CIMAT

Centro de Investigación en Matemáticas, A.C.

Deterministic Eigenvalue processes arising from Random Matrix Theory and Finite Free Probability

T E S I S

Que para obtener el grado de

Maestro en Ciencias con Especialidad en Probabilidad y Estadística

Presenta

Juan Esaul González Rangel

Director de Tesis:

Dr. Octavio Arizmendi Echegaray

Co-director de Tesis:

Dr. José Luis Ángel Pérez Garmendia

Autorización de la versión final

Guanajuato, Guanajuato, Septiembre de 2024

Abstract

In this work, we explore some connections between two areas of Probability Theory: finite free probability and the theory of stochastic processes. The connection is established through the application of both fields to random matrices. Some basic definitions and results are presented in the first chapter. The second chapter introduces the most common application of stochastic processes to the study of random matrices: eigenvalue processes. The third chapter provides an overview of finite free probability theory, focusing on the relationship between expected characteristic polynomials and the convolution of polynomials. In the final chapter, we connect these ideas by considering the expected characteristic polynomial of a matrix-valued stochastic process. We demonstrate that the roots of these characteristic polynomials satisfy differential equations that resemble the stochastic differential equations of matrix-valued stochastic processes discussed in the second chapter.

Palabras clave: Random matrices, Stochastic processes, Finite Free Probability, Orthogonal polynomials, Dyson Brownian motion, Wishart process, Jacobi process.

Contents

Introduction

In modern statistics, random matrix theory can help address two important problems: the study of large covariance matrices and the theoretical justification of deep learning.

Historically, data collection from natural phenomena has been affected by intrinsic noise that arises as a result of imprecision or errors in measurement. Traditional statistical methods are effective in mitigating the effects of this noise when working with small datasets. However, as the number of variables and observations increases, the relationships between variables are modeled with large covariance matrices, making using random matrix theory necessary.

In a related area, the theoretical foundations of new deep learning techniques remain incomplete. Despite the advances that have been achieved using these types of models, their implementation usually follows heuristic rules, which makes it impossible to generalize the results or find optimal methodologies. Currently, there are several attempts to generate a rigorous theory for the convergence of deep learning models. Essentially, a deep learning model is a set of parameters that are updated in each iteration of the training algorithm. An approach to this problem is to represent the training of the model under a certain algorithm as the evolution of a matrix-valued stochastic process.

Random matrix theory, through the study of the evolution of matrix processes, provides tools to solve this problem. Typically, the study of random matrices is based on knowing the behavior of spectral distributions. In particular, when working with matrix-valued stochastic processes, it is of interest to know the behavior of their eigenvalue process.

In Freeman Dyson's seminal work [1], a system of differential equations is obtained to describe the eigenvalues of a Brownian motion in a space of self-adjoint real (GOE), complex (GUE), or quaternionic (GSE) matrices. The process that describes the evolution of each of these eigenvalues is called Dyson's Brownian motion, and it is the only weak solution to the following stochastic differential equation,

$$\mathrm{d}\lambda_j = \sqrt{\frac{2}{\beta}}\mathrm{d}B_j + \sum_{k\neq j}\frac{\mathrm{d}t}{\lambda_j - \lambda_k}, \tag{1}$$

where λ_i is the i-th eigenvalue of the matrix, and $B_1, B_2, \ldots, B_n$ represent independent Brownian motions. In this case, the parameter β depends on the space in which the random matrices take values, and thus $\beta = 1$ represents the orthogonal case (GOE), $\beta = 2$ the unitary case (GUE) and $\beta = 4$ the symplectic case (GSE).

Several subsequent works find differential equations for the eigenvalues of matrix processes. In particular, Graczyk and Maecki [2] manage to generalize Dyson's result by finding a differential equation for the eigenvalues of matrix-valued processes $X = (X_t, t \geq 0)$ that satisfy a stochastic differential equation of the form

$$\mathrm{d}X(t) = g(X(t))\mathrm{d}W(t)h(X(t)) + h(X_t)\mathrm{d}W(t)^T g(X(t)) + b(X(t))\mathrm{d}t, \tag{2}$$

where $W(t)$ is a matrix whose entries are all independent Brownian motions, and where the functions g and h act spectrally on X. Under certain acceptable regularity conditions, Graczyk and Małecki show that the eigenvalues of this process are the only solution to the following differential equation

$$\mathrm{d}\lambda_j = 2g(\lambda_j)h(\lambda_j)\mathrm{d}B_j + \left(b(\lambda_j) + \sum_{k\neq j} \frac{G(\lambda_j,\lambda_k)}{\lambda_j - \lambda_k} \right) \mathrm{d}t, \tag{3}$$

where $B_1, \ldots, B_n$ are independent Brownian motions and $G(x,y) := g^2(x)h^2(y)$.

In a different context, Marcus, Spielman and Srivastava [3] study the convolutions of polynomials and find that, using an analogy with free probability, it is possible to write these convolutions as expected polynomials of sums and products of random matrices invariant under unitary transformations. Specifically, they find that the roots of the characteristic polynomial of a Brownian motion in the space of Hermitian matrices satisfy the equation,

$$\mathrm{d}\lambda_j = \sum_{k\neq j} \frac{\mathrm{d}t}{\lambda_j - \lambda_k}. \tag{4}$$

This equation coincides with the equation satisfied by Dyson's Brownian motion when the diffusion coefficient is removed.

Finally, Holcomb and Paquette [4] constructed a tridiagonal matrix model whose eigenvalues also satisfy (4). The recurrence of this equation in various contexts suggests a deep connection that can be used to translate insights between different fields.

This thesis is a modest attempt to explore how the theory of stochastic processes and finite free probability can be used to approach the study of eigenvalue processes. In the first chapter, we introduce basic concepts of random matrix theory, stochastic calculus, and non-commutative probability. In Chapter 2 we use stochastic calculus tools to study the behavior of eigenvalue processes in a quite general setting. In Chapter 3 we provide a brief introduction to finite free probability theory as a theoretical setting to study expected characteristic polynomials. Finally, in Chapter 4 we construct matrix-valued stochastic processes whose spectrum behaves as the roots of the expected characteristic polynomial for some of the matrix-valued stochastic processes studied in Chapter 2. We then relate these results to Finite Free Probability Theory by showing some connections.

Chapter 1

Preliminaries

This chapter has two purposes; the first is to introduce some concepts and results used throughout the thesis. The second is to standardize the notation coming from different areas, making it easier to read the text. The first section briefly introduces the objects of study of random matrix theory and a few matrix algebra results recalled later. The second section is a brief presentation of stochastic calculus for $\mathbb{R}$-valued processes and then for processes taking values in spaces of matrices. The theory developed here will be used for the dynamics of eigenvalues in Chapters 2 and 4. The last section covers the essential definitions and theorems coming from non-commutative probability theory and constitutes a precedent for Chapter 3.

The use of several kinds of objects commonly denoted with the same symbols makes it necessary to specify the kind of object we are dealing with every time. The time parameter of a stochastic process will always be shown in parentheses (i.e., $W(t)$). The subindexes will represent the entry of a matrix or vector. An integer interval of length k will be represented by the symbol $[k]$, i.e.

$$[k] \coloneqq \{1, 2, \ldots, k-1, k\}.$$

When we place a set S in the combinations symbol $\binom{S}{k}$, we denote the collection of all subsets of S that have exactly k elements.

The determinant of a matrix A will be denoted by $\det[A]$ and if S, T are sets of integers, $\det[A]_{S,T}$ represents the determinant of the submatrix $A_{S,T}$. The transpose of A is denoted as A^T and its adjoint element is A^*. The space of $n \times m$ matrices with entries in a field $\mathbb{F}$ is denoted by $\mathcal{M}_{n,m}(\mathbb{F})$. The space of $n \times n$ symmetric matrices is denoted by $\mathcal{H}_{n,n}(\mathbb{R})$ and the space of $n \times n$ hermitian by $\mathcal{H}_{n,n}(\mathbb{C})$.

The symbol $\langle X, Y \rangle(t)$ denotes the quadratic covariation between the processes X and Y and sometimes it will also be represented by $\mathrm{d}X\mathrm{d}Y$. The derivative concerning a given variable x will be represented either using $\frac{\mathrm{d}}{\mathrm{d}x}$, $\frac{\partial}{\partial x}$ or ∂_x and following the convention in the literature we will sometimes denote $\frac{\mathrm{d}}{\mathrm{d}t} f = g$ as $\mathrm{d}f = g\mathrm{d}t$.

In the next section, we present some basic ideas in the study of random matrices.

1.1 Introduction to main concepts in Random Matrix Theory

This section is meant to introduce the essential objects we study in Random Matrix Theory and illustrate a few techniques that can be used to derive results. However, most of the relevant results are presented later with the tools introduced in the following sections.

1.1.1 Matrix algebra

Before stating the most specific concepts and results related to random matrices, it is important to mention a few purely algebraic well-known facts of matrix algebra as they will be useful in the thesis.

The Cauchy-Binet formula allows us to find the minor of a product of matrices in terms of the minors of the individual matrices.

Theorem 1.1.1 (Cauchy-Binet formula)**.** *Let m, n, p, k be integers, A an $m \times n$ matrix, and B an $n \times p$ matrix, then*

$$\det[AB]_{S,T} = \sum_{|U| \subset \binom{[n]}{k}} \det[A]_{S,U} \det[B]_{U,T},$$

where $S \in \binom{[m]}{k}, T \in \binom{[p]}{k}$.

The following Theorem taken from [3] can be seen as an equivalent to the Cauchy-Binet formula for sums of matrices.

Theorem 1.1.2. *Let k, n be integers such that $k \leq n$, A, B two $n \times n$ matrices, and $S, T \in \binom{[n]}{k}$. Then*

$$\det[A+B]_{S,T} = \sum_{i=0}^{k} \sum_{V \in \binom{[k]}{i}} (-1)^{\|U\|_1 + \|V\|_1} \det[A]_{U(S),V(T)} \det[B]_{\bar{U}(S),\bar{V}(T)},$$

with $\overline{U} = [k] \setminus U$.

A basic linear algebra theorem that has major relevance is the Spectral Theorem.

Theorem 1.1.3 (Spectral Theorem)**.** *Let A be an $n \times n$ self-adjoint matrix. Then there exists an orthonormal basis $v_1, \ldots, v_n \in \mathbb{R}^n$ and real eigenvalues $\lambda_1, \ldots, \lambda_n$ such that, for every $1 \leq i \leq n$.*

$$Av_i = \lambda_i v_i.$$

1.1.2 Random matrix ensembles

A random matrix R is simply a measurable function from a probability space to a space of matrices.

$$R : (\Omega, \mathscr{F}, \mathbb{P}) \to \mathcal{M}_{n,m}(\mathbb{F}).$$

In general Random Matrix Theory, the field for the entries $\mathbb{F}$ can be quite general but for the goals of this work, it is enough to consider $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$.

Given any self adjoint $n \times n$ matrix A, we can associate an empirical probability measure $\hat{\mu}$ to its set of eigenvalues $\lambda_1, \ldots, \lambda_n$ given by

$$\hat{\mu}(B) \coloneqq \frac{1}{n} \sum_{j=1}^{n} \mathbb{1}_B(\lambda_j).$$

We call $\hat{\mu}$ the empirical spectral measure.

If A is random (i.e. its entries are random variables), $\hat{\mu} : \Omega \times \mathscr{F} \to [0,1]$ is a random measure, which means that for every ω, $\hat{\mu}(\omega, \cdot)$ is a probability measure, and for every measurable set $B \in \mathscr{F}$, $\hat{\mu}(\cdot, B)$ is a real random variable. In this case, it is possible to define a deterministic empirical measure associated with A by simply taking the expectation of $\hat{\mu}$ on the measure of A.

$$\hat{\nu}(B) \coloneqq \mathbb{E}\left[\frac{1}{n}\sum_{j=1}^{n} \mathbb{1}_B(\lambda_j)\right] = \frac{1}{n}\sum_{j=1}^{n} P\left(\lambda_j \in B\right).$$

We call $\hat{\nu}$ the mean spectral measure.

We are usually interested in knowing if $\hat{\mu}$ and $\hat{\nu}$ converge to a given law when $n \to \infty$. The following examples show that this happens in some cases.

Example 1.1.1. Let $A = cI$ for $c \in \mathbb{R}$ and I the identity matrix with size $n \times n$. Then $\hat{\mu}_n(\{c\}) = \hat{\nu}_n(\{c\}) = 1$ for every n and

$$\lim_{n\to\infty} \hat{\nu}_n = \lim_{n\to\infty} \hat{\mu}_n = \delta_c.$$

Example 1.1.2. Let A be an $n \times n$ matrix with $A = \operatorname{diag}[u_1, u_2, \ldots, u_n]$ and $u_1, \ldots, u_n$ are i.i.d. random variables with uniform distribution on $[0,1]$. Then both $\hat{\mu}_n$ and $\hat{\nu}_n$ converge to the measure of a uniform random variable on $[0,1]$.

Using that the trace equals the sum of eigenvalues, we have that the expectation over $\hat{\nu}$ is equal to

$$\int_{\mathbb{C}} z\hat{\nu}(\mathrm{d}z) = \mathbb{E}\left[\frac{1}{n}Tr(A)\right].$$

The next identity allows us to compute similar moments of A.

Theorem 1.1.4 (Trace identity)**.** *Let A be a normal $n \times n$ matrix ($A^*A = AA^*$), then*

$$\frac{1}{n}Tr(A^k A^{*j}) = \frac{1}{n}\sum_{i=1}^{n} \lambda_i^k \overline{\lambda}_i^j = \int_{\mathbb{C}} z^k \bar{z}^j \hat{\mu}(\mathrm{d}x),$$

where $\hat{\mu}$ is the empirical spectral measure associated to A. If A is random and we take expectation over its probability law, we have

$$\int_{\mathbb{C}} x^k \bar{z}^j \hat{\nu}(\mathrm{d}z) = \frac{1}{n}\mathbb{E}\left[Tr(A^k A^{*j})\right].$$

In random matrix theory, it is common to work with matrix ensembles. An ensemble is a set of matrices with an associated probability measure.

Example 1.1.3 (Independent identically distributed entries ensemble)**.** If A is an $n \times n$ matrix whose entries $A_{ij}, 1 \le i \le n, 1 \le j \le n$ are all independent identically distributed random variables, we say that A is an i.i.d. ensemble.

Example 1.1.4 (Diagonal i.i.d. ensemble)**.** If D is a diagonal $n \times n$ matrix whose every entry is an i.i.d. random variable, then we say that D is a diagonal i.i.d. ensemble.

Example 1.1.5 (Gaussian invariant ensembles)**.** Let $\mathbb{H}$ denote the field of quaternions and $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$. If R is a hermitian matrix whose entries are standard normal random variables in $\mathbb{F}$ independent except for symmetries, then we say that R is a Gaussian invariant ensemble. Depending on $\mathbb{F}$, we have particular names for each ensemble.

- If $\mathbb{F} = \mathbb{R}$, we call R the Gaussian orthogonal ensemble.
- If $\mathbb{F} = \mathbb{C}$, we call R the Gaussian unitary ensemble.
- If $\mathbb{F} = \mathbb{H}$, we call R the Gaussian symplectic ensemble.

The specific names are given because the distribution of the eigenvalues of R remains unchanged under a conjugation by an orthogonal matrix (respectively unitary or symplectic). This property is analogous to the property of a vector of independent normal variables that preserve their distribution after being transformed by an orthogonal matrix.

Example 1.1.6 (Haar unitary ensemble)**.** If we consider $\mathcal{U}_{n,n}(\mathbb{C})$ the group of complex unitary matrices ($U^*U = I_n = UU^*$). In $\mathbb{U}_{n,n}(\mathbb{C})$, we can define a Haar measure unique up to a constant. If we normalize this measure we have the only Haar probability measure μ_U in $\mathbb{U}_{n,n}(\mathbb{C})$. A Haar unitary ensemble is a matrix sampled from μ_U.

Example 1.1.7 (Wigner example)**.** Let W be a random self-adjoint matrix whose every entry is an i.i.d. random variable except for the symmetries. Then we say that W is a Wigner ensemble. Notice in particular that the Gaussian invariant ensembles are Wigner ensembles. This is one of the first ensembles studied.

Example 1.1.8 (Wishart ensemble)**.** Let R be an $n \times n$ i.i.d. standard Gaussian ensemble and define $E \coloneqq R^T R$, then we say E is a Wishart ensemble. This ensemble is used to model covariance matrices.

The next results give the convergence of an empirical spectral measure to a continuous probability measure in $\mathbb{R}$ when the matrix dimension n tends to infinity. The limit only depends on the first two moments of the variables involved and not on the whole distribution, thus it can be seen as a matrix analogous to the Central Limit Theorem.

Theorem 1.1.5 (Wigner's semicircle law [5])**.** *For each $n \in \mathbb{N}$, let $W^{(n)}$ be a Wigner ensemble and its entries $W_{ij}^{(n)}$ satisfy the following conditions.*

- $\mathbb{E}\left[|W_{ij}^{(n)}|^k\right] < \infty$ *for all* $k \in \mathbb{N}$.
- $\mathbb{E}\left[W_{ij}\right] = 0$ *for every* $1 \leq i \leq n, 1 \leq j \leq n$.
- $\mathbb{E}\left[W_{ij}^{(n)2}\right] = 1/\sqrt{n}$.

Then both $\hat{\mu}$ and $\hat{\nu}$ converge in distribution to the semicircle distribution, i.e. the absolutely continuous distribution with density,

$$f(x) = \frac{\sqrt{4 - x^2}}{2\pi}.$$

Moreover, $\hat{\mu}$ converges almost surely.

1.2 Stochastic Calculus

When we work with continuous-time processes, stochastic calculus is one of the most useful tools. The results provided by the Itô integral and related concepts allow us to study dynamical systems with random behavior. In this section, we introduce some stochastic calculus used in the thesis. The first part deals with the definition and main properties of Itô and Stratonovich integrals in $\mathbb{R}$-valued processes. Later, we generalize the definition to $\mathbb{R}^d$-valued processes, and finally, to processes taking values in spaces of matrices.

1.2.1 Stochastic calculus for $\mathbb{R}^n$-valued processes

We start by defining the Itô integral for $\mathbb{R}$-valued processes and stating its main properties. In all of the following definitions, we consider that we are working on a filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, P)$, and we use the convention that a continuous-time stochastic process is a process indexed by $\mathbb{R}^+ = [0, \infty)$. Most of the definitions and results presented here are taken from [6]. A more general approach to the subject can be found in [7] or [8].

There are several definitions of the Itô integral, some of which are more general. For simplicity, we will use one that resembles the definition of the Lebesgue integral. Consider a simple process $X = (X(t))_{t\geq 0}$, adapted to the filtration $\mathscr{F}_t$, i.e.,

$$X(t) = \xi_0 \delta_{0,t} + \sum_{j=0}^{n-1} \xi_j \mathbb{1}_{(t_j, t_{j+1}]}(t),$$

for some $0 = t_0 < t_1 < \cdots < t_n = T$ and variables ξ_i that are $\mathscr{F}_{t_i}$-measurables. For simple adapted processes, the Itô integral with respect to a Brownian motion can be defined as

Definition 1.2.1. Let X be a simple process adapted to $\mathscr{F}_t$ and square integrable. Let $B = (B(t))_{t\geq 0}$ be a Brownian motion adapted to $\mathscr{F}_t$. The Itô integral of X with respect to B on $[0, T]$ is

$$\int_0^T X(s) \mathrm{d}B(s) \coloneqq \sum_{j=0}^{n-1} \xi_j (B(j+1) - B(j)).$$

It is not hard to verify from the definition that the integral is linear and has zero mean. Also, it can be proven that it satisfies the so-called Itô isometry,

$$\mathbb{E}\left[\left(\int_0^T X(s) \mathrm{d}B(s)\right)^2\right] = \int_0^T \mathbb{E}\left[X^2(s)\right] \mathrm{d}s.$$

Another important fact is that the Itô integral is a martingale. These last two properties are the most important, and one typically aims to preserve them when defining a stochastic integral for more general processes. The proof of these properties can be found in [9], [6] or [7].

If we have a square-integrable continuous time process $X(t)$ adapted to $\mathscr{F}_t$ and a sequence of simple adapted processes $X_n(t)$ that are also square-integrable and converge in probability to $X(t)$, then under proper conditions, one can prove that the integrals $\int_0^T X_n(s) \mathrm{d}B(s)$ also have a limit in probability. We define the Itô integral for general adapted processes as this limit.

Definition 1.2.2. Let X be a square-integrable adapted process to $\mathscr{F}_t$ and $(X_n(t))_{n\geq 0}$ a sequence of simple adapted, square-integrable processes converging in probability to $X(t)$. We define the integral of $X(t)$ with respect to a Brownian motion $B(t)$ as

$$\int_0^T X(s)\mathrm{d}B(s) = \overset{P}{\lim_{n\to\infty}} \int_0^T X_n(s)\mathrm{d}B(s).$$

As we have mentioned before, when extending the definition of the Itô integral to more general processes, one wishes to preserve its good properties. The next theorem allows us to characterize when the integral has these properties

Theorem 1.2.1. *If $X = (X(t), t \geq 0)$ and $Y = (Y(t), t \geq 0)$ are regular adapted process and satisfy*

$$P\left(\int_0^T X^2(s)\mathrm{d}s < \infty\right) = 1, \qquad P\left(\int_0^T Y^2(s)\mathrm{d}s < \infty\right) = 1,$$

then the integrals $\int_0^T X(s)\mathrm{d}B(s), \int_0^T Y(s)\mathrm{d}B(s)$ exist and it satisfy the following properties,

1. *Linearity. For $\alpha, \beta \in \mathbb{R}$,*

$$\int_0^T (\alpha X(s) + \beta Y(s))\mathrm{d}B(s) = \alpha \int_0^T X(s)\mathrm{d}B(s) + \beta \int_0^T Y(s)\mathrm{d}B(s). \tag{1.1}$$

If additionally, the process X satisfies

$$\int_0^T \mathbb{E}\left[X^2(s)\right]\mathrm{d}s < \infty,$$

then the integral $\int_0^T X(s)\mathrm{d}B(s)$ has the following properties

2. *Martingale property. For $t \leq T$*

$$\mathbb{E}\left[\int_0^T X(s)\mathrm{d}B(s)\middle|\mathscr{F}_t\right] = \int_0^t X(s)\mathrm{d}B(s) = \int_0^T X(s)\mathbb{1}_{[0,t]}(s)\mathrm{d}B(s).$$

3. *Itô's isometry.*

$$\mathbb{E}\left[\int_0^T X(s)\mathrm{d}B(s)\right]^2 = \int_0^T \mathbb{E}\left[X^2(s)\right]\mathrm{d}s.$$

For a proof of last theorem, consult [6]. If we consider, for a given adapted process X, the Itô integral process $I = (I(t), t \geq 0)$ defined as

$$I(t) \coloneqq \int_0^t X(s)\mathrm{d}B(s),$$

we know that this process is a martingale. We can also define its quadratic variation $\mathrm{d}\langle I, I\rangle(t)$ as a limit in probability,

$$\mathrm{d}\langle I, I\rangle(t) \coloneqq \overset{P}{\lim_{\delta_n\to 0}} \sum_{i=0}^{n-1} (Y(t_{j+1,n}) - Y(t_{j,n}))^2,$$

with $\{t_{j,n}\}_{j=1}^{n}$ a partition of $[0,t]$ for every n and $\delta_n = \sup_j \{t_{j+1,n} - t_{j,n}\} \to 0$ as $n \to \infty$.

The following theorem gives us a way to find explicitly the quadratic variation of an Itô integral process,

Theorem 1.2.2. *The quadratic variation of an Itô integral is*

$$\mathrm{d}\langle I(t), I(t)\rangle(t) = \mathrm{d}\langle \int_0^t X(s)\mathrm{d}B(s), \int_0^t X(s)\mathrm{d}B(s)\rangle(t) = \int_0^t X^2(s)\mathrm{d}s.$$

We can notice in Definition 1.2.1 that when we define the Itô integral, we care about in which point of the interval $(t_i, t_{i+1}]$ we evaluate $X(t)$. This is not casual and we want to take the lower value so that the resulting process is a martingale. However, this causes several results from classical calculus not to be generalized in Itô calculus. One of the main results in Itô calculus is the following theorem, which allows us to prove some other important facts.

Theorem 1.2.3 (Itô formula for Brownian motion)**.** *Let f be a twice differentiable function and B a Brownian motion, then*

$$f(B(t)) = f(0) + \int_0^t f'(B(s))\mathrm{d}B(s) + \frac{1}{2}\int_0^t f''(B(s))\mathrm{d}s.$$

The proof is done using a second order Taylor expansion for f and taking a limit. We will omit it here, but it is important to notice that (under suitable conditions) the Itô formula allows us to write a function of a Brownian motion as a sum of a martingale and a finite variation process.

We can define the Itô integral with respect to processes other than the Brownian motion. The class of processes for which we can define an Itô integral is rather large, so we will define a smaller class that is of interest for this work.

Definition 1.2.3 (Itô process)**.** We say that an $\mathscr{F}_t$-adapted process $Y = (Y(t), 0 \le t \le T)$ is an Itô process if there exist $\mu = (\mu(t), 0 \le t \le T)$, $\sigma = (\sigma(t), 0 \le t \le T)$ adapted processes such that $\int_0^T |\mu(s)|\mathrm{d}s < \infty$, $\int_0^T \sigma^2(s)\mathrm{d}s < \infty$, and $Y(0)$ an $\mathscr{F}_0$-measurable variable that satisfy

$$Y(t) = Y(0) + \int_0^t \mu(s)\mathrm{d}s + \int_0^t \sigma(s)\mathrm{d}B(s). \tag{1.2}$$

It is a usual convention to write the "differential" of an Itô process as

$$\mathrm{d}Y(s) = \mu(s)\mathrm{d}s + \sigma(s)\mathrm{d}B(s).$$

This notation only means that Y satisfies (1.2). We usually call $\int_0^t \mu(s)\mathrm{d}s$ the finite variation part of Y and $\int_0^t \sigma(s)\mathrm{d}B(s)$ the martingale part of Y.

Using the fact that the covariation of any function with a finite variation function is zero, we can find that the quadratic variation of an Itô process is given by

$$\mathrm{d}\langle Y, Y\rangle(t) = \int_0^t \sigma^2(s)\mathrm{d}s.$$

Now we can define the integral of an adapted process X with respect to an Itô process Y.

Definition 1.2.4 (Itô integral with respect to an Itô process)**.** Let Y be an adapted process such that its Itô integral exists for every t in $[0,T]$. Let Y be an Itô process $\mathrm{d}Y = \mu\mathrm{d}s + \sigma\mathrm{d}B$ and X, Y satisfy

$$\int_0^T |X(s)\mu(s)|\mathrm{d}s < \infty,$$
$$\int_0^T X^2(s)\sigma^2(s)\mathrm{d}s < \infty.$$

Then, the integral of X with respect to Y is defined, for $0 \leq t \leq T$ as

$$\int_0^t X(s)\mathrm{d}Y(s) = \int_0^t X(s)\mu(s)\mathrm{d}s + \int_0^t X(s)\sigma(s)\mathrm{d}B(s).$$

Although the definition of an Itô integral with respect to an Itô process can be given more directly, it coincides with the last one. In a similar spirit as in the definition of the Itô integral with respect to an Itô process, we can also extend Theorem 1.2.3 for Itô processes.

Theorem 1.2.4 (Itô formula for Itô processes)**.** *Let Y be an Itô process satisfying $\mathrm{d}Y = \mu\mathrm{d}s+\sigma\mathrm{d}B$ and f be a twice continuously differentiable function, then the stochastic differential of $f(Y(t))$ is well defined and is given by*

$$\begin{aligned}\mathrm{d}f(Y(t)) &= f'(Y(t))\mathrm{d}Y(t) + \frac{1}{2}f''(Y(t))\mathrm{d}\langle Y,Y\rangle(t),\\ &= \left(f'(Y(t))\mu(t) + \frac{1}{2}f''(Y(t))\sigma^2(t)\right)\mathrm{d}t + f'(Y(t))\sigma(t)\mathrm{d}B(t).\end{aligned}$$

Although the Itô integral is the most common one, it is not the only notion of stochastic integration and some others can be used in certain contexts. One of the most used alternatives is the Stratanovich integral, which preserves several properties of standard calculus. The Stratanovich integral is useful when we deal with random matrix calculus because it allows to simplify calculations.

Definition 1.2.5 (Stratanovich integral)**.** Let X and Y be two continuous adapted processes. The Stratanovich integral of X with respect to Y denoted as $\int_0^t X(s)\partial Y(s)$ is the L^2 limit of the sums

$$\sum_{i=0}^{n-1}\frac{1}{2}(X(t_{i+1,n}) + X(t_{i,n}))(Y(t_{i+1,n}) - Y(t_{i,n})).$$

as $\delta_n = \delta_n = \sup_j \{t_{j+1,n} - t_{j,n}\} \to 0$.

The main difference between the Itô and Stratanovich integrals is the point we take for the evaluation of the integrand process in the interval $(t_{i,n}, t_{i+1,}]$. While we take the left point in the Itô integral, we take the average between the extremes for the Stratanovich one. Both integrals happen to be related by the following result.

Theorem 1.2.5 (Relationship between Itô and Stratanovich integrals)**.** *Let X, Y be two continuous adapted processes such that the Itô integral of X with respect to Y is well-defined. The Stratanovich integral of X with respect to Y is*

$$\int_0^t X(s)\partial Y(s) = \int_0^t X(s)\mathrm{d}Y(s) + \frac{1}{2}\langle X, Y\rangle(t).$$

By the last theorem we can write the Stratanovich differential similarly to the Itô differential as

$$Y(s)\partial X(s) = Y(s)\mathrm{d}X(s) + \frac{1}{2}\mathrm{d}\langle X, Y\rangle(t).$$

Sometimes we write this differential as $Y(s)\partial X(s) = Y(s) \circ \mathrm{d}X(s)$. This notation is especially helpful when we work with matrix-valued processes.

Perhaps the main situation when the Stratanovich integral is used instead of the Itô version is when we want to preserve the classical integration by parts formula. The next theorem uses the relationship between both integrals to compute the differential of the product XY.

Theorem 1.2.6 (Integration by parts for Itô and Stratanovich integral [8])**.** *Let X, Y be two adapted processes such that the integrals $\int_0^t X(s)\mathrm{d}Y(s)$ and $\int_0^t Y(s)\mathrm{d}X(s)$ are well defined, then*

$$\begin{aligned}
\mathrm{d}(XY) &= X\mathrm{d}Y + Y\mathrm{d}X + \mathrm{d}\langle X, Y\rangle, \\
&= X\mathrm{d}Y + \frac{1}{2}\mathrm{d}\langle X, Y\rangle + Y\mathrm{d}X + \frac{1}{2}\mathrm{d}\langle X, Y\rangle = X\partial Y + Y\partial X, \\
&= X \circ \mathrm{d}Y + Y \circ \mathrm{d}X.
\end{aligned}$$

Notice that in the case of the Stratanovich integral, we recover the classical integration by parts formula.

The next results are technical but needed for the proofs in Chapter 2. The first one states the existence of a process called local time and the stochastic differential equation it satisfies, the second one gives a way to prove when this local time process is zero. Both results with their proofs acn be found in [8]

Theorem 1.2.7 (Tanaka's formula)**.** *Let X be a continuous semimartingale. For any real number a, there exists an increasing continuous process L^a called the local time of X in a such that,*

$$\begin{aligned}
|X(t) - a| &= |X(0) - a| + \int_0^t \operatorname{sgn}(X(s) - a)\ \mathrm{d}X(s) + L^a(t), \\
(X(t) - a)^+ &= (X(0) - a)^+ + \int_0^t \mathbb{1}_{\{X(s) > a\}}\ \mathrm{d}X(s) + \frac{1}{2}L^a(t), \\
(X(t) - a)^- &= (X(0) - a)^- - \int_0^t \mathbb{1}_{\{X(s) \leq a\}}\ \mathrm{d}X(s) + \frac{1}{2}L^a(t).
\end{aligned}$$

Theorem 1.2.8. *Let $\rho : (0, \infty) \to (0, \infty)$ a measurable function that satisfies*

$$\int_{0^+} \frac{\mathrm{d}s}{\rho(s)} = \infty.$$

If X is a continuous semimartingale such that, for some $\epsilon > 0$ and every t, the process

$$A_t = \int_0^t \mathbb{1}_{\{0<X(s)\leq\epsilon\}}\rho(X(s))^{-1}\ \mathrm{d}\langle X,X\rangle(s) < \infty \qquad a.s.,$$

then $L^0(X) = 0$.

Gronwall's lemma allows us to bound a function satisfying a differential inequality by the solution of the associated differential equation. It will be useful for the multidimensional version of the Yamada-Watanabe Theorem. The proof is in [9, page 213].

Lemma 1.2.9 (Gronwall's lemma)**.** *Let $T > 0$ and g be any nonnegative bounded measurable function on $[0,T]$. Assume that there exists two constants $a \geq 0$ and $b \geq 0$ such that for every $t \in [0,T]$,*

$$g(t) \leq a + b\int_0^t g(s)\ \mathrm{d}s.$$

Then we also have, for every $t \in [0,T]$,

$$g(t) \leq a\exp(bt).$$

Below it is a generalized version of the well-known Mckean's principle. This result gives solutions for a stopping time to be infinite a.s. In particular, it is used when working with eigenvalue processes to conclude the non-collision of the eigenvalues. This generalization and its proof appear in [10].

Lemma 1.2.10 (Generalized McKean's argument)**.** *Let $Z = (Z_s)_{s\in\mathbb{R}_+}$ be an adapted càdlàg $\mathbb{R}^+\setminus\{0\}$-valued stochastic process on a stochastic interval $[0,\tau_0)$ such that $Z_0 > 0$ a.s. and $\tau_0 = \inf\{0 < s \leq \tau_0 : Z_{s-} = 0\}$. Suppose that $h : \mathbb{R}_+\setminus\{0\} \to \mathbb{R}$ is continuous and satisfies the following:*

1. *For all $t \in [0,\tau_0)$, we have $h(Z_t) = h(Z_0) + M_t + P_t$, where*
   (a) *P is an adapted càdlàg process on $[0,\tau_0)$ such that $\inf_{t\in[0,\tau_0\wedge T]} P_t > -\infty$ a.s. for each $T \in \mathbb{R}_+\setminus\{0\}$,*
   (b) *M is a continuous local martingale on $[0,\tau_0)$ with $M_0 = 0$,*
2. *$\lim_{z\to 0} h(z) = -\infty$.*

Then $\tau_0 = \infty$ a.s.

Stochastic Calculus for $\mathbb{R}^n$-valued processes

If we have a continuous time stochastic process taking values in $\mathbb{R}^n$, it is possible to define the Itô integral with respect to a multivariate Brownian motion. Based on this, we can extend several results of univariate stochastic calculus, this can then be used for introducing stochastic calculus for matrix-valued processes.

An $\mathbb{R}^n$-valued Brownian motion $\vec{B} = \{(B_1(t),\ldots,B_n(t)), t \geq 0\}$ is an n length vector whose every entry B_i is an independent Brownian motion in $\mathbb{R}$. An n-dimensional process $\vec{X}$ is said to be adapted to a filtration $\mathscr{F}$ if each of its entries is. If X_i is the ith entry of $\vec{X}$ and for every i we have that

$$\int_0^T X_i^2(s)\mathrm{d}s < \infty,$$

then we can define the Itô itegral of $\vec{X}$ with respect to $\vec{B}$ in $0 \leq t \leq T$ as

$$\int_0^t \vec{X}(s) \cdot \mathrm{d}\vec{B}(s) \coloneqq \sum_{j=1}^n \int_0^t X_j(s)\mathrm{d}B_j(s).$$

Notice that the integral notation suggests the similarity with a dot product. Similarly, we also denote the multivariate integral as $\vec{X}(s) \cdot \mathrm{d}\vec{B}(s)$.

The process $\sum_{j=1}^n \int_0^t X_j(s)\mathrm{d}B_j(s)$ takes values in $\mathbb{R}$. If we add a finite variation part μ, then we can have a process Y similar to an Itô process but driven by a multidimensional Brownian motion,

$$\mathrm{d}Y(s) = \mu(s)\mathrm{d}s + \sum_{j=1}^n X_j(s)\mathrm{d}B_j(s).$$

If we take $\vec{\mu}(t) = (\mu_1(t), \ldots, \mu_n(t))$ to be a vector of integrable functions and for each $i \in [n]$ we consider a vector-valued process $\vec{\sigma}_i(t) = (\sigma_{i1}(t), \ldots, \sigma_{in}(t))$, then for each $i \in [n]$ we have a single dimensional Itô process driven by a multivariate Brownian motion,

$$\mathrm{d}Y_i = \mu_i(s)\mathrm{d}s + \sum_{j=1}^n \sigma_{ij}(s)\mathrm{d}B_j(s).$$

By taking $\vec{Y} = (Y_1, \ldots, Y_n)$ we have an n-dimensional Itô process, which is denoted in differential form by

$$\mathrm{d}\vec{Y}(s) = \vec{\mu}(s)\mathrm{d}s + \Sigma(s)\mathrm{d}\vec{B}(s),$$

with Σ an $n \times n$ matrix valued function with entries σ_{ij}.

Before giving the Itô formula for multidimensional processes, we need to know the quadratic covariation between entries of a multidimensional Itô process.

Theorem 1.2.11. *Let $\vec{Y}$ be an n-dimensional Itô process, then the quadratic covariation of two of its entries Y_i, Y_j is given by*

$$\langle Y_i, Y_j \rangle(t) = \int_0^t \left(\Sigma\Sigma^T\right)_{ij}(t)\mathrm{d}t.$$

The matrix $\Sigma\Sigma^T$ is often called the diffusion matrix. We generalize the Itô formula for multidimensional Itô processes in the following theorem.

Theorem 1.2.12 (Multidimensional Itô formula)**.** *Let $\vec{Y}$ be an n-dimensional Itô process and $f : \mathbb{R}^n \to R^m$ be a C^2 function. The process $f(Y_1(t), \ldots, Y_n(t))$ is also an Itô process and has a stochastic differential given by*

$$\mathrm{d}f(Y_1(t),\dots,Y_n(t)) = \sum_{i=1}^{n}\frac{\partial}{\partial x_i}f(Y_1(t),\dots,Y_n(t)) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2}{\partial x_i\partial x_j}f(Y_1(t),\dots,Y_n(t))\mathrm{d}\langle Y_i,Y_j\rangle(t).$$

In particular when $n=2$, $Y_1(t)=Y(t)$ for some Itô process $\mathrm{d}Y=\mu\mathrm{d}t+\sigma\mathrm{d}B$, $Y_2(t)=t$ and $f:\mathbb{R}^2\to\mathbb{R}$, we have that

$$\mathrm{d}f(Y(t),t) = \frac{\partial f}{\partial x}(Y(t),t)\mathrm{d}Y(t) + \frac{\partial f}{\partial t}(Y(t),t)\mathrm{d}t + \frac{1}{2}\sigma^2(t)\frac{\partial^2 f}{\partial x^2}(Y(t),t)\mathrm{d}t. \tag{1.3}$$

Infinitesimal generator and harmonic functions

Every Itô process is Markovian and thus it has an associated Markov semigroup and infinitesimal generator. These operators can tell us many of the properties of the processes and particularly the infinitesimal generator can be used to prove that some transformations of an Itô process are martingales.

Definition 1.2.6. Let X be an Itô process with $\mathrm{d}X(t)=\mu(X,t)\mathrm{d}t+\sigma(X,t)\mathrm{d}B(t)$. The infinitesimal generator of X is the second order differential operator $\mathcal{A}_t$,

$$\mathcal{A}_t f(x,t) = (\mathcal{A}_t f)(x,t) = \frac{1}{2}\sigma^2(x,t)\frac{\partial^2 f}{\partial x^2}(x,t) + \mu(x,t)\frac{\partial f}{\partial x}(x,t).$$

With this definition, we can re-write equation (1.3) as

$$\mathrm{d}f(Y(t),t) = \left(\mathcal{A}_t f(Y(t),t) + \frac{\partial f}{\partial t}(Y(t),t)\right)\mathrm{d}t + \frac{\partial f}{\partial x}(Y(t),t)\sigma(Y(t),t)\mathrm{d}B(t).$$

If the integral $\int_0^t \frac{\partial f}{\partial x}(X(s),s)\sigma(X(s),s)\mathrm{d}B(s)$ is a martingale, then the process $f(Y(t),t)-\int_0^t\left(\mathcal{A}_s f(Y(s)) + \frac{\partial f}{\partial s}(Y(s),s)\right)\mathrm{d}s$ is a martingale. This result is stated in the following theorem

Theorem 1.2.13. *Let $Y(t)$ be an Itô process with differential $\mathrm{d}Y(t)=\mu(Y(t),t)\mathrm{d}t+\sigma(Y(t),t)\mathrm{d}B(t)$ such that $\mu(x,t)$ and $\sigma(x,t)$ are Lipschitz in x with the same constant for every t and satisfy*

$$|\mu(x,t)|+|\sigma(x,t)|\le K(1+|x|).$$

If $f(x,t)$ is a twice continuously differentiable function in x and once in t with $\partial_x f$ bounded, then the process

$$M^f(t) \coloneqq f(Y(t),t) - \int_0^t\left(\mathcal{A}_s f(Y(s),s) + \frac{\partial f}{\partial s}(Y(s),s)\right)\mathrm{d}s,$$

is a martingale.

We have in particular that under the same conditions when $\mathcal{A}_t f(Y(t)) + \frac{\partial f}{\partial t}(Y(t),t)=0$, $f(Y(t),t)$ is a martingale. If f only depends on x, this is equivalent to asking it to be a solution to $\mathcal{A}_t f=0$. These functions are known as harmonic functions for the process Y.

Complex stochastic calculus

It is possible to define continuous time stochastic processes in more general fields than $\mathbb{R}$ and then create a notion of stochastic integral for these processes. Particularly, in the case of random matrix theory, we care about processes taking values in the field of complex numbers ($\mathbb{C}$) and the field of quaternions ($\mathbb{H}$). Provided that both spaces can be seen as a vector space with $\mathbb{R}$ as the field of scalars, the extension of the definitions is natural by considering that every process in $\mathbb{C}$ or $\mathbb{H}$ has the form $A + iB$ or $A + iB + jC + kD$, respectively, with A, B, C, D stochastic processes taking values in $\mathbb{R}$.

Example 1.2.1 (Brownian motion in $\mathbb{C}$)**.** Let B_1 and B_2 be independent Brownian motions taking values in $\mathbb{R}$. We say that $Z = B_1 + iB_2$ is a Brownian motion in $\mathbb{C}$.

Some of the matrix-valued processes in this thesis have entries in $\mathbb{C}$. Thus, introducing the following result for complex Brownian motions results useful.

Theorem 1.2.14. *Let Z be a complex Brownian motion, then its quadratic covariation and the quadratic variation with respect to its complex conjugate are given by*

$$\langle Z, Z\rangle(t) = \langle B_1 + iB_2, B_2 + iB_2\rangle(t) = \langle B_1, B_1\rangle(t) - \langle B_2, B_2\rangle(t) = 0,$$
$$\langle Z, \overline{Z}\rangle(t) = \langle B_1 + iB_2, B_2 - iB_2\rangle(t) = \langle B_1, B_1\rangle(t) + \langle B_2, B_2\rangle(t) = 2t.$$

1.2.2 Stochastic calculus for matrix-valued processes

Similarly, as we can generalize the stochastic calculus results for $\mathbb{R}^n$-valued processes, it is possible to extend the definitions and results to matrix-valued processes. Given a filtered probability space $(\Omega, \mathscr{F}\mathscr{F}_t, P())$, an $n \times m$ continuous-time matrix-valued process M is a function

$$M : \mathbb{R}^+ \times \Omega \to \mathcal{M}_{m,n}(\mathbb{F}),$$
$$(t, \omega) \mapsto M(t, \omega),$$

where for every fixed ω^*, $M(t, \omega^*)$ is a function from $\mathbb{R}^+$ to $\mathcal{M}_{m,n}(\mathbb{F})$ and for every fixed t^*, $M(t^*)$ is a random matrix. $\mathbb{F}$ represents an arbitrary field for the entries, $\mathbb{R}, \mathbb{C}$ or $\mathbb{H}$ are usual choices, but in this thesis we are only interested in matrix-valued processes with entries in $\mathbb{R}$ and $\mathbb{C}$.

Usually, we need the matrix-valued process to satisfy some symmetry condition such as being symmetric, hermitian or orthogonal. It is common to restrict the matrix-valued process to take values in a smaller subset of $\mathcal{M}_{m,n}$. Along this work, we are only interested in squared matrix-valued processes.

Example 1.2.2 (Brownian motion in $\mathcal{M}_{n,n}(\mathbb{F})$)**.** We say that a matrix-valued process $B = (B(t), t \geq 0)$ is a standard Brownian motion in $M_{n,n}(\mathbb{F})$ if every entry B_{ij} is an independent Brownian motion in the field $\mathbb{F}$.

Example 1.2.3 (Symmetric Brownian motion)**.** Let W be an $n \times n$ symmetric matrix-valued stochastic process. We say that W is a standard Brownian motion in the space of symmetric matrices if every entry W_{ij} is a real Brownian motion independent of all the other entries, except for the symmetries.

Now we show the definition of the Itô integral for matrix-valued processes.

Definition 1.2.7 (Itô integral with respect to a matrix-valued Itô process)**.** Let $W = (W(t), t \geq 0)$ be a matrix-valued Brownian motion in $\mathcal{M}_{n,m}(\mathbb{F})$ and let X and Y be two adapted matrix-valued processes in $\mathcal{M}_{p,n}(\mathbb{F})$ and $\mathcal{M}_{m,q}(\mathbb{F})$, respectively. The ij entry of the Itô integral $\int_0^t (X(s)\mathrm{d}W(s)Y(s))$ is defined as,

$$\left(\int_0^t (X(s)\mathrm{d}W(s)Y(s))\right)_{ij} = \sum_{k,l} \int_0^t X_{ik}(s)Y_{lj}(s)\mathrm{d}W_{kl}(s),$$

where $1 \leq k \leq n$, $1 \leq l \leq m$, $1 \leq i \leq p$ and $1 \leq j \leq q$.

The definition above applies also when we are integrating with respect to a Brownian motion in a smaller subspace of $\mathcal{M}_{n,m}(\mathbb{F})$. An interesting property of the stochastic matrix integral is that one can integrate by the left or right. This operation needs not to be commutative, even if it is well defined in both cases.

Just as in the $\mathbb{R}$ and $\mathbb{R}^n$ case, we can enlarge the class of process with respect to which we can integrate. It is convenient to define such processes only in spaces of squared matrices.

Definition 1.2.8 (Matrix-valued Itô process)**.** Let B be a Brownian motion in $\mathcal{M}_{n,n}(\mathbb{F})$ and S, R, M be adapted matrix-valued processes taking values in $\mathcal{M}_{n,n}(\mathbb{F})$. Then we say the process X satisfying

$$\mathrm{d}X(t) = S(t)\mathrm{d}B(t)R(t) + M(t)\mathrm{d}t,$$

is an Itô process in $\mathcal{M}_{n,n}(\mathbb{F})$.

In particular, if $\mathbb{F} = \mathbb{R}$ and M is symmetric, we have that Y satisfying

$$\mathrm{d}Y(t) = S(t)\mathrm{d}B(t)R(t) + R(t)\mathrm{d}B(t)^T S(t) + M(t)\mathrm{d}t,$$

is an Itô process in the space of symmetric matrices with real coefficients.

The definition of an Itô integral with respect to an Itô process in $\mathcal{M}_{n,n}(\mathbb{F})$ is a direct extension of the definition of Itô integral with respect to an Itô process in $\mathbb{R}$.

The quadratic covariation between two matrix-valued processes X, Y taking values in $\mathcal{M}_{nm}(\mathbb{F})$ and $\mathcal{M}_{mp}(\mathbb{F})$ is the matrix $\langle X, Y\rangle(t)$ with entries given by

$$\langle X, Y\rangle_{ij}(t) = \sum_{=1}^{m} \langle X_{ik}, Y_{kj}\rangle(t).$$

The same applies when we find the quadratic variation of a matrix-valued process. Now we state the Itô formula for matrix-valued processes. This is taken from [11]

Theorem 1.2.15 (Itô Formula for matrix-valued processes)**.** *Let $U \subset \mathcal{M}_{m,n}(\mathbb{R})$ an open set, X a continuous semimartingale taking values in U and $f : U \to \mathbb{R}$ twice continuously differentiable. Then $f(X)$ is a continuous semimartingale and*

$$\begin{aligned} f(X(t)) = f(X(0)) &+ \mathrm{Tr}\left(\int_0^t Df(X(s))^T \mathrm{d}X(s)\right) \\ &+ \frac{1}{2}\sum_{j,l=1}^{n}\sum_{i,k=1}^{m} \int_0^t \frac{\partial^2}{\partial X_{ij}\partial X_{kl}} f(X(s))\mathrm{d}\langle X_{ij}, X_{kl}\rangle(s). \end{aligned}$$

The next version of matrix-valued integration by parts formula appears in [12] and it is extensively used along the thesis.

Theorem 1.2.16 (Integration by parts for matrix-valued processes)**.** *Let X and Y be two matrix-valued semimartingales taking values in $\mathcal{M}_{nm}(\mathbb{F})$ and $\mathcal{M}_{np}(\mathbb{F})$, respectively. Then the differential of the product X^TY is*

$$\mathrm{d}(X^TY) = X^T(\mathrm{d}Y) + (\mathrm{d}X)^TY + (\mathrm{d}X)^T(\mathrm{d}Y).$$

By extending the definition of the Stratanovich integral first to multivariate processes and then to matrix-valued ones, we can see that in general, if X and Y are matrix-valued continuous semimartingales, then

$$Y^T(\partial X) = Y^T(\mathrm{d}X) + \frac{1}{2}(\mathrm{d}Y)^T(\mathrm{d}X).$$

Using this fact, we can write Theorem 1.2.16 in Stratanovich notation as

Theorem 1.2.17 (Integration by parts for matrix-valued Stratanovich integrals)**.** *Let X and Y be two matrix-valued semimartingales taking values in $\mathcal{M}_{nm}(\mathbb{F})$ and $\mathcal{M}_{np}(\mathbb{F})$, respectively. Then the differential of the product X^TY is*

$$\mathrm{d}(X^TY) = X^T(\partial Y) + (\partial X)^TY.$$

1.3 Free probability

Although random matrices are constructed based on classical definitions of random variables as measurable functions on a given probability space, they differ fundamentally from classical random variables in their basic algebraic properties. Specifically, the product of random matrices does not necessarily commute, whereas the product of real-valued random variables always commutes. This simple fact complicates the use of classical analytical tools in probability to study random matrices. The emerging field of non-commutative probability has recently developed several techniques for the algebraic study of non-commutative random variables. These techniques have significantly impacted the theory of random matrices.

In this section, we briefly introduce the fundamental concepts and ideas of Free probability and its application to the study of random matrices. These concepts will be used later in Chapter 3.

1.3.1 Non-commutative probability spaces

A classical probability space is defined based on its analytical structure. When we work in non-commutative probability spaces, we are mainly concerned with the algebraic relationship between random variables. Before introducing the concept of a non-commutative probability space, we define an unital algebra.

Definition 1.3.1 (Unital algebra)**.** Let $\mathcal{A}$ be a vector space over the field $\mathbb{F}$ equipped with an additional binary operation $\cdot$. We say that $\mathcal{A}$ is an unital algebra if it satisfies the following properties for $a, b, c \in \mathcal{A}$ and $\alpha, \beta \in \mathbb{F}$:

- Right distributivity. $(a + b) \cdot c = a \cdot c + b \cdot c$.

- Left distributivity. $a \cdot (b + c) = a \cdot b + a \cdot c$.
- Compatibility with scalars. $(\alpha a) \cdot (\beta b) = (\alpha\beta)(a \cdot b)$.
- Identity. There is an element $1_{\mathcal{A}} \in \mathcal{A}$ such that $a1_{\mathcal{A}} = a = 1_{\mathcal{A}}a$. We call $1_{\mathcal{A}}$ the identity element in $\mathcal{A}$.

Typically, we denote $a \cdot b$ as ab to simplify notation. Now, we define a non-commutative probability space as found in [13].

Definition 1.3.2 (Non-commutative probability space)**.** A non-commutative probability space $(\mathcal{A}, \varphi)$ consists of a unital algebra $\mathcal{A}$ over $\mathbb{C}$ and a unital linear functional φ, i.e.

$$\varphi : \mathcal{A} \to \mathbb{C}, \qquad \varphi(1_{\mathcal{A}}) = 1.$$

An element a in $\mathcal{A}$ is called a non-commutative random variable in $(\mathcal{A}, \varphi)$ or simply a random variable in $(\mathcal{A}, \varphi)$.

A concept related to non-commutative probability is that of $*$-probability space (star probability space). We introduce the definition of a $*$-algebra.

Definition 1.3.3 ($*$-Algebra)**.** We say that a unital algebra $\mathcal{A}$ is a $*$-algebra if it is equipped with an antilinear operation $*$, $\mathcal{A} \ni a \mapsto a^* \in \mathcal{A}$ that satisfies $(a^*)^* = a$ and $(ab)^* = b^*a^*$ for all $a, b \in \mathcal{A}$.

We call a^* the adjoint of a. If we replace "unital algebra" by "$*$-algebra" in Definition 1.3.2 we get the definition of a $*$-probability space.

There are additional properties that the functional φ in a non-commutative probability space $(\mathcal{A}, \varphi)$ may satisfy. If $\varphi(ab) = \varphi(ba)$ for every $a, b \in \mathcal{A}$, we say that φ is tracial. If $(\mathcal{A}, \varphi)$ is a $*$-probability space and $\varphi(a^*a) \geq 0$ for every $a \in \mathcal{A}$, we say that φ is positive. If $\varphi(a^*a) = 0$ only when $a = 0$, then we say that φ is faithful.

The notion of a non-commutative probability space generalizes the idea of some spaces of random variables, including spaces of random matrices.

We call a^* the adjoint of a. If we replace 'unital algebra' with '$*$-algebra' in Definition 1.3.2, we obtain the definition of a $*$-probability space.

There are additional properties that the functional φ in a non-commutative probability space $(\mathcal{A}, \varphi)$ may satisfy. If $\varphi(ab) = \varphi(ba)$ for every $a, b \in \mathcal{A}$, we say that φ is tracial. If $(\mathcal{A}, \varphi)$ is a $*$-probability space and $\varphi(a^*a) \geq 0$ for every $a \in \mathcal{A}$, we say that φ is positive. If $\varphi(a^*a) = 0$ only when $a = 0$, then we say that φ is faithful.

We can distinguish different types of non-commutative random variables based on their algebraic properties; we say that $a \in \mathcal{A}$ is self-adjoint if $a^* = a$, unitary if $a^*a = 1_{\mathcal{A}} = aa^*$, and normal if it commutes with its adjoint, $a^*a = aa^*$.

The notion of a non-commutative probability space generalizes the concept of certain spaces of random variables, including spaces of random matrices.

Example 1.3.1. Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a classical probability space and $\mathcal{A} = L^\infty(\Omega, \mathbb{P}, \mathbb{R})$ the set of bounded measurable functions from Ω taking values in $\mathbb{R}$. Equip $\mathcal{A}$ with the linear operator φ given by

$$\varphi(a) = \int_\Omega a(\omega) \mathrm{d}P(\omega) = \mathbb{E}[a],$$

for $a \in \mathcal{A}$. Then $(\mathcal{A}, \varphi)$ is a non-commutative probability space.

If we take the space $\mathcal{A}' = L^{\infty}(\Omega, \mathbb{P}, \mathbb{C})$ of bounded measurable functions taking complex values, then we can have a $*$-operation given by the complex conjugate and $(\mathcal{A}', \varphi)$ is a $*$-probability space.

The hypothesis of the variables being bounded is for them to form an algebra. If we can not guarantee a^n is in $\mathcal{A}$ for every n, then $\mathcal{A}$ is not an algebra. This condition causes our non-commutative probability space not to include all of the classical random variables we can define in a classical probability space. However, we can relax hypotheses to enlarge the class of classical random variables that form an algebra.

Example 1.3.2. Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a classical probability space and

$$\mathcal{A} = L^{\infty-}(\Omega, \mathbb{P}, \mathbb{F}) \coloneqq \bigcap_{1 \leq p < \infty} L^p(\Omega, \mathbb{P}, \mathbb{F}),$$

with $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$.

Clearly $L^{\infty}(\Omega, \mathbb{P}, \mathbb{F}) \subset L^{\infty-}(\Omega, \mathbb{P}, \mathbb{F})$, and since we are asking every variable in $L^{\infty-}(\Omega, \mathbb{P}, \mathbb{F})$ to have all the positive moments, we can assure it is an algebra. With this, we have that the space of the classical random variables with finite moments of every order is a non-commutative probability space.

A definition of a non-commutative probability space that includes random variables without finite moments of every order can be provided, but it is not needed for this work.

Example 1.3.3. Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and $\mathcal{M}_{n,n}(\mathbb{F})$ be the algebra of $n \times n$ matrices with entries in $\mathbb{F}$, where the operations are $+$ the sum of matrices and $\cdot$ the usual matrix product. Denote by tr the normalized trace, i.e.

$$\operatorname{tr}(A) \coloneqq \frac{1}{n} Tr(A) = \frac{1}{n} \sum_{j=1}^{n} A_{jj},$$

for $A \in \mathcal{M}_{n,n}(\mathbb{F})$.

Define a $*$-operation by

$$(A^*)_{ij} \coloneqq \overline{A_{ji}}.$$

Then $(\mathcal{M}_{n,n}(\mathbb{F}), \operatorname{tr})$ is a $*$-probability space. In the case where $\mathbb{F} = \mathbb{R}$, the $*$-operation is simply taking the transpose.

The entries of a matrix need not be random for the space of matrices to be a non-commutative probability space, but we can indeed ask every entry to be a random variable and thus we recover the definition of a space of random variables is indeed a non-commutative probability space.

Example 1.3.4. Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and $\mathcal{A} = \mathcal{M}_{n,n}(L^{\infty-}(\Omega, \mathbb{P}, \mathbb{F}))$ the space of $n \times n$ matrices with entries in $L^{\infty-}(\Omega, \mathbb{P}, \mathbb{F})$. Equip $\mathcal{A}$ with the functional $\varphi(\cdot)$ given by

$$\varphi(A) = \mathbb{E}\left[\operatorname{tr}(A)\right] = \frac{1}{n} \sum_{j=1}^{n} \int_{\Omega} A_{jj}(\omega) \mathrm{d}\mathbb{P}(\omega).$$

Then $(\mathcal{A}, \varphi)$ is a $*$-probability space with the $*$-operation being the conjugate transpose or simply the transpose if $\mathbb{F} = \mathbb{R}$.

One of the central concepts in classical probability theory is the distribution of a random variable. In a non-commutative probability space $(\mathcal{A}, \varphi)$ we can define the distribution of a random variable a as how φ acts on powers of a. In analogy to classical probability, we give a definition of the moments of a non-commutative random variable.

Definition 1.3.4. Let a be a random variable in a $*$-probability space $(\mathcal{A}, \varphi)$. A $*$-moment of a is any expression of the form

$$\varphi(a^{e_0} \cdots a^{e_k}),$$

with $k \in \mathbb{N}$ and $e_0, \ldots, e_k \in \{1, *\}$.

Denote by $\mathbb{C}\langle X, X^* \rangle$ the unital algebra freely generated by the indeterminates X and X^*, i.e. $\mathbb{C}\langle X, X^* \rangle$ is the algebra over $\mathbb{C}$ generated by all of the monomials of the form

$$x^{e_0} \cdots x^{e_k},$$

with k and $e_1, \ldots, e_k$ as in Definition 1.3.4. Now we present the definition of a $*$-distribution.

Definition 1.3.5. Let a be a random variable in a $*$-probability space $(\mathcal{A}, \varphi)$. A $*$-distribution is a linear functional

$$\mu : \mathbb{C}\langle X, X^* \rangle \to \mathbb{C},$$

that satisfies

$$\mu(x^{e_0} \cdots x^{e_k}) = \varphi(a^{e_0} \cdots a^{e_k}).$$

with $k \in \mathbb{N}$ and $e_0, \ldots, e_k \in \{1, *\}$.

In the particular case in which a is normal, the distribution is determined by the moments of the form $\varphi(a^k (a^*)^l)$ for $k, l \geq 0$, and if a is self-adjoint it is enough to consider the moments $\varphi(a^k)$ for $k \geq 0$.

For the normal case, it is possible to link the concept of distribution in classical probability with the definition in the non-commutative case.

Definition 1.3.6. Let a be a random variable in a non-commutative probability space $(\mathcal{A}, \varphi)$. If there exists a compactly supported probability measure μ on $\mathbb{C}$ such that for every $k, l \in \mathbb{N}$,

$$\int_{\mathbb{C}} z^k \bar{z}^l \mathrm{d}\mu(z) = \varphi(a^k (a^*)^l),$$

then the distribution of a is uniquely determined by μ and we call μ the analytic distribution of a.

In the case $\mathcal{A} = L^\infty(\Omega, \mathbb{P}, \mathbb{F})$ and $\varphi(\cdot) = \mathbb{E}[\cdot]$, the analytical distribution of $a \in \mathcal{A}$ is the measure on the Borel sets of $\mathbb{F}$ induced by a, i.e.

$$\mu(B) = P\left(\{\omega \in \Omega : a(\omega) \in B\}\right),$$

for B a Borel set of $\mathbb{F}$.

1.3.2 Tensor and Free Independence

The definition of independence in classical probability assumes that the product of random variables is commutative. This property does not hold in general for non-commutative probability spaces and particularly in the case of random matrix spaces. Instead, we have four different notions of independence defined in terms of the moments. In this thesis, we will only consider the notions of tensor and free independence.

Formally, a notion of independence is a rule that allows to compute mixed moments of the form

$$\varphi(a^{m_1}b^{n_1}\cdots a^{m_k}b^{n_k}), \qquad k\in\mathbb{N}, \quad m_i,n_j\in\{1,*\},$$

in terms of the individual moments of a and b.

The classical notion of independence corresponds to tensor independence in non-commutative probability.

Definition 1.3.7 (Tensor independence)**.** Let $(\mathcal{A},\varphi)$ be a non-commutative probability space and I a set of indexes. A set of unital subalgebras $(\mathcal{A}_i)_{i\in I}$ is called tensor independent if the following two conditions are satisfied

- For every $a\in\mathcal{A}_i$ and $b\in\mathcal{A}_j$, a and b commute.
- For all the finite subsets $J\subset I$ and all $a_j\in\mathcal{A}_j$ we can compute $\varphi\left(\prod_{j\in J}a_j\right)$ as

$$\varphi\left(\prod_{j\in J}a_j\right)=\prod_{j\in J}\varphi\left(a_j\right)$$

When we say that two random variables a,b are tensor independent, we mean that the unitary subalgebras $\mathcal{A}_1,\mathcal{A}_2$ generated by a and b, respectively are tensor independent.

Tensor independence is a symmetric relationship. If a is tensor independent of b, then b is tensor independent of a, this is not the case for every notion of independence. We can distinguish three additional notions of independence, we introduce the notion of "Free independence", the definition of the other notions can be found in [14].

Definition 1.3.8 (Free independence)**.** Let $(\mathcal{A},\varphi)$ be a non-commutative probability space and I a set of indexes. A set of unital subalgebras $(\mathcal{A}_i)_{i\in I}$ is called tensor independent if $\varphi(a_1\cdots a_k)=0$ whenever we have that

- $k\in\mathbb{Z}^+$,
- $a_j\in\mathcal{A}_{i(j)}$ with $i(j)\in I$ for every $j\in[k]$,
- $\varphi(a_j)=0$ for every $j\in[k]$,
- Consecutive elements in $a_1\cdots a_k$ come from different algebras, i.e.

$$i(1)\neq i(2), i(2)\neq i(3),\ldots,i(k-1)\neq i(k).$$

When we say that two random variables a,b are freely independent, we mean that the unitary subalgebras $\mathcal{A}_1,\mathcal{A}_2$ generated by a and b respectively, are freely independent.

Convolution

In classical probability, when we need to compute the distribution of a sum of (tensor) independent random variables, we find the convolution of the associated measures. In non-commutative probability spaces, adding two independent random variables in any notion of independence gives place to a different kind of convolution.

Definition 1.3.9 (Non-commutative convolution)**.** Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space and $a, b \in \mathcal{A}$ be self-adjoint and independent (in some notion), then the convolution (in some notion) of a and b is the algebraic distribution of $a + b$, i.e. the linear operator characterizing the moments

$$\varphi\left((a+b)^m\right),$$

with m a sequence of the form $m = e_1 e_2 \cdots e_k$, $k \in \mathbb{N}$ and $e_j \in \{1, *\}$, for all $j \in [k]$.

If a is a classical random variable and its moment generating function $\mathbb{E}\left[e^{ta}\right]$ exists, we can define $K_a(t)$ the cumulant generating function of a as

$$K_a(t) \coloneqq \log \mathbb{E}\left[e^{ta}\right].$$

This function is analytic (at least in a neighborhood of zero) and has a Taylor expansion given by

$$K_a(t) = \sum_{n=1}^{\infty} \frac{t^n}{n!} \kappa_n(a),$$

with $\kappa_n(a)$ being the nth cumulant of a. This function linearizes the (tensor) convolution in the sense that if a, b are (tensor) independent, then

$$K_{a+b}(t) = K_a(t) + K_b(t).$$

This ultimately implies that the cumulants $\kappa_a(n)$ linearize the (tensor) convolution if a, b are (tensor) independent, i.e.

$$\kappa_{a+b}(n) = \kappa_a(n) + \kappa_b(n).$$

Analogous coefficients can be defined for the other notions of independence. The main tool to study them is the Cauchy-Stieltjes transform, sometimes called only "Cauchy transform".

Definition 1.3.10 (Cauchy-Stieltjes transform)**.** Let μ be a probability measure on $\mathbb{R}$, its Cauchy-Stieltjes transform, $G_\mu(t)$ is

$$G_\mu(z) \coloneqq \int_{\mathbb{R}} \frac{1}{z-t} \mu(\mathrm{d}t),$$

for $z \in \mathbb{C}^+ \coloneqq \{z \in \mathbb{C} : Im(z) > 0\}$. The Cauchy transform takes values in $\mathbb{C}^- \coloneqq \{z \in \mathbb{C} : Im(z) > 0\}$.

When μ has bounded support, there is a relationship between their moments and the Cauchy-Stieltjes transform μ given by

$$G_\mu(z) = z^{-1}\int_{\mathbb{R}} \frac{1}{1-\frac{t}{z}}\mu(\mathrm{d}t) = z^{-1}\int_{\mathbb{R}}\sum_{k=0}^{\infty}\left(\frac{t}{z}\right)^k \mu(\mathrm{d}t),$$
$$= \sum_{k=0}^{\infty} z^{-(k+1)}\int_{\mathbb{R}} t^k\mu(\mathrm{d}t) = \sum_{k=0}^{\infty}\frac{m_k(\mu)}{z^{k+1}},$$

with $m_k(\mu)$ being the kth moment of μ and for $|z| \geq \sup\{t : t \in \mathrm{supp}(\mu)\}$.

The collection of functions $f_z(t) = 1/(z-t)$ parametrized by $z \in \mathbb{C}^+$ forms a separating test family for the space of probability measures. This implies that the Cauchy-Stieltjes transform uniquely characterizes the probability measure. Due to the relevance of this proposition, we state it in the next theorem.

Theorem 1.3.1. *Given μ and ν two probability measures on $\mathbb{C}$ with Cauchy-Stieltjes transforms G_μ and G_ν respectively. Then*

$$\mu = \nu,$$

if, and only if

$$G_\mu = G_\nu.$$

The following are Cauchy-Stieltjes transforms for a few well-known distributions

Example 1.3.5.

- A random variable has standard semicircle distribution if it has density f such that

$$f(x) = \frac{\sqrt{4-x^2}}{2\pi}, \qquad x \in [-2,2].$$

In this case, its Cauchy-Stieltjes transform is

$$G_s(z) = \frac{z-\sqrt{z^2-4}}{2}, \quad \forall z \in \mathbb{C}^+.$$

- A random variable has arcsine distribution if it has density g such that

$$g(x) = \frac{1}{\pi}\frac{1}{\sqrt{4-x^2}}, \qquad x \in [-2,2].$$

In this case, its Cauchy-Stieltjes transform is given by

$$G_a(z) = \frac{1}{\sqrt{z^2-4}}.$$

- Let μ be an empirical probability measure over a finite set of points $\lambda_1, \lambda_2, \ldots, \lambda_n$. Then

$$\mu(\{x\}) = \frac{1}{n}\sum_{k=1}^{n}\delta_{\lambda_k}(x),$$

and its Cauchy-Stieltjes transform is given by

$$G_\mu(z) = \frac{1}{n}\sum_{k=0}^{n}\frac{1}{z-\lambda_k}.$$

If μ is absolutely continuous with respect to the Lebesgue measure and has density f_μ, then it is possible to recover f_μ from the Cauchy transform G_μ by using the Stieltjes inversion formula.

Theorem 1.3.2. *Let μ be an absolutely continuous probability measure in $\mathbb{R}$ with density h_μ and Cauchy transform G_μ. Then*

$$f_\mu(x) = \lim_{y\to 0} -\frac{1}{\pi} Im[G_\mu(x+iy)].$$

Using the Cauchy-Stieltjes transform it is possible to define other transforms that linearize convolution in different notions of independence. We include here the ones that are useful for this work

Definition 1.3.11 (Non-commutative linearizing transforms)**.** Let μ be a probability measure on $\mathbb{R}$ with Cauchy-Stieltjes transform G_μ. We can define the following transforms for μ.

- The $\mathcal{K}$ transform $\mathcal{K}_\mu$ is the compositional inverse of G_μ,

$$\mathcal{K}_\mu(z) \coloneqq G_\mu^{-1}(z).$$

- The $\mathcal{R}$ transform $\mathcal{R}_\mu$ is defined in terms of the $\mathcal{K}$ transform

$$\mathcal{R}_\mu(z) \coloneqq \mathcal{K}_\mu(z) - \frac{1}{z}.$$

Let a and b be non-commutative random variables with real support and analytic distributions μ and ν, respectively. The distribution of the sum $a+b$ is the convolution in a notion of independence if a and b satisfy that notion of independence.

- If a and b are tensor independent, the distribution of $a+b$ is the classical (or tensor) convolution of μ and ν denoted by $\mu * \nu$.
- If a and b are freely independent, the distribution of $a+b$ is the free convolution of μ and ν denoted by $\mu \boxplus \nu$.

We say that μ and ν satisfy an independence relationship if the associated random variables satisfy it.

The relationship between the transforms defined above and the convolution of non-commutative random variables is stated in the following theorem.

Theorem 1.3.3. *Let μ, ν be two compactly supported probability measures on $\mathbb{R}$. If they are freely independent, then*

$$\mathcal{R}_{\mu\boxplus\nu}(z) = \mathcal{R}_\mu(z) + \mathcal{R}_\nu(z).$$

In classical probability, the sum of independent random variables under mild assumptions converges to a normal random variable. This result is known as the Central Limit Theorem and there is an analogous result for the case of free independence. In the next section, we state this result for the classical and free independence cases.

1.3.3 Classical and free central limit theorems

Each notion of independence and convolution gives in turn a new limit for sums of the form

$$S_n = \frac{X_1 + X_2 + \cdots + X_n}{\sqrt{n}},$$

for X_i independent (in some sense) and identically distributed random variables with mean 0 and variance 1.

In classical probability, the variables are tensor independent and the limit is the Gaussian distribution. The next theorem gives us a generalization in the case of free independence. A proof can be found in [13].

Theorem 1.3.4. *Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space and $(a_i)_{i\in\mathbb{N}}$ be a sequence of random variables in $\mathcal{A}$ with common distribution and that have zero mean ($\varphi(a_1) = 0$) and variance 1 ($\varphi(a_1^2) = 1$).*

Denote by $\xrightarrow{D}$ the convergence in distribution and define S_n as

$$S_n := \frac{\sum_{j=1}^{n} a_j}{\sqrt{n}}.$$

- *If the random variables are tensor-independent, then*

$$S_n \xrightarrow[n\to\infty]{D} N,$$

with N a standard normal random variable.

- *If the random variables are freely independent, then*

$$S_n \xrightarrow[n\to\infty]{D} s,$$

with s a standard semicircular random variable.

Chapter 2

Eigenvalue processes for matrix-valued processes

In this chapter, we prove the form of the dynamic equation for the Dyson Brownian motion in the real and complex cases. We then show an extension of the result to matrix-valued diffusion processes. Most of the material in the chapter is taken from [2] which in turn uses techniques appearing in [12] for the study of Wishart processes. In the first section, the essential tools are introduced and then the result is proven for the Dyson Brownian motion. In the second section, the results are generalized for matrix-valued Itô processes, and the particular cases of the Wishart and Jacobi processes are given.

At the end of the chapter, we will prove that if $B = (B(t), t \geq 0)$ is a Brownian motion in $\mathcal{M}_{n,n}(\mathbb{R})$ and X is a matrix-valued process satisfying the following system of stochastic differential equations

$$\mathrm{d}X(t) = g(X(t))\mathrm{d}B(t)h(X(t)) + h(X(t))\mathrm{d}B(t)^T g(X(t)) + b(X(t))\mathrm{d}t, \tag{2.1}$$

then its eigenvalues are semimartingales that satisfy the system of stochastic differential equations given by

$$\mathrm{d}\lambda_i = 2g(\lambda_i)h(\lambda_i)\mathrm{d}W_i + \left(b(\lambda_i) + \sum_{k\neq i} \frac{g^2(\lambda_i)h^2(\lambda_k) + g^2(\lambda_k)h^2(\lambda_i)}{\lambda_i - \lambda_k}\right)\mathrm{d}t, \tag{2.2}$$

until the time of the first collision, which means whenever $\lambda_i = \lambda_j$ for any $i \neq j$. We will show, however, that the first collision is impossible in finite time a. s. The analogous result holds when B is a Brownian motion in $\mathcal{M}_{n,n}(\mathbb{C})$ and we take $B(t)^*$ instead of $B(t)^T$ in (2.1).

To illustrate the techniques and introduce the most basic process of this kind, we will first prove the particular cases of the real and complex Dyson Brownian motion, which are simply the eigenvalue processes of Brownian motion in $\mathcal{H}_{n,n}(\mathbb{R})$ and $\mathcal{H}_{n,n}(\mathbb{C})$, respectively. In both cases, the processes are particular cases of (2.2) with $b \equiv 0$ and $g \equiv h \equiv 1/\sqrt[4]{2}$. Notice that this implies that X is a process in $\mathcal{H}_{n,n}(\mathbb{R})$ ($\mathcal{H}_{n,n}(\mathbb{C})$) whose off-diagonal entries are standard real-valued (complex-valued) Brownian motions and the diagonal entries would are real-valued (complex-valued) Brownian motions with variance 2. In the later case (2.2) would turn into

$$\mathrm{d}\lambda_i = \sqrt{2}\mathrm{d}W_i + \sum_{k\neq i} \frac{\mathrm{d}t}{\lambda_i - \lambda_k},$$

which exactly the system of stochastic differential equations we expect to have for the Dyson Brownian motion.

2.1 Dyson Brownian Motion

This section is mainly used to show the results for the Dyson Brownian motion, but some of the general results will be used also to prove the analogous equations for the spectrum of more general matrix diffusions in later sections. In the first subsection we derive the stochastic differential equations for the eigenvalues of a symmetric real Brownian matrix up to the collison time. In the second subsection we do the same for a self adjoint complex Brownian matrix. In the third section we prove that the collision time is almost surely infinite. No proof of the existence and uniqueness of a solution for the equations is given in this section, but it is later given in a more general case in section 2.2.

The first description of the Dyson Brownian motion was given in [1] as a model for a Couloumb gass executing each one a Brownian motion and the repuslive forces between them. In generality, the Dyson Brownian motion is the process that models the spectrum of a Brownian motion in $\mathcal{H}_{n,n}(\mathbb{F})$ and it is described by the system of stochastic differential equations

$$\mathrm{d}\lambda_i = \sqrt{\frac{2}{\beta}}\mathrm{d}W_i + \sum_{k\neq i}\frac{\mathrm{d}t}{\lambda_i - \lambda_k},$$

where the $W_i, 1 \leq i \leq n$ are independent standard Brownian motions and β parametrizes the field in which the entries of the matrix take values. If $\beta = 1$, then $\mathbb{F} = \mathbb{R}$, if $\beta = 2$, then $\mathbb{F} = \mathbb{C}$ and if $\beta = 4$, $\mathbb{F} = \mathbb{H}$. Other matrix models for different values of β have been proposed in [15] and [4]. In the first one for $\beta \in [0, 2]$ and for $\beta \in (0, \infty)$ in the second one.

In this section, we only prove the result for $\beta \in \{1, 2\}$ and we will notice that the difference is essentially the covariance between entries of the matrix, so the extension to $\beta = 4$ would be quite similar.

2.1.1 Real case

The properties of the one dimensional Brownian motion allow to easily extend the definition to Brownian motions in different spaces, such as $\mathbb{R}^n$ or $\mathcal{M}_{n,n}(\mathscr{F})$. In the particular case that B is a Brownian motion in the space of symmetric matrices with real entries, for every t, the process is a multiple of a GOE, i.e. $B(t) = \sqrt{t}R$ with R a GOE. This means the law of B keeps invariant under orthogonal transformations and this property is essential for the results in this case. For the sake of clearness, we give a precise definition of a Brownian motion in the space of symmetric matrices.

Definition 2.1.1 (Brownian motion in $\mathcal{H}_{n,n}(\mathbb{R})$)**.** Let $B = (B(t), t \geq 0)$ be a stochastic process taking values in $\mathcal{M}_{n,n}(\mathbb{R})$ whose entries are standard Brownian motions $\{B_{ij}(t)\}_{1\leq i\leq n, 1\leq j\leq n}$ such that,

$$\mathrm{d}\langle B_{ij}, B_{kl}\rangle(t) = (\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})\,\mathrm{d}t.$$

Then we say that B is a Brownian motion in $\mathcal{H}_{n,n}(\mathbb{R})$.

A process B as defined above clearly has real eigenvalues, so we can order them. Let $\lambda_1(t) > \lambda_2(t) > \cdots > \lambda_n(t)$ be the eigenvalues. Notice they are also time-dependent functions,

we are interested in knowing if for some t_0 the order of some of them is changed. Due to the continuity of the paths, this happens only if at some point the eigenvalues changing the order are equal. The next stopping time gives us the first time of collision of the eigenvalues.

Definition 2.1.2 (First time of collision)**.** Let $\lambda_1(t) > \lambda_2(t) > \cdots > \lambda_n(t)$ be the ordered eigenvalues of a matrix-valued stochastic process. We define the first collision time τ as

$$\tau \coloneqq \inf\{t : \lambda_i(t) = \lambda_j(t) \text{ for some } i \neq j\}. \tag{2.3}$$

In the following theorem, we derive a stochastic differential equation for the behaviour of a Brownian symmetric matrix's spectrum on $[0, \tau)$.

Theorem 2.1.1. *Let $B = (B(t), t \geq 0)$ be a symmetric $n \times n$ matrix-valued Brownian motion in $\mathcal{M}_{n,n}(\mathbb{R})$ with diagonalization $B = H\Lambda H^T$ and eigenvalues $\lambda_1, \ldots, \lambda_n$. Define τ the first time of collision of the eigenvalues as in* (2.3)*.*

Then, for $t < \tau$ the eigenvalue process $\Lambda = (\Lambda(t), t \geq 0)$ verifies the following system of stochastic differential equations:

$$\mathrm{d}\lambda_i = \sqrt{2}\mathrm{d}W_i + \sum_{k \neq i} \frac{\mathrm{d}t}{\lambda_i - \lambda_k}, \tag{2.4}$$

where $(W_i)_{i \in [n]}$ are independent Brownian motions on $\mathbb{R}$.

Proof. For a fixed $t \geq 0$ the matrix $B(t)$ is equal in law to $\sqrt{t}R$ with R a Gaussian orthogonal ensemble. Using that R is invariant under orthogonal transformations this implies that for every orthogonal matrix O, we have

$$OB(t)O^T \overset{d}{=} O\sqrt{t}RO^T \overset{d}{=} \sqrt{t}ORO^T \overset{d}{=} \sqrt{t}R \overset{d}{=} B(t).$$

In particular, $B(t)$ is also invariant under orthogonal transformations.

We have that $H^{-1} = H^T$, its stochastic logarithm L is defined by a stochastic differential equation as

$$\mathrm{d}L \coloneqq H^{-1}\partial H = H^T \partial H = H^T \mathrm{d}H + \frac{1}{2}(\mathrm{d}H^T)\mathrm{d}H.$$

We use the matrix Itô formula on $I = H^T H$,

$$0 = \mathrm{d}I = \mathrm{d}(H^T H) = H^T \mathrm{d}H + (\mathrm{d}H)^T H + (\mathrm{d}H)^T \mathrm{d}H = H^T \partial H + (\partial H)^T H = \mathrm{d}L + \mathrm{d}L^T.$$

This implies that $\mathrm{d}L^T = -\mathrm{d}L$ and the stochastic logarithm of H is skew-symmetric. Now we use that $\Lambda = H^T B H$ and the matrix Itô formula again to get

$$\begin{aligned}
\mathrm{d}\Lambda = \mathrm{d}(H^T B H) &= (\partial H^T B)H + H^T B \partial H = (\partial H)^T B H + H^T (\partial B) H + H^T B \partial H, \\
&= (\partial H)^T H \Lambda + H^T (\partial B) H + \Lambda H^T \partial H = (\partial L)^T \Lambda + H^T (\partial B) H + \Lambda \partial L, \\
&= H^T (\partial B) H - (\partial L)\Lambda + \Lambda \partial L.
\end{aligned}$$

The diagonals of $(\partial L)\Lambda$ and $\Lambda \partial L$ coincide, so $\mathrm{d}\Lambda_{ii} = \left(H^T(\partial B)H\right)_{ii}$. Let $\mathrm{d}N \coloneqq H^T(\partial B)H$. Now, for $i \neq j$, we use that Λ is diagonal to get

$$0 = \mathrm{d}N_{ij} + (\lambda_i - \lambda_j)\mathrm{d}L_{ij}.$$

We can then conclude that $\mathrm{d}L_{ij} = \mathrm{d}N_{ij}/(\lambda_j - \lambda_i)$ whenever $i \neq j$. Now we need to find a more explicit representation for $\mathrm{d}N$. We see that $\mathrm{d}N$ and $H^T(\mathrm{d}B)H$ differ only in a finite variation part, so the martingale term must coincide and it is characterized by the quadratic covariation that we can find using the covariation of B.

$$\begin{aligned}\mathrm{d}N_{ij}\mathrm{d}N_{kl} &= \mathrm{d}\langle (H^T(\partial B)H)_{ij}, (H^T(\partial B)H)_{kl}\rangle(t) = \mathrm{d}\langle (H^T(\mathrm{d}B)H)_{ij}, (H^T(\mathrm{d}B)H)_{kl}\rangle(t),\\ &= \sum_{p,q,r,s} \mathrm{d}\langle H_{pi}\mathrm{d}B_{pq}H_{qj}, H_{rk}\mathrm{d}B_{rs}H_{sl}\rangle(t) = \sum_{p,q,r,s} H_{pi}H_{qj}H_{rk}H_{sl}\mathrm{d}B_{pq}\mathrm{d}B_{rs},\end{aligned}$$

Recall that $\mathrm{d}\langle B_{ij}, B_{kl}\rangle(t) = (\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})\,\mathrm{d}t$,

$$\begin{aligned}&= \sum_{p,q,r,s} H_{pi}H_{qj}H_{rk}H_{sl}(\delta_{rp}\delta_{sq} + \delta_{rq}\delta_{sp})\mathrm{d}t,\\ &= \left(\sum_p H^T_{ip}\delta_{rp}H_{rk}\right)\left(\sum_q H^T_{jq}\delta_{qs}H_{sl}\right)\mathrm{d}t + \left(\sum_p H^T_{ip}\delta_{sp}H_{sl}\right)\left(\sum_q H^T_{jq}\delta_{rq}H_{rk}\right)\mathrm{d}t\\ &= (\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})\mathrm{d}t.\end{aligned}$$

We had previously got that $\mathrm{d}\Lambda_{ii} = \mathrm{d}N_{ii}$, and thus using the covariation of N we can find the martingale term of λ_i,

$$\mathrm{d}\lambda_i\mathrm{d}\lambda_j = \mathrm{d}\Lambda_{ii}\mathrm{d}\Lambda_{jj} = \mathrm{d}N_{ii}\mathrm{d}N_{jj} = 2\delta_{ij}\mathrm{d}t.$$

Then the martingale term of every eigenvalue is $\sqrt{2}$ times a Brownian motion that is independent of the martingale term of any other eigenvalue. Now we need to find the finite variation part of λ_i. Let us call F the finite variation part of N, then

$$\begin{aligned}\mathrm{d}F &= \frac{1}{2}\left(\mathrm{d}H^T\mathrm{d}BH + H^T\mathrm{d}B\mathrm{d}H\right) = \frac{1}{2}\left((\mathrm{d}H^TH)(H^T\mathrm{d}BH) + (H^T\mathrm{d}Bh)(H^T\mathrm{d}H)\right),\\ &= \frac{1}{2}\left((\mathrm{d}N\mathrm{d}L)^T + \mathrm{d}N\mathrm{d}L\right).\end{aligned}$$

The last equality is because $\mathrm{d}N$ and $H^T(\mathrm{d}B)H$ only differ in a finite variation term. Using the previous results we find

$$(\mathrm{d}N\mathrm{d}L)_{ij} = \sum_k \mathrm{d}N_{ik}\mathrm{d}L_{kj} = \sum_{k\neq j}\frac{\mathrm{d}N_{ik}\mathrm{d}N_{kj}}{\lambda_j - \lambda_k} = \delta_{ij}\sum_{k\neq j}\frac{\mathrm{d}t}{\lambda_j - \lambda_k}.$$

We can conclude that

$$\mathrm{d}F_{ii} = \frac{1}{2}\left((\mathrm{d}N\mathrm{d}L)^T_{ii} + (\mathrm{d}N\mathrm{d}L)_{ii}\right) = (\mathrm{d}N\mathrm{d}L)_{ii} = \sum_{k\neq i}\frac{\mathrm{d}t}{\lambda_i - \lambda_k}.$$

Now we know the martingale and finite variation terms of λ_i and we can write the explicit expression for it.

$$\mathrm{d}\lambda_i = \sqrt{2}\mathrm{d}W_i + \sum_{k\neq i} \frac{\mathrm{d}t}{\lambda_i - \lambda_k},$$

where $W_1, \ldots, W_n$ are independent standard Brownian motions. □

2.1.2 Complex case

Before proceeeding with the result, we define a Brownian motion in $\mathcal{H}_{n,n}(\mathbb{C})$, which is totally analogous to the Brownian motion in $\mathcal{H}_{n,n}(\mathbb{R})$.

Definition 2.1.3 (Brownian motion in $\mathcal{H}_{n,n}(\mathbb{C})$)**.** Let $B = (B(t), t \geq 0)$ be a stochastic process taking values in $\mathcal{M}_{n,n}(\mathbb{C})$ whose off-diagonal entries are complex Brownian motions $\{B_{ij}(t)\}_{1\leq i\leq n, 1\leq j\leq n}$ such that,

$$\mathrm{d}\langle B_{ij}, B_{kl}\rangle(t) = 2\delta_{ik}\delta_{jl}\mathrm{d}t, \tag{2.5}$$

and the diagonal entries are n independient real-valued Brownian motions with variance 2, which means

$$\mathrm{d}\langle B_{ii}, B_{jj}\rangle(t) = 2\delta_{ij}\mathrm{d}t. \tag{2.6}$$

Then we say that B is a Brownian motion in $\mathcal{H}_{n,n}(\mathbb{C})$.

Notice that (2.5) implies that $B_{ij}(t) = \overline{B_{ji}}$ and together with (2.6) this means that B is effectively a process taking values in $\mathcal{H}_{n,n}(\mathbb{C})$.

For the case of a Brownian motion in $\mathcal{H}_{n,n}(\mathbb{C})$ there are basically two equivalent formulations of the Dyson Brownian motion. If we take the eigenvalue process defined as in 2.1.3, then the eigenvalues satisfy the system of SDE

$$\mathrm{d}\lambda_i = \mathrm{d}W_i + 2\sum_{k\neq i} \frac{\mathrm{d}t}{\lambda_i - \lambda_k},$$

if instead we consider the process $B' = \sqrt{\frac{1}{2}}B$, then the eigenvalues obey the system

$$\mathrm{d}\lambda_i = \mathrm{d}W_i + \sum_{k\neq i} \frac{\mathrm{d}t}{\lambda_i - \lambda_k}.$$

We are more interested in the latter re-scaled process since it generalizes the $\mathcal{H}_{n,n}(\mathbb{R})$ case in the following sense. If we take $\beta \in \{1, 2\}$, and study the eigenvalues of the process $\sqrt{\frac{1}{\beta}}B(t)$ with B a Brownian motion in $\mathcal{H}_{n,n}(\mathbb{R})$ (resp. $\mathcal{H}_{n,n}(\mathbb{C})$), then before the time of first collision they satisfy

$$\mathrm{d}\lambda_i = \sqrt{\frac{2}{\beta}}\mathrm{d}W_i + \sum_{k\neq i} \frac{\mathrm{d}t}{\lambda_i - \lambda_k}. \tag{2.7}$$

Although we do not prove it here, it is a well known fact that equation (2.7) holds also in the case $\beta = 4$ which is for a self-adjoint random matrix whose entries are quaternionic Brownian motions [1].

Theorem 2.1.2. *Let $B' = (B(t), t \geq 0)$ be a matrix-valued Brownian motion in $\mathcal{H}_{n,n}(\mathbb{C})$ and $B \coloneqq \frac{1}{2}B'$ have diagonalization $B = H\Lambda H^*$ and eigenvalues $\lambda_1, \ldots, \lambda_n$. Define τ the first time of collision of the eigenvalues as in (2.3).*

Then, for $t < \tau$ the eigenvalue process $\Lambda = (\Lambda(t), t \geq 0)$ verifies the following system of stochastic differential equations:

$$\mathrm{d}\lambda_i = \mathrm{d}W_i + \sum_{k \neq i} \frac{\mathrm{d}t}{\lambda_i - \lambda_k}, \tag{2.8}$$

where $(W_i)_{i \in [n]}$ are independent Brownian motions.

Proof. The proof is the same as the real case, but in this case, the entries are complex Brownian motions outside the diagonal and real Brownian motions in the diagonal. Recalling the covariation for a Brownian motion in $\mathcal{H}_{n,n}(\mathbb{C})$ and the re-scaling, we have

$$\mathrm{d}\langle B_{ij}, B_{kl}\rangle(t) = \frac{1}{2}\mathrm{d}\langle B'_{ij}, B'_{kl}\rangle = \delta_{il}\delta_{jk}\mathrm{d}t.$$

Similarly to the real case, for a fixed $t \geq 0$ the matrix $B(t)$ is equal in law to $\sqrt{t}R$ with R a Gaussian unitary ensemble. Using that R is invariant under unitary transformations this implies that for every unitary matrix U, we have

$$UB(t)U^* \overset{d}{=} U\sqrt{t}RU^* \overset{d}{=} \sqrt{t}URU^* \overset{d}{=} \sqrt{t}R \overset{d}{=} B(t).$$

We have that $H^{-1} = H^*$, so repeating the procedure in the real case we define equally the stochastic logarithm as $\mathrm{d}L \coloneqq H^*\partial H$ and use Itô's formula for $I = H^*H$,

$$0 = \mathrm{d}I = \mathrm{d}(H^*H) = H^*\mathrm{d}H + (\mathrm{d}H)^*H + (\mathrm{d}H)^*\mathrm{d}H = H^*\partial H + (\partial H)^*H = \mathrm{d}L + \mathrm{d}L^*.$$

We have that $\mathrm{d}L^* = -\mathrm{d}L$ and L is skew-hermitian. Now we use Itô formula for $\Lambda = H^*BH$,

$$\begin{aligned}\mathrm{d}\Lambda = \mathrm{d}(H^*BH) &= (\partial H^*B)H + H^*B\partial H = (\partial H)^*BH + H^*(\partial B)H + H^*B\partial H,\\ &= (\partial H)^*H\Lambda + H^*(\partial B)H + \Lambda H^*\partial H = (\mathrm{d}L)^*\Lambda + H^*(\partial B)H + \Lambda\mathrm{d}L,\\ &= H^*(\partial B)H - \mathrm{d}L\Lambda + \Lambda\mathrm{d}L.\end{aligned}$$

The processes $\mathrm{d}L\Lambda$ and $\Lambda\mathrm{d}L$ have the same diagonal entries, so the diagonal of Λ coincides with the one of $H^*(\partial B)H$. Define $\mathrm{d}N \coloneqq H^*(\partial B)H$. Outside the diagonal, Λ has zero entries, so we can equate this to the corresponding entries of $\mathrm{d}N$ and $\mathrm{d}L\Lambda, \Lambda\mathrm{d}L$ to get for every $i \neq j$,

$$0 = \mathrm{d}N_{ij} + (\lambda_i - \lambda_j)\mathrm{d}L_{ij},$$

which in turn implies

$$\mathrm{d}L_{ij} = \frac{\mathrm{d}N_{ij}}{\lambda_j - \lambda_i}, \qquad i \neq j.$$

Now, repeating the real case, we find the quadratic covariation of $\mathrm{d}N$ using that $\mathrm{d}N$ and $H^*\mathrm{d}BH$ coincide up to a finite variation term, but using $\mathrm{d}\langle B_{ij}, \overline{B_{kl}}\rangle(t) = 2\delta_{il}\delta_{jk}\mathrm{d}t$.

$$\begin{aligned}
\mathrm{d}N_{ij}\mathrm{d}N_{kl} &= \mathrm{d}\langle (H^*(\partial B)H)_{ij}, (H^*(\partial B)H)_{kl}\rangle(t) = \mathrm{d}\langle (H^*(\mathrm{d}B)H)_{ij}, (H^*(\mathrm{d}B)H)_{kl}\rangle(t), \\
&= \sum_{p,q,r,s} \mathrm{d}\langle H^*_{ip}\mathrm{d}B_{pq}H_{qj}, H^*_{kr}\mathrm{d}B_{rs}H_{sl}\rangle(t) = \sum_{p,q,r,s} H^*_{ip}H_{qj}H^*_{kr}H_{sl}\mathrm{d}B_{pq}\mathrm{d}B_{rs}, \\
&= \sum_{p,q,r,s} H^*_{ip}H_{qj}H^*_{kr}H_{sl}\delta_{rq}\delta_{sp}\mathrm{d}t = \left(\sum_p H^*_{ip}\delta_{sp}H_{sl}\right)\left(\sum_q H^*_{kr}\delta_{rq}H_{qj}\right)\mathrm{d}t, \\
&= \delta_{il}\delta_{kj}\mathrm{d}t.
\end{aligned}$$

We can use this and the previous result that $\mathrm{d}\Lambda$ and $\mathrm{d}N$ coincide in the diagonal to find the covariation of the eigenvalues and this way we find their martingale term

$$\mathrm{d}\lambda_i\mathrm{d}\lambda_j = \mathrm{d}\Lambda_{ii}\mathrm{d}\Lambda_{jj} = \mathrm{d}N_{ii}\mathrm{d}N_{jj} = \delta_{ij}\mathrm{d}t.$$

Just as in the real case, the martingale term of every eigenvalue is a (real) Brownian motion independent of any other eigenvalue. Again, call F the finite variation part of N and use that $\mathrm{d}N$ and $H^T(\mathrm{d}B)H$ only differ in a finite variation term to find,

$$\begin{aligned}
\mathrm{d}F &= \frac{1}{2}\left(\mathrm{d}H^*\mathrm{d}BH + H^*\mathrm{d}B\mathrm{d}H\right) = \frac{1}{2}\left((\mathrm{d}H^*H)(H^*\mathrm{d}BH) + (H^*\mathrm{d}BH)(H^*\mathrm{d}H)\right), \\
&= \frac{1}{2}\left((\mathrm{d}N\mathrm{d}L)^* + \mathrm{d}N\mathrm{d}L\right).
\end{aligned}$$

We have an expression for L in terms of N. We recall it to find the covariation $\mathrm{d}\langle N, L\rangle(t)$,

$$(\mathrm{d}N\mathrm{d}L)_{ij} = \sum_k \mathrm{d}N_{ik}\mathrm{d}L_{kj} = \sum_{k\neq j}\frac{\mathrm{d}N_{ik}\mathrm{d}N_{kj}}{\lambda_j - \lambda_k} = \delta_{ij}\sum_{k\neq j}\frac{\mathrm{d}t}{\lambda_j - \lambda_k}.$$

Thus F is a diagonal matrix, and the ith diagonal term is given by

$$\mathrm{d}F_{ii} = \frac{1}{2}\left((\mathrm{d}N\mathrm{d}L)^*_{ii} + (\mathrm{d}N\mathrm{d}L)_{ii}\right) = (\mathrm{d}N\mathrm{d}L)_{ii} = \sum_{k\neq i}\frac{\mathrm{d}t}{\lambda_i - \lambda_k}.$$

With the martingale and finite variation terms of λ_i, we conclude the stated result.

$$\mathrm{d}\lambda_i = \mathrm{d}W_i + \sum_{k\neq i}\frac{\mathrm{d}t}{\lambda_i - \lambda_k},$$

where $W_1, \ldots, W_n$ are independent standard Brownian motions. □

2.2 Generalization for matrix-valued diffusion processes

Theorem 2.2.1. *Let $B = (B(t), t \geq 0)$ be a Brownian motion in $\mathcal{M}_{n,n}(\mathbb{R})$ and $X(t)$ be a symmetric $n \times n$ matrix-valued stochastic process satisfying the stochastic differential equation*

$$\mathrm{d}X(t) = g(X(t))\mathrm{d}B(t)h(X(t)) + h(X(t))\mathrm{d}B(t)^T g(X(t)) + b(X(t))\mathrm{d}t, \tag{2.9}$$

where g, h, b are real functions acting spectrally, and $X(0)$ is a symmetric $n \times n$ matrix with n different eigenvalues.

Let $G(x, y) = g^2(x)h^2(y) + g^2(y)h^2(x)$, and

$$\tau = \inf\{t : \lambda_i(t) = \lambda_j(t) \text{ for some } i \neq j\}. \tag{2.10}$$

Then, for $t < \tau$ the eigenvalue process $\Lambda(t)$ verifies the following system of stochastic differential equations:

$$\mathrm{d}\lambda_i = 2g(\lambda_i)h(\lambda_i)\mathrm{d}W_i + \left(b(\lambda_i) + \sum_{k\neq i} \frac{G(\lambda_i, \lambda_k)}{\lambda_i - \lambda_k}\right)\mathrm{d}t, \tag{2.11}$$

where $(W_i)_i$ are independent Brownian motions.

Proof. Recall that for every t, the process $X(t)$ admits a decomposition of the form

$$X(t) = H\Lambda H^T,$$

where both Λ and H are matrix-valued stochastic processes, $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ is the diagonal matrix of ordered eigenvalues of $X(t)$ and H is the corresponding matrix of eigenvectors.

Let us define the stochastic logarithm of H as

$$\mathrm{d}A \coloneqq H^{-1}\partial H = H^T\partial H = H^T\mathrm{d}H + \frac{1}{2}(\mathrm{d}H^T)\mathrm{d}H.$$

By using Itô's formula on $I = H^T H$ we find

$$0 = \mathrm{d}I = \mathrm{d}(H^T H) = H^T\mathrm{d}H + (\mathrm{d}H)^T H + (\mathrm{d}H)^T\mathrm{d}H = H^T\partial H + (\partial H)^T H = A + A^T.$$

Which means A is skew-symmetric. Using that $H^T H = I$, we have $\Lambda = H^T H\Lambda H^T H = H^T X H$, by the matrix Itô formula, we find

$$\begin{aligned}
\mathrm{d}\Lambda = \mathrm{d}(H^T X H) &= (\partial H^T X)H + H^T X\partial H,\\
&= (\partial H)^T X H + H^T(\partial X)H + H^T X\partial H,\\
&= (\partial H)^T H\Lambda + H^T(\partial X)H + \Lambda H^T\partial H,\\
&= (\partial A)^T\Lambda + H^T(\partial X)H + \Lambda\partial A,\\
&= H^T(\partial X)H - (\partial A)\Lambda + \Lambda\partial A.
\end{aligned}$$

The entries in the diagonals of $(\partial A)\Lambda$ and $\Lambda\partial A$ coincide, and thus the diagonal of $\Lambda\partial A - (\partial A)\Lambda$ is zero. Let us denote $\mathrm{d}N = H^T(\partial X)H$, then

$$\mathrm{d}\lambda_j = \mathrm{d}N_{jj},$$

and, using that Λ is a diagonal matrix, if $i \neq j$,

$$0 = \mathrm{d}N_{i,j} + (\lambda_i - \lambda_j)\mathrm{d}A_{ij}.$$

This leads to the following representation for A_{ij},

$$\mathrm{d}A_{i,j} = \frac{\mathrm{d}N_{i,j}}{\lambda_j - \lambda_i}, \qquad i \neq j. \tag{2.12}$$

From (4.2) we compute the quadratic covariation $\mathrm{d}X_{ij}\mathrm{d}X_{km}$,

$$\begin{aligned}
\mathrm{d}X_{ij}\mathrm{d}X_{km} = \mathrm{d}\big\langle & (g(X(t))\mathrm{d}B(t)h(X(t)))_{ij} + (h(X_t)\mathrm{d}B^T(t)(g(X(t))))_{ij}, \\
& (g(X(t))\mathrm{d}B(t)h(X(t)))_{km} + (h(X_t)\mathrm{d}B^T(t)(g(X(t))))_{km}\big\rangle, \\
= \mathrm{d}\big\langle & (g(X(t))\mathrm{d}B(t)h(X(t)))_{ij}, (g(X(t))\mathrm{d}B(t)h(X(t)))_{km}\big\rangle \\
+ \mathrm{d}\big\langle & (g(X(t))\mathrm{d}B(t)h(X(t)))_{ij}, (h(X_t)\mathrm{d}B^T(t)(g(X(t))))_{km}\big\rangle \\
+ \mathrm{d}\big\langle & (h(X_t)\mathrm{d}B^T(t)(g(X(t))))_{ij}, (h(X_t)\mathrm{d}B^T(t)(g(X(t))))_{km}\big\rangle \\
+ \mathrm{d}\big\langle & (h(X_t)\mathrm{d}B^T(t)(g(X(t))))_{ij}, (g(X(t))\mathrm{d}B(t)h(X(t)))_{km}\big\rangle
\end{aligned}$$

Let us first find $\mathrm{d}\big\langle (g(X(t))\mathrm{d}B(t)h(X(t)))_{ij}, (g(X(t))\mathrm{d}B(t)h(X(t)))_{km}\big\rangle$, the other summands are analogous,

$$\begin{aligned}
& \mathrm{d}\big\langle (g(X(t))\mathrm{d}B(t)h(X(t)))_{ij}, (g(X(t))\mathrm{d}B(t)h(X(t)))_{km}\big\rangle, \\
& \qquad = \mathrm{d}\Big\langle \sum_{p,q} g(X(t))_{ip}\mathrm{d}B(t)_{pq}h(X(t))_{qj}, \sum_{r,s} g(X(t))_{kr}\mathrm{d}B(t)_{rs}h(X(t))_{sm} \Big\rangle
\end{aligned}$$

using the independence between the entries in the brownian matrix,

$$\begin{aligned}
& = \sum_{p,q} \mathrm{d}\big\langle g(X(t))_{ip}\mathrm{d}B(t)_{pq}h(X(t))_{qj}, g(X(t))_{kp}\mathrm{d}B(t)_{pq}h(X(t))_{qm}\big\rangle \\
& = \sum_{pq} g(X(t))_{ip}h(X(t))_{qj}, g(X(t))_{kp}h(X(t))_{qm}\mathrm{d}t, \\
& = \Big(\sum_{p} g(X(t))_{ip}g(X(t))_{kp}\Big)\Big(\sum_{q} h(X(t))_{qj}h(X(t))_{qm}\Big)\mathrm{d}t, \\
& = \big(g(X(t))g(X(t))^T\big)_{ik}\big(h(X(t))^T h(X(t))\big)_{jm}\mathrm{d}t, \\
& = \big(Hg(\Lambda)H^T Hg(\Lambda)H^T\big)_{ik}\left(Hh(\Lambda)H^T Hh(\Lambda)H^T\right)_{jm}\mathrm{d}t, \\
& = \big(Hg^2(\Lambda)H^T\big)_{ik}\left(Hh^2(\Lambda)H^T\right)_{jm}\mathrm{d}t = g^2(X)_{ik}h^2(X)_{jm}\mathrm{d}t.
\end{aligned}$$

Proceeding similarly with the other four summands we find that

$$\mathrm{d}X_{ij}\mathrm{d}X_{km} = (g^2(X)_{ik}h^2(X)_{jm} + g^2(X)_{im}h^2(X)_{jk} + g^2(X)_{jk}h^2(X)_{im} + g^2(X)_{jm}h^2(X)_{ik})\mathrm{d}t.$$

Since $\mathrm{d}N = H^T(\partial X)H$ only differs in a finite variation part of $H^T(\mathrm{d}X)H$, the martingale part of both processes coincide and then the quadratic covariation of the entries of N is

$$\begin{aligned}
\mathrm{d}N_{ij}\mathrm{d}N_{km} &= \mathrm{d}\big\langle (H^T\mathrm{d}XH)_{ij}, (H^T\mathrm{d}XH)_{km}\big\rangle = \sum_{pqrs}\mathrm{d}\big\langle H^T_{ip}\mathrm{d}X_{pq}H_{qj}, H^T_{kr}\mathrm{d}X_{rs}H_{sm}\big\rangle,\\
&= \sum_{pqrs} H^T_{ip}H_{qj}H^T_{kr}H_{sm}\mathrm{d}X_{pq}\mathrm{d}X_{rs},\\
&= \sum_{pqrs} H^T_{ip}H_{qj}H^T_{kr}H_{sm}\big(g^2(X)_{pr}h^2(X)_{qs} + g^2(X)_{ps}h^2(X)_{qr} + g^2(X)_{qs}h^2(X)_{pr}\\
&+ g^2(X)_{qr}h^2(X)_{ps}\big)\mathrm{d}t.
\end{aligned}$$

We find first $\sum_{pqrs} H^T_{ip}H_{qj}H^T_{kr}H_{sm}g^2(X)_{pr}h^2(X)_{qs}$ and the other terms are similar,

$$\begin{aligned}
\sum_{pqrs} H^T_{ip}H_{qj}H^T_{kr}H_{sm}g^2(X)_{pr}h^2(X)_{qs} &= \left(\sum_{pr} H^T_{ip}g^2(X)_{pr}H_{rk}\right)\left(\sum_{qs} H^T_{jq}h^2(X)_{qs}H_{sm}\right),\\
&= \big(H^THg^2(\Lambda)H^TH\big)_{ik}\big(H^THh^2(\Lambda)H^TH\big)_{jm},\\
&= g^2(\Lambda)_{ik}h^2(\Lambda)_{jm}.
\end{aligned}$$

Repeating the analogous procedure with all of the terms we find that the covariation is

$$\mathrm{d}N_{ij}\mathrm{d}N_{km} = (g^2(\Lambda)_{ik}h^2(\Lambda)_{jm} + g^2(\Lambda)_{im}h^2(\Lambda)_{jk} + g^2(\Lambda)_{jk}h^2(\Lambda)_{im} + g^2(\Lambda)_{jm}h^2(\Lambda)_{ik})\mathrm{d}t.$$

It follows that the quadratic variation in the diagonal is

$$\mathrm{d}N_{ii}\mathrm{d}N_{jj} = 4\delta_{ij}g^2(\lambda_i)h^2(\lambda_j)\mathrm{d}t.$$

Now, in order to compute F, the finite variation part of N, we use (4.2),

$$\begin{aligned}
\mathrm{d}F &= H^Tb(X)H\mathrm{d}t + \frac{1}{2}(\mathrm{d}H^T\mathrm{d}XH + H^T\mathrm{d}X\mathrm{d}H),\\
&= b(\Lambda)\mathrm{d}t + \frac{1}{2}\left((\mathrm{d}H^TH)(H^T\mathrm{d}XH) + (H^T\mathrm{d}XH)(H^T\mathrm{d}H)\right),
\end{aligned}$$

using that the martingale part of $H^T\mathrm{d}H$ and $H^T\partial H$ coincide and the same with $H^T(\partial X)H$ and $H^T(\mathrm{d}X)H$,

$$= b(\Lambda)\mathrm{d}t + \frac{1}{2}((\mathrm{d}N\mathrm{d}A)^T + \mathrm{d}N\mathrm{d}A).$$

Now we can use (2.12) and (2.2) to find $\mathrm{d}N\mathrm{d}A$,

$$(\mathrm{d}N\mathrm{d}A)_{ij} = \sum_{k\neq j}\mathrm{d}N_{ik}\mathrm{d}A_{kj} = \sum_{k\neq j}\frac{\mathrm{d}N_{ik}\mathrm{d}N_{kj}}{\lambda_j-\lambda_k} = \delta_{ij}\sum_{k\neq j}\frac{g^2(\lambda_i)h^2(\lambda_k) + g^2(\lambda_k)h^2(\lambda_i)}{\lambda_i-\lambda_k}\mathrm{d}t.$$

Recalling that $G(x,y) = g^2(x)h^2(x) + g^2(y)h^2(y)$, we have that

$$(\mathrm{d}N\mathrm{d}A)_{ij} = \delta_{ij}\sum_{k\neq j}\frac{G(\lambda_i,\lambda_k)}{\lambda_i-\lambda_k}\mathrm{d}t.$$

From (2.2) we have that the martingale part of N_{ii} has the form $2g(\lambda_i)h(\lambda_i)\mathrm{d}W_i$ for some Brownian motion W_i. Putting together the martingale and finite variation parts of N we have that

$$\mathrm{d}N_{ii} = 2g(\lambda_i)h(\lambda_i)\mathrm{d}W_i + \sum_{k\neq j}\frac{G(\lambda_i,\lambda_k)}{\lambda_i-\lambda_k}\mathrm{d}t.$$

Since $\mathrm{d}\lambda_i = \mathrm{d}N_{ii}$, this finishes the proof. □

Theorem 2.2.2. *Let $W(t)$ be a complex $n\times n$ Brownian matrix. Suppose that $X = (X(t), t \geq 0)$ is a matrix-valued process taking values in the group of self-adjoint matrices and it satisfies the following matrix stochastic differential equation:*

$$\mathrm{d}X(t) = g(X(t))\mathrm{d}W(t)h(X(t)) + h(X(t))\mathrm{d}W(t)^*g(X(t)) + b(X(t))\mathrm{d}t, \tag{2.13}$$

with $g, h, b : \mathbb{R}\to\mathbb{R}$ and X_0 is an Hermitian $n\times n$ random matrix with n different eigenvalues.

Let $G(x,y) = g^2(x)h^2(y) + g^2(y)h^2(x)$, and

$$\tau = \inf\{t : \lambda_i(t) = \lambda_j(t) \text{ for some } i\neq j\}.$$

Then, for $t < \tau$ the eigenvalue process Λ_t verifies the following system of stochastic differential equations:

$$\mathrm{d}\lambda_i = 2g(\lambda_i)h(\lambda_i)\mathrm{d}W_i + \left(b(\lambda_i) + 2\sum_{k\neq i}\frac{G(\lambda_i,\lambda_k)}{\lambda_i-\lambda_k}\right)\mathrm{d}t, \tag{2.14}$$

where $(W_i)_i$ are independent Brownian motions.

Proof. Recall that for a complex Brownian motion Z we have that

$$\mathrm{d}\langle Z, Z\rangle(t) = 0, \qquad \mathrm{d}\langle Z, \overline{Z}\rangle(t) = 2\mathrm{d}t.$$

Then we can compute the quadratic covariation $\mathrm{d}X_{ij}\mathrm{d}X_{kl}$ using (2.13),

$$\begin{aligned}
\mathrm{d}X_{ij}\mathrm{d}X_{kl} &= \mathrm{d}\langle X_{ij}, X_{kl}\rangle(t),\\
&= \mathrm{d}\langle (g(X)\mathrm{d}W h(X) + h(X)\mathrm{d}W^* g(X))_{ij}, (g(X)\mathrm{d}W h(X) + h(X)\mathrm{d}W^* g(X))_{kl}\rangle(t),\\
&= \mathrm{d}\langle (g(X)\mathrm{d}W h(X))_{ij}, (h(X)\mathrm{d}W^* g(X))_{kl}\rangle\\
&\quad + \mathrm{d}\langle (g(X)\mathrm{d}W h(X))_{kl}, (h(X)\mathrm{d}W^* g(X))_{ij}\rangle(t),\\
&= 2g^2(X)_{il}h^2(X)_{jk}\mathrm{d}t + 2g^2(X)_{kj}h^2(X)_{li}\mathrm{d}t,\\
&= 2\left(g^2(X)_{il}h^2(X)_{kj} + g^2(X)_{jk}h^2(X)_{il}\right)\mathrm{d}t.
\end{aligned}$$

Analogously to the real case, we define A, the stochastic logarithm of H, as

$$A := H^{-1}\partial H = H^*\partial H.$$

By using Itô's formula we find,

$$0 = \mathrm{d}I = \mathrm{d}(H^*H) = H\partial H^* + (\partial H)H^* = A^* + A,$$

which means A is skew-Hermitian. This implies that the real parto of the terms in the diagonal of A is zero. Let us now apply Itô's formula to $\Lambda = H^*XH$,

$$\begin{aligned}
\mathrm{d}\Lambda = \mathrm{d}(H^*XH) &= H^*(\mathrm{d}(XH)) + (\mathrm{d}H^*)XH + \mathrm{d}(H^*)\mathrm{d}(XH),\\
&= H^*(\mathrm{d}X)H + H^*X\mathrm{d}H + H^*(\mathrm{d}X\mathrm{d}H) + (\mathrm{d}H^*)XH + \mathrm{d}(H^*)(\mathrm{d}X)H\\
&\qquad + \mathrm{d}(H^*)X\mathrm{d}H + \mathrm{d}H^*\mathrm{d}X\mathrm{d}H,\\
&= H^*(\partial X)H + H^*X\partial H + (\partial H^*)XH = H^*(\partial X)H + \Lambda H^*\partial H + (\partial H^*)H\Lambda,\\
&= H^*(\partial X)H + \Lambda\partial A + \partial A^T\Lambda = H^*(\partial X)H + \Lambda\partial A - \partial A\Lambda.
\end{aligned}$$

By the relationship between Itô's and Stratanovich's integrals,

$$H^*(\partial X)H = H^*(\mathrm{d}X)H + \frac{1}{2}(\mathrm{d}H^*(\mathrm{d}X)H + H^*\mathrm{d}X\mathrm{d}H),$$

so using that X is hermitian, we have that $H^*(\partial X)H$ is also hermitian and its diagonal elements are real. The process $\Lambda\partial A - (\partial A)\Lambda$ is zero in the diagonal and thus $\mathrm{d}\lambda_i = (H^*(\partial X)H)_{ii}$. If $i \neq j$, we have

$$0 = (H^*(\partial X)H)_{ij} + \lambda_i\partial A_{ij} - \lambda_j\partial A_{ji} = (H^*(\partial X)H)_{ij} + (\lambda_i - \lambda_j)\partial A_{ij}.$$

The last part implies $\partial A_{ij} = \frac{(H^*(\partial X)H)_{ij}}{\lambda_j - \lambda_i}$, whenever $i \neq j$.

Define $\mathrm{d}N = \mathrm{d}H^*(\partial X)H$. The martingale part of N and $H^*(\mathrm{d}X)H$ is the same, since they differ only in a finite variation term. We can find $\mathrm{d}N_{ij}\mathrm{d}N_{kl}$ using $\mathrm{d}X_{ij}\mathrm{d}X_{kl}$,

$$\mathrm{d}N_{ij}\mathrm{d}Nkl = 2(g^2(\Lambda)_{il}h^2(\Lambda)_{jk} + g^2(\Lambda)_{jk}h^2(\Lambda)_{il})\mathrm{d}t.$$

Then, for the elements in the diagonal we have

$$\mathrm{d}N_{ii}\mathrm{d}N_{jj} = 4\delta_{ij}(g^2(\lambda_i)h^2(\lambda_i))\mathrm{d}t. \tag{2.15}$$

Now we compute the finite variation part of $\mathrm{d}N$ from (2.13). Let us denote it as $\mathrm{d}F$.

$$\begin{aligned}
\mathrm{d}F &= H^*b(X)H\mathrm{d}t + \frac{1}{2}(\mathrm{d}H^*(\mathrm{d}X)H + H^*\mathrm{d}X\mathrm{d}H),\\
&= b(\Lambda)\mathrm{d}t + \frac{1}{2}\big((\mathrm{d}H^*H)(H^*\mathrm{d}XH) + (H^*\mathrm{d}XH)(H^*\mathrm{d}H)\big),\\
&= b(\Lambda)\mathrm{d}t + \frac{1}{2}\big((\mathrm{d}N\mathrm{d}A)^* + \mathrm{d}N\mathrm{d}A\big).
\end{aligned}$$

Using the quadratic variation of $\mathrm{d}N$ and $\mathrm{d}A$ we find their covariation,

$$(\mathrm{d}N\mathrm{d}A)_{ij} = \sum_k (\mathrm{d}N)_{ik}(\mathrm{d}A)_{kj} = \sum_k \frac{(\mathrm{d}N)_{ik}(\mathrm{d}N)_{kj}}{\lambda_j - \lambda_i},$$
$$= 2\delta_{ij} \sum_{k\neq j} \frac{g^2(\lambda_i)h^2(\lambda_k) + g^2(\lambda_k)h^2(\lambda_j)}{\lambda_j - \lambda_k} + \mathrm{d}N_{ij}\mathrm{d}A_{jj}.$$

By the properties shown above for $\mathrm{d}N$ and $\mathrm{d}A$, if$i = j$, $\mathrm{d}N_{jj}$ is real and $\mathrm{d}A_{jj}$ is purely imaginary. By independence of the real and imaginary parts of the complex Brownian motion, this implies that $\mathrm{d}N_{jj}\mathrm{d}A_{jj} = 0$. We have

$$\mathrm{d}F_{ii} = \left(b(\lambda_i) + 2\sum_{k\neq i} \frac{G(\lambda_i, \lambda_k)}{\lambda_i - \lambda_j} \right) \mathrm{d}t,$$

where $G(x, y) = g^2(x)h^2(y) + g^2(y)h^2(x)$.

Using the quadratic variation of $\mathrm{d}N$, we find that the martingale part of $\mathrm{d}N_{ii}$ is

$$\mathrm{d}M_{ii} = 2g(\lambda_i)h(\lambda_i)\mathrm{d}W_i,$$

for some Brownian motion. Recall that $\mathrm{d}\lambda_i = \mathrm{d}N_{ii}$, then we have that there exist $W_1, \ldots, W_p$ independent Brownian motions such that

$$\mathrm{d}\lambda_i = \mathrm{d}N_{ii} = 2g(\lambda_i)h(\lambda_i)\mathrm{d}W_i + \left(b(\lambda_i) + 2\sum_{k\neq i} \frac{G(\lambda_i, \lambda_k)}{\lambda_i - \lambda_j} \right) \mathrm{d}t.$$

This ends the proof. □

2.2.1 Non-collision of the eigenvalues

We have found that the eigenvalues of a matrix-valued diffusion process in $\mathcal{M}_{n,n}(\mathbb{R})$ and $\mathcal{M}_{n,n}(\mathbb{C})$ satisfy a system of stochastic differential equations until the first time of collision. Now we prove that this time τ is infinite a.s.

The proof we give is taken from [2] and makes use a generalized version of the so called McKean's argument appearing first in [16] and used in [12]. The original McKean's argument makes use of a multivariate function $U(x_1, \ldots, x_n)$ that has continuous derivatives and such that its evaluation on $\lambda_1, \cdots, \lambda_n$ is a local martingale. With a change of time a Girsanov theorem, one is able to conclude that the time of collision τ is infinite (for details see [12]). The generalization allows us to drop the local martingale hypothesis, by considering that the non-martingale part of the process is "well-behaved" (see hypotheses in Lemma 1.2.10).

Theorem 2.2.3. *Let $\Lambda = (\lambda_i)_{i=1,\ldots,n}$ be an n-dimensional stochastic process starting at the open Weyl chamber $\Delta_n = (\lambda_1(0) < \cdots < \lambda_n(0)) \subset \mathbb{R}^n$ and satisfying* (2.11) *with functions $g, h, b : \mathbb{R} \to \mathbb{R}$ such that g^2, h^2, b are Lipschitz continuous and g^2h^2 is convex or is continuously differentiable with derivative locally Lipschitz on $\mathbb{R}$. Then the first collision time τ defined as in* (2.10) *is infinite a.s.*

Proof. Start by defining $U := -\sum_{i<j} \log(\lambda_j - \lambda_i)$ for $t \in [0, \tau]$. Notice that, in concordance to the hypotheses of Lemma 1.2.10, $\log(x)$ is a continuous function from $\mathbb{R}^+ \setminus \{0\}$ to $\mathbb{R}$ such that $\lim_{x\downarrow 0} \log(x) = -\infty$. For each pair of different $i, j \in [n]$, such that $\lambda_i > \lambda_j$, we evaluate

$\log(\cdot)$ on the difference $Z_{ij}(t) = \lambda_i(t) - \lambda_j(t)$. The initial condition at the open Weyl chamber guarantees that $Z_{ij}(0) > 0$ for every i, j. We need to prove that, for every i, j

$$\log(Z_{ij}(t)) = Z_{ij}(0) + M_{ij}(t) + P_{ij}(t),$$

where Z_{ij} is a local martingale on $[0, \tau)$, and P_{ij} is an adapted càdlàg process such that

$$\inf_{t\in[0,\tau\wedge T)} P(t) > -\infty, \qquad \text{a.s.}$$

for every $T \in \mathbb{R}^+ \setminus \{0\}$. Working with the sum of the variables accounts for working with all the differences at once. The sign affecting U is a convention. If we change the hypothesis in 1.2.10 to consider a function h with limit $+\infty$ near 0^+, the inequality bounding P would be reversed. So we will find $P(t) < \infty$.

By the multidimensional Itô formula, we have

$$\mathrm{d}U = \sum_{i=1}^{n} \frac{\partial U}{\partial \lambda_i}\mathrm{d}\lambda_i + \frac{1}{2}\sum_{i,j} \frac{\partial^2 U}{\partial\lambda_i\partial\lambda_j}\mathrm{d}\langle\lambda_i, \lambda_j\rangle(t),$$

by the zero covariation between eigenvalues,

$$= \sum_{i=1}^{n} \frac{\partial U}{\partial x_i}\mathrm{d}\lambda_i + \frac{1}{2}\sum_{j=1}^{n} \frac{\partial^2 U}{\partial x_j^2}\mathrm{d}\langle\lambda_j, \lambda_j\rangle(t).$$

For a fixed λ_i the term $\frac{\partial U}{\partial\lambda_i}\mathrm{d}\lambda_i$ is

$$\begin{aligned}\frac{\partial U}{\partial\lambda_i}\mathrm{d}\lambda_i = \frac{\partial}{\partial\lambda_i}\left[-\sum_{k<j}\log(\lambda_j - \lambda_k)\right]\mathrm{d}\lambda_i &= \frac{\partial}{\partial\lambda_i}\left[-\sum_{k<i}\log(\lambda_i - \lambda_k) - \sum_{j>i}\log(\lambda_j - \lambda_i)\right]\mathrm{d}\lambda_i,\\ &= -\sum_{k<i}\frac{\mathrm{d}\lambda_i}{\lambda_i - \lambda_k} + \sum_{j>i}\frac{\mathrm{d}\lambda_i}{\lambda_j - \lambda_i}.\end{aligned}$$

Summing the last for every $i \in [n]$, we have

$$\sum_{i=1}^{n}\frac{\partial U}{\partial\lambda_i}\mathrm{d}\lambda_i = \sum_{i=1}^{n}\left[\sum_{j>i}\frac{\mathrm{d}\lambda_i}{\lambda_j - \lambda_i} - \sum_{k<i}\frac{\mathrm{d}\lambda_i}{\lambda_i - \lambda_k}\right] = \sum_{i<j}\frac{\mathrm{d}\lambda_i - \mathrm{d}\lambda_j}{\lambda_j - \lambda_i}.$$

For the second derivative we can use the previously found results,

$$\begin{aligned}\frac{\partial^2 U}{\partial\lambda_i^2}\mathrm{d}\langle\lambda_i, \lambda_i\rangle(t) &= \frac{\partial}{\partial\lambda_i}\left[-\sum_{k<i}\frac{1}{\lambda_i - \lambda_k} + \sum_{j>i}\frac{1}{\lambda_j - \lambda_i}\right]\mathrm{d}\langle\lambda_i, \lambda_i\rangle(t),\\ &= \sum_{k<i}\frac{\mathrm{d}\langle\lambda_i, \lambda_i\rangle(t)}{(\lambda_i - \lambda_k)^2} + \sum_{j>i}\frac{\mathrm{d}\langle\lambda_i, \lambda_i\rangle(t)}{(\lambda_j - \lambda_i)^2}.\end{aligned}$$

Summing over all the values $i \in [n]$, we have

$$\sum_{i=1}^{n}\sum_{i=1}^{n}\frac{\partial U}{\partial \lambda_i}\mathrm{d}\lambda_i = \sum_{i<j}\frac{\mathrm{d}\langle\lambda_i,\lambda_i\rangle(t)+\mathrm{d}\langle\lambda_j,\lambda_j\rangle(t)}{(\lambda_j-\lambda_i)^2}$$

Using what we found for the derivatives and the fact that $\mathrm{d}\lambda_i\mathrm{d}\lambda_i = 4g^2(\lambda_i)h^2(\lambda_i)$, we find

$$\begin{aligned}\mathrm{d}U &= \sum_{i<j}\left[\frac{\mathrm{d}\lambda_i-\mathrm{d}\lambda_j}{\lambda_j-\lambda_i}+\frac{1}{2}\frac{\mathrm{d}\langle\lambda_i,\lambda_i\rangle+\mathrm{d}\langle\lambda_j,\lambda_j\rangle}{(\lambda_j-\lambda_i)^2}\right],\\ &= \sum_{i<j}\left[\frac{\mathrm{d}\lambda_i-\mathrm{d}\lambda_j}{\lambda_j-\lambda_i}+2\frac{g^2(\lambda_i)h^2(\lambda_i)+g^2(\lambda_j)h^2(\lambda_j)}{(\lambda_j-\lambda_i)^2}\mathrm{d}t\right].\end{aligned}$$

Now we expand the terms $\mathrm{d}\lambda_i$ in the first summand,

$$\begin{aligned}&\frac{\mathrm{d}\lambda_i-\mathrm{d}\lambda_j}{\lambda_j-\lambda_i} =\\ &=\frac{2g(\lambda_i)h(\lambda_i)\mathrm{d}W_i+\left(\sum_{k\neq i}\frac{G(\lambda_i,\lambda_k)}{\lambda_i-\lambda_k}+b(\lambda_i)\right)\mathrm{d}t-2g(\lambda_j)h(\lambda_j)\mathrm{d}W_j-\left(\sum_{k\neq j}\frac{G(\lambda_j,\lambda_k)}{\lambda_j-\lambda_k}+b(\lambda_j)\right)\mathrm{d}t}{\lambda_j-\lambda_i}.\end{aligned}$$

Denote by $\mathrm{d}M$ the martingale part of U. Summing the martingale part of for every pair of values i, j in the last expression we find

$$\mathrm{d}M = 2\sum_{i<j}\frac{g(\lambda_i)h(\lambda_i)\mathrm{d}W_i-g(\lambda_j)h(\lambda_j)\mathrm{d}W_j}{\lambda_j-\lambda_i}.$$

The finite variation part of U, $\mathrm{d}P$ is given by $\sum_{i<j}2\frac{g^2(\lambda_i)h^2(\lambda_i)+g^2(\lambda_j)h^2(\lambda_j)}{(\lambda_j-\lambda_i)^2}\mathrm{d}t$ together with the finite variation part of $\sum_{i<j}\frac{\mathrm{d}\lambda_i-\mathrm{d}\lambda_j}{\lambda_j-\lambda_i}$. This last term is

$$\begin{aligned}&\sum_{i<j}\frac{b(\lambda_i)-b(\lambda_j)}{\lambda_j-\lambda_i}\mathrm{d}t+\sum_{i<j}\frac{1}{\lambda_j-\lambda_i}\left(\sum_{k\neq i}\frac{G(\lambda_i,\lambda_k)}{\lambda_i-\lambda_k}-\sum_{k\neq j}\frac{G(\lambda_j,\lambda_k)}{\lambda_j-\lambda_k}\right)\mathrm{d}t,\\ &=\sum_{i<j}\frac{b(\lambda_i)-b(\lambda_j)}{\lambda_j-\lambda_i}\mathrm{d}t+\sum_{i<j}\frac{1}{\lambda_j-\lambda_i}\left[\sum_{k\neq i,j}\left(\frac{G(\lambda_i,\lambda_k)}{\lambda_i-\lambda_k}-\frac{G(\lambda_j,\lambda_k)}{\lambda_j-\lambda_k}\right)-\frac{2G(\lambda_i,\lambda_j)}{\lambda_j-\lambda_i}\right]\mathrm{d}t.\end{aligned}$$

In order to organize the finite variation part, we define the following processes:

$$
\begin{aligned}
\mathrm{d}A_1 &:= \sum_{i<j} \frac{b(\lambda_i) - b(\lambda_j)}{\lambda_j - \lambda_i}\mathrm{d}t, \\
\mathrm{d}A_2 &:= 2\sum_{i<j} \frac{g^2(\lambda_i)h^2(\lambda_i) + g^2(\lambda_i)h^2(\lambda_i) - G(\lambda_i,\lambda_j)}{(\lambda_j - \lambda_i)^2}\mathrm{d}t, \\
\mathrm{d}A_3 &:= \sum_{i<j} \frac{1}{\lambda_j - \lambda_i} \sum_{k\neq i, k\neq j} \left(\frac{G(\lambda_i,\lambda_k)}{\lambda_i - \lambda_k} - \frac{G(\lambda_j,\lambda_k)}{\lambda_j - \lambda_k} \right) \mathrm{d}t.
\end{aligned}
$$

Notice that $\mathrm{d}P = \mathrm{d}A_1 + \mathrm{d}A_2 + \mathrm{d}A_3$. The Lipschitz condition on b implies that

$$
\begin{aligned}
|A_1(t)| &\leq |\sum_{i<j} \int_0^t \frac{b(\lambda_i) - b(\lambda_j)}{\lambda_j - \lambda_i}\mathrm{d}s| \leq \sum_{i<j} \int_0^t \frac{|b(\lambda_i) - b(\lambda_j)|}{\lambda_j - \lambda_i}\mathrm{d}s \leq \sum_{i<j} \frac{K|\lambda_i - \lambda_j|}{\lambda_j - \lambda_i}t, \\
&= \sum_{i<j} Kt = Kt\frac{n(n-1)}{2},
\end{aligned}
$$

where K is the cpnstant coming from the Lipschitz continuity. Now, for A_2, we have

$$
\begin{aligned}
\mathrm{d}A_2 &= 2\sum_{i<j} \frac{g^2(\lambda_i)h^2(\lambda_i) + g^2(\lambda_i)h^2(\lambda_i) - g^2(\lambda_i)h^2(\lambda_j) - g^2(\lambda_j)h^2(\lambda_i)}{(\lambda_j - \lambda_i)^2}\mathrm{d}t, \\
&= 2\sum_{i<j} \frac{\left(g^2(\lambda_j) - g^2(\lambda_i)\right)\left(h^2(\lambda_j) - h^2(\lambda_i)\right)}{(\lambda_j - \lambda_i)^2}\mathrm{d}t.
\end{aligned}
$$

Again, we use the Lipschitz continuity of h^2 and g^2 to find a bound. We suppose that the constant for the Lipschitz continuity is also K.

$$
|A_2(t)| \leq 2\sum_{i<j} K^2 t = K^2 t n(n-1).
$$

For the term $\mathrm{d}A_3$, we have

$$
\begin{aligned}
\mathrm{d}A_3 &= \sum_{i<j, k\neq i,j} \frac{G(\lambda_i,\lambda_k)(\lambda_j - \lambda_k) - G(\lambda_j,\lambda_k)(\lambda_i - \lambda_k)}{(\lambda_i - \lambda_k)(\lambda_j - \lambda_k)(\lambda_j - \lambda_i)}\mathrm{d}t, \\
&= \sum_{i<j<k} \frac{G(\lambda_i,\lambda_k)(\lambda_j - \lambda_k) - G(\lambda_j,\lambda_k)(\lambda_i - \lambda_k)}{(\lambda_i - \lambda_k)(\lambda_j - \lambda_k)(\lambda_j - \lambda_i)}\mathrm{d}t \\
&\qquad + \sum_{i<k<j} \frac{G(\lambda_i,\lambda_k)(\lambda_j - \lambda_k) - G(\lambda_j,\lambda_k)(\lambda_i - \lambda_k)}{(\lambda_i - \lambda_k)(\lambda_j - \lambda_k)(\lambda_j - \lambda_i)}\mathrm{d}t \\
&\qquad\qquad + \sum_{k<i<j} \frac{G(\lambda_i,\lambda_k)(\lambda_j - \lambda_k) - G(\lambda_j,\lambda_k)(\lambda_i - \lambda_k)}{(\lambda_i - \lambda_k)(\lambda_j - \lambda_k)(\lambda_j - \lambda_i)}\mathrm{d}t.
\end{aligned}
$$

By reassigning the indices in the last two sums, we arrive to the equivalent expression;

$$\begin{aligned}
\mathrm{d}A_3 = \sum_{i<j<k} \Bigg[& \frac{G(\lambda_i,\lambda_k)(\lambda_j-\lambda_k) - G(\lambda_j,\lambda_k)(\lambda_i-\lambda_k)}{(\lambda_i-\lambda_k)(\lambda_j-\lambda_k)(\lambda_j-\lambda_i)} \\
& + \frac{G(\lambda_i,\lambda_j)(\lambda_k-\lambda_j) - G(\lambda_k,\lambda_j)(\lambda_i-\lambda_j)}{(\lambda_i-\lambda_j)(\lambda_k-\lambda_j)(\lambda_k-\lambda_i)} \\
& + \frac{G(\lambda_j,\lambda_i)(\lambda_k-\lambda_i) - G(\lambda_k,\lambda_i)(\lambda_j-\lambda_i)}{(\lambda_j-\lambda_i)(\lambda_k-\lambda_i)(\lambda_k-\lambda_j)} \Bigg] \mathrm{d}t, \\
= \sum_{i<j<k} & \frac{G(\lambda_j,\lambda_k)(\lambda_k-\lambda_j) - G(\lambda_i,\lambda_k)(\lambda_i-\lambda_k) + G(\lambda_i,\lambda_j)(\lambda_j-\lambda_i)}{(\lambda_j-\lambda_i)(\lambda_k-\lambda_i)(\lambda_k-\lambda_j)} \mathrm{d}t.
\end{aligned}$$

Now, define the funtion $H(x,y,z)$ as

$$\begin{aligned}
H(x,y,z) &\coloneqq \left[(g^2(x)-g^2(z))(h^2(y)-h^2(z)) + (g^2(y)-g^2(z))(h^2(x)-h^2(z))\right](y-x), \\
&= \left(G(x,y) - G(x,z) - G(y,z) + G(z,z)\right)(y-x).
\end{aligned}$$

By the Lipschitz conditions on g^2 and h^2 we find that

$$\begin{aligned}
|H(x,y,z| &\leq |y-x|[K|x-z|K|y-z| + K|y-z|K|x-z|], \\
&= 2K^2|(y-x)(x-z)(y-z)|.
\end{aligned}$$

Also, we have the equality

$$\begin{aligned}
H(x,y,z) + H(y,z,x) - H(x,z,y) = & \left(G(x,y) - G(x,z) - G(y,z) + G(z,z)\right)(y-x) \\
& + \left(G(y,z) - G(y,x) - G(z,x) + G(x,x)\right)(z-y) \\
& - \left(G(x,z) - G(x,y) - G(z,y) + G(y,y)\right)(z-x), \\
= & \, 2(z-y)G(y,z) - 2(z-x)G(x,z) + 2(y-x)G(x,y) \\
& + G(x,x)(z-y) - G(y,y)(z-x) + G(z,z)(y-x).
\end{aligned}$$

Using the last expression, we can write $2\mathrm{d}A_3$ as

$$\begin{aligned}
2\mathrm{d}A_3 &= 2\sum_{i<j<k} \frac{G(\lambda_j,\lambda_k)(\lambda_k-\lambda_j) - G(\lambda_i,\lambda_k)(\lambda_i-\lambda_k) + G(\lambda_i,\lambda_j)(\lambda_j-\lambda_i)}{(\lambda_j-\lambda_i)(\lambda_k-\lambda_i)(\lambda_k-\lambda_j)} \mathrm{d}t, \\
&= \sum_{i<j<k} \frac{H(\lambda_i,\lambda_j,\lambda_k) + H(\lambda_j,\lambda_k,\lambda_i) - H(\lambda_i,\lambda_k,\lambda_j)}{(\lambda_j-\lambda_i)(\lambda_k-\lambda_i)(\lambda_k-\lambda_j)} \mathrm{d}t \\
&\quad + \sum_{i<j<k} \frac{-G(\lambda_i,\lambda_i)(\lambda_k-\lambda_j) + G(\lambda_j,\lambda_j)(\lambda_k-\lambda_i) - G(\lambda_k,\lambda_k)(\lambda_j-\lambda_i)}{(\lambda_j-\lambda_i)(\lambda_k-\lambda_i)(\lambda_k-\lambda_j)} \mathrm{d}t.
\end{aligned}$$

Assign the names $\mathrm{d}A_4$ and $\mathrm{d}A_5$ to the last summands. For A_4 we have, using the bound on $H(x,y,z)$,

$$|A_4| \leq \sum_{i<j<k} 6K^2 t = \frac{6K^2 n(n-1)(n-2)t}{6} = K^2 n(n-1)(n-2)t.$$

Finally, we can re-write $\mathrm{d}A_5$ as

$$\begin{aligned}\mathrm{d}A_5(t) &= \sum_{i<j<k} \frac{G(\lambda_j,\lambda_j)(\lambda_k-\lambda_i) - G(\lambda_i,\lambda_i)(\lambda_k-\lambda_j) - G(\lambda_k,\lambda_k)(\lambda_j-\lambda_i)}{(\lambda_j-\lambda_i)(\lambda_k-\lambda_i)(\lambda_k-\lambda_j)}\mathrm{d}t, \\ &= \sum_{i<j<k} \left(\frac{G(\lambda_j,\lambda_j)-G(\lambda_i,\lambda_i)}{\lambda_j-\lambda_i} - \frac{G(\lambda_k,\lambda_k)-G(\lambda_j,\lambda_j)}{\lambda_k-\lambda_j}\right)\frac{1}{\lambda_k-\lambda_i}\mathrm{d}t. \qquad (2.16)\end{aligned}$$

If $G(x,x)$ is convex, then

$$\frac{G(x,x)-G(y,y)}{x-y},$$

is monotone decreasing in every variable letting the other fixed and thus A_5 is non-positive. If $G(x,x)$ is continuously differentiable with derivative locally Lipschitz, then

$$|G'(x,x) - G'(y,y)| \leq C|x-y|,$$

and we conclude that every summand in (2.16) is bounded by C, which means $|A_5(t)| \leq Ctn(n-1)(n-2)/6$.

We have found that the finite variation part of U is bounded for every finite t, then all the hypotheses of Lemma 1.2.10 are satisfied and we can conclude that $\tau = \infty$ a.s.

□

The following result is the first part of a multidimensional version of Yamada-Watanabe theorem. It is the main result in [2] and we will use it to prove the uniqueness of the solutions to (2.11).

Theorem 2.2.4 (Multidimensional Yamada-Watanabe Theorem [2])**.** *Let $n \in \mathbb{N}$ and*

$$b_i : \mathbb{R}^n \to \mathbb{R}, \qquad i = 1, \ldots, n,$$

be real-valued bounded continuous functions satisfying the following Lipschitz conditions for $C > 0$,

$$|b_i(y_1) - b_i(y_2)| \leq C \left\|y_1 - y_2\right\|, \quad i = 1, \ldots, n,$$

for every $y_1, y_2 \in \mathbb{R}^n$.

Further, let $\sigma_i : \mathbb{R} \to \mathbb{R}, i = 1, \ldots, n$ be a set of bounded measurable functions such that

$$|\sigma_i(x) - \sigma_i(y)|^2 \leq \rho_i(|x-y|), \quad x, y \in \mathbb{R},$$

where $\rho_i : (0,\infty) \to (0,\infty)$ are measurable functions such that

$$\int_{0^+} \rho_i^{-1}(x)\ \mathrm{d}x = \infty.$$

Then the pathwise uniqueness holds for the following system of stochastic differential equations

$$\mathrm{d}Y_i = \sigma_i(Y_i)\mathrm{d}B_i + b_i(Y)\mathrm{d}t, \qquad i = 1, \dots, n, \tag{2.17}$$

where $B_1, \dots, B_n$ are independent Brownian motions.

Proof. Let Y and $\hat{Y}$ be two solutions with respect to the same multidimensional Brownian motion $B = (B_i)_{i\leq n}$ such that $Y(0) = \hat{Y}(0)$ a.s., for $i \leq n$ we have

$$Y_i(t) - \hat{Y}_i(t) = \int_0^t \left(\sigma_i(Y_i) - \sigma_i(\hat{Y}_i)\right)\ \mathrm{d}B_i(s) + \int_0^t \left(b_i(Y_i) - b_i(\hat{Y}_i)\right)\ \mathrm{d}s. \tag{2.18}$$

We can then see that

$$\int_0^t \frac{\mathbb{1}_{\{Y_i(s)>\hat{Y}_i(s)\}}}{\rho_i(Y_i(s) - \hat{Y}_i(s))}\mathrm{d}\langle Y_i - \hat{Y}_i, Y_i - \hat{Y}_i\rangle = \int_0^t \frac{(\sigma_i(Y_i(s)) - \sigma_i(\hat{Y}_i(s)))^2}{\rho_i(Y_i(s) - \hat{Y}_i(s))}\mathbb{1}_{\{Y_i(s)>\hat{Y}_i(s)\}}\ \mathrm{d}s \leq t.$$

Applying Theorem 1.2.8 we have that the local time of $Y_i - \hat{Y}_i$ at 0 is 0. Then, we can use the Tanaka formula to find

$$\begin{aligned}
|Y_i(t) - \hat{Y}_i(t)| &= \int_0^t \operatorname{sgn}(Y_i(s) - \hat{Y}_i(s))\ \mathrm{d}(Y_i(s) - \hat{Y}_i(s)), \\
&= \int_0^t \operatorname{sgn}(Y_i(t) - \hat{Y}_i(t))(\sigma_i(Y_i) - \sigma_i(\hat{Y}_i))\ \mathrm{d}B_i(s) \\
&\qquad + \int_0^t \operatorname{sgn}(Y_i(s) - \hat{Y}_i(s))(b_i(Y_i(s)) - b_i(\hat{Y}_i(s)))\ \mathrm{d}s.
\end{aligned}$$

Since σ_i is bounded, we have that $\operatorname{sgn}(Y_i(t) - \hat{Y}_i(t))(\sigma_i(Y_i(t)) - \sigma_i(\hat{Y}_i(t)))$ is bounded and therefore the first integral in the last expression is a martingale with mean 0, which in turns implies that

$$|Y_i(t) - \hat{Y}_i(t)| - \int_0^t \operatorname{sgn}(Y_i(s) - \hat{Y}_i(s))(b_i(Y_i(s)) - b_i(\hat{Y}_i(s)))\ \mathrm{d}s,$$

is a zero-mean martingale. Then, by using the Lipschitz properties of b_i we have

$$\begin{aligned}\left\|Y_i(t)-\hat{Y}_i(t)\right\|_1 &= \mathbb{E}\left[\int_0^t \operatorname{sgn}(Y_i(s)-\hat{Y}_i(s))(b_i(Y_i(s))-b_i(\hat{Y}_i(s)))\,\mathrm{d}s\right],\\ &\leq \mathbb{E}\left[\int_0^t |b_i(Y_i(s))-b_i(\hat{Y}_i(s))|\,\mathrm{d}s\right],\\ &= \int_0^t \left\|b_i(Y_i(s))-b_i(\hat{Y}_i(s))\right\|_1\,\mathrm{d}s \leq C\int_0^t \left\|Y_i(s)-\hat{Y}_i(s)\right\|_1\,\mathrm{d}s.\end{aligned}$$

The L^1 norm in the last term is taken integrating with respect to ω. Summing for every i we get

$$\left\|Y(t)-\hat{Y}(t)\right\|_1 \leq C\int_0^t \left\|Y(t)-\hat{Y}(s)\right\|_1\,\mathrm{d}s.$$

Using Gronwall's lemma (1.2.9) and the equalilty in the initial conditions of Y and $\hat{Y}$, we get that

$$\left\|Y(t)-\hat{Y}(t)\right\|_1 \leq \left\|Y(0)-\hat{Y}(0)\right\|_1 e^{ct} = 0.$$

Thus, we can conclude that

$$\left\|Y(t)-\hat{Y}(t)\right\|_1 = 0,$$

which implies $Y(t)=\hat{Y}(t)$ a.s. for every fixed $t>0$. The pathwise uniqueness follows from the continuity. □

With the next theorem, the result on the pathwise uniqueness is particularized for the cases of systems of stochastic differential equations coming from eigenvalue systems.

Theorem 2.2.5 (Spectral matrix Yamada-Watanabe theorem)**.** *Let $X(t)$ be an $n\times n$ symmetric matrix-valued process satisfying the equation* (4.2) *with initial condition $X(0)$ that is a symmetric $n\times n$ matrix with n different eigenvalues. Suppose further that*

$$|g(x)h(x)-g(y)h(y)|^2 \leq \rho(|x-y|), \qquad x,y\in\mathbb{R}, \tag{2.19}$$

with $\rho:(0,\infty)\to(0,\infty)$ a measurable function satisfying

$$\int_{0^+}\rho^{-1}(x)\mathrm{d}x = \infty,$$

that $G(x,y) \coloneqq g^2(x)h^2(y)+g^2(y)h^2(x)$ is locally Lipschitz and strictly positive on the set $\{x\neq y\}$ and that b is locally Lipschitz. Then if τ is defined as in (2.10)*, for $t<\tau$, the process of eigenvalues satisfying* (2.11) *has a pathwise unique solution.*

Proof. Let $H_0\Lambda_0H_0^T$ be a diagonalization for $X(0)$. We need to show that a unique strong solution exists for (2.11) when $\Lambda(0)=\Lambda_0$. The functions

$$a_i(\lambda_1,\ldots,\lambda_n) = b(\lambda_i)+\sum_{k\neq i}\frac{G(\lambda_i,\lambda_k)}{\lambda_i-\lambda_k},$$

are locally Lipschitz continuous on $\Delta_n = \{0 \leq \lambda_1 < \lambda_2 < \cdots < \lambda_n\}$ so they can be extended from the compact sets

$$D_m = \{0 \leq \lambda_1 < \lambda_2 < \cdots < \lambda_n < m, \lambda_{i+1} - \lambda_i \geq 1/n\},$$

to bounded Lipschitz functions on $\mathbb{R}^n$. Let a_i^m denote such extension for $m \in \mathbb{N} \setminus \{0\}$. For $i = 1, \ldots, n$, we consider the system of stochastic differential equations,

$$\mathrm{d}\lambda_i^m = 2g(\lambda_i^m)h(\lambda_i^m)\mathrm{d}W_i + a_i^m(\Lambda^m)\mathrm{d}t. \tag{2.20}$$

We have that $|g(x)h(x) - g(y)h(y)|^2 \leq \rho(|x-y|)$ and $\int_{0^+} \rho(x)^{-1}\mathrm{d}x = \infty$, and using Theorem 2.2.4 we get that there is a unique strong solution for the system (2.20) for every m. Due to the path uniqueness, and since $D_m \subset D_{m+1}$, for every pair m, m' we have that λ^m and $\lambda^{m'}$ coincide on $D_{m \wedge m'}$. Letting $m \to \infty$ we have that $D_m \to \Delta_n$, and $\lim_{n\to} \lambda_i^m$ is a solution on Δ_n. If λ_i is another solution on Δ_n, they must coincide in every D_m and thus they coincide in Δ_n.

So there is a unique strong solution $\Lambda(t)$ for the system of stochastic differential equations up to the first exit time from Δ_n. This time is τ, the first collision time of the eigenvalues. □

From Theorem 2.2.3 we already knwo that the time of first collision τ is infinite a.s. In the next corollary, we summarize both results.

Corollary 2.2.6. *Suppose that b, g^2, h^2 are Lipschitz continuous, g^2h^2 is convex or continuously differentiable with derivative locally Lipschitz on $\mathbb{R}$ and that $G(x,y) \coloneqq g^2(x)h^2(y) + g^2(y)h^2(x)$ is strictly positive on $\{x \neq y\}$. Then the system of SDEs* (2.11) *for the eigenvalue process satisfying* (4.2) *has a unique strong solution on $[0, \infty)$.*

Proof. Let f is a non-negative Lipschitz continuous function, define $\sqrt{f}$ as the positive square root of f. Using that $\sqrt{x}$ is 1/2-Hölder continuous, we have that

$$|\sqrt{f(x)} - \sqrt{f(y)}| \leq |f(x) - f(y)|^{\frac{1}{2}} \leq K|x-y|^{\frac{1}{2}}.$$

Then $\sqrt{f}$ is 1/2-Hölder continuous. Since g^2 and h^2 are Lipschitz continuous, then g^2h^2 is locally Lipschitz continuous and gh is 1/2-Hölder continuous. Then

$$|g(x)h(x) - g(y)h(y)|^2 \leq \left(K|x-y|^{\frac{1}{2}}\right)^2 = K^2|x-y|.$$

Taking $\rho(|x-y|) = K^2|x-y|$ we see that the conditions of Theorem 2.2.5 are satisfied and then the uniqueness and existence of the strong solution applies on $[0, \tau)$. By Theorem 2.2.3 we have that $\tau = \infty$ a.s., thus the existence and uniqueness is satisfied on $[0, \infty)$. □

With these results, we are are ready to apply to two processes of interest.

2.2.2 Wishart process

The Wishart process is a dynamic version of the Wishart matrix, first described by John Wishart [17]. If we assume a data population consisting of n features observed in k individuals, we can form a rectangular $n \times k$ array with this data; let us name this array X. If we further assume that the variables and individuals are completely uncorrelated, and that individual data points (the ith feature of the jth individual) follow a standard normal distribution, then

X is a standard independent Gaussian matrix of size $n \times k$. The matrix X can be thought of as a size k sample observation of an independent normal vector $\vec{v}$ in $\mathbb{R}^k$. It is a well-known fact in statistics that an estimator for the covariance matrix of $\vec{v}$ is $W \coloneqq X^T X$. Under the aforementioned assumptions, W follows a Wishart distribution.

Besides estimating a covariance matrix, the Wishart matrix distribution has other applications in multivariate statistics. A notable one is its use in Principal Component Analysis (PCA). In [18], Principal Components are described as "a sequence of projections of the data, mutually uncorrelated and ordered in variance". The key idea is that if we find the eigenvalues of a covariance matrix and order them, we can identify the most influential features in the data variance, i.e., the eigenvector associated with the largest eigenvalue carries the most variance of the data. Each eigenvector is a "projection of the data" and is uncorrelated with any other eigenvector, as they form an orthogonal basis.

A natural question that arises when performing PCA is what happens to the Principal Components if we add a mild perturbation (e.g., Gaussian noise). Bru [12] addressed this problem by considering the addition of Brownian motion as noise. This approach allowed her to use stochastic calculus techniques to study the behavior of the Principal Components when the added variance fluctuated.

In this subsection, we use the theorems proven previously to replicate Bru's results. It is worth mentioning that the techniques in this section taken from [2] were originally inspired by the work of Bru in [12].

Once we have Theorems 2.2.1 and 2.2.2, proving the form of the eigenvalues in a Wishart process is quite straightforward.

Corollary 2.2.7. *Let $\tilde{B} = (\tilde{B}(t), t \geq 0)$ be a Brownian motion in $\mathcal{M}_{n,m}(\mathbb{R})$ with $n \geq m$ and define $X = \tilde{B}^T \tilde{B}$. Then the eigenvalues of X, $\lambda_1 > \lambda_2 > ... > \lambda_n$ are given by the unique strong solution to the following system of stochastic differential equations*

$$\mathrm{d}\lambda_i = 2\sqrt{\lambda_i}\mathrm{d}W_i + \left(m + \sum_{k \neq i} \frac{|\lambda_i| + |\lambda_k|}{\lambda_i - \lambda_k} \right) \mathrm{d}t.$$

Morover, if Y is any matrix-valued stochastic process satisfying the stochastic differential equation

$$\mathrm{d}Y(t) = \sqrt{Y(t)}\mathrm{d}B(t) + \mathrm{d}B^T(t)\sqrt{Y(t)} + \alpha I \mathrm{d}t,$$

with respect to $B = (B(t), t \geq 0)$ a Brownian motion in $\mathcal{M}_{n \times n}(\mathbb{R})$ and $n \geq p - 1$, then its eigenvalues are the unique strong solution to the system of stochastic differential equations

$$\mathrm{d}\lambda_i = 2\sqrt{\lambda_i}\mathrm{d}W_i + \left(\alpha + \sum_{k \neq i} \frac{|\lambda_i| + |\lambda_k|}{\lambda_i - \lambda_k} \right) \mathrm{d}t. \tag{2.21}$$

Proof. We prove first that X satisfies

$$\mathrm{d}X(t) = \sqrt{X(t)}\mathrm{d}B(t) + \mathrm{d}B^T(t)\sqrt{X(t)} + \alpha I \mathrm{d}t,$$

for an $n \times n$ matrix-valued Brownian motion, and then use Theorem 2.2.1.

By the matrix Itô formula we have for X,

$$\mathrm{d}X(t) = (\mathrm{d}\tilde{B}^T(t))\tilde{B}(t) + (\tilde{B}^T(t))\mathrm{d}\tilde{B}(t) + \mathrm{d}\langle \tilde{B}^T, \tilde{B} \rangle(t).$$

For the covariation term, we can find

$$\mathrm{d}\langle \tilde{B}^T, \tilde{B}\rangle(t)_{ij} = \sum_{k=1}^{m} \mathrm{d}\langle \tilde{B}_{ki}, \tilde{B}_{kj}\rangle(t) = m\delta_{ij}\mathrm{d}t.$$

This means $\mathrm{d}\langle \tilde{B}^T, \tilde{B}\rangle(t) = mI\mathrm{d}t$.

For the remaining terms, we find the covariation,

$$(\tilde{B}^T \mathrm{d}\tilde{B})_{ij}((\mathrm{d}\tilde{B})^T \tilde{B})_{kl} = X_{il}\delta_{jk}\mathrm{d}t,$$

which in total accounts for

$$\mathrm{d}\langle X_{ij}, X_{kl}\rangle(t) = (X_{ik}\delta_{jl} + X_{il}\delta_{jk} + X_{jk}\delta_{il} + X_{jl}\delta_{ik})\mathrm{d}t.$$

With this, we can find the quadratic variation for the diagonal and off-diagonal entries of X.

$$\mathrm{d}\langle X_{ij}, X_{ij}\rangle(t) = \begin{cases} (X_{ii} + X_{jj})\mathrm{d}t & \text{if } i \neq j, \\ 4X_{ii}\mathrm{d}t & \text{if } i = j. \end{cases}$$

With this covariations, we find that the entries of $\mathrm{d}X$ coincide with those of $\sqrt{X(t)}\mathrm{d}B(t) + \mathrm{d}B^T(t)\sqrt{X(t)} + \alpha I\mathrm{d}t$. Now, in Theorem 2.2.1 substitute $g(x) = \sqrt{x}, h(x) \equiv 1$ and $b(x) \equiv \alpha$ to find that (2.21) is satisfied. □

Using Brownian motions in $\mathcal{M}_{n,m}(\mathbb{C})$ instead and repeating all the steps we find the corresponding equation for the complex Wishart process. Also, the matrix can be rescaled as in the Dyson Brownian motion case to find a version where the β parameter affects the martingale part.

2.2.3 Jacobi process

Similarly to the Wishart case, the Jacobi process is a dynamical generalization of a random matrix used in statistics. The Jacobi process typically appears in two contexts: one is in the multivariate analysis of variance (MANOVA) [19], where it is defined as the "quotient" of a Wishart matrix and the sum of that matrix with an independent Wishart matrix, i.e.,

$$J \coloneqq (W_1 + W_2)^{-1} W_1,$$

where W_1, W_2 are indepenent Wishart matrices. In this context, the Jacobi matrix (known as MANOVA matrix) is a generalization of an F distribution in the context of univariate ANOVA.

The second context, and the one we are interested in is the Generalized Singular Value Decomposition (GSVD) algorithm [20] which is used for a block matrix. The construction we give here is found in [21] and it coincides with the construction of a Beta-matrix in [22].

Let M be an independent Gaussian matrix in $\mathcal{M}_{m,n}(\mathbb{R})$ with $n \leq m$. We can decompose M as a block matrix in the way

$$M = \begin{bmatrix} M_1 \\ M_2 \end{bmatrix}.$$

So $M_1 \in \mathcal{M}_{n_1,n}(\mathbb{R})$ and $M_2 \in \mathcal{M}_{n_2,n}(\mathbb{R})$ are independent Gaussian matrices in its respective spaces, and $n_1 + n_2 = m$. With the GSVD algorithm we can find simultaneously singular value decompositions for M_1 and M_2 such that $M_1 = U_1 C H, M_2 = U_2 S H$ where $U_1 \in \mathcal{M}_{n_1,n_1}(\mathbb{R}), U_2 \in \mathcal{M}_{n_2,n_2}(\mathbb{R})$ are orthogonal matrices, $C \in \mathcal{M}_{n_1,n}, S \in \mathcal{M}_{n_2,n}(\mathbb{R})$ are pseudo diagonals satisfying $C^T C + S^T S = I_{n\times n}$, and $H \in \mathcal{M}_{n,n}$ is invertible. Although the decomposition is not unique, it can be taken so that U_1, U_2, H are Haar distributed, mutually independent and independent from C, S.

Then we take $W_1 = M_1^T M_1$ and $W_2 = M_1^T M_1$, we have that W_1 and W_2 are $n \times n$ Wishart matrices with shape parameters n_1 and n_2, respectively. Our matrix of interest is

$$J \coloneqq (W_1 + W_2)^{-\frac{1}{2}} W_1 (W_1 + W_2)^{-\frac{1}{2}}.$$

With the singular values decomposition of M_1 and M_2, we notice that

$$\begin{aligned}
\det[J] &= \det[(W_1 + W_2)^{-\frac{1}{2}} W_1 (W_1 + W_2)^{-\frac{1}{2}}], \\
&= \det[(H^T C^T C H + H^T S^T S H)^{-\frac{1}{2}} H^T C^T C H (H^T C^T C H + H^T S^T S H)^{-\frac{1}{2}}], \\
&= \det[(C^T C + S^T S)^{\frac{1}{2}} C^T C (C^T C + S^T S)^{\frac{1}{2}}], \\
&= \det C^T C.
\end{aligned}$$

So J has the same eigenvalues of $C^T C$. If we give a singular values decomposition for M, $M = VDU$ with $D \in \mathcal{M}_{m,n}(\mathbb{R})$ pseudodiagonal $D = (\Delta, 0)^T$, Δ diagonal, and $U \in \mathcal{M}_{n,n}(\mathbb{R}), V \in \mathcal{M}_{m,m}(\mathbb{R})$ Haar unitaries and independent. Then $M^T M = U^T \Delta^2 U$, and

$$(W_1 + W_2)^{\frac{1}{2}} = (U^T \Delta U)^{\frac{1}{2}}.$$

Letting X be the $m \times n_1$ upper left corner of V we have that $M_1 = U \Delta X$, and then

$$M_1^T M_1 = U^T \Delta X^T X \Delta U = (M^T M)^{\frac{1}{2}} (U^T X^T X U) (M^T M)^{\frac{1}{2}}.$$

Substituting this in our previous definition for J we have

$$\begin{aligned}
J &= (W_1 + W_2)^{-\frac{1}{2}} W_1 (W_1 + W_2)^{-\frac{1}{2}}, \\
&= (M^T M)^{-\frac{1}{2}} (M^T M)^{\frac{1}{2}} (U^T X^T X U) (M^T M)^{\frac{1}{2}} (M^T M)^{-\frac{1}{2}} = (XU)^T XU.
\end{aligned}$$

Since X is invariant under multiplication by orthogonal matrices, and U, X are independent, then the law of J is equal to the law of $X^T X$. The conclusion is that we can build the Jacobi matrix in two different ways, one using the standard definition as in the MANOVA case and another as the square of the upper left corner in a Haar-distributed random matrix. This was discovered by Collins [23] and is used by Doumerc to construct the dynamical version of this matrix [22].

For the stochastic process case, it is shown in [22] that if X is the upper corner of a Haar unitary Brownian motion, then $J \coloneqq X^T X$ satisfies the following stochastic differential equation

$$\mathrm{d}J(t) = \sqrt{J(t)}\mathrm{d}B(t)\sqrt{I_n - J(t)} + \sqrt{I_n - J(t)}\mathrm{d}B(t)\sqrt{J(t)} + (n_2 I_n - (n_1 + n_2)J(t))\mathrm{d}t.$$

With this differential equation, it is easy again to use Theorem 2.2.1 to conclude the next Corollary:

Corollary 2.2.8. *Let X be an $n \times n$ matrix-valued process satisfying the following stochastic differential equation*

$$\mathrm{d}X(t) = \sqrt{X(t)}\mathrm{d}B(t)\sqrt{I_n - X(t)} + \sqrt{I_n - X(t)}\mathrm{d}B(t)\sqrt{X(t)} + (n_2 I_n - (n_1 + n_2)X(t))\mathrm{d}t.$$

Then its eigenvalues are the unique strong solution in $[0, \infty)$ to the system of stochastic differential equations

$$\mathrm{d}\lambda_i = 2\sqrt{\lambda_i(1-\lambda_i)}\mathrm{d}W_i + \left(n_2 - (n_1 + n_2)\lambda_i + \sum_{k \neq i} \frac{\lambda_i(1-\lambda_k) + \lambda_k(1-\lambda_i)}{\lambda_i - \lambda_k} \right) \mathrm{d}t, \quad (2.22)$$

where $\{W_i\}_{i=1}^n$ are n independent standard Brownian motions.

Proof. Let $g(x) = \sqrt{|x|}, h(x) = \sqrt{|1-x|}$ and $b(x) = n_2 - (n_1+n_2)x$ in 2.2.1. The uniqueness of the solution comes from Corollary 2.2.6. □

2.3 Path simulations

A standard technique for the simulation of solutions to stochastic differential equations is the Euler-Maruyama method. Details about the method can be found in [24][Chapter 10]. The code used for the following figures can be found in the Appendix A.

These path simulations have the purpose of visualizing the behavior of the eigenvalue process and how the presence of the Brownian motion term affects the trajectories. Comparing with the simulations of the deterministic processes in Chapter 4 will allow us to appreciate how much the finite-variation part of the process "determines" its evolution.

In Figure 2.1, we can see the path of a nine-dimensional Dyson Brownian motion compared to the finite-variation term. The color of each line represents the correspondence between the drift and the stochastic visualization. We can notice that both the random and the deterministic version do not collide. In the deterministic version, the values separate over time, while in the stochastic one, they ar affected by a noise that causes them to deviate slightly.

In Figure 2.2 there are four different path simulations for a Dyson Brownian motion with dimensions equal to four, five, seven, and nine. The initial condition in the third last cases is with n equally spaced points around zero. In the first case, the initial condition is slightly different because the two points in the middle have a separation of size one, while the surrounding points are separated by 0.1 of them. This appears to have little effect on this system, but in Chapter 4 we will see that in the deterministic case, it has an effect.

Although the distances between paths in every simulation of Figure 2.2 are stochastic, we can notice that the spacing between consecutive points is more or less uniform, which is what we would expect as the particles reject each other with similar strength. Also, it is interesting to notice that, although the number of paths increases in every chart of the figure, the total path spacing until $t = 10$ is more or less uniform in all of them.

The "uniform spacing" behavior of the Dyson process contrasts with the path simulations for the Wishart process in Figure 2.3. In this figure, we have the path for a Wishart process

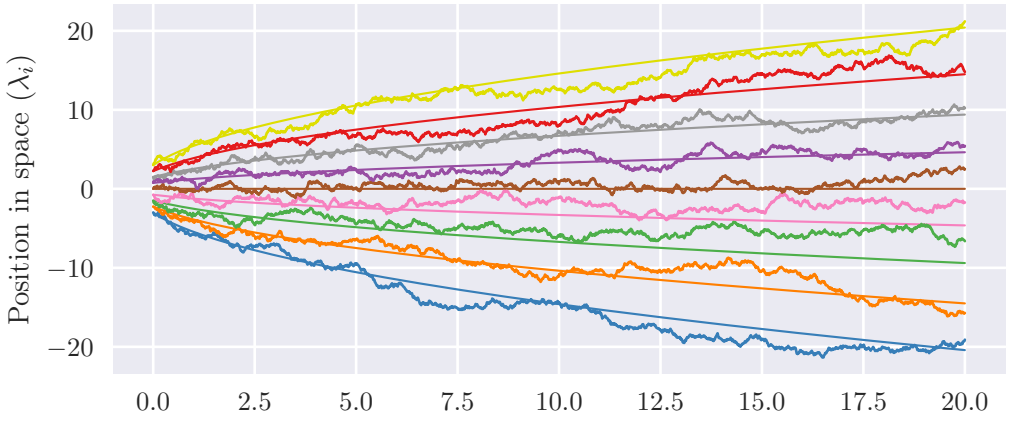

Figure 2.1: Superposition of simulation of a Dyson Brownian motion and its finite variation part. The initial conditions are $(-3, -2.25, -1.5, -0.75, 0, 0.75, 1.5, 2.25, 3)$.

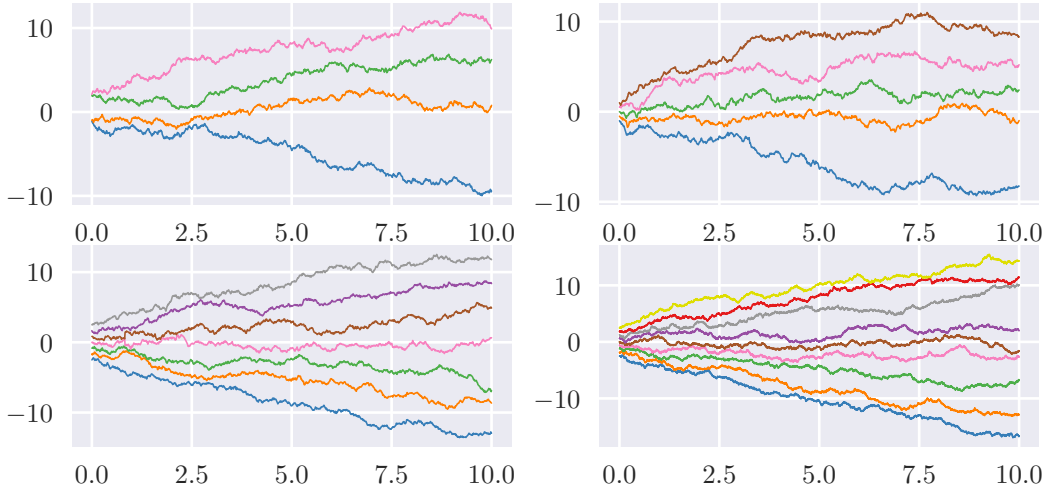

Figure 2.2: Simulation of four different Dyson Brownian motions with different dimensions. The initial conditions from left to right and top tp bottom are $(-1.1, -1, 2, 2.1)$, $(-1, -0.5, 0, 0.5, 1)$, $(-2.5, -1.67, -0.83, 0, 0.83, 1.67, 2.5)$ and $(-2.5, -1.875, -1.25, -0.625, 0, 0.625, 1.25, 1.875, 2.5)$.

of dimension nine with the corresponding finite variation path. As we can see, the spacing of the particles with the biggest values tends to grow faster. This can be explained because, for the Wishart process, the interaction term is proportional to the value of both functions.

For the Jacobi process path simulations in Figure 2.4, we notice a more stable behavior. In Chapter 4 we will see that this process has a stationary distribution and thus it is expected that, given any initial condition, the process converges to it. Notice that in this simulation the convergence appears to have been rather quick, as the points seem to be roughly evenly separated at time $t = 0.05$. The stationary behavior can be more evident with the path simulations of the deterministic version in Chapter 4.

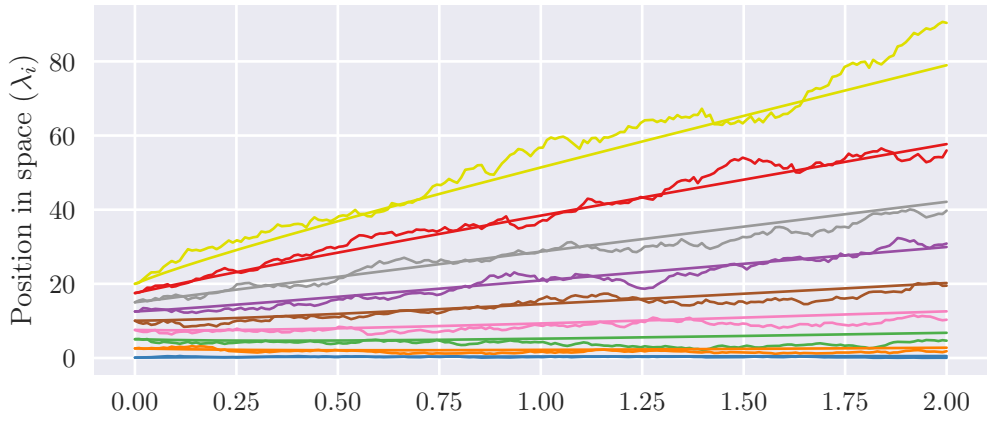

Figure 2.3: Superposition of path simulations of a Wishart process and its finite variation part. The initial conditions are $(0.1, 2.6, 5.1, 7.6, 10.1, 12.6, 15.1, 17.6, 20.1)$.

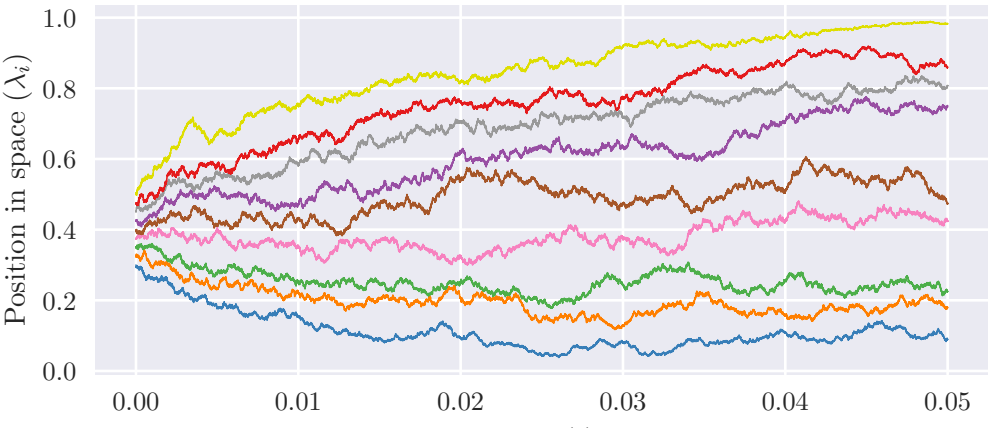

Figure 2.4: Path simulation for a Jacobi process. The initial conditions are $(0.3, 0.325, 0.35, 0.375, 0.4, 0.425, 0.45, 0.475, 0.5)$.

Chapter 3

Finite Free Probability

This chapter introduces the main ideas in finite free probability theory. In the first section, we give the definition of three polynomial convolutions and some classic orthogonal polynomial ensembles. This section is purely algebraic and it is not related to probability. In section 2, we present the definition of minor orthogonality, one of the central ideas in finite free probability. This is also the section in which we state the initial relations between random matrices and polynomial convolutions. In the third section, we link finite free probability to free probability by defining the $\mathcal{R}_d$ transform, then we use the precedent contents to find polynomials related to some important distributions in finite free probability and the corresponding limit theorems. The material in this chapter comes mainly from [3], [25] and [14].

3.1 Polynomials and convolution

3.1.1 Convolution of polynomials

In this subsection, we present three notions of polynomial convolution, the first two were introduced around a century ago, [26], [27]. Their study began using tools outside of probability theory, but we do not include any of those results here, instead, we are merely interested in introducing them to relate them to expected characteristic polynomials of random matrices. In the next subsection we introduce three ensembles of orthogonal polynomials, namely the Hermite, Laguerre and Jacobi polynomials. We prove some nice properties of these polynomials, especially related to the notions of convolution introduced previously. The third notion of convolution was presented in the context of Finite Free Probability Theory [3] as it was found to share similar properties to the other two, linking it to Random Matrix Theory.

The three notions of convolution are defined for complex polynomials. Although one could get the convolution between any two polynomials, we are interested uniquely in monic polynomials as our main object of interest are the roots and their behavior under convolution. In full generality, we can write a monic complex polynomial $p(z)$ as

$$p(z) = \sum_{j=0}^{n} z^{n-j}(-1)^j a_j.$$

The three notions of convolution are defined in function of the polynomial degree, but the polynomials need not to have the same degree. In the case the degree is different, we can take the convolution with the highest degree, and notice that this equivalent to have zero

coefficients for higher powers of the polynomial with the minor degree.

Symmetric additive convolution

This notion of convolution will be the most used along the text. Several of the results we find for the relationship of this convolution with random matrix theory can be extrapolated to the other notions, but the space and time required would be longer than the dedicated to the present work.

Definition 3.1.1 (Symmetric additive convolution)**.** Let $p(z), q(z)$ be two complex polynomials of z, with degree less or equal to d,

$$p(z) = \sum_{j=0}^{n} z^{n-j}(-1)^j a_j,$$
$$q(z) = \sum_{j=0}^{n} z^{n-j}(-1)^j b_j.$$

The nth symmetric additive convolution of p and q is

$$\begin{aligned} p(z) \boxplus_n q(z) &:= \sum_{k=0}^{n} z^{n-k}(-1)^k \sum_{i+j=k} \frac{(n-i)!(n-j)!}{n!(n-k)!} a_i b_j, \\ &= \frac{1}{n!} \sum_{k=0}^{n} \partial_z^k p(z) \partial_z^{n-k} q(0), \\ &= \frac{1}{n!} \sum_{k=0}^{n} \partial_z^k q(z) \partial_z^{n-k} p(0), \end{aligned}$$

with ∂_z denoting the differentiation with respect to z.

It is straightforward from the definition to prove that the symmetric additive convolution is linear: Let p, q, r be degree n polynomials with a_i, b_i, c_i their respective coefficients and $\alpha \in \mathbb{R}$, then

$$\begin{aligned} p \boxplus_n (\alpha q + r) &= \sum_{k=0}^{n} z^{n-k}(-1)^k \sum_{i+j=k} \frac{(n-i)!(n-j)!}{n!(n-k)!} a_i(\alpha b_j + c_j), \\ &= \alpha \sum_{k=0}^{n} z^{n-k}(-1)^k \sum_{i+j=k} \frac{(n-i)!(n-j)!}{n!(n-k)!} a_i b_j \\ &\quad + \sum_{k=0}^{n} z^{n-k}(-1)^k \sum_{i+j=k} \frac{(n-i)!(n-j)!}{n!(n-k)!} a_i c_j, \\ &= \alpha(p \boxplus_n q) + (p \boxplus_n r). \end{aligned}$$

Theorem 3.1.1. *Let $P(\partial_z), Q(\partial_z)$ be linear differential operators and p, q be polynomials of degree at most n such that*

$$p(z) = P(\partial_z)[z^n], \qquad q(z) = Q(\partial_z)[z^n].$$

Then $p(z) \boxplus_n q(z) = P(\partial_z)Q(\partial_z)[z^n]$.

Proof. We prove first that $z^n \boxplus z^n = z^n$. By definition of the convolution,

$$\begin{aligned} z^n \boxplus_n z^n &= \sum_{k=0}^{n} z^{n-k}(-1)^k \sum_{i+j=k} \frac{(n-i)!(n-j)!}{n!(n-k)!} a_i b_j, \\ &= z^{n-0}(-1)^0 \left(\frac{(n-0)!(n-0)!}{n!(n-0)!} \right) = z^n. \end{aligned}$$

Now, we prove the multiplicative property of P and Q,

$$(P(\partial_z)[r(z)] \boxplus_n Q(\partial_z)[s(z)]) = P(\partial_z)Q(\partial_z)(r(z) \boxplus s(z)).$$

For this proof, let us do it for operators of the form $P(\partial_z)[z^n] = z^{n-k}$ and then use the linearity of the symmetric additive convolution to extend linearly to polynomials of such operators. For $j + k \leq n$, take $P(\partial_z) \coloneqq \frac{(n-k)!}{n!}\partial_z^k$ and $Q(\partial_z) \coloneqq \frac{(n-j)!}{n!}\partial_z^j$ and notice

$$\begin{aligned} P(\partial_z)[z^n] &= \frac{(n-k)!}{n!}\partial_z^k[z^n] = z^{n-k}, \\ Q(\partial_z)[z^n] &= \frac{(n-j)!}{n!}\partial_z^j[z^n] = z^{n-j}. \end{aligned}$$

The symmetric additive convolution of these polynomials is

$$\begin{aligned} P(\partial_z)[z^n] \boxplus_n Q(\partial_z)[z^n] &= z^{n-k} \boxplus_n z^{n-j}, \\ &= \frac{(n-k)!(n-j)!}{n!n!} \frac{n!}{(n-k-j)!} z^{n-k-j} = \frac{(n-k)!(n-j)!}{n!(n-k-j)!} z^{n-k-j}. \end{aligned}$$

On the other hand, the product of the linear operators is

$$\begin{aligned} P(\partial_z)Q(\partial_z)[z^n] &= \left(\frac{(n-k)!}{n!}\partial_z^k \frac{(n-j)!}{n!}\partial_z^j \right)[z^n] = \frac{(n-k)!(n-j)!}{n!n!}\partial_z^{j+k}[z^n], \\ &= \frac{(n-k)!(n-j)!}{n!n!} \frac{n!}{(n-k-j)!} z^{n-k-j} = \frac{(n-k)!(n-j)!}{n!(n-k-j)!} z^{n-k-j}. \end{aligned}$$

To conclude, we extend this property linearly to any P, Q polynomials on ∂_z. □

Symmetric multiplicative convolution

Later in the thesis it is explained how the symmetric additive convolution is related to the additive convolution in Free Probability, i.e. the convolution of two measures when you add two freely independent random variables. In a similar fashion, when you multiply two freely independent random variables, there is a way to find the law of this product in terms of the

laws o the original random variables. This operation between two probability distribution is called "free multiplicative convolution". The symmetric multiplicative convolution is analogously related to the free multiplicative convolution. Although most of the results we will show are related to linking the symmetric additive convolution to sums of random variables, we also show a theorem linking products of random matrices to symmetric multiplicative convolution and several of the later additive results can be extended to the multiplicative case (see [28], [3]).

Definition 3.1.2 (Symmetric multiplicative convolution)**.** Let p and q be as in Definition 3.1.1 with degree at most n. The nth symmetric multiplicative convolution of p and q is

$$p(z) \boxtimes_n q(z) \coloneqq \sum_{i=0}^{n} z^{n-i}(-1)^i \frac{a_i b_i}{\binom{n}{i}}.$$

Asymmetric additive convolution

This notion of convolution was found for the first time by Marcus, Spielman and Srivastava in the seminal paper on Finite Free Probability Theory [3]. Although this convolution is not related to a notion of Free convolution of measures, we introduce it because it appears naturally in the results found in Chapter 4.

Definition 3.1.3 (Asymetric additive convolution)**.** Let p and q be two polynomials of degree at most n as in Definition 3.1.1. The nth asymmetric additive convolution of q and p is defined as

$$p(z) \boxplus\!\!\!\!\boxplus q(z) \coloneqq \sum_{k=0}^{n} z^{n-k}(-1)^k \sum_{j=0}^{k} \left(\frac{(n-j)!(n-k+j)!}{n!(n-k)!} \right)^2 a_j b_{k-j}.$$

Although many properties of these convolutions can be found without using Finite Free Probability Theory, we are mainly interested in describing their relationship to random matrices, so we will not include them. In the next subsection we present three ensembles of orthogonal polynomials and find that two of them have good properties related to the convolutions described above. The polynomials are also associated to Random Matrix Theory, as it is shown in later sections.

3.1.2 Some classical orthogonal polynomial ensembles

In order to define a property of "orthogonality" between elements of a given space, it is required to have some notion of inner product. If we are specifically working with spaces of square integrable functions with respect to some measure $L^2(\mu)$, the inner product between $f, g \in L^2(\mu)$ is given by

$$\langle f, g \rangle = \int_{\Omega} f(x) g(x) \mathrm{d}\mu(x),$$

where Ω is the space where the measure μ is defined.

It is clear that a set of orthogonal polynomials does not exist for every $L^2(\mu)$ space that we take. For example, if we take μ to be the Lebesgue measure on $\mathbb{R}$, then polynomic functions

are not integrable and thus the class of polynomials in $L^2(\mu)$ is empty. If we restrict only to probability measures associated to random variables in $L^{\infty-}(\Omega, \mathbb{P}, \mathbb{F})$, then every polynomial function is in square integrable.

The following is the precise definition of a set of orthogonal functions.

Definition 3.1.4. Let μ be a measure in $\mathbb{R}^n$, $I \subseteq \mathbb{N}$ be a set of integer indices and $(f_i(x))_{i\in I}$ be a collection of functions indexed by I. We say that the functions $(f_i)_{i\in I}$ are a family of orthogonal functions, if they satisfy the relationship,

$$\int_{\mathbb{R}^d} f_j(x) f_k(x) \mathrm{d}\mu(x) = \|f_j\|_2^2 \, \delta_{jk}.$$

When the functions $(f_i)_{i\in I}$ are polynomials, we call the set an *ensemble of orthogonal polynomials*.

It is important to notice that if a set of polynomials is orthogonal under any finite measure, then it will be orthogonal under any rescaling of the same measure. This leads to different definitions of famous polynomial ensembles literature. Some definitions are more common than others, however, we will give here the definitions that allow us to reduce our use of notation. What we call here "Laguerre polynomials" of "Jacobi polynomials" are rescaled versions of the most common definitions. For a classical text on these polynomials, see [27]. For a more thorough study of orthogonal polynomials related to the theory of stochastic processes, see [29].

Hermite polynomials

There are essentially two definitions of the Hermite polynomials, the first one is mostly used in physics [30], and the second one is related to Probability Theory as appearing in [28]. The main difference between them is the measure under which they are orthogonal. The "physicist Hermite polynomials" are orthogonal under the measure $e^{-x^2}\mathrm{d}x$, the heat kernel. The "probabilist Hermite polynomials" are orthogonal under $e^{-\frac{x^2}{2}}\mathrm{d}x$, the Gaussian kernel. We will restrict to the probabilist Hermite polynomials and will use the term "Hermite polynomials" to talk about them.

The nth Hermite polynomial, denoted by $H_n(z)$ is defined by a linear differential operator applied to z^n,

$$H_n(z) \coloneqq e^{-\frac{\partial_z^2}{2}}[z^n] \coloneqq \sum_{k=0}^{\infty} \frac{(-1)^k}{2^k k!} \partial^{2k}[z^n]. \tag{3.1}$$

Both the physicist and the probabilist Hermite polynomials have a generalization to bivariate polynomials on z and t. In the case of the probabilist Hermite polynomials, this generalization has the nice interpretation as being the polynomials orthogonal under the measure $e^{-t\frac{x^2}{2}}\mathrm{d}x$, i.e. the measure of a Gaussian random variable with variance t. Notice that this implies that the physicist Hermite polynomials are the probabilist Hermite polynomials with variance 2. The generalized Hermite polynomials also known as time dependent Hermite polynomials, or Hermite polynomials with variance are polynomials on z and t generated by the analogous linear operator

$$H_n(z,t) \coloneqq e^{-\frac{t\partial_z^2}{2}}(z^n) \coloneqq \sum_{k=0}^{\infty} \frac{(-1)^k t^k}{2^k k!} \partial^{2k}[z^n]. \tag{3.2}$$

Notice that $H_n(z,1) = H_n(z)$. Using the former definitions we can find explicit expressions for both $H_n(z)$ and $H_n(z,t)$,

$$H_n(z) = \sum_{k=0}^{\infty} \frac{(-1)^k}{2^k k!} \partial^{2k}[z^n] = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \frac{(-1)^k}{2^k k!} \frac{n!}{(n-2k)!} z^{n-2k} = n! \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \frac{(-1)^k z^{n-2k}}{2^k k!(n-2k)!},$$
$$H_n(z,t) = \sum_{k=0}^{\infty} \frac{(-t)^k}{2^k k!} \partial^{2k}[z^n] = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \frac{(-t)^k}{2^k k!} \frac{n!}{(n-2k)!} z^{n-2k} = n! \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \frac{(-t)^k z^{n-2k}}{2^k k!(n-2k)!}.$$

An easy substitution allows to see that the coefficient of z^m in $H_n(z,t)$ is

$$a_m = \begin{cases} \frac{n!(-t)^{\frac{n-m}{2}}}{2^{\frac{n-m}{2}}(\frac{n-m}{2})!m!}, & \text{if } m \text{ and } n \text{ have the same parity,} \\ 0, & \text{otherwise.} \end{cases} \tag{3.3}$$

The last expression gives us a way to find the first few Hermite polynomials,

$$\begin{aligned}
H_1(z,t) &= z, \\
H_2(z,t) &= z^2 - t, \\
H_3(z,t) &= z^3 - 3tz, \\
H_4(z,t) &= z^4 - 6tz^2 + 3t^2, \\
H_5(z,t) &= z^5 - 10tz^3 + 15t^2 z, \\
H_6(z,t) &= z^6 - 15tz^4 + 45t^2 z^2 - 15t^3, \\
H_7(z,t) &= z^7 - 21tz^5 + 105t^2 z^3 - 105t^3 z, \\
H_8(z,t) &= z^8 - 28tz^6 + 210t^2 z^4 - 420t^3 z^2 + 105t^4, \\
H_9(z,t) &= z^9 - 36tz^7 + 378t^2 z^5 - 1260t^3 z^3 + 945t^4 z, \\
H_{10}(z,t) &= z^{10} - 45tz^8 + 630t^2 z^6 - 3150t^3 z^4 + 4725t^4 z^2 - 945t^5.
\end{aligned}$$

Replacing $t = 1$, we can find the corresponding standard Hermite polynomials.

The Hermite polynomials are characterized by the following recursion together with the initial conditions $H_1(x,t)$ and $H_2(x,t)$.

$$H_n(x,t) = xH_{n-1}(x,t) - t(n-1)H_{n-2}(x,t). \tag{3.4}$$

The next proposition shows us that the Hermite polynomials with variance are well-behaved with respect to the symmetric additive convolution. This result will be more obvious once we have developed the tools of Finite Free Probability Theory and we will be able to give an easier proof after Section 3.2.

Proposition 3.1.2. *The symmetric additive convolution between two Hermite polynomials with the same order $H_n(z,t_1), H_n(z,t_2)$ is another Hermite polynomial with variance $t_1 + t_2$.*

Proof. We proceed directly by definition of the symmetric additive convolution and the closed form for the polynomials.

$$
\begin{aligned}
&H_n(z,t_1) \boxplus_n H_n(z,t_2) = \\
&= \sum_{k=0}^{n} z^{n-k}(-1)^k \sum_{i=0}^{k} \frac{(n-i)!(n-k+i)!}{n!(n-k)!} b_i a_{k-i}, \\
&= \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} z^{n-2k} \sum_{i=0}^{2k} \frac{(n-i)!(n-2k+i)!}{n!(d-2k)!} \frac{n!(-t_1)^{i/2}}{2^{i/2}(i/2)!(n-i)!} \frac{n!(-t_2)^{2k-i}}{2^{k-i/2}(k-i/2)!(n-2k+i)!}, \\
&= \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} z^{n-2k} \frac{n!}{2^k} \sum_{i=0}^{2k} \frac{(-t_1)^{i/2}(-t_2)^{k-i/2}}{(i/2)!(k-i/2)!} = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} z^{n-2k} \frac{n!}{k!2^k} \sum_{i=0}^{2k} \frac{k!(-t_1)^{i/2}(-t_2)^{k-i/2}}{(i/2)!(k-i/2)!}, \\
&= \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} z^{n-2k} \frac{n!}{k!2^k} \sum_{i=0}^{k} \binom{k}{i} (-t_1)^{i/2}(-t_2)^{k-i/2} = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} z^{n-2k}(-1)^k \frac{n!(t_1+t_2)^k}{k!2^k}, \\
&=: H_n(z, t_1+t_2).
\end{aligned}
$$

□

Laguerre polynomials

As mentioned before, the ensembles of orthogonal polynomials we work with are not always monic, but we are interested on monic versions of them, so the definitions given here are proper re-scalings. The Laguerre polynomials are the set of orthogonal polynomials under the measure $e^{-x}x^{\alpha}\mathrm{d}x$. This is a scaling of the measure associated to a chi-squared distribution with $2(\alpha+1)$ degrees of freedom. When $\alpha = 0$, it corresponds to the polynomials orthogonal under the law of a chi-squared random variable with 2 degrees of freedom. We will call this particular case the "standard Laguerre polynomials", and when $\alpha > 0$ we will call them the "generalized Laguerre polynomials".

The Laguerre polynomials are usually defined as generated by the linear operator

$$\frac{1}{n!}\left(\partial_z - 1\right)[z^n]. \tag{3.5}$$

In order to get the monic polynomials proportional to the ones generated to the expression in (3.5), we need to cancel the quotient term $n!$. Adopting this convention, the standard Laguerre polynomials are defined by the linear differential operator

$$L_n(z) \coloneqq (1-\partial_z)^n[z^n] \coloneqq \sum_{k=0}^{n} \binom{n}{k}(-1)^k \partial_z^k[z^n] = \sum_{k=0}^{n} \binom{n}{k}(-1)^k \frac{n!}{(n-k)!} z^{n-k},$$

while the generalized Laguerre polynomials are defined by the differential operator

$$L_n^{\alpha}(z) \coloneqq x^{-\alpha}\left(1-\partial_z\right)^n\left[x^{n+\alpha}\right] \coloneqq \sum_{k=0}^{n}(-1)^k(n+\alpha)_{n+\alpha-k}x^{n-k},$$

where $(n+\alpha)_{n+\alpha-k}$ denotes the Pochhammer symbol.

Notice that this is equivalent to the definition by the linear differential operator

$$L_n^{\alpha}(z) \coloneqq (1-\partial_z)^{n+\alpha}\,[x^n] = \sum_{k=0}^{n} \binom{n}{k} (-1)^k (n+\alpha)_{n+\alpha-k} x^{n-k}.$$

Just as in the case of the Hermite polynomials, we can generalize to bi-variate versions which corresponds to a chi-squared law with variance t^2. The time dependent Laguerre polynomials also known as Laguerre polynomials with variance are polynomials on z and t defined by the analogous linear operator

$$L_n(z,t) \coloneqq (1-t\partial_z)^n[z^n] \coloneqq \sum_{k=0}^{n} \binom{n}{k} (-t)^k \partial_z^k[z^n] = \sum_{k=0}^{n} \binom{n}{k} (-t)^k \frac{n!}{(n-k)!} z^{n-k}. \tag{3.6}$$

Likewise, the time dependent generalized Laguerre polynomials are defined by the operator one would expect

$$L_n^{\alpha}(z,t) \coloneqq (1-t\partial_z)^{n+\alpha}\,[x^n] \coloneqq \sum_{k=0}^{n} \binom{n}{k} (-t)^k (n+\alpha)_{n+\alpha-k} x^{n-k}.$$

Expanding the definition, we can list the first few time-dependent standard Laguerre polynomials,

$$\begin{aligned}
L_1(z,t) &= z - t,\\
L_2(z,t) &= z^2 - 4tz + 2t^2,\\
L_3(z,t) &= z^3 - 9tz^2 - 18t^2 z + 6t^3,\\
L_4(z,t) &= z^4 - 16tz^3 + 72t^2z^2 - 96t^3 z + 24t^4,\\
L_5(z,t) &= z^5 - 25tz^4 + 200t^2z^3 - 600t^3z^2 + 600t^4 z - 120t^5,\\
L_6(z,t) &= z^6 - 36tz^5 + 450t^2z^4 - 2400t^3z^3 + 5400t^4z^2 - 4320t^5 z + 720t^6,\\
L_7(z,t) &= z^7 - 49tz^6 + 882t^2z^5 - 7350t^3z^4 + 29400t^4z^3 - 52920t^5z^2 + 35280t^6 z - 5040t^7,\\
L_8(z,t) &= z^8 - 64tz^7 + 1568t^2z^6 - 18816t^3z^5 + 117600t^4z^4 - 376320t^5z^3 + 564480t^6z^2\\
&\quad - 322560t^7 z + 40320t^8,\\
L_9(z,t) &= z^9 - 81tz^8 + 18144t^2z^7 - 42336t^3z^6 + 381024t^4z^5 - 1905120t^5z^4 + 5080320t^6z^3\\
&\quad - 6531840t^7z^2 + 3265920t^8 z - 362880t^9.
\end{aligned}$$

Replacing $t = 1$, we get the standard Laguerre polynomials.

The Laguerre polynomials are characterized by the following recursion together with the first two polynomials $L_1(z) = z - 1$ and $L_2(z) = z^2 - 4tz + 2t^2$,

$$L_n(z) = (z - 2n + 1)L_{n-1}(z) - (n-1)^2 L_{n-2}(z).$$

Jacobi polynomials

The Jacobi polynomials depend on two parameters α, β and they are orthogonal under the measure $(1-z)^{\alpha}(1+z)^{\beta}\mathrm{d}z$. The standard Jacobi polynomials are defined by the following differential operator

$$P_n^{(\alpha,\beta)}(z) = \frac{(-1)^n}{2^n n!}(1-z)^{-\alpha}(1+z)^{-\beta}\partial_z^n\left\{(1-z)^\alpha(1+z)^\beta\left(1-z^2\right)^n\right\}.$$

The parameters α, β take values in $\mathbb{R}^+$. In the particular case when α and β are positive integers, we have

$$P_n^{(\alpha,\beta)}(x) = \sum_{k=0}^{n}\binom{n+\alpha}{n-k}\binom{n+\beta}{k}\left(\frac{x-1}{2}\right)^k\left(\frac{x+1}{2}\right)^{n-k} \tag{3.7}$$

If $\alpha = \beta = 0$, then we have the Legendre polynomials defined by the differential operator

$$P_n(z) = \frac{1}{2^n n!}\partial_z^n[(z^2-1)^n].$$

Once again, we are interested in a monic scaling of these polynomials, but now we also want a "transaltion of them". The original Jacobi (and Legendre) polynomials are defined in $[-1,1]$, but we are interested in polynomials that take values in $[0,1]$. We get these polynmials by evaluating a Jacobi polynomial in $1-2z$, so that they are orthogonal under the measure $\mathbb{1}_{(0,1)}(z)z^\alpha(1-z)^\beta \mathrm{d}z$, i.e. the measure of a beta distribution. In the particular case when $\alpha = \beta = 0$, they are orthogonal under the measure of an uniform random variable in $(0,1)$. The closed form for the Legendre polynomials evaluated in $1-2z$ and scaled to be monic is

$$P_n(z) = \frac{n!n!}{(2n)!}\sum_{k=0}^{n} z^k(-1)^{n-k}\binom{n}{k}\binom{n+k}{k}.$$

The first few Legendre polynomials are

$$\begin{aligned}
P_1(z) &= z,\\
P_2(z) &= z^2 - \frac{z}{2} + \frac{1}{6},\\
P_3(z) &= z^3 - \frac{2z^2}{5} + \frac{z}{2} - \frac{1}{10},\\
P_4(z) &= z^4 - \frac{z^3}{2} + \frac{3z^2}{7} - \frac{5z}{28} + \frac{1}{35},\\
P_5(z) &= z^5 - \frac{7z^4}{10} + \frac{7z^3}{3} - \frac{14z^2}{15} + \frac{7z}{30} - \frac{1}{42},\\
P_6(z) &= z^6 - \frac{z^5}{2} + \frac{2z^4}{3} - \frac{4z^3}{7} + \frac{5z^2}{21} - \frac{z}{42} + \frac{1}{231},\\
P_7(z) &= z^7 - \frac{z^6}{2} + \frac{5z^5}{7} - \frac{2z^4}{3} + \frac{10z^3}{21} - \frac{z^2}{14} + \frac{z}{42} - \frac{1}{429},\\
P_8(z) &= z^8 - \frac{5z^7}{8} + \frac{5z^6}{4} - \frac{5z^5}{6} + \frac{5z^4}{14} - \frac{z^3}{28} + \frac{z^2}{56} - \frac{z}{56} + \frac{1}{6435},\\
P_9(z) &= z^9 - \frac{3z^8}{4} + \frac{z^7}{2} - \frac{z^6}{2} + \frac{3z^5}{7} - \frac{2z^4}{21} + \frac{z^3}{28} - \frac{z^2}{56} + \frac{z}{315} - \frac{1}{12155},\\
P_{10}(z) &= z^{10} - \frac{z^9}{2} + \frac{3z^8}{5} - \frac{3z^7}{5} + \frac{7z^6}{15} - \frac{z^5}{15} + \frac{z^4}{35} - \frac{z^3}{70} + \frac{z^2}{231} - \frac{z}{11547} + \frac{1}{46189}.
\end{aligned}$$

Just as in the previous case, we will use the name "Legendre polynomials" or "Jacobi

polynomials" to talk about these monic scalings.

The Jacobi polynomials are a generalization of several other polynomial ensembles and they are closely related to several topics in the theory of orthogonal polynomials. To read more about them, we recomend to revise [27, Chapter IV].

3.2 Expected characteristic polynomials

In this section we show the relationship between the polynomial convolutions defined before and Random Matrix Theory, specifically through the expected characteristic polynomial. Although several results can be found, we will restrict to the symmetric additive case and its relationship to sums of random matrices. The main result in this section is Theorem 3.2.8, a similar theorem can also be stated for the product of matrices and symmetric multiplicative convolution. The first subsection introduces minor orthogonality, which is a key concept to prove the following results. In the second subsection we prove the main result.

3.2.1 Minor orthogonality

The next is a rather technical definition, but that allows to prove several of our results of interest.

Definition 3.2.1 (Minor orthogonality)**.** Let R be an $m \times n$ random matrix. We say R is minor orthogonal if for every $k, l \in \mathbb{Z}$ such that $k, l \leq \max\{m, n\}$ and all sets S, T, U, V with $|S| = |T| = k$ and $|U| = |V| = l$, it satisfies

$$\mathbb{E}_R \left[\det[R]_{S,T} \det[R^*]_{U,V}\right] = \frac{\delta_{S,V}\delta_{T,U}}{\binom{\max\{m,n\}}{k}}.$$

In the last definition, $\mathbb{E}_R$ denotes taking the expectation with respect to R. We use this convention because in what follows we will have computations of expectations with several random matrices involved and it is convenient to know what is the matrix we are taking the expectation of in every step.

Later we will prove that some well-known matrix ensembles are minor orthogonal, and that property will help us to make use of the related results. Before doing that, we will prove a lemma that allows us to conclude that some transformations of minor orthogonal matrices are also minor orthogonal.

Lemma 3.2.1. *If R is minor orthogonal and Q is a constant matrix such that $QQ^* = I$, then Q is minor orthogonal. If $Q^*Q = I$, then RQ is minor orthogonal.*

Proof. Recall that by the Cauchy-Binet formula (Theorem 1.1.1), for $|S| = |T| = k$ we have

$$\det[QR]_{S,T} = \sum_{|W|=k} \det[Q]_{S,W} \det[R]_{W,T},$$

so with $|S| = |T| = k, |U| = |V| = l$,

$$\begin{aligned}
\mathbb{E}_R\left[\det[QR]_{S,T}\det[R^*Q^*]_{U,V}\right] &= \mathbb{E}_R\left[\sum_{|W|=k}\sum_{|Z|=l}\det[Q]_{S,W}\det[R]_{W,T}\det[R^*]_{U,Z}\det[Q^*]_{Z,V}\right], \\
&= \sum_{|W|=k}\sum_{|Z|=l}\det[Q]_{S,W}\det[Q^*]_{Z,V}\mathbb{E}_R\left[\det[R]_{W,T}\det[R^*]_{U,Z}\right], \\
&= \sum_{|W|=k}\sum_{|Z|=l}\det[Q]_{S,W}\det[Q^*]_{Z,V}\mathbb{E}_R\left[\frac{\delta_{W,Z}\delta_{T,U}}{\binom{\max\{m,n\}}{k}}\right], \\
&= \sum_{|W|=k}\det[Q]_{S,W}\det[Q^*]_{W,V}\frac{\delta_{T,U}}{\binom{\max\{m,n\}}{k}}, \\
&= \det[QQ^*]_{S,V}\frac{\delta_{T,U}}{\binom{\max\{m,n\}}{k}}, \\
&= \det[I]_{S,V}\frac{\delta_{T,U}}{\binom{\max\{m,n\}}{k}},
\end{aligned}$$

Notice that $\det[I]_{S,V} = 1$ if and only if $S = V$, so we conclude that

$$\mathbb{E}_R\left[\det[QR]_{S,T}\det[R^*Q^*]_{U,V}\right] = \frac{\delta_{S,V}\delta_{T,U}}{\binom{\max\{m,n\}}{k}}.$$

The case $Q^*Q = I$ is proven in the same way. □

The former theorem implies that minor orthogonality is preserved under unitary transformations. Now, we will introduce the signed permutation matrix ensemble and prove that it is minor orthogonal. This ensemble will be later used to conclude minor orthogonality in a broader class of random matrices.

Definition 3.2.2 (Signed permutation matrix)**.** A signed permutation matrix is a matrix that can be written EP where E is a diagonal matrix with entries ± 1 and P is a permutation matrix.

Lemma 3.2.2. *A random matrix sampled uniformly from the set of signed permutation matrices is minor-orthogonal.*

Proof. Let Q be a signed permutation matrix, we can write $Q = EP$, where E is a diagonal random matrix with entries ± 1 taken uniformly and P is a matrix chosen uniformly from the permutation matrices, and both are independent. Then for $|S| = |T| = k$ and $|U| = |V| = l$, we have

$$\begin{aligned}
\mathbb{E}_Q\left[\det[Q]_{S,T}\det[Q^*]_{U,V}\right] &= \mathbb{E}_{E,P}\left[\det[EP]_{S,T}\det[P^*E]_{U,V}\right], \\
&= \sum_{|W|=k}\sum_{|Z|=l}\mathbb{E}_{E,P}\left[\det[E]_{S,W}\det[P]_{W,T}\det[P^*]_{U,Z}\det[E]_{Z,V}\right],
\end{aligned}$$

every $[E]_{S,W}$ is a diagonal or shift matrix and the determinant would be zero if $S \neq W$, so

$$\mathbb{E}_Q\left[\det[Q]_{S,T}\det[Q^*]_{U,V}\right] = \mathbb{E}_{E,P}\left[\det[E]_{S,S}\det[P]_{W,T}\det[P^*]_{U,Z}\det[E]_{V,V}\right],$$

Let $\{\chi_i\}_{1\leq i\leq n}$ be the diagonal entries of E, then

$$\mathbb{E}_Q\left[\det[Q]_{S,T}\det[Q^*]_{U,V}\right] = \mathbb{E}_E\left[\prod_{i\in S}\chi_i\prod_{j\in V}\chi_j\right]\mathbb{E}_P\left[[P]_{S,T}[P^*]_{U,V}\right].$$

Now we use that the variables χ_i are independent and uniform in $\{-1,1\}$, so that $\mathbb{E}[\chi_i]=0$, but $\mathbb{E}[\chi_i^2]=1$ for all i, and this means

$$\mathbb{E}_E\left[\prod_{i\in S}\chi_i\prod_{j\in V}\chi_j\right] = \delta_{S,V}.$$

The last equality leads to

$$\begin{aligned}\mathbb{E}_R\left[\det[QR]_{S,T}\det[R^*Q^*]_{U,V}\right] &= \delta_{S,V}\mathbb{E}_P\left[\det[P]_{S,T}\det[P^*]_{U,V}\right],\\ &= \delta_{S,V}\mathbb{E}_P\left[\det[P]_{S,T}\det[P^*]_{U,S}\right],\\ &= \delta_{S,V}\mathbb{E}_P\left[\det[P]_{S,T}\det[P]_{S,U}\right].\end{aligned}$$

The submatrix $P_{S,T}$ can be transformed in a diagonal matrix by a permutation matrix because it has at most a non zero entry for each row and each column. If the diagonal matrix has a zero entry in the diagonal, then the determinant $\det[P]_{S,T}$ is zero, in other case, it is different that zero. The only case when all of the diagonal entries of the diagonal matrix are not zero is when $T=\pi(S)$ with π the permutation function corresponding to P. This means that in order to have a non-zero determinant we need $T=\pi(S)=U$, and $\det[P]_{S,U}\in\{-1,1\}$, so

$$\begin{aligned}\mathbb{E}_R\left[\det[QR]_{S,T}\det[R^*Q^*]_{U,V}\right] &= \delta_{S,V}\delta_{T,U}\mathbb{E}_P\left[\det[P]_{S,T}\det[P]_{S,T}\right],\\ &= \delta_{S,V}\delta_{T,U}\mathbb{E}_P\left[\det[P]_{S,T}^2\right],\\ &= \delta_{S,V}\delta_{T,U}\mathbb{E}_P\left[\delta_{T=\pi(S)}\right],\\ &= \delta_{S,V}\delta_{T,U}P\left(T=\pi(S)\right).\end{aligned}$$

We are supposing that we are sampling uniformly from the permutation matrices of size $n\times n$, so the probability that $T=\pi(S)$ when π is a permutation of n elements and $|S|=|T|=k$ is $1/\binom{n}{k}$. So, we can conclude

$$\mathbb{E}_R\left[\det[QR]_{S,T}\det[R^*Q^*]_{U,V}\right] = \frac{\delta_{S,V}\delta_{T,U}}{\binom{n}{k}}.$$

This is the definition of being minor-orthogonal. □

Repeating the former proof but with E having size $m\times m$ and P being $m\times n$, we have the corollary for rectangular signed permutations.

Corollary 3.2.3. *Let E be an $m\times m$ diagonal random matrix with independent entries such that $P(E_{ii}=1)=P(E_{ii}=-1)=1/2$, and let P be an $m\times n$ random matrix taken*

uniformly from the set of $m \times n$ permutation matrices independent from E. Then EP is minor orthogonal.

Now that we know that the signed permutations ensemble is minor orthogonal, we can use this and Lemma 3.2.1 to prove that Haar unitary ensembles are minor orthogonal.

Corollary 3.2.4. *An $m \times n$ random matrix sampled from the Haar measure on $\mathcal{M}_{n,m}(\mathbb{C})$ is minor-orthogonal.*

Proof. Let R be a Haar distributed random $m \times n$ matrix with $m \leq n$ and Q a random signed permutation matrix. Any random signed permutation matrix is unitary, so RQ is Haar distributed for fixed Q, and by Lemmas 3.2.1 and 3.2.2 we have that it is also minor-orthogonal. Then, if Q is uniformly sampled from the signed permutation matrices,

$$\mathbb{E}_R \left[\det[R]_{S,T} \det[R^*]_{U,V} \right] = \mathbb{E}_{R,Q} \left[\det[RQ]_{S,T} \det[(RQ)^*]_{U,V} \right].$$

Since Q is minor orthogonal, RQ is also minor orthogonal for fixed R and

$$\mathbb{E}_R \left[\mathbb{E}_Q \left[\det[RQ]_{S,T} \det[(RQ)^*]_{U,V} \right] \right] = \mathbb{E}_R \left[\frac{\delta_{S,V}\delta_{T,U}}{\binom{n}{k}} \right] = \frac{\delta_{S,V}\delta_{T,U}}{\binom{n}{k}},$$

where $k = |S| = |T|$. □

Let us denote by $\sigma_k(A)$ the coefficient of of $(-1)^k x^{d-k}$ in the characteristic polynomial of a d-dimensional matrix A. We will use the fact that

$$\sigma_k(A) = \sum_{|S|=k} \det[A]_{S,S}.$$

The following two lemmas will help us to find explicit coefficients for expected characteristic polynomials with the aid of minor orthogonality properties.

Lemma 3.2.5. *Let $m \leq n$, B an $n \times n$ random matrix and R an $m \times n$ minor-orthogonal matrix independent from B. For all sets $S, T \subset \binom{[m]}{k}$ we have*

$$\mathbb{E}_{B,R} \left[\det[RBR^*]_{S,T} \right] = \mathbb{E}_B \left[\delta_{S,T} \frac{\sigma_k(B)}{\binom{n}{k}} \right].$$

Proof. Using the Cauchy-Binet formula we have

$$\begin{aligned}
\mathbb{E}_{B,R}\left[\det[RBR^*]_{S,T}\right] &= \mathbb{E}_B\left[\sum_{X,Y\in\binom{[n]}{k}} \mathbb{E}_R\left[\det[R]_{S,X}\det[B]_{X,Y}\det[R^*]_{Y,T}\right]\right], \\
&= \mathbb{E}_B\left[\sum_{X,Y\in\binom{[n]}{k}} \det[B]_{X,Y}\mathbb{E}_R\left[\det[R]_{S,X}\det[R^*]_{Y,T}\right]\right], \\
&= \mathbb{E}_B\left[\sum_{X,Y\in\binom{[n]}{k}} \det[B]_{X,Y}\frac{\delta_{S,T}\delta_{X,Y}}{\binom{n}{k}}\right], \\
&= \mathbb{E}_B\left[\sum_{X\in\binom{[n]}{k}} \det[B]_{X,X}\frac{\delta_{S,T}}{\binom{n}{k}}\right], \\
&= \mathbb{E}_B\left[\delta_{S,T}\frac{\sigma_k(B)}{\binom{n}{k}}\right].
\end{aligned}$$

□

Lemma 3.2.6. *Let $a > d$, A an $a \times a$ random matrix and Q a random $a \times d$ matrix sampled from the Haar measure on $\mathcal{M}_{a,d}(\mathbb{C})$, then*

$$\mathbb{E}_A\left[\mathbb{E}_Q\left[\chi_x\left(QAQ^*\right)\right]\right] = \mathbb{E}_A\left[\frac{d!}{a!}\partial_x^{(a-d)}\chi_x(A)\right],$$

where $\chi_z(\cdot)$ denotes the characteristic polynomial of $\cdot$ with z as a variable.

Proof. Let A be a fixed matrix and Q a Haar unitary matrix on $\mathcal{M}_{a,b}(\mathbb{C})$. The kth coefficient of the expected characteristic polynomial of QAQ^* is

$$\begin{aligned}
\mathbb{E}_Q\left[\sigma_k(QAQ^*)\right] &= \sum_{|S|=k}\mathbb{E}_Q\left[\det[QAQ^*]_{S,S}\right], \\
&= \sum_{|S|=k}\frac{\sigma_k(A)}{\binom{a}{k}} = \frac{\binom{d}{k}\sigma_k(A)}{\binom{a}{k}}.
\end{aligned}$$

Taking expectation with respect to A in the last expression we find

$$\mathbb{E}_A\left[\mathbb{E}_Q\left[\chi_x\left(QAQ^*\right)\right]\right] = \mathbb{E}_A\left[\frac{\binom{d}{k}\sigma_k(A)}{\binom{a}{k}}\right] = \frac{\binom{d}{k}\mathbb{E}_A[\sigma_k(A)]}{\binom{a}{k}},$$

which is the kth coefficient of $\frac{d!}{a!}\frac{\mathrm{d}^{(a-d)}}{\mathrm{d}x}\mathbb{E}_A\left[\chi_x(A)\right]$. □

With these results, now we are ready to prove the main theorems relating polynomial convolutions to the expected characteristic polynomials of sums and products of random matrices.

3.2.2 Relation to polynomial convolution

Until now, we have used minor orthogonality properties to find some expressions for expected characteristic polynomials. In this section, we will see that under suitable circumstances, those expressions are actually related to polynomial convolutions. We prove the relationship for the symmetric additive and symmetric multiplicative cases. An analogous for the asymmetric additive case can be found in [3].

For both the additive and multiplicative cases, the results are written in two parts; one for the statement of the expected characteristic polynomial coefficients and another stating explicitly the convolution of expected characteristic polynomials. The relationship with Free Probability Theory will be more evident for the additive case in the next section where we introduce the $\mathcal{R}_n$ finite transform. For the multiplicative case a similar transform can be introduced, but it is not included in this work.

Theorem 3.2.7. *Let A, B be $n \times n$ random matrices and R a $n \times n$ minor-orthogonal matrix, such that A, B, R are jointly independent, then we have*

$$\mathbb{E}_{A,B,R}\left[\sigma_k(A + RBR^*)\right] = \sum_{i=0}^{k} \frac{\binom{n-i}{k-i}}{\binom{n}{k-i}} \mathbb{E}_A\left[\sigma_i(A)\right] \mathbb{E}_A\left[\sigma_{k-i}(B)\right],$$

where $\sigma_j(X)$ represents the jth coefficient in the characteristic polynomial of the matrix X.

Proof. We use that

$$\sigma_k(A) = \sum_{|S|=k} \det[A]_{S,S},$$

together with Theorem 1.1.2 and Lemma 3.2.5 to get

$$\mathbb{E}_{A,B,R}\left[\sigma_k(A + RBR^*)\right] = \sum_{S \in \binom{[n]}{k}} \mathbb{E}_{A,B,R}\left[\det[A + RBR^*]_{S,S}\right].$$

Denote by $\overline{U}$ the complement of U, then

$$\begin{aligned}
&\mathbb{E}_{A,B,R}\left[\sigma_k(A + RBR^*)\right] \\
&\quad = \sum_{S \in \binom{[n]}{k}} \sum_{i=0}^{k} \sum_{U,V \in \binom{[k]}{i}} (-1)^{\|U\|_1 + \|V\|_1} \mathbb{E}_A\left[\det[A]_{U(S),V(S)}\right] \mathbb{E}_{B,R}\left[\det[RBR^*]_{\overline{U}(S),\overline{V}(S)}\right], \\
&\quad = \sum_{S \in \binom{[n]}{k}} \sum_{i=0}^{k} \sum_{U,V \in \binom{[k]}{i}} (-1)^{\|U\|_1 + \|V\|_1} \mathbb{E}_A\left[\det[A]_{U(S),V(S)}\right] \delta_{\overline{U}(S),\overline{V}(S)} \frac{\mathbb{E}_B\left[\sigma_{k-i}(B)\right]}{\binom{n}{k-i}}.
\end{aligned}$$

Using that $U(S) = V(S)$ if and only if $\overline{U}(S) = \overline{V}(S)$,

$$\mathbb{E}_{A,B,R}\left[\sigma_k(A + RBR^*)\right] = \sum_{i=0}^{k} \frac{\mathbb{E}_B\left[\sigma_{k-i}(B)\right]}{\binom{n}{k-i}} \sum_{S \in \binom{[n]}{k}} \sum_{U,V \in \binom{[k]}{i}} \mathbb{E}_A\left[\det[A]_{U(S),U(S)}\right].$$

To finish the proof we need to find

$$\sum_{S\in\binom{[n]}{k}}\sum_{U\in\binom{[k]}{i}}\mathbb{E}_A\left[\det[A]_{U(S),U(S)}\right]. \tag{3.8}$$

We are summing over all of the sets $V\in\binom{[n]}{i}$, but they appear more than once in the sum. To find the number of times every element $V\in\binom{[n]}{i}$ appears in the sum, we can count the total number of terms we are summing in (3.8) and divide by the total number of elements in $\binom{[n]}{i}$. We have that $|\binom{[n]}{i}|=\binom{n}{i}$ and the number of summands is $\binom{n}{k}\binom{k}{i}$, so

$$\frac{\binom{n}{k}\binom{k}{i}}{\binom{n}{i}}=\frac{\frac{n!}{k!(n-k)!}\frac{k!}{i!(k-i)!}}{\frac{n!}{i!(n-i)!}}=\frac{(n-i)!}{(n-k)!(k-i)!}=\binom{n-i}{k-i}.$$

Thus, we have

$$\sum_{S\in\binom{[n]}{k}}\sum_{U\in\binom{[k]}{i}}\mathbb{E}_A\left[\det[A]_{U(S),U(S)}\right]=\binom{n-i}{k-i}\sum_{V\in\binom{[n]}{i}}\mathbb{E}_A\left[\det[A]_{V,V}\right]=\binom{n-i}{k-i}\mathbb{E}_A\left[\sigma_i(A)\right].$$

With this, we can conclude

$$\mathbb{E}_{A,B,R}\left[\sigma_k(A+RBR^*)\right]=\sum_{i=0}^{k}\frac{\binom{n-i}{k-i}}{\binom{n}{k-i}}\mathbb{E}_A\left[\sigma_i(A)\right]\mathbb{E}_A\left[\sigma_{k-i}(B)\right].$$

□

Theorem 3.2.8. *If $p(z)$ is the characteristic polynomial of A and $q(z)$ is the characteristic polynomial of B, where A and B are $n\times n$ normal matrices with complex entries, then*

$$p(z)\boxplus_n q(z)=\mathbb{E}_Q\left[\chi_z(A+QBQ^*)\right],$$

where $\chi_z(\cdot)$ denotes the characteristic polynomial of $\cdot$ with z as a variable and $\mathbb{E}_Q$ denotes taking expectation over Q where Q is sampled from the Haar measure on the unitary complex $n\times n$ matrices.

Proof. It follows directly from Theorem 3.2.7 and definition of the symmetric additive convolution. □

The next two theorems are the analogous of the previous ones for the symmetric multiplicative convolution. As it was stated previously, these convolutions can be seen as "approximations" to the free additive and free multiplicative convolutions.

Theorem 3.2.9. *Let A and B be $n\times n$ random matrices and R a minor-orthogonal $n\times n$ matrix, such that A,B,R are jointly independent, then*

$$\mathbb{E}_{A,B,R}\left[\sigma_k(ARBR^*)\right]=\frac{\mathbb{E}_A[\sigma_k(A)]\mathbb{E}_B[\sigma_k(B)]}{\binom{n}{k}}.$$

Proof. By the Cauchy-Binet formula and independence

$$\begin{aligned}\mathbb{E}_{A,B,R}\left[\sigma_k(ARBR^*)\right] &= \sum_{S\in\binom{[n]}{k}} \mathbb{E}_{A,B,R}\left[\det[ARBR^*]_{S,S}\right], \\ &= \sum_{S,T\in\binom{[n]}{k}} \mathbb{E}_{A}\left[\det[A]_{S,T}\right]\mathbb{E}_{B,R}\left[\det[RBR^*]_{T,S}\right],\end{aligned}$$

Now we use Lemma 3.2.5 to conclude

$$\begin{aligned}\mathbb{E}_{A,B,R}\left[\sigma_k(ARBR^*)\right] &= \sum_{S,T\in\binom{[n]}{k}} \mathbb{E}_{A}\left[\det[A]_{S,T}\right]\delta_{T,S}\frac{\mathbb{E}_B\left[\sigma_k(B)\right]}{\binom{n}{k}}, \\ &= \frac{\mathbb{E}_A[\sigma_k(A)]\mathbb{E}_B[\sigma_k(B)]}{\binom{n}{k}}.\end{aligned}$$

□

Theorem 3.2.10. *Let $p(z)$ be the characteristic polynomial of A and $q(z)$ be the characteristic polynomial of B where A and B are $n\times n$ normal matrices with complex entries, then*

$$p(z)\boxtimes_n q(z) = \mathbb{E}_Q\left[\chi_z(AQBQ^*)\right],$$

with χ_z and $\mathbb{E}_Q$ as in Theorem 3.2.8.

Proof. It follows directly from Theorem 3.2.9 and definition of the symmetric multiplicative convolution. □

With the results in this section we related polynomial convolutions to expected characteristic polynomials. In the following section we will be able to link the convolutions to Free Probability Theory by assigning an empirical measure to the roots of a polynomial in a similar fashion one can assign a measure to the eigenvalues of a matrix.

3.3 The finite $\mathcal{R}_n$ transform

In finite probability, the $\mathcal{R}$ transform linearizes free additive convolution, which means the $\mathcal{R}$ transform of a free additive convolution of measures is simply the sum of the $\mathcal{R}$ tranforms of each measure. The $\mathcal{R}$ transform of a measure μ is often defined as a formal series

$$\mathcal{R}_\mu(s) = \sum_{k=0}^{\infty} s^k r_k,$$

with r_k being the free cumulants of the measure. For the finite free case, a similar transform can be defined, and we will prove that it converges to the $\mathcal{R}$ transform.

In the first part of this section we introduce some technical definitions and lemmas, then we define the finite $\mathcal{R}_n$ transform and prove its two main properties, namely that it converges to an $\mathcal{R}$ transform and that it linearizes symmetric additive convolution of polynomials. In the following two subsections we use the transform to find the polynomials associated to finite free distributions and later to prove the analogous in Finite Free Probability Theory to the most classical limit theorems in Probability Theory. The content of this section is taken from [25].

Before stating the results in this section, it is important to introduce the notation for "congruent polynomials". Let p, q be two polynomials on z, we say that p is congruent with q modulo z^n if the first $n+1$ coefficients of p and q coincide. We will denote this relationship by

$$p(z) \equiv q(z), \qquad \mod [z^n].$$

If we think of a formal power series as an "infinite polynomial", the definition of polynomial congruence can be extended. Notice that if $f(z)$ is a formal power series, then $f(z) \mod [z^n]$ is a polynomial of degree at most n.

3.3.1 Definition of the transform

Given any polynomial $p(z)$ with order p we can associate an empirical measure μ_p to its roots z_i given by

$$\mu_p(\{x\}) = \frac{1}{n}\sum_{j=1}^{p} \delta_{x,z_j}.$$

This measure is similar to the spectral empirical measure of a random matrix. We can find its Cauchy transform in terms of the polynomial with the following lemma.

Lemma 3.3.1. *Let p be a monic polynomial of order p with roots $\{z_i\}_{i=1}^n$, then the Cauchy transform of the empirical measure associated to the roots z_i is given by*

$$G_{\mu_p}(z) \coloneqq \frac{1}{n}\sum_{j=1}^{n} \frac{1}{z - z_j} = \frac{\partial_z p}{np}(z) = \frac{1}{n}\partial_z \ln p(z).$$

Proof. $p(z)$ is a monic polynomial with roots $\{z_j\}_{j\in[n]}$, then we can write,

$$p(z) = \prod_{j=1}^{n}(z - z_j).$$

By the Leibnitz rule we find

$$\partial_z p(z) = \sum_{j=1}^{n}\prod_{k\neq j}(z - z_k).$$

Using the last equation we have

$$\frac{\partial_z p}{np}(z) = \frac{1}{n}\sum_{j=1}^{n}\frac{\prod_{k\neq j}(z-z_k)}{\prod_{l=1}^{n}(z-z_l)} = \frac{1}{n}\sum_{j=1}^{n}\frac{1}{z-z_j} =: G_{\mu_p}(z).$$

□

Now we will define two objects that are auxiliary in the following proofs. Both are named transforms, but it is important to remark that one of them is an operator acting on functions while the other is a multiset which is uniquely related to another multiset in some way.

The existence of the U transform can be stated by the following lemma that we state without proof. The proof can be found in [14].

Lemma 3.3.2 (U transform)**.** *Let S be a multiset of complex numbers and denote by $|S| = n$ its number of elements with multiplicity. Then there exists an unique multiset of complex numbers T such that $|T| = n$ and*

$$\prod_{s_i \in S} (x - s_i) = \frac{1}{n} \sum_{t_i \in T} (x - t_i)^n.$$

We call such multiset the U transform of S.

We will use the U transform of a set of eigenvalues (or roots of a polynomial) in order to find explicit expressions for the $\mathcal{R}_n$ transform and prove some of its properties.

The Legendre transform, and the properties that will be proven later, will be useful to prove that the finite $\mathcal{R}_n$ transform converges to an $\mathcal{R}$ transform.

Definition 3.3.1 (Legendre's transform)**.** Let f a convex function in a domain $D \subset \mathbb{R}$ and define

$$D^* := \left\{ x^* \in \mathbb{R} : \sup_{x \in D} \{x x^* - f(x)\} < \infty \right\}. \tag{3.9}$$

We define f^* the Legendre transform of f as the function

$$\begin{aligned} f^* : D^* &\to \mathbb{R}. \\ s &\mapsto \sup_{x \in D} \{xs - f(x)\}. \end{aligned}$$

Lemma 3.3.3. *Let f be a strictly convex function in a domain $D \subset \mathbb{R}$ and such that its derivative exists in a point $x_0 \in D$. Then $\partial_x [f(x)]|_{x=x_0} \in D^*$ and*

$$f^*(f'(x_0)) = x_0 f'(x_0) - f(x_0).$$

If additionally, f has a second derivative, then the following two results are satisfied

$$\begin{aligned} (f')^{-1}(x_0) &= (f^*)'(x_0), \\ f''((f^*)'(x_0)) &= \frac{1}{(f^*)''(x_0)}. \end{aligned}$$

Proof. Since f is strictly convex and differentiable at x_0 we have for $x \in D, x \neq x_0$ that

$$f(x) > f(x_0) + (x - x_0) f'(x_0),$$

resting $x f'(x_0)$ in both sides leads to

$$f(x) - x f'(x_0) > f(x_0) - x_0 f'(x_0),$$

the inequality is reversed when we multiply by -1;

$$x f'(x_0) - f(x) < x_0 f'(x_0) - f(x_0).$$

This means that

$$\sup_{x\in D}\{xf'(x_0) - f(x)\} = x_0f'(x_0) - f(x_0) < \infty.$$

Then, by the definition of D^*, we have that $f'(x_0) \in D^*$ and the Legendre transform in this point is $f^*(f'(x_0)) = x_0f'(x_0) - f(x_0)$.

For the second part, if f has a second derivative at x_0, differentiate the last equation to find

$$(f^*)'(f'(x_0))f''(x_0) = x_0f''(x_0) + f'(x_0) - f'(x_0) = x_0f''(x_0).$$

The second derivative $f''(x_0)$ can not be zero because f is strictly convex, and therefore $(f^*)'(f'(x_0)) = z$, which means f^* and f' are inverse under composition (in any point where f is twice differentiable). For the second equation we use the fact that these functions are inverse and derive,

$$f'((f^*)'(z)) = z,$$

deriving with respect to z gives us

$$f''((f^*)(x_0))(f^*)''(x_0) = 1,$$

when we divide both sides by $(f^*)''(x_0)$

$$f''((f^*)(x_0)) = \frac{1}{(f^*)''(x_0)}.$$

□

The following theorem relates the Legendre transform to an L^∞ norm and will be used in the proof of the convergence of the finite $\mathcal{R}_n$ transform to an $\mathcal{R}$ transform.

Lemma 3.3.4. *Let $D \subset \mathbb{R}$ and μ a measure that is absolutely continuous with respect to the Lebesgue measure, then for any continuous function $f : D \to \mathbb{R}$ we have that*

$$f^*(s) = \ln \left\| e^{xs-f(x)} \right\|_\infty,$$

for all $s \in D^*$ where the Legendre transform and the norm are taken over D.

Proof. We can write f^* as $f^*(s) = \ln(\exp(f^*(s)))$, then

$$f^*(s) = \ln(\exp(f^*(s))) = \ln\left(\exp\left\{\sup_{x\in D}\{xs - f(x)\}\right\}\right),$$

using that $\exp(x)$ is monotone increasing,

$$= \ln\left(\sup_{x\in D}\exp(xs - f(x))\right) = \ln\left\| e^{xs-f(x)} \right\|_\infty.$$

□

Just as in the free case, the finite $\mathcal{R}_n$ transform can be defined in terms of another transform that we will call the finite $\mathcal{K}_n$ transform.

Definition 3.3.2 (The $\mathcal{K}_n$ transform [14])**.** Let A be an $n \times n$ symmetric matrix with real entries. We define the $\mathcal{K}_n$ transform of its empirical spectral measure μ_A as

$$\mathcal{K}_n^{\mu_A}(s) \coloneqq -\frac{\partial}{\partial s} \ln \left\| e^{xs} \Delta^+[xI - A] \right\|_n ,$$

where Δ^+ represents the normalized determinant

$$\Delta^+[A] = \begin{cases} \det[A]^{\frac{1}{n}} & \text{if } A \text{ is positive definite,} \\ 0 & \text{otherwise.} \end{cases}$$

and the integration domain for the norm is (ρ_A, ∞) with ρ_A the spectral radius of A.

The operator Δ^+ acting on A is a particular case of a much more general object called the Fuglede-Kadison determinant, which acts on a broad set of operators and satisfies many of the properties of determinants. More information about it can be found in [31].

With the finite $\mathcal{K}_n$ transform it is possible to define the finite $\mathcal{R}_n$ transform.

Definition 3.3.3 (The $\mathcal{R}_n$ transform)**.** Let A be an $n \times n$ symmetric matrix with real entries. We define the $\mathcal{R}_n$ transform of its empirical spectral measure μ_A as

$$\mathcal{R}_n^{\mu_A}(s) = \mathcal{K}_n^{\mu_A}(s) - \left(1 + \frac{1}{n}\right) \frac{1}{s}.$$

Notice that for bigger n, the definition above is more similar to the definition of the $\mathcal{R}$ transform in terms of the $\mathcal{K}$ transform given in Chapter 1. If we assume that $\mathcal{K}_n^{\mu_A}(s)$ converges to some $\mathcal{K}$ transform when $n \to \infty$, then $\mathcal{R}_n^{\mu_A}(s)$ would converge to the corresponding $\mathcal{R}$ transform. In the next theorem we will prove exactly this convergence.

Theorem 3.3.5. *Let A be a self-adjoint $n \times n$ matrix with empirical spectral distribution μ_A, then*

$$\lim_{n \to \infty} \mathcal{K}_n^{\mu_A}(s) = G_{\mu_A}^{-1}(s),$$

with $s \in (\rho_A, \infty)$ and where $G_{\mu_A}^{-1}(s)$ is the inverse under composition of $G_{\mu_A}(s)$.

Proof. We begin defining the function $g(x) \coloneqq -\ln \Delta^+[xI - A]$ and let $\lambda_1, \ldots, \lambda_n$ be the ordered eigenvalues of A. Notice that without loss of generality we can assume $I - A$ to be positive definite. When we differentiate g with respect to x, we find

$$\begin{aligned} \partial_x[g(x)] = \partial_x \left[-\ln \Delta^+[xI - A] \right] &= -\partial_x \left\{ \ln \left[\left(\prod_{j=1}^{n} (x - \lambda_j) \right)^{\frac{1}{n}} \right] \right\}, \\ &= -\partial_x \left\{ \ln \left[\prod_{j=1}^{n} (x - \lambda_j)^{\frac{1}{n}} \right] \right\} = -\partial_x \left[\sum_{j=1}^{n} \frac{1}{n} \ln(x - \lambda_j) \right], \\ &= -\frac{1}{n} \sum_{j=1}^{n} \frac{1}{x - \lambda_j} = -G_{\mu_A}(x). \end{aligned}$$

For the second derivative we have

$$\partial_{xx}[g(x)] = \partial_x \left\{ \left[-\frac{1}{n} \sum_{j=1}^{n} \frac{1}{x - \lambda_j} \right] \right\} = \frac{1}{n} \sum_{j=1}^{n} \frac{1}{(x - \lambda_j)^2} > 0.$$

So $g(x)$ is strictly convex and has a second derivative. Using Lemma 3.3.3 we get

$$(g^*)'(x) = (g')^{-1}(x) = (-G_{\mu_A})^{-1}(x) = G_{\mu_A}^{-1}(-x). \tag{3.10}$$

Now we find $(g^*)'(x)$ with the help of Lemma 3.3.4.

$$(g^*)'(s) = \frac{\partial}{\partial s} \ln \|e^{xs} - g(x)\|_\infty = \frac{\partial}{\partial s} \ln \left\|e^{xs} \Delta^+(xI - A)\right\|_\infty .$$

Substituting the last in (3.10) we get

$$G_{\mu_A}^{-1}(s) = (g^*)'(-s) = -\frac{\partial}{\partial s} \ln \left\|e^{-xs} \Delta^+(xI - A)\right\|_\infty .$$

On the other hand, the limit of $\mathcal{K}_n^{\mu_A}$ is

$$\lim_{n\to\infty} \mathcal{K}_n^{\mu_A}(s) = -\lim_{n\to\infty} \frac{\partial}{\partial s} \ln \left\|e^{xs} \Delta^+(xI - A)\right\|_n = -\frac{\partial}{\partial s} \ln \left\|e^{-xs} \Delta^+(xI - A)\right\|_\infty .$$

This gives the desired result. □

The simplest finite $\mathcal{K}_n$ transform we can find is the one related to the measure $\mu_0 = \delta_0$. Knowing this transform will also be useful for proving further results.

Lemma 3.3.6. *Let μ_0 be the probability measure corresponding to an atom in 0 with probability 1, i.e.*

$$\mu_0(\{x\}) = \begin{cases} 1, & si\ x = 0, \\ 0, & otherwise. \end{cases}$$

Then its finite $\mathcal{K}_n$ transform is

$$\mathcal{K}_n^{\mu_0}(s) = \left(1 + \frac{1}{n}\right) \frac{1}{s}.$$

Proof. Notice that for every m, the measure μ_0 is the empirical spectral measure associated to the matrix $0_{n\times n}$, so we can write

$$\left\|e^{-xs} \Delta^+(xI)\right\|_n = \left\|e^{-xs} (x^n)^{\frac{1}{n}}\right\|_n .$$

Remember the integral is taken from ρ_A to ∞. In this case all the eigenvalues are 0, so

$$\left\|e^{-xs} \Delta^+(xI)\right\|_n = \left(\int_0^\infty (e^{-xs} x)^n\right)^{\frac{1}{n}} = (\mathcal{L}[x^n](ns))^{\frac{1}{n}},$$

where $\mathcal{L}[f]$ represents the Laplace transform of the function f. With this, we have

$$\left\|e^{-xs}\Delta^+(xI)\right\|_n = \left(\frac{n!}{(ns)^{n+1}}\right)^{\frac{1}{n}}.$$

The logarithm of the previous expression is

$$\ln\left(\frac{n!}{(ns)^{n+1}}\right)^{\frac{1}{n}} = \frac{1}{n}\ln\left(\frac{n!}{(ns)^{n+1}}\right) = \frac{1}{n}\log\left(\frac{n!}{(n)^{n+1}}\right) - \frac{n+1}{n}\log s.$$

Finally, we differentiate to find the desired expression.

$$\ln\left(\frac{1}{n}\log\left(\frac{n!}{(n)^{n+1}}\right) - \frac{n+1}{n}\log s\right) = -\partial_s\left[-\frac{n+1}{n}\log s\right] = \left(\frac{n+1}{n}\right)\frac{1}{s}.$$

□

Notice that once we have this, an alternative definition for $\mathcal{R}_n^{\mu_A}(s)$ would be $\mathcal{R}_n^{\mu_A}(s) = \mathcal{K}_n^{\mu_A}(s) - \mathcal{K}_n^{\mu_0}(s)$.

Our next goal is to prove that the finite $\mathcal{R}_n$ transform linearizes the symmetric additive convolution. In order to do so, we will first prove a couple basic lemmas.

Lemma 3.3.7. *Let p, q be polynomials with degree n and U_p, U_q be the U transforms of its sets of roots, then*

$$[p \boxplus_n q](x) = \frac{1}{n^2}\sum_{u_j\in U_p, u_k\in U_q}(x - u_j - u_k)^n.$$

Proof. Let $p(x) = \sum_{j=0}^n x^{n-j}(-1)^j p_j$. Let x_i be its roots, by definition of the U transform we have,

$$p(x) = \prod_{j=0}^n (x - x_j) = \frac{1}{n}\sum_{u_j\in U_p}(x - u_j)^n = \sum_{j=0}^n x^{n-j}(-1)^j\binom{n}{j}\frac{1}{n}\sum_{u_k\in U_p}u_k^j.$$

Equating coefficients we find that $p_j = \binom{n}{j}\frac{1}{n}\sum_{u_k\in U_p}u_k^j$ and the analogous happens for q_k. Finally, we find the convolution

$$
\begin{aligned}
[p \boxplus_n q](x) &= \sum_{k=0}^{n} x^{n-k}(-1)^k \sum_{i+j=k} \frac{(n-i)!(n-j)!}{(n-k)!n!} \binom{n}{i}\binom{n}{j} \frac{1}{n} \sum_{u_m \in U_p} u_m^i \frac{1}{n} \sum_{u_l \in U_q} u_l^j, \\
&= \sum_{k=0}^{n} x^{n-k}(-1)^k \sum_{i+j=k} \frac{n!}{i!j!(n-k)!} \frac{1}{n^2} \sum_{u_m \in U_p, u_l \in U_q} u_m^i u_l^j, \\
&= \frac{1}{n^2} \sum_{u_m \in U_p, u_l \in U_q} \sum_{k=0}^{n} \binom{n}{k} x^{n-k}(-1)^k \sum_{j=0}^{k} \binom{k}{j} u_m^{k-j} u_l^j, \\
&= \frac{1}{n^2} \sum_{u_m \in U_p, u_l \in U_q} \sum_{k=0}^{n} \binom{n}{k} x^{n-k}(-1)^k (u_m + u_l)^k, \\
&= \frac{1}{n^2} \sum_{u_m \in U_p, u_l \in U_q} (x - u_m - u_l)^n.
\end{aligned}
$$

□

Lemma 3.3.8. *Let A be an $n \times n$ real symmetric matrix. Denote by $\lambda(A)$ its spectrum and by U_A the U transform of $\lambda(A)$, then*

$$
\frac{\|e^{-xs}\Delta^+(xI - A)\|_n^n}{\|e^{-xs}\Delta^+(xI)\|_n^n} \equiv \frac{1}{n} \sum_{u_i \in U_A} e^{-nsu_i}, \qquad \mod [s^{n+1}].
$$

Proof. Suppose, without loss of generality that A is positive definite. By definition of the U transform, for any $n \times n$ such symmetric matrix A we have

$$
\Delta^+(xI - A)^n = \det[xI - A] = \prod_{j=1}^{n}(x - \lambda_j) = \frac{1}{n} \sum_{u_i \in U_A} (x - u_i)^n.
$$

Using this and the previously found result for $\|e^{-xs}\Delta^+(xI)\|_n^n$ we find

$$
\frac{\|e^{-xs}\Delta^+(xI - A)\|_n^n}{\|e^{-xs}\Delta^+(xI)\|_n^n} = \frac{(ns)^{n+1}}{n!} \left\|e^{-xs}\Delta^+(xI - A)\right\|_n^n = \frac{(ns)^{n+1}}{n!} \int_{\rho_A}^{\infty} e^{-nxs}\Delta^+(xI - A)^n \mathrm{d}x.
$$

By definition of the U transform

$$
\frac{\|e^{-xs}\Delta^+(xI - A)\|_n^n}{\|e^{-xs}\Delta^+(xI)\|_n^n} = \frac{(ns)^{n+1}}{n!} \int_{\rho_A}^{\infty} e^{-nxs} \frac{1}{n} \sum_{u_i \in U_A} (x - u_i)^n \mathrm{d}x.
$$

Use the change of variables $y = x + \rho_A$

$$
\frac{\|e^{-xs}\Delta^+(xI - A)\|_n^n}{\|e^{-xs}\Delta^+(xI)\|_n^n} = \frac{(ns)^{n+1}e^{-ns\rho_A}}{n!} \int_{0}^{\infty} e^{-nsy} \frac{1}{n} \sum_{u_i \in U_A} (y + \rho_A - u_i)^n \mathrm{d}y.
$$

Using linearity and the definition of the Laplace transform

$$\frac{\|e^{-xs}\Delta^+(xI-A)\|_n^n}{\|e^{-xs}\Delta^+(xI)\|_n^n} = \frac{1}{n}\sum_{u_i\in U_A}\frac{(ns)^{n+1}}{n!}e^{-ns\rho_A}\mathcal{L}\{(y+\rho_A-U_A)^n\}(ns).$$

By linearity of the Laplace transform

$$\frac{\|e^{-xs}\Delta^+(xI-A)\|_n^n}{\|e^{-xs}\Delta^+(xI)\|_n^n} = \frac{1}{n}\sum_{u_i\in U_A}\frac{(ns)^{n+1}}{n!}e^{-ns\rho_A}\sum_{k=0}^{n}\binom{n}{k}(\rho_A-u_i)^{n-k}\mathcal{L}\{y^k\}(ns).$$

Finally, we expand and cancel terms to arrive to

$$\begin{aligned}\frac{\|e^{-xs}\Delta^+(xI-A)\|_n^n}{\|e^{-xs}\Delta^+(xI)\|_n^n} &= \frac{1}{n}\sum_{u_i\in U_A}\frac{(ns)^{n+1}}{n!}e^{-ns\rho_A}\sum_{k=0}^{n}\frac{n!}{k!(n-k)!}(\rho_A-u_i)^{n-k}\frac{k!}{(ns)^{k+1}},\\ &= \frac{1}{n}\sum_{u_i\in U_A}e^{-ns\rho_A}\sum_{k=0}^{n}\frac{(ns)^{n-k}(\rho_A-u_i)^{n-k}}{(n-k)!}.\end{aligned}$$

We can use the polynomial congruence relation for the series expansion of the exponential function and the truncated series $\sum_{k=0}^{n}x^n/n!$.

$$\begin{aligned}\frac{\|e^{-xs}\Delta^+(xI-A)\|_n^n}{\|e^{-xs}\Delta^+(xI)\|_n^n} &\equiv \frac{1}{n}\sum_{u_i\in U_A}e^{-ns\rho_A}e^{ns(\rho_A-u_i)}, \qquad \mod [s^{n+1}],\\ &\equiv \frac{1}{n}\sum_{u_i\in U_A}e^{-nsu_i}, \qquad \mod [s^{n+1}].\end{aligned}$$

□

With the previous result in hand, it is now easier to find the finite $\mathcal{R}_n$ transform of a spectral measure in terms of the U transform of its spectrum.

Corollary 3.3.9. *Let A be a an $n\times n$ symmetric matrix with real entries, with spectrum $\lambda(A)$ and U_A be the U transform of $\lambda(A)$, then*

$$\mathcal{R}_n^{\mu_A}(s) \equiv -\frac{1}{n}\frac{\partial}{\partial s}\ln\left(\frac{1}{n}\sum_{u_i\in U_A}e^{-nsu_i}\right), \qquad \mod [s^n].$$

Proof. Taking logarithm in both sides of Lemma 3.3.8 we get

$$\ln\left(\frac{\|e^{-xs}\Delta^+(xI-A)\|_n^n}{\|e^{-xs}\Delta^+(xI)\|_n^n}\right) \equiv \ln\left(\frac{1}{n}\sum_{u_i\in U_A}e^{-nsu_i}\right), \qquad \mod [s^{n+1}].$$

The first $n+1$ coefficients of the power series coincide, so the first k coefficients of the derivatives also coincide. Now we use the definition of the $\mathcal{R}_n$ transform

$$\begin{aligned}\mathcal{R}_n^{\mu_A}(s) &= \mathcal{K}_n^{\mu_A}(s) - \mathcal{K}_n^{\mu_0}(s),\\ &= -\frac{\partial}{\partial s}\ln\left\|e^{-xs}\Delta^+(xI-A)\right\|_n + \frac{\partial}{\partial s}\ln\left\|e^{-xs}\Delta^+(xI)\right\|_n,\\ &= -\frac{1}{n}\frac{\partial}{\partial s}\ln\left(\frac{\|e^{-xs}\Delta^+(xI-A)\|_n^n}{\|e^{-xs}\Delta^+(xI)\|_n^n}\right),\\ &\equiv -\frac{1}{n}\frac{\partial}{\partial s}\ln\left(\frac{1}{n}\sum_{u_i\in U_A} e^{-nsu_i}\right), \qquad \bmod [s^n].\end{aligned}$$

□

One of the central results in this section is the equivalence of the symmetric additive convolution of two polynomials and the sum of the finite $\mathcal{R}_n$ transform of the measures associated to its roots (which is in turn the empirical measure associated to the eigenvalues of some matrix). The following theorem shows us that the analogy between the finite $\mathcal{R}_n$ transform and the $\mathcal{R}$ transform is not only in terms of their convergence, but also in the role of linearizing convolution.

Theorem 3.3.10. *Let A and B be two $n\times n$ real symmetric matrices. The following are equivalents:*

- $$\mathcal{R}_n^{\mu_A}(s) + \mathcal{R}_n^{\mu_B}(s) = \mathcal{R}_n^{\mu_{A+B}}(s), \qquad \bmod [s^n].$$
- $$\det[xI-A] \boxplus_n \det[xI-B] = \det[xI-A-B].$$

Proof. Let us denote by U_A, U_B and U_{A+B} the U transforms of the $\lambda(A), \lambda(B)$ and $\lambda(A+B)$, respectively. By Lemma 3.3.7, the second statement is equivalent to

$$\mathbb{E}\left[(x-U_A-U_B)^n\right] = \mathbb{E}\left[(x-U_{A+B})^n\right].$$

Expanding the power and equating terms, the last happens if and only if the first m moments of U_A+U_B and U_{A+B} coincide. This is in turn equivalent to

$$\mathbb{E}\left[e^{-ms(U_A+U_B)}\right] = \mathbb{E}\left[e^{-ms(U_A+U_B)}\right] = \mathbb{E}\left[e^{-msU_{A+B}}\right], \qquad \bmod [s^{n+1}].$$

Define the function $f_A(s) \coloneqq -\frac{1}{n}\ln\mathbb{E}\left[e^{-msU_A}\right]$ and similarly for B and $A+B$. Using this function, the second statement is equivalent to

$$f_A(s) + f_B(s) \equiv f_{A+B}(s), \qquad \bmod [s^{n+1}]. \tag{3.11}$$

Using Corollary 3.3.9, the first statement is equivalent to

$$\frac{\partial}{\partial_s} f_A(s) + \frac{\partial}{\partial_s} f_B(s) \equiv \frac{\partial}{\partial_s} f_{A+B}(s), \qquad \bmod [s^n]. \tag{3.12}$$

So proving the equivalence of the two statements reduces to prove the equivalence between (3.11) and (3.12). Because the formal series in (3.11) coincide up to the $(n+1)$th term,

the derivatives coincide up to the nth term, so one implication is trivial. For the second implication, it suffices to prove that $f_A(0) + f_B(0) = f_{A+B}(0)$, but we have that $f_A(0) = f_B(0) = f_{A+B}(0)$, so the result follows. □

As it was mentioned several times before in this chapter, the relationship between symmetric additive convolution and free additive convolution shown by these results is not exclusive. Similar results can be given for the symmetric multiplicative convolution and free multiplicative convolution. Although we are not including them here, they can be consulted in [28] and [3].

For the last part of this chapter, we will use the recently exposed tools to find the polynomials corresponding to three classical laws in the finite free setting; namely the constant, Gaussian and Poisson distributions. After that, we will prove the finite free version of the three corresponding limit theorems; the law of large numbers, the central limit theorem and the Poisson limit theorem.

3.3.2 Basic distributions in Finite Free Probability

We have seen that the finite $\mathcal{R}_n$ transform is a power series that linearizes finite free convolution. In free probability terms, we can say that the corresponding coefficients in the expansion are the finite free cumulants. Using this and the fact that we know what are the cumulants for several well known distributions, we can find the polynomials corresponding to these laws in Finite Free Probability Theory. As a prerequisite, we prove a theorem that allows us to recover a polynomial from the finite $\mathcal{R}_n$ transform of its roots.

Theorem 3.3.11. *Let $P(n^{-1}\partial_z)$ be a polynomial on $n^{-1}\partial_z$ such that the linear differential operator applied to x^n, $P(n^{-1}\partial_z)[z^n]$ is a monic polynomial. Then $P(n^{-1}\partial_z) = \det[zI - A]$ for a matrix A if and only if*

$$\frac{1}{n}\frac{\partial_s P(s)}{P(s)} \equiv -\mathcal{R}_n^{\mu_A}(s), \qquad \mod [s^{n+1}]. \tag{3.13}$$

Proof. Recall that

$$\mathcal{R}_n^{\mu_A}(s) \equiv -\frac{1}{n}\frac{\partial}{\partial s}\ln\left(\frac{1}{n}\sum_{u_i\in U_A} e^{-nsu_i}\right), \qquad \mod [s^n].$$

Also

$$\frac{1}{n}\frac{\partial_s P(s)}{P(s)} = \frac{1}{n}\partial_s[\ln P(s)].$$

Then (3.13) is satisfied if and only if

$$P(s) \equiv \frac{1}{n}\sum_{u_i\in U_A} e^{-nsu_i}, \qquad \mod [s^{n+1}].$$

The last relationship in turn is equivalent to

$$P(n^{-1}\partial_z)[z^n] = \left(\frac{1}{n}\sum_{u_i\in U_A} e^{-u_i\partial_z}\right)[z^n] = \frac{1}{n}\sum_{u_i\in U_A} e^{-u_i\partial_z}[z^n].$$

Expanding the last term and using the definition of U_A, we find

$$\begin{aligned}\frac{1}{n}\sum_{u_i\in U_A} e^{-u_i\partial_z}[z^n] &= \frac{1}{n}\sum_{u_i\in U_A}\sum_{j=0}^{n}\frac{(-u_i)^j}{j!}\partial_z^j[z^n],\\ &= \frac{1}{n}\sum_{u_i\in U_A}\sum_{j=0}^{n}(-u_i)^j\frac{n!}{j!(n-j)!}z^{n-j} = \frac{1}{n}\sum_{u_i\in U_A}(z-u_i)^n = \det[zI-A].\end{aligned}$$

□

With this theorem, we are ready to find the polynomials corresponding to the basic finite free distributions.

Constant distribution

A constant random variable has a first cumulant equal to $c \in \mathbb{R}$ and all the remaining cumulants equal to zero. Then its $\mathcal{R}_n^{\mu_A}$ transform is the constant formal series

$$\mathcal{R}_n^{\mu_A}(s) = c.$$

We can use Theorem 3.3.11 to find the polynomial p with this $\mathcal{R}_n^{\mu_A}$ transform.

$$\begin{aligned}-c &= \frac{1}{n}\partial_s[\ln P(s)],\\ \Rightarrow -\int ncds &= \ln P(s),\\ \Rightarrow \exp\{-ncs + c_0\} &= P(s),\end{aligned}$$

for some constant c_0. Then $p(z)$ is the monic scaling of the polynomial $P(n^{-1}\partial_z)[z^n] = e^{-c\partial_z}[z^n]$. Expanding we find

$$\begin{aligned}p(z) = e^{-c\partial_z}[z^n] &= \sum_{k=0}^{n}\frac{(-c)^k}{k!}\partial_z^k[z^n],\\ &= \sum_{k=0}^{n}\frac{n!}{k!(n-k)!}(-c)^k z^{n-k} = (z-c)^n.\end{aligned}$$

Thus, as we would expect, p is the monic polynomial with all the roots equal to c which is also the expected characteristic polynomial of the deterministic matrix cI.

Gaussian distribution

In free and classical probability, the central limit random variable is characterized for having the first two cumulants non-zero and the rest of the cumulants equal to zero. We already know what is the polynomial corresponding to a constant distribution in finite free probability. To find the general polynomial corresponding to a Gaussian, we can find the polynomial with the second cumulant non-zero and the rest equal to zero. If the second cumulant is equal to σ^2, then the formal series expansion of $\mathcal{R}_n^{\mu_A}$ is

$$\mathcal{R}_n^{\mu_A}(s) = s\sigma^2.$$

We can use Theorem 3.3.11 again to find

$$\begin{aligned}
\frac{1}{n}\partial_s[\ln P(s)] &= -s\sigma^2, \\
\Rightarrow \ln P(s) &= -n\sigma^2 \int s\mathrm{d}s, \\
\Rightarrow P(s) &= \exp\left\{-n\sigma^2\frac{s^2}{2} + c_0\right\},
\end{aligned}$$

where the constant c_0 accounts for making the polynomial monic. Now, evaluating P in $n^{-1}\partial_z$ and applying to z^n, we have

$$p(z) = P(n^{-1}\partial_z)[z^n] = e^{-\sigma^2\frac{\partial_z^2}{2n}}[z^n] =: H_n(z, \sigma^2/n).$$

The last expression corresponds to the Hermite polynomial with variance σ^2/n.

Poisson distribution

The Poisson distribution has all of its cumulants equal to a constant $\nu \in \mathbb{R}$, the formal series expansion of its $\mathcal{R}_n^{\mu_A}$ transform is

$$\mathcal{R}_n^{\mu_A}(s) = \sum_{k=0}^{\infty} \nu s^n = \nu \sum_{k=0}^{\infty} s^n = \frac{\nu}{1-s}.$$

Where the convergence is taken in $|s| < 1$. Another use of Theorem 3.3.11 leads to

$$\begin{aligned}
\frac{\nu}{1-s} = \mathcal{R}_n^{\mu_A}(s) &= -\frac{1}{n}\partial_s[\ln P(s)], \\
\Rightarrow -\int \frac{n\nu}{s-1}\mathrm{d}s &= \ln P(s), \\
\Rightarrow \exp\{n\nu\ln(1-s)\} &= P(s).
\end{aligned}$$

Finally, we evaluate in $n^{-1}\partial_z$ and apply to z^n and find the polynomial that has the given cumulants.

$$P(n^{-1}\partial_z)[z^n] = \left(1 - \frac{\partial_z}{n}\right)^{n\nu}[z^n] =: L_n^{(\nu-1)n}(z). \tag{3.14}$$

So the finite free Poisson distribution corresponds to a generalized Laguerre polynomial. Notice that when $\nu = 1$, we recover the original definition of the Laguerre polynoamials given in Section 3.1.

3.3.3 Finite Free Limit Theorems

One of the central topics in Probability Theory is the convergence of sums of random variables to some laws. In Free Probability we have analogous convergence theorems for the main distributions. Three of the most famous limit theorems have the previously shown laws as limits (constant, Gaussian, Poisson). In non-commutative Probability Theory in general, these laws are characterized by their cumulants and happen to be limits of the corresponding sums of random variables. In what follows, we will show that when we re-scale and convolute polynomials under standard assumptions, we also recover our three limit laws.

Theorem 3.3.12 (Finite Free Law of Large Numbers)**.** *Let $(p_i(z))_{i=1}^n$ be a sequence of degree n monic polynomials with real roots r_{ij} such that*

$$p_i(z) = \prod_{j=1}^{n}(z - r_{ij}).$$

And for every fixed i,

$$\frac{1}{n}\sum_{j=1}^{n} r_{ij} = m,$$
$$\frac{1}{n}\sum_{j=1}^{n} r_{ij}^2 < c,$$

for some positive constant c. Define $q_i(k, z) \coloneqq k^{-n}p_i(kz)$. Then

$$\lim_{k\to\infty}[q_1(k, z) \boxplus_n q_2(k, z) \boxplus_n \cdots \boxplus_n q_k(k, z)] = (x - m)^n.$$

Proof. Write for every i the polynomial $p_i(z)$ as

$$p_i(z) = z^n + a_{i1}z^{n-1} + \cdots$$

If $P_i(\partial_z)$ is a linear operator that acts on z^n to generate $p_i(z)$, then

$$P_i(\partial_z) = 1 + a_{i1}n\partial_z + \cdots$$

Since the coefficient a_{i1} is the sum of the roots, by hypothesis we have that $a_{i1} = m$ for every i. Now, the last equation implies that if q_i is generated by a linear differential operator $Q_i(\partial_z)$, then

$$Q_i(\partial_z) = 1 - \frac{m}{k}\partial_z + O(k^{-2}).$$

Using the multiplicative Theorem 3.1.1 we find

$$\begin{aligned}\lim_{k\to\infty}[q_1(k,z)\boxplus_n q_2(k,z)\boxplus_n\cdots\boxplus_n q_k(k,z)] &= \lim_{k\to\infty}\prod_{j=1}^{k} Q_j(\partial_z)[z^n],\\ &= \lim_{k\to\infty}\left(1-\frac{m}{k}\partial_z+O(k^{-2})\right)^k[z^n],\\ &= e^{-m\partial_z}[z^n]=(z-m)^n.\end{aligned}$$

□

Theorem 3.3.13 (Finite Free Central Limit Theorem [28])**.** *Let $p_1, p_2, \dots$ be a sequence of degree n real rooted polynomials with $p_i=\prod_j(x-r_{i,j})$ such that*

$$\sum_{j=1}^{n} r_{i,j}=0, \qquad \frac{1}{n}\sum_{j=1}^{n} r_{i,j}^2=\sigma^2, \tag{3.15}$$

for all i. Define $q_i(x)=k^{-n/2}p_i(\sqrt{k}x)$, then

$$\lim_{k\to\infty}(q_1\boxplus_n\cdots\boxplus_n q_k)=H_n(z,\sigma^2/(n-1)).$$

with $H_n(z,\sigma^2/(n-2))$ represents the nth Hermite polynomial with variance $\sigma^2/(n-1)$.

Proof. Using the Vieta's formulas and the hypotheses (3.15), we have that for every i it is satisfied

$$\begin{aligned}a_1 &= \sum_{j=1}^{n} r_{ij}=0,\\ a_2 &= \sum_{1\le j<m\le n} r_{ij}r_{im}=\frac{1}{2}\sum_{j=1}^{n} r_{ij}\sum_{m\neq j} r_{im}=\frac{1}{2}\sum_{j=1}^{n} r_{ij}\left(a_1-r_{ij}\right),\\ &= -\frac{1}{2}\sum_{j=1}^{n} r_{ij}^2=-\frac{1}{2}n\sigma^2.\end{aligned}$$

So every p_i has the form

$$p_i(z)=z^n+(0)z^{n-1}-\frac{n\sigma^2}{2n(n-1)}z^{n-2}+\cdots$$

Multiplying by the factors to get q, we get

$$\begin{aligned}q_i(z) &= z^n+(0)z^{n-1}-\frac{n\sigma^2k^{-n/2}}{2n(n-1)}k^{\frac{n-2}{2}}z^{n-2}+\cdots,\\ &= z^n-\frac{\sigma^2}{2(n-1)k}z^{n-2}.\end{aligned}$$

This means that the linear differential operator $Q_i(\partial_z)$ that generates q_i must have the form

$$Q_i(\partial_z)[z^n] = \left(1 - \frac{\sigma^2}{2(n-1)k}\partial_z^2 + O(k^{-3/2})\right)[z^n].$$

We recall the multiplicative Theorem 3.1.1 to find,

$$\begin{aligned} q_1 \boxplus_n \cdots \boxplus_n q_k &= \prod_{j=1}^{k} Q_j(\partial_z)[z^n], \\ &= \left(1 - \frac{\sigma^2}{2(n-1)k}\partial_z^2 + O(k^{-3/2})\right)^k [z^n]. \end{aligned}$$

Finally, letting $k \to \infty$, we can find

$$\lim_{k\to\infty} (q_1 \boxplus_n \cdots \boxplus_n q_k) = e^{-\frac{\sigma^2}{2(n-1)}}[z^n].$$

The last expression is the definition given previously for $H_n(z, \sigma^2/(n-1))$. □

Theorem 3.3.14 (Finite Free Poisson Limit Theorem)**.** *Let $p(z) = z^{n-1}(z-1)$. And for $\nu n \in \mathbb{N}$, the νn times symmetric additive convolution of $p(z)$ with itself is a polynomial corresponding to the finite free Poisson distribution.*

$$\underbrace{p(z) \boxplus_n p(z) \boxplus_n \cdots \boxplus_n p(z)}_{\nu n \text{ times}} = \left(1 - \frac{1}{n}\partial_z\right)^{\nu n}[z^n].$$

Proof. Notice that $z^{n-1}(z-1) = z^n - z^{n-1}$ can be written as generated by a linear differential operator in the following way

$$\left(1 - \frac{1}{n}\partial_z\right)[z^n] = z^n - z^{n-1}.$$

Now, applying once again Theorem 3.1.1 we get

$$\underbrace{p(z) \boxplus_n p(z) \boxplus_n \cdots \boxplus_n p(z)}_{\nu n \text{ times}} = \left(1 - \frac{1}{n}\partial_z\right)^{\nu n}[z^n].$$

Which is exactly the polynomial corresponding to the finite free Poisson distribution given in (3.14). □

Chapter 4

Deterministic eigenvalue processes for matrix-valued processes

In this last chapter, we give a relationship between finite free probability and the eigenvalues of matrix-valued stochastic processes. In the first section, we construct a matrix process whose eigenvalues evolve according to the dynamics of the Dyson Brownian motion without the martingale part. We call this process the deterministic Dyson Brownian motion. In the second section, we construct a similar matrix-valued process whose spectrum evolves according to the finite variation part of equation (2.11). We are especially interested in the deterministic versions of the Wishart and Jacobi processes. In the third section, we relate these processes with finite free probability by showing that the convolution of certain polynomials satisfies differential equations that ultimately lead to the conclusion that their roots follow similar dynamics to the deterministic version of the eigenvalue processes. This constitutes the main result of this work.

4.1 Deterministic Dyson Brownian motion

In this section we prove that a given matrix-valued stochastic process has a deterministic spectrum and follows the dynamics of the finite variation part in the Dyson Brownian motion. The proof uses the same techniques as the former results for the stochastic differential equations of eigenvalue processes.

Theorem 4.1.1. *Let Z be a process with covariation $\mathrm{d}Z_{ij}\mathrm{d}Z_{kl} = (\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk}-2\delta_{ij}\delta_{kl}\delta_{ik})\mathrm{d}t$ and no finite variation part, which means Z is a symmetric matrix with independent Brownian motions in its entries, except for the diagonal, where $Z_{ii} = 0$. Let X be a matrix valued process such that $X = H^T\Lambda H$ and that satisfies the stochastic differential equation*

$$H^T\mathrm{d}XH = \mathrm{d}Z.$$

Then the eigenvalue process Λ satisfies

$$\mathrm{d}\lambda_i = \sum_{k\neq i}\frac{\mathrm{d}t}{\lambda_i-\lambda_k}. \tag{4.1}$$

Proof. Define $\mathrm{d}A = H^T\partial H$ and $\mathrm{d}N = H^T\partial ZH$.

The same procedure as in 2.2.1 leads to

$$\mathrm{d}\Lambda = \mathrm{d}N + \Lambda \mathrm{d}A - \mathrm{d}A\Lambda.$$

We conclude that $\mathrm{d}\lambda_i = \mathrm{d}N_{ii}$ and for $i \neq j$,

$$\begin{aligned} 0 &= \mathrm{d}N_{ij} + \lambda_i \mathrm{d}A_{ij} - \lambda_j \mathrm{d}A_{ij}, \\ \Rightarrow \mathrm{d}A_{ij} &= \frac{\mathrm{d}N_{ij}}{\lambda_j - \lambda_i}. \end{aligned}$$

The quadratic covariation of N is the same as the one for Z because they only differ in a finite variation term, so

$$\mathrm{d}N_{ij}\mathrm{d}N_{kl} = \mathrm{d}\big\langle (H^T \mathrm{d}X H)_{ij}, (H^T \mathrm{d}X H)_{kl} \big\rangle(t) = \mathrm{d}\big\langle Z_{ij}, Z_{kl} \big\rangle = \big(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{kj} - 2\delta_{ij}\delta_{kl}\delta_{ik}\big)\mathrm{d}t.$$

Particularly, we have that $\mathrm{d}N_{ii}\mathrm{d}N_{jj} = 0$ for every choice of j and i. Thus every entry in the diagonal of N is a finite variation process and so it is λ_i. Let us finally compute the finite variation part F of N.

$$\begin{aligned} \mathrm{d}F &= \frac{1}{2}\big(H^T \mathrm{d}X \mathrm{d}H + \mathrm{d}H^T \mathrm{d}X H\big), \\ &= \frac{1}{2}\big(H^T \mathrm{d}X H H^T \mathrm{d}H + \mathrm{d}H^T H H^T \mathrm{d}X H\big), \\ &= \frac{1}{2}\big(\mathrm{d}N \mathrm{d}A + (\mathrm{d}N \mathrm{d}A)^T\big). \end{aligned}$$

For $\mathrm{d}N\mathrm{d}A$ we have

$$\begin{aligned} (\mathrm{d}N\mathrm{d}A)_{ij} &= \sum_{k\neq j} \mathrm{d}N_{ik}\mathrm{d}A_{kj} = \sum_{k\neq j} \frac{\mathrm{d}N_{ik}\mathrm{d}A_{kj}}{\lambda_j - \lambda_k}, \\ &= \sum_{k\neq j} \frac{\delta_{ik}\delta_{kj} + \delta_{ij}\delta_{kk} - 2\delta_{ik}\delta_{kj}\delta_{ij}}{\lambda_j - \lambda_k}\mathrm{d}t = \delta_{ij} \sum_{k\neq j} \frac{\mathrm{d}t}{\lambda_j - \lambda_k}. \end{aligned}$$

Then F is diagonal with $\mathrm{d}F_{ii} = \sum_{k\neq i} \frac{\mathrm{d}t}{\lambda_i - \lambda_k}$. We conclude that

$$\mathrm{d}\lambda_i = \sum_{k\neq i} \frac{\mathrm{d}t}{\lambda_i - \lambda_k}.$$

□

4.2 Deterministic eigenvalue processes for matrix-valued diffusions

Now that we have shown the construction of a matrix-valued process whose eigenvalue is the deterministic Dyson Brownian motion, we generalize the result to get processes with a deterministic spectrum that can follow the dynamics of any eigenvalue process with the form (2.11).

Theorem 4.2.1. *Let $B = (B(t), t \geq 0)$ be a Brownian motion in $\mathcal{M}_{n,n}(\mathbb{R})$ and $Y(t) = QMQ^T$ be a symmetric $n \times n$ matrix-valued stochastic process satisfying the stochastic differential equation*

$$\mathrm{d}Y(t) = g(Y(t))\mathrm{d}B(t)h(Y(t)) + h(Y(t))\mathrm{d}B(t)^T g(Y(t)) + b(Y(t))\mathrm{d}t, \tag{4.2}$$

where g, h, b are real functions acting spectrally, and $Y(0)$ is a symmetric $n \times n$ matrix with n different eigenvalues.

Let $G(x, y) = g^2(x)h^2(y) + g^2(y)h^2(x)$, τ be defined as in (2.10), and take a process $X = (X(t), t \geq 0)$ with diagonalization $X = H\Lambda H^T$ satisfying:

- $(H^T(\mathrm{d}X)H)_{ij} = (Q^T(\mathrm{d}Y)Q)_{ij}$ *for* $i \neq j$.
- $(H^T(\mathrm{d}X)H)_{ii} = 0$ *for every* $i \in [n]$.

Then, for $t < \tau$ the eigenvalue process $\Lambda(t)$ verifies the following system of stochastic differential equations:

$$\mathrm{d}\lambda_i = \left(b(\lambda_i) + \sum_{k \neq i} \frac{G(\lambda_i, \lambda_k)}{\lambda_i - \lambda_k}\right)\mathrm{d}t. \tag{4.3}$$

Proof. We define again L to be the stochastic logarithm of H, $\mathrm{d}L \coloneqq H^T \circ \mathrm{d}H$ and by the same techniques as in the proof of Theorem 2.2.1 we have that,

$$\mathrm{d}\Lambda = H^T \circ (\partial X) \circ H - (\mathrm{d}L) \circ \Lambda + \Lambda \circ \mathrm{d}L.$$

Define $\mathrm{d}N \coloneqq H^T \circ (\partial X) \circ H$. Using that $\Lambda \circ \partial L - (\partial L) \circ \Lambda$ has zero diagonal, we get that $\mathrm{d}\Lambda_i = \mathrm{d}N_{ii}$. The martingale part of $\mathrm{d}N_{ii}$ is the martingale part of $(H^T(\mathrm{d}X)H)_{ii}$, and by hypothesis, this is zero.

From the above equation, for $i \neq j$ we have that

$$0 = \mathrm{d}N_{ij} - (\mathrm{d}L \circ \Lambda)_{ij} + (\Lambda \circ \mathrm{d}L)_{ij} = \mathrm{d}N_{ij} - \mathrm{d}L_{ij}(\lambda_j - \lambda_i).$$

This implies that $\mathrm{d}L_{ij} = \mathrm{d}N_{ij}/(\lambda_j - \lambda_i)$. Now we compute the finite variation part $\mathrm{d}F$ of $\mathrm{d}N$,

$$\begin{aligned}\mathrm{d}F &= H^T b(X) H \mathrm{d}t + \frac{1}{2}(\mathrm{d}H^T \mathrm{d}X H + H^T \mathrm{d}X H),\\ &= b(\Lambda)\mathrm{d}t + \frac{1}{2}((\mathrm{d}N\mathrm{d}L)^T + \mathrm{d}N\mathrm{d}L).\end{aligned}$$

For $\mathrm{d}N\mathrm{d}A$ we find

$$(\mathrm{d}N\mathrm{d}L)_{ij} = \sum_{k \neq j} \mathrm{d}N_{ik}\mathrm{d}L_{kj} = \sum_{k \neq j} \frac{\mathrm{d}N_{ik}\mathrm{d}N_{kj}}{\lambda_j - \lambda_k}.$$

Now we use that the martingale part of $\mathrm{d}N$ has the same entries as $Q^T(\mathrm{d}Y)Q$ and by the results in Theorem 2.2.1 we know that

$$(Q^T M Q)_{ik}(Q^T M Q)_{kj} = \delta_{ij} G(\lambda_i, \lambda_k)\mathrm{d}t,$$

so substituting the last equality we get

$$\mathrm{d}\lambda_i = \mathrm{d}F_{ii} = b(\lambda_i)\mathrm{d}t + \sum_{k\neq j} \frac{G(\lambda_i, \lambda_k)}{\lambda_j - \lambda_k}\mathrm{d}t,$$

which is the desired result. □

These results can be particularized for any matrix-valued diffusions. Especially, we are interested in the Wishart and Jacobi processes. We state the result in these two cases as a corollary to the last theorem.

4.2.1 Wishart process

Corollary 4.2.2. *Let $Y = (Y(t), t \geq 0)$ be an $n \times n$ Wishart process with parameter m and diagonalization $Y = QMQ^T$. Let X be an $n \times n$ self-adjoint matrix process with diagonalization $X = H\Lambda H^T$, such that the off-diagonal part of $H^T \mathrm{d}XH$ and $Q^T \mathrm{d}YQ$ coincide, and $(H^T \mathrm{d}XH)_{ii} = 0$ for every $i \in [n]$. Then the eigenvalues of X, $\lambda_1 < \lambda_2 < \cdots < \lambda_n$ satisfy the following system of stochastic differential equations*

$$\mathrm{d}\lambda_i = \left(m + \sum_{k\neq i} \frac{|\lambda_i| + |\lambda_k|}{\lambda_i - \lambda_k} \right) \mathrm{d}t. \tag{4.4}$$

Proof. Apply Theorem 4.2.1 with $b(x) \equiv m, h(x) \equiv 1, g(x) = \sqrt{x}$. □

4.2.2 Jacobi process

Corollary 4.2.3. *Let $Y = (Y(t), t \geq 0)$ be an $n \times n$ matrix Jacobi process with parameters n_1, n_2 and diagonalization $Y = QMQ^T$. Let X be an $n \times n$ self-adjoint matrix process with diagonalization $X = H\Lambda H^T$, such that the off-diagonal part of $H^T \mathrm{d}XH$ and $Q^T \mathrm{d}YQ$ coincide, and $(H^T \mathrm{d}XH)_{ii} = 0$ for every $i \in [n]$. Then the eigenvalues of X, $\lambda_1 < \lambda_2 < \cdots < \lambda_n$ satisfy the following system of stochastic differential equations*

$$\mathrm{d}\lambda_i = \left(n_2 - (n_1 + n_2)\lambda_i + \sum_{k\neq i} \frac{\lambda_k(1 - \lambda_i) + \lambda_i(1 - \lambda_k)}{\lambda_i - \lambda_k} \right) \mathrm{d}t. \tag{4.5}$$

Proof. Apply Theorem 4.2.1 with $b(x) = n_2 - (n_1 + n_2)x, h(x) = \sqrt{|1 - x|}, g(x) = \sqrt{|x|}$. □

The systems of eigenvalues for the processes in this chapter follow a deterministic behavior. In the next section, we study some properties derived from the dynamics for the deterministic versions of the Dyson Brownian motion and the Wishart process.

4.3 Dynamical behavior of the deterministic eigenvalue processes

This section aims to derive some results of the dynamical behavior for the deterministic Dyson Brownian motion and the deterministic Wishart process. In particular, we are interested in

studying the possibility of collision. Since the Jacobi process is stationary, the behavior is a bit different. We do not include results about it here.

An easy result for a system of functions satisfying (4.3), is that the distance between the biggest and the smallest function always grows.

Proposition 4.3.1. *Let $\lambda_1, \dots, \lambda_n$ be a system of n functions satisfying (4.3) with b monotone increasing and initial condition $\lambda_1(0) < \cdots < \lambda_n(0)$. Then the distance between λ_n and λ_1 grows for every time before the first collision $t \leq \tau$.*

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\lambda_n(t) - \lambda_1(t)\right) > 0, \qquad t \geq 0.$$

Proof. Compute the derivative of the difference using (4.3):

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left(\lambda_n(t) - \lambda_1(t)\right) &= \frac{\mathrm{d}}{\mathrm{d}t}\lambda_n(t) - \frac{\mathrm{d}}{\mathrm{d}t}\lambda_1(t), \\ &= b(\lambda_n) + \sum_{k\neq n} \frac{G(\lambda_n, \lambda_k)}{\lambda_n - \lambda_k} - b(\lambda_1) - \sum_{k\neq 1} \frac{G(\lambda_1, \lambda_k)}{\lambda_1 - \lambda_k}, \\ &= b(\lambda_n) - b(\lambda_1) + \sum_{k\neq n} \frac{G(\lambda_n, \lambda_k)}{\lambda_n - \lambda_k} + \sum_{k\neq 1} \frac{G(\lambda_1, \lambda_k)}{\lambda_1 - \lambda_k}.\end{aligned}$$

Due to the monotonicity of b, the difference $b(\lambda_n) - b(\lambda_1)$ is non-negative. The two last terms are positive because $\lambda_n - \lambda_k$ and $\lambda_k - \lambda_1$ are positive for every $k \in [n] \setminus \{1, n\}$. □

In particular, in a system of functions following the equations of the deterministic Dyson Brownian motion or the deterministic Wishart process, the distance between the biggest and the smallest functions always grows.

4.3.1 Dyson Brownian motion

If we have a system of functions $\lambda_1, \lambda_2, \dots, \lambda_n$ that satisfy a system like (4.1) or (4.4), we will say that the functions λ_i and λ_j repel each other at time t_0 if the derivative $\frac{\mathrm{d}}{\mathrm{d}t}|\lambda_i - \lambda_j|$ exists and is positive at time t_0. The next Lemma gives us a necessary and sufficient condition for this to happen in the case of a deterministic Dyson Brownian motion.

Lemma 4.3.2. *Let $\lambda_1, \dots, \lambda_n$ be a system of n functions satisfying the system of differential equations (4.1). Take $\lambda_i(t_0) > \lambda_j(t_0)$ and such that there is no λ_k that satisfies $\lambda_i(t_0) > \lambda_k(t_0) > \lambda_j(t_0)$. Then, λ_i and λ_k repel each other at time $t = t_0$ if and only if*

$$\frac{2}{(\lambda_i(t_0) - \lambda_j(t_0))^2} > \sum_{k\neq i,j} \frac{1}{(\lambda_i(t_0) - \lambda_k(t_0))(\lambda_j(t_0) - \lambda_k(t_0))}. \tag{4.6}$$

Proof. Since there is no confusion, let us write $\lambda_i(t_0)$ as λ_i for every i. The distance between two functions grows if and only if its derivative its positive, so

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(\lambda_i - \lambda_j) &= \sum_{k\neq i} \frac{1}{\lambda_i - \lambda_k} - \sum_{k\neq j} \frac{1}{\lambda_j - \lambda_k} = \sum_{k\neq i,j} \left(\frac{1}{\lambda_i - \lambda_k} - \frac{1}{\lambda_j - \lambda_k}\right) + \frac{2}{\lambda_i - \lambda_j}, \\ &= \sum_{k\neq i,j} \frac{\lambda_j - \lambda_k - \lambda_i + \lambda_k}{(\lambda_i - \lambda_k)(\lambda_j - \lambda_k)} + \frac{2}{\lambda_i - \lambda_j}, = \frac{2}{\lambda_i - \lambda_j} - \sum_{k\neq i,j} \frac{\lambda_i - \lambda_j}{(\lambda_i - \lambda_k)(\lambda_j - \lambda_k)}.\end{aligned}$$

So this derivative is positive if and only if

$$\frac{2}{\lambda_i - \lambda_j} > \sum_{k \neq i,j} \frac{\lambda_i - \lambda_j}{(\lambda_i - \lambda_k)(\lambda_j - \lambda_k)}.$$

Finally, we divide both sides of the equality by $\lambda_i - \lambda_j$ to find the equivalent condition:

$$\frac{2}{(\lambda_I - \lambda_j)^2} > \sum_{k \neq i,j} \frac{1}{(\lambda_i - \lambda_k)(\lambda_j - \lambda_k)}.$$

□

Theorem 4.3.3. *Let $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ be a system of functions moving according to (4.1) and let λ_i, λ_j be such that for a given t_0 it is satisfied $\lambda_i(t_0) > \lambda_j(t_0)$ and*

$$\lambda_i(t_0) - \lambda_j(t_0) = \min_{k,l \in [n]} |\lambda_k(t_0) - \lambda_l(t_0)|.$$

Then, λ_i and λ_k repel.

Proof. Define the quotients c_{ik}, c_{jk} as

$$c_{ik} = \frac{\lambda_i(t_0) - \lambda_k(t_0)}{\lambda_i(t_0) - \lambda_j(t_0)}, \qquad c_{jk} = \frac{\lambda_j(t_0) - \lambda_k(t_0)}{\lambda_i(t_0) - \lambda_j(t_0)}.$$

Notice that the condition that the distance between λ_i and λ_j at time t_0 is minimal means that there is no λ_k between them. Therefore, for a fixed k_0, c_{ik_0} and c_{jk_0} have the same sign and $c_{ik_0} c_{jk_0} > 0$.

Again, let us write $\lambda_i(t_0)$ as λ_i for every i. We can write the right-hand side of (4.6) as

$$\sum_{k \neq i,j} \frac{1}{(\lambda_i - \lambda_k)(\lambda_j - \lambda_k)} = \sum_{k \neq i,j} \frac{1}{c_{ik}(\lambda_i - \lambda_j)c_{jk}(\lambda_i - \lambda_j)} = \sum_{k \neq i,j} \frac{(c_{ik}c_{jk})^{-1}}{(\lambda_i - \lambda_j)^2}.$$

With this, condition (4.6) can be written as

$$\frac{2}{(\lambda_i - \lambda_j)^2} > \sum_{k \neq i,j} \frac{(c_{ik}c_{jk})^{-1}}{(\lambda_i - \lambda_j)^2},$$

this condition is equivalent to

$$2 > \sum_{k \neq i,j} \frac{1}{c_{ik}c_{jk}}. \tag{4.7}$$

We know that the distance between λ_i and λ_j is minimal, in the worst-case scenario, all of the other distances are the same as $\lambda_i - \lambda_j$, so $|c_{ik}|, |c_{ij}| \geq 1$. Notice that if λ_{i-1} is the function located immediately below λ_i, this would mean that in the worst case scenario $c_{i,i-1} = -1$ and $c_{j,i-1} = -2$. Similarly, for λ_{i-2} we would have $c_{i,i-1} = -2, c_{j,i-2} = -3$. In general, $c_{i,i-l} = -l, c_{j,i-l} = -(l+1)$. Analogously, if λ_{j+l} is the function located l positions below j, then $c_{j,j+l} = l, c_{i,j+l} = l + 1$.

Now, for the arrangement of the functions, we have two extreme cases. If λ_i and λ_j are the functions in one of the extremes (i.e. the two biggest or two smallest ones), this would mean that the right-hand side of (4.7) can be written as

$$\sum_{k\neq i,j}\frac{1}{c_{ik}c_{jk}}=\sum_{k=1}^{n-2}\frac{1}{k(k+1)}=\sum_{k=1}^{n-2}\frac{k+1-k}{k(k+1)}=\sum_{k=1}^{n-2}\left(\frac{1}{k}-\frac{1}{k+1}\right),$$
$$=1-\frac{1}{n-1}=\frac{n-2}{n-1}<2.$$

So in this case, applying Lemma 4.3.2, λ_i, λ_j would repel each other. The other extreme case is when exactly half of the functions are located at each side of λ_i and λ_j. Let us suppose first that n is even, in this case we would have,

$$\sum_{k\neq i,j}\frac{1}{c_{ik}c_{jk}}=\sum_{k=1}^{\frac{n-2}{2}}\frac{1}{c_{i,i-k}c_{j,i-k}}+\sum_{k=1}^{\frac{n-2}{2}}\frac{1}{c_{i,j+k}c_{j,j+k}}=\sum_{k=1}^{\frac{n-2}{2}}\frac{1}{k(k+1)}+\sum_{k=1}^{\frac{n-2}{2}}\frac{1}{k(k+1)},$$
$$=2\sum_{k=1}^{\frac{n-2}{2}}\frac{1}{k(k+1)}=2\left(1-\frac{2}{n}\right)=2-\frac{4}{n}<2.$$

There is also repulsion in this case. Finally, if n is odd, $n-1$ is even and

$$\sum_{k\neq i,j}\frac{1}{c_{ik}c_{jk}}=\sum_{k=1}^{\frac{n+1-2}{2}}\frac{1}{c_{i,i-k}c_{j,i-k}}+\sum_{k=1}^{\frac{n+1-2}{2}}\frac{1}{c_{i,j+k}c_{j,j+k}}+\frac{1}{\left(\frac{n+1}{2}\right)\left(\frac{n+3}{2}\right)},$$
$$=2\left(1-\frac{2}{n+1}\right)+\frac{4}{(n+1)(n+3)}<2.$$

So, even in the worst-case scenario, criterion (4.6) is satisfied, and we conclude that the minimal distance grows for every initial condition of the system. □

If at some time t_0, it happens that $\lambda_i(t_0)=\lambda_j(t_0)$, we say that the functions λ_i and λ_j collide at time t_0. An easy consequence of Theorem 4.3.3 is that there are no collisions for a deterministic Dyson Brownian motion.

Corollary 4.3.4. *If a system of functions satisfies* (4.1) *with initial condition* $\lambda_1(0)<\lambda_2(0)<\cdots<\lambda_n(0)$*, then there are no collisions.*

Proof. Let δ be the minimal distance between two functions at time 0, i.e.,

$$\delta=\min_{k,l\in[n]}|\lambda_k(0)-\lambda_l(0)|.$$

By Theorem 4.3.3, the distance between any two functions is bigger than δ for every $t>0$. □

Although the minimal distance between two functions always grows in a system that satisfies (4.1), this is not the case for all the distances in every configuration. This result is stated in the following remark.

Remark 1. In a system that satisfies (4.1), for $n \geq 3$, not all functions repel for every given initial condition. For $n = 2$, the functions always repel each other.

Proof. We prove the first part by providing a counterexample. Suppose that λ_i, λ_j are the two biggest or two smallest functions at time t_0, and their separation is 1, while all the other separations are 0.1. The same computations that the proof of Theorem 4.3.3, but with a re-scaling of the $c_{ik}c_{jk}$ lead to

$$\sum_{k\neq i,j} \frac{1}{c_{ik}c_{jk}} = \frac{1}{0.01}\left(1 - \frac{1}{n-1}\right) = 100 - \frac{100}{n-1}.$$

Clearly, for $n \geq 3$, condition (4.6) is not satisfied. The second part is a direct consequence of Theorem 4.3.3. □

4.3.2 Wishart process

Lemma 4.3.5. *Let $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ be a system of n functions satisfying (4.4). Let λ_i, λ_j be such that $\lambda_i(t_0) > \lambda_j(t_0)$ and there is no λ_k such that $\lambda_i(t_0) > \lambda_k(t_0) > \lambda_j(t_0)$. Then λ_i and λ_j repel each other at time t_0 if and only if*

$$\frac{\lambda_i(t_0) + \lambda_j(t_0)}{(\lambda_i(t_0) - \lambda_j(t_0))^2} < \sum_{k\neq i,j} \frac{\lambda_k(t_0)}{(\lambda_i(t_0) - \lambda_k(t_0))(\lambda_j(t_0) - \lambda_k(t_0))}. \tag{4.8}$$

Proof. Take the derivative of the separation and use linearity with (4.4) to get

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}(\lambda_i - \lambda_j) &= \sum_{k\neq i} \frac{\lambda_i + \lambda_k}{\lambda_i - \lambda_k} - \sum_{k\neq j} \frac{\lambda_j + \lambda_k}{\lambda_j - \lambda_k}, \\
&= \sum_{k\neq i,j} \frac{(\lambda_i + \lambda_k)(\lambda_j - \lambda_k) - (\lambda_j + \lambda_k)(\lambda_i - \lambda_k)}{(\lambda_i - \lambda_k)(\lambda_j - \lambda_k)} + 2\frac{\lambda_i + \lambda_j}{\lambda_i - \lambda_j}, \\
&= \sum_{k\neq i,j} \frac{2\lambda_k(\lambda_j - \lambda_i)}{(\lambda_i - \lambda_k)(\lambda_j - \lambda_k)} + 2\frac{\lambda_i + \lambda_j}{\lambda_i - \lambda_j}.
\end{aligned}$$

Thus, the derivative is positive if and only if

$$\sum_{k\neq i,j} \frac{2\lambda_k(\lambda_j - \lambda_i)}{(\lambda_i - \lambda_k)(\lambda_j - \lambda_k)} > 2\frac{\lambda_i + \lambda_j}{\lambda_j - \lambda_i}.$$

Dividing both sides by $2(\lambda_j - \lambda_i)$ we get

$$\sum_{k\neq i,j} \frac{\lambda_k}{(\lambda_i - \lambda_k)(\lambda_j - \lambda_k)} > \frac{\lambda_i + \lambda_j}{(\lambda_j - \lambda_i)^2}.$$

□

Lemma 4.3.6. *Let $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ be a system of n functions satisfying (4.4) with initial condition $0 < \lambda_1(0) < \lambda_2(0) < \cdots \lambda_n(0)$. Define τ as the first collision time:*

$$\tau = \inf\{t \geq 0 \colon \lambda_i(t) = \lambda_j(t) \text{ for some } i \neq j\}.$$

Then, if $m > n-1$, the function $\lambda_1(t)$ remains greater than zero for every $0 \leq t < \tau$.

Proof. Define the function $f(t)$ as

$$f(t) \coloneqq \sup_{k \geq 2} \left\{ \frac{\lambda_k(t) + \lambda_1(t)}{\lambda_k(t) - \lambda_1(t)} \right\}.$$

For every $2 \leq k \leq n$, the quotient $(\lambda_k(t) + \lambda_1(t))/(\lambda_k(t) - \lambda_1(t))$ converges to 1 as $\lambda_1(t)$ converges to 0. There are only finitely many eigenvalues, so they converge uniformly and $f(t)$ also converges to 1 as $\lambda_1(t)$ converges to 0.

The derivative of λ_1 is

$$\frac{\mathrm{d}\lambda_1}{\mathrm{d}t} = m + \sum_{2 \leq k \leq n} \frac{\lambda_1 + \lambda_k}{\lambda_1 - \lambda_k} = m - \sum_{2 \leq k \leq n} \frac{\lambda_1 + \lambda_k}{\lambda_k - \lambda_1} \geq m - (n-1) f(t).$$

Suppose that $\lambda_1(t)$ reaches 0 at time t_1. Since $f(t)$ is continuous before the first collision time and converges to 1 as $\lambda_1(t) \to 0$, we use that $m > n-1$, and conclude that exists a t_0 such that, for $t_0 \leq t \leq t_1$, $(n-1)f(t)$ is smaller than m, i.e.

$$\frac{\lambda_1 + \lambda_k}{\lambda_k - \lambda_1} \geq m - (n-1) f(t) \geq 0, \qquad t \geq t_0.$$

So the derivative of $\lambda_1(t)$ is positive for every $t_0 \leq t \leq t_1$, but this contradicts the fact that the function decreases in this interval. The contradiction comes form supposing that $\lambda_1(t)$ reaches zero. □

Corollary 4.3.7. *In a system of n functions satisfying* (4.4) *there is no collision of the functions in finite time.*

Proof. Proposition 4.3.1 guarantees that the n functions do not collide together simultaneously. Suppose that the first collision of the system happens between $k < n$ functions at time t_1 i the point c. Let $K = \{n_1, n_2, \ldots, n_k\}$ denote the collection of indices of the functions colliding at time t_1 with $n_1 < n_2 < \cdots < n_k$. The set K consists of only consecutive numbers because t_1 is the first collision time.The derivative of the distance between the λ_{n_k} and λ_{n_1} is

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left[\lambda_{n_k}(t) - \lambda_{n_1}(t)\right] &= m + \sum_{j \neq n_k} \frac{\lambda_{n_k} + \lambda_j}{\lambda_{n_k} - \lambda_j} - m - \sum_{j \neq n_1} \frac{\lambda_{n_1} + \lambda_j}{\lambda_{n_1} - \lambda_j}, \\
&= \sum_{j \neq n_k} \frac{\lambda_{n_k} + \lambda_j}{\lambda_{n_k} - \lambda_j} - \sum_{j \neq n_1} \frac{\lambda_{n_1} + \lambda_j}{\lambda_{n_1} - \lambda_j}, \\
&= \sum_{j \in K} \left\{ \frac{\lambda_{n_k} + \lambda_j}{\lambda_{n_k} - \lambda_j} + \frac{\lambda_{n_1} + \lambda_j}{\lambda_j - \lambda_{n_1}} \right\} + \sum_{j \notin K} \left\{ \frac{\lambda_{n_k} + \lambda_j}{\lambda_{n_k} - \lambda_j} + \frac{\lambda_{n_1} + \lambda_j}{\lambda_j - \lambda_{n_1}} \right\}.
\end{aligned}$$

In the last expression, the first sum is positive because for $j \in K'$, the function $\lambda_j(t)$ is greater than $\lambda_{n_1}(t)$ and smaller than $\lambda_{n_k}(t)$ before the first collision.

Denote by $d(t)$ the distance between λ_{n_k} and λ_{n_1}:

$$d(t) = \lambda_{n_k}(t) - \lambda_{n_1}(t).$$

Since we are supposing the collision at time t_1, the limit of d as t tends to t_1 is 0. Then for every $\epsilon > 0$, there exists a $t_\epsilon < t_1$ such that $d(t) < \epsilon$ for every $t_\epsilon \leq t \leq t_1$. The distance between any other functions $\lambda_j(t)$ with index $j \in K$ is bounded by $d(t)$ because $d(t)$ is the distance of the extremes. So, using the last equality we find:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left[\lambda_{n_k}(t) - \lambda_{n_1}(t)\right] &\geq \sum_{j\in K}\left\{\frac{\lambda_{n_k}+\lambda_j}{d(t)} + \frac{\lambda_{n_1}+\lambda_j}{d(t)}\right\} + \sum_{j\notin K}\left\{\frac{\lambda_{n_k}+\lambda_j}{\lambda_{n_k}-\lambda_j} + \frac{\lambda_{n_1}+\lambda_j}{\lambda_j-\lambda_{n_1}}\right\},\\ &\geq \frac{4\lambda_1(t)(k-1)}{d(t)} + \sum_{j\notin K}\left\{\frac{\lambda_{n_k}+\lambda_j}{\lambda_{n_k}-\lambda_j} + \frac{\lambda_{n_1}+\lambda_j}{\lambda_j-\lambda_{n_1}}\right\}.\end{aligned}$$

Now, take $b(t)$ to be the minimal distance between a function outside the collision and a function in the collision, i.e.,

$$b(t) \coloneqq \inf_{j\in K, i\notin K}\{|\lambda_j(t) - \lambda_i(t)|\}.$$

Using the continuity and the fact that the functions λ_j for $j \notin K$ do not collide until time t_1, we have that there exists a $\delta > 0$ such that $b(t) > \delta$ for every $t \in [0, t_1]$. By the same continuity argument, all the functions $\lambda_i(t)$ for $i \in [k]$ are bounded until time t_1. Take $M > 0$ to be a universal bound for all of them. With this, we can conclude

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left[\lambda_{n_k}(t) - \lambda_{n_1}(t)\right] &\geq \frac{4\lambda_1(t)(k-1)}{d(t)} - \sum_{j\notin K}\left\{\frac{\lambda_{n_k}+\lambda_j}{b(t)} + \frac{\lambda_{n_1}+\lambda_j}{b(t)}\right\},\\ &\geq \frac{4\lambda_1(t)(k-1)}{d(t)} - \frac{4M(n-k)}{\delta}.\end{aligned}$$

Taking an appropriate ϵ, we can make the last quantity bigger than zero for t between t_ϵ and t_1. This means that the derivative of the difference between λ_{n_k} and λ_{n_1} is positive in (t_ϵ, t_1) and this contradicts the fact that the distance decreases to zero. □

Proposition 4.3.8. *Let $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ be a system of n functions moving according to* (4.4) *and let λ_i, λ_j be such that for a given t_0 it is satisfied $\lambda_i(t_0) > \lambda_j(t_0)$ and*

$$\lambda_i(t_0) - \lambda_j(t_0) = \min_{k,l\in[n]} |\lambda_k(t_0) - \lambda_l(t_0)|.$$

Then, λ_i and λ_j do not necessarily repel.

Proof. We will prove the result by providing an initial condition for which the minimal distance will have a negative derivative. Let $\lambda_j(t_0) = \min_{k\leq n} \lambda_k(t_0)$ and so $\lambda_i(t_0) = \min_{k\neq j} \lambda_k(t_0)$. Similarly to the proof of Theorem 4.3.3, let us denote $\lambda_i(t_0)$ as λ_i along the proof and define the quotients c_{ik}, c_{jk} as

$$c_{ik} = \frac{\lambda_i - \lambda_k}{\lambda_i - \lambda_j}, \qquad c_{jk} = \frac{\lambda_j - \lambda_k}{\lambda_i - \lambda_j}.$$

For every fixed k, the quotients c_{ik} and c_{jk} have the same sign and given the minimality of $\lambda_i - \lambda_j$ we have that $c_{ik}c_{jk} > 1$. Using these quantities, the separation condition (4.8) is reduced to

$$\lambda_i + \lambda_j > \sum_{k \neq i,j} \frac{\lambda_k}{c_{ik}c_{jk}}.$$

Take λ_i, λ_j such that $\lambda_i - \lambda_j = 1$ at $t = t_0$, and that $\lambda_k > \lambda_i$ for all $k \notin \{i, j\}$. Take the separation between all the $(\lambda_k)_{k \neq i,j}$ to be $1 + \epsilon$ for some $\epsilon > 0$. Then the coefficients $c_{i,i+1}, c_{j,i+1}$ are equal to

$$c_{i,i+1} = \frac{\lambda_i - \lambda_{i+1}}{\lambda_i - \lambda_j} = -\frac{1+\epsilon}{1}, \qquad c_{j,i+1} = \frac{\lambda_j - \lambda_{i+1}}{\lambda_i - \lambda_j} = -\frac{2+\epsilon}{1}.$$

In general, $|c_{i,i+l}| = l(1+\epsilon), |c_{j,i+l}| = 1 + l(1+\epsilon)$. Furthermore, we have that λ_{i+l} can be expressed as

$$\lambda_{i+l} = \lambda_i + l(1+\epsilon) = \lambda_j + 1 + l(1+\epsilon).$$

With this, the separating condition for λ_i and λ_j can be written as

$$2\lambda_j + 1 > \sum_{k=2}^{n} \frac{\lambda_j + 1 + k(1+\epsilon)}{k(1+\epsilon)(1+k(1+\epsilon))}.$$

The left-hand side is fixed for fixed $\lambda_j(t_0)$. For the left-hand side, we have

$$\sum_{k=2}^{n} \frac{\lambda_j + 1 + k(1+\epsilon)}{k(1+\epsilon)(1+k(1+\epsilon))} = \sum_{k=2}^{n} \frac{\lambda_j + 1}{k(1+\epsilon)(1+k(1+\epsilon))} + \sum_{k=2}^{n} \frac{1}{1+k(1+\epsilon)}.$$

For the second element in the sum and ϵ sufficiently small we have

$$\sum_{k=2}^{n} \frac{1}{1+k(1+\epsilon)} > \sum_{k=2}^{n} \frac{1}{1+2k} > \frac{1}{2}\sum_{k=2}^{n} \frac{1}{1+k}.$$

The last expression can be made arbitrarily big for n big enough. We conclude that under these conditions, for a system with enough functions, the minimal distance can be made smaller with a specific initial condition.

□

4.4 Path simulations

With these simulations, we exemplify the dynamical behavior of the system of functions explained in the last section. In Figure 4.1 we have the analogous to Figure 2.2. The initial conditions are uniformly spaced points around zero for the last three examples. For the first example, the distance between the two points in the middle is two, while the top and bottom points are separated by 0.1 of the two in the middle. We can notice that in concordance with Remark 1, the bigger value decreases for some time until the particle spacing is more

uniform and then it starts growing. For the functions with a small separation, this distance grows very quickly at the beginning, as expected by Theorem 4.3.3.

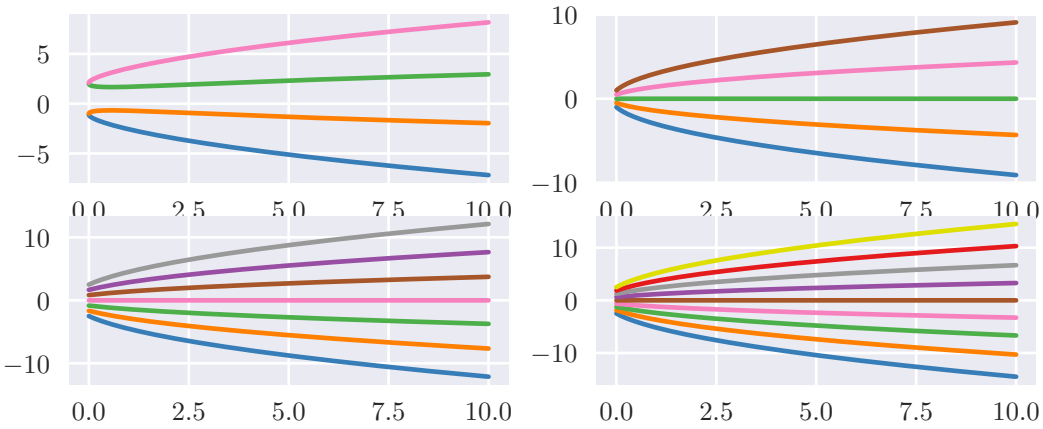

Figure 4.1: Simulation of four different deterministic Dyson Brownian motions with different dimensions. The initial conditions from left to right are $(-1.1, -1, 2, 2.1)$, $(-1, -0.5, 0, 0.5, 1)$, $(-2.5, -1.67, -0.83, 0, 0.83, 1.67, 2.5)$ and $(-2.5, -1.875, -1.25, -0.625, 0, 0.625, 1.25, 1.875, 2.5)$.

In the remaining three panels of Figure 4.1, we also see a behavior according to 4.3.3. Also notice that (4.1) implies that if we start with a configuration that is evenly spaced, then the points will roughly stay evenly spaced through time, which is precisely what we can see.

In Figure 4.2, we have the simulation analogous to 2.3 with a nine-dimensional deterministic Wishart process. Apart from observing that the total distance of the particles grows, we check that this example confirms Proposition 4.3.8, as the distance between the two paths in the bottom is minimal at some point and it continues to decrease. However, the biggest is the value of the function and the surrounding ones, the faster their spacing grows, as we can notice for the two functions at the top.

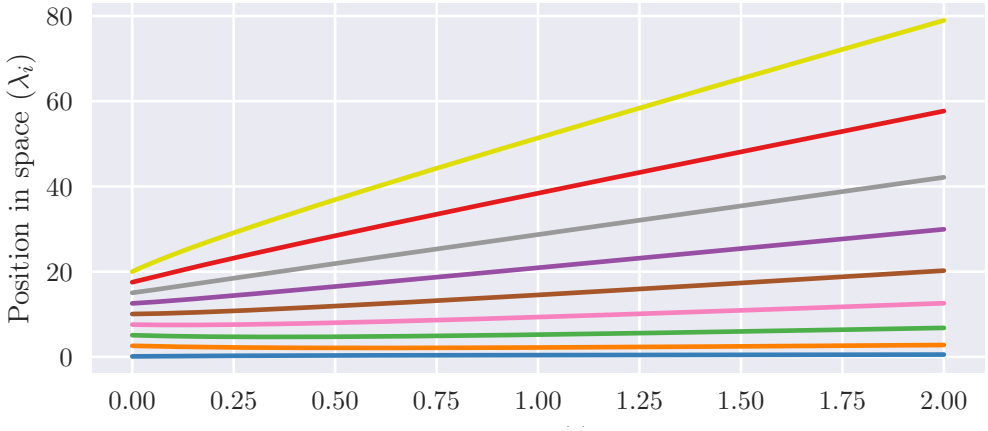

Figure 4.2: Simulation of the path of a deterministic Wishart process. The initial conditions are $(0.1, 2.6, 5.1, 7.6, 10.1, 12.6, 15.1, 17.6, 20.1)$.

For the path simulations of the deterministic Jacobi process, we observe that for $t = 0.05$

they are almost evenly spaced in $[0, 1]$. We would expect this deterministic process to converge to the equilibrium which would be the roots of the ninth Jacobi polynomial with parameters $n_1 = 1, n_2 = 2$.

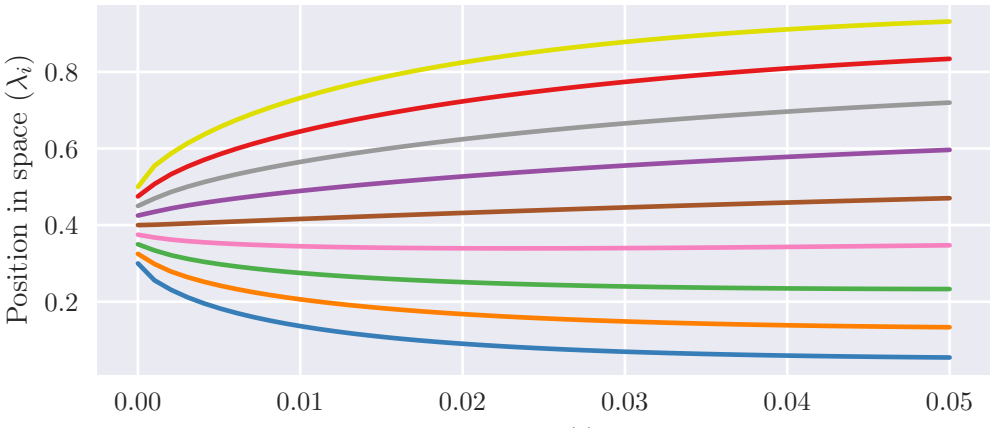

Figure 4.3: Simulation of the path of a deterministic Jacobi process. The initial conditions are $(0.3, 0.325, 0.35, 0.375, 0.4, 0.425, 0.45, 0.475, 0.5)$.

4.5 Connections with finite free probability

In this section we relate finite free probability to eigenvalue processes. To do so, we first find the expected characteristic polynomial of the random matrices. We show that in some cases they are well-known polynomials whose convolutions satisfy good properties. We start with the case of a self-adjoint Brownian matrix and then get similar results for the Wishart and Jacobi processes.

Before proceeding with the processes, let us state a lemma that will help us to find the expected characteristic polynomials by using a tridiagonal model.

Lemma 4.5.1. *Let A be an $n \times n$ tridiagonal symmetric matrix with diagonal elements $\{a_i\}_{1\leq i\leq n}$ and subdiagonal elements $\{b_j\}_{2\leq j\leq n}$, such that*

$$A = \begin{bmatrix} a_n & b_{n-1} & 0 & \cdots & 0 \\ b_{n-1} & a_{n-1} & b_{n-2} & \cdots & 0 \\ 0 & b_{n-2} & a_{n-2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & b_1 & a_1 \end{bmatrix}.$$

Then its characteristic polynomial $\chi_z^n(A) = \det(zI_n - A)$ satisfies the following recursion

$$\chi_z^n(A) = (z - a_n)\chi_z^{n-1}(A) - b_{n-1}^2\chi_z^{n-2}(A),$$

where $\chi_z^{n-1}(A)$ represents the characteristic polynomial of the $(n-1) \times (n-1)$ lower-right block of A. Notice that, in particular, this block is also tridiagonal.

Proof. Let us write

$$A = \begin{bmatrix} a_n & b_{n-1}e_1^T \\ b_{n-1}e_1 & B \end{bmatrix},$$

with e_1 the first element in the canonical base of $\mathbb{R}^{n-1}$, and write B as

$$B = \begin{bmatrix} a_{n-1} & b_{n-2}e_1^{(n-2)^T} \\ b_{n-2}e_1^{(n-2)} & C \end{bmatrix}.$$

The determinant of $zI_n - A$ is

$$\begin{aligned} \det(zI_n - A) &= (zI_n - A)_{11}\det(zI_{n-1} - B) - (zI_n - A)_{12}(zI_{n-2} - A)_{21}\det(zI_n - C), \\ &= (z - a_n)\chi_z^{n-1}(A) - b_{n-1}^2\chi_z^{n-2}(A). \end{aligned}$$

□

With this lemma and the tools developed in the former chapters, we are ready to study the expected characteristic polynomial of some matrix-valued stochastic processes. In the next subsection, we start with the Dyson Brownian motion.

4.5.1 Dyson Brownian motion

In Chapter 3 we saw that the tools devloped by Marcus and collaborators [3], [32] allow us to express expected characteristic polynomials of sums and products of random matrices as convolutions of the original polynomials. However, we still did not see any tool useful for finding such polynomials when we can not express them in terms of another characteristic polynomial. This topic has been covered in the literature several times [33]–[35] but we will use the approach by Ioanna Dumitriu and Alan Edelman found in [36] because it requires to introduce less technical concepts. Once we find the expected characteristic polynomials by these means, we will relate them to Finite Free Probability Theory using the previous developed tools.

As a first step, we will make use of the invariance under orthogonal (or unitary) transformations of the GOE (or GUE) in order to find a tridiagonal model with the same matrix distribution.

Lemma 4.5.2. *Let A_β denote an $n \times n$ Gaussian invariant ensemble, GOE for $\beta = 1$, GUE for $\beta = 2$ and GSE for $\beta = 4$. Then the eigenvalues of the tridiagonal matrix H_β have the same joint law as the eigenvalues of A_β, with G_β defined as*

$$G_\beta = \frac{1}{\sqrt{2}} \begin{bmatrix} N_1 & \xi_2 & 0 & \cdots & 0 \\ \xi_2 & N_2 & \xi_3 & \cdots & 0 \\ 0 & \xi_3 & N_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \xi_n & N_n \end{bmatrix}. \tag{4.9}$$

The diagonal entries $N_i, 0 \leq i \leq n$ are independent normal random variables with mean 0 and variance 2, while the subdiagonal entries ξ_i are independent random variables distributed as

$$\xi_i \sim \chi_{\beta(n+1-i)}.$$

Note that in this case χ_ν denotes the chi distribution which is the squared root of a chi-squared random variable or the absolute value of a normal random variable in the case ν is integer.

We call H_β the tridiagonal β-Hermite ensemble.

Proof. Write A_β as

$$A_\beta = \begin{bmatrix} N_1 & \vec{x}^T \\ \vec{x} & B_\beta \end{bmatrix},$$

with N_1 a normal random variable, $\vec{x}$ an $n-1$-dimensional Gaussian vector with independent standard Gaussian entries in $\mathbb{R}, \mathbb{C}$ or $\mathbb{H}$ depending on β, and B_β an $(n-1)\times(n-1)$ GOE, GUE or GSE, for $\beta = 1, 2, 4$, respectively. All of the elements are independent of each other.

Now we take $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{n-2}$ to be any vectors independent of x with $\|\vec{v}_i\|_2 = \|\vec{x}\|_2$ for every i. Define a matrix Q such that

$$Q \coloneqq \frac{1}{\|\vec{x}\|} \begin{bmatrix} \vec{x}^T \\ \vec{v}_1^T \\ \vdots \\ \vec{v}_n^T \end{bmatrix}.$$

Then Q is an $(n-1)\times(n-1)$ orthogonal (or unitary, symplectic, according to β) matrix such that $Q\vec{x}^T = \|\vec{x}\|_2 e_1$, where $e_1 = (1, 0, \cdots, 0)$ is the first element in the canonical basis of $\mathbb{R}^{n-1}$. With this, we have

$$\begin{bmatrix} 1 & \vec{0}^T \\ \vec{0} & Q \end{bmatrix} \begin{bmatrix} N_1 & \vec{x}^T \\ \vec{x} & B_\beta \end{bmatrix} \begin{bmatrix} 1 & \vec{0}^T \\ \vec{0} & Q^T \end{bmatrix} = \begin{bmatrix} N_1 & \vec{x}^T \\ \|\vec{x}\|_2 e_1 & QB_\beta \end{bmatrix} \begin{bmatrix} 1 & \vec{0}^T \\ \vec{0} & Q^T \end{bmatrix} = \begin{bmatrix} N_1 & \|x\|_2 e_1^T \\ \|x\|_2 e_1 & QB_\beta Q^T \end{bmatrix}.$$

Now we can find the distribution of each of the blocks of the new matrix. The variable N_1 has not changed, it is a standard normal variable. The term $\|\vec{x}\|_2$ is the norm of a Gaussian vector of length $n-1$ with real (complex or quaternionic) entries, non-correlated and with variance $1/2$, so it is distributed like a $\frac{1}{\sqrt{2}}\chi_{\beta(n-1)}$ random variable, where β indicates the number of normal variables in each entry of the matrix. Since B_β is a GOE (GUE, GSE), it is invariant under orthogonal (unitary, symplectic) transformations and thus $QB_\beta Q^T$ is a GOE (GUE, GSE).

The matrix $\begin{bmatrix} 1 & \vec{0}^T \\ \vec{0} & Q \end{bmatrix}$ is orthogonal (unitary, symplectic), so the eigenvalue distribution of A_β remains unchanged under this transformation. By repeating the procedure with B_β, we find the tridiagonal matrix (4.9), which finishes the proof. □

Since G_β has the same distribution as an invariant Gaussian ensemble, the expected characteristic polynomials are the same, and we can find them using that G_β is tridiagonal and Lemma 4.5.1.

Theorem 4.5.3. *Let X be an $n \times n$ GUE, then its expected characteristic polynomial, $P_n(z)$ is the nth Hermite polynomial.*

Proof. Let A_2 be a tridiagonalization of X. Due to Lemma 4.5.2 we know that the eigenvalues of A_2 have the same joint distribution of those of X, so the expected characteristic polynomial must coincide. We can use Lemma 4.5.1 to find the expected characteristic polynomial of A. Let Q_n denote its characteristic polynomial, then

$$\begin{aligned} P_n(z) &= \mathbb{E}\left(\left(z - \frac{N_1}{\sqrt{2}}\right) Q_{n-1}(z) - \frac{\xi_{2(n-1)}^2}{2} Q_{n-2}(z)\right), \\ &= zP_{n-1}(z) - (n-1)P_{n-2}(z). \end{aligned}$$

So $P_n(z)$ satisfies the recursion that determines the Hermite polynomials. We need to check the initial conditions $P_1(z)$ and $P_2(z)$. For $P_1(z)$ the condition is trivial

$$P_1(z) = \mathbb{E}\left(z - \frac{N_1}{\sqrt{2}}\right) = z.$$

For $P_2(z)$, we have

$$\begin{aligned} P_2(z) &= \mathbb{E}\left(\det \frac{1}{\sqrt{2}} \begin{bmatrix} z - N_1 & \xi_2 \\ \xi_2 & z - N_2 \end{bmatrix}\right) = \mathbb{E}\left[\left(z - \frac{N_1}{\sqrt{2}}\right)\left(z - \frac{N_2}{\sqrt{2}}\right) - \frac{\xi_2^2}{2}\right], \\ &= z^2 - 1. \end{aligned}$$

So $P_1(z) = H_1(z)$ and $P_2(z) = H_2(z)$. Using the recursion, we can conclude that $P_n(z) = H_n(z)$. □

Recall that a self adjoint Brownian motion has the same law as $\sqrt{t}$ times a Gaussian invariante ensemble. Once again, we use an equality in distribution and Lemma 4.5.1 in order to find the expected characteristic polynomial of a self adjoint Brownian motion.

Corollary 4.5.4. *Let $B(t)$ be an $n \times n$ self-adjoint complex Brownian matrix, then its expected characteristic polynomial $P_n(z,t)$ is the nth generalized Hermite polynomial with variance t, $H_n(z,t)$, i.e. the Hermite polynomials which are orthogonal with respect to a Gaussian random variable of variance t.*

Proof. Find the expected characteristic polynomial by definition and use the equality in distribution of $B(t)$ with A_2, a GUE;

$$\mathbb{E}\left[\det[zI - B(t)]\right] = \mathbb{E}\left[\det[zI - \sqrt{t}A_2]\right] = t^{-\frac{n}{2}}\mathbb{E}\left[\det[zt^{\frac{1}{2}}I - A_2]\right],$$

using Theorem 4.5.3

$$\begin{aligned} &= t^{-\frac{n}{2}} H_n(zt^{\frac{1}{2}}) = t^{-\frac{n}{2}} n! \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \frac{(-1)^k (zt^{\frac{1}{2}})^{n-2k}}{2^k k!(n-2k)!}, \\ &= n! \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \frac{(-t)^k z^{n-2k}}{2^k k!(n-2k)!} = H_n(z,t). \end{aligned}$$

□

Now that we know that the expected characteristic polynomial of a self adjoint Brownian motion is a Hermite polynomial with variance t, we are ready to give a much easier proof of Proposition 3.1.2 using the tools of Finite Free Probability Theory.

Alternate proof of Proposition 3.1.2. Let $H_n(z,t_1)$ and $H_n(z,t_2)$ be the expected characteristic polynomials of two independent self-adjoints Brownian motions $B_1 = (B_1(t), t \geq 0)$ and $B_2 = (B_2(t), t \geq 0)$ evaluated at the times t_1 and t_2, respectively. By the Lévy property of the Brownian motion, we know that $B \coloneqq B_1 + B_2$ is another self-adjoint Brownian motion, evaluated at the time $t_1 + t_2$. By Theorem 3.2.8, and using the invariance under unitary transformations of B_1, B_2,

$$H_n(z,t_1) \boxplus_n H_n(z,t_2) = \mathbb{E}\left[\chi_z(B_1(t) + QB_2(t)Q^*)\right] = \mathbb{E}\left[\chi_z(B_1(t) + B_2(t))\right] = \mathbb{E}\left[\chi_z(B(t))\right].$$

A last use of Corollary 4.5.4 allows us to conclude that this polynomial is $H_n(z, t_1 + t_2)$. □

The Hermite polynomials not only are the expected characteristic polynomials for our Gaussian invariant ensembles, they somehow also carry all the "non-stochastic information" of the eigenvalues in the sense that their roots satisfy (4.1). We will prove this by first showing that the Hermite polynomials are "harmonic" in some sense, i.e., they are a solution for the martingale problem of a Brownian motion in reversed time.

Theorem 4.5.5. *The Hermite polynomials $H_n(z,t)$ solve the differential equation*

$$\partial_t[H_n(z,t)] + \frac{1}{2}\partial_{zz}[H_n(z,t)] = 0, \tag{4.10}$$

and the Cauchy transform $G_{H_n}(y)$ of the empirical measure associated to its roots $\{z_j(t)\}_{j\in[n]}$,

$$G_{H_n}(z,t) \coloneqq \frac{1}{n}\sum_{j=1}^{n} \frac{1}{z_j(t) - z},$$

solves the viscous Burgers equation with diffusion coefficient $-1/2$,

$$\partial_t[\partial G_{H_n}(z,t)] + nG_{H_n}(z,t)\partial_z[G_{H_n}(z,t)] = -\frac{1}{2}\partial_{zz}[G_{H_n}(z,t)].$$

Proof. Lets us write $H_n(z,t) = H_n$, $G_{H_n}(z,t) = G$ and $\frac{\partial f}{\partial z} = \partial_z f$ to simplify the notation.

First, we prove that $H_n(z,t)$ satisfies (4.10). Write $H_n(z,t) = \exp\left\{-\frac{t\partial_z^2}{2}\right\}[z^n]$, where

$$\exp\left\{-\frac{t\partial_z^2}{2}\right\}[z^n] \coloneqq \sum_{j=0}^{\infty} \frac{\left(-t\partial_z^2\right)^j}{2^j j!}[z^n].$$

Then

$$\begin{aligned}
\partial_t H_n &= \partial_t \exp\left\{-\frac{t\partial_z^2}{2}\right\}(z^n) = \partial_t \sum_{j=0}^{\lfloor n/2\rfloor}(-1)^j \frac{t^j \partial_z^{2j}(z^n)}{2^j j!},\\
&= \sum_{j=0}^{\lfloor n/2\rfloor}(-1)^j \frac{\partial_t(t^j)\partial_z^{2j}(z^n)}{2^j j!} = \sum_{j=0}^{\lfloor n/2\rfloor}(-1)^j \frac{j(t^{j-1})\partial_z^{2j}(z^n)}{2^j j!},\\
&= \sum_{j=1}^{\lfloor n/2\rfloor}(-1)^j \frac{j(t^{j-1})\partial_z^{2j}(z^n)}{2^j j!} = -\frac{1}{2}\sum_{j=1}^{\lfloor n/2\rfloor}(-1)^{j-1} \frac{(t^{j-1})\partial_z^{2(j-1)}\partial_{zz}(z^n)}{2^{j-1}(j-1)!},\\
&= -\frac{1}{2}\partial_{zz}\sum_{k=0}^{\lfloor n/2-1\rfloor}(-1)^k \frac{(t^k)\partial_z^{2k}(z^n)}{2^k k!} = -\frac{1}{2}\partial_{zz}H_n.
\end{aligned}$$

For the Cauchy transform part, we recall from Lemma 3.3.1 that since H_n is monic on z, then $G = \frac{\partial_z H_n}{nH_n}$.

Now, let us show that G satisfies (4.5.5), for this we compute $\partial_t G$, $G\partial_z G$ and $\partial_{zz} G$,

$$\begin{aligned}
\partial_t G &= \frac{1}{n}\partial_t\left(\frac{\partial_z H_n}{H_n}\right) = \frac{1}{n}\frac{H_n\partial_t\partial_z H_n - \partial_z H_n \partial_t H_n}{H_n^2},\\
\partial_z G &= \frac{1}{n}\partial_z\left(\frac{\partial_z H_n}{H_n}\right) = \frac{1}{n}\frac{H_n\partial_{zz} H_n - (\partial_z H_n)^2}{H_n^2},\\
\partial_{zz} G &= \frac{1}{n}\partial_z\left(\frac{H_n\partial_{zz}H_n - (\partial_z H_n)^2}{H_n^2}\right),\\
&= \frac{1}{n}\frac{H_n^2\partial_z H_n\partial_{zz}H_n + H_n^3\partial_{zzz}H_n - 2H_n^2\partial_z H_n\partial_{zz}H_n}{H_n^4}\\
&\qquad\qquad - \frac{1}{n}\frac{2H_n^2\partial_z H_n\partial_{zz}H_n - 2H_n(\partial_z H_n)^3}{H_n^4},\\
&= \frac{1}{n}\frac{-3H_n\partial_z H_n\partial_{zz}H_n + H_n^2\partial_{zzz}H_n + 2(\partial_z H_n)^3}{H_n^3}\\
G\partial_z G &= \frac{\partial_z H_n}{nH_n}\left(\frac{H_n\partial_{zz}H_n - (\partial_z H_n)^2}{nH_n^2}\right) = \frac{1}{n^2}\frac{H_n\partial_z H_n\partial_{zz}H_n - (\partial_z H_n)^3}{H_n^3}.
\end{aligned}$$

Finally, we can use the above results to find

$$\begin{aligned}
\partial_t G + nG\partial_z G &= \frac{1}{n}\left(\frac{H_n\partial_t\partial_z H_n - \partial_z H_n\partial_t H_n}{H_n^2} + \frac{H_n\partial_z H_n\partial_{zz}H_n - (\partial_z H_n)^3}{H_n^3}\right),\\
&= \frac{1}{n}\left(\frac{-\frac{1}{2}H_n^2\partial_{zzz}H_n + \frac{1}{2}H_n\partial_z H_n\partial_{zz}H_n + H_n\partial_z H_n\partial_{zz}H_n - (\partial_z H_n)^3}{H_n^3}\right),\\
&= \frac{1}{n}\left(\frac{\frac{3}{2}H_n^2\partial_z H_n\partial_{zz}H_n - \frac{1}{2}H_n^2\partial_{zzz}H_n - (\partial_z H_n)^3}{H_n^3}\right) = -\frac{1}{2}\partial_{zz}G.
\end{aligned}$$

□

Using Theorem 4.5.5 its is easy to prove that the roots of the time dependent Hermite polynomials evolve according to (4.1).

Corollary 4.5.6. *The roots $\{z_i(t)\}_{i\leq n}$ of the Hermite polynomials $H_n(z,t)$ satisfy the deterministic Dyson's equation,*

$$\mathrm{d}z_i = \sum_{k\neq i} \frac{\mathrm{d}t}{z_i - z_k}.$$

Proof. Let $z_i(t)$ be the roots of $H_n(z,t)$, this means that $H_n(z_i(t),t) = 0$, which implies

$$\partial_t[H_n(z_i(t),t)] = 0.$$

By the chain rule and the heat kind equation (4.10) in Theorem 4.5.5 we have that

$$0 = \partial_t[H_n(z_i(t),t)] = (\partial_t[z_i(t)])\partial_z[H_n(z_i(t),t)] - \frac{1}{2}\partial_{zz}[H_n(z_i(t),t)].$$

Using the product rule, for the derivative of products, we have

$$\begin{aligned}
\partial_z[H_n(z_i(t),t)] &= \partial_z\left[\prod_{j=1}^{n}(z-z_j)\right]\Bigg|_{z=z_i(t)} = \sum_{k=1}^{n}\prod_{j\neq k}(z-z_j)\Bigg|_{z=z_i(t)} \\
&= \prod_{j\neq i}(z_i - z_j). \\
\partial_{zz}[H_n(z_i(t),t)] &= \partial_z\left[\sum_{k=1}^{n}\prod_{j\neq k}(z-z_j)\right]\Bigg|_{z=z_i(t)} = \sum_{k=1}^{n}\sum_{j\neq k}\prod_{l\neq j,k}(z-z_l)\Bigg|_{z=z_i(t)}, \\
&= 2\sum_{j\neq i}\prod_{l\neq j,i}(z_i - z_l).
\end{aligned}$$

Finally, if we substitute this in the previously found expression, we find

$$\frac{\mathrm{d}}{\mathrm{d}t}z_i(t) = \frac{\partial_{zz}[H_n(z_i(t),t)]}{2\partial_z[H_n(z_i(t),t)]} = \frac{2\sum_{k\neq i}\prod_{j\neq i,j\neq k}(z_i - z_j)}{2\prod_{j\neq i}(z_i - z_j)} = \sum_{k\neq i}\frac{1}{z_i - z_k}.$$

□

As a final part for this subsection, let us prove that if you have an arbitrary (real valued) initial condition for a self adjoint Brownian motion, then its expected characteristic polynomial will also satisfy (4.10). Notice that as a consequence, this means that its roots will move according to (4.1).

Theorem 4.5.7. *Let A be an $n\times n$ fixed self-adjoint matrix and W an $n\times n$ self-adjoint Brownian matrix, then $q_A(z,t)$ defined as*

$$q_A(z,t) \coloneqq \mathbb{E}\left[\chi_z(A+W)\right],$$

satisfies the following differential equation

$$\partial_t[q_A(z,t)] + \frac{1}{2}\partial_{zz}[q_A(z,t)] = 0.$$

Proof. Let $p(z) = \mathbb{E}[\chi_z(A)]$ and $r(z,t) = \mathbb{E}[\chi_z(W)]$. Corollary 4.5.4 tells us that $r(z,t)$ is the nth Hermite polynomial $H_n(z,t)$. Write these polynomials as

$$p(z) = \sum_{j=0}^{n} a_j z^j,$$
$$H_n(z,t) = \sum_{j=0}^{n} b_j t^{j/2} z^{d-j}.$$

Notice that, from the explicit representation of the Hermite polynomials in Section 3.1, if n is odd all of the b_j are zero for even j and if n is even, all of the b_j are zero for odd j. Further, using the explicit expression for the coefficients we have the following recursion for b_j

$$b_j = \frac{n!(-1)^{j/2}}{2^{j/2}(j/2)!(n-j)!},$$
$$b_{j-2} = \frac{n!(-1)^{\frac{j-2}{2}}}{2^{\frac{j-2}{2}}\left(\frac{j-2}{2}\right)!(n-j+2)!},$$

this leads to

$$\begin{aligned} b_j &= \frac{n!(-1)^{\frac{j-2}{2}}}{2^{\frac{j-2}{2}}\left(\frac{j-2}{2}\right)!(n-j+2)!}\frac{(-1)(n-j+2)(n-j+1)}{j}, \\ &= b_{j-2}\frac{(-1)(n-j+2)(n-j+1)}{j}. \end{aligned} \tag{4.11}$$

Using the invariance of W under unitary transforms and Theorem 3.2.8 we have that $q_A(z,t) = H_n(z,t) \boxplus_n p(z)$. By definition,

$$q_A(z,t) = H_n(z,t) \boxplus_n p(z) \coloneqq \sum_{k=0}^{n} z^{n-k}(-1)^k \sum_{i=0}^{k} c_{k,i,n} b_i a_{k-i} t^{i/2},$$

with $c_{k,i,n} = \frac{(n-i)!(n-k+i)!}{n!(n-k)!}$.

We compute first $\partial_{zz}[q_A(z,t)]$,

$$\begin{aligned}
\partial_{zz}[q_A(z,t)] &= \partial_{zz}\left[\sum_{k=0}^{n} z^{n-k}(-1)^k \sum_{i=0}^{k} \frac{(n-i)!(n-k+i)!}{n!(n-k)!} b_i a_{k-i} t^{i/2}\right], \\
&= \sum_{k=0}^{n-2}(n-k)(n-k-1)z^{n-k-2}(-1)^k \sum_{i=0}^{k} \frac{(n-i)!(n-k+i)!}{n!(n-k)!} b_i a_{k-i} t^{i/2}, \\
&= \sum_{k=0}^{n-2} z^{n-k-2}(-1)^k \sum_{i=0}^{k} \frac{(n-i)!(n-k+i)!}{n!(n-k-2)!} b_i a_{k-i} t^{i/2}, \\
&= \sum_{k=2}^{n} z^{n-k}(-1)^k \sum_{i=0}^{k-2} \frac{(n-i)!(n-(k-2)+i)!}{n!(n-(k-2)-2)!} b_i a_{k-2-i} t^{i/2}, \\
&= \sum_{k=2}^{n} z^{n-k}(-1)^k \sum_{i=0}^{k-2} \frac{(n-i)!(n-k+i+2)!}{n!(n-k)!} b_i a_{k-2-i} t^{i/2}.
\end{aligned}$$

Now the derivative with respect to t is

$$\begin{aligned}
\partial_t[q_A(z,t)] &= \partial_t\left[\sum_{k=0}^{n} z^{n-k}(-1)^k \sum_{i=0}^{k} c_{k,i,n} b_i a_{k-i} t^{i/2}\right], \\
&= \sum_{k=0}^{n} z^{n-k}(-1)^k \sum_{i=0}^{k} \frac{(n-i)!(n-k+i)!}{n!(n-k)!} b_i a_{k-i} \frac{i}{2} t^{\frac{i-2}{2}}, \\
&= \sum_{k=0}^{n} z^{n-k}(-1)^k \sum_{j=0}^{k-2} \frac{(n-j-2)!(n-k+j+2)!}{n!(n-k)!} b_{j+2} a_{k-j-2} \left(\frac{j+2}{2}\right) t^{j/2}, \\
&= \sum_{k=0}^{n} z^{n-k}(-1)^k \sum_{j=0}^{k-2} \frac{(n-j)!(n-k+j+2)!}{n!(n-k)!} b_j a_{k-j-2} \frac{(j+2)t^{j/2}}{2},
\end{aligned}$$

using (4.11) for b_{j+2}

$$\begin{aligned}
&= \frac{1}{2}\sum_{k=0}^{n} z^{n-k}(-1)^{k+1} \sum_{j=0}^{k-2} \frac{(n-j)!(n-k+j+2)!}{n!(n-k)!} b_j a_{k-j-2} t^{j/2}, \\
&= -\frac{1}{2}\sum_{k=2}^{n} z^{n-k}(-1)^k \sum_{j=0}^{k-2} \frac{(n-j)!(n-k+j+2)!}{n!(n-k)!} b_j a_{k-j-2} t^{j/2} = -\frac{1}{2}\partial_{zz} q_A(z,t).
\end{aligned}$$

□

We conclude with this theorem that, independently of the initial condition of the self-adjoint Brownian motion, the roots of its expected characteristic polynomial will move according to the deterministic Dyson equation. In the next subsection we get similar results for the Wishart process.

4.5.2 Wishart process

Similarly to the Dyson Brownian motion case, we will first construct a tridiagonal model that will help us to get the expected characteristic polynomial.

Lemma 4.5.8. *Let W_β be an $n \times n$ matrix from the β-Wishart ensemble ($\beta \in \{1, 2, 4\}$). The eigenvalues of W_β have the same joint law as those of the tridiagonal matrix $L_\beta = B_\beta^T B_\beta$ with B_β is an $m \times n$ bidiagonal matrix defined as*

$$B_\beta = \begin{bmatrix} \xi_{n\beta} & 0 & 0 & \cdots & 0 \\ \xi_{\beta(m-1)} & \xi_{n\beta-\beta} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \xi_\beta & \xi_{n\beta-\beta(m-1)} \end{bmatrix}. \tag{4.12}$$

The variables ξ_j being independent random variables with distribution $\xi_j \sim \chi_j$. We call L_β the tridiagonal β-Laguerre ensemble.

Proof. Take G to be an $m \times n$ matrix with independent standard β-Gaussian random variables as entries (real if $\beta = 1$, complex if $\beta = 2$ and quaternions if $\beta = 4$). Write G as

$$\begin{bmatrix} \vec{x}^T \\ G_1 \end{bmatrix}.$$

We have that $\vec{x}$ is a vector distributed as a multivariate β-normal variable with mean $\vec{0}$ and covariance matrix $\Sigma = I$, while G_1 is an $(m-1) \times n$ matrix of independent standard β-Gaussian random variables.

Take R to be a "right reflector" corresponding to $\vec{x}$ independent of G_1, which means $\vec{x}^T R = \|\vec{x}\|_2 e_1$. Due to being a reflector, R is orthogonal (unitary, simplectic) and this means that $G_1 R$ is an $(m-1) \times n$ matrix with independent standard β-Gaussian matrices as entries.

Now take $G_1 R = [\vec{y}, G_2]$ with $\vec{y}$ being a β-Gaussian vector with mean $\vec{0}$ and covariance matrix $\Sigma = I$. Then G_2 is an $(m-1) \times (n_1)$ matrix of independent standard β-Gaussian random variables. Let L be a left reflector corresponding to $\vec{y}$ ($Ly = \|\vec{y}\|_2 e_1$) independent of G_2. Again by the orthogonality (unitarity, simplecticity) of L, LG_2 is still an $(m-1) \times (n-1)$ matrix of independent standard β-Gaussian random variables. This means that

$$\begin{bmatrix} 1 & 0 \\ 0 & L \end{bmatrix} GR = \begin{bmatrix} \|\vec{x}\|_2 & 0 \\ \|\vec{y}\|_2 e_1 & LG_2 \end{bmatrix}.$$

We proceed with this procedure now for LG_2. The product by an orthogonal (unitary, simplectic) matrix does not affect the singular values of a matrix, so the singular values of the bidiagonal matrix B_β and the original matrix G are the same. The eigenvalues of $W = GG^T$ are the squares of the singular values of G and the same happens with the eigenvalues of $L_\beta = B_\beta B_\beta^T$ which is a tridiagonal matrix.

The distribution of the entries follows from the definition of the complex and simplectic normal distributions as sums of real normal random variables. □

Another use of Lemma 4.5.1 together with the last Lemma help us to prove that the Laguerre polynomials are the expected characteristic polynomials of the Wishart ensemble.

Theorem 4.5.9. *Let R be an $n\times n$ matrix and R_{ij} be independent Gaussian random variables with mean 0 and variance 1, then*

$$E\left[\chi_x\left(R^TR\right)\right] = L_n(z),$$

where $L_n(x) = (1-\partial_z)^n\left[z^n\right]$ is the nth standard Laguerre polynomial.

Proof. We use the tridiagonal model defined before together with Lemma 4.5.1. Let $P_n(z)$ be the characteristic polynomial of L_β with L_β being a tridiagonal β-Laguerre ensemble, then

$$P_n(z) = (z-(L_\beta)_{11})P_{n-1}(z) - (L_\beta)_{12}^2P_{n-2}(z).$$

The corresponding entries are

$$(L_\beta)_{11} = \xi_{n\beta}^2 + \xi_{\beta(m-1)}^2, \qquad (L_\beta)_{12} = \xi_{\beta(n-1)}\xi_{\beta(m-1)},$$

with this the expected characteristic polynomial satisfies the recursion

$$\begin{aligned}\mathbb{E}\left[P_n(z)\right] &= \mathbb{E}\left[(z-\xi_{n\beta}^2-\xi_{\beta(m-1)}^2)P_{n-1}(z) - \xi_{\beta(n-1)}^2\xi_{\beta(m-1)}^2P_{n-2}(z)\right],\\ &= (z-\beta(n-m+1))\mathbb{E}\left[P_{n-1}(z)\right] - \beta^2(n-1)(m-1)\mathbb{E}\left[P_{n-2}(z)\right], \qquad (4.13)\end{aligned}$$

which is satisfied by the Laguerre polynomials. By finding the cases $n=1$ and $n=2$, then using a recursive argument, we finish the proof.

The case $n=1$ is trivial because in this case G is an univariate Gaussian random variable, so $G^TG = G^2$ is a chi-squared random variable and $P_1(z) = z-1$. For the case $n=2$ we have

$$\begin{aligned}P_2(z) = \det[zI - G^TG] &= \det\begin{bmatrix} z-g_1^2-g_2^2 & -g_1g_3-g_2g_4\\ -g_1g_3-g_2g_4 & z-g_3^2-g_4^2\end{bmatrix},\\ &= z^2 - z(g_1^2+g_2^2+g_3^2+g_4^2) + g_1^2g_3^2 + g_2^2g_4^2 - 2g_1g_2g_3g_4.\end{aligned}$$

Taking expectation leads to $P_z(z) = z^2-4z+2$. The rest of the polynomials are found using recursion (4.13). □

The dynamical case can be found using the static case and the fact that a Wishart matrix has the same law as tG^TG, for G a matrix of independent standard normal variables.

Theorem 4.5.10. *Let B be an $n\times n$ matrix and B_{ij} be independent standard Brownian motions, then*

$$E\left[\chi_z\left(B^TB\right)\right] = L_n(z,t),$$

where $L_n(z,t) = (1-t\partial_z)^n z^n$ is the nth Laguerre polynomial with variance t.

Proof. Let G be an $n\times n$ matrix of independent standard Gaussian random variables, so $B \overset{d}{=} \sqrt{t}G$ and this means $B^TB \overset{d}{=} tG^TG$, it follows that

$$\begin{aligned}\mathbb{E}\left[\det[zI - B^TB]\right] &= \mathbb{E}\left[\det[zI - tG^TG]\right] = t^n\mathbb{E}\left[\frac{z}{t}I - G^TG]\right], \\ &= t^n\left(1 - \frac{\mathrm{d}}{\mathrm{d}(z/t)}\right)^n\left[\frac{z}{t}^n\right] = t^n\sum_{k=0}^{n}(-1)^k\binom{n}{k}\frac{n!}{(n-k)!}\left(\frac{z}{t}\right)^{n-k}, \\ &= \sum_{k=0}^{n}(-t)^k\binom{n}{k}\frac{n!}{(n-k)!}z^{n-k} = \left(1 - t\frac{\mathrm{d}}{\mathrm{d}z}\right)^n[z^n].\end{aligned}$$

□

As defined before, the time dependant Laguerre polynomials are generated by the operator $(1 - t\partial_z)^n[z^n]$. On the other hand, the Finite Free Poisson distribution is represented by the roots of the polynomial $(1 - \frac{1}{n}\partial_z)^n[z^n]$, this is the expected characteristic polynomial of a Wishart process rescales by $1/n$ and evaluated at time $t = 1$. It turns out that we can find a heat kind equation for every polynomial satisfying $p(z,t) = (1 - ct\partial_z)^n[z^n]$ with $c \in \mathbb{R}$. Every such equation would imply a unique dynamics for the polynomial roots which is basically a rescaling of the space-time relationship.

Theorem 4.5.11. *Let $c \in \mathbb{R}$ and $P(t,z) \coloneqq (1 - ct\partial_z)^n[z^n]$, with $t \geq 0, z \in \mathbb{C}$, then $P(t,z)$ satisfies the following differential equation*

$$c\partial_z P(t,z) + cz\partial_{zz}P(t,z) + \partial_t P(t,z) = 0.$$

Proof. By definition of $P(t,z)$,

$$P(t,z) = (1 - ct\partial_z)^n[z^n] = \sum_{k=0}^{n}\binom{n}{k}(-ct)^k\partial_z^k[z^n] = \sum_{k=0}^{n}\binom{n}{k}(-ct)^k\frac{n!}{(n-k)!}z^{n-k}.$$

Now we find the derivatives

$$\begin{aligned}\partial_z[P(t,z)] &= \partial_z\left[\sum_{k=0}^{n}\binom{n}{k}\frac{(-ct)^k\,n!}{(n-k)!}z^{n-k}\right] = \sum_{k=0}^{n-1}\binom{n}{k}\frac{(-ct)^k\,n!}{(n-k-1)!}z^{n-k-1}, \\ \partial_{zz}[P(t,z)] &= \partial_z\left[\sum_{k=0}^{n-1}\binom{n}{k}\frac{(-ct)^k\,n!}{(n-k-1)!}z^{n-k-1}\right] = \sum_{k=0}^{n-2}\binom{n}{k}\frac{(-ct)^k\,n!}{(n-k-2)!}z^{n-k-2}, \\ \partial_t[P(t,z)] &= \partial_t\left[\sum_{k=0}^{n}\binom{n}{k}\frac{(-ct)^k\,n!}{(n-k)!}z^{n-k}\right] = \sum_{k=1}^{n}\binom{n}{k}k\frac{(-c)^kt^{k-1}n!}{(n-k)!}z^{n-k}, \\ &= \sum_{k=0}^{n-1}\binom{n}{k+1}(k+1)(-c)^{k+1}t^k\frac{n!}{(n-k-1)!}z^{n-k-1}.\end{aligned}$$

For the sum we can separate the last terms with index $n-1$. Thus we have

$$
\begin{aligned}
(c\partial_z + cz\partial_{zz} + \partial_t)[P] = \sum_{k=0}^{n-2} \Bigg[& c\binom{n}{k}(-ct)^k \frac{n!}{(n-k-1)!} z^{n-k-1} \\
& + cz\binom{n}{k}(-ct)^k \frac{n!}{(n-k-2)!} z^{n-k-2} \\
& + \binom{n}{k+1}(k+1)(-c)^{k+1}t^k \frac{n!}{(n-k-1)!} z^{n-k-1} \Bigg] \\
& + cn(-ct)^{n-1}n! + n(-c)^n t^{n-1} n!,
\end{aligned}
$$

reagruping the terms we find

$$
\begin{aligned}
&= \sum_{k=0}^{n-2} \binom{n}{k}(-ct)^k n! \left(\frac{(n-k-1)cz^{n-k-1} - c(n-k-1)z^{n-k-1}}{(n-k-1)!} \right), \\
&= \sum_{k=0}^{n-2} \binom{n}{k} \frac{c^{k+1}(-t)^k n! z^{n-k-1}}{(n-k-1)!} (1 + n - k - 1 - n + k) = 0.
\end{aligned}
$$

□

Now, analogously to the Hermite polynomials and the Dyson Brownian motion, we prove the dynamical behavior for the roots of the polynomial defined by $(1 - ct\partial_z)^n[z^n]$. In this case, the dynamics will depend on the constant c. For a negative c, the time will be reversed and in general for a bigger c the movement will be faster.

Theorem 4.5.12. *Let $P(t, z)$ be a monic polynomial with degree n satisfying the equation*

$$c\partial_z[P(t,z)] + cz\partial_{zz}[P(t,z)] + \partial_t[P(t,z)] = 0.$$

Then its roots $(z_i(t))_{i=1}^n$ satisfy the equation of motion

$$\mathrm{d}z_i = c\left(\sum_{k \neq i} \frac{z_i + z_k}{z_i - z_k} + n \right) \mathrm{d}t.$$

Proof. Let $z_i(t)$ be such that $P(t, z_i(t)) = 0$ for every t, this means in particular that $\partial_z P(t, z_i(t)) = 0$, so

$$
\begin{aligned}
0 = \partial_t[P(t, z_i(t))] &= \partial_t[z_i(t)]\partial_z[P(t,z)]|_{z=z_i} + \partial_t[P(t,z)]|_{z=z_i}, \\
&= \partial_t[z_i(t)]\partial_z[P(t,z)]|_{z=z_i} - [c\partial_z[P(t,z)] - cz\partial_{zz}[P(t,z)]]_{z=z_i},
\end{aligned}
$$

now we use that $P(t,z)$ is monic and recall the derivatives of a monic polynomials evaluated at $z = z_i(t)$ from the proof of Corollary 4.5.6

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}z_i(t) &= c\left[\frac{\partial_z[P(t,z)] + z\partial_{zz}[P(t,z)]}{\partial_z[P(t,z)]}\right]_{z=z_i} = c\left[1 + \frac{2z_i\sum_{k\neq i}\prod_{j\neq i,j\neq k}(z_i - z_j)}{\prod_{j\neq i}(z_i - z_j)}\right], \\
&= c\left[1 + \sum_{k\neq i}\frac{2z_i}{z_i - z_k}\right], \\
&= c\left[1 + \sum_{k\neq i}\frac{2z_i}{z_i - z_k} - \sum_{k\neq i}\frac{z_i - z_k}{z_i - z_k} + \sum_{k\neq i}\frac{z_i - z_k}{z_i - z_k}\right], \\
&= c\left[1 + \sum_{k\neq i}\frac{2z_i - z_i + z_k}{z_i - z_k} + n - 1\right] = c\left[\sum_{k\neq i}\frac{z_i + z_k}{z_i - z_k} + n\right].
\end{aligned}$$

□

In total analogy to Proposition 3.1.2, we have a similar result for the Laguerre polynomials using asymmetric additive convolution.

Proposition 4.5.13. *Let $L_n(z,t_1) = (1 - ct_1\partial_z)^n[z^n]$ and $L_n(z,t_2) = (1 - ct_2\partial_z)^n[z^n]$ be two Laguerre polynomials of degree n, then their asymmetric additive convolution $L_n(z,t_1) \boxplus\!\!\boxplus L_n(z,t_2)$ is a Laguerre polynomial of degree n with variance $t_1 + t_2$, $L_n(z, t_1 + t_2)$.*

Proof.

$$\begin{aligned}
&L_n(z,t_1) \boxplus\!\!\boxplus L_n(z,t_2) \\
&\quad= \sum_{k=0}^{n} z^{n-k}(-1)^k \sum_{j=0}^{k}\left(\frac{(n-j)!(n-k+j)!}{n!(n-k)!}\right)^2 \binom{n}{j}\frac{(ct_1)^j n!}{(n-j)!}\binom{n}{k-j}\frac{(ct_1)^{k-j}n!}{(n-k+j)!} \\
&\quad= \sum_{k=0}^{n} z^{n-k}(-1)^k \sum_{j=0}^{k}\frac{((n-j)!)^2((n-k+j)!)^2(n!)^4(ct_1)^j(ct_2)^{k-j}}{(n!)^2((n-k)!)^2 j!(k-j)!((n-j)!)^2((n-k+j)!)^2}, \\
&\quad= \sum_{k=0}^{n}\binom{n}{k}\frac{(-1)^k z^{n-k}n!}{(n-k)!}\sum_{j=0}^{k}\binom{k}{j}(ct_1)^j(ct_2)^{k-j}, \\
&\quad= \sum_{k=0}^{n}\binom{n}{k}\frac{(-1)^k z^{n-k}n!}{(n-k)!}[c(t_1+t_2)]^k = L_n(z, t_1+t_2).
\end{aligned}$$

□

Finally, we state a theorem that allows us to conclude that, for every initial conditions, if convoluted with a Laguerre polynomial, they will always move according to the deterministic Wishart process.

Theorem 4.5.14. *Let $L_n(z,t)$ be a Laguerre polynomial of order n with variance t and let $p(z)$ be any monic polynomial. Then the asymmetric additive convolution of $L_n(z,t)$ and $p(z)$, $L_n(z,t) \boxplus\!\!\boxplus p(z)$ satisfies the following differential equation*

$$c\partial_z[L_n(z,t) \boxplus\!\!\boxplus p(z)] + cz\partial_{zz}[L_n(z,t) \boxplus\!\!\boxplus p(z)] + \partial_t[L_n(z,t) \boxplus\!\!\boxplus p(z)] = 0.$$

Proof. We find the derivatives first

$$
\begin{aligned}
&\partial_z[L_n(z,t) \boxplus\!\!\boxplus p(z)] \\
&\quad = \partial_z \left[\sum_{k=0}^{n} z^{n-k}(-1)^k \sum_{j=0}^{k} \left(\frac{(n-j)!(n-k+j)!}{n!(n-k)!} \right)^2 \binom{n}{j} \frac{(ct_1)^j n!}{(n-j)!} b_{k-j} \right], \\
&\quad = \sum_{k=0}^{n-1} (n-k) z^{n-k-1}(-1)^k \sum_{j=0}^{k} \left(\frac{(n-j)!(n-k+j)!}{n!(n-k)!} \right)^2 \binom{n}{j} \frac{(ct_1)^j n!}{(n-j)!} b_{k-j},
\end{aligned}
$$

we use the first derivative to find the second one

$$
\begin{aligned}
&\partial_{zz}[L_n(z,t) \boxplus\!\!\boxplus p(z)] \\
&\quad = \partial_{zz} \left[\sum_{k=0}^{n} z^{n-k}(-1)^k \sum_{j=0}^{k} \left(\frac{(n-j)!(n-k+j)!}{n!(n-k)!} \right)^2 \binom{n}{j} \frac{(ct_1)^j n!}{(n-j)!} b_{k-j} \right], \\
&\quad = \sum_{k=0}^{n-2} (n-k)(n-k-1) z^{n-k-2}(-1)^k \sum_{j=0}^{k} \left(\frac{(n-j)!(n-k+j)!}{n!(n-k)!} \right)^2 \binom{n}{j} \frac{(ct_1)^j n!}{(n-j)!} b_{k-j},
\end{aligned}
$$

finally, for the derivative in time

$$
\begin{aligned}
&\partial_t[L_n(z,t) \boxplus\!\!\boxplus p(z)] \\
&\quad = \partial_t \left[\sum_{k=0}^{n} z^{n-k}(-1)^k \sum_{j=0}^{k} \left(\frac{(n-j)!(n-k+j)!}{n!(n-k)!} \right)^2 \binom{n}{j} \frac{(ct_1)^j n!}{(n-j)!} b_{k-j} \right], \\
&\quad = \sum_{k=0}^{n} z^{n-k}(-1)^k \sum_{j=1}^{k} \left(\frac{(n-j)!(n-k+j)!}{n!(n-k)!} \right)^2 \binom{n}{j} \frac{c^j j t_1^{j-1} n!}{(n-j)!} b_{k-j}, \\
&\quad = \sum_{k=1}^{n} z^{n-k}(-1)^k \sum_{j=0}^{k-1} \left(\frac{(n-j-1)!(n-k+j+1)!}{n!(n-k)!} \right)^2 \binom{n}{j+1} \frac{c^{j+1}(j+1) t_1^j n!}{(n-j-1)!} b_{k-j-1},
\end{aligned}
$$

with a shift in the coefficients we can have the same index as in the two other derivatives

$$
\begin{aligned}
&\quad = \sum_{k=0}^{n-1} z^{n-k-1}(-1)^{k+1} \sum_{j=0}^{k} \left(\frac{(n-j-1)!(n-k+j)!}{n!(n-k-1)!} \right)^2 \binom{n}{j+1} \frac{c^{j+1}(j+1) t_1^j n!}{(n-j-1)!} b_{k-j}, \\
&\quad = \sum_{k=0}^{n-1} (n-k)^2 z^{n-k-1}(-1)^{k+1} \sum_{j=0}^{k} \left(\frac{(n-j)!(n-k+j)!}{n!(n-k)!} \right)^2 \binom{n}{j} \frac{c^{j+1} t_1^j n!}{(n-j)!} b_{k-j},
\end{aligned}
$$

Now we can sum all over the same set onf indexes up to $n-2$ and sum the two terms wth index $n-1$ outside

$$
\begin{aligned}
&(c\partial_z + cz\partial_{zz} + \partial_t)[L_n(z,t) \boxplus\boxplus p(z)] \\
&= \sum_{k=0}^{n-2} z^{n-k-1}(-1)^k \sum_{j=0}^{k} \left(\frac{(n-j)!(n-k+j)!}{n!(n-k-1)!}\right)^2 \binom{n}{j} \frac{c^{j+1}t_1^j n!}{(n-j)!} b_{k-j}\Big((n-k)^2 - (n-k)^2\Big) \\
&\qquad + (-1)^{n-1} \sum_{j=0}^{n-1} \left(\frac{(j+1)!}{n!}\right)^2 \binom{n}{j} \frac{c^{j+1}t_1^j n!}{(n-j)!} b_{n-j+1} \\
&\qquad\qquad + (-1)^{n} \sum_{j=0}^{n-1} \left(\frac{(j+1)!}{n!}\right)^2 \binom{n}{j} \frac{c^{j+1}t_1^j n!}{(n-j)!} b_{n-j+1} \\
&= 0.
\end{aligned}
$$

□

All the results in this section were analogous to the Hermite polynomials relationship to the Dyson Brownian motion. As we will see in the next section, the results are not as easy to generalize for the matrix Jacobi process. Thus we will need to make a more extensive use of Finite Free Probability Theory.

4.5.3 Jacobi process

For the last process we study in this thesis, we will use a slightly different approach. Although some useful matrix models have been proposed for the Jacobi ensamble [37], [38], we will make use of the definition of the Jacobi matrix in terms of Wishart matrices and what we already know about the Laguerre relationship to Wishart. As we will see, the Jacobi process is stationary, which makes it impossible to define something as the "dynamical Jacobi polynomials" (at least as we did with Hermite and Laguerre), however, it is possible to recover the deterministic dynamics when we have an initial law that does not make the process stationary.

Before proceeding with the proper Jacobi process, let us find the expected characteristic polynomial of a static Jacobi matrix. This process is more challenging than the two previous processes, but the static case is greatly simplified by using the tools of Finite Free Probability theory.

Recall from Chapter 2 that there are essentially two definitions for the Jacobi matrix, one coming from the Multivariate Analysis of Variance and the other coming from the upper left corner of a Haar unitary Brownian matrix. For the static case, we work with the MANOVA case.

Theorem 4.5.15. *Let A, B be two $n \times n$ random matrices with i.i.d. standard Gaussian entries. Define the matrix M as*

$$M \coloneqq (A^T A + B^T B)^{-1} A^T A,$$

then the expected characteristic polynomial of M, $\chi_z(M)$ is the nth Legendre polynomial,

$$\chi_z(M) = \frac{n!n!}{(2n)!} \sum_{k=0}^{n} z^k (-1)^{n-k} \binom{n}{k} \binom{n+k}{k}.$$

Proof. Denote $p_M(z) \coloneqq \chi_z(M)$. To obtain $p_M(z)$, we can first find that of its inverse, $p_{M^{-1}}(z)$. We have

$$M^{-1} = (A^TA)^{-1}(A^TA + B^TB) = I + (A^TA)^{-1}B^TB.$$

Under the assumptions that I and $(A^TA)^{-1}B^TB$ are normal and that $(A^TA)^{-1}B^TB$ is invariant under conjugation by an orthogonal matrix, it holds that

$$p_{M^{-1}}(z) = p_I(z) \boxplus_n p_{(A^TA)^{-1}B^TB}(z).$$

Furthermore, under the assumption that $(A^TA)^{-1}$ and B^TB are normal and invariant under conjugation by orthogonal matrices, it holds that

$$p_{(A^TA)^{-1}B^TB}(z) = p_{(A^TA)^{-1}} \boxtimes p_{B^TB}(z).$$

We also know that $p_{B^TB}(z) = L_n(z)$ and $p_{(A^TA)^{-1}}(z) = \frac{z^n}{L_n(0)}L_n(1/z)$. With this, we can find $p_{M^{-1}}(z)$, by noticing that $L_n(0) = (-1)^n n!$,

$$\begin{aligned}
p_{(A^TA)^{-1}B^TB}(z) = p_{(A^TA)^{-1}} \boxtimes_n p_{B^TB}(z) &= \left(\frac{(-1)^n}{n!}z^n p_{A^TA}(1/z) \boxtimes p_{B^TB}(z)\right), \\
&= \frac{(-1)^n}{n!}\sum_{k=0}^{n} z^{n-k}(-1)^k \frac{a_k b_k}{\binom{n}{k}}, \\
&= \frac{(-1)^n}{n!}\sum_{k=0}^{n} z^{n-k}(-1)^k \frac{\binom{n}{k}\frac{n!}{(n-k)!}\binom{n}{k}(-1)^{n-2k}\frac{n!}{k!}}{\binom{n}{k}}, \\
&= \frac{(-1)^n}{n!}\sum_{k=0}^{n} z^{n-k}(-1)^k \binom{n}{k}\frac{n!}{(n-k)!}(-1)^{n-2k}\frac{n!}{k!}, \\
&= \sum_{k=0}^{n} z^{n-k}(-1)^k \binom{n}{k}^2 (-1)^{2n-2k}.
\end{aligned}$$

Now we calculate $p_{M^{-1}}(z)$, remembering that $p_I(z) = (z-1)^n$

$$\begin{aligned}
p_I \boxplus_n p_{(A^TA)^{-1}B^TB} &= \sum_{k=0}^{n} z^{n-k}(-1)^k \sum_{j=0}^{k} \frac{(n-j)!(n-k+j)!}{n!(n-k)!}\binom{n}{j}\binom{n}{k-j}^2 (-1)^{2n-2(k-j)}, \\
&= \sum_{k=0}^{n} z^{n-k}(-1)^{2n-k}\binom{n}{k}\sum_{j=0}^{k}\frac{n!n!k!}{j!(k-j)!(k-j)!(n-k+j)!}, \\
&= \sum_{k=0}^{n} z^{n-k}(-1)^{2n-k}\binom{n}{k}\sum_{j=0}^{k}\binom{n}{k-j}\binom{k}{j}, \\
&= \sum_{k=0}^{n} z^{n-k}(-1)^{2n-k}\binom{n}{k}\binom{n+k}{k}.
\end{aligned}$$

Finally, we use $p_M(z) = \frac{z^n}{p_{M^{-1}}(0)}p_{M^{-1}}(1/z)$ to obtain

$$p_M(z) = \frac{n!n!z^n}{(2n)!}\sum_{k=0}^{n} z^{k-n}(-1)^{n-k}\binom{n}{k}\binom{n+k}{k} = \frac{n!n!}{(2n)!}\sum_{k=0}^{n} z^{k}(-1)^{n-k}\binom{n}{k}\binom{n+k}{k}.$$

This corresponds to the n-th Legendre polynomial evaluated at $1-2z$, which is precisely the desired expected characteristic polynomial. □

Now we will approach the dynamical version. Let N_1, N_2 be two $n \times n$ random matrices with i.i.d standard Gaussian entries and B_1, B_2 be two independent Brownian motions in $\mathcal{M}_{n,n}(\mathbb{R})$.

Notice that if we substitute these matrices in the first definition of the Jacobi matrix, we get

$$(B_1^T B_1 + B_2^T B_2)^{-1} B_1^T B_1 \overset{d}{=} (tN_1^T N_1 + tN_2^T N_2)^{-1} tN_1^T N_1 = (N_1^T N_1 + N_2^T N_2)^{-1} N_1^T N_1,$$

while for the second definition, we get

$$\begin{aligned}(B_1^T B_1 + B_2^T B_2)^{-1/2} B_1^T B_1 (B_1^T B_1 + B_2^T B_2)^{-1/2} \\ &\overset{d}{=} (tN_1^T N_1 + tN_2^T N_2)^{-1/2} tN_1^T N_1 (tN_1^T N_1 + tN_2^T N_2)^{-1/2}, \\ &= (N_1^T N_1 + N_2^T N_2)^{-1/2} N_1^T N_1 (N_1^T N_1 + N_2^T N_2)^{-1/2}.\end{aligned}$$

Thus, in both definitions of the Jacobi matrix, if we substitute the Gaussian random variables with Brownian motions, we get the same distribution. This means that the matrix Jacobi process is stationary with the law of the zeroes of the associated Jacobi polynomial. Then the expected characteristic polynomial doest not depend on t and we can not formulate something as the time dependent Jacobi polynomials or Jacobi polynomials with variance. We will see, however, that if we start the process with another distribution, we can recover the deterministic Jacobi dynamics. All of the following content is taken from [21], where the author manages to generalize the results we got in the previous subsections together with the Jacobi case. We will work with the Jacobi matrix definition in [22] where the Jacobi process can be expressed by the following matrix

$$W = (M_1^T M_1 + M_2^T M_2)^{-1} M_1^T M_1 (M_1^T M_1 + M_2^T M_2)^{-1},$$

where M_1, M_2 are independent Gaussian matrices with dimensions $n_1 \times k$ and $n_2 \times k$, respectively. We can write $W_i = M_i^T M_i$ and $\Delta = \det\left[(W_1+W_2)^{-1}\right]$ for a shorter notation. Now to compute the characteristic polynomial we have,

$$\begin{aligned}p_W(z) &= \det\left[zI - (W_1+W_2)^{-\frac{1}{2}} W_1 (W_1+W_2)^{-\frac{1}{2}}\right], \\ &= \Delta \det\left[z(W_1+W_2) - (W_1+W_2)^{\frac{1}{2}} W_1 (W_1+W_2)^{-\frac{1}{2}}\right], \\ &= \Delta \det\left[z(W_1+W_2)(W_1+W_2)^{\frac{1}{2}}(W_1+W_2)^{-\frac{1}{2}} - (W_1+W_2)^{\frac{1}{2}} W_1 (W_1+W_2)^{-\frac{1}{2}}\right], \\ &= \Delta \det\left[(W_1+W_2)^{\frac{1}{2}}\right] \det\left[z(W_1+W_2) - W_1\right] \det\left[(W_1+W_2)^{-\frac{1}{2}}\right], \\ &= \Delta \det\left[(z-1)W_1 + zW_2\right].\end{aligned}$$

Thus finding the expected characteristic polynomial (up to a normalization constant) can be done without the need to use the inverse of $W_1 + W_2$.

Replicating the procedure in [21], we will first replicate the former results of the static Wishart and Jacobi matrices with an operator that generalized characteristic polynomials, then we will show that the Wishart process can be built in the base of this operator by adding a time parameter. Finally, we will extend this same procedure to the Jacobi process. In the process of building these generalizations we will make extensive use of Finite Free Probability Theory.

Definition 4.5.1 (Generalized characteristic polynomial)**.** Let A, B be two random independent Gaussian matrices in $\mathcal{M}_{n_1,k}(\mathbb{R})$ and $\mathcal{M}_{n_2,k}(\mathbb{R})$ respectively. We define the generalized characteristic polynomial $p_{A,B}(x, y, z)$ as

$$p_{A,B}(x, y, z) \coloneqq \det\left[xI + yA^TA + zB^TB\right].$$

Similarly, we define the reciprocal generalized reciprocal polynomial as $q_{A,B}(x, y, z) \coloneqq y^{n_1}z^{n_2}P_{A,B}(x, y, z)$.

Notice that this generalizes the characteristic polynomial of the (static and dynamical) Wishart and the static Jacobi cases.

Similarly to the univariate cases (Theorem 3.1.1), we have a multiplicative theorem in terms of the operators that generate the new multivariate polynomials.

Theorem 4.5.16. *Let F and G be two-variable polynomials and $A, B \in \mathcal{M}_{k,n_1}, C, D \in \mathcal{M}_{k,n_2}$ such that C and D are invariant under product by signed permutation matrices. Suppose that*

$$\mathbb{E}\left[q_{A,B}(x, y, z)\right] = F(\partial_x\partial_y, \partial_x\partial_z)[x^k y^{n_1} z^{n_2}],$$
$$\mathbb{E}\left[q_{C,D}(x, y, z)\right] = G(\partial_x\partial_y, \partial_x\partial_z)[x^k y^{n_1} z^{n_2}].$$

Then $\mathbb{E}\left[q_{A+C,B+D}(x, y, z)\right] = F(\partial_x\partial_y, \partial_x\partial_z)G(\partial_x\partial_y, \partial_x\partial_z)[x^k y^{n_1} z^{n_2}]$

Proof. If F', G' are the linear differential operators (multivariate polynomials on $\partial_x\partial_y, \partial_x\partial_z$) generating $p_{A,B}$ and $p_{C,D}$ and they satisfy the former multiplicative property, then it can be extended to F, G by linearity. Thus we prove for F', G'.

By definition, we have that

$$\mathbb{E}\left[p_{A,B}(x, y, z)\right] = \mathbb{E}\left[\det[xI + yA^TA + zB^TB]\right] = \mathbb{E}\left[\chi_x(-yA^TA - zB^TB)\right],$$
$$\mathbb{E}\left[p_{C,D}(x, y, z)\right] = \mathbb{E}\left[\det[xI + yC^TC + zD^TD]\right] = \mathbb{E}\left[\chi_x(-yC^TC - zD^TD)\right].$$

Notice that by Theorem 3.2.8 we have that

$$\mathbb{E}\left[p_{A+C,B+D}(x, y, z)\right] = \mathbb{E}\left[\chi_x(-yA^TA - zB^TB)\right] \boxplus_n \mathbb{E}\left[\chi_x(-yC^TC - zD^TD)\right].$$

Applying Theorem 3.1.1 we get to the desired result by considering the derivatives in y, z as constant coefficients when evaluating in x. □

With the next theorem we get a limit for the products of differential operators, similarly to what we did in Chapter 3 when proving the finite free limit theorems. We will use this

theorem to prove properties of the generalized characteristic polynomial of Gaussian matrices, by see these matrices as limits.

Theorem 4.5.17. *Let $\{A_i\}_{i=0}^{\infty}$ and $\{B_i\}_{i=0}^{\infty}$ be sequences of matrices invariant under transformation by signed permutation matrices such that*

$$\mathbb{E}\left[Tr(A_i^T A_i)\right] = \sigma_1 n_1 k,$$
$$\mathbb{E}\left[Tr(B_i^T B_i)\right] = \sigma_2 n_2 k.$$

If we define C_m, D_m as

$$C_m \coloneqq \sum_{i=0}^{m} \frac{A_i}{\sqrt{m}},$$
$$D_m \coloneqq \sum_{i=0}^{m} \frac{D_i}{\sqrt{m}}.$$

Then the matrices C_m, D_m satisfy

$$\lim_{m\to\infty} \mathbb{E}\left[q_{C,D}(x,y,z)\right] = e^{\sigma_1 \partial_x\partial_y + \sigma_2 \partial_x\partial_z}[x^k y^{n_1} z^{n_2}].$$

Proof. First we use the following expansion for the determinant,

$$\det(A + hB) = \det(A) + h\, Tr\left(\mathrm{adj}(A)\, B\right) + O\left(h^2\right).$$

This means for $q_{A_i/\sqrt{m},B_i/\sqrt{m}}(x,y,z)$ that

$$\begin{aligned}
&\mathbb{E}\left[q_{A_i/\sqrt{m},B_i/\sqrt{m}}(x,y,z)\right] \\
&= y^{n_1} z^{n_2} \mathbb{E}\left[\det[xI + \frac{1}{m}\left(y^{-1}A_i^T A_i + z^{-1} B_i^T B I\right)]\right], \\
&= y^{n_1} z^{n_2}\left[x^k + \frac{x^{k-1}y^{-1}}{m}\mathbb{E}\left[Tr(A_i^T A_i)\right] + \frac{x^{k-1}z^{-1}}{m}\mathbb{E}\left[Tr(B_i^T B_i)\right] + O\left(\frac{1}{m^2}\right))\right], \\
&= x^k y^{n_1} z^{n-2} + \frac{\sigma_1 n_1 k}{m} x^{k-1} y^{n_1 - 1} z^{n-2} + \frac{\sigma_2 n_2 k}{m} x^{k-1} y^{n_1} z^{n_2 - 1} + O\left(\frac{1}{m^2}\right), \\
&= \left(1 + \frac{\sigma_1}{m}\partial_y\partial_x + \frac{\sigma_2}{m}\partial_z\partial_x + O\left(\frac{1}{m^2}\right)\right)[x^k y^{n_1} z^{n_2}]
\end{aligned}$$

Once we have found $q_{A_i/\sqrt{m},B_i/\sqrt{m}}(x,y,z)$ as a polynomial on the operators $\partial_x\partial_y$ and $\partial_x\partial_z$, we apply Theorem 4.5.16 to get, for $q_{C_m,D_m}(x,y,z)$,

$$\mathbb{E}\left[q_{C_m,D_m}(x,y,z)\right] = \left(1 + \frac{\sigma_1}{m}\partial_y\partial_x + \frac{\sigma_2}{m}\partial_z\partial_x + O\left(\frac{1}{m^2}\right)\right)^m [x^k y^{n_1} z^{n_2}]$$

Tanking $m \to \infty$ in the last expression we get

$$\lim_{m\to\infty}\mathbb{E}\left[q_{C_m,D_m}(x,y,z)\right]=\lim_{m\to\infty}\left(1+\frac{\sigma_1}{m}\partial_y\partial_x+\frac{\sigma_2}{m}\partial_z\partial_x+O\left(\frac{1}{m^2}\right)\right)^m[x^k y^{n_1}z^{n_2}],$$
$$=e^{\sigma_1\partial_x\partial_y+\sigma_2\partial_x\partial_z}[x^k y^{n_1}z^{n_2}],$$

which is the desired result. □

Applying the previous theorem, we find that the differential operator that generates the generalized characteristic polynomial of Gaussian matrices, has an exponential form.

Corollary 4.5.18. *Let A and B be equal in law to $\sigma_1 N_1$ and $\sigma_2 N_2$ where N_1 and N_2 are matrices with all of the entries being independent standard normal random variables, then*

$$\mathbb{E}\left[q_{A,B}(x,y,z)\right]=e^{\sigma_1\partial_x\partial_y+\sigma_2\partial_x\partial_z}[x^k y^{n_1}z^{n_2}].$$

Proof. Let $\{A_i\}_{i\in\mathbb{N}}$ and $\{B_i\}_{i\in\mathbb{N}}$ be as in Theorem 4.5.17 and define C_m, D_m in the same way. On the one hand we have, as shown previously

$$\lim_{m\to\infty}\mathbb{E}\left[q_{C_m,D_m}(x,y,z)\right]=e^{\sigma_1\partial_x\partial_y+\sigma_2\partial_x\partial_z}[x^k y^{n_1}z^{n_2}].$$

Due to the existence of the second moment of the entries, on the other hand, we have that C_m, D_m converge in law to a Gaussian independent matrix whose entries have variance σ_1^2 and σ_2^2, respectively, i.e.

$$\lim_{m\to\infty}\mathbb{E}\left[q_{C_m,D_m}(x,y,z)\right]=\mathbb{E}\left[q_{\sigma_1 N_1,\sigma_2 N_2}(x,y,z)\right].$$

□

With this we have an operator that allows us to compute $q_{A,B}(x,y,z)$ for A, B independent Gaussian matrices with zero mean and variances σ_1, σ_2, respectively. Our goal is to be able to use this to study matrix-valued stochastic processes by letting the variances vary linearly in time. Hence, the entries of the matrices have the same law as a standard Brownian motion starting at zero. An interesting question is how this behavior would be affected if we start our processes at points different than zero. The answer is given by Theorem 4.5.16 and noticing that any polynomial $r(x,y,z)$ of orders k, n_1, n_2 can be seen as a polynomial differential operator acting on $x^k y^{n_1} y^{n_2}$, so if $S\in\mathcal{M}_{k,n_1}(\mathbb{R}), T\in\mathcal{M}_{k,n_2}(\mathbb{R})$ are fixed matrices and A, B are Gaussian independent matrices with variances σ_1 and σ_2, respectively, then

$$\mathbb{E}\left[q_{S+A,T+B}(x,y,z)\right]=e^{\sigma_1\partial_x\partial_y+\sigma_2\partial_x\partial_z}[q_{S,T}(x,y,z)].$$

We are interested in $p_{A,B}(x,y,)$, the reciprocal polynomial in y and z. Let us denote by $R_k^z(\cdot)$ the reciprocal polynomial operator of order k in the variable z, which means, for $r(x,y,z)$ a polynomial,

$$R_k^z\left(r(x,y,z)\right)=z^k r\left(x,y,\frac{1}{z}\right).$$

Now notice that for A, B Gaussian independent matrices with variances σ_1, σ_2 and given matrices S, T we have $q_{S+A,T+B}(x,y,z)=R_{n_2}^z(R_{n_1}^y(p_{S+A,T+B}(x,y,z))$, so

$$\begin{aligned}
\mathbb{E}\left[p_{A,B}(x,y,z)\right] &= R^z_{n_2}\left(R^y_{n_1}\left(\mathbb{E}\left[q_{S+A,T+B}(x,y,z)\right]\right)\right), \\
&= R^z_{n_2}\left(R^y_{n_1}\left(e^{\sigma_1\partial_x\partial_y+\sigma_2\partial_x\partial_z}q_{S,T}(x,y,z)\right)\right), \\
&= R^z_{n_2}\left(R^y_{n_1}\left(e^{\sigma_1\partial_x\partial_y+\sigma_2\partial_x\partial_z}R^z_{n_2}(R^y_{n_1}(p_{S,T}(x,y,z)\right)\right), \\
&= \left(R^z_{n_2}\circ R^y_{n_1}\circ e^{\sigma_1\partial_x\partial_y+\sigma_2\partial_x\partial_y}\circ R^z_{n_2}\circ R^y_{n_1}\right)[p_{S,T}(x,y,z)].
\end{aligned}$$

We will find an expression for this operator applied to $x^iy^jz^l$ with $j \le n_1, l \le n_2$. Then it is extended linearly to a general polynomial.

$$\begin{aligned}
&(R^z_{n_2}\circ R^y_{n_1}\circ e^{\sigma_1\partial_x\partial_y+\sigma_2\partial_x\partial_y}\circ R^z_{n_2}\circ R^y_{n_1})[x^iy^jz^l] \\
&\quad= (R^z_{n_2}\circ R^y_{n_1}\circ e^{\sigma_1\partial_x\partial_y+\sigma_2\partial_x\partial_y})[x^iy^{n_1-j}z^{n_2-l}], \\
&\quad= (R^z_{n_2}\circ R^y_{n_1}\circ e^{\sigma_2\partial_x\partial_z})\sum_{r=0}^{\infty}\frac{\sigma_1^r}{r!}\partial_x^r\partial_y^r[x^iy^{n_1-j}z^{n_2-l}], \\
&\quad= (R^z_{n_2}\circ R^y_{n_1}\circ e^{\sigma_2\partial_x\partial_z})\sum_{r=0}^{\infty}\frac{\sigma_1^r}{r!}\partial_x^r[x^i]\frac{(n_1-j)!}{(n_1-j-r)!}y^{n_1-j-r}z^{n_2-l}, \\
&\quad= (R^z_{n_2}\circ R^y_{n_1}\circ e^{\sigma_2\partial_x\partial_z})z^{-l}\sum_{r=0}^{\infty}\binom{n_1-j}{r}y^{n_1-j-r}\sigma_1^r\partial_x^r[x^i],
\end{aligned}$$

doing the same procedure for $e^{\sigma_2\partial_x\partial_z}$ we get

$$\begin{aligned}
&= (R^z_{n_2}\circ R^y_{n_1}\circ e^{\sigma_2\partial_x\partial_z})z^{n_2-l}\left(y+\sigma_1\partial_x\right)^{n_1-j}[x^i], \\
&= R^z_{n_2}\circ R^y_{n_1}\left\{(z+\sigma_2\partial_x)^{n_2-l}\left(y+\sigma_1\partial_x\right)^{n_1-j}[x^i]\right\}, \\
&= (1+z\sigma_2\partial_x)^{n_2-l}\left(1+y\sigma_1\partial_x\right)^{n_1-j}[x^iy^jz^l].
\end{aligned}$$

In the case when we start the processes at the origin we have that

$$p_{S,T}(x,y,z) = \det[xI + y0_{n_1,k} + z0_{n_2,k}] = x^k\det[I] = x^k,$$

which leads to

$$\mathbb{E}\left[p_{A,B}(x,y,z)\right] = (1+z\sigma_2\partial_x)^{n_2-l}\left(1+y\sigma_1\partial_x\right)^{n_1-j}[x^k]. \tag{4.14}$$

If we let $\sigma_1 = \sigma_2 = 1$ and replace $y = -1, z = 0$ in (4.14), for $n_1 = k$, we recover the static Wishart matrix and we can conclude that

$$\mathbb{E}\left[\det[xI - A^TA]\right] = (1-\partial_x)^{n_1}[x^k], \tag{4.15}$$

which is exactly the associated Laguerre polynomial previously defined in Section 3.1. Thus, we have successfully replicated the result in Subsection 4.5.2.

Keeping the unitary variances and replacing $x = 0, y = (x-1), z = x$, we get the static Jacobi matrix, as shown previously, thus

$$\begin{aligned}\mathbb{E}\left[\det[(x-1)W_1+xW_2]\right] &= (1+x\partial_x)^{n_2}\,(1+(x-1)\partial_x)^{n_1}\,[x^k]\\ &= \left\{\left(\sum_{j=0}^{n_2}\binom{n_2}{j}x^j\partial_x^j\right)\left(\sum_{l=0}^{n_1}\binom{n_1}{j}x^l\partial_x^l\right)\right\}[x^k],\\ &= \sum_{j=0}^{k}\sum_{l\leq n-j}\binom{n_2}{j}\binom{n_1}{l}x^j(x-1)^{n-l}\partial_x^{j+l}[x^k]\\ &= \sum_{l=0}^{k}\binom{n_2}{k-l}\binom{n_1}{l}x^{n-l}(x-1)^l =: P_k^{n_2-k,n_1-k}(2x-1).\end{aligned}$$

This is the standard Jacobi polynomial in (3.7) with order k and parameters n_2-k, n_1-k evaluated in $2x-1$. When we normalize to make it monic, we get the result appearing in [34].

Now let $\sigma_1=\sigma_2=t$, then A,B are equal in law to standard Brownian motions in $\mathcal{M}_{n_1,k}(\mathbb{R})$ and $\mathcal{M}_{n_2,k}(\mathbb{R})$, respectively. We should recover the expected characteristic polynomial of the Wishart and Jacobi processes. The Wishart case is evident by replacing -1 with $-t$ in (4.15) and comparing it to equation (3.6). For the Jacobi process with initial condition $p(z)$, we have

$$\mathbb{E}\left[\det[(z-1)W_1(t)+zW_2(t)]\right] = (1+xt\partial_x)^{n_2}\,(1+(x-1)t\partial_x)^{n_1}\,[p(z)]. \tag{4.16}$$

In analogy to what we found for the Dyson Brownian motion and the Wishart process, we need to verify if the roots of the polynomial defined in (4.16) satisfy a differential equation given by the finite variation part of (2.22). Let us work in full generality and define for polynomials of the form $x^iy^jz^l$, the operator

$$Q^t_{n_1,n_2}[x^iy^jz^l] = (1+zt\partial_x)^{n_2-l}\,(1+yt\partial_x)^{n_1-j}\,[x^iy^jz^l],$$

that is later extended to general polynomials linearly.

Denote the time-dependant polynomial $Q^t_{n_1,n_2}[p(x,y,z)]$ by $\hat{p}(x,y,z,t)$. In analogy to what we have done with the Hernmite and Laguerre cases, we want to find for every t the values $x(t),y(t),z(t)$ such that $\hat{p}(x(t),y(t),z(t),t)=0$. By derivating with respect to t, we get

$$\begin{aligned}0 &= \partial_t[\hat{p}(x(t),y(t),z(t),t)]\\ &= [(\partial_t x(t))\partial_x\hat{p}(x,y,z,t)+(\partial_t y(t))\partial_y\hat{p}(x,y,z,t)\\ &\qquad +(\partial_t z(t))\partial_z\hat{p}(x,y,z,t)+\partial_t\hat{p}(x,y,z,t)]_{(x,y,z)=(x(t),y(t),z(t))}.\end{aligned}$$

Now, again in full generality, let us find some properties satisfied by the time-dependant polynomial with arbitrary parameters.

Lemma 4.5.19. *The time-dependant polynomial $\hat{p}(x,y,z,t)$ satisfies*

$$\partial_t[\hat{p}(x,y,z,t)]_{t=0} = (n_1 y + n_2 z)\partial_x[\hat{p}(x,y,z,t)]_{t=0} - y^2\partial_y\partial_x[\hat{p}(x,y,z,t)]_{t=0} - z^2\partial_z\partial_x[\hat{p}(x,y,z,t)]_{t=0}. \tag{4.17}$$

Proof. A general polynomial p on x, y, z, can be written as

$$p(x,y,z) = \sum_{i,j,l} c_{ijl} x^i y^j z^l,$$

where c_{ijl} is the coefficiente associated to the term $x^i y^j z^l$. When we apply $Q^t_{n_1,n_2}$ to $p(x,y,z)$ we have

$$\begin{aligned}
Q^t_{n_1,n_2}[p(x,y,z)] &= (1+zt\partial_x)^{n_2-l}(1+yt\partial_x)^{n_1-j}\left[\sum_{i,j,l} c_{ijl}x^i y^j z^l\right], \\
&= \sum_{i,j,l} c_{ijl}(1+zt\partial_x)^{n_2-l}(1+yt\partial_x)^{n_1-j}\left[x^i y^j z^l\right] \\
&= \sum_{i,j,l} c_{ijl}\sum_{r=0}^{n_2-l}\binom{n_2-l}{r}(zt\partial_x)^r \sum_{s=0}^{n_1-j}\binom{n_1-j}{s}(yt\partial_x)^{n_1-j}[x^i y^j k^l], \\
&= \sum_{i,j,l} c_{ijl}\left\{(n_2-l)zt\partial_x + (n_1-j)yt\partial_x + O(t^2)\right\}[x^i y^j k^l], \\
&= \sum_{i,j,l} c_{ijl}\left\{(n_2-l)itx^{i-1}y^j z^{l+1} + (n_1-j)itx^{i-1}y^{j+1}z^l\right\} + O(t^2).
\end{aligned}$$

Then, differentiating in t and evaluating at $t=0$ gives us

$$\begin{aligned}
\partial_t\left\{Q^t_{n_1,n_2}[p(x,y,z)]\right\}\big|_{t=0} &= \partial_t\left\{\sum_{i,j,l} c_{ijl}\left\{(n_2-l)itx^{i-1}y^j z^{l+1}\right.\right. \\
&\qquad \left.\left. +(n_1-j)itx^{i-1}y^{j+1}z^l\right\} + O(t^2)\right\}\Bigg|_{t=0}, \\
&= \sum_{ijl} c_{ijl}\{(n_2-l)ix^{i-1}y^j z^{l+1} + (n_1-j)ix^{i-1}y^{j+1}z^l\}.
\end{aligned}$$

We can re-write the last two terms in every summand as

$$\begin{aligned}
(n_2-l)ix^{i-1}y^j z^{l+1} &= n_2 z\partial_x[x^i y^j z^l] - z^2\partial_x\partial_z[x^i y^j z^l], \\
(n_1-j)ix^{i-1}y^{j+1}z^l &= n_1 y\partial_x[x^i y^j z^l] - y^2\partial_x\partial_y[x^i y^j z^l].
\end{aligned}$$

Extending by linearity to p we have that

$$\partial_t[\hat{p}]_{t=0} = (n_1 y + n_2 z)\partial_x[p] - y^2\partial_y\partial_x[p] - z^2\partial_z\partial_x[p].$$

This finishes the proof since $p(x,y,z) = \hat{p}(x,y,z,t)|_{t=0}$. □

This last result together with (4.17) give us that

$$\begin{aligned}\partial_t[x(t)]\partial_x[p(x(t),y(t),z(t))] + \partial_t[y(t)]\partial_y[p(x(t),y(t),z(t))] + \partial_t[z(t)]\partial_z[p(x(t),y(t),z(t))\\ = -(n_1 y + n_2 z)\partial_x[p] + y^2\partial_y\partial_x[p] + z^2\partial_z\partial_x[p].\end{aligned}$$

As a last step, we will use that $\hat{p}$ is homogeneous in order to find more convenient expressions for the derivatives. We state this result as a general Lemma for homogeneous polynomials.

Lemma 4.5.20. *Let $r(x,y,z)$ be a homogeneous polynomial of order k in x,y,z, then*

$$y^2\partial_y\partial_x r + z\overset{\partial}{z}\partial_x r = (y+z)(k-1)\partial_x r - yz(\partial_z\partial_x r + \partial_y\partial_x r) - (y+z)x\partial_{xx} r.$$

Proof. The fact that r is homogeneous with degree k implies that $\partial_x r$ is homogeneus with degree $k-1$, then

$$z\partial_z\partial_x r + y\partial_y\partial_x r + x\partial_{xx} r = (k-1)\partial_x r.$$

The last equality implies

$$\begin{aligned}y^2\partial_y\partial_x r &= y(k-1)\partial_x r - yz\partial_z\partial_x r - yx\partial_{xx} r,\\ z^2\partial_z\partial_x r &= z(k-1)\partial_x r - zy\partial_y\partial_x r - zx\partial_{xx} r.\end{aligned}$$

And summing up these two expressions leads to

$$y^2\partial_y\partial_x r + z^2\partial_z\partial_x r = (y+z)(k-1)\partial_x r - yz(\partial_z\partial_x r + \partial_y\partial_x r) - (y+z)x\partial_{xx} r.$$

□

Then we have for $\hat{p}$

$$\begin{aligned}&(\partial_t x)(\partial_x\hat{p}) + (\partial_t y)(\partial_y\hat{p}) + (\partial_t z)(\partial_z\hat{p})|_{t=0}\\ &\qquad = -(n_1 y + n_2 z)\partial_x\hat{p} + (y+z)(k-1)\partial_x\hat{p} - yz(\partial_z\partial_x\hat{p} + \partial_y\partial_x\hat{p}) - (y+z)x\partial_{xx}\hat{p}|_{t=0}.\end{aligned}$$

With this, we are ready to prove that, for an arbitrary initial condition, we can recover the deterministic Jacobi dynamics.

Theorem 4.5.21. *Let $J = (J(t), t \geq 0)$ be an $n \times n$ matrix valued Jacobi process with parameters $n_1 - k$ and $n_2 - k$, respectively. Let $A \in \mathcal{M}_{n,n}$ be any matrix with roots in $[0,1]$. Then the roots of the expected characteristic polynomial of $A + J(t)$ satisfy the deterministic Jacobi equation* (4.5).

Proof. In the above results, we can substitute $y = w, z = (w-1)$ and define the polynomial $j(x,w,t)$ as

$$j(x,w,t) = \hat{p}(x(t), w(t), w(t)-1, t).$$

Using the above results, we have for $j(w,t)$

$$\begin{aligned}&(\partial_t x)(\partial_x j) + (\partial_t w)(\partial_w j)|_{t=0} \\ &\quad = -(n_1 w + n_2 w - n_2))\partial_x j + (2w-1)(k-1)\partial_x j - w(w-1)(\partial_x \partial_w j) - (2w-1)x\partial_{xx} j|_{t=0}.\end{aligned}$$

Letting $x = 0$, we find the evolution equation for the matrix Jacobi process

$$(\partial_t w)\frac{\partial_w j}{\partial_x j}\bigg|_{t=0} = -(n_1 w + n_2 w) + (2w-1)(k-1) - w(w-1)\frac{\partial_x \partial_w j}{\partial_x j}\bigg|_{t=0},$$

then

$$\begin{aligned}\partial_t w &= -[n_1 w + n_2(w-1) + (2w-1)(k-1)]\frac{\partial_x j}{\partial_w j} - w(w-1)\frac{\partial_x \partial_w j}{\partial_w j}, \\ &= [-(n_1 - k + 1)w - (n_2 - k + 1)(w-1)]\frac{\partial_x j}{\partial_w j} - w(w-1)\frac{\partial_x \partial_w j}{\partial_w j}.\end{aligned}$$

To further simplify the last expression we need some hypothesis about the relationship between $\partial_x j$ and $\partial_w j$. We recover this from the definition as the generalized characteristic polynomial.

$$\begin{aligned}j(x,w,t) &= \mathbb{E}\left[\det[xI + (u-1)A_1^T(t)A_1(t) + uA_2^T(t)A_2(t)]\right], \\ &= \mathbb{E}\left[\det[xI + (u-1)A_1^T(t)A_1(t) + u(I - A_1^T(t)A_1(t))]\right], \\ &= \mathbb{E}\left[\det[(x+u)I - A_1^T(t)A_1(t) + (u-u)A_1^T(t)A_1(t)]\right], \\ &= \mathbb{E}\left[\det[(x+u)I - A_1^T(t)A_1(t)]\right].\end{aligned}$$

So this means that the derivatives of j with respect to x are the same as its derivatives with respect to w and then we can equate $\partial_w j$ to $\partial_x j$ to get

$$
\begin{aligned}
\partial_t w_i &= -(n_1-k+1)w_i - (n_2-k+1)(w_i-1) - w_i(w_i-1)\frac{\partial_{xx} j}{\partial_x j}, \\
&= -(n_1-k+1)w_i - (n_2-k+1)(w_i-1) - w_i(w_i-1)\sum_{j\neq i}\frac{2}{w_i-w_j}, \\
&= n_2 - (n_1+n_2)w_i + (w_i+w_i-1)(k-1) - \sum_{j\neq i}\frac{2w_i(w_i-1)}{w_i-w_j}, \\
&= n_2 - (n_1+n_2)w_i + (w_i+w_i-1)\sum_{j\neq i}\frac{w_i-w_j}{w_i-w_j} - \sum_{j\neq i}\frac{2w_i(w_i-1)}{w_i-w_j}, \\
&= n_2 - (n_1+n_2)w_i + \sum_{j\neq i}\frac{(w_i+w_i-1)(w_i-w_j) - 2w_i(w_i-1)}{w_i-w_j}, \\
&= n_2 - (n_1+n_2)w_i + \sum_{j\neq i}\frac{w_i^2 + w_i(w_i-1) - w_i w_j - w_j(w_i-1) - 2w_i(w_i-1)}{w_i-w_j}, \\
&= n_2 - (n_1+n_2)w_i + \sum_{j\neq i}\frac{-w_i w_j - w_j(w_i-1) + w_i}{w_i-w_j}, \\
&= n_2 - (n_1+n_2)w_i + \sum_{j\neq i}\frac{w_j(1-w_i) + w_i(1-w_j)}{w_i-w_j}.
\end{aligned}
$$

□

Conclusions

The contents presented in this thesis come from different areas that are seemingly unrelated but connected through the application to random matrix theory. The thesis has no pretensions of arriving at new results through the connections of these topics, but a deeper work in this direction can lead to mutual enrichment. The study of convolution from the stochastic processes point of view could allow us to get new results. Also, the use of algebraic and combinatorics tools of Free Probability Theory is still not very well explored and could be a starter for new ways to explore these objects.

Two of the most important results in this work are Theorem 4.5.3 in conjunction with Proposition 3.1.2 and Theorem 4.5.10 together with 4.5.13. Each pair of these results tells us that, for the specific case of Hermite and Laguerre polynomials, we can relate the symmetric (correspondingly asymmetric) additive convolution to the evolution of an Itô process over time. This fact has the advantage that allows us to use the several tools of stochastic calculus applied to a different topic. As an example, the proof of Proposition 3.1.2 could be made much simpler using the Lévy process properties of the Brownian motion. A slightly less trivial example is that, by applying 4.3.3 with the former theorem and proposition relating the convolution to the evolution of the process, we can conclude the original result of [39] that the minimal distance of the roots in a polynomial always grows when doing symmetric additive convolution with another polynomial. In the same line, an implication of Theorem **??** is that the asymmetric additive convolution does not satisfy the same property, i.e. we have found an example of an asymmetric additive convolution of polynomials in which the mesh (minimal distance between roots) does not grow.

Given the existence of several algorithms in Machine Learning and multivariate statistics that are modeled by the transformation of a random matrix [40], [41], it is natural to expect that stochastic calculus could be applied to deduce results about the behavior of the system over time. This approach has already been proved useful in the work of Bru to study the behavior of Principal Analysis Components under a random perturbation [12].

On the other hand, the proofs of some of the results about the matrix-valued processes are greatly simplified with the tools of Finite Free Probability Theory. As an example, the relationship of Jacobi polynomials to the Jacobi matrix distribution is much shorter using the properties of convolution than the one found in [35]. Similar ideas could be applied to get analogous results for other matrix-valued stochastic processes.

There are several problems related to the relationship between the topics in the thesis that have not been covered or even mentioned. In the first place, most of the work was focused on symmetric additive convolution. While this is arguably the notion of convolution that can be the most easily related to stochastic processes, similar results could be obtained for the other two notions and they could be applied to processes with multiplicative increments, as an example. The exploration of more matrix-valued stochastic models and how they can be studied from stochastic analysis and finite free probability points of view is still a pending

topic.

Finally, some questions arose during the realization of this work that were left behind because they surpassed the goals and time limitations of the thesis. As an example, it would be of interest what is the behavior of the stochastic expected polynomial of a matrix-valued diffusion process. This more general approach would deal not only with expected values but also with more general properties of the eigenvalues and maybe it could be related to Finite Free Probability Theory using convolution of random polynomials. A natural question concerning the last section would be to find an explicit representation for the Jacobi polynomials considering different starting positions, prove that their roots converge to the stable distribution, and evaluate the velocity of convergence.

In sum, this master's thesis serves as an exploration of how two different branches of probability can be related through the study of a common object of study. While some of the results presented here, can be interesting, there are a lot of questions and opportunities to explore deeper connections between these areas.

Bibliography

[1] F. J. Dyson, "A brownian-motion model for the eigenvalues of a random matrix," *Journal of Mathematical Physics*, vol. 3, no. 6, pp. 1191–1198, Jun. 22, 1962, ISSN: 0022-2488. DOI: {10.1063/1.1703862}. eprint: {https://pubs.aip.org/aip/jmp/article-pdf/3/6/1191/8158355/1191_1_online.pdf}. [Online]. Available: {https://doi.org/10.1063/1.1703862}.

[2] P. Graczyk and J. Małecki, "Multidimensional yamada-watanabe theorem and its applications," *Journal of Mathematical Physics*, vol. 54, Aug. 2011. DOI: 10.1063/1.4790507.

[3] A. W. Marcus, D. A. Spielman, and N. Srivastava, "Finite free convolutions of polynomials," *Probability Theory and Related Fields*, vol. 182, no. 3-4, pp. 807–848, 2022.

[4] D. Holcomb and E. Paquette, "Tridiagonal models for dyson brownian motion," *arXiv preprint arXiv:1707.02700*, 2017.

[5] J. A. Mingo and R. Speicher, *Free probability and random matrices*. Springer, 2017, vol. 35.

[6] F. C. Klebaner, *Introduction to stochastic calculus with applications*. World Scientific Publishing Company, 2012.

[7] I. Karatzas and S. Shreve, *Brownian motion and stochastic calculus*. springer, 2014, vol. 113.

[8] D. Revuz and M. Yor, *Continuous martingales and Brownian motion*. Springer Science & Business Media, 2013, vol. 293.

[9] J.-F. Le Gall, *Brownian motion, martingales, and stochastic calculus*. Springer, 2016.

[10] E. Mayerhofer, O. Pfaffel, and R. Stelzer, "On strong solutions for positive definite jump diffusions," *Stochastic processes and their applications*, vol. 121, no. 9, pp. 2072–2086, 2011.

[11] E. A. Trujillo Rivera, "El proceso de Wishart y la dinámica de sus eigenvalores y propiedades distribucionales vía cálculo estocástico," M.S. thesis, CIMAT, 2011.

[12] M.-F. Bru, "Diffusions of perturbed principal component analysis," *Journal of multivariate analysis*, vol. 29, no. 1, pp. 127–136, 1989.

[13] A. Nica and R. Speicher, *Lectures on the Combinatorics of Free Probability* (London Mathematical Society Lecture Note Series). Cambridge University Press, 2006.

[14] D. Perales Anaya, "Cumulantes en probabilidad libre finita," M.S. thesis, Universidad Nacional Autónoma de México, 2016.

[15] R. Allez and A. Guionnet, "A diffusive matrix model for invariant β-ensembles," *Electronic Journal of Probability*, vol. 18, no. none, pp. 1 –30, 2013. DOI: 10.1214/EJP.v18-2073. [Online]. Available: https://doi.org/10.1214/EJP.v18-2073.

[16] H. P. McKean, *Stochastic integrals*. American Mathematical Soc., 1969, vol. 353.

[17] J. Wishart, "The generalised product moment distribution in samples from a normal multivariate population," *Biometrika*, vol. 20A, no. 1-2, pp. 32–52, Dec. 1928, ISSN: 0006-3444. DOI: 10.1093/biomet/20A.1-2.32. eprint: https://academic.oup.com/biomet/article-pdf/20A/1-2/32/530655/20A-1-2-32.pdf. [Online]. Available: https://doi.org/10.1093/biomet/20A.1-2.32.

[18] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning* (Springer Series in Statistics). New York, NY, USA: Springer New York Inc., 2001.

[19] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*. Wiley-Interscience, 2005.

[20] C. F. Van Loan, "Generalizing the singular value decomposition," *SIAM Journal on Numerical Analysis*, vol. 13, no. 1, pp. 76–83, 1976. DOI: 10.1137/0713009. eprint: https://doi.org/10.1137/0713009. [Online]. Available: https://doi.org/10.1137/0713009.

[21] A. W. Marcus, "Finite free point processes," *arXiv preprint arXiv:2205.00495*, 2022.

[22] Y. Doumerc, "Matrices aléatoires, processus stochastiques et groupes de réflexions," Ph.D. dissertation, Toulouse 3, 2005.

[23] B. Collins, "Intégrales matricielles et Probabilités Non-Commutatives," Theses, Université Pierre et Marie Curie - Paris VI, Jan. 2003. [Online]. Available: https://theses.hal.science/tel-00004306.

[24] S. Asmussen and P. W. Glynn, *Stochastic simulation: algorithms and analysis*. Springer, 2007, vol. 57.

[25] A. W. Marcus, *Polynomial convolutions and (finite) free probability*, 2021. arXiv: 2108.07054 [math.CO].

[26] J. L. Walsh, "On the location of the roots of certain types of polynomials," *Transactions of the American Mathematical Society*, vol. 24, no. 3, pp. 163–180, 1922.

[27] G. Szegö, *Orthogonal polynomials*. American Mathematical Soc., 1939, vol. 23.

[28] A. W. Marcus, "Polynomial convolutions and (finite) free probability," *arXiv preprint arXiv:2108.07054*, 2021.

[29] W Schoutens, "Orthogonal polynomials and stochastic processes," *Lecture Notes in Statistics, Springer*, vol. 960, 2000.

[30] D. Babusci, G. Dattoli, S. Licciardi, and E. Sabia, *Mathematical methods for physicists*. World Scientific, 2019.

[31] P. de la Harpe, *The fuglede-kadison determinant, theme and variations*, 2012. arXiv: 1107.1059 [math.OA]. [Online]. Available: https://arxiv.org/abs/1107.1059.

[32] O. Arizmendi and D. Perales, "Cumulants for finite free convolution," *Journal of Combinatorial Theory, Series A*, vol. 155, pp. 244–266, 2018, ISSN: 0097-3165. DOI: https://doi.org/10.1016/j.jcta.2017.11.012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0097316517301668.

[33] P. Deift, *Orthogonal Polynomials and Random Matrices: A Riemann-Hilbert Approach: A Riemann-Hilbert Approach*. American Mathematical Soc., 2000, vol. 3.

[34] A. Edelman, "Eigenvalues and condition numbers of random matrices," *SIAM journal on matrix analysis and applications*, vol. 9, no. 4, pp. 543–560, 1988.

[35] K. Aomoto, "Jacobi polynomials associated with selberg integrals," *SIAM Journal on Mathematical Analysis*, vol. 18, no. 2, pp. 545–549, 1987.

[36] I. Dumitriu and A. Edelman, "Matrix models for beta ensembles," *Journal of Mathematical Physics*, vol. 43, no. 11, pp. 5830–5847, 2002.

[37] R. A. Lippert, "A matrix model for the β-Jacobi ensemble," *Journal of Mathematical Physics*, vol. 44, no. 10, pp. 4807–4816, Oct. 2003, ISSN: 0022-2488. DOI: 10.1063/1.1604184. eprint: https://pubs.aip.org/aip/jmp/article-pdf/44/10/4807/19006721/4807_1_online.pdf. [Online]. Available: https://doi.org/10.1063/1.1604184.

[38] A. Edelman and B. D. Sutton, "The beta-jacobi matrix model, the cs decomposition, and generalized singular value problems," *Foundations of Computational Mathematics*, vol. 8, no. 2, pp. 259–285, 2008.

[39] P. Brändén, I. Krasikov, and B. Shapiro, "Elements of pólya-schur theory in the finite difference setting," *Proceedings of the American Mathematical Society*, vol. 144, no. 11, pp. 4831–4843, 2016.

[40] J. Pennington and Y. Bahri, "Geometry of neural network loss surfaces via random matrix theory," in *Proceedings of the 34th International Conference on Machine Learning*, D. Precup and Y. W. Teh, Eds., ser. Proceedings of Machine Learning Research, vol. 70, PMLR, 2017, pp. 2798–2806. [Online]. Available: https://proceedings.mlr.press/v70/pennington17a.html.

[41] M. S. Srivastava, *Methods of multivariate statistics*. John Wiley & Sons, 2002, vol. 419.

Appendix A

Code for the path simulations

```
#%% Euler-Maruyama method for eigenvalue SDE solutions
import numpy as np
import matplotlib.pyplot as plt
import matplotlib

def dyson(dt=0.1,time=10,n=5,starting=np.linspace(start=0, stop=5, num
=5)):
    num_points = int(time / dt)
    simulation = np.zeros((num_points,n))
    for l in range(n):
        simulation[0,l] = starting[l]
    sqrtime = np.sqrt(dt)
    for i in range(num_points-1):
        for j in range(n):
            mu = 0
            for k in range(n):
                diff = simulation[i,j] - simulation[i,k]
                if diff != 0:
                    mu += 1/diff
            simulation[i+1,j] = simulation[i,j] + dt*mu + sqrtime*np.
random.randn()
    return simulation

def deterministic_dyson(dt=0.1,time=10,n=5,starting=np.linspace(start=0,
 stop=5, num=5)):
    num_points = int(time / dt)
    simulation = np.zeros((num_points,n))
    for l in range(n):
        simulation[0,l] = starting[l]
    sqrtime = np.sqrt(dt)
    for i in range(num_points-1):
        for j in range(n):
            mu = 0
            for k in range(n):
                diff = simulation[i,j] - simulation[i,k]
                if diff != 0:
                    mu += 1/diff
            simulation[i+1,j] = simulation[i,j] + dt*mu
    return simulation

def wishart(dt=0.1,time=10,n=5,starting=np.linspace(start=0, stop=5, num
=5)):
```

```
        num_points = int(time / dt)
        simulation = np.zeros((num_points,n))
        for l in range(n):
            simulation[0,l] = starting[l]
        sqrtime = np.sqrt(dt)
        for i in range(num_points-1):
            for j in range(n):
                mu = 0
                for k in range(n):
                    diff = simulation[i,j] - simulation[i,k]
                    if diff != 0:
                        mu += (simulation[i,k]+simulation[i,j])/diff
                simulation[i+1,j] = simulation[i,j] + dt*(mu+n) + sqrtime*np.random.randn()*np.sqrt(simulation[i,j])
        return simulation

    def deterministic_wishart(dt=0.1,time=10,n=5,starting=np.linspace(start=0, stop=5, num=5)):
        num_points = int(time / dt)
        simulation = np.zeros((num_points,n))
        for l in range(n):
            simulation[0,l] = starting[l]
        sqrtime = np.sqrt(dt)
        for i in range(num_points-1):
            for j in range(n):
                mu = 0
                for k in range(n):
                    diff = simulation[i,j] - simulation[i,k]
                    if diff != 0:
                        mu += (simulation[i,k]+simulation[i,j])/diff
                simulation[i+1,j] = simulation[i,j] + dt*(mu+n)
        return simulation

    def jacobi(dt=0.1,time=10,n=5,starting=np.linspace(start=0.1, stop=0.9, num=5),n_1=5,n_2=5):
        num_points = int(time / dt)
        simulation = np.zeros((num_points,n))
        for l in range(n):
            simulation[0,l] = starting[l]
        sqrtime = np.sqrt(dt)
        for i in range(num_points-1):
            for j in range(n):
                mu = 0
                for k in range(n):
                    diff = simulation[i,j] - simulation[i,k]
                    if diff != 0:
                        mu += (simulation[i,j]*(1-simulation[i,k]) + simulation[i,k]*(1-simulation[i,j]))/diff
                simulation[i+1,j] = simulation[i,j] + dt*(mu+n_2-(n_1+n_2)*simulation[i,j]) + 2*sqrtime*np.random.randn()*np.sqrt(simulation[i,j]*(1-simulation[i,j]))
        return simulation

    def deterministic_jacobi(dt=0.1,time=10,n=5,starting=np.linspace(start=0.1, stop=0.9, num=5),n_1=5,n_2=5):
        num_points = int(time / dt)
        simulation = np.zeros((num_points,n))
        for l in range(n):
            simulation[0,l] = starting[l]
```

```
91          sqrtime = np.sqrt(dt)
92          for i in range(num_points-1):
93              for j in range(n):
94                  mu = 0
95                  for k in range(n):
96                      diff = simulation[i,j] - simulation[i,k]
97                      if diff != 0:
98                          mu += (simulation[i,j]*(1-simulation[i,k]) +
    simulation[i,k]*(1-simulation[i,j]))/diff
99                  simulation[i+1,j] = simulation[i,j] + dt*(mu+n_2-(n_1+n_2)*
    simulation[i,j])
100         return simulation
```

Listing A.1: Euler-Maruyama algorithm for simulation of eigenvalue processes