total of all systems. Hence, if we assume a random system, then it is more likely to be controllable. Nevertheless, there is a necessity to evaluate if a random system is controllable or uncontrollable; because the systems are not all the same. Furthermore, $\varepsilon-$ controllability is a broader notion of controllability meaning because in the systems we do not only have two states i.e. controllable or uncontrollable but there are also intermediate states of controllability. In other words, we can think about $\varepsilon$ - controllability as a scale notion of controllability. In addition, in this thesis we define the notion of $\varepsilon$ - controllability which has a very special meaning in random systems because we can evaluate how controllable such a system is, to put it in another way, it shows how difficult it is to change the behavior of the system through control theory techniques like the feedback mechanism. In economic science there are systems which are characterized by randomness like networks or links that are unstable, residuals, noise, laws under uncertainty etc. Finally, in such systems controllability in the classical point of view is not applicable thus we are using $\varepsilon$ - controllability which is a scale or an approximation of controllability as a more appropriate tool. In the ensemble we use to parameterize random systems we calculated the distribution of $\varepsilon$-controllability of these systems.

## References

[45] Wikipedia. https://en.wikipedia.org/wiki/Directed_graph.
[2] A., F. (1855). Poggendorff's annalen. Phil. Mag., 94:59-86.
[3] ABED, E. H. and Jyu-Horng-FU (1986). Local feedback stabilization and bifurcation control, I. Hopf bifurcation. Systems \& Control Letters, 7:11-17.
[4] Anderson, G. W., Guionnet, A., and Zeitouni, O. (2009). An Introduction to Random Matrices. Cambridge Studies in Advanced Mathematics. Cambridge University Press.
[5] Appel, K. and Haken, W. (1977). The Solution of the Four-Color-Map Problem. Scientific American, a division of Nature America, Inc., 237(4):108-121.
[6] Baillieul, J. (1980). Controllability and observability of polynomial dynamical systems. Nonlinear Analysis, Theory, Methods \& Applications, 5(5):543-552.
[7] Baillieul, J. (1990). The geometry of homogeneous polynomial dynamical systems. Nonlinear Analysis, Theory, Methods \& Applications, 4(5):879-900.
[8] Ball, K. (1997). An elementary introduction to modern convex geometry. Flavors of Geometry, 31(4):1-55.
[9] Bellman, R. and Kalaba, R. (1960). Dynamic Programming and Adaptive Processes: Mathematical Foundation. IRE Transactions on Automatic Control, 5:5-10.
[10] Berge, C. (1958). Theorie des graphes et ses applications. Mathematisch Centrum, pages 68-69.
[11] Bollodas, B. (1998). Modern Graph Theory, volume 1. Springer,New York,NY.
[12] Boyce, W. E. and DiPrima, R. C. (2004). Elementary Differential Equations and Boundary Value Problems. JOHN WILEY SONS.
[13] Brouwer, A. E. and Haemers, W. H. (2011). Spectra of graphs. Springer: Amsterdam.
[14] Chung, F. R. K. (1997). Spectral Graph Theory. American Mathematical Society.
[15] Dijkstra, E. W. (1959). A Note on Two Problems in Connexion with Graphs. Numerische Mathematik, 1:269-271.
[16] Donath, W. E. and Hoffman, A. J. (1973). Lower bounds for the partitioning of graphs. IBM J. Res. Develop., pages 420-425.
[17] Dyson, F. J. (1953). The Dynamics of a Disordered Linear Chain. The Physical Review, 92(6):1331-1338.
[18] Fick, A. (1855). On liquid diffusion. Philos Mag. J. Sci., 10:31-99.
[19] Fiedler, M. (1973). Algebraic connectivity of graphs. Czech. Math. J., pages 298-305.
[20] Finke, G., Burkard, R. E., and Rendl, F. (1987). Quadraticassignmentproblem. Ann.Discrete Math., pages 61-82.
[21] Folland, G. (1999). Real Analysis: Modern techniques and their applications. John Wiley \& Sons.
[22] Ford, L. R. (1956). Network flow theory. Rand Corp. Paper, page 923.
[23] G., K. (1847). Uber die auflosung der gleichungen, auf welche man bei der untersuchung der lineaaren verteilung galvanischer strome gefuhrt wird. Ann Phys. Chem, pages 497508.
[24] Graham, T. (1829). Literature and Art. Quaterly Journal of Science, 74.
[25] Grone, R. (1991). On the geometry and Laplacian of a graph. Linear Algebra and its Applications, pages 167-178.
[26] Grone, R. and Merris, R. (1990). Ordering trees by algebraic connectivity. Graphs and Combinatorics, pages 229-237.
[27] Grone, R., Merris, R., and Sunder, V. (1990). The laplacian spectrum of a graph. SIAM J.Matrix Analysis and its Applications, pages 218-238.
[28] Halikias, G., Galanis, G., Karcanias, N., and Milonidis, E. (2012). Nearest common root of polynomials, approximate greatest common divisor and the structured singular value. IMA Journal of Mathematical Control and Information, 30(4):423-442.
[29] Horn, R. A. and Johnson, C. A. (1985). Matrix analysis. Cambridge: Cambridge University Press.
[30] Hurwitz, A. (1895). On the conditions under which an equation has only roots with negative real parts. Mathematishe Annalen, 46:273-284.
[31] Izenman, Alan, and Julian (2020). Random matrix theory and its applications. AMS.
[32] J., T. R. (1993). Introduction to Graph Theory. Dover.
[33] Karcanias, N. (2007). The greatest common divisor of a set of polynomials, control theory and approximate algebraic computations. Bull. Greek Math. Soc., 54:29-46.
[34] Karcanias, N., Fatouros, S., Mitrouli, M., and Halikias, G. (2006). Approximate greatest common divisor of many polynomials, generalised resultants, and strength of approximation. Computers Mathematics with Applications, 51(12):1817-1830.
[35] Karcanias, N. and Giannakopoulos, C. (1984). Grassmann invariants, almost zeros and the determinantal zero, pole assignment problems of linear multivariable systems. Internat. J. Control, 40(4):673-698.
[36] Karcanias, N. and Giannakopoulos, C. (1989). Necessary and sufficient conditions for zero assignment by constant squaring down. Linear Algebra and its Applications, 122-124:415-446.
[37] Karcanias, N., Giannakopoulos, C., and Hubbard, M. (1983). Almost zeros of a set of polynomials. International Journal of Control, 38(6):1213-1238.
[38] Karcanias, N. and Halikias, G. (2013). Approximate zero polynomials of polynomial matrices and linear systems. Linear Algebra Appl., 4:1091-1103.
[39] Karkanias, N. Milonidis, E. (2007). Mathematical Methods for Robust and Nonlinear Control, volume 367. Springer.
[40] Kruskal jr., J. B. (1956). On the Shortest Spanning Subtree of a Graph and the Travelling Salesman Problem. Proc. Amer. Math. Soc., 7:48-50.
[41] L. R. Ford, J. and Fulkerson, D. R. (1956). Maximal Flow Through a Network. Canadian Journal of Mathematics, 8:399-404.
[42] LaSalle, J. P. (1959). Time Optimal Control. Proceedings of the National Academy of Sciences, 45:573-577.
[43] Loberman, H. and Weinberger, A. (1957). Formal Procedures for Connecting Terminals with a Minimum Total Wire Length. J. Ass. Comp. Mach., 4:428-437.
[44] Marčenko, V. A. and Pastur, L. A. (1967). Mathematics of the USSR-Sbornik Distribution of Eigenvalues for Some Sets of Random Matrices. Mathematics of the USSR-Sbornik, 1(4):457-483.
[45] Maxwell, J. C. (1868). On Governors. M.A., F.R.SS.L \& E., 20.
[46] Meghanathan, N. (2015). Probabilistic Diffusion in Random Network Graphs. https://arxiv.org/pdf/1511.06613.pdf.
[47] Merris, R. (1991). The number of eigenvalues greater than two in the laplacian spectrum of a graph. Portugal. Math., pages 345-349.
[48] Merris, R. (1994). Laplacian matrices of graphs: a survey. Linear Algebra and its Applications, pages 143-176.
[49] Merris, R. (1995). A survey of graph Laplacians. Linear and Multilinear Algebra, pages 19-31.
[50] Merris, R. (1998a). A note on Laplacian graph eigenvalues. Linear and its Applications, pages 33-35.
[51] Merris, R. (1998b). Doubly stochastic graph matrices II. Linear and Multilinear Algebra, pages 275-285.
[52] Mohar, B. (1991). Eigenvalues, diameter, and means distance in graphs. Graphs and Combinatorics, pages 53-64.
[53] Mohar, B. (1992). Laplace eigenvalues of graphs-a survey. Discrete Mathematics, pages 171-183.
[54] Mohar, B. (1997). Some applications of Laplace eigenvalues of graphs. Kluwer AC. Press, pages 225-275.
[55] Mohar, B. and et al., Y. A. (1991). The Laplacian spectrum of graphs. Wiley, New York.
[56] Nyquist, H. (1932). Regeneration Theory. Bell System Technical Journal, 11:126-147.
[57] Ormerod, P. (2008). Random Matrix Theory and Macro-Economic Time-Series: An Illustration Using the Evolution of Business Cycle Synchronisation, 1886-2006. Economics: The Open-Access, 2(2008-26):1-10.
[58] Ormerod, P. and Mounfield, C. (2000). Random matrix theory and the failure of macroeconomic forecasts. Physica A: Statistical Mechanics and its Applications, 280(3):497504.
[59] Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L., and Stanley, H. (2000). A random matrix theory approach to financial cross-correlations. Physica A: Statistical Mechanics and its Applications, 287(3):374-382.
[60] Sontag, E. (1979). On the observability of polynomial systems, I:Finite-time problems. SIAM, J. Cont. and Opt., 17(1):139-151.
[61] Sontag, E. and Rouchal, Y. (1976). On discrete-time polynomial systems. Nonlinear Analysis, Theory, Methods \& Applications, 1(1):55-64.
[62] Wigner, E. P. (1958). On the Distribution of the Roots of Certain Symmetric Matrices. Annals of Mathematics, 67(2):325-327.
[63] Willems, J. and Polderman, J. (1998). Introduction to Mathematical Systems Theory. Springer.

## Part IV

## Appendix

## . 1 Chapter 3



Fig. 1 Plot of the vector field for the system in Example 4.

## . 2 Chapter 4

Sort of leverages in descending order on an interbank loan network


Fig. 2 Sort list of leverages in descending order.


Fig. 3 The mean initial leverage $\left(l_{M}\right)$ is equal to $7 / 9$. At the point of equilibrium $l_{i}=l_{j}=$ $l \in \mathbb{R}, i \neq j$ and $l \geq l_{M}$


Fig. 4 From dark blue to red we are moving from the highest level of leverage per bank to the lowest and furthermore from the heighest level of weights to the lowest, for the nodes and edges, respectively.

## . 3 Chapter 5



Fig. 5 Mean of p-index ${ }^{2}$ vs $k, k=1,2, \ldots, 25$


Fig. 6 Standard deviation of p -index ${ }^{2}$ vs $k, k=1,2, \ldots, 25$


Fig. 7 Mean of the size of almost zero vs $k, k=1,2, \ldots, 25$


Fig. 8 Standard deviation of the size of almost zero vs $k, k=1,2, \ldots, 25$

(a) PDF of the p-index ${ }^{2} P_{M}\left(s_{0}, \overline{s_{0}}\right)$ of degree 1

(c) PDF of the p-index ${ }^{2} P_{M}\left(s_{0}, \overline{s_{0}}\right)$ of degree 3

(b) PDF of the p-index ${ }^{2} P_{M}\left(s_{0}, \overline{s_{0}}\right)$ of degree 2

(d) PDF of the p-index ${ }^{2} P_{M}\left(s_{0}, \overline{s_{0}}\right)$ of degree 4

Fig. 9 Probability distribution functions of the p -index ${ }^{2}$ for degrees ranging from 1 up to 4

| Statistic | Value | Percentile | Value |
| :---: | :---: | :---: | :---: |
| Sample Size | 2000 | Min | 3.7874 |
| Range | 10.57 | 5\% | 5.703 |
| Mean | 7.9657 | 10\% | 6.1758 |
| Variance | 2.2216 | 25\% (Q1) | 6.9208 |
| Std. Deviation | 1.4905 | 50\% (Median) | 7.8683 |
| Coef. of Variation | 0.18711 | 75\% (Q3) | 8.8806 |
| Std. Error | 0.03333 | 90\% | 9.8809 |
| Skewness | 0.42373 | 95\% | 10.629 |
| Excess Kurtosis | 0.29537 | Max | 14.357 |

Fig. 10 Descriptive statistics of polynomials degree 1


Fig. 11 The distribution of the p-index ${ }^{2}$ of $s_{0}, P_{M}\left(\overline{s_{0}}, s_{0}\right)$ follows Gamma distribution and flattens with respect to the size of $M$.


Fig. 12 The distribution of almost zero $s_{0}$ of $\left\|s_{0}\right\|^{2}$ follows Gamma distribution and sharpens with respect to the size of $M$.

| $\#$ | Distribution |  | Kolmogorov <br> Smirnov |  | Anderson <br> Darling |  | Chi-Squared |  |
| :---: | :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | Statistic | Rank | Statistic | Rank | Statistic | Rank |  |
| 48 | Pearson 6 | 0.01017 | 1 | 0.21144 | 9 | 6.3371 | 13 |  |
| 23 | Gen. Gamma (4P) | 0.01023 | 2 | 0.21106 | 8 | 5.5692 | 3 |  |
| 1 | Beta | 0.01049 | 3 | 0.21427 | 10 | 5.2194 | 1 |  |
| 19 | Gamma | 0.01064 | 4 | 0.26813 | 13 | 6.0832 | 9 |  |
| 22 | Gen. Gamma | 0.01075 | 5 | 0.26038 | 12 | 6.0422 | 8 |  |
| 20 | Gamma (3P) | 0.01079 | 6 | 0.20358 | 7 | 5.6324 | 4 |  |
| 21 | Gen. Extreme Value | 0.01096 | 7 | 0.25248 | 11 | 5.5327 | 2 |  |
| 38 | Log-Pearson 3 | 0.01099 | 8 | 0.20063 | 6 | 6.2552 | 12 |  |
| 16 | Fatigue Life (3P) | 0.01158 | 9 | 0.19461 | 3 | 6.0265 | 7 |  |
| 30 | Johnson SB | 0.01172 | 10 | 0.19345 | 1 | 5.7952 | 5 |  |
| 41 | Lognormal (3P) | 0.012 | 11 | 0.19373 | 2 | 6.2537 | 11 |  |
| 49 | Pearson 6 (4P) | 0.01228 | 12 | 0.19479 | 4 | 6.0885 | 10 |  |
| 47 | Pearson 5 (3P) | 0.01236 | 13 | 0.19528 | 5 | 5.9357 | 6 |  |
| 3 | Burr (4P) | 0.01402 | 14 | 0.41143 | 14 | 7.4827 | 14 |  |
| 15 | Fatigue Life | 0.01766 | 15 | 0.60417 | 16 | 8.9919 | 15 |  |
| 40 | Lognormal | 0.01803 | 16 | 0.57055 | 15 | 9.2336 | 16 |  |
| 42 | Nakagami | 0.02078 | 17 | 1.8313 | 21 | 13.379 | 18 |  |
| 2 | Burr | 0.02238 | 18 | 1.4744 | 18 | 17.477 | 21 |  |
| 8 | Dagum (4P) | 0.02403 | 19 | 1.4355 | 17 | 14.751 | 19 |  |
| 10 | Erlang (3P) | 0.02578 | 20 | 2.1405 | 22 | 10.591 | 17 |  |
|  |  |  |  |  |  |  |  |  |

Fig. 13 Fit of various distributions for the size or p -index ${ }^{2}$ for the ensemble of degree 1

(a) Polynomials of sum of squares degree 1

(c) Polynomials of sum of squares degree 3
(b) Polynomials of sum of squares degree 2

(d) Polynomials of sum of squares degree 4

Fig. 14 Probability distribution functions of sum of squares of polynomials degrees 1 up to 4

## . 4 Chapter 6

The diagram of the upper bound for $\varepsilon$-uncontrollability given in Theorem 6.5.4 (Fig. 15);


Fig. 15 Moving from the lower to the upper value of $\varepsilon$ the probability bound increases with n.

The diagram of the upper bound for the derivative $\left.\frac{d P_{\varepsilon}}{d \varepsilon}\right|_{\varepsilon=0}$ given in Corollary 6.5.4.3 (Fig.16).
growth vs n


Fig. 16 The growth of probability bound increases quadraticaly with n .

## Index

A
Adjacency, 59
Adjacent, 57
Arcs, 59
Autonomous, 35

## B

Behavior, 34

## C

Cardinality, 57
Characteristic matrix, 28
Characteristic polynomial, 28, 31
Characteristic values, 28
Connectivity, 58
Constant coefficient, 26
Constructibility, 41
Controllability, 35, 36
Cubic, 58

D
Degree, 58
Diagonalizable, 27
Diameter, 58
Diffusion, 14
Directed, 58
Distance, 58
Dynamical system, 34

## E

Eigenvalues, 28
Empty set, 57

## F

Finite, 57
Fundamental matrix, 31, 35

## G

GOE, 54, 106
Graph, 57
Graph theory, 12, 57
GSE, 54
GUE, 54

## H

Homogeneous, 26
Hurwitz criterion, 44, 46

## I

I/O systems, 34
Incident, 57, 59
Indegree, 59
Infinite, 57
Initial, 59
Initial value problem, 26
Input, 34
Integral transforms, 21
Integrating fuctor, 20

Inverting the transform, 21
Isoleted, 58
Isomorphic, 57

## K

K-regular, 58
Kernel, 21
L
Laplace Transform method, 19
Latent variables, 34
Linear systems, 34
Lyapunov criterion, 44, 48
M
Mathematical model, 34
Maximum degree, 58
Minimum degree, 58

## N

Neighbours, 57
Nonhomogeneous, 26

## O

Observability, 41
Outdegree, 59
Output, 34
$\mathbf{P}$
p-index, 94
Path, 58
Phase portrait, 30
Piecewise continuous, 20
Pole polynomial, 49
Poles, 48, 49

## R

Random matrix ensemble, 105
Reachability Gramian, 38, 40
Reachable, 36, 40
Resolvent matrix, 28
RMT, 8, 50
Routh criterion, 44, 45
Routh Matrix, 45

## S

Similarity transformation, 27
Spanning subgraph, 58
State reachability, 35
State space, 35
State transition matrix, 35
State variables, 34
State-feedback, 34
Subgraph, 58
Symmetric, 59

## T

Terminal, 59
Toeplitz, 42
Transition matrix, 35

## U

Uncertainty principle, 101
Universum, 34

## W

Wronskian, 31

## Z

Zero polynomial, 50
Zeros, 48, 49

