CIMAT Centro de Investigación en Matemáticas, A.C.

BERRY-ESSEEN TYPE THEOREMS FOR BOOLEAN AND MONOTONE CENTRAL LIMIT THEOREMS

T E S I S

Que para obtener el grado de

Doctor en Ciencias

con Orientación en

Matemáticas Básicas

Presenta

Mauricio Salazar Méndez

Director de Tesis:

Dr. Octavio Arizmendi Echegaray

Autorización de la versión final

Guanajuato, Gto., Julio de 2019

Abstract

In this thesis we study the speed of convergence in the Boolean and monotone central limit theorems. We also investigate some ergodic properties for certain transformations on the real line.

In the Boolean central limit theorem, we obtain a bound of order $\frac{1}{\sqrt[3]{n}}$ of the speed of convergence for measures with finite fourth moment. The proof is based on the Boolean cumulants. When the measures have bounded support, we get a bound of order $\frac{1}{\sqrt{n}}$ of the speed of convergence, and we show that this bound is sharp. We derive this result from a more general theorem describing very explicitly the convergence in the Boolean central limit theorem. These results are in terms of the Lévy metric.

For the monotone central limit theorem, we obtain a bound of order $\frac{1}{\sqrt[8]{n}}$ of the speed of convergence for measures with finite fourth moment. We improve this bound for the case of measures with finite sixth moment. We get a bound of order $\frac{1}{\sqrt[4]{n}}$, and we prove that this bound is sharp. The proofs of these results are based on the complex analysis methods of non-commutative probability.

Finally, the F-transform of measures singular to the Lebesgue measure induces a transformation on the real line which preserves the Lebesgue measure. We prove that for measures of zero mean and unit variance these transformations are pointwise dual ergodic with return sequence $\frac{\sqrt{2n}}{\pi}$, extending a previous result of Aaronson.

Acknowledgments

Primeramente, agradezco a mi asesor el Dr. Octavio Arizmendi por invitarme y guiarme en el estudio de un área tan interesante y fértil como lo es la probabilidad no-conmutativa.

Me gustaría agradecer a Octavio Arizmendi y Jiun-Chau Wang por permitirme incluir los resultados en [7] y [6] en esta tesis.

Agradezco a Cesar Maldonado, Jiun-Chau Wang, Víctor Pérez, y Carlos Vargas por ser parte del jurado en mi defensa de tesis y por sus comentarios tan valiosos.

Quiero hacer un agradecimiento especial a la maestra Helga Fetter por su gran apoyo en los primeros años de mi doctorado.

Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACYT) por la beca otorgada durante mi doctorado. También agradezco a CIMAT por la beca de elaboración de tesis otorgada en los últimos meses de mi doctorado.

Para el departamento de servicios escolares un enorme agradecimiento. Siempre atendían con amabilidad y eficacia.

Agradezco a mis padres por su apoyo incondicional y por haberme impulsado a estudiar matemáticas, ¡originalmente quería estudiar electrónica! También agradezco a mis hermanos Saúl, Isaías y Viridiana por apoyarme en distintas situaciones.

Agradezco infinitamente a mi hermosa novia Thalía por darme su apoyo, paciencia y comprensión durante este proceso del doctorado, sé que no fue ¡fácil! Un especial agradecimiento a Pancho, mi gran compañero de parrandas. Y también, por supesto, un agradecimiento sincero a Terán, Francisco, PPL y Chiza por la amistad brindada durante estos años.

Contents

Chapter 1

Introduction

This thesis fits into the areas of non-commutative probability and infinite ergodic theory. In non-commutative probability theory, we obtain some Berry-Essen type estimates for the Boolean and monotone central limit theorems. On the side of infinite ergodic theory, we improve a result of Aaronson about the ergodicity of a class of transformations on the real line. This class includes most of the known transformations on the real line that preserve the Lebesgue measure.

Non-Commutative Probability

Non-commutative (or quantum) probability has its origins in the works of Cushen and Hudson [14], Hudson [20], Giri and von Waldenfels [17], von Waldenfels [33], and Voiculescu [30]. Their work leads to generalizing the concepts of random variable, expectation, and distribution to a non-commutative setting. More importantly, it was realized that in the non-commutative world the notion of independence is a broader concept and that other types of independence exist besides the classical one. The existence of new types of independence attracted much attention. This because, at least theoretically, a new notion of independence may lead to new central limit theorems and other important analogs to the ones in classical probability such as stochastic processes. Over time several notions of independence have been introduced. Some of the most studied are the following: free independence [30], Boolean independence [29], operator-valued free independence [31], c-free independence [12], q-independence [27], monotone independence [25], and bi-free independence [32]. For each of the new notions of independence a central limit theorem was also obtained.

Now, there is an important result that gives a more quantitative version of the classical central limit theorem. This result is known as the Berry-Esseen Theorem ([11], [15]), and it establishes a bound for the rate of convergence in the classical central limit theorem. Let μ be a probability measure of zero mean and unit variance. The Berry-Esseen theorem states that if $\int_{\mathbb{R}} |x|^3 d\mu < \infty$, then the distance to the standard Gaussian distribution $\mathcal{N}$ is bounded for n big enough as follows

$$d_{kol}(D_{\frac{1}{\sqrt{n}}}\mu^{*n}, \mathcal{N}) \leq C\frac{\int_{\mathbb{R}} |x|^3 d\mu}{\sqrt{n}},$$

where $*$ denotes the classical convolution, d_{kol} stands for the Kolmogorov distance, $D_b\mu$ designates the dilation of a measure μ by a factor $b > 0$, and C is an absolute constant. Moreover, this estimate is sharp, so there is a measure μ_0 such that $d_{kol}(D_{\frac{1}{\sqrt{n}}}\mu_0^{*n}, \mathcal{N}) \geq \frac{c}{\sqrt{n}}$ as $n \to \infty$.

The great importance of the Berry-Esseen theorem is that, as in any limit theorem, for many applications of the classical central limit theorem one needs a quantitative version of it; a version where an estimation of the speed of convergence is obtained. So, as non-commutative probability brought new notions of independence and its respective central limit theorem, it is important to obtain quantitative versions (Berry-Esseen type estimates) for them. Next, we collect some of the main results obtained in this direction. In the free central limit theorem, a Berry-Esseen type estimate was given by Kargin [21] for the bounded case and then broadly improved by Chistyakov and Götze [13] for measures with finite fourth moment; if μ is a probability measure with $m_1(\mu) = 0$, $m_2(\mu) = 1$, and $m_4(\mu) < \infty$, then the distance to the standard semicircle distribution $\mathcal{S}$ satisfies for n large

$$d_{kol}(D_{\frac{1}{\sqrt{n}}}\mu^{\boxplus n}, \mathcal{S}) \leq C' \frac{|m_3(\mu)| + |m_4(\mu)|^{1/2}}{\sqrt{n}},$$

where the symbol $\boxplus$ denotes the free convolution, and C' is an absolute constant. Here, as in the classical Berry-Esseen theorem, the estimate is sharp.

Now for the operator-valued free central limit theorem, Speicher [24] obtained that the rate of convergence in is bounded by the order $\frac{1}{\sqrt{n}}$. In the finite-free probability setting, Arizmendi and Perales [5] also show that the rate of convergence in is bounded by the order $\frac{1}{\sqrt{n}}$.

This thesis enters in this line of research.

Main Objective: To obtain Berry-Esseen type estimates for the Boolean and monotone central limit theorems.

For the Boolean central limit theorem, the first thing one notes is that convergence in the Kolmogorov distance does not hold. Thus, we consider the Lévy distance instead, which seems the most appropriate.

Our first contribution considers measures with finite fourth moment.

Theorem 5. *Let μ be a probability measure such that $m_1(\mu) = 0$, $m_2(\mu) = 1$, and $m_4(\mu) < \infty$. Then for the measure $\mu_n = D_{\frac{1}{\sqrt{n}}}\mu^{\uplus n}$, where $\uplus$ denotes the Boolean convolution, we have that*

$$d_{lev}(\mu_n, \mathbf{b}) \leq \frac{7}{2}\sqrt[3]{\frac{m_4(\mu) - 1}{n}} \quad \text{for } n \geq 1.$$

The proof of this theorem is a direct consequence of a general quantitative estimate of the distance to Bernoulli distribution in terms of the fourth moment: $d_{lev}(\mu, \mathbf{b}) \leq \frac{7}{2}\sqrt[3]{m_4(\mu) - 1}$.

We think that the estimate in the above theorem is not sharp. We believe that the sharp rate must be of order $\frac{1}{\sqrt{n}}$.

Our second contribution specializes in the case of measures with bounded support. In this case we give a sharp bound for the rate of convergence.

Theorem 6. *Let μ be a probability measure such that $m_1(\mu) = 0$, $m_2(\mu) = 1$, and $supp(\mu) \subset [-K, K]$. Then the measure $\mu_n := D_{\frac{1}{\sqrt{n}}}\mu^{\uplus n}$ satisfies for $\sqrt{n} > K$ that:*

1) *supp $\mu_n \subset [\frac{-K}{\sqrt{n}}, \frac{K}{\sqrt{n}}] \cup \{x_1, x_2\}$, where $|(-1) - x_1| \leq \frac{K}{\sqrt{n}}$ and $|1 - x_2| \leq \frac{K}{\sqrt{n}}$.*

2) *For $p = \mu_n(\{x_1\})$, $q = \mu_n(\{x_2\})$ and $r = \mu_n([\frac{-K}{\sqrt{n}}, \frac{K}{\sqrt{n}}])$, we have that $p, q \in [\frac{1}{2} - \frac{2K}{\sqrt{n}}, \frac{1}{2} + \frac{K}{2\sqrt{n}}]$ and $r < \frac{4K}{\sqrt{n}}$.*

In particular, the Lévy distance between μ_n and $\mathbf{b}$ is bounded by

$$d_{lev}(\mu_n, \mathbf{b}) \leq \frac{2K}{\sqrt{n}}.$$

Now, we present our results for the monotone case. In this case we consider the Kolmogorov distance as in the classical and free cases. Recall that the arcsine distribution $\mathbf{a}$ is given by the density $d\mathbf{a}(t) := \frac{1}{\pi\sqrt{2-t^2}}$ for $t \in (-\sqrt{2}, \sqrt{2})$.

Our first estimation treats measures with finite fourth moment.

Theorem 7. *Let μ be a probability measure. Suppose that $m_1(\mu) = 0$, $m_2(\mu) = 1$, and $m_4(\mu) < \infty$. Let $\mu_n = D_{1/\sqrt{n}}\mu^{\triangleright n}$, $n \geq 1$. Then there exists a constant k depending only on μ such that we have the Kolmogorov distance*

$$d_{kol}(\mu_n, \mathbf{a}) \leq kn^{-1/8}$$

We do not know if the bound in this theorem is sharp.

In our following theorem we obtain a sharp bound for the rate of convergence, but we weaken a bit the generality.

Theorem 8. *Let μ be a probability measure. Suppose that $m_1(\mu) = 0$, $m_2(\mu) = 1$, and $m_6(\mu) < \infty$.*

Let $\mu_n = D_{1/\sqrt{n}}\mu^{\triangleright n}$, $n \geq 1$. Then there exists a constant k depending only on μ such that we have the Kolmogorov distance

$$d_{kol}(\mu_n, \mathbf{a}) \leq kn^{-1/4}$$

Infinite Ergodic Theory

Infinite ergodic theory studies the ergodic properties of measure-preserving transformations when the measure of the underlying space is infinite. The transformations on the real line that preserve the Lebesgue measure were one of the first situations studied. Next, we discuss some of the main results obtained for this situation.

In 1973 Adler and Weiss [3] proved that the Boole transform $T(x) = x - \frac{1}{x}$, for $x \in \mathbb{R} \setminus \{0\}$, preserve the Lebesgue measure and is ergodic. Later, Li and Schweiger [23] proved that the generalized Boole transform $T(x) = x - \sum_{i=1}^{n} \frac{p_i}{x-a_i}$, for $x \in \mathbb{R} \setminus \{a_i \mid i = 1, ..., n\}$, where $p_i \geq 0$ and $a_i \in \mathbb{R}$, also preserves the Lebesgue measure and is ergodic. In [28] Schweiger showed the same conclusions are valid for $\tan(x)$, and in [22] similar results were obtained for other real transformations preserving the Lebesgue.

Aaronson [1] obtains a generalization that includes all the past results. Consider the F-transform $F_\mu(z) = [\int_{\mathbb{R}} \frac{1}{z-t} d\mu(t)]^{-1}$ of a positive measure μ over $\mathbb{R}$. For singular measures μ with respect to the Lebesgue measure, it turns out that we can extend the F-transform F_μ to $\mathbb{R}$, that is, the limit $T(x) := \lim_{y \to 0} F_\mu(x + iy)$ exists and is real a.e. in $x \in \mathbb{R}$. Moreover, if μ is a probability measure, then $T(x)$ preserves the Lebesgue measure. The nice observation is that all the transformations that we mention above are obtained in this way. For example, the Boole transform is the extension to $\mathbb{R}$ of the F-transform $F_{\mathbf{b}}(z) = z - \frac{1}{z}$ of the Bernoulli distribution $\mathbf{b} = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$.

Now, Aaronson proved that under additional hypothesis on the measure μ, the transformation on $\mathbb{R}$ obtained as above is ergodic.

Theorem (Aaronson). *Let μ be a probability measure of zero mean and unit variance. Suppose that μ is singular with respect to the Lebesgue measure and that the support of μ is bounded. Then the real restriction of F_μ,*

$$T(x) := \lim_{y \to 0} F_\mu(x + iy)$$

is conservative and ergodic transform. Moreover, T is pointwise dual ergodic with return sequence $\frac{\sqrt{2n}}{\pi}$.

In [34], Wang proves that the conclusions of T being conservative and ergodic still holds if we omit the hypothesis of bounded support. We extend the result of Wang, by proving that it is also true that T is pointwise dual ergodic with return sequence $\frac{\sqrt{2n}}{\pi}$.

Theorem 12. *Let μ be a probability measure of zero mean and unit variance. Suppose that μ is singular with respect to the Lebesgue measure. Then the real restriction of F_μ*

$$T(x) := \lim_{y \to 0} F_\mu(x + iy)$$

is pointwise dual ergodic with return sequence $\frac{\sqrt{2n}}{\pi}$.

This thesis is organized as follows:

- In Chapter 2 we give a brief introduction to non-commutative probability theory passing through the relevant concepts and properties involved in our research about the speed of convergence of the Boolean and monotone central limit theorem.
- In Chapter 3 we establish a theorem that bounds the Lévy distance of a measure to the Bernoulli distribution in terms of the fourth moment. We use this theorem and the theory of Boolean cumulants to prove the theorem of the rate of convergence in the Boolean central limit theorem for measures of finite fourth moment. We also prove the estimate for measures of bounded support and give an example that shows that the bound is sharp.
- In Chapter 4 we give the proofs of the theorems concerning the speed of convergence in the monotone central limit theorem. The proofs are based on the complex analytic tools of non-commutative probability.
- In Chapter 5 we first review the basics of infinite ergodic theory and the work of Aaronson about inner function on the upper half-plane. Then we establish Theorem 12.

Chapter 2

Non-Commutative Probability

Non-commutative probability theory generalizes classical probability theory. In this new framework, the notion of independence loses the intrinsic role that it has in classical probability, and new types of independence appear. For some of those new types of independence it is possible to develop parallel theories with the ones in classical probability.

In this chapter we give a brief introduction to the theory of non-commutative probability focusing on the Boolean and monotone notions of independence. We first review the main analytic tools of the theory. Then we explain the main ingredients and the general objective of the theory. Finally, we study the sum of random variables that are Boolean or monotone independent.

2.1 Analytic Tools

In this section we introduce some transforms that assign to a measure on $\mathbb{R}$ an analytic complex function. This transforms form the analytic machinery of non-commutative probability theory. We also define the Kolmogorov and Lévy distances, and we recall Bai's theorem to estimate the Kolmogorov distance. These concepts are used to study the convergence of measures in the subsequent chapters.

2.1.1 Notation

We denote by $\mathcal{M}$ the set of all Borel probability measures and by $\mathcal{M}_0^1$ the subset of $\mathcal{M}$ of measures with zero mean and unit variance. Let μ be a probability measure. We write $m_n(\mu)$ for the n-th moment of μ. The notation $\mathcal{F}\mu$ denotes the cumulative distribution function of μ, that is $\mathcal{F}\mu(x) = \mu((-\infty, x])$ for $x \in \mathbb{R}$. The support of a measure $\mu \in \mathcal{M}$ is denoted by $supp(\mu)$. For $\mu \in \mathcal{M}$ let $D_a\mu$ denote the dilation of a measure μ by a factor $a > 0$; this means that $D_a\mu(B) = \mu(a^{-1}B)$ for all Borel sets $B \subset \mathbb{R}$.

By $\mathbb{C}^+$ and $\mathbb{C}^-$ we denote the open upper and lower complex half-planes, respectively.

2.1.2 The Cauchy transform

The Cauchy transform (a.k.a. Stieltjes transform) associates to each Borel measure on $\mathbb{R}$ an analytic complex function. Many features of the measure like the support or the moments become now information of an analytic function that we can study using the tools from complex analysis.

Definition 1. *The **Cauchy transform** of a positive measure* μ *is defined as*

$$G_\mu(z) := \int_{\mathbb{R}} \frac{1}{z-t} d\mu(t) \quad for\, z \in \mathbb{C} \setminus supp(\mu). \tag{2.1}$$

For a measure $\mu \in \mathcal{M}$, the Cauchy transform is analytic in $\mathbb{C} \setminus supp(\mu)$ and maps $\mathbb{C}^+$ to $\mathbb{C}^-$ (and vice versa). We can recover a measure $\mu \in \mathcal{M}$ from its Cauchy transform. This result is known as the **Stieltjes inversion formula**:

$$\mu((a,b]) = \lim_{\delta \downarrow 0} \lim_{\epsilon \downarrow 0} -\frac{1}{\pi} \int_{a+\delta}^{b+\delta} Im(G_\mu(x+i\epsilon))dx. \tag{2.2}$$

This formula implies that the Cauchy transform G_μ determines completely the measure μ. Next, we present some important properties of the Cauchy transform.

Proposition 1. *Let* μ *be a Borel measure on* $\mathbb{R}$*. Then the Cauchy transform of the dilation* $D_a\mu$ *is given by*

$$G_{D_a\mu}(z) = \frac{1}{a} G_\mu(\frac{z}{a}) \quad for\, z \in \mathbb{C}^+. \tag{2.3}$$

Proposition 2. *Let* μ *be a Borel measure on* $\mathbb{R}$*. Then*

$$|G_\mu(z)| \leq \frac{\mu(\mathbb{R})}{Imz}, \quad for\ z \in \mathbb{C}^+. \tag{2.4}$$

Proposition 3. *Let* μ *be a Borel measure on* $\mathbb{R}$*. If* $supp(\mu) \subset [-K, K]$ *, then*

$$G_\mu(x) \neq 0 \quad for\, x \in \mathbb{R} \setminus [-K, K]. \tag{2.5}$$

The following proposition is essentially Proposition 5.1 ii) in [10]. It tells us that the asymptotic behavior of the Cauchy transform is quite simple.

Proposition 4. *Let* $z_n \in \mathbb{C}^+$ *such that* $z_n \to \infty$ *non-tangentially, that is,* $|z_n| \to \infty$ *and for some* $\alpha > 0$ *we have* $|Re(z_n)| \leq \alpha Im(z)$ *for all* n*. Then*

$$\lim_{n\to\infty} z_n G_\mu(z_n) \to \mu(\mathbb{R}). \tag{2.6}$$

Proposition 5. *If* $\int |t| d\mu(t) < \infty$*, then*

$$G_\mu(z) = \frac{\mu(\mathbb{R})}{z} + \frac{\mu(\mathbb{R})}{z} \int_{\mathbb{R}} \frac{t}{z-t} d\mu(t) \quad for\ z \in \mathbb{C}^+. \tag{2.7}$$

Proposition 6. *Let* μ *be a finite Borel measure over* $\mathbb{R}$*. Suppose that* $\int_{\mathbb{R}} |t|^n \mu < \infty$*. Then we have the expansion*

$$G_\mu(z) = \sum_{i=0}^{n} m_i(\mu) \frac{1}{z^{i+1}} + o(\frac{1}{z^{n+1}}), \quad as\ z \to \infty\ non\text{-}tangentially,$$

where $m_i(\mu) := \int_{\mathbb{R}} t^i \mu$*.*

Conversely, if n *is even and* $G_\mu(z)$ *admits the previous expansion for real numbers* m_i $i = 0, ..., n$*, then* $m_i = \int_{\mathbb{R}} t^i \mu$ *for* $i = 0, ..., n$*.*

2.1.3 The F-transform

The F-transform is defined as the multiplicative inverse of the Cauchy transform. This transform plays a central role in non-commutative probability theory. Many important operations of the theory (as the sum of "independent" random variables) can be expressed naturally in terms of the F-transform.

Definition 2. *The F-**transform** of a positive measure μ is the holomorphic function $F_\mu : \mathbb{C}^+ \to \mathbb{C}^+$ defined as*

$$F_\mu(z) := \frac{1}{G_\mu(z)} \quad \text{for } z \in \mathbb{C}^+. \tag{2.8}$$

One reason why the F-transform is so useful is that it is, in particular, a Nevanlinna function, and these functions are well understood.

Now, we present some properties of the F-transform relevant for us.

Proposition 7. *For $b > 0$ we have*

$$F_{D_b\mu}(z) = bF_\mu(\frac{z}{b}) \quad \text{for } z \in \mathbb{C}^+. \tag{2.9}$$

The F-transform can be used to identify potential atoms of a measure.

Proposition 8. *If $a \in \mathbb{R}$ is an isolated atom of $\mu \in \mathcal{M}$, then $F_\mu(a) = 0$.*

The following proposition establishes a very helpful connection between the asymptotics at infinity of the Cauchy transform and the F-transform.

Proposition 9. *Let μ be a positive measure. Let G_μ and F_μ be the corresponding Cauchy and F-transform respectively. Then G_μ admits the representation*

$$G_\mu(z) = \frac{m_0}{z} + \frac{m_1}{z^2} + \cdots + \frac{m_n}{z^{n+1}} + o(\frac{1}{z^{n+1}}), \quad \text{as } z \to \infty, \tag{2.10}$$

where $m_0, m_1, ..., m_n$ are real numbers, if and only if F_μ admits the representation

$$F_\mu(z) = z + s_1 + \frac{s_2}{z} + \frac{s_3}{z^2} + \cdots + \frac{s_n}{z^{n-1}} + o(\frac{1}{z^{n-1}}), \quad \text{as } z \to \infty, \tag{2.11}$$

where $s_1, ..., s_n$ are real numbers.

Proof. We proof the forward implication. The proof of the backward implication is analogous. Without loss of generality, we assume that $m_0 = 1$. All the asymptotics in this proof are for $z \to \infty$. We have that

$$\begin{aligned} F_\mu(z) &= \frac{1}{G_\mu(z)} \\ &= \frac{1}{\frac{1}{z} + \frac{m_1}{z^2} + \cdots + \frac{m_n}{z^{n+1}} + o(\frac{1}{z^{n+1}})} \\ &= \frac{1}{\frac{1}{z}(1 + \frac{m_1}{z} + \cdots + \frac{m_n}{z^n} + o(\frac{1}{z^n}))}. \end{aligned}$$

Put $p(z) = \frac{m_1}{z} + \cdots + \frac{m_n}{z^n}$. We have that

$$\frac{1}{1 + p(z) + o(\frac{1}{z^n})} = \frac{1}{1 + p(z)} - \frac{o(\frac{1}{z^n})}{(1 + p(z))(1 + p(z) + o(\frac{1}{z^n}))}. \tag{2.12}$$

For $|z|$ large enough $|p(z)|$ is small, so we have that

$$\begin{aligned}\frac{1}{1+p(z)} &= 1 - p(z) + p(z)^2 - p(z)^3 + p(z)^4 + \cdots \\ &= 1 + \sum_{j=1}^{\infty} s_j z^{-j},\end{aligned}$$

where $s_1, s_2, s_3, \ldots$ are some real numbers. Note that $s_1 = -m_1$. Now, we have that as $z \to \infty$

$$-\frac{o(\frac{1}{z^n})}{(1+p(z))(1+p(z)+o(\frac{1}{z^n}))} = o(\frac{1}{z^n}). \tag{2.13}$$

Thus

$$\frac{1}{1+p(z)+o(\frac{1}{z^n})} = 1 + \sum_{j=1}^{n} s_j z^{-j} + o(\frac{1}{z^n}). \tag{2.14}$$

Finally, we conclude that

$$\begin{aligned}F_\mu(z) &= z(1 + \sum_{j=1}^{n} s_j z^{-j} + o(\frac{1}{z^n})) \\ &= z + s_1 + \frac{s_2}{z} + \frac{s_3}{z^2} + \cdots + \frac{s_n}{z^{n-1}} + o(\frac{1}{z^{n-1}}).\end{aligned}$$

□

Next we present the Nevanlinna representation for F-transforms. This is a very useful formula.

Proposition 10. *Let μ be a positive measure. Suppose that $m_2(\mu) < \infty$. Then there exists a positive measure ν such that $\nu(\mathbb{R}) = var(\mu)$ and*

$$F_\mu(z) = z - \alpha - G_\nu(z) \quad \text{for } z \in \mathbb{C}^+, \tag{2.15}$$

where $\alpha = m_1(\mu)$.

In particular, if $\mu \in \mathcal{M}_0^1$, then there exists $\nu \in \mathcal{M}$ such that $F_\mu(z) = z - G_\nu(z)$ for $z \in \mathbb{C}^+$.

Proposition 11. *Let μ be a probability measure. Suppose that $\int_{\mathbb{R}} t^{2n} d\mu < \infty$. According to the previous proposition, there exists a positive measure ν such that $\nu(\mathbb{R}) = var(\mu)$ and $F_\mu(z) = z - \alpha - G_\nu(z)$. Then we have that $m_{2n-2}(\nu) < \infty$.*

Proof. The hypothesis $\int_{\mathbb{R}} t^{2n} d\mu < \infty$ and the Proposition 6 implies that

$$G_\mu(z) = \sum_{i=0}^{2n} m_i(\mu) \frac{1}{z^{i+1}} + o(\frac{1}{z^{2n+1}}), \quad \text{as } z \to \infty \text{ non-tangentially.}$$

By Proposition 9 we have that

$$F_\mu(z) = z + s_1 + \frac{s_2}{z} + \frac{s_3}{z^2} + \cdots + \frac{s_{2n}}{z^{2n-1}} + o(\frac{1}{z^{2n-1}}), \quad \text{as } z \to \infty,$$

where $s_1 = -m_1$. So we conclude that

$$G_\nu(z) = \frac{-s_2}{z} + \frac{-s_3}{z^2} + \cdots + \frac{-s_{2n}}{z^{2n-1}} + o(\frac{1}{z^{2n-1}}), \quad \text{as } z \to \infty.$$

Thus Proposition 6 implies that $m_{2n-2}(\nu) < \infty$.

□

The following result is Lemma 2.1 in [18]. It describes the relationship of the supports of the measures of the previous formula.

Proposition 12. *Let μ and ν be probability measures on $\mathbb{R}$ such that $F_\mu(z) = z - G_\nu(z)$ for $z \in \mathbb{C}^+$. Then:*

1) $\mathbb{C} \setminus supp(\mu)$ is the maximal domain where $G_\mu(z)$ is analytic and satisfies the integral representation (2.1)*.*

2) $\mathbb{C} \setminus supp(\nu)$ is the maximal domain where $F_\mu(z)$ is analytic and satisfies the integral representation (2.8)*.*

3) $\{x \in \mathbb{R} \setminus supp(\mu) \mid G_\mu(x) \neq 0\} \subset \mathbb{R} \setminus supp(\nu)$, and $\{x \in \mathbb{R} \setminus supp(\nu) \mid F_\mu(x) \neq 0\} \subset \mathbb{R} \setminus supp(\mu)$.

2.1.4 The Kolomogorov and Lévy distances for probability measures

In this section we define the metrics that we use in our main theorems to quantify the rate of convergence on the Boolean and monotone central limit theorems.

Definition 3. *Let μ and ν probability measures over $\mathbb{R}$. Let $\mathcal{F}\mu$ and $\mathcal{F}\nu$ its respective cumulative distribution functions. Then*

- *the **Kolomogorov distance** between μ and ν is defined as*

$$d_{kol}(\mu, \nu) := \sup_{x \in \mathbb{R}} |\mathcal{F}\mu(x) - \mathcal{F}\nu(x)|.$$

- *the **Lévy distance** between μ and ν is defined as*

$$d_{lev}(\mu, \nu) := inf\{\epsilon > 0 \mid \mathcal{F}\mu(x - \epsilon) - \epsilon \leq \mathcal{F}\nu(x) \leq \mathcal{F}\mu(x + \epsilon) + \epsilon \ \ for\ all\ x \in \mathbb{R}\}.$$

.

For two measures μ and ν the Kolmogorov distance $d_{kol}(\mu, \nu)$ is the greatest vertical separation between the graphs of the functions $\mathcal{F}\mu$ and $\mathcal{F}\nu$. Meanwhile, the Lévy distance $d_{lev}(\mu, \nu)$ is the side-length of the largest square that can be inscribed between the graphs, adding vertical segments where there is a discontinuity, of the functions $\mathcal{F}\mu$ and $\mathcal{F}\nu$. This, in particular, implies that

$$d_{lev}(\mu, \nu) \leq d_{kol}(\mu, \nu).$$

The next theorem due to Bai [8] allows us to estimate the Kolmogorov distance of two probability measures using its Cauchy transforms.

Theorem 1 (Bai's inequality)**.** *Let μ and ν probability measures. Suppose that $\int_{\mathbb{R}} |\mathcal{F}_\mu(x) - \mathcal{F}_\nu(x)|dx < \infty$. Then we have that for all $y > 0$*

$$d_{kol}(\mu, \nu) \leq \frac{1}{\pi(2\gamma - 1)}[\int_{\mathbb{R}} |G_\mu(z) - G_\nu(z)|dx + \frac{1}{y} \sup_{r \in \mathbb{R}} \int_{|t| \leq 2ay} |\mathcal{F}_\nu(r + t) - \mathcal{F}_\nu(r)|dt], \qquad (2.16)$$

where $z = x + iy$ and a and γ are constants such that $\gamma = \frac{1}{\pi} \int_{|t|<a} \frac{1}{t^2+1} dt > \frac{1}{2}$.

2.2 Non-Commutative Probability

The starting observation of non-commutative probability theory is that certain elements in some algebras endowed with a functional behave like independent random variables. That is, those elements have a natural distribution defined with respect to the functional, and it makes sense to define the notion of independence, with respect to the functional also, for them. In this new context it is realized that the notion of independence is a broader concept and that other types of independence exist. More interestingly, for some of the new types of independence we can develop a parallel theory with classical probability.

Non-commutative random variables

A non-commutative probability space (A, ϕ) consists of a unital $*$-algebra A, $*$ is an adjoint operation, and a functional $\phi : A \to \mathbb{C}$ such that sends the unity 1_A to 1 and is positive: $\phi(a^*a) \geq 0$. A selfadjoint element $a \in A$ is an element such that $a = a^*$. The positivity assumption on the functional implies that for selfadjoint elements we have that $\phi(a^n) \in \mathbb{R}$ for a positive integer n. We call $\phi(a^n)$ the n-th **moment** of a. It may happen that for some selfadjoint element a there exist a probability measure μ such that $\phi(a^n) = \int_{\mathbb{R}} t^n d\mu(t)$ for $n = 0, 1, 2, 3, ...$ (if we do not ask that the functional to send 1_A to 1, then such a probability measure can not exist). If such distribution is unique (i.e. it is determined by moments) then we call μ the **distribution** of a and call to a a **non-commutative random variable**. By example, $(M_n(\mathbb{C}), Tr)$ any selfadjoint matrix has a unique distribution μ given by $\mu = \frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i}$, where λ_i are the eigenvalues of M and $Tr(M^n) = \int_{\mathbb{R}} t^n d\mu(t)$.

Some important theories that fall within this framework are graph theory, where we consider a graph by its adjacency matrix and the functional may be the normalized trace or the element 1-1 of the matrix, random matrices, and algebras of operators on Hilbert spaces where the functional is a vacuum state.

Independence

If X and Y are two classical random variables whose distributions are determined by moments then independence is characterized by the conditions

$$E(X^nY^n) = E(X^n)E(Y^n).$$

In a similar way, we can characterize any set of independent random variables $\{X_i\}_{i \in I}$ whose distributions are determined by moments.

In a non-commutative probability space (A, ϕ) the functional ϕ takes the role of the expectation. We define the notion of classical independence in terms of the functional as follows. We say that the random variables $\{a_i\}_{i \in I}$ in a non-commutative probability space (A, ϕ) are classical independent with respect to ϕ if they commute and for any product of elements $a_{j_1}a_{j_2}\cdots a_{j_m}$, $j_m \in I$, we have that the reduced product $a_{j_{i_1}}^{r_1}\cdots a_{j_{i_k}}^{r_k}$, where $a_{j_1}a_{j_2}\cdots a_{j_m} = a_{j_{i_1}}^{r_1}\cdots a_{j_{i_k}}^{r_k}$ and $j_{i_l} \neq j_{i_s}$ for $s \neq l$, can be factorized by ϕ as follows

$$\phi(a_{j_{i_1}}^{r_1}\cdots a_{j_{i_k}}^{r_k}) = \phi(a_{j_{i_1}}^{r_1})\cdots\phi(a_{j_{i_k}}^{r_k}).$$

Note that independence is just a rule for factorizing the functional evaluated in products of random variables. In non-commutative probability new rules for making such factorizations are defined. These new factorizations are new notions of independence.

We say that the random variables $\{a_i\}_{i\in I}$ are Boolean independent with respect to ϕ if for any product of elements $a_{j_1}a_{j_2}\cdots a_{j_m}$, $j_m \in I$, we have that the reduced product $a_{j_{i_1}}^{r_1}\cdots a_{j_{i_k}}^{r_k}$, where $a_{j_1}a_{j_2}\cdots a_{j_m} = a_{j_{i_1}}^{r_1}\cdots a_{j_{i_k}}^{r_k}$ and $j_{i_l} \neq j_{i_{l+1}}$, can be factorized by ϕ as follows

$$\phi(a_{j_{i_1}}^{r_1}\cdots a_{j_{i_k}}^{r_k}) = \phi(a_{j_{i_1}}^{r_1})\cdots\phi(a_{j_{i_k}}^{r_k}).$$

We say that the random variables $\{a_i\}_{i\in I}$, where I is an ordered set, say $I \subset \mathbb{N}$, are monotone independent with respect to ϕ if for any product of elements $a_{j_1}a_{j_2}\cdots a_{j_m}$, $j_m \in I$, we have that the reduced product $a_{j_{i_1}}^{r_1}\cdots a_{j_{i_k}}^{r_k}$, where $a_{j_1}a_{j_2}\cdots a_{j_m} = a_{j_{i_1}}^{r_1}\cdots a_{j_{i_k}}^{r_k}$ and $j_{i_l} \neq j_{i_{l+1}}$, can be factorized by ϕ as follows

$$\phi(a_{j_{i_1}}^{r_1}\cdots a_{j_{i_k}}^{r_k}) = \phi(a_{j_{i_s}}^{r_s})\phi(a_{j_{i_1}}^{r_1}\cdots a_{j_{i_{s-1}}}^{r_{s-1}}a_{j_{i_{s+1}}}^{r_{s+1}}\cdots a_{j_{i_k}}^{r_k}),$$

where $a_{j_{i_s}}^{r_s} \geq a_{j_{i_l}}^{r_l}$ for $l = 1, ..., k$.

We present some examples of non-commutative probability spaces with independent random variables in the new senses. Consider the system $(M_{n\times n}(\mathbb{C}), \psi_1)$ where $\psi_1(M) = M_{11}$ for a matrix M. Let $A \in M_{m\times m}(\mathbb{C})$ and $B \in M_{k\times k}(\mathbb{C})$ such that $n = mk$. Denote by I_m the identity matrix of $m \times m$, and by P_m the matrix such that $P_m(1,1) = 1$ and the other entries are zero. Then we have that $(A \otimes I_m)$ and $(I_k \otimes B)$ are tensorial independent; $(A \otimes P_m)$ and $(P_k \otimes B)$ are Boolean independent; and $(A \otimes P_m)$ and $(I_k \otimes B)$ are monotone independent with the order $(A \otimes P_m) < (I_k \otimes B)$.

Broadly speaking, non-commutative probability studies meaningful process

$$f(a_1, a_2, a_3, ...)$$

of Boolean (monotone) independent random variables a_i. Usually, the outcome of the process is some random variable, and we want to compute their moments or distribution. Most of theses processes are motivated by the processes studied in classical probability.

2.3 Sum of Random Variables

The simplest process to study in non-commutative probability is the sum of independent random variables. In this section we discuss how to obtain the moments and distribution of the sum of Boolean (monotone) independent random variables. We also present the main theorems, which are limit theorems, concerning the sum of random variables: the law of large numbers and the central limit theorem.

Moments and Distribution

Following the ideas of classical probability, to deal with the problem of obtaining the moments and distribution of the sum of independent random variables, we restate it in a more convenient way. The Boolean (monotone) independence induces an operation called the Boolean (monotone) convolution $\uplus\,(\triangleright) : \mathcal{M} \times \mathcal{M} \to \mathcal{M}$, that associates to the couple (μ, ν) the distribution $\mu \uplus \nu$ $(\mu \triangleright \nu)$ of the sum x+y of Boolean (monotone) independent random variables x and y with distribution μ and ν respectively. For a detailed definition we refer to [29], [26], and [16]. So, the problem of interest becomes now to find a way of computing the Boolean (monotone) convolution and its moments.

In classical probability, the problem of computing the moments of the convolution of measures (i.e. the moments of the sum of independent random variables) was settled with the discovery of the cumulants. The cumulants of a measure are quantities defined in terms of its moments with the property that the cumulants of the convolution of measures are the sum of the cumulants of the measures convoluted. The n-th cumulant of a measure is defined by a certain function of the first n moments. This formula has an inverse, so the cumulants solve the problem of computing the moments of the convolution. For the Boolean (monotone) convolution there also exist some quantities that are additive respect the Boolean (monotone) convolution. Naturally, these quantities are called the Boolean (monotone) cumulants.

Let μ be a probability measure. Suppose that the n-th moment of μ exists. We define the **Boolean cumulants** $r_k(\mu)$ for $k = 1, 2, ..., n$ by the recurrence formula

$$m_k(\mu) = \sum_{i=1}^{k} r_i(\mu) m_{k-i}(\mu). \tag{2.17}$$

Explicitly the formula for the first four cumulants gives

$$\begin{aligned} m_1(\mu) &= r_1(\mu), \\ m_2(\mu) &= r_1(\mu)^2 + r_2(\mu), \\ m_3(\mu) &= r_1(\mu)^3 + 2r_1(\mu)r_2(\mu) + r_3(\mu), \text{ and} \\ m_4(\mu) &= r_1(\mu)^4 + 3r_1(\mu)^2 r_2(\mu) + r_2(\mu)^2 + 2r_3(\mu)r_1(\mu) + r_4(\mu). \end{aligned}$$

Note that $r_1(\mu) = m_1(\mu)$ and $r_2(\mu) = m_2(\mu) - m_1(\mu)^2 = Var(\mu)$. Also note that if $\mu \in \mathcal{M}_0^1$, then

$$m_4(\mu) = 1 + r_4(\mu). \tag{2.18}$$

The Boolean cumulants behave like the classical cumulants: they are additive for Boolean independent random variables, and the n-th Boolean cumulant is homogeneous of degree n on the random variable. We prove this in Proposition 14 below in a more general context.

On the other hand, the problem of computing the Boolean (monotone) convolution is attacked using a similar idea to the characteristic function in classical probability. A certain transform (the "characteristic function") that sends measures to complex analytic functions is found, mainly using generating functions, such that a simple operation on the transforms of two given measures correspond to their Boolean (monotone) convolution. We next define some transforms that have the mentioned properties.

Theorem 2. *The **monotone convolution** $\mu \triangleright \nu$ is determined by the equation*

$$F_{\mu \triangleright \nu}(z) = F_\mu(F_\nu(z)) \quad \textit{for } z \in \mathbb{C}^+.$$

Definition 4. *The K**-transform** (or self-energy) of a measure $\mu \in \mathcal{M}$ is defined as*

$$K_\mu(z) := z - F_\mu(z) \quad \textit{for}\, z \in \mathbb{C}^+.$$

Theorem 3. *Let μ and ν be probability measures. The **Boolean convolution** $\mu \uplus \nu$ is determined by the equation*

$$K_{\mu \uplus \nu}(z) = K_\mu(z) + K_\nu(z) \quad \textit{for}\, z \in \mathbb{C}^+.$$

To end with this section, we study some properties of the K-transform that are relevant for Chapter 3. Directly from the definitions above we obtain that

$$F_{\mu^{\uplus n}}(z) = (1-n)z + nF_\mu(z), \quad z \in \mathbb{C}^+. \tag{2.19}$$

The next proposition gives an asymptotic expansion of the K-transform in terms of the Boolean cumulants.

Proposition 13. *Let μ be a probability measure. Suppose that $m_n(\mu) < \infty$. Then K_μ admits the expansion*

$$K_\mu(z) = r_1 + \frac{r_2}{z} + \cdots + \frac{r_n}{z^{n-1}} + o(\frac{1}{z^{n-1}}), \quad \textit{as } z \to \infty \textit{ non-tangentially}$$

where $r_i = r_i(\mu)$ is the i-th Boolean cumulant of μ for $i = 1, 2, ..., n$. Conversely, if K_μ admits the previous expansion for even n, then $m_n(\mu) < \infty$.

Proof. Suppose that $m_n(\mu) < \infty$. Proposition 6 implies that

$$G_\mu(z) = \sum_{i=0}^{n} m_i(\mu)\frac{1}{z^{i+1}} + o(\frac{1}{z^{n+1}}), \quad \text{as } z \to \infty \text{ non-tangentially},$$

By Proposition 9 we obtain that

$$F_\mu(z) = z + s_1 + \frac{s_2}{z} + \frac{s_3}{z^2} + \cdots + \frac{s_n}{z^{n-1}} + o(\frac{1}{z^{n-1}}), \quad \text{as } z \to \infty \text{ non-tangentially},$$

for some real numbers $s_1, s_2, ..., s_n$. Since $K_\mu(z) = z - F_\mu(z)$ we obtain that

$$K_\mu(z) = r_1 + \frac{r_2}{z} + \frac{r_3}{z^2} + \cdots + \frac{r_{2n}}{z^{n-1}} + o(\frac{1}{z^{n-1}}), \quad \text{as } z \to \infty \text{ non-tangentially},$$

where $r_i = -s_i$ for $i = 1, ..., n$. Now, we want to prove that the numbers r_i are the Boolean cumulants of μ. That is, we want to prove that

$$m_k(\mu) = \sum_{i=1}^{k} r_i(\mu) m_{k-i}(\mu),$$

for $k = 1, 2, ..., n$. Note that

$$G_\mu(z)K_\mu(z) = zG_\mu(z) - 1 \tag{2.20}$$

The second side of this equation can be expressed as

$$\begin{aligned} zG_\mu(z) - 1 &= z(\sum_{i=0}^{n} m_i(\mu)\frac{1}{z^{i+1}} + o(\frac{1}{z^{n+1}})) - 1 \\ &= \sum_{i=1}^{n} m_i(\mu)\frac{1}{z^i} + o(\frac{1}{z^n}). \end{aligned}$$

Expanding the first side of the Equation 2.20, we obtain

$$\begin{aligned} G_\mu(z) \cdot K_\mu(z) &= (\sum_{i=0}^{n} m_i(\mu)\frac{1}{z^{i+1}} + o(\frac{1}{z^{n+1}}))(\sum_{k=1}^{n} \frac{r_k}{z^{k-1}} + o(\frac{1}{z^{n-1}})) \\ &= (\sum_{i=0}^{n} m_i(\mu)\frac{1}{z^{i+1}})(\sum_{k=1}^{n} \frac{r_k}{z^{k-1}}) + (\sum_{i=0}^{n} m_i(\mu)\frac{1}{z^{i+1}})(o(\frac{1}{z^{n-1}})) \\ &+ (\sum_{k=1}^{n} \frac{r_k}{z^{k-1}})(o(\frac{1}{z^{n+1}})) + (o(\frac{1}{z^{n+1}}))(o(\frac{1}{z^{n-1}})) \\ &= (\sum_{i=0}^{n} m_i(\mu)\frac{1}{z^{i+1}})(\sum_{k=1}^{n} \frac{r_k}{z^{k-1}}) + o(\frac{1}{z^n}). \end{aligned}$$

Putting $m_{-1} = 0$, $m_i = m_i(\mu)$ for $i = 0, 1, ..., n$, $m_i = 0$ for $i = n+1, ..., 2n$, and $r_i = 0$ for $i = n+1, ..., 2n$, we obtain that

$$(\sum_{i=0}^{n} m_i(\mu)\frac{1}{z^{i+1}})(\sum_{k=1}^{n}\frac{r_k}{z^{k-1}}) = (\sum_{i=0}^{n+1} m_{i-1}\frac{1}{z^i})(\sum_{k=0}^{n-1}\frac{r_{k+1}}{z^k})$$
$$= \sum_{k=0}^{2n}(\sum_{i=0}^{k} r_{i+1}m_{k-i-1})\frac{1}{z^i}$$

Therefore, we arrive to

$$G_\mu(z) \cdot K_\mu(z) = \sum_{k=0}^{n}(\sum_{i=0}^{k} r_{i+1}m_{k-i-1})\frac{1}{z^i} + o(\frac{1}{z^n}). \tag{2.21}$$

The Equation 2.20 and the previous calculations imply that

$$\sum_{k=1}^{n} m_k\frac{1}{z^k} + o(\frac{1}{z^n}) = \sum_{k=0}^{n}(\sum_{i=0}^{k} r_{i+1}m_{k-i-1})\frac{1}{z^i} + o(\frac{1}{z^n}).$$

We conclude that for $k = 1, ..., n$

$$m_k = \sum_{i=0}^{k} r_{i+1}m_{k-i-1}$$
$$= \sum_{i=0}^{k-1} r_{i+1}m_{k-i-1}$$
$$= \sum_{i=1}^{k} r_i m_k$$

Therefore $r_1, ..., r_n$ are the first n Boolean cumulants of μ.

On the other hand. Suppose that for an even integer n K_μ admits the expansion

$$K_\mu(z) = r_1 + \frac{r_2}{z} + \cdots + \frac{r_n}{z^{n-1}} + o(\frac{1}{z^{n-1}}), \quad \text{as } z \to \infty \text{ non-tangentially}$$

It follows that

$$F_\mu(z) = z - r_1 - \frac{r_2}{z} + \cdots - \frac{r_n}{z^{n-1}} + o(\frac{1}{z^{n+1}}), \quad \text{as } z \to \infty \text{ non-tangentially.}$$

Then by Propositions 9 and 6 we conclude that $m_n(\mu) < \infty$. □

Finally, the next proposition shows that the Boolean cumulants defined previously have the right properties.

Proposition 14. *Let μ and ν be probability measures. Then we have the following:*

- *If $m_n(\mu) < \infty$ and $m_n(\nu) < \infty$ for even n, then we have that $m_n(\mu \uplus \nu) < \infty$ and*

$$r_i(\mu \uplus \nu) = r_i(\mu) + r_i(\nu), \tag{2.22}$$

for $i = 1, 2, ..., n$.

- *If* $m_k(\mu) < \infty$, *then*

$$r_k(D_a\mu) = a^k r_k(\mu). \tag{2.23}$$

Proof. Suppose that for even n we have $m_n(\mu) < \infty$. Then Proposition 13 implies that

$$K_\mu(z) = r_1(\mu) + \frac{r_2(\mu)}{z} + \cdots + \frac{r_n(\mu)}{z^{n-1}} + o(\frac{1}{z^{n-1}}), \quad \text{as } z \to \infty \text{ non-tangentially,}$$

and

$$K_\nu(z) = r_1(\nu) + \frac{r_2(\nu)}{z} + \cdots + \frac{r_n(\nu)}{z^{n-1}} + o(\frac{1}{z^{n-1}}), \quad \text{as } z \to \infty \text{ non-tangentially.}$$

Thus we obtain that

$$K_{\mu \uplus \nu}(z) = K_\mu(z) + K_\nu(z) = r_1(\nu) + r_1(\nu) + \frac{r_2(\mu) + r_2(\nu)}{z} + \cdots + \frac{r_n(\mu) + r_n(\nu)}{z^{n-1}} + o(\frac{1}{z^{n-1}}),$$

as $z \to \infty$ non-tangentially. Since n is even, Proposition 13 implies that $m_n(\mu \uplus \nu) < \infty$ and that the Boolean cumulants of $\mu \uplus \nu$ are given by $r_i(\mu \uplus \nu) = r_i(\mu) + r_i(\nu)$ for $i = 1, 2, ..., n$. Now, we prove Equation 2.23. By definition

$$m_k(\mu) = \sum_{i=1}^{k} r_i(\mu) m_{k-i}(\mu) = \sum_{i=1}^{k-1} r_i(\mu) m_{k-i}(\mu) + r_k,$$

because $m_0(\mu) = 1$. So, we have that

$$r_k(\mu) = m_k(\mu) - \sum_{i=1}^{k} r_i(\mu) m_{k-i}(\mu). \tag{2.24}$$

Recall that the n-th moment of a measure μ satisfies $m_n(D_a\mu) = a^n m_n(\mu)$. We proceed by induction. For $n = 1$ we have that $r_1(D_a\mu) = m_1(D_a\mu) = am_1(\mu) = ar_1(\mu)$. Now, suppose that $r_n(D_a\mu) = a^n r_n(\mu)$ for $n = 1, 2, ..., k-1$. By Equation 2.24 we have that

$$\begin{aligned} r_k(D_a\mu) &= m_k(D_a\mu) - \sum_{i=1}^{k} r_i(D_a\mu) m_{k-i}(D_a\mu) \\ &= a^k m_k(\mu) - \sum_{i=1}^{k} a^i r_i(\mu) a^{k-i} m_{k-i}(\mu) \\ &= a^k (m_k(\mu) - \sum_{i=1}^{k} r_i(\mu) m_{k-i}(\mu)). \end{aligned}$$

Therefore $r_k(D_a\mu) = a^k r_k(\mu)$. ◻

Chapter 3

A Boolean Berry-Esseen Type Theorem

In this Chapter we study the rate of convergence in the Boolean central limit theorem with respect to the Lévy distance. In the next section we establish our results. In Sections 3.2 and 3.3 we give the proofs of our theorems.

3.1 Main Results

Let μ be a probability measure of zero mean and unit variance. Define $\mu_n := D_{\frac{1}{\sqrt{n}}}\mu^{\uplus n}$ where $\uplus$ denotes the Boolean convolution. Our objective is to obtain a Berry-Essen type estimate for the Boolean central limit theorem. Now, the Berry-Essen theorem and the Berry-Esseen type theorem for the Free independence are in terms of the Kolmogorov distance. But in the Boolean central limit theorem there is not convergence in the Kolmogorov distance, as one can easily see from almost any example (see Example 1 for a particular one). Here we consider the Lévy distance instead, which seems the most appropriate. Thus, in other words, our objective is to estimate the Lévy distance $d_{lev}(\mu_n, \mathbf{b})$, where $\mathbf{b}$ is the Bernoulli distribution $\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$.

The first problem we face is to find a way of estimating the Lévy distance other than the definition. Using just the definition to estimate $d_{lev}(\mu_n, \mathbf{b})$ seems a very difficult approach. Our first theorem gives us a way to estimate the more particular situation of $d_{lev}(\nu, \mathbf{b})$ for some measure $\nu \in \mathcal{M}_0^1$, and we use it to estimate $d_{lev}(\mu_n, \mathbf{b})$. The theorem says that the distance $d_{lev}(\nu, \mathbf{b})$ depends uniquely on the fourth moment. This theorem is a version of a *fourth moment theorem* with respect to the Lévy distance. In [4] Arizmendi and Gaxiola prove a similar result, but with respect to a different distance given in terms of the Cauchy transform.

Theorem 4. *Let μ be a probability measure such that $m_1(\mu) = 0$, $m_2(\mu) = 1$, and $m_4(\mu) < \infty$. Then*

$$d_{lev}(\mu, \mathbf{b}) \leq \frac{7}{2}\sqrt[3]{m_4(\mu) - 1}. \tag{3.1}$$

Note that the radical $m_4(\mu) - 1$ is positive for a measure with $\mu_2(\mu) = 1$ (since $E(X^4) \geq E(X^2)^2$) . This theorem says in particular that if $m_4(\mu) = 1$ for a measure $\mu \in \mathcal{M}_0^1$, then μ is the Bernoulli distribution.

Now, this theorem gives an useful way to estimate $d_{lev}(\mu_n, \mathbf{b})$. That is because we can easily obtain the fourth moment of μ_n through the Boolean cumulants. We obtain the following theorem that gives a bound of the rate of convergence in the Boolean CLT for measures with finite fourth moment.

Theorem 5. *Let μ be a probability measure such that $m_1(\mu) = 0$, $m_2(\mu) = 1$, and $m_4(\mu) < \infty$. Then for the measure $\mu_n = D_{\frac{1}{\sqrt{n}}}\mu^{\uplus n}$ we have that*

$$d_{lev}(\mu_n, \mathbf{b}) \leq \frac{7}{2}\sqrt[3]{\frac{m_4(\mu) - 1}{n}} \quad \text{for } n \geq 1.$$

We think that the estimate in the above theorem is not sharp. Our guess is that the sharp rate must be of order $\frac{1}{\sqrt{n}}$, as it is in the case of our next theorem.

Our next theorem is not the result of finding a way of estimating the Lévy distance $d_{lev}(\mu_n, \mathbf{b})$. It is the direct consequence of the surprising fact that the measure μ_n when μ has bounded support becomes quite simple for big n. It consists of one atom about -1, one atom about +1, and a negligible mass around 0. It is possible to estimate such atoms using the F-transform of μ_n, and with that estimations directly obtain an estimation of $d_{lev}(\mu_n, \mathbf{b})$. The theorem is the following.

Theorem 6. *Let μ be a probability measure such that $m_1(\mu) = 0$, $m_2(\mu) = 1$, and $supp(\mu) \subset [-K, K]$. Then the measure $\mu_n := D_{\frac{1}{\sqrt{n}}}\mu^{\uplus n}$ satisfies for $\sqrt{n} > K$ that:*

1) *$supp\, \mu_n \subset [\frac{-K}{\sqrt{n}}, \frac{K}{\sqrt{n}}] \cup \{x_1, x_2\}$, where $|(-1) - x_1| \leq \frac{K}{\sqrt{n}}$ and $|1 - x_2| \leq \frac{K}{\sqrt{n}}$.*

2) *For $p = \mu_n(\{x_1\})$, $q = \mu_n(\{x_2\})$ and $r = \mu_n([\frac{-K}{\sqrt{n}}, \frac{K}{\sqrt{n}}])$, we have that $p, q \in [\frac{1}{2} - \frac{2K}{\sqrt{n}}, \frac{1}{2} + \frac{K}{2\sqrt{n}}]$ and $r < \frac{4K}{\sqrt{n}}$.*

In particular, the Lévy distance between μ_n and $\mathbf{b}$ is bounded by

$$d_{lev}(\mu_n, \mathbf{b}) \leq \frac{2K}{\sqrt{n}}.$$

Note that the theorem says that as soon as $\sqrt{n} > K$ we can guarantee the limit measure μ_n has two atoms, one close to -1 and the other to 1, with a mass about $\frac{1}{2}$, and the rest of the measure is concentrated around 0. This gives us a very explicit description of the convergence in the Boolean central limit theorem.

If the support is not bounded, then we can not assure the appearance of the two atoms. By instance, for the standard Normal distribution $\mathbf{g}$ one can see that, Proposition 7.4 in [19], the distribution $\mathbf{g}^{\uplus n}$ is given by the density

$$\frac{n}{|(1-n)xG_{\mathbf{g}}(x) + n|^2} \frac{1}{2\pi} \exp -x^2/2,$$

so no atoms show up in $D_{\frac{1}{\sqrt{n}}}\mathbf{g}^{\uplus n}$.

The main results in this chapter have been published in collaboration with Octavio Arizmendi [6].

3.2 Unbounded Case

The objective of this section is to prove Theorems 4 and 5. Theorem 5 essentially establishes a Berry-Esseen type theorem for the Boolean central limit. We first give two technical lemmas. Then we prove Theorem 4. After that, We prove in a proposition that the bound of this theorem is sharp. At last, we prove Theorem 5.

The next lemma says that for a measure ν of zero mean and unit variance the Lévy distance $d_{lev}(\nu, \mathbf{n})$ is small if ν concentrates evenly around the points +1 and -1.

Lemma 1. *If $\mu \in \mathcal{M}_0^1$ and $\mu((-1-\epsilon,-1+\epsilon)\cup(1-\epsilon,1+\epsilon)) \geq 1-\epsilon$ for some $\epsilon \in (0,1)$, then*

$$d_{lev}(\mu,\mathbf{b}) \leq \frac{7}{2}\epsilon. \tag{3.2}$$

Proof. Define $R_1 := (-\infty,-1-\epsilon]$, $R_2 := (-1-\epsilon,-1+\epsilon)$, $R_3 := [-1+\epsilon,1-\epsilon]$, $R_4 := (1-\epsilon,1+\epsilon)$, and $R_5 := [1+\epsilon,\infty)$. Let $p_i := \mu(R_i)$ for $i = 1,2,3,4,5$. Clearly, there exists $t_i \in R_i$ such that

$$\int_{R_i} t d\mu(t) = t_i p_i.$$

Note that $p_2 + p_4 \geq 1-\epsilon$ by hypothesis. So we have

$$p_i \leq \epsilon \quad \text{for } i = 1,3,5. \tag{3.3}$$

Observe that

$$\begin{aligned}
|t_1p_1| + |t_5p_5| &\leq \int_{R_1} t^2 d\mu(t) + \int_{R_5} t^2 d\mu(t) \\
&= 1 - \int_{R_2} t^2 d\mu(t) - \int_{R_3} t^2 d\mu(t) - \int_{R_4} t^2 d\mu(t) \\
&\leq 1 - (1-\epsilon)^2(p_2+p_4) \\
&\leq 1 - (1-\epsilon)^2(1-\epsilon) \\
&\leq 3\epsilon,
\end{aligned}$$

where the first equality is because of $m_2(\mu) = 1$. It is clear that $|t_3p_3| \leq \epsilon$. Thus, we obtain from $m_1(\mu) = 0$ that

$$|t_2p_2 + t_4p_4| = |t_1p_1 + t_3p_3 + t_5p_5| \leq 4\epsilon.$$

Also note that

$$\begin{aligned}
|p_2 - p_4| - |t_2p_2 + t_4p_4| &\leq |t_2p_2 + t_4p_4 + p_2 - p_4| \\
&\leq |t_2+1|p_2 + |t_4-1|p_4 \\
&\leq \epsilon(p_2+p_4) \\
&\leq \epsilon.
\end{aligned}$$

It follows that $|p_2 - p_4| \leq 5\epsilon$, and since $1-\epsilon \leq p_2+p_4 \leq 1$, then

$$\frac{1}{2} - 3\epsilon \leq p_2, p_4 \leq \frac{1}{2} + \frac{5}{2}\epsilon. \tag{3.4}$$

Using the estimates (3.3) and (3.4), it is easy to see that

$$d_{lev}(\mu,\mathbf{b}) \leq \frac{7}{2}\epsilon.$$

□

Let μ be a probability measure, and let X be a random variable with distribution μ. By μ^2 we denote the distribution of X^2. Note that for $\epsilon > 0$ we have

$$\mu((-1-\epsilon,-1+\epsilon)\cup(1-\epsilon,1+\epsilon)) \geq \mu^2((1-\epsilon,1+\epsilon)). \tag{3.5}$$

Lemma 2. *If $\mu \in \mathcal{M}_0^1$ and $m_4(\mu) < \infty$, then:*

(i) $Var(\mu^2) = r_4(\mu)$.

(ii) $\mu^2((1-\epsilon, 1+\epsilon)) \geq 1 - \frac{r_4(\mu)}{\epsilon^2}$.

Proof. For part *(i)* observe that $Var(\mu^2) = m_2(\mu^2) - m_1(\mu^2)^2 = m_4(\mu) - 1$. Hence, by (2.18) we obtain that $Var(\mu^2) = r_4(\mu)$.

For part *(ii)* we see that by the Chebyshev inequality we have

$$P(|X^2 - E(X^2)| < \epsilon) \geq 1 - \frac{Var(X^2)}{\epsilon^2},$$

and using *(i)* we conclude that $\mu^2((1-\epsilon, 1+\epsilon)) \geq 1 - \frac{r_4(\mu)}{\epsilon^2}$. □

We use the previous lemmas to prove Theorem 4

Proof of Theorem 4. By Lemma 2 and inequality (3.5), we see that

$$\mu((-1-\epsilon, -1+\epsilon) \cup (1-\epsilon, 1+\epsilon)) \geq \mu^2((1-\epsilon, 1+\epsilon)) \geq 1 - \frac{r_4(\mu)}{\epsilon^2}.$$

Taking $\epsilon = \sqrt[3]{r_4(\mu)}$, we obtain

$$\mu((-1-\sqrt[3]{r_4(\mu)}, -1+\sqrt[3]{r_4(\mu)}) \cup (1-\sqrt[3]{r_4(\mu)}, 1+\sqrt[3]{r_4(\mu)})) \geq 1 - \sqrt[3]{r_4(\mu)}.$$

By Lemma 1 we conclude that when $r_4(\mu) < 1$, then

$$d_{lev}(\mu, \mathbf{b}) \leq \frac{7}{2}\sqrt[3]{r_4(\mu)} = \frac{7}{2}\sqrt[3]{m_4(\mu) - 1}.$$

□

The following proposition shows that the bound in Theorem 4 is sharp.

Proposition 15. *For all $\alpha > \frac{1}{3}$ and for all $C > 0$ there exists $\mu \in \mathcal{M}_0^1$ such that*

$$d_{lev}(\mu, \mathbf{b}) > C \cdot r_4(\mu)^\alpha$$

Proof. Fix $\alpha > \frac{1}{3}$ and $C > 0$. Let $\epsilon \in (0, 1/2)$. Define

$$\mu_\epsilon = \frac{\epsilon}{2}\delta_{-\sqrt{1+\epsilon}} + (\frac{1}{2} - \epsilon)\delta_{-1} + \frac{\epsilon}{2}\delta_{-\sqrt{1-\epsilon}} + \frac{\epsilon}{2}\delta_{\sqrt{1-\epsilon}} + (\frac{1}{2} - \epsilon)\delta_1 + \frac{\epsilon}{2}\delta_{\sqrt{1+\epsilon}}$$

Clearly $\mu_\epsilon \in \mathcal{M}_0^1$. We also have that

$$\begin{aligned} m_4(\mu_\epsilon) &= \epsilon(1+\epsilon)^2 + \epsilon(1-\epsilon)^2 + (1-2\epsilon) \\ &= 1 + 2\epsilon^3. \end{aligned}$$

So by (2.18) we obtain that $r_4(\mu) = 2\epsilon^3$.

On the other hand, since μ_ϵ is atomic, then, for $\epsilon < 1$, one has

$$d_{lev}(\mu_\epsilon, \mathbf{b}) \geq min\{\frac{\epsilon}{2}, 1 - \sqrt{1-\epsilon}, \sqrt{1+\epsilon} - 1\} \geq \frac{\epsilon}{4}.$$

It is not hard to see that for ϵ small enough we have that $\frac{\epsilon}{4} > C(2^\alpha \epsilon^{3\alpha})$. For such ϵ

$$d_{lev}(\mu_\epsilon, \mathbf{b}) > C \cdot r_4(\mu)^\alpha.$$

□

Now we are able to prove Theorem 5.

Proof of Theorem 5. It follows from (2.22) and (2.23) that

$$r_4(\mu_n) = r_4(D_{1/\sqrt{n}}\mu^{\uplus n}) = \frac{1}{n^2} n r_4(\mu) = \frac{r_4(\mu)}{n}.$$

By Theorem 4 we conclude that

$$d_{lev}(\mu_n, \mathbf{b}) \leq \frac{7}{2}\sqrt[3]{\frac{r_4(\mu)}{n}} = \frac{7}{2}\sqrt[3]{\frac{m_4(\mu)-1}{n}}.$$

□

3.3 Bounded case

In this section we prove Theorem 6. This theorem establishes a Berry-Esseen type theorem for the Boolean central limit theorem for measures of bounded support. We also give an example that shows that the estimate in this theorem is sharp.

Proof of Theorem 2. Using (2.9) and (2.19), we obtain

$$F_{\mu_n}(z) = (1-n)z + \sqrt{n}F_\mu(\sqrt{n}z) \quad \text{for } z \in \mathbb{C}^+.$$

By Proposition 10 there exists a measure $\nu \in \mathcal{M}$ such that $F_\mu(z) = z - G_\nu(z)$ for $z \in \mathbb{C}^+$. It follows that

$$\begin{aligned} F_{\mu_n}(z) &= (1-n)z + \sqrt{n}(\sqrt{n}z - G_\nu(\sqrt{n}z)) \\ &= z - \sqrt{n}G_\nu(\sqrt{n}z)) \\ &= z - G_{D_{\frac{1}{\sqrt{n}}}\nu}(z) \qquad for\ z \in \mathbb{C}^+, \end{aligned}$$

where the third equality is due to (2.3).

Note that *supp* $(\nu) \subset [-K, K]$. Indeed, suppose that $x \in \mathbb{R} \setminus [-K, K]$. Since *supp* $(\mu) \subset [-K, K]$, then $x \in \mathbb{R} \setminus$ *supp* (μ). Therefore, by (2.5) we obtain that $G_\mu(x) \neq 0$. Finally, part (3) of Proposition 12 implies that $x \in \mathbb{R} \setminus$ *supp* (ν).

Let us write $\hat{\nu} = D_{\frac{1}{\sqrt{n}}}\nu$. So we have that *supp*$(\hat{\nu}) \subset [\frac{-K}{\sqrt{n}}, \frac{K}{\sqrt{n}}]$ and

$$F_{\mu_n}(z) = z - G_{\hat{\nu}}(z) \quad \text{for } z \in \mathbb{C}^+. \tag{3.6}$$

By the third part of Proposition 12, we conclude that *supp*$(\mu_n) \subset [\frac{-K}{\sqrt{n}}, \frac{K}{\sqrt{n}}] \cup \{x \in \mathbb{R} \setminus [\frac{-K}{\sqrt{n}}, \frac{K}{\sqrt{n}}] \mid F_{\mu_n}(x) = 0\}$. To conclude the proof of part 1), it is left to prove that there are only two zeros x_1 and x_2 which satisfy the conditions $|(-1) - x_1| \leq \frac{K}{\sqrt{n}}$ and $|1 - x_2| \leq \frac{K}{\sqrt{n}}$. The second part of Proposition 12 implies for $z \in \mathbb{C} \setminus [\frac{-K}{\sqrt{n}}, \frac{K}{\sqrt{n}}]$ that $F_{\mu_n}(z) = z - G_{\hat{\nu}}(z)$ and that $F_{\mu_n}(z)$ is analytic. Therefore, we obtain from the definition of Cauchy transform that

$$F'_{\mu_n}(x) = 1 + \int_{-\infty}^{\infty} \frac{1}{(t-x)^2} d\hat{\nu}(t) \quad \text{for } x \in \mathbb{R} \setminus [\frac{-K}{\sqrt{n}}, \frac{K}{\sqrt{n}}].$$

In particular, $F_{\mu_n}(x)$ is increasing in $(\frac{K}{\sqrt{n}}, \infty)$ and can have at most one zero there. It is clear that

$$F_{\mu_n}(x) > x - \frac{1}{x - K/\sqrt{n}} > K/\sqrt{n} \quad \text{for } x > 1 + K/\sqrt{n}$$

and

$$F_{\mu_n}(x) < x - \frac{1}{x + K/\sqrt{n}} < -K/\sqrt{n} \quad \text{for } K/\sqrt{n} < x < 1 - K/\sqrt{n}.$$

Since $F_{\mu_n}(x)$ is continuous in $(\frac{K}{\sqrt{n}}, \infty)$, it must have a zero x_2 in $[1 - \frac{K}{\sqrt{n}}, 1 + \frac{K}{\sqrt{n}}]$.

A similar argument shows that $F_{\mu_n}(x)$ has only a zero x_1 in $(-\infty, \frac{-K}{\sqrt{n}})$ bounded in $[-1 - \frac{K}{\sqrt{n}}, -1 + \frac{K}{\sqrt{n}}]$. We conclude the proof of part 1).

Using (3.6) and Proposition 10, we obtain that $m_1(\mu_n) = 0$ and $m_2(\mu_n) = 1$. The idea of the rest of the proof is that these two moments force the mass of μ_n to concentrate evenly in x_1 and x_2.

We put $p = \mu_n(\{x_1\})$, $q = \mu_n(\{x_2\})$, and $r = \mu_n([\frac{-K}{\sqrt{n}}, \frac{K}{\sqrt{n}}])$. Note that $p+q+r = 1$ by part 1). It is clear that there exist $y_1 \in [\frac{-K}{\sqrt{n}}, \frac{K}{\sqrt{n}}]$ and $y_2 \in [0, \frac{K}{\sqrt{n}}]$ such that $\int_{-K/\sqrt{n}}^{K/\sqrt{n}} x \, d\mu_n(x) = y_1 r$ and $\int_{-K/\sqrt{n}}^{K/\sqrt{n}} x^2 \, d\mu_n(x) = y_2^2 r$. Define $\epsilon := 1 + x_1$ and $\delta := x_2 - 1$. Since $m_1(\mu_n) = 0$, then we have that $x_1 p + y_1 r + x_2 q = 0$; it follows that $p(-1+\epsilon) + r y_1 + q(1+\delta) = 0$. Therefore, we deduce the inequalities

$$|q - p| \le p|\epsilon| + r|y_1| + q|\delta| \le (p+q+r)\frac{K}{\sqrt{n}} = \frac{K}{\sqrt{n}}. \tag{3.7}$$

Since $p + q \le 1$, then $p + p - \frac{K}{\sqrt{n}} \le 1$. Therefore we obtain that $p \le \frac{1}{2} + \frac{K}{2\sqrt{n}}$. Similarly, we can conclude that $q \le \frac{1}{2} + \frac{K}{2\sqrt{n}}$. Now, we have from $m_2(\mu_n) = 1$ that $x_1^2 p + y_2^2 r + x_2^2 q = 1$. It follows that $(1 + 2\epsilon + \epsilon^2)p + y_2^2 r + (1 + 2\delta + \delta^2)q = 1$, and we get the estimate

$$p + q = 1 - (\epsilon^2 p + y_2^2 r + \delta^2 q) - 2(\epsilon p + \delta q) \ge 1 - \frac{K^2}{n} - 2\frac{K}{\sqrt{n}} \ge 1 - 3\frac{K}{\sqrt{n}}.$$

Since $q \le p + \frac{K}{\sqrt{n}}$, then $2p + \frac{K}{\sqrt{n}} \ge 1 - 3\frac{K}{\sqrt{n}}$. It follows that $p \ge \frac{1}{2} - \frac{2K}{\sqrt{n}}$ and $q \ge \frac{1}{2} - \frac{2K}{\sqrt{n}}$. Finally, we conclude that $r \le 4\frac{K}{\sqrt{n}}$ because of $p + q + r = 1$.

It follows from the estimates obtained for p, q, and r that

$$d_{lev}(\mu_n, \mathbf{b}) \le \frac{2K}{\sqrt{n}}.$$

□

In the next example we compute the

Example 1. *Let n be a positive integer. Define $p_n := \frac{1}{2}\frac{\sqrt{1+4n}+1}{\sqrt{1+4n}}$, $q_n := \frac{1}{2}\frac{\sqrt{1+4n}-1}{\sqrt{1+4n}}$, $x_n := \frac{1-\sqrt{1+4n}}{\sqrt{4n}}$, and $y_n := \frac{1+\sqrt{1+4n}}{\sqrt{4n}}$. Let μ_n be the probability measure given by $\mu_n := p_n\delta_{x_n} + q_n\delta_{y_n}$. Then $\mu_n = D_{\frac{1}{\sqrt{n}}}\mu_1^{\uplus n}$, where $\mu := \mu_1$, and $d_{lev}(\mu_n, \mathbf{b}) \ge \frac{1}{6\sqrt{n}}$.*

To prove $\mu_n = D_{\frac{1}{\sqrt{n}}}\mu^{\uplus n}$, it is sufficient to show that $F_{\mu_n}(z) = F_{D_{\frac{1}{\sqrt{n}}}\mu^{\uplus n}}(z)$ for $z \in \mathbb{C}^+$. First we compute the Cauchy transform of μ_n:

$$\begin{aligned} G_{\mu_n}(z) &= \frac{p_n}{z - x_n} + \frac{q_n}{z - y_n} = \frac{p_n(z - y_n) + q_n(z - x_n)}{(z - x_n)(z - y_n)} \\ &= \frac{p_n(z - y_n) + q_n(z - x_n)}{(z - x_n)(z - y_n)} = \frac{z - y_n p_n - x_n q_n}{(z - x_n)(z - y_n)} \\ &= \frac{z - x_n - y_n}{z^2 - (x_n + y_n)z + x_n y_n} = \frac{z - 1/\sqrt{n}}{z^2 - z/\sqrt{n} - 1}. \end{aligned}$$

Hence, we have that $F_{\mu_n}(z) = \frac{\sqrt{n}z^2 - z - \sqrt{n}}{\sqrt{n}z - 1}$. *In particular* $F_\mu(z) = \frac{z^2 - z - 1}{z - 1}$.

On the other hand, we compute

$$\begin{aligned} F_{D_{\frac{1}{\sqrt{n}}}\mu^{\uplus n}}(z) &= \frac{1}{\sqrt{n}}((1-n)\sqrt{n}z + nF_\mu(\sqrt{n}z) = (1-n)z + \sqrt{n}F_\mu(\sqrt{n}z) \\ &= (1-n)z + \sqrt{n}(\frac{nz^2 - \sqrt{n}z - 1}{\sqrt{n}z - 1}) = \frac{(z - nz)(\sqrt{n}z - 1) + \sqrt{n}nz^2 - nz - \sqrt{n}}{\sqrt{n}z - 1} \\ &= \frac{\sqrt{n}z^2 - z - \sqrt{n}}{\sqrt{n}z - 1} = F_{\mu_n}(z). \end{aligned}$$

Now it is easy to see, that $d_{lev}(\mu_n, \mathbf{b}) = max\{|-1-x_n|, |1-y|, |1/2-p_n|\}$, *and since* $|1/2-p_n| \geq \frac{1}{6\sqrt{n}}$, *then* $d_{lev}(\mu_n, \mathbf{b}) \geq \frac{1}{6\sqrt{n}}$.

Chapter 4

A Monotone Berry-Essen Type Theorem

In this section we study the speed of convergence in the monotone central limit theorem with respect to the Kolmogorov distance. We obtain two theorems. In the next section we describe them. In Section 4.2 we prove the theorems. Since the proof requires many estimations, we opt for proving the estimations in the next two sections. In Section 4.3 we obtain the estimations concerning the monotone convolution, and in Section 4.4 we give the calculations of the integrals.

4.1 Main Results

Let μ be a probability measure of zero mean and unit variance. Let n be a positive integer, and define $\mu_n := D_{\frac{1}{\sqrt{n}}}\mu^{\triangleright n}$. We denote by $\mathbf{a}$ to the arcsine distribution:

$$d\mathbf{a}(t) := \begin{cases} \frac{1}{\pi\sqrt{2-t^2}} & t \in (-\sqrt{2}, \sqrt{2}), \\ 0 & \text{elsewhere.} \end{cases} \tag{4.1}$$

Our goal is to study the convergence on the monotone central limit theorem with respect to the Kolmogorov distance. In other words, we want to estimate the Kolmogorov distance $d_{kol}(\mu_n, \mathbf{a}) = \sup_{x\in\mathbb{R}} |\mathcal{F}_{\mu_n}(x) - \mathcal{F}_{\mathbf{a}}(x)|$ as n goes to infinity.

Our first result gives an estimation of the rate of convergence in the monotone central limit theorem for measures of finite fourth moment.

Theorem 7. *Let μ be a probability measure with $m_1(\mu) = 0$ and $m_2(\mu) = 1$. Let $\mu_n = D_{1/\sqrt{n}}\mu^{\triangleright n}$, $n \geq 1$. Suppose that $m_4(\mu) < \infty$. Then there exists a constant k depending only on μ such that we have the Kolmogorov distance*

$$d_{kol}(\mu_n, \mathbf{a}) \leq kn^{-1/8}$$

The following theorem improves the bound of the previous one, but it slightly less general being only for measures of finite sixth moment.

Theorem 8. *Let μ be a probability measure. Suppose that $m_1(\mu) = 0$, $m_2(\mu) = 1$, and $m_6(\mu) < \infty$. Let $\mu_n = D_{1/\sqrt{n}}\mu^{\triangleright n}$, $n \geq 1$. Then there exists a constant k depending only on μ such that we have the Kolmogorov distance*

$$d_{kol}(\mu_n, \mathbf{a}) \leq kn^{-1/4}$$

In fact, the rate in Theorem 8 is sharp as we see in the example of the next section.

Before passing to the proofs, we make some comments about the difficult parts of this problem. The first problem we face is that the Kolmogorov distance is defined at the level of cumulative distribution functions and the definition of monotone convolution is at the level of F-transforms. So, a very natural strategy is to estimate the Kolmogorov distance in terms of the F-transforms (equivalently of Cauchy transforms) of the involved measures. This task is done by Bai's inequality. This inequality was also the approach used in the proofs of the free Berry-Esseen theorem given by Chistyakov and Götze [13] and Kargin [21].

Bai's inequality reduces the problem to the level of F-transforms. The task now is essentially to estimate $F_{\mu^{\triangleright n}}(z)$ with precision for z near the real axis. This is the most difficult part because it is not trivial to obtain even a basic expansion of $F_{\mu^{\triangleright n}}(z)$, and any such expansion is quite messy.

The results of this chapter were published in [7] with the collaboration of Arizmendi and Wang.

4.2 Proofs of Main Results

In this section we prove our two theorems from a general argument. The idea is to use Bai's inequality. It implies to obtain nontrivial estimates of some integrals and some other expressions involving F-transforms. For the sake of clarity, we put the main estimations into lemmas, and we prove them in the following sections.

Proof of Theorems 7 *and* 8. Let μ be a probability measure of zero mean and unit variance. Let n be a positive integer, and define $\mu_n := D_{\frac{1}{\sqrt{n}}}\mu^{\triangleright n}$. Our objective is to estimate the Kolmogorov distance $d_{kol}(\mu_n, \mathbf{a})$ as n goes to infinity. As we mention in the previous section, the strategy is to use Bai's inequality.

The following lemma allows us to use Bai's inequality.

Lemma 3.

$$\int_{\mathbb{R}} |\mathcal{F}_{\mu_n}(x) - \mathcal{F}_{\mathbf{a}}(x)|dx < \infty. \tag{4.2}$$

We may now conclude by Bai's inequality that

$$d_{kol}(\mu_n, \mathbf{a}) \leq \frac{1}{\pi(2\gamma - 1)} \left(\int_{\mathbb{R}} |G_{\mu_n}(z) - G_{\mathbf{a}}(z)|dx + \frac{1}{y} \sup_{r\in\mathbb{R}} \int_{|t|\leq 2ay} |\mathcal{F}_{\mathbf{a}}(r+t) - \mathcal{F}_{\mathbf{a}}(r)|dt \right), \tag{4.3}$$

where $z = x + iy$, $y > 0$, and a and γ are constants such that

$$\gamma = \frac{1}{\pi} \int_{|t|<a} \frac{1}{t^2+1} dt > \frac{1}{2}. \tag{4.4}$$

The second integral is not difficult to handle. We get that

Lemma 4.

$$\frac{1}{y} \sup_{r\in\mathbb{R}} \int_{|t|\leq 2ay} |\mathcal{F}_{\mathbf{a}}(r+t) - \mathcal{F}_{\mathbf{a}}(r)|dt \leq \frac{16a^3}{3\sqrt[4]{2}\pi}\sqrt{y}.$$

As a consequence we obtain for $\gamma = \frac{2}{3}$ and $a = \sqrt{3}$ that (4.3) becomes

$$d_{kol}(\mu_n, \mathbf{a}) \leq \frac{3}{\pi} \left(\int_{\mathbb{R}} |G_{\mu_n}(z) - G_{\mathbf{a}}(z)|dx + 8\sqrt{y} \right). \tag{4.5}$$

The problem is reduced now to bound the difference of the Cauchy transforms $|G_{\mu_n}(z) - G_{\mathbf{a}}(z)|$. Note that the right side of the inequality only depends on y: For any $y > 0$ the right side of the inequality takes a value, and that value is greater or equal than $d_{kol}(\mu_n, \mathbf{a})$. The problem here is that as y goes to zero we lose control of the integral $\int_{\mathbb{R}} |G_{\mu_n}(z) - G_{\mathbf{a}}(z)| dx$. Since the monotone convolution is defined in terms of the F-transform, it is more convenient to work with them instead of the Cauchy transforms. Consider the F-transforms $F_{\mu_n}(z)$ and $F_{\mathbf{a}}$ of the measures μ_n and $\mathbf{a}$ respectively. Let $\sqrt{\cdot} : \mathbb{C}\backslash[0,\infty) \to \mathbb{C}^+$ denote the principal root square. Define the complex function $\epsilon_n(z) : \mathbb{C}^+ \to \mathbb{C}$ as $\epsilon_n(z) := z^2 - 2 - F_{\mu_n}^2$, that is $F_{\mu_n}(z) = \sqrt{z^2 - 2 + \epsilon_n(z)}$. So we have that

$$\int_{\mathbb{R}} |G_{\mu_n}(z) - G_{\mathbf{a}}(z)| dx = \int_{\mathbb{R}} \Big| \frac{1}{\sqrt{z^2 - 2 + \epsilon_n(z)}} - \frac{1}{\sqrt{z^2 - 2}} \Big| dx. \tag{4.6}$$

We may think of $|\epsilon_n(z)|$ as an error between $F_{\mu_n}(z)$ and $F_{\mathbf{a}}(z)$. The next lemma gives an estimation of the last integral.

Lemma 5. *Let $\epsilon : \mathbb{C}^+ \to \mathbb{C}$ such that $|\epsilon(z)| < \frac{3y}{2}$. Then*

$$\int_{\mathbb{R}} \Big| \frac{1}{\sqrt{z^2 - 2 + \epsilon(z)}} - \frac{1}{\sqrt{z^2 - 2}} \Big| dx \le C\sqrt{y}.$$

We make some comments about this lemma. For fixed $y > 0$ the point $z^2 - 2$ draws a parabola being the x-axis its axis of symmetry. The vertex is at the point $(-2, -y^2)$. It is easy to see that the parabola is always at a distance greater than 2y of the semi-axis $[0, \infty)$. From this fact we conclude first that the hypothesis of the lemma must ask at least that $|\epsilon(z)| < 2y$, so we guaranteed that the term $z^2 - 2 + \epsilon(z)$ is in the domain of the principal root square. Second, by asking the error to be smaller, say $|\epsilon(z)| < \frac{3y}{2}$, we assure that $|z^2 - 2 + \epsilon(z)| > y/2$, and thus the term $\frac{1}{\sqrt{z^2-2+\epsilon(z)}}$ can not be too big and its size is depending essentially on how small is y. Thus, one may expect an estimation of the integral (4.6) in terms only of y. We could get a better estimation than the one given in the lemma by taking the error term much smaller (compared to y), but for our purposes with this estimation is enough.

Let us go back to the estimation of $d_{kol}(\mu_n, \mathbf{a})$. By lemma 5, (4.6), and (4.5) we have that if $|\epsilon_n(z)| < \frac{3y}{2}$, then we have the Kolmogorov distance

$$d_{kol}(\mu_n, \mathbf{a}) \le \frac{3}{\pi} \left(C\sqrt{y} + 8\sqrt{y} \right). \tag{4.7}$$

We next conclude Theorem 7. First, we give an expression for the error $\epsilon_n(z)$ from which we derive a bound of $|\epsilon_n(z)|$ in terms only of y and n. Then by choosing $y = \frac{1}{n^\alpha}$ for some adequate α, we may conclude that $|\epsilon_n(z)| < \frac{3y}{2}$ and apply the last inequality.

goes to zero we lose control in the precision of the estimate of $F_{\mu_n}(z)$, because in general we do not have good estimates of Cauchy transform $G(z)$ as y goes to zero, and we would not have a good estimate of $|\epsilon_n(z)|$. So, we need to find an optimal choice of $y = \frac{1}{n^\alpha}$ that minimize $|\epsilon_n(z)|$.

We left the necessary expansions and estimations for the next section. Suppose that $m_4(\mu) < \infty$. By equation (4.24) we have that

$$\epsilon_n(z) = -\frac{2}{n} \sum_{j=1}^{j=n} \int_{\mathbb{R}} \frac{t}{\sqrt{n} F_n^{\circ n-1}(z) - t} d_\nu(t) + \frac{1}{n} \sum_{j=1}^{j=n} G_\nu(\sqrt{n} F_n^{j-1}(z))^2, \tag{4.8}$$

where $F_n(z) = F_{D_{\frac{1}{\sqrt{n}}}\mu}(z)$ and ν is the probability measure such that $F_\mu(z) = z - G_\nu(z)$. Note that by Proposition 10 we have that $m_2(\nu) < \infty$.

Note that by definition of $B_n(z)$, see next section, we have $B_n(z) = \frac{1}{n}\sum_{j=1}^{j=n} G_\nu(\sqrt{n}F_n^{j-1}(z))^2$. If $n \geq \max\{4\alpha^2, 4w(\mathbb{R})^2\}$, then by the proof of lemma 7 we have

$$|\frac{1}{n}\sum_{j=1}^{j=n} G_\nu(\sqrt{n}F_n^{j-1}(z))^2| \leq \frac{8}{\sqrt{n}}. \tag{4.9}$$

Observe that the stronger hypothesis $m_6(\mu) < \infty$ of the lemma is not used to conclude this.

Now, we have the following estimation

$$\begin{aligned}\left| -\frac{2}{n}\sum_{j=1}^{j=n}\int_{\mathbb{R}} \frac{t}{\sqrt{n}F_n^{\circ j-1}(z) - t} d_\nu(t)\right| &\leq \frac{2}{n}\sum_{j=1}^{j=n} \frac{m}{Im(\sqrt{n}F_n^{\circ j-1}(z))} \\ &\leq \frac{2}{n}\sum_{j=1}^{j=n} \frac{m}{\sqrt{n}Im((z))} \\ &= \frac{2m}{\sqrt{n}y},\end{aligned}$$

where $m = \int_{\mathbb{R}} |t| d\nu(t)$. Observe that m is finite because of $m_2(\nu) < \infty$.

By the above estimations we have that

$$|\epsilon_n(z)| \leq \frac{2m}{\sqrt{n}y} + \frac{8}{n^{1/2}}. \tag{4.10}$$

Take $y = \frac{2m}{n^{1/4}}$. Then we have that

$$|\epsilon_n(z)| \leq \frac{1}{n^{1/4}} + \frac{8}{n^{1/2}}. \tag{4.11}$$

Clearly for n big enough we have that $|\epsilon_n(z)| \leq \frac{3y}{2}$, and then equation (4.7) implies

$$d_{kol}(\mu_n, \mathbf{a}) \leq \frac{3}{\pi}\left(\frac{\sqrt{2m}C}{n^{1/8}} + \frac{8\sqrt{2m}}{n^{1/8}}\right). \tag{4.12}$$

We conclude Theorem 7.

Now assume that $m_6(\mu) < \infty$. Take $y = \frac{h}{n^{1/2}}$, where h is as in lemma 7. Then by the same lemma we obtain that for n big enough

$$|\epsilon_n(z)| \leq \frac{h}{n^{1/2}} < \frac{3y}{2}. \tag{4.13}$$

Hence, the equation (4.7) implies

$$d_{kol}(\mu_n, \mathbf{a}) \leq \frac{3}{\pi}\left(\frac{\sqrt{h}C}{n^{1/4}} + \frac{8\sqrt{h}}{n^{1/4}}\right). \tag{4.14}$$

We conclude Theorem 8. □

Example of sharpness for Theorem 8.

We prove that for some measure μ with $m_6(\mu) < \infty$ we have that $d_{kol}(\mu_n, \gamma) \geq \frac{1}{5n^{1/4}}$. We need the next proposition, which is derived from Section 5 of [18].

Proposition 16. *Let $c \geq 0$. Let μ the measure with F-transform given by $F_\mu(z) = c + \sqrt{(z-c)^2 - 2}$. Then:*

1) We have that $\mu = \mu_{ac} + \mu_{sing}$ with $Supp(\mu_{ac}) \subset [c - \sqrt{2}, c + \sqrt{2}]$ and $\mu_{sing} = \frac{|c|}{c^2+2}\delta_{c-\sqrt{c^2+2}}$.

2) The F-transform of $\mu^{\triangleright n}$ is $F_{\mu^{\triangleright n}}(z) = c + \sqrt{(z-c)^2 - 2n}$.

Let μ the measure given by the F-transform $F_\mu(z) = 1 + \sqrt{(z-1)^2 - 2}$. The proposition implies that $F_{\mu^{\triangleright n}}(z) = 1 + \sqrt{(z-1)^2 - 2n}$. By equation (2.9) the F-transform of $\mu_n := D_{\frac{1}{\sqrt{n}}}\mu^{\triangleright n}$ is $F_{\mu_n}(z) = \frac{1}{\sqrt{n}}F_{\mu^{\triangleright n}}(\sqrt{n}z) = \frac{1}{\sqrt{n}} + \sqrt{(z - \frac{1}{\sqrt{n}})^2 - 2}$. Take l to be the infimum of the $Supp(\mu_n)$. Again by the previous proposition we see that $l = \frac{1}{\sqrt{n}} - \sqrt{-2 + \frac{1}{n}} > -\sqrt{2} + \frac{0.5}{\sqrt{n}}$. It follows that $d_{kol}(\mu_n, \mathbf{a}) = \sup_{x\in\mathbb{R}} |\mu_n(-\infty, x] - \mathbf{a}(-\infty, x]| \geq |\mu_n(-\infty, -\sqrt{2} + \frac{0.5}{\sqrt{n}}] - \mathbf{a}(-\infty, -\sqrt{2} + \frac{0.5}{\sqrt{n}}]| = \mathbf{a}((-\sqrt{-2}, -\sqrt{2} + \frac{0.5}{\sqrt{n}}])$. Finally, note that

$$
\begin{aligned}
\mathbf{a}((-\sqrt{-2}, -\sqrt{2} + \frac{0.5}{\sqrt{n}}]) &= \frac{1}{\pi}\int_{-\sqrt{2}}^{-\sqrt{2}+\frac{0.5}{\sqrt{n}}} \frac{1}{\sqrt{2-x^2}}dx \\
&= \frac{1}{\pi}\int_{-\sqrt{2}}^{-\sqrt{2}+\frac{0.5}{\sqrt{n}}} \frac{1}{\sqrt{(\sqrt{2}-x)(\sqrt{2}+x)}}dx \\
&\geq \frac{1}{2\pi}\int_{-\sqrt{2}}^{-\sqrt{2}+\frac{0.5}{\sqrt{n}}} \frac{1}{\sqrt{(\sqrt{2}+x)}}dx \\
&= \frac{1}{2\pi}\int_{0}^{\frac{0.5}{\sqrt{n}}} \frac{1}{\sqrt{x}}dx \\
&= \frac{1}{2\pi}2(\frac{0.5}{\sqrt{n}})^{\frac{1}{2}} \geq \frac{1}{5n^{1/4}}.
\end{aligned}
$$

4.3 Properties of F_{μ_n}

Let $\mu \in \mathcal{M}_0^1$ such that $m_4(\mu) < \infty$. Define $\mu_n = D_{\frac{1}{\sqrt{n}}}\mu^{\triangleright n}$. In this section we obtain an estimate of $F_{\mu_n}(z)$ of the form $F_{\mu_n} = \sqrt{z^2 - 2 - \epsilon_n(z)}$ where $e_n(z)$ is an error term that goes to zero as n goes to infinity (have in mind that the F-transform of the limiting distribution is $F_{\mathbf{A}}(z) = \sqrt{z^2-2}$). We also estimate $|\epsilon_n(z)|$ for $Im(z)$ small.

To give a better perspective of the problem of computing $F_{\mu_n}(z)$ we first briefly review what is known about it. That problem is equivalent to compute $F_{\mu^{\triangleright n}}(z)$. We find two main contributions in the literature. The first is in the context of monotone infinitely divisibility. In [25] and [9] is established that a measure μ is monotone infinitely divisibility if and only if there exists a composition semigroup of F-transforms $\{F_t(z)\}_{t\geq 0}$ with $F_0 = Id$ and $F_1 = \mu$. In particular $F_n(z) = F_{\mu^{\triangleright n}}(z)$. In [25] Muraki computes $F_{\mu^{\triangleright n}}(z)$ for the cases where μ is a point measure δ_a, the arcsine distribution, the monotone Poisson distribution, the deformed arcsine distribution (our example of sharpness for Theorem 8), and the Cauchy distribution. Also in the context of obtaining examples of monotone infinitely divisible distributions, in [18] Hasebe obtains $F_{\mu^{\triangleright n}}(z)$ for the measure given by $F_\mu(z) = c + [(z-c)^\alpha + b]^{\frac{1}{\alpha}}$. This is a generalization of the deformed arcsine distribution.

The second contribution is given in [34] by Wang. In example 3.5 it is computed $F_{\mu_n}(z)$ for the Bernoulli distribution $\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$. He obtains that

$$F_{\mu_n}(z) = \sqrt{z^2 - 2 + \frac{1}{n^2}\sum_{j=0}^{n-1} \frac{1}{F^{\circ j}_{D_{\frac{1}{\sqrt{n}}}\mu}(z)^2}}.$$

His method to obtain this was essentially instead of obtaining directly $F_{\mu_n}(z) = F^{\circ n}_{D_{\frac{1}{\sqrt{n}}}\mu}(z)$ he computed $g^{-1} \circ F^{\circ n}_{D_{\frac{1}{\sqrt{n}}}\mu} \circ g(z)$ where $g(z) = \sqrt{z}$. This method can be generalized to obtain the same result that we obtained here, but here we use another approach that we describe next.

By Proposition 10 there exists $\nu \in \mathcal{M}$ such that $m_2(\nu) < \infty$ and $F_\mu(z) = z - G_\nu(z)$ for $z \in \mathbb{C}^+$. Let us define $F_n(z)$ as (see Proposition 2.9)

$$F_n(z) := F_{D_{\frac{1}{\sqrt{n}}}\mu}(z) = z - \frac{1}{\sqrt{n}} G_\nu(\sqrt{n}z) \quad \text{for } z \in \mathbb{C}^+ \tag{4.15}$$

By definition of monotonic convolution, we have that $F_{\mu_n}(z) = F_{D_{\frac{1}{\sqrt{n}}}\mu^{\triangleright n}}(z) = F_n^{\circ n}(z)$.

Observe that by (2.7) we have that

$$G_\nu(z) = \frac{1}{z} + \frac{1}{z}\int_{\mathbb{R}} \frac{t}{z-t} d\nu(t), \tag{4.16}$$

and so

$$zG_\nu(z) = 1 + \int_{\mathbb{R}} \frac{1}{z-t} d\nu(t). \tag{4.17}$$

The j-th iteration of $F_n(z)$ is given by

$$F_n^{\circ j}(z) = F_n^{j-1}(z) - \frac{1}{\sqrt{n}} G_\nu(\sqrt{n} F_n^{j-1}(z)). \tag{4.18}$$

Squaring this equation we obtain that

$$F_n^{\circ j}(z)^2 = F_n^{j-1}(z)^2 - \frac{1}{n}\sqrt{n} F_n^{j-1}(z) G_\nu(\sqrt{n}F_n^{j-1}(z)) + \frac{1}{n} G_\nu(\sqrt{n}F_n^{j-1}(z))^2. \tag{4.19}$$

By equation (4.17) it follows that

$$F_n^{\circ j}(z)^2 = F_n^{j-1}(z)^2 - \frac{2}{n} - \frac{2}{n}\int_{\mathbb{R}} \frac{t}{\sqrt{n}F_n^{\circ n-1}(z) - t} d_\nu(t) + \frac{1}{n} G_\nu(\sqrt{n}F_n^{j-1}(z))^2. \tag{4.20}$$

We sum the equation from j=1 to j=n as follows:

$$\sum_{j=1}^{j=n} F_n^{\circ j}(z)^2 = \sum_{j=1}^{j=n} F_n^{j-1}(z)^2 - 2 - \frac{2}{n}\sum_{j=1}^{j=n}\int_{\mathbb{R}} \frac{t}{\sqrt{n}F_n^{\circ n-1}(z) - t} d_\nu(t) + \frac{1}{n}\sum_{j=1}^{j=n} G_\nu(\sqrt{n}F_n^{j-1}(z))^2. \tag{4.21}$$

This implies that

$$F_n^{\circ n}(z)^2 = z^2 - 2 - \frac{2}{n}\sum_{j=1}^{j=n}\int_{\mathbb{R}} \frac{t}{\sqrt{n}F_n^{\circ n-1}(z) - t} d_\nu(t) + \frac{1}{n}\sum_{j=1}^{j=n} G_\nu(\sqrt{n}F_n^{j-1}(z))^2. \tag{4.22}$$

We conclude that

$$F_n^{\circ n}(z) = \sqrt{z^2 - 2 + \epsilon(z)}, \tag{4.23}$$

where

$$\epsilon(z) = -\frac{2}{n}\sum_{j=1}^{j=n}\int_{\mathbb{R}}\frac{t}{\sqrt{n}F_n^{\circ n-1}(z) - t}d_\nu(t) + \frac{1}{n}\sum_{j=1}^{j=n}G_\nu(\sqrt{n}F_n^{j-1}(z))^2. \tag{4.24}$$

This expression is used for proving Theorem 7. But for obtaining Theorem 8 we need a more sophisticated expansion of $\epsilon(z)$.

Recall that $m_2(\nu) < \infty$. By Proposition 10 we can expand $G_\nu(z)$ as

$$G_\nu(z) = \frac{1}{z - \hat{G}_w(z)}, \tag{4.25}$$

where $\hat{G}_w(z) = \alpha + G_w(z)$, $\alpha \in \mathbb{R}$, and G_w is the Cauchy transform of some positive finite measure w with $w(\mathbb{R}) = var(\nu)$.

A straight forward computation gives

$$\int_{\mathbb{R}}\frac{t}{z-t}d\nu(t) = zG_\nu(z) - 1 = \frac{\hat{G}_w(z)}{z - \hat{G}_w(z)}. \tag{4.26}$$

Using (4.25) and (4.26) we rewrite the error as

$$\epsilon_n(z) = -\frac{2}{n}\sum_{j=0}^{n-1}\frac{\hat{G}_w(\sqrt{n}F_n^{\circ j}(z))}{\sqrt{n}F_n^{\circ j}(z) - \hat{G}_w(\sqrt{n}F_n^{\circ j}(z))} + \frac{1}{n}\sum_{j=0}^{n-1}\frac{1}{[\sqrt{n}F_n^{\circ j}(z) - \hat{G}_w(\sqrt{n}F_n^{\circ j}(z))]^2}. \tag{4.27}$$

For a brief reference to $\epsilon_n(z)$, we write it as $\epsilon_n(z) = A_n(z) + B_n(z)$ where

$$A_n(z) := -\frac{2}{n}\sum_{j=0}^{n-1}\frac{G_w(\sqrt{n}F_n^{\circ j}(z))}{\sqrt{n}F_n^{\circ j}(z) - G_w(\sqrt{n}F_n^{\circ j}(z))},$$

and

$$B_n(z) := \frac{1}{n}\sum_{j=0}^{n-1}\frac{1}{[\sqrt{n}F_n^{\circ j}(z) - G_w(\sqrt{n}F_n^{\circ j}(z))]^2}.$$

In the next lemma we obtain some useful estimates that we use for bounding $|\epsilon_n(z)|$.

Lemma 6. *Suppose that $n \geq \max\{4\alpha^2, 4w(\mathbb{R})^2\}$. Then for all z such that $1/\sqrt{n} \leq Im(z) \leq 1$ we have that*

i) $|F_n^{\circ n}(z) - z| \leq 8.$

ii) $|\sum_{j=0}^{n-1} Im[G_\nu(\sqrt{n}F_n^{\circ j}(z))]| \leq 8\sqrt{n}.$

Proof. Since $F_n(z) = z - \frac{1}{\sqrt{n}}G_\nu(\sqrt{n}z)$, then by (2.4) we have

$$|F_n(z) - z| \leq \frac{1}{\sqrt{n}} \leq 1 \quad \text{for } Im(z) \geq \frac{1}{\sqrt{n}}. \tag{4.28}$$

Now rewriting $F_n(z)$ with (4.25) we get

$$F_n(z) = z - \frac{1}{\sqrt{n}} \frac{1}{\sqrt{n}z - \alpha - G_w(\sqrt{n}z)}$$
$$= z - \frac{1}{n} \frac{1}{z - \frac{\alpha}{\sqrt{n}} - \frac{1}{\sqrt{n}} G_w(\sqrt{n}z)}.$$

Hence, we have that

$$|F_n(z) - z| \leq \frac{1}{n} \quad \text{for } Im(z) \geq \frac{1}{\sqrt{n}},\ |z| \geq 2, \text{ and } n \geq \max\{4\alpha^2, 4w(\mathbb{R})^2\}. \tag{4.29}$$

We can now prove part i). Suppose $Im(z) \geq \frac{1}{\sqrt{n}}$ and $n \geq \max\{4\alpha^2, 4w(\mathbb{R})^2\}$. We want to show that $|F_n^{\circ n}(z) - z| \leq 7$. If $|z| \geq 3$ (4.29) implies that $|F_n^{\circ n}(z) - z| \leq 1$. Now assume that $|z| < 3$. Define $J = \{j \in \{1, 2, ..., n\} \mid |F_n^{\circ j}(z)| \geq 3\}$. If $J = \varnothing$ then $|F_n^{\circ n}(z) - z| \leq |F_n^{\circ n}(z)| + |z| < 6$. If $J \neq \varnothing$, then take $k = \min J$. By (4.28) and (4.29) we have

$$|F_n^{\circ n}(z) - z| \leq |F_n^{\circ n}(z) - F_n^{\circ k}(z)| + |F_n^{\circ k}(z) - F_n^{\circ k-1}(z)| + |F_n^{\circ k-1}(z) - z| \leq 8.$$

Now we prove part ii). Since

$$F_n(z) = z - \frac{1}{\sqrt{n}} G_\nu(\sqrt{n}z),$$

then

$$\sum_{j=1}^{n} F_n^{\circ j}(z) = \sum_{j=0}^{n-1} F_n^{\circ j}(z) - \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} G_\nu(\sqrt{n} F_n^{\circ j}(z)).$$

It follows that

$$F_n^{\circ n}(z) = z - \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} G_\nu(\sqrt{n} F_n^{\circ j}(z)),$$

and hence

$$|\sum_{j=0}^{n-1} G_\nu(\sqrt{n} F_n^{\circ j}(z)| = \sqrt{n} |F_n^{\circ n}(z) - z|.$$

If we suppose that $Im(z) \geq \frac{1}{\sqrt{n}}$ and $n \geq \max\{4\alpha^2, 4w(\mathbb{R})^2\}$, then by part i) we conclude that

$$|\sum_{j=0}^{n-1} G_\nu(\sqrt{n} F_n^{\circ j}(z)| \leq 8\sqrt{n}.$$

□

In the following lemma we obtain a bound for $|\epsilon(z)|$.

Lemma 7. *Suppose $m_6(\mu) < \infty$. Then there exists a constant h only depending on μ such that for all z with $Im(z) \geq 1/\sqrt{n}$ we have that $|\epsilon(z)| \leq \frac{h}{\sqrt{n}}$.*

Proof. Let z a complex number such that $Im(z) \geq \frac{1}{\sqrt{n}}$. Recall that $\epsilon(z) = A_n(z) + B_n(z)$. We first bound $B_n(z)$ and then $A_n(z)$. In the proof of Lemma 6 we see that F_n increases the imaginary part. So

$$Im[F_n^{\circ j}(z)] \geq Im(z) \geq \frac{1}{\sqrt{n}}. \tag{4.30}$$

Now, by lemma 6 part ii) and (4.25) it follows that

$$\begin{aligned}
3\sqrt{n} \geq |\sum_{j=0}^{n-1} G_\nu(\sqrt{n}F_n^{\circ j}(z))| &\geq Im[\sum_{j=0}^{n-1} G_\nu(\sqrt{n}F_n^{\circ j}(z))] \\
&= \sum_{j=0}^{n-1} \frac{Im[\sqrt{n}F_n^{\circ j}(z)] - Im[\hat{G}_w(\sqrt{n}F_n^{\circ j}(z))]}{|\sqrt{n}F_n^{\circ j}(z) - \hat{G}_w(\sqrt{n}F_n^{\circ j}(z))|^2} \\
&\geq \sum_{j=0}^{n-1} \frac{1}{|\sqrt{n}F_n^{\circ j}(z) - \hat{G}_w(\sqrt{n}F_n^{\circ j}(z))|^2},
\end{aligned} \tag{4.31}$$

where the last inequality is by (4.30) and that $Im(\hat{G_w}(\cdot)) < 0$ over $\mathbb{C}^+$. We conclude that $B_n(z) \leq \frac{3}{\sqrt{n}}$.

On the other hand, suppose $\int_{\mathbb{R}} t^6 d\mu(t) < \infty$. Proposition (x) implies that $\int_{\mathbb{R}} t^4 dv(t) < \infty$, and by the same reason $\int_{\mathbb{R}} t^2 dw(t) < \infty$. Hence $\int_{\mathbb{R}} |t| dw(t) =: M < \infty$ and $w(\mathbb{R}) = var(\nu) < \infty$. Using the definition of $\hat{G_w}$ and (4.30) we have that

$$\begin{aligned}
|\hat{G}_w(\sqrt{n}F_n^{\circ j}(z)))| &\leq |\alpha| + \int |\frac{1}{\sqrt{n}F_n^{\circ j}(z)) - t}| dw(t) \\
&\leq |\alpha| + \frac{w(\mathbb{R})}{Im[\sqrt{n}F_n^{\circ j}(z))]} \\
&\leq |\alpha| + w(\mathbb{R}) =: K.
\end{aligned} \tag{4.32}$$

From (4.26) we obtain the following estimate

$$|\sqrt{n}F_n^{\circ j}(z)G_w(\sqrt{n}F_n^{\circ j}(z)))| \leq \int |\frac{t}{\sqrt{n}F_n^{\circ j}(z)) - t}| dw(t) \leq \frac{\int_{\mathbb{R}} |t| dw(t)}{Im[\sqrt{n}F_n^{\circ j}(z))]} \leq M.$$

We can now bound $A_n(z)$ as follows:

$$\begin{aligned}
|\sum_{j=0}^{n-1} \frac{\hat{G}_w(\sqrt{n}F_n^{\circ j}(z))}{\sqrt{n}F_n^{\circ j}(z) - \hat{G}_w(\sqrt{n}F_n^{\circ j}(z))}| &= |\sum_{j=0}^{n-1} \frac{\alpha + G_w(\sqrt{n}F_n^{\circ j}(z))}{\sqrt{n}F_n^{\circ j}(z) - G_w(\sqrt{n}F_n^{\circ j}(z))]}| \\
\leq |\sum_{j=0}^{n-1} \frac{\alpha}{\sqrt{n}F_n^{\circ j}(z) - G_w(\sqrt{n}F_n^{\circ j}(z))]}| &+ |\sum_{j=0}^{n-1} \frac{G_w(\sqrt{n}F_n^{\circ j}(z))}{\sqrt{n}F_n^{\circ j}(z) - G_w(\sqrt{n}F_n^{\circ j}(z))]}|.
\end{aligned}$$

The first sum of the last expression is $|\alpha|| \sum_{j=0}^{n-1} G_\nu(\sqrt{n}F_n^{\circ j}(z))|$, and lemma 6 implies that it is bounded by $9|\alpha|\sqrt{n}$. Now from the previous estimates note that

$$
\begin{aligned}
|\sum_{j=0}^{n-1} \frac{G_w(\sqrt{n}F_n^{\circ j}(z))}{\sqrt{n}F_n^{\circ j}(z) - \hat{G}_w(\sqrt{n}F_n^{\circ j}(z))}| &= |\sum_{j=0}^{n-1} \frac{G_w(\sqrt{n}F_n^{\circ j}(z))[\sqrt{n}F_n^{\circ j}(z) - \hat{G}_w(\sqrt{n}F_n^{\circ j}(z))]}{\sqrt{n}F_n^{\circ j}(z) - G_w(\sqrt{n}F_n^{\circ j}(z))]^2}| \\
&= |\sum_{j=0}^{n-1} \frac{\sqrt{n}F_n^{\circ j}(z)G_w(\sqrt{n}F_n^{\circ j}(z)) - \alpha G_w(\sqrt{n}F_n^{\circ j}(z)) - G_w^2(\sqrt{n}F_n^{\circ j}(z))}{\sqrt{n}F_n^{\circ j}(z) - G_w(\sqrt{n}F_n^{\circ j}(z))]^2}| \\
&\leq \sum_{j=0}^{n-1} \frac{|\sqrt{n}F_n^{\circ j}(z)G_w(\sqrt{n}F_n^{\circ j}(z))| + |\alpha G_w(\sqrt{n}F_n^{\circ j}(z))| + |G_w^2(\sqrt{n}F_n^{\circ j}(z))|}{|\sqrt{n}F_n^{\circ j}(z) - G_w(\sqrt{n}F_n^{\circ j}(z))|^2} \\
&\leq \sum_{j=0}^{n-1} \frac{M + |\alpha|K + K^2}{|\sqrt{n}F_n^{\circ j}(z) - G_w(\sqrt{n}F_n^{\circ j}(z))|^2}
\end{aligned}
$$

By (4.31) we conclude that $|A_n(z)| = O(\frac{1}{\sqrt{n}})$. □

4.4 Integral Estimates

In the proof of the main theorems of Section 4.2, we put all the estimations concerning integrals into lemmas. These are the Lemmas 3,4 and 8. In this section we prove them.

Proof of Lemma 3. We have that

$$
\begin{aligned}
&\int_{\mathbb{R}} |\mathcal{F}\mu_n(x) - \mathcal{F}_{\mathbf{a}}(x)|dx \\
&= \int_{-\infty}^{-\sqrt{2}} \mathcal{F}_{\mu_n}(x)dx + \int_{-\sqrt{2}}^{\sqrt{2}} |\mathcal{F}_{\mu_n}(x) - \mathcal{F}_{\mathbf{a}}(x)|dx + \int_{\sqrt{2}}^{\infty} (1 - \mathcal{F}_{\mu_n}(x))dx \\
&= \int_{-\infty}^{-\sqrt{2}} \mu_n\left((-\infty, x]\right)dx + \int_{-\sqrt{2}}^{\sqrt{2}} |\mathcal{F}_{\mu_n}(x) - \mathcal{F}_{\mathbf{a}}(x)|dx + \int_{\sqrt{2}}^{\infty} \mu_n\left([x, \infty)\right)dx \\
&= \int_{x \geq \sqrt{2}} \mu_n\{|t| \geq x\}dx + \int_{-\sqrt{2}}^{\sqrt{2}} |\mathcal{F}_{\mu_n}(x) - \mathcal{F}_{\mathbf{a}}(x)|dx
\end{aligned}
$$

We want to bound the last two integrals. The second one is clearly bounded. To see that the first integral is also bounded just note that since $m_2(\mu_n) = 1$, then $\int_{|t| \geq x} t^2 d\mu_n \leq 1$, which implies that $\mu_n\{|t| \geq x\} \leq \frac{1}{x^2}$ for $x \geq 1$.

We conclude that

$$\int_{\mathbb{R}} |\mathcal{F}\mu_n(x) - \mathcal{F}_{\mathbf{a}}(x)|dx < \infty. \tag{4.33}$$

□

Proof of Lemma 4. Without loss of generality let us suppose $|ay| < 1/4$. We have that $|\mathcal{F}_{\mathbf{a}}(r + t) - \mathcal{F}_{\mathbf{a}}(r)| = |\int_r^{r+t} d\mathbf{a}(x)|$ and $\mathbf{a}$ is supported in $[-\sqrt{2}, \sqrt{2}]$ with density $d\mathbf{a}(x) = \frac{1}{\pi\sqrt{2-x^2}}$. It is not hard to see that $|\mathcal{F}_{\mathbf{a}}(r + t) - \mathcal{F}_{\mathbf{a}}(r)| \leq \mathcal{F}_{\mathbf{a}}(-\sqrt{2} + |t|) - \mathcal{F}_{\mathbf{a}}(-\sqrt{2})$ for $r \in \mathbb{R}$ and $|t| \leq \frac{1}{2}$. This implies that

$$\int_{|t|\le 2ay} |\mathcal{F}_{\mathbf{a}}(r+t)-\mathcal{F}_{\mathbf{a}}(r)|dt \le 2\int_0^{2ay} (\mathcal{F}_{\mathbf{a}}(-\sqrt{2}+t)-\mathcal{F}_{\mathbf{a}}(-\sqrt{2}))dt$$
$$= 2\int_0^{2ay}\int_{-\sqrt{2}}^{-\sqrt{2}+t} \frac{1}{\pi\sqrt{2-s^2}}dsdt$$
$$\le \frac{2}{\pi\sqrt{2}\sqrt[4]{2}}\int_0^{2ay}\int_{-\sqrt{2}}^{-\sqrt{2}+t} \frac{1}{\sqrt{\sqrt{2}+s}}dsdt$$
$$= \frac{\sqrt{2}}{\pi\sqrt[4]{2}}\int_0^{2ay}\int_0^{t} \frac{1}{\sqrt{s}}dsdt$$
$$= \frac{4\sqrt{2}(2a)^{3/2}}{3\sqrt[4]{2}\pi}y^{3/2},$$

and the desired result follows. □

Lemma 8. *Suppose that* $|\epsilon(z)| < \frac{3y}{2}$. *Then*

$$\int_{\mathbb{R}} \Big|\frac{1}{\sqrt{z^2-2+\epsilon(z)}} - \frac{1}{\sqrt{z^2-2}}\Big|dx \le C\sqrt{y},$$

where C *is an absolute constant.*

Proof. Fix $y>0$. Recall $z = x+iy$. The proof consists in bounding the integrand. Note that

$$\Big|\frac{1}{\sqrt{z^2-2+\epsilon(z)}} - \frac{1}{\sqrt{z^2-2}}\Big| = \Big|\frac{\sqrt{z^2-2}-\sqrt{z^2-2+\epsilon(z)}}{\sqrt{z^2-2}\cdot\sqrt{z^2-2+\epsilon(z)}}\Big|. \tag{4.34}$$

We have that

$$\sqrt{z^2-2}-\sqrt{z^2-2+\epsilon(z)} = \int_L \frac{1}{\sqrt{w}}dw, \tag{4.35}$$

where L is the straight line joining z^2-2 and $z^2-2-\epsilon(z)$. Indeed, we know that $\sqrt{\cdot}$ is analytical in $\mathbb{C}-[0,\infty)$. Note that

- If $|x|\ge 1$, then $|Im(z^2-2)| = |2xy| \ge 2y$.
- If $|x|<1$, then $Re(z^2-2) = x^2-y^2-2 < -1-y^2 < -2y$.

This implies that $L \subset [0,\infty)$, since $|\epsilon(z)| \le \frac{3}{2}y$, so equation (4.35) is right. It also shows that $|z^2-2| \ge 2y$.

Claim: $|z^2-2+\delta| \ge |z^2-2|/4$ for all $\delta\in\mathbb{C}$ with $|\delta| < \frac{3y}{2}$.

Note that $|z^2-2+\delta| \ge |z^2-2| - \frac{3}{2}y$. Since $|z^2-2|\ge 2y$, then $y \le |z^2-2|/2$. We conclude that $|z^2-2+\delta| \ge |z^2-2|/4$.

Now, we can bound the numerator of (4.34) as follows

$$|\sqrt{z^2-2}-\sqrt{z^2-2+\epsilon(z)}| \le |L|\cdot \sup_{w\in L} 1/\sqrt{|w|}$$
$$= |\epsilon(z)|\cdot \sup_{w\in L} 1/\sqrt{|w|}$$
$$\le \frac{2|\epsilon(z)|}{|z^2-2|^{\frac{1}{2}}} < \frac{3y}{|z^2-2|^{\frac{1}{2}}},$$

where the second inequality is by the claim. Using this and again the claim, we conclude that

$$\Big|\frac{\sqrt{z^2-2}-\sqrt{z^2-2+\epsilon(z)}}{\sqrt{z^2-2}\cdot\sqrt{z^2-2+\epsilon(z)}}\Big| \le \frac{3y}{|z^2-2|^{1/2}}\cdot\frac{1}{|z^2-2|^{1/2}}\cdot\frac{1}{|\frac{z^2-2}{4}|^{1/2}}.$$

So we arrive to

$$\int_{\mathbb{R}}\Big|\frac{1}{\sqrt{z^2-2+\epsilon(z)}}-\frac{1}{\sqrt{z^2-2}}\Big|dx \le \int_{\mathbb{R}}\frac{6y}{|z^2-2|^{3/2}}dx.$$

Finally, lemma 9 gives the desired result. $\square$

Lemma 9. *There exists* $C' > 0$ *such that*

$$\int_{\mathbb{R}}\frac{1}{|z^2-2|^{3/2}}dx \le \frac{C'}{\sqrt{y}}.$$

Proof of Lemma. Since the integrand is symmetric with respect to x, then we have

$$\int_{\mathbb{R}}\frac{1}{|z^2-2|^{3/2}}dx = 2\int_0^\infty\frac{1}{|z^2-2|^{3/2}}dx.$$

So, it is enough to show that for some absolute $c > 0$

$$\int_0^\infty\frac{1}{|z^2-2|^{3/2}}dx \le \frac{c}{\sqrt{y}}. \tag{4.36}$$

It is convenient to split the integral into two parts:

$$\int_0^\infty\frac{1}{|z^2-2|^{3/2}}dx = \int_0^2\frac{1}{|z^2-2|^{3/2}}dx + \int_2^\infty\frac{1}{|z^2-2|^{3/2}}dx.$$

It is easy to see that $|z^2-2| > (x-1)^2$ for $x > 2$ and $y > 0$. Thus

$$\int_2^\infty\frac{1}{|z^2-2|^{3/2}}dx \le \int_2^\infty\frac{1}{(x-1)^3}dx < \infty. \tag{4.37}$$

On the other hand, we have that

$$\begin{aligned}\int_0^2\frac{1}{|z^2-2|^{3/2}}dx &= \int_0^2\frac{dx}{|z-\sqrt{2}|^{3/2}\cdot|z+\sqrt{2}|^{3/2}}\\ &\le \int_0^2\frac{dx}{|z-\sqrt{2}|^{3/2}} = \int_0^2\frac{dx}{\left[\sqrt{(x-\sqrt{2})^2+y^2}\right]^{3/2}}\\ &= \int_{-\sqrt{2}}^{2-\sqrt{2}}\frac{dx}{(x^2+y^2)^{3/4}} = \frac{1}{y^{3/2}}\int_{-\sqrt{2}}^{2-\sqrt{2}}\frac{dx}{[(x/y)^2+1]^{3/4}}\\ &\le \frac{1}{\sqrt{y}}\int_{\mathbb{R}}\frac{d\theta}{[\theta^2+1]^{3/4}}.\end{aligned}$$

Since $\int_{\mathbb{R}}\frac{d\theta}{[\theta^2+1]^{3/4}} < \infty$, then this estimation and (4.37) implies (4.36). $\square$

Chapter 5

Ergodic Properties of F-transforms

The F-transforms of certain measures belong to a class of analytic functions on the upper half-plane, called inner functions, that admit a real-valued extension to the real line. Sometimes that extension preserves the Lebesgue measure. Therefore, we are in the scope of infinite ergodic theory, and it is of interest to determine the ergodic properties of such transformations. In this context, we obtain a theorem that extends a previous work of Aaronson.

In Section 2.1 we give some preliminaries on infinite ergodic theory and the theory of inner functions on the upper half-plane. In Section 2.2 we establish our main theorem, and we see an interesting application of the monotone central limit theorem to prove a weak form of this theorem. In Section 2.3 we give a proof of our main theorem.

5.1 Preliminaries

In this section we briefly discuss the main ideas of infinite ergodic theory and the theory of inner functions on the upper-half plane. The latter theory studies the (infinite) ergodic properties of certain transforms on $\mathbb{R}$ which are induced by some analytic functions on the upper half-plane.

5.1.1 Pointwise Dual Ergodic Transform

One of the central questions in ergodic theory is to understand the frequency at which some orbit passes through some region of the space. When the measure of the underlying space is finite, the answer is well understood. But, when the measure of the underlying space is infinite, the situation becomes very complex and only in certain cases we can say something reasonable about that frequency. The objective of this section is to introduce the notion of pointwise dual ergodic transform that defines one of those situations when we can say something.

Let us start with the basic setting of ergodic theory. Let $(X, \mathcal{A}, m)$ be a measure space and $T : X \to X$ be a transformation such that T preserves the measure: $m(T^{-1}(A)) = m(A)$ for all $A \in \mathcal{A}$. Infinite ergodic theory study the case $m(X) = \infty$. Now, consider the following simple and natural situation. Given a set $A \in \mathcal{A}$ and a point $x \in A$, does the orbit of x $O(x) := \{T^k(x) \mid k = 0, 1, 2, 3, ...\}$ return to the set A, and if the orbit return to A at which frequency does it?.

To study those questions is convenient to introduce the occupation times function of a set $A \in \mathcal{A}$:

$$S_A := \sum_{k=0}^{n-1} 1_A \circ T^k. \tag{5.1}$$

Note that $S_n(A)(x)$ counts how many times before time n the orbit $O(x)$ visited the set A.

In general, it may happen that the orbit return infinitely many times or that the orbit wanders in some region of the set X. More exactly, the situation is as follows. The space X can be separated into a conservative part C and a dissipative part D such that $X = C \cup D$. In the conservative part, for any $A \in \mathcal{A}$ we have that $\lim_{n\to\infty} S_n(A)(x) = \infty$ a.e. for $x \in A$. This means that the orbit of almost every point in A will return to A infinitely often. In the dissipative part, we have the opposite behavior; the orbits that enter a set will never go back to it. We call a transformation conservative if $X = C$.

The previous argument settles the first question of our situation of study. Now, for the question about the frequency at which the orbit returns, we restrict the discussion only for conservative transformation, where it makes more sense. Note that the frequency at which the orbit returns means the asymptotic growth of the occupation time function $S_n(A)$.

Before proceeding, we also consider another natural restriction on our transformation T. Note that if exists a set A invariant by T, that is $T^{-1}(A) = A$, it implies that $TA \subset A$ and $TA^c \subset A^c$. So we could study T separately in A and A^c. We say that T is ergodic if we have a set A such that $T^{-1}(A) = A$, then $m(A) = 0$ or $m(A^c) = 0$. We restrict the situation of study to ergodic transforms.

We first consider the finite case $m(X) < \infty$. By the Poincaré recurrence theorem we have that every transform is conservative. Now, if besides it is ergodic, the Birkoff theorem tells us that $\frac{1}{n}S_A(x) \to \frac{m(A)}{m(X)}$ a.e. for $x \in X$. So, the asymptotic growth of the occupation times function is asymptotically of order some constant times n.

For the infinite case $m(X) = \infty$, Birkoff theorem says that $\frac{1}{n}S_A(x) \to 0$ a.e. for $x \in X$. One may wonder if there is some sequence $(a_n)_{n\geq 1}$ that captures the correct rate of growth of $S_n(A)(x)$. The next theorem shows that such a sequence does not exist.

Theorem 9. *For any positive sequence (a_n) and any $A \in \mathcal{A}$ we have that*

$$\overline{\lim}_{n\to\infty} \frac{1}{a_n} S_n(A)(x) = \infty \quad a.e. \text{ on } X,$$

or

$$\underline{\lim}_{n\to\infty} \frac{1}{a_n} S_n(A)(x) = 0 \quad a.e. \text{ on } X.$$

This means that any sequence $(a_n)_{n\geq 1}$ either overestimate or underestimate the growth of the occupation times function a.e. in X.

We conclude that the asymptotic pointwise behavior of the occupation times function is very complex in the infinite case, and that it seems to depend strongly on the point x where we evaluate the function. Thus, in general we can not say something reasonable. However, for some transformations there is still possible to say something about the asymptotic behavior of the occupation time function but only at the distributional level.

For some transforms T happens that we can still capture the asymptotic of $S_n(A)$ but only at the dual level. Denote $L_p := L_p(X, \mathcal{A}, m)$. Each T induces a linear isometry L in L_∞ given by $f \to f \circ T$. The dual operator (or transfer operator) $\hat{T} : L_1 \to L_1$, is defined by

$$\int_X \hat{T} f \cdot g dm = \int_X f \cdot (g \circ T) dm \quad f \in L_1,\ g \in L_\infty. \tag{5.2}$$

Some very special transforms T have the property that for any $f \in L_1$

$$\frac{1}{a_n} \sum_{k=0}^{n-1} \hat{T}^k f(x) \xrightarrow[n\to\infty]{} \int_X f dm \quad \text{a.e. for } x \in X.$$

They are called **pointwise dual ergodic transforms**. The following result of Aaronson tells us that we can capture the asymptotic of $S_n(A)$ in the distributional sense. Here, we are thinking $S_n(A)$ as a random variable.

Theorem 10. *Suppose that T is pointwise dual ergodic such that its return sequence a_n is regularly varying of index α, that is $\frac{a_{(tn)}}{a_n} \xrightarrow[n\to\infty]{} t^\alpha$ for some real α, inside the interval $\in [0,1]$. Then*

$$\frac{S_n(f)}{a_n} \xrightarrow{d} Y_\alpha, \tag{5.3}$$

where Y_α is the normalized Mittag-Leffer distribution of order α.

5.1.2 Inner Functions on the Upper Half-Plane

Some transforms $T : \mathbb{R} \to \mathbb{R}$ that preserve the Lebesgue measure can be seen as the real extension of a certain holomorphic function $F : \mathbb{C}^+ \to \mathbb{C}^+$. By instance, the Boole transform $T(x) = x - \frac{1}{x}$ is the real extension of the holomorphic function $F(z) = z - \frac{1}{z}$. The interesting thing is that the ergodic properties of these transforms T can be studied in terms of the dynamics of the function F. In this section we characterize and determine the ergodic properties of such transforms. We are based on Aaronson [2] Chapter 6.

Let $F : \mathbb{C}^+ \to \mathbb{C}^+$ be a holomorphic function. We say that T is an inner function if $\lim_{y\downarrow 0} F(x+iy)$ exist a.e. $x \in \mathbb{R}$ and belongs to $\mathbb{R}$. We denote this limit by $T_F(x)$, and we call it the real restriction of F. Next we obtain a characterization of the inner functions.

Recall that the Nevanlinna representation says that any holomorphic function $F : \mathbb{C}^+ \to \mathbb{C}^+$ can be written as

$$F(z) = \alpha_F z + \beta_F + \int_{\mathbb{R}} \frac{1+tz}{t-z} d\mu_F, \tag{5.4}$$

where $\alpha_F, \beta_F \in \mathbb{R}$, $\alpha \geq 0$, and μ_F is a positive measure on $\mathbb{R}$.

Proposition 17. *$F : \mathbb{C}^+ \to \mathbb{C}^+$ is an inner function if and only if μ_F is singular with respect the Lebesgue measure.*

Now, the next situation of interest is when an inner function preserves the Lebesgue measure.

Proposition 18. *Suppose $F : \mathbb{C}^+ \to \mathbb{C}^+$ is an inner function, then $m \circ T_F^{-1} = \frac{1}{\alpha_F}\lambda$.*

This implies that T_F preserves the Lebesgue measure if and only if $\alpha_F = 1$.

When we have a measure-preserving transformation on an infinite space, the first ergodic properties to determine are conservativity and ergodicity.

Proposition 19. *Suppose that $F : \mathbb{C}^+ \to \mathbb{C}^+$ is an inner function with $\alpha_F = 1$. Then T_F is conservative and ergodic if and only if*

$$\sum_{k=1}^{n} Im \frac{-1}{F^k(z)} = \infty \tag{5.5}$$

for some $z \in \mathbb{C}^+$.

The next ergodic property of interest is the pointwise dual ergodicity. The following theorem gives a criterion for it.

Theorem 11. *Let $F : \mathbb{C}^+ \to \mathbb{C}^+$ be an inner function. Suppose that $\alpha_F = 1$ and that its real restriction T_F is conservative and ergodic. Then T_F is pointwise dual ergodic with return sequence*

$$a_n(T_F) \sim \frac{1}{\pi}\sum_{k=1}^{n} Im\frac{-1}{F^k(z)} \quad \textit{for } \in \mathbb{C}^+. \tag{5.6}$$

5.2 Main Result

The F-transform of a measure μ is an analytic function $F_\mu : \mathbb{C}^+ \to \mathbb{C}^+$. If μ is singular respect to the Lebesgue measure, then by Proposition 17 F_μ is an inner function. Thus, its real restriction $T_\mu(x) := \lim_{y\downarrow 0} F_\mu(x+iy)$ exists a.e. on $\mathbb{R}$ and preserves the Lebesgue measure. The next theorem establish some ergodic properties of the transform T_μ.

Theorem (Aaronson). *Suppose that μ has zero mean, unit variance, and bounded support. Then T_μ is conservative and ergodic. Moreover, it is pointwise dual ergodic with return sequence $a_n(T_\mu) \sim \frac{\sqrt{2n}}{\pi}$.*

In [34] this result is generalized.

Theorem (Wang). *Suppose that μ has zero mean and unit variance. Then T_μ is conservative and ergodic.*

The following theorem is the main result of this chapter. We complete the generalization started by Wang.

Theorem 12. *Suppose that μ has zero mean and unit variance. Then T_μ is pointwise dual ergodic with return sequence $a_n(T_\mu) \sim \frac{\sqrt{2n}}{\pi}$.*

Recall that $S_n(A) = \sum_{k=0}^{n-1} 1_A \circ T^k$ is the occupation times function. Since $\frac{a_{(tn)}(T_\mu)}{a_n(T_\mu)} = t^{1/2}$, then a_n is regularly varing of index 1/2. Thus, by the Theorem 10 we have that for any Borel set A

$$\lim_{n\longrightarrow\infty} P([\frac{\pi}{\sqrt{2n}}S_n(A) \le x]) = \frac{2}{\pi}\int_0^x e^{-t^2/\pi}dt \quad \text{for } x \ge 0,$$

where P is any absolutely continuous probability measure on $\mathbb{R}$. So, we obtain a good asymptotic understanding of $S_n(A)$ in the distributional sense.

Remark 1. *Since the main object of the theorem of this section is the F-transform of some measure, one may expect a connection between the theory of Chapter 4 and the results of this section. Indeed, that is the case, and in the following lines we obtain a simple proof of our main theorem but under the weak assumption that the fourth moment is finite. By Theorem 11 we have that for any $z \in \mathbb{C}^+$*

$$a_n(T_F) \sim \frac{1}{\pi}\sum_{m=1}^{n} Im\frac{-1}{F^m(z)}. \tag{5.7}$$

For a positive integer m define $F_m(z) := F_{D_{1/\sqrt{m}}\mu}(z)$. Formula (2.9) implies that $F_m(z) = F(\sqrt{m}z)/\sqrt{m}$. Let us rewrite it as $F(z) = \sqrt{m}F_m(z/\sqrt{m})$. Therefore

$$F^{\circ m}(z) = \sqrt{m}F_m^{\circ m}(z/\sqrt{m}). \tag{5.8}$$

By (4.23) we have that $F_m^{\circ m}(z)$ admits the approximation

$$F_m^{\circ m}(z) = \sqrt{z^2 - 2 + \epsilon_m(z)} \quad \text{for } z \in \mathbb{C}^+. \tag{5.9}$$

Since we are assuming that $m_6(\mu) < \infty$, the inequality (4.13) says that for some $h > 0$ we have that $|\epsilon_m(z)| \leq \frac{h}{\sqrt{n}}$ when $Im(z) \geq \frac{h}{\sqrt{n}}$.

It follows that for $z = hi$

$$\begin{aligned}
\sum_{m=1}^{n} Im \frac{-1}{F^m(z)} &= \sum_{m=0}^{n-1} Im(\frac{-1}{F^{\circ m}(hi)}) \\
&= \sum_{m=1}^{n-1} Im(\frac{-1}{\sqrt{m}F_m^{\circ m}(hi/\sqrt{m})}) \\
&= \sum_{m=1}^{n-1} Im(\frac{-1}{\sqrt{m}\sqrt{(hi/\sqrt{m})^2 - 2 + \epsilon_m(hi/\sqrt{m})}} \\
&= \sum_{m=1}^{n-1} Im(\frac{-1}{\sqrt{m}\sqrt{-2 + \delta_m}}) \\
&\sim \frac{1}{\sqrt{2}} \sum_{m=1}^{n-1} \frac{1}{\sqrt{m}} \sim \frac{1}{\sqrt{2}}(2\sqrt{n}) \sim \sqrt{2n},
\end{aligned}$$

where $\delta_m = (hi/\sqrt{m})^2 + \epsilon_m(hi/\sqrt{m})$. The only property of δ_m that matters for the above calculations is that $\delta_m \to 0$.

We conclude that

$$a_n(T_F) \sim \frac{\sqrt{2n}}{\pi}. \tag{5.10}$$

This theorem was published in [7] with the collaboration of Arizmendi and Wang.

5.3 Proof

In this section we prove the main theorem. But before, we need a technical lemma.

Lemma 10. *Let μ a probability measure with zero mean and unit variance. Let ν be the probability measure such that $F_\mu(z) = z - G_\nu(z)$. Let k a positive number big enough so that $\nu([-k,k]) \geq 9/10$, and define $\Gamma := \{z : y \geq 2(k+1), |x| \leq y\}$. If $z \in \Gamma$, then we have that*

$$F_\mu^{\circ n}(z) \in \Gamma \quad \forall n \in \mathbb{N}. \tag{5.11}$$

Proof. It is enough to prove that if $z \in \Gamma$, then $F_\mu(z) \in \Gamma$. First, we need to establish a technical result. Let $\Gamma_+ := \{z : y \geq 2(k+1), y-1 \leq x \leq y\}$, $\Gamma_- := \{z : y \geq 2(k+1), -y \leq x \leq -(y-1)\}$, and $\Gamma_0 := \Gamma \setminus \{\Gamma_+ \cup \Gamma_-\}$.

Claim:

i) $Re(G_\nu(z)) > 0$ if $z \in \Gamma_+$,

ii) $Re(G_\nu(z)) < 0$ if $z \in \Gamma_-$.

We only prove i). The proof of ii) is similar. Note that

$$Re(G_\nu(z)) = \int_{\mathbb{R}} \frac{(x-t)}{(x-t^2)+y^2} d\nu(t) = \frac{1}{y}\int_{\mathbb{R}} \frac{\frac{(x-t)}{y}}{(\frac{x-t}{y})^2+1} d\nu(t)$$
$$= \frac{1}{y}\int_{-k}^{k} \frac{\frac{(x-t)}{y}}{(\frac{x-t}{y})^2+1} d\nu(t) + \frac{1}{y}\int_{|t|>k} \frac{\frac{(x-t)}{y}}{(\frac{x-t}{y})^2+1} d\nu(t).$$

Now if $z \in \Gamma_+$, then

$$\frac{y-\frac{1}{2}y}{y} \leq \frac{y-1-k}{y} \leq \frac{x-t}{y} \leq \frac{y+k}{y} \leq \frac{y+\frac{1}{2}y}{y} \quad \text{for } t \in [-k,k].$$

So we have that

$$\frac{1}{2} \leq \frac{x-t}{y} \leq \frac{3}{2} \quad \text{for } t \in [-k,k].$$

It is not hard to see that

$$\inf_{\frac{1}{2}\leq\theta\leq\frac{3}{2}} \frac{\theta}{\theta^2+1} = \frac{1/2}{(1/2)^2+1} = 0.4$$

It follows that for $z \in \Gamma_+$

$$\frac{1}{y}\int_{-k}^{k} \frac{\frac{(x-t)}{y}}{(\frac{x-t}{y})^2+1} d\nu(t) \geq \frac{1}{y}\int_{-k}^{k} 0.4 d\nu(t) \geq 0.36\frac{1}{y},$$

where the last inequality is because by hypothesis $\nu([-k,k]) \geq 0.9$. We conclude that for $z \in \Gamma_+$

$$\begin{aligned} Re(G_\nu(z)) &\geq 0.36\frac{1}{y} + \frac{1}{y}\int_{|t|>k} \frac{\frac{(x-t)}{y}}{(\frac{x-t}{y})^2+1} d\nu(t) \\ &\geq 0.36\frac{1}{y} - \frac{1}{y}\int_{|t|>k} d\nu(t) \\ &\geq 0.36\frac{1}{y} - 0.1\frac{1}{y} \\ &\geq 0.26\frac{1}{y} > 0. \end{aligned}$$

So the claim is true.

We are now able to prove that if $z \in \Gamma$, then $F_\mu(z) \in \Gamma$. Take $z \in \Gamma$. There are 3 cases. Case 1: $z \in \Gamma_0$. Note that the distance of z to $\delta\Gamma$ (the boundary of Γ) is at least $\frac{1}{\sqrt{2}}$. Since $|F_\mu(z) - z| = |G_\nu(z)| \leq \frac{1}{y} < \frac{1}{2}$, then clearly $F_\mu(z) \in \Gamma$.

Case 2: $z \in \Gamma_+$. We have that $1 < y - 1 \leq x \leq y$. We want to prove that $F_\mu(z) \in \Gamma$: $Im(F_\mu(z)) \geq 2(k+1)$ and $|Re(F_\mu(z)| \leq Im(F_\mu(z))$. Since $Im(F_\mu(z)) = y + \int_{\mathbb{R}} \frac{y}{(x-t)^2+y^2} d\nu(t)$, then $Im(F_\mu(z)) \geq y \geq 2(k+1)$. Now, we have that $Re(F_\mu(z)) = x - Re(G_\nu(z))$ and $Re(G_\nu(z)) > 0$ for $z \in \Gamma_+$ (by the claim), so $Re(F_\mu(z)) = x - Re(G_\nu(z)) < x \leq y \leq Im(F_\mu(z))$. Since $x \geq y - 1 \geq 1$ and $|G_\nu(z)| \leq \frac{1}{y} < \frac{1}{2}$, then $Re(F_\mu(z)) > 1 - 1/2 = 1/2$. We conclude that $|Re(F_\mu(z)| \leq Im(F_\mu(z))$.

Case 3: $z \in \Gamma_-$. The argument is analogous to the case 2.

□

Proof of Theorem 12. By Theorem 11, it is enough to prove that $\frac{1}{\pi}\sum_{k=1}^{n} Im\frac{-1}{F^k(z)} \sim \frac{\sqrt{2n}}{\pi}$.

Let μ be a probability measure with zero mean and $0 < var(\mu) < +\infty$. By Proposition 10 there exists a measure ν such that $F_\mu(z) = z - G_\nu(z)$ with $\nu(\mathbb{R}) = var(\mu)$. Let $z_n = F_\mu^{\circ n}(c \cdot i)$ with $c = 2(k+1)$ and k is as in Lemma 10. So, by this lemma we have $|Re(z_n)| \le Im(z_n)$. Next note that the sequence z_n satisfies

$$\lim_{n\to\infty} |z_n| = \infty.$$

This is because the Denjoy-Wolff point of F_{μ_n} must be ∞. Indeed, note that $Im(F_{\mu_n}(z)) = y + \int_{\mathbb{R}} \frac{y}{(x-t)^2+y^2} d\nu(t)$ where $z = x + iy$. Since $\nu(\mathbb{R}) = var(\mu) > 0$, then $Im(F_\mu(z)) > Im(z)$ for $z \in \mathbb{C}^+$. So, F_{μ_n} has no fixed point in $\mathbb{C}^+$ and $Im(z_n)$ is increasing. Therefore the D-W point of F_μ is ∞.

We conclude that $z_n \to \infty$ non-tangentially and by Proposition 4 we have

$$z_n G_\nu(z_n) \to \nu(\mathbb{R}).$$

It follows that

$$\begin{aligned} z_{n+1}^2 - z_n^2 &= (z_{n+1} - z_n)(-\int_{\mathbb{R}} \frac{d\nu(t)}{z_n - t}) \\ &= (\frac{z_{n+1}}{z_n} + 1)(-z_n G_\nu(z_n)) \\ &\to -2\nu(\mathbb{R}) = -2var(\mu) \end{aligned}$$

as $n \to \infty$. Consequently, we obtain the convergence of averages

$$\frac{1}{n+1}\sum_{j=1}^{n-1}(z_{j+1}^2 - z_j^2) \to -2var(\mu) \quad (n \to \infty).$$

Observe now that

$$\frac{z_n^2}{n} = \frac{z_1^2}{n} + \frac{n-1}{n}\left[\sum_{j=1}^{n-1}(z_{j+1}^2 - z_j^2\right], \; n \ge 2,$$

and so we finally have

$$\frac{z_n^2}{n} = -2var(\mu).$$

The point here is that the number $-2var(\mu) < 0$. Hence we can apply the principal branch of $\sqrt{\cdot}$ to $\frac{z_n^2}{n}$, and the continuity of $\sqrt{\cdot}$ on $(-\infty, 0)$ says that

$$\lim_{n\to\infty} \frac{z_n}{\sqrt{n}} = i\sqrt{2var(\mu)}.$$

This shows that

$$Im\left(\frac{-1}{z_n}\right) \sim \sqrt{\frac{1}{2var(\mu)n}} \quad (n \to \infty)$$

and naturally,

$$\sum_{n=1}^{N} Im\left(\frac{-1}{z_n}\right) \sim \sqrt{\frac{2N}{var(\mu)}} \quad (N \to \infty).$$

□

Bibliography

[1] Jon Aaronson. Ergodic theory for inner functions of the upper half plane. *Annales de l'IHP Probabilités et statistiques*, 14(3):233–253, 1978.

[2] Jon Aaronson. *An Introduction to Infinite Ergodic Theory.* Mathematical surveys and monographs. American Mathematical Society, 1997.

[3] Roy L. Adler and Benjamin Weiss. The ergodic infinite measure preserving transformation of boole. *Israel Journal of Mathematics*, 16(3):263–278, 1973.

[4] Octavio Arizmendi and Tulio Gaxiola. Asymptotic spectral distributions of distance-k graphs of star product graphs. In *XI Symposium on Probability and Stochastic Processes*, pages 47–58. Springer, 2015.

[5] Octavio Arizmendi and Daniel Perales. A berry-esseen type theorem for finite free convolution. *Preprint.*

[6] Octavio Arizmendi and Mauricio Salazar. A berry-esseen type limit theorem for boolean convolution. *Archiv der Mathematik*, 111(1):101–111, 2018.

[7] Octavio Arizmendi, Mauricio Salazar, and Jiun-Chau Wang. Berry-esseen type estimate and return sequence for parabolic iteration in the upper half-plane. *arXiv preprint arXiv:1811.12545*, 2018.

[8] Zhi Dong Bai. Convergence rate of expected spectral distributions of large random matrices part i: Wigner matrices. In *Advances In Statistics*, pages 60–83. World Scientific, 2008.

[9] Serban Teodor Belinschi. *Complex analysis methods in noncommutative probability.* PhD thesis, Citeseer, 2005.

[10] Hari Bercovici and Dan Voiculescu. Free convolution of measures with unbounded support. *Indiana University Mathematics Journal*, 42(3):733–773, 1993.

[11] Andrew C. Berry. The accuracy of the gaussian approximation to the sum of independent variates. *Transactions of the American Mathematical Society*, 49(1):122–136, 1941.

[12] Marek Bożejko, Michael Leinert, and Roland Speicher. Convolution and limit theorems for conditionally free random variables. *Pacific Journal of Mathematics*, 175(2):357–388, 1996.

[13] Gennadii P. Chistyakov and Friedrich Götze. Limit theorems in free probability theory. i. *The Annals of Probability*, 36(1):54–90, 2008.

[14] Clive D. Cushen and Robin L. Hudson. A quantum-mechanical central limit theorem. *Journal of Applied Probability*, 8(3):454–469, 1971.

[15] Carl-Gustav Esseen. On the liapunoff limit of error in the theory of probability. *Arkiv för matematik, astronomi och fysik.*, A28(9):1–19, 1942.

[16] Uwe Franz. Monotone and boolean convolutions for non-compactly supported probability measures. *Indiana University Mathematics Journal*, 58(3):1151–1185, 2009.

[17] Narasimhan Giri and Wilhelm von Waldenfels. An algebraic version of the central limit theorem. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 42(2):129–134, 1978.

[18] Takahiro Hasebe. Monotone convolution semigroups. *Studia Mathematica*, 200(2):175–199, 2010.

[19] Takahiro Hasebe. Free infinite divisibility for beta distributions and related ones. *Electronic Journal of Probability*, 19, 2014.

[20] Robin L. Hudson. A quantum-mechanical central limit theorem for anti-commuting observables. *Journal of Applied Probability*, 10(3):502–509, 1973.

[21] Vladislav Kargin. Berry–esseen for free random variables. *Journal of Theoretical Probability*, 20(2):381–395, 2007.

[22] J.H.B. Kemperman. The ergodic behavior of a class of real transformations. In *Stochastic Processes and Related Topics I (Proc. Summer Research Inst. on Statistical Inference for Stochastic Processes, Indiana Univ., Bloomington, 1974)*, pages 249–258, 1975.

[23] Tien-Yien Li and Fritz Schweiger. The generalized boole's transformation is ergodic. *Manuscripta Mathematica*, 25(2):161–167, 1978.

[24] Tobias Mai and Roland Speicher. Operator-valued and multivariate free berry-esseen theorems. In Kösters H. Löwe M. Merkl F. Rolles S. Eichelsbacher P., Elsner G., editor, *Limit Theorems in Probability, Statistics and Number Theory*, volume 42, pages 113–140, Berlin, Heidelberg, 2013. Springer.

[25] Naofumi Muraki. Monotonic convolution and monotonic lévy-hincin formula. *preprint*, 2000.

[26] Naofumi Muraki. Monotonic independence, monotonic central limit theorem and monotonic law of small numbers. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 4(01):39–58, 2001.

[27] Michael Schürman. Quantumq-white noise and aq-central limit theorem. *Communications in Mathematical Physics*, 140(3):589–615, 1991.

[28] F. Schweiger. \tanx is ergodic. *Proceedings of the American Mathematical Society*, pages 54–56, 1978.

[29] Roland Speicher and Reza Woroudi. Boolean convolution. In *OF FIELDS INST. COMMUN.* Citeseer, 1993.

[30] Dan Voiculescu. Symmetries of some reduced free product c*-algebras. In *Operator algebras and their connections with topology and ergodic theory*, pages 556–588. Springer, 1985.

[31] Dan Voiculescu. *Operations on certain non-commutative operator-valued random variables*. Institutul National Pentru Creatie Stiintifica si Tehnica [INCREST]. Institutul de Matematica, 1986.

[32] Dan-Virgil Voiculescu. Free probability for pairs of faces i. *Communications in Mathematical Physics*, 332(3):955–980, 2014.

[33] Wilhem von Waldenfels. An algebraic central limit theorem in the anti-commuting case. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 42(2):135–140, 1978.

[34] Jiun-Chau Wang. The central limit theorem for monotone convolution with applications to free lévy processes and infinite ergodic theory. *Indiana University Mathematics Journal*, 63(2):303–327, 2014.