The Theory of Free Probability and Some Applications in the Study of Random Matrices

Pierre Yves Gaudreau Lamarre

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies in partial fulfillment of the requirements for the degree of Master of Science in Mathematics[1]

Department of Mathematics and Statistics
Faculty of Science
University of Ottawa

[1]The M.Sc. program is a joint program with Carleton University, administered by the Ottawa-Carleton Institute of Mathematics and Statistics.

2010 *Mathematics Subject Classification.* Primary 46L54; Secondary 60B20

Abstract. In this thesis, we provide a review of the theories of C^*-algebras and free probability, and we offer a survey of some applications of free probability in the study of the spectrum of random matrices. Then, we study the occurence of $*$-freeness in tensor products of families of noncommutative random variables, with possible applications in the study of the spectrum of tensor products of large random matrices.
Main Original Results: In this context, the main original results we obtained are as follows: general sufficient conditions for the $*$-freeness in tensor products (Corollary 6.11 and Proposition 6.15); a complete characterization for the occurence of $*$-freeness in the tensor product of two families that contain $*$-free variables (Theorem 6.20); and a nontrivial necessary condition for the occurence of $*$-freeness in tensor products in group algebras (Corollary 6.42).

Résumé. La présente thèse contient une introduction aux théories des C^*-algèbres et des probabilités libres, ainsi qu'un résumé de quelques applications des probabilités libres dans l'étude des valeurs propres de matrices aléatoires. Ensuite, nous étudions la présence de $*$-liberté dans les produits tensoriels de variables aléatoires non commutatives dans l'espoir d'appliquer ces avancées dans l'étude du spectre de produits tensoriels de matrices aléatoires.
Principaux Résultats Inédits: Dans ce contexte, les principaux résultats inédits que nous avons obtenus sont les suivants: des conditions suffisantes générales pour l'apparition de $*$-liberté dans les produits tensoriels (Corollaire 6.11 et Proposition 6.15); une description complète de la présence de $*$-liberté dans les produits tensoriels de deux familles de variables aléatoires qui contiennent des variables $*$-libres (Théorème 6.20); et une condition nécessaire non triviale pour l'apparition de $*$-liberté dans les produits tensoriels dans les algèbres de groupes (Corollaire 6.42).

Contents

Notation

Symbol	Meaning
$1_{\mathcal{A}}$	Unit element of the algebra $\mathcal{A}$
$\mathbf{1}_X$	Indicator function of the set X
a°	Centering of the variable a (that is, $a - \varphi(a) \cdot 1_{\mathcal{A}}$)
$\mathcal{A}^+$	Positive cone of the algebra $\mathcal{A}$
$\mathscr{B}(X)$	Borel σ-algebra on the topological space X
$\mathrm{B}(x, \varepsilon)$	Open ball centred at x with radius $\varepsilon > 0$
$\mathbb{C}\big[(a_i : i \in I)\big]$	Ring of commutative polynomials over $\mathbb{C}$ in the indeterminates $(a_i : i \in I)$
$\mathbb{C}\big\langle(a_i : i \in I)\big\rangle$	Ring of noncommutative polynomials over $\mathbb{C}$ in the indeterminates $(a_i : i \in I)$
$C^*(a_i : i \in I)$	C^*-algebra generated by the collection $(a_i : i \in I)$
$\mathrm{C}(X, Y)$	Space of continuous functions from X to Y
$\mathrm{C}_0(X, Y)$	Space of continuous functions from X to Y that vanish at infinity
$\Delta(\mathcal{A})$	Multiplicative linear functionals on the algebra $\mathcal{A}$
$\det(X)$	Determinant of the matrix $X \in \mathcal{M}_n(R)$
$\mathrm{diag}(x_1, \ldots, x_n)$	diagonal matrix $D = (D_{ij} : 1 \leqslant i, j \leqslant n)$ such that $D_{ii} = x_i$ and $D_{ij} = 0$ if $i \neq j$
δ_x	Dirac mass at the point x
$\mathbf{E}[X]$	Expected value of the random variable X
$\mathbb{F}_n$	Free group with n generators
$\mathcal{G}$	Gelfand transform
$\langle g \rangle$	Group generated by g
Id	Identity function
Id_n	$n \times n$ identity matrix
$\mathrm{Inv}(\mathcal{A})$	Invertible elements in the algebra $\mathcal{A}$
$K(\pi)$	Kreweras complement of the partition π
$\mathrm{Ker}(\pi)$	Kernel of the homomorphism π

Symbol	Meaning
$L^p(\Omega, \mathscr{A}, \mu)$	L^p-space on the measure space $(\Omega, \mathscr{A}, \mu)$
$L^{\infty-}$	Intersection of every L^p-space on a measure space $(\Omega, \mathscr{A}, \mu)$
$\mathcal{M}_{m\times n}(R)$	Set of $m \times n$ matrices with entries in R
$\mathcal{M}_n(R)$	Set of square $n \times n$ matrices with entries in R
$\mu(\sigma, \pi)$	Möbius function evaluated at the partitions σ and μ
NC(n)	Noncrossing partitions of $\{1, \ldots, n\}$
$\|\cdot\|_{\text{op}}$	Operator norm
$P(n)$	Partitions of $\{1, \ldots, n\}$
Φ_a	Functional calculus with continuous functions at the element a
R_a	Resolvent function of the operator a
$\rho(a)$	Spectral radius of the operator a
$*$-alg$(a_i : i \in I)$	$*$-algebra generated by the collection $(a_i : i \in I)$
$\mathrm{Sp}_{\mathcal{A}}(a)$	Spectrum of the operator a in the algebra $\mathcal{A}$
Span_K	Linear span over the field K
$\mathrm{supp}(\mu)$	Support of the measure μ
$\mathrm{tr}(X)$	Trace of the matrix $X \in \mathcal{M}_n(R)$
$\mathbb{U}_n$	Group of unitary $n \times n$ matrices in $\mathcal{M}_n(\mathbb{C})$
$\mathbf{Var}[X]$	Variance of the random variable X
V^*	Dual space of the vector space V
$X^\top$	Transpose of the matrix $X \in \mathcal{M}_{m\times n}(R)$
X^*	Conjugate transpose of the matrix $X \in \mathcal{M}_{m\times n}(\mathbb{C})$

Acknowledgements

The research in this thesis was supported financially by the Natural Sciences and Engineering Research Council of Canada,[1] the Ontario Student Assistance Program[2] and the University of Ottawa.[3]

I thank the Department of Mathematics at Kyoto University, especially Knorico Arisawa, for their hospitality during my stay in the fall of 2014, during which I completed a significant portion of this thesis. Thanks to the dedication of the people at Kyoto University, my visit in Japan was remarkably productive and culturally enriching.

I thank Roland Speicher and the researchers of the Arbeitsgruppe Freie Wahrscheinlichkeit at Universität des Saarlandes, as well as James Mingo and the members of the Research Group in Free Probability at Queen's University, for the warm welcome I have received, the opportunity to share my research findings with them, and insightful conversations on free probability and my research.

I express my sincere gratitude to my master's thesis supervisor, Benoit Collins, for his outstanding supervision of my thesis research, as well as previous undergraduate projects. Over the past four years, I feel that I have benefitted tremendously from his vast knowledge of mathematics, as well as his career advice. I recall he once said to me that having students under one's supervision is much like having children, and I believe that his conscientious and dedicated guidance corroborates this view.

I thank Camille Male, who, despite not having any obligation to do so, spent many hours teaching me mathematics and carefully appraising my work. I have learned a great deal from his engaging teaching style, and his meticulous review of my work has made my mathematical writing more rigorous and thorough.

Je remercie mes parents, Marie Josée et Yves, pour leur soutien indéfectible tout au cours de ma vie et de mes études universitaires. Je leur suis infiniment reconnaissant de m'avoir sensibilisé à l'importance de l'éducation si tôt dans ma vie, de m'avoir inculqué leur goût du

[1] Alexander Graham Bell Canada Graduate Scholarship (CGS-M).
[2] Ontario Graduate Scholarship.
[3] Teaching Assistantships and Ontario Graduate Scholarship.

travail soigné (« prends ton temps, fais une bonne job »), de m'avoir encouragé à poursuivre mes ambitions, et par dessus tout, de m'avoir encouragé à être résilient et à persévérer malgré les difficultés qui au moment paraissaient insurmontables. Une grande partie du succès que j'ai eu jusqu'à date leur est certainement attribuable.

I thank Jessyka Gunville from the bottom of my heart for being a faithful and loving life companion during the past years. I am grateful for her tireless support and encouragements in moments of weakness and doubts, for her understanding during the long nights I spent working on mathematics and for filling my life with joy and fulfillment. Life is not as scary with a partner, and I look forward to our next adventures.

Introduction

In very succinct terms, the theory of free probability may be described as a theoretical framework designed to analyze noncommutative random variables using a generalization of the moments in classical probability, together with the notion of **free independence**, which can be thought of as a noncommutative analog of the concept of statistical independence for complex-valued random variables.

Consider a collection $(X_i : i \in I)$ of $\mathbb{C}$-valued random variables (i.e., measurable maps from a probability space $(\Omega, \mathscr{A}, P)$ to $\big(\mathbb{C}, \mathscr{B}(\mathbb{C})\big)$, see Appendix B.1) with respective distributions $(\mu_{X_i} : i \in I)$. If one intends to study the X_i exclusively through their moments (provided they exist), that is, the quantities defined as

$$\alpha_{m,n}^{(i)} = \mathbf{E}\Big[X_i^m \overline{X_i}^n\Big] = \int z^m \overline{z}^n \, \mathrm{d}\mu_{X_i}(z), \qquad\qquad m, n \in \mathbb{N}$$

then the measure-theoretic machinery underlying the formal definition of these variables arguably becomes superfluous. A more natural framework in this context consists of defining a probability space as a pair $(\mathcal{A}, \varphi)$, where $\mathcal{A}$ is an algebra of $\mathbb{C}$-valued random variables of interest, and $\varphi : \mathcal{A} \to \mathbb{C}$ is a linear functional from which the moments of the variables in $\mathcal{A}$ can be extracted. In this particular case, the obvious choice for $\mathcal{A}$ and φ is to let $\mathcal{A}$ be a subalgebra of $L^{\infty-}$, which we define as

$$L^{\infty-} := \bigcap_{p=1}^{\infty} L^p(\Omega, \mathscr{A}, P),$$

that is, the algebra of all $\mathbb{C}$-valued random variables on $(\Omega, \mathscr{A}, P)$ with finite moments of all orders; and to let φ be the usual expected value, that is

$$\varphi(X) = \mathbf{E}[X] = \int X(\omega) \, \mathrm{d}P(\omega), \qquad\qquad X \in \mathcal{A}.$$

Requiring that $\mathcal{A}$ be an algebra ensures that for every $X \in \mathcal{A}$ and $n \in \mathbb{N}$, the functional φ can be evaluated at X^n, and thus every moment of the form $\varphi(X^n)$ becomes available. If it is also required that $\mathcal{A}$ be a $*$-algebra (see Definition 1.1 and Example 1.2), then every moment of the form $\varphi(X^m \overline{X}^n)$ (with $m, n \in \mathbb{N}$) becomes available.

This algebraic approach to probability with moments has a surprising advantage: thanks to the theories of noncommutative geometry and operator algebras (more specifically, C^*-algebras), there is a fairly straightforward way to extend the ideas used in classical probability to algebras of noncommutative random variables, and the theory developed from this extension has had very impressive applications in diverse areas of mathematics, such as the study of the eigenvalue distributions of random matrices.

In this context, the objectives and organization of the present thesis are as follows:

The first objective is to complete a survey of the main foundational works in the theory of free probability and some applications to the study of random matrices, namely:

(1) Offer a partial overview of the basic theory of C^*-algebras, upon which the noncommutative probability framework in question is built (Chapter 1).

(2) Introduce the theoretical framework used to study noncommutative random variables with **generalized moments** (Chapter 2).

(3) Revisit the notion of independence for $\mathbb{C}$-valued random variables in the context of noncommutative probability spaces and use it as an inspiration to define free independence. Then, study the properties and potential usefulness of free independence (Chapters 3 and 4).

(4) Introduce some of the current applications of free probability in the study of the spectrum of random matrices (Chapter 5).

The second objective is to provide a significant contribution to the advancement of knowledge in the subject of the thesis. All of the original research in this thesis is contained in Chapter 6, and was conducted in collaboration with/under the supervision of Professor Benoit Collins (thesis supervisor) and Camille Male (CNRS, Université Paris V). The problem that we set out to work on is the following (refer to the index for definitions, and to Section 6.1 for context):

Problem 0.1. Given $K \in \mathbb{N}$ families

$$\mathbf{a}_k = (a_{k;i} : i \in I) \subset (\mathcal{A}_k, *, \varphi_k), \qquad 1 \leqslant k \leqslant K$$

of noncommutative random variables (where $(\mathcal{A}_k, *, \varphi_k)$ are $*$-probability spaces), characterize the $*$-freeness of the elements in the tensor product collection

$$\otimes_k \mathbf{a}_k = (\otimes_k a_{k;i} : i \in I) \subset (\otimes_k \mathcal{A}_k, *, \otimes_k \varphi_k)$$

in terms of our knowledge of the behaviour of the individual factor collections $\mathbf{a}_k$.[4]

The main results we obtained concerning the above problem are the following:

(1) Sufficient conditions for the $*$-freeness in tensor products (Corollary 6.11 and Proposition 6.15) that are general enough to answer questions about the asymptotic freeness of tensor products of unitary random matrices (Question 6.1 and Proposition 6.12);

(2) a complete characterization of a special case of Problem 0.1 (Theorem 6.20); and

(3) a necessary condition for a special case of Problem 0.1 (Corollary 6.42) achieved with some insights in the theory of free groups, which we believe could be of independent interest and lead to interesting further research (Lemma 6.40 and Proposition 6.41).

[4] We commit the following slight abuse of notation: $\otimes_k \mathbf{a}_k = \otimes_{k \leqslant K} \mathbf{a}_k = \mathbf{a}_1 \otimes \cdots \otimes \mathbf{a}_K$.

Part I. Noncommutative Probability Theory

Chapter 1

C^*-algebras

Most of the concepts and proofs featured in this chapter were covered in less detail in the course *Topics in Analysis: Introduction to C^*-algebras*, which was taught by professor David Handelman at the Department of Mathematics and Statistics of the University of Ottawa during the autumn of 2013. Other notable sources include **[Da96]**, **[Ka09]**, and **[Di77]**.

1.1. Algebras over a Field

Let F be a field. An **algebra** $\mathcal{A} = (V, \cdot)$ over the field F consists of a vector space V over F together with a bilinear product on V:

$$\begin{array}{ccc} V \times V & \to & V \\ (a, b) & \mapsto & a \cdot b. \end{array}$$

An algebra $\mathcal{A}$ is said to be **unital** if there exists a vector $1_{\mathcal{A}}$ that acts as a unit with respect to the bilinear product, that is, for every vector a, one has $1_{\mathcal{A}} \cdot a = a = a \cdot 1_{\mathcal{A}}$. Unless otherwise mentioned, every algebra considered in this thesis will be over the field of complex numbers $\mathbb{C}$.

Given an algebra $\mathcal{A} = (V, \cdot)$, the elements of the underlying vector space $a \in V$ are usually referred to as elements of the algebra itself $a \in \mathcal{A}$. Furthermore, to alleviate notation, for any $a, b \in \mathcal{A}$, the product $a \cdot b$ is written as ab.

Definition 1.1. A **$*$-algebra** $(\mathcal{A}, *)$ consists of an algebra $\mathcal{A}$ equipped with an operation

$$\begin{array}{ccc} \mathcal{A} & \to & \mathcal{A} \\ a & \mapsto & a^* \end{array}$$

satisfying the following conditions:

(1) for every $a \in \mathcal{A}$, one has $(a^*)^* = a$ ($*$ is an involution);

(2) for every $a, b \in \mathcal{A}$, one has $(ab)^* = b^*a^*$ ($*$ is an antihomomorphism with respect to the bilinear product); and

(3) for every $k \in \mathbb{C}, a, b \in \mathcal{A}$, one has $(ka)^* = \overline{k}a^*$ and $(a + b)^* = a^* + b^*$ ($*$ is antilinear).

Example 1.2. As briefly mentioned in the introduction of this thesis, given a probability space $(\Omega, \mathscr{A}, P)$, the set

$$L^{\infty-} = \bigcap_{p=1}^{\infty} L^p(\Omega, \mathscr{A}, P)$$

equipped with the usual function operations and the $*$-operation of complex conjugate, that is, for every $X, Y \in L^{\infty-}$, $k \in \mathbb{C}$ and $\omega \in \Omega$, one has

(1) $(X+Y)(\omega) = X(\omega) + Y(\omega)$;
(2) $(XY)(\omega) = X(\omega)Y(\omega)$;
(3) $(kX)(\omega) = k \cdot X(\omega)$; and
(4) $\overline{X}(\omega) = \overline{X(\omega)}$,

is a $*$-algebra.

Example 1.3. The example that is arguably at the origin of the definition of a $*$-algebra is the following: Let $\big(H, \langle \cdot, \cdot \rangle\big)$ be a Hilbert space, and let $B(H)$ be the space of bounded linear operators on H. If $B(H)$ is equipped with the (unique) $*$-operation that satisfies the relation

$$\langle Ta, b\rangle = \langle a, T^*b\rangle, \qquad\qquad T \in B(H),\ a, b \in H,$$

(that is, the adjoint operator) then $\big(B(H), *\big)$ a $*$-algebra.

Remark 1.4. As one would expect, unit vectors are invariant with respect to the $*$-operation. Indeed, if $(\mathcal{A}, *)$ is a unital $*$-algebra, then $1_{\mathcal{A}}^* = 1_{\mathcal{A}}$, as for any given $a \in \mathcal{A}$, one has $1_{\mathcal{A}}^* a = (a^* 1_{\mathcal{A}})^* = (a^*)^* = a$, and similarly, $a 1_{\mathcal{A}}^* = a$.

Let $(\mathcal{A}, *)$ be a $*$-algebra. An element $a \in \mathcal{A}$, is said to be **self-adjoint** if $a = a^*$, **normal** if $aa^* = a^*a$, and **unitary** if $aa^* = a^*a = 1_{\mathcal{A}}$. This vocabulary is reminiscent of that which is commonly used to describe properties of matrices. As one would expect, the space $\mathcal{M}_n(\mathbb{C})$ of $n \times n$ matrices with entries in $\mathbb{C}$ (equipped with the usual matrix operations and conjugate transpose $*$) is a $*$-algebra (as a special case of Example 1.3).

Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be algebras. A map $\Phi : \mathcal{A}_1 \to \mathcal{A}_2$ is called an **algebra homomorphism** if for every $a, b \in \mathcal{A}_1$ and $k \in \mathbb{C}$, one has

(1) $\Phi(a+b) = \Phi(a) + \Phi(b)$;
(2) $\Phi(ab) = \Phi(a)\Phi(b)$; and
(3) $\Phi(ka) = k\Phi(a)$.

Let $(\mathcal{A}_1, *_1)$ and $(\mathcal{A}_2, *_2)$ be $*$-algebras. An algebra homomorphism Φ from $\mathcal{A}_1$ to $\mathcal{A}_2$ is called a **$*$-homomorphism** if $\Phi(a^{*_1}) = \Phi(a)^{*_2}$ for every $a \in \mathcal{A}_1$.

Proposition 1.5. *Let $(\mathcal{A}_1, *_1)$, $(\mathcal{A}_2, *_2)$ and $(\mathcal{A}_3, *_3)$ be $*$-algebras. If the functions $\Psi : \mathcal{A}_1 \to \mathcal{A}_2$ and $\Phi : \mathcal{A}_2 \to \mathcal{A}_3$ are $*$-isomorphisms (bijective $*$-homomorphisms), then $\Psi^{-1} : \mathcal{A}_2 \to \mathcal{A}_1$ and $\Phi \circ \Psi : \mathcal{A}_1 \to \mathcal{A}_3$ are $*$-isomorphisms.*

Proof. The proof that Ψ^{-1} and $\Phi \circ \Psi$ are algebra isomorphisms is identical to the proof that inverses and compositions of ring isomorphisms are also ring isomorphisms.[1] It now remains to prove that they are $*$-homomorphisms. Let $b \in \mathcal{A}_2$ and $a \in \mathcal{A}_1$ be arbitrary. Since Ψ is an

[1] See for instance [**Be99**] Proposition 1.2.2.

$*$-isomorphism, there exists a unique $b_0 \in \mathcal{A}_1$ such that $\Psi(b_0) = b$, and thus $\Psi(b_0^{*_1}) = \Psi(b_0)^{*_2} = b^{*_2}$. Therefore, $\Psi^{-1}(b) = b_0$ and $\Psi^{-1}(b^{*_2}) = b_0^{*_1}$, hence Ψ^{-1} is a $*$-isomorphism. Since Ψ and Φ are $*$-homomorphisms,

$$\Phi(\Psi(a^{*_1})) = \Phi(\Psi(a)^{*_2}) = \Phi(\Psi(a))^{*_3},$$

which concludes the proof that $\Phi \circ \Psi$ is a $*$-isomorphism. □

Remark 1.6. Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be algebras and $0_{\mathcal{A}_1}$ and $0_{\mathcal{A}_2}$ be the zero vectors in $\mathcal{A}_1$ and $\mathcal{A}_2$ respectively. If $\Phi : \mathcal{A}_1 \to \mathcal{A}_2$ is an algebra homomorphism, then it is linear, hence $\Phi(0_{\mathcal{A}_1}) = 0_{\mathcal{A}_2}$.

A **Banach algebra** $(\mathcal{A}, \|\cdot\|)$ is an algebra $\mathcal{A} = (V, \cdot)$ equipped with a norm $\|\cdot\|$ such that

(1) $(V, \|\cdot\|)$ is a Banach space;[2] and
(2) for every $a, b \in \mathcal{A}$, one has $\|ab\| \leqslant \|a\|\|b\|$ (or, (almost) equivalently,[3] the product operation $(a, b) \mapsto ab$ is continuous in both coordinates).

Unless otherwise stated, a Banach algebra is equipped with the metric topology induced by its norm $\|\cdot\|$.

Definition 1.7. Let $\mathcal{A}$ be an algebra, $\|\cdot\|$ be a norm on $\mathcal{A}$, and $*$ be an operator on $\mathcal{A}$ such that

(1) $(\mathcal{A}, \|\cdot\|)$ is a Banach algebra;
(2) $(\mathcal{A}, *)$ is a $*$-algebra; and
(3) for every $a \in \mathcal{A}$, one has $\|aa^*\| = \|a\|^2$.

Then, $(\mathcal{A}, *, \|\cdot\|)$ is called a ***C**-algebra**.

Remark 1.8. Let $(\mathcal{A}, *, \|\cdot\|)$ be a C^*-algebra that is not unital. Then, if one defines $\mathcal{A}_{\mathrm{u}} = \mathcal{A} \oplus \mathbb{C}$ together with the operations and norm defined as

$$\begin{aligned}
(a, k)(b, t) &\coloneqq (ab + ta + kb, kt), && a, b \in \mathcal{A},\ k, t \in \mathbb{C} \\
(a, k)^* &\coloneqq (a^*, \overline{k}), && a \in \mathcal{A},\ k \in \mathbb{C} \\
\|(a, k)\| &\coloneqq \sup\big\{\|ab + kb\| : \|b\| = 1\big\}, && a \in \mathcal{A},\ k \in \mathbb{C}
\end{aligned}$$

(see **[Da96]** Proposition I.1.3 for the details), then $\mathcal{A}_{\mathrm{u}}$ is a unital C^*-algebra (whose unit is given by $0_{\mathcal{A}} \oplus 1$), and $\mathcal{A} \oplus 0$ (which is trivially isomorphic to $\mathcal{A}$ as a C^*-algebra) is a maximal ideal of $\mathcal{A}_{\mathrm{u}}$ of codimension 1. The space $(\mathcal{A}_{\mathrm{u}}, *, \|\cdot\|)$ is called the **unitalization** of $(\mathcal{A}, *, \|\cdot\|)$.

Example 1.9. Let $\big(H, \langle\cdot,\cdot\rangle\big)$ be a Hilbert space, let $B(H)$ be the space of bounded linear operators on H, and define the adjoint $*$-operation on $B(H)$ as in Example 1.3. If one defines the **operator norm** as

$$\|T\|_{\mathrm{op}} = \sup\left\{\left(\frac{\langle Ta, Ta\rangle}{\langle a, a\rangle}\right)^{1/2} : a \in H,\ a \neq 0\right\}, \qquad T \in B(H),$$

[2]Recall that a Banach space is defined as a normed vector space that is complete (with respect to the topology induced by the norm).

[3]The fact that $\|ab\| \leqslant \|a\|\|b\|$ implies continuity is obvious. As for the converse, it can be shown that if the product operation in a complete normed algebra $(\mathcal{A}, \|\cdot\|)$ is continuous, then there exists a norm $\|\cdot\|_B$ such that $C\|\cdot\| = \|\cdot\|_B$ for some constant C and $\|ab\|_B \leqslant \|a\|_B\|b\|_B$ for all a and b (see **[EMT04]** Theorem 10.1.1).

then $\big(B(H), *, \|\cdot\|_{\text{op}}\big)$ is a C^*-algebra.[4] In particular, for any $n \in \mathbb{N}$, the space $\mathcal{M}_n(\mathbb{C})$ of $n \times n$ matrices with entries in $\mathbb{C}$ equipped with the conjugate transpose $*$ and the matrix operator norm $\|\cdot\|_{\text{op}}$ is a C^*-algebra.

Example 1.10. Let (X, τ) be a locally compact Hausdorff topological space. Consider the space $\mathrm{C}_0(X, \mathbb{C})$ of continuous functions $f : X \to \mathbb{C}$ that *vanish at infinity*, that is, for every $\varepsilon > 0$, there exists a compact subset $K \subset X$ such that if $x \notin K$, then $|f(x)| < \varepsilon$. If $\mathrm{C}_0(X, \mathbb{C})$ is equipped with the usual function operations and the complex conjugate, then one could easily show that this forms a $*$-algebra over $\mathbb{C}$. If $\mathrm{C}_0(X, \mathbb{C})$ is also equipped with the **supremum norm** $\|\cdot\|_\infty$ (i.e., $\|f\|_\infty = \sup\{|f(x)| : x \in X\}$), then it forms a C^*-algebra. Note that, if X is not compact, then $\big(\mathrm{C}_0(X, \mathbb{C}), \bar{\cdot}, \|\cdot\|_\infty\big)$ is not unital, since the indicator function $\mathbf{1}_X$ does not vanish at infinity.

Example 1.11. Let (X, τ) be a compact Hausdorff topological space. Then, the space $\mathrm{C}_0(X, \mathbb{C})$ is actually equal to $\mathrm{C}(X, \mathbb{C})$ (the space of continuous functions from X to $\mathbb{C}$), since every function from X to $\mathbb{C}$ vanishes at infinity. Thus, if $\mathrm{C}(X, \mathbb{C})$ is equipped with the $*$-operation of complex conjugate and the supremum norm, it is a unital C^*-algebra.

Remark 1.12. If $(\mathcal{A}, \|\cdot\|)$ is a unital Banach algebra, then $\|1_\mathcal{A}\| \geqslant 1$. Indeed, this clearly follows from the inequality $\|1_\mathcal{A}\| = \|1_\mathcal{A} 1_\mathcal{A}\| \leqslant \|1_\mathcal{A}\|\|1_\mathcal{A}\|$. Moreover, If $(\mathcal{A}, *, \|\cdot\|)$ is a unital C^*-algebra, then $\|1_\mathcal{A}\| = 1$. This is a consequence of Remark 1.4: $\|1_\mathcal{A}\| = \|1_\mathcal{A} 1_\mathcal{A}\| = \|1_\mathcal{A} 1_\mathcal{A}^*\| = \|1_\mathcal{A}\|^2$.

Remark 1.13. If $(\mathcal{A}, *, \|\cdot\|)$ is a C^*-algebra and $a \in \mathcal{A}$ is self-adjoint, then for every $n \in \mathbb{N}$, $\|a^{2^n}\| = \|a\|^{2^n}$. This can easily be shown by induction on n: If $n = 1$, then, according to the definition of a C^*-algebra, one has $\|a\|^2 = \|a^* a\| = \|a^2\|$. Similarly, if the result holds for some $n \in \mathbb{N}$, then

$$\|a\|^{2^{n+1}} = (\|a\|^{2^n})^2 = (\|a^{2^n}\|)^2 = \|(a^{2^n})^* a^{2^n}\| = \|a^{2^{n+1}}\|.$$

Definition 1.14. Let $(\mathcal{A}, *)$ be a $*$-algebra and $(a_i : i \in I)$ be a collection of elements of $\mathcal{A}$. The ***-algebra generated by $(a_i : i \in I)$***, which is denoted by $*\text{-alg}(a_i : i \in I)$, is defined as the linear span

$$\mathrm{Span}_\mathbb{C}\Big\{a_{i(1)}^{n(1)} \cdots a_{i(p)}^{n(p)} : p \in \mathbb{N},\ i(1), ..., i(p) \in I,\ n(1), ..., n(p) \in \{1, *\}\Big\}.$$

Let $(\mathcal{A}, *, \|\cdot\|)$ be a C^*-algebra and $(a_i : i \in I)$ be a collection of elements of $\mathcal{A}$. The ***C*-algebra generated by $(a_i : i \in I)$***, which is denoted by $C^*(a_i : i \in I)$, is defined as the closure of $*\text{-alg}(a_i : i \in I)$ in $(\mathcal{A}, \|\cdot\|)$.

Remark 1.15. It is necessary to define $C^*(a_i : i \in I)$ as the closure of $*\text{-alg}(a_i : i \in I)$ rather than $*\text{-alg}(a_i : i \in I)$ itself: Consider the field $\mathbb{C}$ of complex numbers equipped with the complex modulus $|\cdot|$ and the complex conjugate $\bar{\cdot}$. Then, $(\mathbb{C}, \bar{\cdot}, |\cdot|)$ is a C^*-algebra. Let $(a_i : i \in \mathbb{Q}^2) \subset \mathbb{C}$ be defined as $a_{(p,q)} = p + \mathrm{i}q$ for every $(p, q) \in \mathbb{Q}^2$. Then, $*\text{-alg}(a_i : i \in \mathbb{Q}^2) = \mathbb{Q} + \mathrm{i}\mathbb{Q}$ is not a C^*-algebra since it is not complete.

The following proposition motivates the appellation C^*-algebra *generated by a collection*. Indeed, given a collection of objects $(a_i : i \in I)$ in a C^*-algebra $(\mathcal{A}, *, \|\cdot\|)$, $C^*(a_i : i \in I)$ is the *smallest* C^*-subalgebra of $\mathcal{A}$ that contains $(a_i : i \in I)$.

[4] See Chapter 4 of [**EMT04**] for the details.

Proposition 1.16. *Let $(\mathcal{A}, *, \|\cdot\|)$ be a C^*-algebra and $(a_i : i \in I)$ be a collection of elements in $\mathcal{A}$. If $\mathcal{B} \subset \mathcal{A}$ is a C^*-subalgebra of $\mathcal{A}$ that contains every a_i, then $\mathcal{B}$ contains $C^*(a_i : i \in I)$.*

Proof. Since $\mathcal{B}$ is a C^*-algebra, it is closed under the addition, the product, the scalar product and the $*$ operation in $\mathcal{A}$. Thus, given that $\mathcal{B}$ contains every a_i, it is clear that $*$-alg$(a_i : i \in I) \subset \mathcal{B}$. Since $\mathcal{B}$ is itself a C^*-algebra, it is closed, hence $C^*(a_i : i \in I) \subset \mathcal{B}$. □

One of the interesting features of C^*-algebras generated by a collection of elements is that they *preserve commutativity*. More precisely:

Proposition 1.17. *Let $(\mathcal{A}, *, \|\cdot\|)$ be a unital C^*-algebra and $(a_i : i \in I)$ be a collection of normal elements of $\mathcal{A}$ that commute, that is, for every $i, j \in I$, one has $a_i a_j = a_j a_i$. Then, $C^*\big(1_{\mathcal{A}}, (a_i : i \in I)\big)$ is a commutative unital C^*-algebra.*

Proof. Consider two finite collections of objects $(b_i : i \leqslant n)$ and $(c_j : j \leqslant m)$ in $\mathcal{A}$ such that if $b, c \in \{b_1, ..., b_n, c_1, ..., c_m\}$, then $bc = cb$. Then, for any scalars $s_1, ..., s_n, t_1, ..., t_m \in \mathbb{C}$, it clearly is the case that

$$\left(\sum_{i=1}^{n} s_i b_i\right)\left(\sum_{j=1}^{m} t_j c_j\right) = \left(\sum_{j=1}^{m} t_j c_j\right)\left(\sum_{i=1}^{n} s_i b_i\right).$$

Consequently, to prove that the span[5]

$$\begin{aligned} &*\text{-alg}\big(1_{\mathcal{A}}, (a_i : i \in I)\big) \\ =&\text{Span}_{\mathbb{C}}\left\{a_{i(1)}^{n(1)} \cdots a_{i(p)}^{n(p)} : p \in \mathbb{N},\ i(1), ..., i(p) \in I,\ n(1), ..., n(p) \in \{0, 1, *\}\right\} \end{aligned}$$

is commutative, it suffices to show that for every choice of $p, q \in \mathbb{N}$,

$$n(1), ..., n(p), m(1), ..., m(q) \in \{0, 1, *\},$$

and $i(1), ..., i(p), j(1), ..., j(q) \in I$, one has

$$\left(a_{i(1)}^{n(1)} \cdots a_{i(p)}^{n(p)}\right)\left(a_{j(1)}^{m(1)} \cdots a_{j(q)}^{m(q)}\right) = \left(a_{j(1)}^{m(1)} \cdots a_{j(q)}^{m(q)}\right)\left(a_{i(1)}^{n(1)} \cdots a_{i(p)}^{n(p)}\right).$$

The above equality clearly follows from the normality of the a_i and the fact that the a_i commute, hence $*$-alg$\big(1_{\mathcal{A}}, (a_i : i \in I)\big)$ is commutative.

Let $b, c \in C^*\big(1_{\mathcal{A}}, (a : i \in I)\big)$ be arbitrary. Then, there exists sequences $\{b_n : n \in \mathbb{N}\}$ and $\{c_n : n \in \mathbb{N}\}$ in $*$-alg$\big(1_{\mathcal{A}}, (a_i : i \in I)\big)$ such that $\|b_n - b\|$ and $\|c_n - c\|$ converge to zero. Therefore,

$$\begin{aligned} bc &= \left(\lim_{n\to\infty} b_n\right)\left(\lim_{n\to\infty} c_n\right) \\ &= \lim_{n\to\infty} b_n c_n, && \text{(both limits exist)} \\ &= \lim_{n\to\infty} c_n b_n, && (\forall n,\ b_n, c_n \in *\text{-alg}\big(1_{\mathcal{A}}, (a_i : i \in I)\big)) \\ &= \left(\lim_{n\to\infty} c_n\right)\left(\lim_{n\to\infty} b_n\right) \\ &= cb, \end{aligned}$$

concluding the proof that $C^*\big(1_{\mathcal{A}}, (a_i : i \in I)\big)$ is commutative. □

[5]We use the convention $b^0 = 1_{\mathcal{A}}$ for every nonzero $b \in \mathcal{A}$.

1.2. Spectrum and Resolvent Set

An element a in a unital algebra $\mathcal{A}$ is said to be **invertible** if there exists $a^{-1} \in \mathcal{A}$ such that $a^{-1}a = aa^{-1} = 1_{\mathcal{A}}$. $\mathrm{Inv}(\mathcal{A})$ is used to denote the set of invertible elements in $\mathcal{A}$. For any $a \in \mathcal{A}$, the **spectrum** of a is defined as the set

$$\mathrm{Sp}_{\mathcal{A}}(a) \coloneqq \{\lambda \in \mathbb{C} : \lambda \cdot 1_{\mathcal{A}} - a \notin \mathrm{Inv}(\mathcal{A})\},$$

the **resolvent set** of a is defined as $\mathbb{C} \setminus \mathrm{Sp}_{\mathcal{A}}(a)$, and the **resolvent function** at a is defined as

$$\begin{array}{rccc} \mathrm{R}_a : & \mathbb{C} \setminus \mathrm{Sp}_{\mathcal{A}}(a) & \to & \mathrm{Inv}(\mathcal{A}) \\ & \lambda & \mapsto & (\lambda \cdot 1_{\mathcal{A}} - a)^{-1}. \end{array}$$

The **spectral radius** of a is defined as

$$\rho(a) = \sup\{|\lambda| : \lambda \in \mathrm{Sp}_{\mathcal{A}}(a)\}.$$

Example 1.18. Let $(\Omega, \mathscr{A}, P)$ be a probability space and let $L^{\infty-}$ be defined as in Example 1.2. Then, for every $X \in L^{\infty-}$ and $\lambda \in \mathbb{C}$, the function $X - \lambda \mathbf{1}_{\Omega}$ is invertible if and only if λ is not in the image of X. Thus, in this particular case, one has

$$\mathrm{Sp}_{L^{\infty-}}(X) = \big\{X(\omega) : \omega \in \Omega\big\}.$$

Remark 1.19. Let $\mathcal{A}$ be a unital algebra, $k \in \mathbb{C}$ be nonzero, and $a \in \mathcal{A}$ be arbitrary. Then, one can easily show that $\lambda \in \mathrm{Sp}_{\mathcal{A}}(ka)$ if and only if $k^{-1}\lambda \in \mathrm{Sp}_{\mathcal{A}}(a)$. Consequently, $\mathrm{Sp}_{\mathcal{A}}(ka) = k\mathrm{Sp}_{\mathcal{A}}(a)$.

Remark 1.20. Let $\mathcal{A}$ and $\mathcal{B}$ be unital algebras, and let $\Phi : \mathcal{A} \to \mathcal{B}$ be an algebra homomorphism. Then, for every $a \in \mathcal{A}$, $\mathrm{Sp}_{\mathcal{B}}\big(\Phi(a)\big) \subset \mathrm{Sp}_{\mathcal{A}}(a)$. Indeed, if $(\lambda \cdot 1_{\mathcal{A}} - a)b = 1_{\mathcal{A}}$, then $\big(\lambda 1_{\mathcal{B}} - \Phi(a)\big)\Phi(b) = 1_{\mathcal{B}}$.

The fundamental properties of the spectrum, from which many important facts about algebras will be derived, are summarized in the following theorem:

Theorem 1.21. *Let $(\mathcal{A}, \|\cdot\|)$ be a unital Banach algebra and $a \in \mathcal{A}$ be arbitrary.*

(1) $\mathrm{Sp}_{\mathcal{A}}(a) \subset \{\lambda \in \mathbb{C} : |\lambda| \leqslant \|a\|\}$*;*

(2) $\mathrm{Sp}_{\mathcal{A}}(a)$ *is compact; and*

(3) $\mathrm{Sp}_{\mathcal{A}}(a)$ *is nonempty.*

In order to show Theorem 1.21, a few preliminary results must first be established.

Lemma 1.22. *Let $(\mathcal{A}, \|\cdot\|)$ be a unital Banach algebra. Then,*

(1) if $a \in \mathcal{A}$ is such that $\|a\| < 1$, then $1_{\mathcal{A}} - a \in \mathrm{Inv}(\mathcal{A})$, and, in particular,

$$(1_{\mathcal{A}} - a)^{-1} = \sum_{n=0}^{\infty} a^n;$$

(2) the set $\mathrm{Inv}(\mathcal{A})$ *is an open subset of $(\mathcal{A}, \|\cdot\|)$; and*

(3) the resolvent set $\mathbb{C} \setminus \mathrm{Sp}_{\mathcal{A}}(a)$ *of a is an open subset of* $\mathbb{C}$*.*

Proof. (1). Since $(\mathcal{A}, \|\cdot\|)$ is a Banach algebra, the metric induced by the norm $\|\cdot\|$ is complete. Thus, if a series is absolutely convergent, it converges to a limit in $\mathcal{A}$.[6] Let $a \in \mathcal{A}$ be such that $\|a\| < 1$. Consider the sequence $\{a^n : n \in \mathbb{N} \cup \{0\}\}$.[7] Since $(\mathcal{A}, \|\cdot\|)$ is a Banach algebra, then $\|a^n\| \leqslant \|a\|^n$ for every integer $n \geqslant 0$. Consequently, since $\|a\| < 1$, the series $1_{\mathcal{A}} + a + a^2 + \cdots$ is absolutely convergent and thus converges to a limit $b \in \mathcal{A}$. For every integer $n \geqslant 0$, let $b_n = 1_{\mathcal{A}} + a + \cdots + a^n$. Clearly, $b_n \to b$. Moreover, For every $n \in \mathbb{N}$, one has

$$(1_{\mathcal{A}} - a)b_n = b_n(1_{\mathcal{A}} - a) = \sum_{j=0}^{n} a^j - \sum_{j=0}^{n} a^{j+1} = 1_{\mathcal{A}} - a^{n+1}.$$

Since a^n clearly converges to $0_{\mathcal{A}}$, taking the limit $n \to \infty$ yields

$$(1_{\mathcal{A}} - a)b = b(1_{\mathcal{A}} - a) = 1_{\mathcal{A}},$$

which implies that $1_{\mathcal{A}} - a$ is invertible and that

$$(1_{\mathcal{A}} - a)^{-1} = \sum_{n=0}^{\infty} a^n.$$

(2). Let $a \in \mathrm{Inv}(\mathcal{A})$ be arbitrary. It suffices to prove that there exists a radius $r > 0$ such that $\mathrm{B}(a, r) \subset \mathrm{Inv}(\mathcal{A})$.[8] Let $b \in \mathrm{B}\left(a, \frac{1}{\|a^{-1}\|}\right)$. Then,

$$\|1_{\mathcal{A}} - a^{-1}b\| = \|a^{-1}(a - b)\| \leqslant \|a^{-1}\|\|a - b\| < \frac{\|a^{-1}\|}{\|a^{-1}\|} = 1.$$

From part (1) of this lemma, this implies that $1_{\mathcal{A}} - (1_{\mathcal{A}} - a^{-1}b) = a^{-1}b$ is invertible, and thus $a(a^{-1}b) = b$ is also invertible.[9]

(3). Define the function $f_a : \mathbb{C} \to \mathcal{A}$ as follows:

$$\begin{array}{rccc} f_a: & \mathbb{C} & \to & \mathcal{A} \\ & \lambda & \mapsto & \lambda \cdot 1_{\mathcal{A}} - a. \end{array}$$

This function is clearly continuous. According to part (2) of this lemma, $\mathrm{Inv}(\mathcal{A})$ is open, and thus $f_a^{-1}(\mathrm{Inv}(\mathcal{A})) = \mathbb{C} \setminus \mathrm{Sp}_{\mathcal{A}}(a)$ is open. □

Proposition 1.23. *Let $(\mathcal{A}, \|\cdot\|)$ be a unital Banach algebra and $a \in \mathcal{A}$ be arbitrary. Then, the resolvent function R_a at a is analytic on the open set $\mathbb{C} \setminus \mathrm{Sp}_{\mathcal{A}}(a)$, which implies in particular that R_a is continuous on the resolvent set.*[10]

Proof. Let $\lambda_0 \in \mathbb{C} \setminus \mathrm{Sp}_{\mathcal{A}}(a)$ and $\lambda \in \mathrm{B}\left(\lambda_0, \|\mathrm{R}_a(\lambda_0)\|^{-1}\right)$ be arbitrary. Then,

$$\| - (\lambda - \lambda_0)\mathrm{R}_a(\lambda_0)\| = \|(\lambda - \lambda_0)\mathrm{R}_a(\lambda_0)\| = |(\lambda - \lambda_0)| \cdot \|\mathrm{R}_a(\lambda_0)\| < 1,$$

---

[6] See Proposition B.11.

[7] For each $n \in \mathbb{N}$, let $a^n = \underbrace{a \cdots a}_{n \text{ times}}$.

[8] $\mathrm{B}(x, \varepsilon) = \{y : \|x - y\| < \varepsilon\}$ denotes the **open ball** of radius ε at x.

[9] Given two invertible elements $a_1, a_2 \in \mathcal{A}$ with respective inverses $a_1^{-1}, a_2^{-1} \in \mathcal{A}$, their product $a_1a_2 \in \mathcal{A}$ is clearly invertible by $a_2^{-1}a_1^{-1}$. Thus, if $a_1, a_2 \in \mathrm{Inv}(\mathcal{A})$, then $a_1a_2 \in \mathrm{Inv}(\mathcal{A})$.

[10] Every analytic function is continuous (see section B.4).

which, according to part (1) of Lemma 1.22, implies that $1_{\mathcal{A}} + (\lambda - \lambda_0)\mathrm{R}_a(\lambda_0)$ is invertible and that

$$(1_{\mathcal{A}} + (\lambda - \lambda_0)\mathrm{R}_a(\lambda_0))^{-1} = \sum_{n=0}^{\infty}(-1)^n(\lambda - \lambda_0)^n\mathrm{R}_a(\lambda_0)^n. \tag{1.1}$$

Notice that

$$\begin{aligned}\lambda \cdot 1_{\mathcal{A}} - a &= (\lambda - \lambda_0)1_{\mathcal{A}} + (\lambda_0 1_{\mathcal{A}} - a) \\ &= (\lambda_0 1_{\mathcal{A}} - a)((\lambda - \lambda_0)(\lambda_0 1_{\mathcal{A}} - a)^{-1} + 1_{\mathcal{A}}), \qquad (\text{since } \lambda_0 \notin \mathrm{Sp}_{\mathcal{A}}(a)) \\ &= (\lambda_0 1_{\mathcal{A}} - a)(1_{\mathcal{A}} + (\lambda - \lambda_0)\mathrm{R}_a(\lambda_0)). \end{aligned} \tag{1.2}$$

According to part (3) of Lemma 1.22, the resolvent set is open. Therefore, since $\lambda_0 \in \mathbb{C} \setminus \mathrm{Sp}_{\mathcal{A}}(a)$, there exists $\varepsilon > 0$ such that $\mathrm{B}(\lambda_0, \varepsilon) \subset \mathbb{C} \setminus \mathrm{Sp}_{\mathcal{A}}(a)$. Let $\eta = \min\{\varepsilon, \|\mathrm{R}_a(\lambda_0)\|^{-1}\}$. If $\lambda \in \mathrm{B}(\lambda_0, \eta)$, then $\lambda \notin \mathrm{Sp}_{\mathcal{A}}(a)$, that is, $\lambda \cdot 1_{\mathcal{A}} - a$ is invertible, and thus by taking inverses on both sides of equation (1.2), it follows from equation (1.1) that

$$\mathrm{R}_a(\lambda) = (1_{\mathcal{A}} + (\lambda - \lambda_0)\mathrm{R}_a(\lambda_0))^{-1}\mathrm{R}_a(\lambda_0) = \sum_{n=0}^{\infty}(-1)^n(\lambda - \lambda_0)^n\mathrm{R}_a(\lambda_0)^{n+1}.$$

Since $\lambda_0 \in \mathbb{C}$ and $\lambda \in \mathrm{B}(\lambda_0, \eta)$ were arbitrary, it follows that R_a is analytic on the resolvent set of a. □

All the necessary ingredients are now in place to prove Theorem 1.21.

Proof of Theorem 1.21. (1). Let $\lambda \in \mathrm{Sp}_{\mathcal{A}}(a)$ be such that $|\lambda| > \|a\|$. Then, $\|\lambda^{-1}a\| < 1$, which, according to part (1) of Lemma 1.22, implies that $1_{\mathcal{A}} - \lambda^{-1}a$ is invertible. Given that $1_{\mathcal{A}} - \lambda^{-1}a = \lambda^{-1}(\lambda \cdot 1_{\mathcal{A}} - a)$, then $\lambda \cdot 1_{\mathcal{A}} - a$ is also invertible, which contradicts that $\lambda \in \mathrm{Sp}_{\mathcal{A}}(a)$.

(2). Since $\mathrm{Sp}_{\mathcal{A}}(a)$ is a subset of $\mathbb{C}$, it is compact if and only if it is closed and bounded. Part (1) of this theorem established that $\mathrm{Sp}_{\mathcal{A}}(a)$ is bounded, and, according to part (3) of Lemma 1.22, $\mathbb{C} \setminus \mathrm{Sp}_{\mathcal{A}}(a)$ is open, hence $\mathrm{Sp}_{\mathcal{A}}(a)$ is closed.

(3). Suppose by contradiction that $\mathrm{Sp}_{\mathcal{A}}(a)$ is empty. Then, the resolvent function R_a at a is defined for every $\lambda \in \mathbb{C}$. If

$$\lim_{\lambda \to \infty} \mathrm{R}_a(\lambda) = 0 \tag{1.3}$$

and R_a is constant, then it will necessarily follow that $\mathrm{R}_a(\lambda) = 0$ for every $\lambda \in \mathbb{C}$, which is a contradiction since 0 is not invertible and R_a takes values in $\mathrm{Inv}(\mathcal{A})$.

Suppose that $|\lambda| > \|a\|$ (i.e., $1 > \|\lambda^{-1}a\|$). Then, part (1) of Lemma 1.22 implies that

$$\begin{aligned}\|\mathrm{R}_a(\lambda)\| = |\lambda^{-1}|\|(1_{\mathcal{A}} - \lambda^{-1}a)^{-1}\| &= |\lambda^{-1}|\left\|\sum_{n=0}^{\infty}(\lambda^{-1}a)^n\right\| \\ &\leqslant |\lambda^{-1}|\sum_{n=0}^{\infty}\left\|(\lambda^{-1}a)^n\right\| \leqslant |\lambda^{-1}|\sum_{n=0}^{\infty}\left\|\lambda^{-1}a\right\|^n = \frac{|\lambda^{-1}|}{1 - \|\lambda^{-1}a\|}.\end{aligned}$$

It then clearly follows from the above inequality that (1.3) holds.

Since R_a is defined on all of $\mathbb{C}$, it follows from Proposition 1.23 that it is entire. Therefore, if R_a is uniformly bounded on $\mathbb{C}$, then it will follow from Liouville's Theorem that R_a is constant,

concluding the proof of the theorem. For each $n \in \mathbb{N}$, let $K_n = \{\lambda \in \mathbb{C} : |\lambda| \leqslant n\}$. Then, $K_1 \subset K_2 \subset \cdots$ and $\mathbb{C} = \bigcup_n K_n$. Furthermore, the Heine-Borel Theorem implies that each K_n is compact. Therefore, since R_a is continuous and the image of a compact set through a continuous function is compact, each $\mathrm{R}_a(K_n)$ is compact, hence bounded (that is, for every $n \in \mathbb{N}$, one has $\sup\{\|\mathrm{R}_a(\lambda)\| : \lambda \in K_n\} < \infty$). We claim that there exists $N \in \mathbb{N}$ such that

$$\sup\{\|\mathrm{R}_a(\lambda)\| : \lambda \in \mathbb{C}\} = \sup\{\|\mathrm{R}_a(\lambda)\| : \lambda \in K_N\} < \infty, \tag{1.4}$$

or, in other words, the nondecreasing sequence of real nonnegative numbers $\big\{\sup\{\|\mathrm{R}_a(\lambda)\| : \lambda \in K_n\} : n \in \mathbb{N}\big\}$ eventually becomes constant. Suppose by contradiction that such is not the case. Then, it clearly is possible to construct a sequence $\{\lambda_i : i \in \mathbb{N}\} \subset \mathbb{C}$ such that for each $i \in \mathbb{N}$, $\lambda_i \in K_{n_i}$, where $n_1 < n_2 < \cdots$ and $\mathrm{R}_a(\lambda_1) < \mathrm{R}_a(\lambda_2) < \cdots$, which contradicts (1.3). Therefore, (1.4) holds, hence R_a is uniformly bounded. □

As a first application of the study of the properties of the spectrum of algebras, consider the following result, which characterizes *division* unital Banach algebras completely:

Corollary 1.24 (Gelfand-Mazur Theorem)**.** *Let $(\mathcal{A}, \|\cdot\|)$ be a unital Banach algebra such that every nonzero element in $\mathcal{A}$ is invertible. Then, $\mathcal{A}$ is isometrically isomorphic to $\mathbb{C}$.*

Proof. Consider an arbitrary element $a \in \mathcal{A}$. If $\lambda \in \mathbb{C}$ is in the spectrum of a, then it must be the case that $\lambda \cdot 1_{\mathcal{A}} - a = 0$, that is, $\lambda 1_{\mathcal{A}} = a$ (otherwise, there would exist a nonzero element which is not invertible). Let $\lambda_1, \lambda_2 \in \mathbb{C}$ be such that $\lambda_1, \lambda_2 \in \mathrm{Sp}_{\mathcal{A}}(a)$. Then, $\lambda_1 1_{\mathcal{A}} - a = 0 = \lambda_2 1_{\mathcal{A}} - a$, which implies that $\lambda_1 = \lambda_2$. Therefore, $\mathrm{Sp}_{\mathcal{A}}(a)$ contains at most one element. According to part (3) of Theorem 1.21, $\mathrm{Sp}_{\mathcal{A}}(a)$ contains at least one element, which implies that every element of $\mathcal{A}$ is of the form $\lambda \cdot 1_{\mathcal{A}}$ for some $\lambda \in \mathbb{C}$. Given that any unital algebra contains every scalar multiple of the unit, it is clear that the isometric isomorphism from $\mathcal{A}$ to $\mathbb{C}$ is achieved by mapping every element a to the unique scalar in $\mathrm{Sp}_{\mathcal{A}}(a)$ divided by $\|1_{\mathcal{A}}\|$. □

Another application of Theorem 1.21 (more precisely, part (1)) consists in providing an upper bound for the spectral radius of an element of a unital Banach algebra in terms of the norm of that element (or, conversely, a lower bound on the norm in terms of the spectrum). The following theorem provides yet another link between the spectrum and the norm:

Theorem 1.25 (Gelfand-Beurling formula)**.** *Let $(\mathcal{A}, \|\cdot\|)$ be a unital Banach algebra. For every $a \in \mathcal{A}$, one has*

$$\rho(a) = \lim_{n\to\infty} \|a^n\|^{1/n}.$$

Proof. Let $a \in \mathcal{A}$ be arbitrary. Assume for now that the limit

$$L = \lim_{n\to\infty} \|a^n\|^{1/n}. \tag{1.5}$$

exists and is finite.

Let $\lambda \in \mathbb{C}$ be such that $|\lambda| > L$. Then, according to the Cauchy-Hadamard Theorem,[11] the power series

$$\sum_{n=0}^{\infty} (\lambda^{-1} - 0)^n a^n = \sum_{n=0}^{\infty} (\lambda^{-1})^n a^n$$

[11] See section B.4.

converges to an element $b \in \mathcal{A}$. From the proof of part (1) of Lemma 1.22, it follows that b is such that $b(1_{\mathcal{A}} - \lambda^{-1}a) = (1_{\mathcal{A}} - \lambda^{-1}a)b = 1_{\mathcal{A}}$, which implies that $\lambda \cdot 1_{\mathcal{A}} - a$ is invertible, hence $\lambda \notin \mathrm{Sp}_{\mathcal{A}}(a)$. Therefore, if $\lambda \in \mathrm{Sp}_{\mathcal{A}}(a)$, then $|\lambda| \leqslant L$, and thus $\rho(a) \leqslant L$.

Suppose by contradiction that $\rho(a) < L$. Then, there exists a complex number λ such that $\rho(a) < |\lambda| < L$. For every $n \in \mathbb{N}$, the fact that $(\mathcal{A}, \|\cdot\|)$ is a Banach algebra implies that $\|a^n\|^{1/n} \leqslant \|a\|$, and thus $L \leqslant \|a\|$. Therefore, $|\lambda| < \|a\|$, which implies by part (1) of Lemma 1.22 that the power series

$$\sum_{n=0}^{\infty} (\lambda^{-1})^n a^n \tag{1.6}$$

converges to $(1 - \lambda^{-1}a)^{-1}$. Since

$$L = \lim_{n\to\infty} \|a^n\|^{1/n} = \limsup_{n\to\infty} \|a^n\|^{1/n}$$

and $|\lambda| < L$ (i.e., $|\lambda^{-1}| > L^{-1}$), the convergence of the series (1.6) contradicts the Cauchy-Hadamard Theorem. Consequently, $\rho(a) \geqslant L$.

According to what was established so far in this proof, if the limit L in (1.5) exists and is finite, then it is equal to $\rho(a)$. Thus, to conclude the proof of the theorem, it only remains to show that (1.5) exists and is finite. Since $(\mathcal{A}, \|\cdot\|)$ is a Banach algebra, $\|a^n\| \leqslant \|a\|^n$ for each n. Therefore, every member of the sequence $\{\|a^n\|^{1/n} : n \in \mathbb{N}\}$ is bounded above by $\|a\|$. This implies that if L exists, then it is lower or equal to $\|a\|$ and thus finite. Fix an arbitrary positive integer $k \in \mathbb{N}$. For each $n \in \mathbb{N}$, it follows from the Quotient-Remainder Theorem[12] that there exists two integers $q_n \geqslant 0$ and $0 \leqslant r_n \leqslant k-1$ such that $n = q_n k + r_n$, and thus

$$\frac{q_n}{n} = \frac{1}{k}\left(1 - \frac{r_n}{n}\right).$$

Since $r_n \leqslant k-1$ for each n, it is clear that

$$\lim_{n\to\infty} \frac{r_n}{n} = 0 \text{ and } \lim_{n\to\infty} \frac{q_n}{n} = \frac{1}{k}. \tag{1.7}$$

For every real number $\alpha \geqslant 0$, the function

$$\begin{array}{ccc} [0,\infty) & \to & [0,\infty) \\ b & \mapsto & \alpha^b \end{array}$$

is continuous. Therefore, according to (1.7), for every real numbers $\alpha, \beta \geqslant 0$, one has

$$\lim_{n\to\infty} \alpha^{r_n/n} = 1 \text{ and } \lim_{n\to\infty} \beta^{q_n/n} = \beta^{1/k},$$

hence

$$\lim_{n\to\infty} \beta^{q_n/n}\alpha^{r_n/n} = \beta^{1/k}. \tag{1.8}$$

Applying (1.8) with $\beta = \|a^k\|$ and $\alpha = \|a\|$ yields

$$\begin{aligned} \limsup_{n\to\infty} \|a^n\|^{1/n} &= \limsup_{n\to\infty} \|a^{q_n k + r_n}\|^{1/n} \leqslant \limsup_{n\to\infty} (\|a^{q^n k}\| \|a^{r_n}\|)^{1/n} \\ &\leqslant \limsup_{n\to\infty} \|a^k\|^{q_n/n} \|a\|^{r_n/n} = \|a^k\|^{1/k}. \end{aligned}$$

[12] Otherwise known as the Division Algorithm.

Since $k \in \mathbb{N}$ was arbitrary, $\limsup_n \|a^n\|^{1/n} \leqslant \inf\{\|a^n\|^{1/n} : n \geqslant k\}$ for every $k \in \mathbb{N}$, hence

$$\limsup_{n\to\infty} \|a^n\|^{1/n} \leqslant \lim_{k\to\infty} \inf\{\|a^n\|^{1/n} : n \geqslant k\} = \liminf_{n\to\infty} \|a^n\|^{1/n}.$$

In conclusion, the limit (1.5) exists and is finite. □

Two interesting features of the self-adjoint elements of the C^*-algebra $(\mathcal{M}_n(\mathbb{C}), *, \|\cdot\|_{\mathrm{op}})$ mentioned in Example 1.9 (i.e., Hermitian matrices) are that

(1) their norm is given by

$$\|A\|_{\mathrm{op}} = \max\{|\lambda| : \lambda \text{ is an eigenvalue of } A\}; \text{ and}$$

(2) their spectrum is a subset of the real numbers.

Thanks to the Gelfand-Beurling Formula and the other properties of the spectrum that were established up to this point, these results may be extended to *any* C^*-algebra.

Corollary 1.26. *Let $(\mathcal{A}, *, \|\cdot\|)$ be a unital C^*-algebra and $a \in \mathcal{A}$ be an arbitrary self-adjoint element. Then, $\|a\| = \rho(a)$.*

Proof. This result is a direct consequence of the Gelfand-Beurling formula and Remark 1.13: Given that $\|a^{2^n}\|^{2^{-n}}$ is a subsequence of $\|a^n\|^{1/n}$, one has

$$\rho(a) = \lim_{n\to\infty} \|a^n\|^{1/n} = \lim_{n\to\infty} \|a^{2^n}\|^{2^{-n}} = \|a\|,$$

which concludes the proof. □

Proposition 1.27. *Let $(\mathcal{A}, *, \|\cdot\|)$ be a unital C^*-algebra. If $a \in \mathcal{A}$ is self-adjoint, then $\mathrm{Sp}_{\mathcal{A}}(a) \subset \mathbb{R}$.*

Proof. Let $a \in \mathcal{A}$ be self-adjoint. We want to prove that if $\lambda = \alpha + \mathrm{i}\beta$ with $\alpha, \beta \in \mathbb{R}$ is such that $\beta \neq 0$, then $\lambda \notin \mathrm{Sp}_{\mathcal{A}}(a)$, that is, $(\alpha + \mathrm{i}\beta)1_{\mathcal{A}} - a \in \mathrm{Inv}(\mathcal{A})$.

Let $\lambda = \alpha + \mathrm{i}\beta$, where $\alpha, \beta \in \mathbb{R}$ and $\beta \neq 0$. Then,

$$\lambda \cdot 1_{\mathcal{A}} - a = (\alpha + \mathrm{i}\beta)1_{\mathcal{A}} - a = \beta\left(\frac{\alpha \cdot 1_{\mathcal{A}} - a}{\beta} + \mathrm{i} \cdot 1_{\mathcal{A}}\right).$$

Note that

$$\left(\frac{\alpha \cdot 1_{\mathcal{A}} - a}{\beta}\right)^* = \frac{\overline{\alpha} \cdot 1_{\mathcal{A}}^* - a^*}{\overline{\beta}} = \frac{\alpha \cdot 1_{\mathcal{A}} - a}{\beta}.$$

Thus, if $b + \mathrm{i} \cdot 1_{\mathcal{A}}$ is invertible for every self-adjoint element $b \in \mathcal{A}$, then the result will follow.

Suppose by contradiction that there exists a self-adjoint element $b \in \mathcal{A}$ such that $\mathrm{i} \cdot 1_{\mathcal{A}} + b$ is not invertible. Then, $-\mathrm{i} \cdot 1_{\mathcal{A}} - b \notin \mathrm{Inv}(\mathcal{A})$, which implies that $-\mathrm{i} \in \mathrm{Sp}_{\mathcal{A}}(b)$. According to Remark 1.19, this implies that $-\mathrm{i}^2 = 1 \in \mathrm{Sp}_{\mathcal{A}}(\mathrm{i} \cdot b)$. Consequently, for every $\lambda \in \mathbb{R}$, the element

$$1_{\mathcal{A}} - \mathrm{i} \cdot b = \lambda \cdot 1_{\mathcal{A}} + 1_{\mathcal{A}} - \lambda \cdot 1_{\mathcal{A}} - \mathrm{i} \cdot b = (\lambda + 1)1_{\mathcal{A}} - (\lambda \cdot 1_{\mathcal{A}} + \mathrm{i} \cdot b)$$

is not invertible, hence $\lambda + 1 \in \mathrm{Sp}_{\mathcal{A}}(\lambda \cdot 1_{\mathcal{A}} + \mathrm{i} \cdot b)$. By part (1) of Theorem 1.21, this implies that $|\lambda + 1| \leqslant \|\lambda \cdot 1_{\mathcal{A}} + \mathrm{i} \cdot b\|$, and therefore, since $(\mathcal{A}, *, \|\cdot\|)$ is a C^*-algebra,

$$\begin{aligned}(\lambda+1)^2 = |\lambda+1|^2 &\leqslant \|\lambda \cdot 1_{\mathcal{A}} + \mathrm{i} \cdot b\|^2 = \|(\lambda \cdot 1_{\mathcal{A}} + \mathrm{i} \cdot b)^*(\lambda \cdot 1_{\mathcal{A}} + \mathrm{i} \cdot b)\| \\ &= \|(\overline{\lambda} \cdot 1_{\mathcal{A}}^* + \bar{\mathrm{i}} b^*)(\lambda \cdot 1_{\mathcal{A}} + \mathrm{i} \cdot b)\| = \|(\lambda \cdot 1_{\mathcal{A}} - \mathrm{i} \cdot b)(\lambda \cdot 1_{\mathcal{A}} + \mathrm{i} \cdot b)\| \\ &= \|\lambda^2 \cdot 1_{\mathcal{A}} + b^2\| \leqslant \lambda^2 + \|b^2\|.\end{aligned}$$

Consequently, as $\lambda \in \mathbb{R}$ was arbitrary, it follows that $2\lambda + 1 \leqslant \|b^2\|$ for every $\lambda \in \mathbb{R}$, which is obviously a contradiction since $\|b^2\|$ is a fixed real number. □

1.3. The Gelfand-Naimark Theorems

Some of the most important theorems concerning C^*-algebras are a series of results usually referred to as the Gelfand-Naimark Theorems, which essentially state that Examples 1.9, 1.10, and 1.11 are (up to C^*-algebra isomorphism) the only possible realizations of C^*-algebras:

Gelfand-Naimark Theorem I ([**DB86**] page 3)**.** *Let $(\mathcal{A}, *, \|\cdot\|)$ be an arbitrary C^*-algebra. Then, $(\mathcal{A}, *, \|\cdot\|)$ is isometrically $*$-isomorphic to a C^*-subalgebra of $\big(B(H), *, \|\cdot\|_{\mathrm{op}}\big)$ for some Hilbert space $\big(H, \langle\cdot,\cdot\rangle\big)$.*

Gelfand-Naimark Theorem II ([**DB86**] page 3)**.** *Let $(\mathcal{A}, *, \|\cdot\|)$ be a commutative C^*-algebra. Then, there exists a locally compact Hausdorff topological space X such that $\big(\mathrm{C}_0(X, \mathbb{C}), \bar{\cdot}, \|\cdot\|_\infty\big)$ is isometrically $*$-isomorphic to $(\mathcal{A}, *, \|\cdot\|)$.*

Gelfand-Naimark Theorem III. *Let $(\mathcal{A}, *, \|\cdot\|)$ be a commutative unital C^*-algebra. Then, there exists a compact Hausdorff topological space X such that $\big(\mathrm{C}(X, \mathbb{C}), \bar{\cdot}, \|\cdot\|_\infty\big)$ is isometrically $*$-isomorphic to $(\mathcal{A}, *, \|\cdot\|)$.*

From these results, one may argue that, from an abstract perspective, any distinction between the presentation of two C^*-algebras $\mathcal{A}_1$ and $\mathcal{A}_2$ is *in principle artificial*, as both can be realized as C^*-algebras of bounded linear operators on a suitable Hilbert space.[13] Furthermore, the Gelfand-Naimark Theorem I can be seen as the foundational result of the theory of C^*-algebras (that is, the result that *justifies* the importance of the field), since the study of algebras of bounded linear maps on Hilbert spaces can be reduced to the study of abstract C^*-algebras from the axioms presented in Definition 1.7.

The Gelfand-Naimark theorems also enables one to make sense of the concept of a *noncommutative topological space*: Given a compact Hausdorff topological space X, the open subsets of X are in a one to one correspondance with the closed ideals of the algebra $\mathrm{C}(X, \mathbb{C})$ of continuous functions on X (see Section B.5.1). As a consequence, it turns out that many *topological concepts* have an equivalent *algebraic concept*, and thus certain statements about topological spaces can be *translated* into statement about algebras. As explained in more details in **[Va06]** page 3, one thus obtains Table 1 (see below), highlighting the correspondances between topological and algebraic notions.

Thus, one may think of a *noncommutative locally compact topological space* as a noncommutative C^*-algebra, and the noncommutative analogs of topological concepts are given by

[13] However, as will become clear in Theorem 1.28, the *bounded linear maps on a Hilbert space* presentation of some C^*-algebras are sometimes so abstract that it may be more useful to think of those spaces in other terms.

Table 1. Correspondances between topological and algebraic notions.

Topology	Algebra
Locally Compact Topological Space	Commutative C^*-algebra
Compact Topological Space	Commutative Unital C^*-algebra
Compactification	Unitalization
Open Set	Closed Ideal
Homeomorphism	Automorphism
Regular Measure	Positive Linear Functional
Image (of a function f)	Spectrum (of the element f)

extending Table 1 in the noncommutative case. This idea of understanding *generalized geometrical spaces* in terms of noncommutative algebras of functions defined on said spaces is the foundation of a mathematical theory called **Noncommutative Geometry** (see for instance **[Va06]** and **[Co94]**).

For our purposes in later sections and chapters of this thesis, the most *useful* version of the Gelfand-Naimark Theorem is version III (that is, that which relates to commutative unital C^*-algebras). The objective of this section is to provide a complete proof of this theorem. Though a proof of versions I and II of the theorem will not be provided in this thesis (the reader is referred to Chapters 2 and 3 of **[DB86]** or Section I.9 of **[Da96]** for complete proofs), for the purpose of showcasing the *general idea* behind the proof, we provide the following special case:[14]

Theorem 1.28. *Let* $\big(\mathrm{C}(X,\mathbb{C}),\bar{\cdot},\|\cdot\|_\infty\big)$ *be defined as in Example 1.11 (that is, X is a compact Hausdorff topological space).* $\big(\mathrm{C}(X,\mathbb{C}),\bar{\cdot},\|\cdot\|_\infty\big)$ *is isometrically $*$-isomorphic to a C^*-subalgebra of* $\big(B(H),*,\|\cdot\|_{\mathrm{op}}\big)$ *for some Hilbert space* $\big(H,\langle\cdot,\cdot\rangle\big)$.

Proof. Let Π be the set of positive linear functionals $\varphi:\mathrm{C}(X,\mathbb{C})\to\mathbb{C}$ such that $\varphi(\mathbf{1}_X)=1$. According to the Riesz Representation Theorem, Π is in a one to one correspondance with the set of regular probability measures on the space X.

Let $\varphi\in\Pi$ be arbitrary, and let μ_φ be the regular probability measure associated to φ, that is,

$$\varphi(f)=\int f(x)\ \mathrm{d}\mu_\varphi(x), \qquad\qquad f\in\mathrm{C}(X,\mathbb{C}).$$

Define the set

$$\begin{aligned} N_\varphi &= \big\{f\in\mathrm{C}(X,\mathbb{C}):\varphi(f\overline{f})=0\big\}\\ &= \big\{f\in\mathrm{C}(X,\mathbb{C}):f=0\ \mu_\varphi\text{-almost everywhere}\big\}, \end{aligned}$$

and let H_φ be the closure of $\mathrm{C}(X,\mathbb{C})/N_\varphi$ under the inner product defined as

$$\langle f,g\rangle_\varphi \coloneqq \varphi(\overline{g}f)=\int \overline{g(x)}f(x)\ \mathrm{d}\mu_\varphi(x),$$

[14] The general proof of the Gelfand-Naimark Theorem I follows more or less the same steps as in the provided proof of Theorem 1.28, but many of the results used here do not trivially extend to the general case.

(that is, under the norm $\|f\|_\varphi = \sqrt{\langle f, f\rangle_\varphi}$). Then, H_φ clearly is a Hilbert space (in fact, H_φ is the space $L^2(X, \mathscr{B}(X), \mu_\varphi)$ of square-integrable functions with respect to μ_φ, and $\|\cdot\|_\varphi$ is the L^2-norm).

For every $f \in \mathrm{C}(X, \mathbb{C})$, define the map

$$\begin{array}{rccc} \pi_f^{(\varphi)}: & H_\varphi & \to & H_\varphi \\ & g & \mapsto & fg. \end{array}$$

Every such map is clearly linear, and since $|fg|^2 \leqslant \|f\|_\infty^2 |g|^2$ (as X is compact and f is continuous), one has

$$\left\|\pi_f^{(\varphi)}\right\|_{\mathrm{op}} = \sup\left\{\|fg\|_\varphi : \|g\|_\varphi = 1\right\} \leqslant \sup\left\{\|f\|_\infty \|g\|_\varphi : \|g\|_\varphi = 1\right\} = \|f\|_\infty,$$

and thus every $\pi_f^{(\varphi)}$ is a bounded linear operator. Moreover, one easily checks that the map

$$\begin{array}{ccc} \mathrm{C}(X, \mathbb{C}) & \to & B(H_\varphi) \\ f & \mapsto & \pi_f^{(\varphi)} \end{array} \tag{1.9}$$

defines a $*$-algebra homomorphism,[15] as

$$\begin{aligned} \left\langle \pi_f^{(\varphi)}(g), h \right\rangle_\varphi &= \int \overline{h(x)} f(x) g(x)\, \mathrm{d}\mu_\varphi(x) = \int \overline{\overline{f(x)} h(x)} g(x)\, \mathrm{d}\mu_\varphi(x) \\ &= \left\langle g, \pi_{\overline{f}}^{(\varphi)}(h) \right\rangle_\varphi, \end{aligned}$$

where $g, h \in H_\varphi$ are arbitrary.

Let $H = \bigoplus_{\varphi\in\Pi} H_\varphi$, that is, H is the completion of the set

$$V = \left\{(f_\varphi : \varphi \in \Pi) : f_\varphi \in H_\varphi;\ \text{and } f_\varphi \neq 0 \text{ for finitely many } \varphi\right\}$$

under the norm induced by the inner product

$$\left\langle (f_\varphi : \varphi \in \Pi), (g_\varphi : \varphi \in \Pi) \right\rangle = \sum_{\varphi\in\Pi} \langle f_\varphi, g_\varphi\rangle_\varphi.$$

For every $f \in \mathrm{C}(X, \mathbb{C})$, define the map $\pi_f = \bigoplus_{\varphi\in\Pi} \pi_f^{(\varphi)}$, that is,

$$\begin{array}{rccc} \pi_f: & H & \to & H \\ & (g_\varphi : \varphi \in \Pi) & \mapsto & (f g_\varphi : \varphi \in \Pi). \end{array}$$

Clearly, every π_f is linear, and

$$\begin{aligned} \|\varphi_f\|_{\mathrm{op}} &= \sup\left\{ \left(\sum_{\varphi\in\Pi} \|fg\|_\varphi^2\right)^{1/2} : \left(\sum_{\varphi\in\Pi} \|g\|_\varphi^2\right)^{1/2} = 1 \right\} \\ &\leqslant \sup\left\{ \left(\sum_{\varphi\in\Pi} \|f\|_\infty^2 \|g\|_\varphi^2\right)^{1/2} : \left(\sum_{\varphi\in\Pi} \|g\|_\varphi^2\right)^{1/2} = 1 \right\} \\ &= \|f\|_\infty, \end{aligned} \tag{1.10}$$

[15] The general version of the homomorphism defined in (1.9) is known as the **GNS construction**, named after Gelfand, Naimark, and Segal. See for instance **[Da96]** Theorem I.9.6).

hence every π_f is a bounded linear map. In addition, one easily checks that the map

$$\begin{array}{ccc} \mathrm{C}(X,\mathbb{C}) & \to & B(H) \\ f & \mapsto & \pi_f \end{array} \tag{1.11}$$

is a $*$-homomorphism. Thus, if it is shown that (1.11) is injective and norm-preserving, the proof will be complete.

Let $f \in \mathrm{C}(X,\mathbb{C})$ be arbitrary. Since X is compact and f is continuous, there exists $x \in X$ such that $|f(x)| = \|f\|_\infty$. Let δ_x be the Dirac mass at that point. Since δ_x is a regular probability measure on X, there exists $\varphi_x \in \Pi$ such that $\varphi_x(g) = \int g(y)\ \mathrm{d}\delta_x(y)$ for every g. Define $(g_\varphi : \varphi \in \Pi) \in H$ to be such that $g_\varphi = \mathbf{1}_X$ if $\varphi = \varphi_x$ and $g_\varphi = 0$ otherwise. Then, one clearly has

$$\|\pi_f\|_{\mathrm{op}} \geqslant \big\|\pi_f\big((g_\varphi : \varphi \in \Pi)\big)\big\| = \left(\int |f(y)|^2\ \mathrm{d}\delta_x(y)\right)^{1/2} = |f(x)| = \|f\|_\infty,$$

which, together with (1.10), implies that (1.11) is norm-preserving.

To prove that $f \mapsto \pi_f$ is injective, simply notice that if $f \neq g$, say at a point $x \in X$ (i.e., $f(x) \neq g(x)$), then f and g disagree almost everywhere with respect to the Dirac mass δ_x at the point x, and thus π_f and π_g disagree on any element $(h_\varphi : \varphi \in \Pi)$ such that $h_{\varphi_x} = \mathbf{1}_X$, where φ_x is defined as in the previous paragraph, that is, as the linear functional associated to δ_x. □

We now show the Gelfand-Naimark Theorem III in full generality. Given a Banach algenra $(\mathcal{A}, \|\cdot\|)$, let $\Delta(\mathcal{A})$ denote the collection of nonzero[16] multiplicative linear functionals on $\mathcal{A}$, that is, the set of all nonzero algebra homomorphisms from $\mathcal{A}$ to $\mathbb{C}$.

Remark 1.29. Let $\mathcal{A}$ be a unital algebra. If $\varphi : \mathcal{A} \to \mathbb{C}$ is a nonzero algebra homomorphism, then $\varphi(1_\mathcal{A}) = 1$. Indeed, for any $a \in \mathcal{A}$ such that $\varphi(a) \neq 0$, one has $\varphi(a) = \varphi(1_\mathcal{A} a) = \varphi(1_\mathcal{A})\varphi(a)$.

Let $(\mathcal{A}, \|\cdot\|)$ be a unital Banach algebra such that $\Delta(\mathcal{A})$ is nonempty. The **Gelfand transform** of an arbitrary element $a \in \mathcal{A}$, denoted $\mathcal{G}(a)$, is defined as the following function:

$$\begin{array}{rccc} \mathcal{G}(a): & \Delta(\mathcal{A}) & \to & \mathbb{C} \\ & \varphi & \mapsto & \varphi(a). \end{array}$$

The **Gelfand topology** is the intersection of every topology on $\Delta(\mathcal{A})$ such that all the Gelfand transforms on elements of $\mathcal{A}$ are continuous (as will be shown in Proposition 1.30, the Gelfand topology is not the trivial topology). Unless otherwise mentioned, given any Banach algebra $(\mathcal{A}, \|\cdot\|)$, the space $\Delta(\mathcal{A})$ is assumed to be equipped with the Gelfand topology. The **Gelfand transform** on the algebra $\mathcal{A}$ is defined as the function

$$\begin{array}{rccc} \mathcal{G}: & \mathcal{A} & \to & \mathrm{C}\big(\Delta(\mathcal{A}),\mathbb{C}\big) \\ & a & \mapsto & \mathcal{G}(a). \end{array}$$

Proposition 1.30. *Let $(\mathcal{A}, \|\cdot\|)$ be a Banach algebra. The Gelfand topology is the subspace topology that $\Delta(\mathcal{A})$ inherits from the w^*-topology[17] as a subset of the dual space $\mathcal{A}^*$.*

[16] Nonzero in the sense that $\varphi(a) \neq 0$ for at least one $a \in \mathcal{A}$.

[17] See section C.4.

Proof. For every $\lambda \in \mathbb{C}$, $a \in \mathcal{A}$ and $\varepsilon > 0$, let

$$U(\lambda, a, \varepsilon) = \{\psi \in \mathcal{A}^* : |\psi(a) - \lambda| < \varepsilon\}.$$

If the set

$$\{U(\lambda, a, \varepsilon) \cap \Delta(\mathcal{A}) : a \in \mathcal{A},\ \lambda \in \mathbb{C},\ \varepsilon > 0\} \tag{1.12}$$

is a sub-basis for $\Delta(\mathcal{A})$ and a topology τ on $\Delta(\mathcal{A})$ contains every element in (1.12) if and only if it makes every Gelfand transform continuous, then the lemma will be proved. (Indeed, the topology generated by a sub-basis is the intersection of every topology containing that sub-basis. See section C.1.)

If $\varphi \in \Delta(\mathcal{A})$, then $\varphi \in U(\varphi(a), a, \varepsilon) \cap \Delta(\mathcal{A})$ for every $\varepsilon > 0$ and $a \in \mathcal{A}$, hence

$$\bigcup_{\varphi \in \Delta(\mathcal{A}), a \in \mathcal{A}, \varepsilon > 0} U(\varphi(a), a, \varepsilon) \cap \Delta(\mathcal{A}) = \Delta(\mathcal{A}).$$

Therefore, (1.12) is a sub-basis on $\Delta(\mathcal{A})$.

Let τ be a topology on $\Delta(\mathcal{A})$ that contains (1.12), i.e., $U(\lambda, a, \varepsilon) \cap \Delta(\mathcal{A})$ is an open set in τ for every $\lambda \in \mathbb{C}$, $a \in \mathcal{A}$ and $\varepsilon > 0$. Let $a \in \mathcal{A}$ be arbitrary, and let $\mathcal{G}(a)$ be the Gelfand transform at that point. Since the open balls form a basis for the Euclidean topology on $\mathbb{C}$, to prove that $\mathcal{G}(a)$ is continuous, it suffices to prove that $\mathcal{G}(a)^{-1}\big(\mathrm{B}(\lambda, \varepsilon)\big)$ is open for every $\lambda \in \mathbb{C}$ and $\varepsilon > 0$.[18] For every $\lambda \in \mathbb{C}$ and $\varepsilon > 0$, one has

$$\mathcal{G}(a)^{-1}\big(\mathrm{B}(\lambda, \varepsilon)\big) = \{\psi \in \Delta(\mathcal{A}) : |\psi(a) - \lambda| < \varepsilon\} = U(\lambda, a, \varepsilon) \cap \Delta(\mathcal{A}),$$

and thus $G(a)$ is continuous with respect to τ.

Let τ be a topology on $\Delta(\mathcal{A})$ that makes every Gelfand transform continuous. Let $a \in \mathcal{A}$, $\lambda \in \mathbb{C}$ and $\varepsilon > 0$ be arbitrary. Knowing that the Gelfand transform $\mathcal{G}(a)$ is continuous and that $\mathrm{B}(\lambda, \varepsilon)$ is open in $\mathbb{C}$, it follows from the fact that $\mathcal{G}(a)$ is continuous that $\mathcal{G}(a)^{-1}\big(\mathrm{B}(\lambda, \varepsilon)\big) = U(\lambda, a, \varepsilon) \cap \Delta(\mathcal{A})$ is an open set in τ, hence τ contains (1.12). □

Let $\mathcal{A} = (V, \cdot)$ be an algebra and I be a vector subspace of V. I is called a **right ideal** (respectively, **left ideal**) of $\mathcal{A}$ if, for every $a \in \mathcal{A}$ and $b \in I$, one has $ba \in I$ (respectively, $ab \in I$). If $I \subset V$ is both a right ideal and a left ideal, it is called an **ideal** of $\mathcal{A}$. An ideal $I \subset V$ is said to be **proper** if it is a proper subset of V, that is $I \neq V$. A proper ideal $M \subset V$ is said to be **maximal** if every ideal $I \subset V$ which contains M is either M itself or V.

Let $\mathcal{A}$ be an algebra and I be an ideal of $\mathcal{A}$. The quotient space $\mathcal{A}/I$ is defined using the equivalence relation

$$a \sim b \text{ if and only if } a - b \in I.$$

This relation yields the equivalence classes $a + I = \{a + b : b \in I\}$. For every $k \in \mathbb{C}$ and $a, b \in \mathbb{C}$, let $(a + I) + (b + I) = (a + b) + I$, $k(a + I) = ka + I$ and $(a + I)(b + I) = ab + I$. One could easily show that these operations are well defined and that $\mathcal{A}/I$ equipped with them is in fact an algebra.

[18] Recall that $\mathrm{B}(x, \varepsilon)$ denotes the open ball of centre x and radius ε.

Remark 1.31. Let $\mathcal{A}$ be an algebra and I be an ideal of $\mathcal{A}$. Given that $\mathcal{A}/I$ is an algebra, the **quotient map** π_I, defined as

$$\begin{array}{rccc} \pi_I : & \mathcal{A} & \to & \mathcal{A}/I \\ & a & \mapsto & a+I \end{array}$$

is clearly an algebra homomorphism. Furthermore, $\mathrm{Ker}(\pi_I) = I$.

Given a Banach algebra $(\mathcal{A}, \|\cdot\|)$ and an ideal I, the **quotient norm** $\|\cdot\|_I : \mathcal{A}/I \to [0,\infty)$ is defined as

$$\|a+I\|_I = \inf\{\|c\| : c \in a+I\}, \qquad a \in \mathcal{A}.$$

Clearly, the quotient norm is a well defined function. Furthermore, it can be shown that if the ideal I is a closed subset of $\mathcal{A}$, then $\|\cdot\|_I$ is a norm and $(\mathcal{A}/I, \|\cdot\|_I)$ is a Banach algebra.

Lemma 1.32. *Let $(\mathcal{A}, \|\cdot\|)$ be a commutative unital Banach algebra. Then, every element of $\mathcal{A}$ which is not invertible is contained in a maximal ideal of $\mathcal{A}$.*

Proof. Let $a \in \mathcal{A}$ be such that $a \notin \mathrm{Inv}(\mathcal{A})$, and let $\mathscr{I}_a$ be the set of proper ideals of $\mathcal{A}$ that contain a. Consider the partial order $(\mathscr{I}_a, \subset)$, where $\subset$ denotes the traditional set relation *proper subset of or equal to*. If $\mathscr{I}_a$ is not empty and every totally ordered subset of $(\mathscr{I}_a, \subset)$ has an upper bound in $\mathscr{I}_a$, it will follow from Zorn's Lemma that $\mathscr{I}_a$ has a maximal element M. It would then be easy to show that M is a maximal ideal: Suppose that I is an ideal of $\mathcal{A}$ such that $M \subset I$. If $I \neq \mathcal{A}$, then $I \in \mathscr{I}_a$, since $a \in M \subset I$. Therefore, since M is the maximal element of $\mathscr{I}_a$, it must be the case that $I \subset M$, which implies that $M = I$ by double inclusion.

Consider the ideal generated by a, which is defined as $a\mathcal{A} = \{ab : b \in \mathcal{A}\}$. ($a\mathcal{A} = \mathcal{A}a$ in this case since $\mathcal{A}$ is commutative.) $a\mathcal{A}$ is clearly an ideal, but it is also a proper ideal. To see this, suppose by contradiction that $1_{\mathcal{A}} \in a\mathcal{A}$, that is, there exists $b \in \mathcal{A}$ such that $ab = 1_{\mathcal{A}}$. This contradicts that a is not invertible, hence $1_{\mathcal{A}} \notin a\mathcal{A}$. Therefore, $a\mathcal{A}$ is a proper ideal of $\mathcal{A}$ which contains a, hence $\mathscr{I}_a$ is not empty.

Let $(I_j : j \in J)$ be an arbitrary totally ordered subset of $\mathscr{I}_a$. We claim that $\bigcup_j I_j$ is an upper bound of $(I_j : j \in J)$. It clearly is the case that $I_j \subset \bigcup_{j\in J} I_j$ for every $j \in J$. However, to show that $\bigcup_j I_j$ is an upper bound of $(I_j : j \in J)$ in the context of the partial order $(\mathscr{I}_a, \subset)$, it must be proved that $\bigcup_j I_j \in \mathscr{I}_a$, that is, $\bigcup_j I_j$ is a proper ideal of $\mathcal{A}$ (the fact that $\bigcup_j I_j$ contains a is obvious). Let $b, c \in \bigcup_j I_j$ be arbitrary. Then, there exists $n, m \in J$ such that $b \in I_n$ and $c \in I_m$. Since $(I_j : j \in J)$ is totally ordered, then $I_m \subset I_n$ or $I_n \subset I_m$. In any case, there exists $k \in J$ (k is either m or n) such that $b, c \in I_k$. Since I_k is by definition an ideal, then $b+c \in I_k$ and $db \in I_k$ for all $d \in \mathcal{A}$, which implies in particular that $b+c \in \bigcup_j I_j$ and that $db \in \bigcup_j I_j$ for all $d \in \mathcal{A}$. Consequently, $\bigcup_j I_j$ is an ideal of $\mathcal{A}$. To show that $\bigcup_j I_j$ is proper, suppose by contradiction that $1_{\mathcal{A}} \in \bigcup_j I_j$. This is obviously contradictory, since it implies that $1_{\mathcal{A}} \in I_m$ for some $m \in J$, which contradicts that I_m is proper. □

Lemma 1.33. *Let $(\mathcal{A}, \|\cdot\|)$ be a commutative unital algebra. $M \subset \mathcal{A}$ is a maximal ideal of $\mathcal{A}$ if and only if $M = \mathrm{Ker}(\varphi)$ for some $\varphi \in \Delta(\mathcal{A})$.*

Proof. Let M be a maximal ideal of $\mathcal{A}$, and let $a \in \mathcal{A}$ be such that $a \notin M$ (that is, $a + M$ is non-zero in $\mathcal{A}/M$). Consider the set

$$I_a = \{ab + c : b \in \mathcal{A} \text{ and } c \in M\}.$$

Clearly, given that $\mathcal{A}$ is commutative, I_a is an ideal of $\mathcal{A}$. Furthermore, since $a \notin M$, that is, $a1_{\mathcal{A}} + 0 \notin M$, then $I_a \neq M$, which implies that $I_a = \mathcal{A}$ since M is maximal. Consequently, there exists $b \in \mathcal{A}$ and $c \in M$ such that $1_{\mathcal{A}} = ab + c$, which implies that

$$(a + M)(b + M) + (0_{\mathcal{A}} + M) = (ab + c + M) = (1_{\mathcal{A}} + M).$$

Therefore, $a + M$ is invertible in $\mathcal{A}/M$. Since $a \in \mathcal{A} \setminus M$ was arbitrary, the Gelfand-Mazur Theorem (see Corollary 1.24) implies that there exists an isomorphism φ_0 from $\mathcal{A}/M$ to $\mathbb{C}$. Then, the composition $\varphi = \varphi_0 \circ \pi_M$ (where π_M is the quotient map of $\mathcal{A}/M$, see Remark 1.31) is a nonzero algebra homomorphism from $\mathcal{A}$ to $\mathbb{C}$ (hence an element of $\Delta(\mathcal{A})$) and it is such that

$$\mathrm{Ker}(\varphi) = \big\{a \in \mathcal{A} : \varphi_0\big(\pi_M(a)\big) = 0\big\} = \{a \in \mathcal{A} : \pi_M(a) = 0_{\mathcal{A}} + M\} = M.$$

Let $\varphi \in \Delta(\mathcal{A})$. By the properties of algebra homomorphisms and the zero vector, it is clear that $\mathrm{Ker}(\varphi)$ is an ideal. Let I be an ideal of $\mathcal{A}$ which contains $\mathrm{Ker}(\varphi)$ as a proper subset. Then, there exists $b \in I$ such that $\varphi(b) \neq 0$. Let $a \in \mathcal{A}$ be arbitrary. Then,

$$\varphi\left(a - \tfrac{\varphi(a)}{\varphi(b)}b\right) = \varphi(a) - \frac{\varphi(a)}{\varphi(b)}\varphi(b) = 0,$$

which implies that $a - \frac{\varphi(a)}{\varphi(b)}b \in \mathrm{Ker}(\varphi) \subset I$. Since I is an ideal and $\frac{\varphi(a)}{\varphi(b)}b \in I$, then $a = a - \frac{\varphi(a)}{\varphi(b)}b + \frac{\varphi(a)}{\varphi(b)}b \in I$. As $a \in \mathcal{A}$ was arbitrary, it follows that $\mathcal{A} = I$, which proves that $\mathrm{Ker}(\varphi)$ is a maximal ideal. □

Theorem 1.34. *Let $(\mathcal{A}, \|\cdot\|)$ be a unital Banach algebra. Then, for every $a \in \mathcal{A}$, one has $\{\varphi(a) : \varphi \in \Delta(\mathcal{A})\} \subset \mathrm{Sp}_{\mathcal{A}}(a)$. If $\mathcal{A}$ is commutative, then $\{\varphi(a) : \varphi \in \Delta(\mathcal{A})\} = \mathrm{Sp}_{\mathcal{A}}(a)$.*

Proof. Let $a \in \mathcal{A}$ be arbitrary.

Suppose that $\lambda \in \mathbb{C}$ is such that $\lambda \notin \mathrm{Sp}_{\mathcal{A}}(a)$, that is, $\lambda \cdot 1_{\mathcal{A}} - a \in \mathrm{Inv}(\mathcal{A})$. Then, for every $\varphi \in \Delta(\mathcal{A})$, one has $\varphi(\lambda{\cdot}1_{\mathcal{A}}{-}a)\varphi\big((\lambda{\cdot}1_{\mathcal{A}}{-}a)^{-1}\big) = \varphi(1_{\mathcal{A}}) = 1$.[19] Thus, $\varphi(\lambda{\cdot}1_{\mathcal{A}}{-}a) = \lambda{-}\varphi(a) \neq 0$, which implies that $\lambda \notin \{\varphi(a) : \varphi \in \Delta(\mathcal{A})\}$. Since λ was arbitrary, then

$$\{\varphi(a) : \varphi \in \Delta(\mathcal{A})\} \subset \mathrm{Sp}_{\mathcal{A}}(a). \tag{1.13}$$

Suppose that $\mathcal{A}$ is commutative. Let $\lambda \in \mathbb{C}$ be such that $\lambda \in \mathrm{Sp}_{\mathcal{A}}(a)$, that is, $\lambda \cdot 1_{\mathcal{A}} - a \notin \mathrm{Inv}(\mathcal{A})$. Then, according to Lemma 1.32, there exists a maximal ideal $M \subset \mathcal{A}$ that contains $\lambda{\cdot}1_{\mathcal{A}}{-}a$, which implies by Lemma 1.33 that there exists $\varphi \in \Delta(\mathcal{A})$ such that $\lambda{\cdot}1_{\mathcal{A}}{-}a \in \mathrm{Ker}(\varphi)$. Thus, $\varphi(\lambda{\cdot}1_{\mathcal{A}}{-}a) = \lambda{-}\varphi(a) = 0$, which implies that $\lambda \in \{\varphi(a) : \varphi \in \Delta(\mathcal{A})\}$. Since $\lambda \in \mathrm{Sp}_{\mathcal{A}}(a)$ was arbitrary, it follows that $\{\varphi(a) : \varphi \in \Delta(\mathcal{A})\} \supset \mathrm{Sp}_{\mathcal{A}}(a)$, which, together with (1.13), implies that $\{\varphi(a) : \varphi \in \Delta(\mathcal{A})\} = \mathrm{Sp}_{\mathcal{A}}(a)$. □

Theorem 1.35. *Let $(\mathcal{A}, \|\cdot\|)$ be a commutative unital Banach algebra. Then, $\Delta(\mathcal{A})$ is compact.*

Proof. If it is shown that

(1) for every $\varphi \in \Delta(\mathcal{A})$, $\|\varphi\|_{\mathrm{op}} \leqslant 1$; and

(2) $\Delta(\mathcal{A})$ is a closed subset of the dual space $\mathcal{A}^*$ (with respect to the w^*-topology on $\mathcal{A}^*$),

[19] See Remark 1.29.

then the result will follow from the Banach-Alaoglu Theorem (which states that $\{\varphi \in \mathcal{A}^* : \|\varphi\|_{\text{op}} \leqslant 1\}$ is compact in the w^*-topology on $\mathcal{A}^*$, see Subsection B.5).

(1). Let $\varphi \in \Delta(\mathcal{A})$ be arbitrary. According to Theorems 1.34 and 1.21, for every $a \in \mathcal{A}$, one has $|\varphi(a)| \leqslant \rho(a) \leqslant \|a\|$. Therefore,

$$\|\varphi\|_{\text{op}} = \sup\left\{\frac{|\varphi(a)|}{\|a\|} : a \neq 0\right\} \leqslant \sup\left\{\frac{\|a\|}{\|a\|} : a \neq 0\right\} = 1.$$

(2). For every $\varphi \in \mathcal{A}^*$, $a \in \mathcal{A}$ and $\varepsilon > 0$, define the set

$$U(\varphi, a, \varepsilon) = \{\psi \in \mathcal{A}^* : |\psi(a) - \varphi(a)| < \varepsilon\}.$$

The collection of every such set forms a sub-basis on $\mathcal{A}^*$ for the w^*-topology. Consequently, it can be proved that $\mathcal{A}^* \setminus \Delta(\mathcal{A})$ is open by demonstrating that every element in $\mathcal{A}^* \setminus \Delta(\mathcal{A})$ has a neighbourhood of the form

$$N = \bigcap_{i \leqslant n} U(\varphi_i, a_i, \varepsilon_i), \qquad \varphi_i \in \mathcal{A}^*,\ a_i \in \mathcal{A},\ \varepsilon_i > 0,\ i = 1, 2, ..., n \tag{1.14}$$

such that $N \subset \mathcal{A}^* \setminus \Delta(\mathcal{A})$. Let $\varphi \in \mathcal{A}^* \setminus \Delta(\mathcal{A})$. Then, it must be the case that $\varphi(a) = 0$ for all $a \in \mathcal{A}$, or that there exists $a, b \in \mathcal{A}$ such that $\varphi(ab) \neq \varphi(a)\varphi(b)$. Consider the two cases:

(2.1) Suppose that $\varphi(a) = 0$ for every $a \in \mathcal{A}$. For every $\psi \in U(\varphi, 1_{\mathcal{A}}, \frac{1}{2})$, one has $|\psi(1_{\mathcal{A}}) - \varphi(1_{\mathcal{A}})| = |\psi(1_{\mathcal{A}})| < \frac{1}{2}$, hence $\psi(1_{\mathcal{A}}) \neq 1$. According to Remark 1.29, this implies that $\psi \in \mathcal{A}^* \setminus \Delta(\mathcal{A})$. Therefore, $\varphi \in U(\varphi, 1_{\mathcal{A}}, \frac{1}{2}) \subset \mathcal{A}^* \setminus \Delta(\mathcal{A})$.

(2.2) Suppose that there exists $a, b \in \mathcal{A}$ such that $\varphi(ab) \neq \varphi(a)\varphi(b)$. Then, there exists $\varepsilon > 0$ such that $|\varphi(ab) - \varphi(a)\varphi(b)| > \varepsilon$. Let

$$\psi \in U\left(\varphi, a, \frac{\varepsilon}{3\|b\|}\right) \cap U\left(\varphi, b, \frac{\varepsilon}{3|\varphi(a)|}\right) \cap U\left(\varphi, ab, \frac{\varepsilon}{3}\right),$$

or, in other words, $|\psi(a) - \varphi(a)| < \frac{\varepsilon}{3\|b\|}$, $|\psi(b) - \varphi(b)| < \frac{\varepsilon}{3|\varphi(a)|}$ and $|\psi(ab) - \varphi(ab)| < \frac{\varepsilon}{3}$. Suppose by contradiction that $\psi \in \Delta(\mathcal{A})$. Then, it must be the case that $\psi(ab) = \psi(a)\psi(b)$, and according to Theorems 1.34 and 1.21, one has $\psi(b) \leqslant \|b\|$. Consequently,

$$\begin{aligned}
&|\varphi(ab) - \varphi(a)\varphi(b)| \\
=&|\varphi(ab) - \psi(ab) + \psi(ab) - \psi(b)\varphi(a) + \varphi(a)\psi(b) - \varphi(a)\varphi(b)| \\
=&|\varphi(ab) - \psi(ab) + \psi(a)\psi(b) - \psi(b)\varphi(a) + \varphi(a)\psi(b) - \varphi(a)\varphi(b)| \\
\leqslant&|\varphi(ab) - \psi(ab)| + |\psi(b)||\psi(a) - \varphi(a)| + |\varphi(a)||\psi(b) - \varphi(b)| \\
<&\frac{\varepsilon}{3} + \|b\|\frac{\varepsilon}{3\|b\|} + |\varphi(a)|\frac{\varepsilon}{3|\varphi(a)|} \\
=&\varepsilon,
\end{aligned}$$

which is clearly a contradiction. Therefore, φ is contained in a neighbourhood contained in $\mathcal{A}^* \setminus \Delta(\mathcal{A})$.

Thus, in all cases, there exists a neighbourhood N of φ of the form (1.14) such that $N \subset \mathcal{A}^* \setminus \Delta(\mathcal{A})$, which concludes the proof of the theorem. □

Remark 1.36. By combining Theorem 1.35 and Example 1.11, it can be noticed that, given a commutative unital Banach algebra $(\mathcal{A}, \|\cdot\|)$, the space $\mathrm{C}\big(\Delta(\mathcal{A}), \mathbb{C}\big)$ equipped with the usual

function operations, the complex conjugate and the uniform norm $\|\cdot\|_\infty$ is a C^*-algebra. Furthermore, the Gelfand transform provides a natural map from $\mathcal{A}$ to $\mathrm{C}\big(\Delta(\mathcal{A}),\mathbb{C}\big)$. In fact, this provides the key to the proof of the Gelfand-Naimark Theorem for commutative unital C^*-algebras, which we are now ready to tackle.

Proof of the Gelfand-Naimark Theorem III. Let $(\mathcal{A},*,\|\cdot\|)$ be a commutative unital C^*-algebra, and let $\mathcal{G}$ be the Gelfand transform on $\mathcal{A}$. As hinted in Remark 1.36, it will be shown that $\mathcal{G}$ is in fact an isometric $*$-isomorphism from $\mathcal{A}$ to $\mathrm{C}\big(\Delta(\mathcal{A}),\mathbb{C}\big)$. The proof is separated into four steps:

(1) $\mathcal{G}$ is an algebra homomorphism;

(2) $\mathcal{G}$ is a $*$-homomorphism;

(3) $\|\mathcal{G}(a)\|_\infty = \|a\|$ for every $a\in\mathcal{A}$; and

(4) $\mathcal{G}$ is bijective.

(1). Let $a,b\in\mathcal{A}$ and $k\in\mathbb{C}$ be arbitrary. Then, for every $\varphi\in\Delta(\mathcal{A})$, one has

$$\mathcal{G}(a+b)(\varphi)=\varphi(a+b)=\varphi(a)+\varphi(b)=\mathcal{G}(a)(\varphi)+\mathcal{G}(b)(\varphi),$$

$$\mathcal{G}(ab)(\varphi)=\varphi(ab)=\varphi(a)\varphi(b)=\mathcal{G}(a)(\varphi)\mathcal{G}(b)(\varphi)$$

and

$$\mathcal{G}(ka)(\varphi)=\varphi(ka)=k\varphi(a)=k\mathcal{G}(a)(\varphi),$$

hence $\mathcal{G}(a+b)=\mathcal{G}(a)+\mathcal{G}(b)$, $\mathcal{G}(ab)=\mathcal{G}(a)\mathcal{G}(b)$ and $\mathcal{G}(ka)=k\mathcal{G}(a)$, which proves that $\mathcal{G}$ is an algebra homomorphism.

(2). Let $a\in\mathcal{A}$ be arbitrary, $\mathrm{Re}(a)=\frac{a+a^*}{2}$, and $\mathrm{Im}(a)=\frac{a-a^*}{2\mathrm{i}}$. Then, $\mathrm{Re}(a)$ and $\mathrm{Im}(a)$ are self-adjoint, $a=\mathrm{Re}(a)+\mathrm{i}\cdot\mathrm{Im}(a)$, and $a^*=\mathrm{Re}(a)-\mathrm{i}\cdot\mathrm{Im}(a)$. For any fixed $\varphi\in\Delta(\mathcal{A})$, one has

$$\mathcal{G}(a)(\varphi)=\varphi(a)=\varphi\big(\mathrm{Re}(a)+\mathrm{i}\cdot\mathrm{Im}(a)\big)=\varphi\big(\mathrm{Re}(a)\big)+\mathrm{i}\cdot\varphi\big(\mathrm{Im}(a)\big),$$

and

$$\mathcal{G}(a^*)(\varphi)=\varphi(a^*)=\varphi\big(\mathrm{Re}(a)-\mathrm{i}\cdot\mathrm{Im}(a)\big)=\varphi\big(\mathrm{Re}(a)\big)-\mathrm{i}\cdot\varphi\big(\mathrm{Im}(a)\big).$$

According to Theorem 1.34, $\varphi\big(\mathrm{Re}(a)\big)$ and $\varphi\big(\mathrm{Im}(a)\big)$ are contained in the spectrum of $\mathrm{Re}(a)$ and $\mathrm{Im}(a)$ respectively. Furthermore, according to Lemma 1.27, the spectrum of any self-adjoint element is contained in $\mathbb{R}$, and thus it follows that $\varphi\big(\mathrm{Re}(a)\big),\varphi\big(\mathrm{Im}(a)\big)\in\mathbb{R}$. Therefore, $\mathcal{G}(a^*)(\varphi)=\overline{\mathcal{G}(a)(\varphi)}$, which proves that $\mathcal{G}$ is a $*$-homomorphism.

(3). Knowing that for self-adjoint elements $a \in \mathcal{A}$, one has $\|a\| = \rho(a)$ (see Corollary 1.26), then for every $a \in \mathcal{A}$ (self-adjoint or not),

$$\begin{aligned}
\|\mathcal{G}(a)\|_\infty^2 &= \|\overline{\mathcal{G}(a)}\mathcal{G}(a)\|_\infty && \text{(see Remark 1.36)}\\
&= \|\mathcal{G}(a^*)\mathcal{G}(a)\|_\infty && \text{(part (2) of this theorem)}\\
&= \|\mathcal{G}(a^*a)\|_\infty && \text{(part (1) of this theorem)}\\
&= \sup\{|\varphi(a^*a)| : \varphi \in \Delta(\mathcal{A})\}\\
&= \rho(a^*a) && (\mathcal{A} \text{ is commutative; Theorem 1.34})\\
&= \|a^*a\| && (a^*a \text{ is self-adjoint})\\
&= \|a\|^2, && ((\mathcal{A}, *, \|\cdot\|) \text{ is a } C^*\text{-algebra})
\end{aligned}$$

from which the result follows.

(4). For every $a, b \in \mathcal{A}$, part (3) of this theorem implies that $\|a - b\| = \|\mathcal{G}(a) - \mathcal{G}(b)\|_\infty$, hence $\mathcal{G}$ is injective.

Up to this point, it was established that $\mathcal{G}$ is an injective $*$-homomorphism that preserves the norms from $\mathcal{A}$ to $\mathrm{C}(\Delta(\mathcal{A}), \mathbb{C})$. Therefore, $\mathcal{G}$ is an isometric $*$-isomorphism from $\mathcal{A}$ to its image $\mathcal{G}(\mathcal{A}) \subset \mathrm{C}(\Delta(\mathcal{A}), \mathbb{C})$, which implies in particular that $\mathcal{G}(\mathcal{A})$ is a C^*-algebra. Given that $\mathrm{C}(\Delta(\mathcal{A}), \mathbb{C})$ is a C^*-algebra, and thus complete, it follows that $\mathcal{G}(\mathcal{A})$ is closed, since complete subsets of complete metric spaces are closed. Therefore, if $\mathcal{G}(\mathcal{A})$ contains all the constant functions from $\Delta(\mathcal{A})$ to $\mathbb{C}$ and separates points in $\Delta(\mathcal{A})$, it will follow from the Stone-Weierstrass Approximation Theorem that $\mathcal{G}(\mathcal{A}) = \mathrm{C}(\Delta(\mathcal{A}), \mathbb{C})$, which will conclude the proof that $\mathcal{G}$ is bijective.

For every $\varphi, \psi \in \Delta(\mathcal{A})$ such that $\varphi \neq \psi$, there exists $a \in \mathcal{A}$ such that $\varphi(a) \neq \psi(a)$, which implies that $\mathcal{G}(a)(\varphi) \neq \mathcal{G}(a)(\psi)$, hence $\mathcal{G}(\mathcal{A})$ separates points.

Let $\lambda \in \mathbb{C}$ be arbitrary and $f_\lambda : \Delta(\mathcal{A}) \to \mathbb{C}$ be the constant function which takes the value $f_\lambda(\varphi) = \lambda$ for every $\varphi \in \Delta(\mathcal{A})$. According to Remark 1.29, $\varphi(1_\mathcal{A}) = 1$ for every $\varphi \in \Delta(\mathcal{A})$. Consequently, for any $\varphi \in \Delta(\mathcal{A})$, one has $\mathcal{G}(\lambda \cdot 1_\mathcal{A})(\varphi) = \varphi(\lambda \cdot 1_\mathcal{A}) = \lambda\varphi(1_\mathcal{A}) = \lambda = f_\lambda(\varphi)$, which implies that $\mathcal{G}(\lambda \cdot 1_\mathcal{A}) = f_\lambda$. Thus, $\mathcal{G}(\mathcal{A})$ contains every constant function. □

As a first application of the Gelfand-Naimark Theorem, we prove a result sometimes referred to as the *Spectral Permanence Theorem*, which essentially states that the spectrum of an element a is the same in every C^*-algebra that contains it.

Theorem 1.37 (Spectral Permanence)**.** *Let $(\mathcal{A}, *, \|\cdot\|)$ be a unital C^*-algebra and $a \in \mathcal{A}$ be an arbitrary element. For every unital C^*-subalgebra $\mathcal{B} \subset \mathcal{A}$ that contains a, one has* $\mathrm{Sp}_\mathcal{A}(a) = \mathrm{Sp}_\mathcal{B}(a)$.

Proof. Given that every element which is invertible in $\mathcal{B}$ is also invertible in $\mathcal{A}$, it clearly is the case that $\mathrm{Sp}_\mathcal{B}(a) \subset \mathrm{Sp}_\mathcal{A}(a)$. If the inclusion

$$\mathrm{Inv}(\mathcal{A}) \cap \mathcal{B} \subset \mathrm{Inv}(\mathcal{B}), \tag{1.15}$$

holds, then $\mathrm{Sp}_\mathcal{B}(a) \supset \mathrm{Sp}_\mathcal{A}(a)$ will follow and the result will be proved. Furthermore, it can be shown that it is only necessary to prove that (1.15) holds when restricted to self-adjoint elements: Suppose that (1.15) is true for every self-adjoint element, and let $a \in \mathrm{Inv}(\mathcal{A}) \cap \mathcal{B}$ be arbitrary. Then, a^* is invertible (with inverse $(a^*)^{-1} = (a^{-1})^*$), which implies that a^*a is

invertible (with inverse $(a^*a)^{-1} = a^{-1}(a^*)^{-1}$). Since $\mathcal{B}$ is a C^*-algebra, $a^*a \in \mathrm{Inv}(\mathcal{A}) \cap \mathcal{B}$, and since a^*a is self-adjoint, it follows that $a^*a \in \mathrm{Inv}(\mathcal{B})$, that is, $(a^*a)^{-1} \in \mathcal{B}$. Now,

$$(a^*a)^{-1}a^*a = a^{-1}(a^*)^{-1}a^*a = 1_{\mathcal{A}} = aa^{-1}(a^*)^{-1}a^* = a(a^*a)^{-1}a^*,$$

which implies that $a^{-1} = (a^*a)^{-1}a^*$. Since $a^*, (a^*a)^{-1} \in \mathcal{B}$, it follows that $a^{-1} \in \mathcal{B}$, as desired.

We now prove that (1.15) holds when restricted to self-adjoint elements. Let $a \in \mathrm{Inv}(\mathcal{A}) \cap \mathcal{B}$ be a self-adjoint element, and define $\mathcal{C} \coloneqq C^*(1_{\mathcal{A}}, a, a^{-1})$ and $\mathcal{D} \coloneqq C^*(1_{\mathcal{A}}, a)$. If $\mathcal{C} = \mathcal{D}$, then the result will follow, since the equality $\mathcal{C} = \mathcal{D}$ implies that a unital C^*-subalgebra of $\mathcal{A}$ contains a if and only if it contains its inverse a^{-1}. According to Proposition 1.17, the fact that a is self-adjoint and therefore normal implies that $\mathcal{C}$ and $\mathcal{D}$ are commutative. Therefore, it follows from the Gelfand-Naimark Theorem III that the Gelfand transform $\mathcal{G}$ on $\mathcal{C}$ is an isometric $*$-isomorphism onto $\mathrm{C}\big(\Delta(\mathcal{C}), \mathbb{C}\big)$. Consequently, it clearly is the case that $\mathrm{C}\big(\Delta(\mathcal{C}), \mathbb{C}\big) = C^*\big(\mathbf{1}_{\Delta(\mathcal{C})}, \mathcal{G}(a), \mathcal{G}(a)^{-1}\big)$ and that the preimage of $C^*\big(\mathbf{1}_{\Delta(\mathcal{C})}, \mathcal{G}(a)\big)$ through $\mathcal{G}$ is $\mathcal{D}$. If $C^*\big(\mathbf{1}_{\Delta(\mathcal{C})}, \mathcal{G}(a)\big)$ contains the constant functions from $\Delta(\mathcal{C})$ to $\mathbb{C}$ and separates points in $\Delta(\mathcal{C})$, it will follow from the Stone-Weierstrass Approximation Theorem that $\mathcal{G}(\mathcal{C}) = C^*\big(\mathbf{1}_{\Delta(\mathcal{C})}, \mathcal{G}(a)\big)$, which will prove that $\mathcal{C} = \mathcal{D}$ since $\mathcal{G}^{-1}$ is bijective.

Let $f : \Delta(\mathcal{C}) \to \mathbb{C}$ be a constant function. Then, f is a scalar multiple of $\mathbf{1}_{\Delta(\mathcal{C})}$, which is clearly contained in $C^*\big(\mathbf{1}_{\Delta(\mathcal{C})}, \mathcal{G}(a)\big)$.

Let $\varphi, \psi \in \Delta(\mathcal{C})$ such that $\varphi \neq \psi$, that is, there exists $b \in \mathcal{C}$ such that $\varphi(b) \neq \psi(b)$. Therefore, $\mathcal{G}(b)(\varphi) = \varphi(b) \neq \psi(b) = \mathcal{G}(b)(\psi)$. We claim that this implies that $\mathcal{G}(a)(\varphi) \neq \mathcal{G}(a)(\psi)$, which would in turn imply that $C^*(\mathbf{1}_{\Delta(\mathcal{C})}, \mathcal{G}(a))$ separates points. We proceed by contraposition: Let $\varphi, \psi \in \Delta(\mathcal{C})$ be such that $\mathcal{G}(a)(\varphi) = \varphi(a) = \psi(a) = \mathcal{G}(a)(\psi)$. Then,

$$\mathcal{G}(a)^{-1}(\varphi) = \frac{1}{\varphi(a)} = \frac{1}{\psi(a)} = \mathcal{G}(a)^{-1}(\psi).$$

Consequently, for every $f \in *\text{-alg}(\mathbf{1}_{\Delta(\mathcal{C})}, \mathcal{G}(a), \mathcal{G}(a)^{-1})$, it clearly is the case that $f(\varphi) = f(\psi)$. According to the definition of a C^*-algebra generated by a collection of elements, the fact that $b \in \mathcal{C}$ (and hence $\mathcal{G}(b) \in C^*(\mathbf{1}_{\Delta(\mathcal{C})}, \mathcal{G}(a), \mathcal{G}(a)^{-1})$) implies that there exists a sequence of functions $\{f_n : n \in \mathbb{N}\}$ in $*\text{-alg}(\mathbf{1}_{\Delta(\mathcal{C})}, \mathcal{G}(a), \mathcal{G}(a)^{-1})$ such that $\|f_n - \mathcal{G}(b)\|_\infty \to 0$. Therefore, for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that if $n \geqslant N$, then $\|f_n - \mathcal{G}(b)\|_\infty < \frac{\varepsilon}{2}$. Thus, if $n \geqslant N$, one has

$$\begin{aligned}
&|\mathcal{G}(b)(\varphi) - \mathcal{G}(b)(\psi)| \\
=&|\mathcal{G}(b)(\varphi) - f_n(\varphi) + f_n(\varphi) - \mathcal{G}(b)(\psi)| \\
=&|\mathcal{G}(b)(\varphi) - f_n(\varphi) + f_n(\psi) - f(\psi)| && (f_n \in *\text{-alg}(\mathbf{1}_{\Delta(\mathcal{C})}, \mathcal{G}(a), \mathcal{G}(a)^{-1})) \\
\leqslant&|\mathcal{G}(b)(\varphi) - f_n(\varphi)| + |f_n(\psi) - \mathcal{G}(b)(\psi)| \\
<&\varepsilon.
\end{aligned}$$

Taking $\varepsilon \to 0$ gives the desired result. $\square$

1.4. Functional Calculus with Continuous Functions

Let a be a normal element in a unital C^*-algebra $(\mathcal{A}, *, \|\cdot\|)$. Given a commutative polynomial $P \in \mathbb{C}[z, \overline{z}]$, that is, there exists $\alpha_i \in \mathbb{C}$ and $k(i), l(i) \in \mathbb{N} \cup \{0\}$ $(i \leqslant n)$ such that

$$P(z) = \sum_{i=1}^{n} \alpha_i z^{k(i)} \overline{z}^{l(i)},$$

there is a very natural way in which an element $P(a) \in \mathcal{A}$ may be defined, namely, as

$$P(a) \coloneqq \sum_{i=1}^{n} \alpha_i a^{k(i)} (a^*)^{l(i)}$$

(using the convention $a^0 = 1_{\mathcal{A}}$). According to this definition, one easily notices that the element $P(a)$ is contained in the commutative (since a is normal) unital C^*-algebra $\mathcal{B} = C^*(1_{\mathcal{A}}, a)$ generated by a. Let $\mathcal{G}$ be the Gelfand transform on $\mathcal{B}$. Then, the composition $P \circ \mathcal{G}(a)$ defines a function from $\mathrm{Sp}_{\mathcal{B}}(a)$ to $\mathrm{Sp}_{\mathcal{B}}(a)$. Furthermore, since $\mathcal{G}$ is a $*$-algebra homomorphism, it follows that $P \circ \mathcal{G}(a) = \mathcal{G}\big(P(a)\big)$. Consequently, it clearly is the case that $\mathcal{G}^{-1} \circ P \circ \mathcal{G}$ maps a to $P(a)$ in $\mathcal{B}$. This leads to the following proposition, which provides a means of calculating the norm of any polynomial evaluated at a normal element (provided its spectrum is well-known).

Proposition 1.38. *Let $(\mathcal{A}, *, \|\cdot\|)$ be a unital C^*-algebra and let $a \in \mathcal{A}$ be normal. For any polynomial $P \in \mathbb{C}[z, \overline{z}]$, one has*

$$\|P(a)\| = \sup\big\{|P(\lambda)| : \lambda \in \mathrm{Sp}_{\mathcal{A}}(a)\big\}.$$

Proof. Let $\mathcal{G}$ be the Gelfand transform on $\mathcal{B} = C^*(1_{\mathcal{A}}, a)$. As explained in the previous paragraph, one has $\|P(a)\| = \|\mathcal{G}^{-1} \circ P \circ \mathcal{G}(a)\|$. Thus, according to the Gelfand-Naimark Theorem III,

$$\begin{aligned} \|P(a)\| = \|P \circ \mathcal{G}(a)\|_\infty &= \sup\Big\{\big|P\big(\varphi(a)\big)\big| : \varphi \in \Delta(\mathcal{B})\Big\} \\ &= \sup\big\{|P(\lambda)| : \lambda \in \mathrm{Sp}_{\mathcal{B}}(a)\big\}. \qquad \text{(Theorem 1.34)} \end{aligned}$$

According to the spectral permanence (Theorem 1.37), $\mathrm{Sp}_{\mathcal{A}}(a)$ is equal to $\mathrm{Sp}_{\mathcal{B}}(a)$, which concludes the proof of the proposition. □

Let a be a normal element in a unital C^*-algebra $(\mathcal{A}, *, \|\cdot\|)$. Given that $\mathrm{Sp}_{\mathcal{A}}(a)$ is compact, it follows from the Stone-Weierstrass Approximation Theorem that any continuous map $f : \mathrm{Sp}_{\mathcal{A}}(a) \to \mathbb{C}$ can be uniformly approximated by a sequence of polynomial functions $\{P_n : n \in \mathbb{N}\} \subset \mathbb{C}[z, \overline{z}]$ from $\mathrm{Sp}_{\mathcal{A}}(a)$ to $\mathbb{C}$. Letting $\mathcal{G}$ denote the Gelfand transform on $\mathcal{B} = C^*(1_{\mathcal{A}}, a)$, it is then clear that for each $n \in \mathbb{N}$, one has

$$\begin{aligned} \|P_n(a) - \mathcal{G}^{-1} \circ f \circ \mathcal{G}(a)\| &= \|\mathcal{G}^{-1} \circ P_n \circ \mathcal{G}(a) - \mathcal{G}^{-1} \circ f \circ \mathcal{G}(a)\| \\ &= \|P_n \circ \mathcal{G}(a) - f \circ \mathcal{G}(a)\|_\infty \\ &= \sup\Big\{\big|P_n\big(\varphi(a)\big) - f\big(\varphi(a)\big)\big| : \varphi \in \Delta(\mathcal{B})\Big\} \\ &= \|P_n - f\|_\infty. \end{aligned}$$

Consequently, a natural way of defining the element $f(a)$ is the following:

$$f(a) \coloneqq \lim_{n \to \infty} P_n(a) = \mathcal{G}^{-1} \circ f \circ \mathcal{G}(a).$$

According to this definition, it is clear that $f(a) \in \mathcal{B}$ since $\mathcal{B}$ is complete, and by using the same argument as in Proposition 1.38, one obtains that $\|f(a)\| = \sup\{|f(\lambda)| : \lambda \in \mathrm{Sp}_{\mathcal{A}}(a)\}$.

Definition 1.39. Let $(\mathcal{A}, *, \|\cdot\|)$ be a unital C^*-algebra and $a \in \mathcal{A}$ be normal. The **functional calculus with continuous functions** on a (or simply the functional calculus on a), denoted Φ_a, is defined as the function

$$\Phi_a(f) = \mathcal{G}^{-1} \circ f \circ \mathcal{G}(a) = f(a), \qquad f \in \mathrm{C}\big(\mathrm{Sp}_{\mathcal{A}}(a), \mathbb{C}\big).$$

The elementary properties of the functional calculus with continuous functions, some of which have already been established, are summarized in the following theorem.

Functional Calculus Theorem. *Let $(\mathcal{A}, *, \|\cdot\|)$ be a unital C^*-algebra and $a \in \mathcal{A}$ be normal. The functional calculus with continuous function Φ_a is a $*$-algebra homomorphism from $\mathrm{C}\big(\mathrm{Sp}_{\mathcal{A}}(a), \mathbb{C}\big)$ to $C^*(1_{\mathcal{A}}, a)$. Furthermore, it is the only $*$-algebra homomorphism having the following properties:*

(1) $\Phi_a\left(\mathbf{1}_{\mathrm{Sp}_{\mathcal{A}}(a)}\right) = 1_{\mathcal{A}}$*;*

(2) for every $f \in \mathrm{C}\big(\mathrm{Sp}_{\mathcal{A}}(a), \mathbb{C}\big)$, $\|\Phi_a(f)\| = \sup\{|f(\lambda)| : \lambda \in \mathrm{Sp}_{\mathcal{A}}(a)\}$*; and*

(3) the identity function $\mathrm{Id} : \mathrm{Sp}_{\mathcal{A}}(a) \to \mathbb{C}$ *defined as* $\mathrm{Id}(\lambda) = \lambda$ *for every* $\lambda \in \mathrm{Sp}_{\mathcal{A}}(a)$ *is such that* $\Phi_a(\mathrm{Id}) = a$*.*

Proof. The proof that Φ_a is a $*$-homomorphism is a matter of straightforward computations, and the fact that Φ_a satisfies conditions (1), (2) and (3) was already established in the previous paragraphs. Thus, it only remains to prove the uniqueness of Φ_a.

Let $\Theta_a : \mathrm{C}\big(\mathrm{Sp}_{\mathcal{A}}(a), \mathbb{C}\big) \to \mathbb{C}$ be a $*$-homomorphism that satisfies conditions (1), (2) and (3) of this theorem. Let $P \in \mathbb{C}[z, \overline{z}]$ an arbitrary polynomial. The fact that Φ_a and Θ_a are $*$-homomorphisms and condition (3) of this theorem imply that

$$\Phi_a(P) = \Theta_a(P). \tag{1.16}$$

Let $f \in \mathrm{C}\big(\mathrm{Sp}_{\mathcal{A}}(a), \mathbb{C}\big)$ and $\varepsilon > 0$ be arbitrary. According to the Stone-Weierstrass Approximation Theorem, there exists a polynomial function $P_\varepsilon \in \mathbb{C}[z, \overline{z}]$ such that

$$\sup\big\{\, |P_\varepsilon(\lambda) - f(\lambda)| : \lambda \in \mathrm{Sp}_{\mathcal{A}}(a)\big\} < \frac{\varepsilon}{2}.$$

Given condition (2) of this theorem, equation (1.16), and the fact that Φ_a and Θ_a are $*$-homomorphisms, one has

$$\begin{aligned}
\|\Phi_a(f) - \Theta_a(f)\| &= \|\Phi_a(f) - \Phi_a(P_\varepsilon) + \Phi_a(P_\varepsilon) - \Theta_a(f)\| \\
&\leqslant \|\Phi_a(f) - \Phi_a(P_\varepsilon)\| + \|\Phi_a(P_\varepsilon) - \Theta_a(f)\| \\
&= \|\Phi_a(f) - \Phi_a(P_\varepsilon)\| + \|\Theta_a(P_\varepsilon) - \Theta_a(f)\| \\
&= \|\Phi_a(f - P_\varepsilon)\| + \|\Theta_a(P_\varepsilon - f)\| \\
&= 2\sup\big\{\, |P_\varepsilon(\lambda) - f(\lambda)| : \lambda \in \mathrm{Sp}_{\mathcal{A}}(a)\big\} \\
&< \varepsilon.
\end{aligned}$$

Taking $\varepsilon \to 0$ implies that $\Phi_a = \Theta_a$ since f was arbitrary, which completes the proof of the theorem. □

Remark 1.40. Let a be a normal element in a C^*-algebra $(\mathcal{A}, *, \|\cdot\|)$, and let $f : A \subset \mathbb{C} \to \mathbb{C}$ be analytic on A and such that $\mathrm{Sp}_{\mathcal{A}}(a) \subset A$, that is, there exists $\{\alpha_n : n \in \mathbb{N} \cup \{0\}\} \subset \mathbb{C}$ and $z_0 \in \mathbb{C}$ such that

$$f(z) = \sum_{n=0}^{\infty} \alpha_n (z - z_0)^n$$

for every $z \in A$. Given that power series converge uniformly on any compact set contained in their radius of convergence, it follows from part (2) of the Functional Calculus Theorem that

$$f(a) = \sum_{n=0}^{\infty} \alpha_n (a - z_0)^n.$$

Remark 1.41. Up to this point, two notations for the element

$$\mathcal{G}^{-1} \circ f \circ \mathcal{G}(a) \tag{1.17}$$

for a normal and $f \in \mathrm{C}\big(\mathrm{Sp}_{\mathcal{A}}(a), \mathbb{C}\big)$ have been introduced, namely, $f(a)$ and $\Phi_a(f)$. In most instances, it will be more convenient to think of (1.17) as the function f evaluated at a, hence $f(a)$ will be used. Φ_a will typically be used in technical proofs (such as in the proof of the Spectral Mapping Theorem, which follows), where it is more useful to think of (1.17) as a $*$-homomorphism evaluated at continuous functions.

Spectral Mapping Theorem ([**NS06**] Theorem 3.4)**.** *Let $(\mathcal{A}, *, \|\cdot\|)$ be a unital C^*-algebra, a be a normal element of $\mathcal{A}$, Φ_a be the functional calculus on a, and f be a continuous function from $\mathrm{Sp}_{\mathcal{A}}(a)$ to $\mathbb{C}$. Then,*

$$\mathrm{Sp}_{\mathcal{A}}\big(f(a)\big) = \mathrm{Sp}_{\mathcal{A}}\big(\Phi_a(f)\big) = f\big(\mathrm{Sp}_{\mathcal{A}}(a)\big) = \big\{f(\lambda) : \lambda \in \mathrm{Sp}_{\mathcal{A}}(a)\big\}.$$

Proof. If the statement

$$g(a) \in \mathrm{Inv}(\mathcal{A}) \text{ if and only if } 0 \notin g\big(\mathrm{Sp}_{\mathcal{A}}(a)\big) \tag{1.18}$$

holds for every $g \in \mathrm{C}\big(\mathrm{Sp}_{\mathcal{A}}(a), \mathbb{C}\big)$, then the result will follow. Indeed, if one defines

$$g(\lambda) = \lambda' - f(\lambda), \qquad \lambda \in \mathrm{Sp}_{\mathcal{A}}(a)$$

for some $\lambda' \in \mathbb{C}$, then the statement (1.18) implies that

$$g(a) = \Phi_a(g) = \lambda' 1_{\mathcal{A}} - \Phi_a(f) = \lambda' 1_{\mathcal{A}} - f(a)$$

is invertible (i.e., λ' is not an element of the spectrum of $f(a)$) if and only if $g(\lambda) = \lambda' - f(\lambda) \neq 0$ for every $\lambda \in \mathrm{Sp}_{\mathcal{A}}(a)$ (i.e., $\lambda' \notin f\big(\mathrm{Sp}_{\mathcal{A}}(a)\big)$).

Let $g \in \mathrm{C}\big(\mathrm{Sp}_{\mathcal{A}}(a), \mathbb{C}\big)$ be arbitrary. Suppose that $0 \notin g\big(\mathrm{Sp}_{\mathcal{A}}(a)\big)$. Then, the inverse function

$$\begin{array}{rccc} g^{-1}: & \mathrm{Sp}_{\mathcal{A}}(a) & \to & \mathbb{C} \\ & \lambda & \mapsto & g(\lambda)^{-1}. \end{array}$$

can be defined on all of $\mathrm{Sp}_{\mathcal{A}}(a)$. Since the functional calculus is a $*$-homomorphism, it follows that $\Phi_a(g)\Phi_a(g^{-1}) = \Phi_a(gg^{-1}) = 1_{\mathcal{A}}$, and thus $\Phi_a(g) = g(a)$ is invertible.

Let $g \in \mathrm{C}\big(\mathrm{Sp}_{\mathcal{A}}(a), \mathbb{C}\big)$ be such that $\Phi_a(g)$ is invertible. Suppose by contradiction that $0 \in g\big(\mathrm{Sp}_{\mathcal{A}}(a)\big)$. Then, there exists $\lambda_0 \in \mathrm{Sp}_{\mathcal{A}}(a)$ such that $g(\lambda_0) = 0$. Choose $k > 0$ such

that $k > \|\Phi_a(g)^{-1}\|$. Since g is continuous, there exists $\delta > 0$ such that if $|\lambda - \lambda_0| < \delta$, then $|g(\lambda) - g(\lambda_0)| = |g(\lambda)| < \frac{1}{k}$. Define $h : \mathrm{Sp}_{\mathcal{A}}(a) \to \mathbb{C}$ as follows:

$$h(\lambda) = k \max\left\{0, 1 - \frac{2|\lambda - \lambda_0|}{\delta}\right\}, \qquad \lambda \in \mathrm{Sp}_{\mathcal{A}}(a).$$

One notices that

(1) $h(\lambda_0) = k$;

(2) if $|\lambda - \lambda_0| \geqslant \frac{\delta}{2}$, then $h(\lambda) = 0$;

(3) if $0 < |\lambda - \lambda_0| < \frac{\delta}{2}$, then $h(\lambda) = k\left(1 - \frac{2|\lambda - \lambda_0|}{\delta}\right) \leqslant k$; and

(4) $h \in \mathrm{C}\big(\mathrm{Sp}_{\mathcal{A}}(a), \mathbb{C}\big)$.

Furthermore, for every $\lambda \in \mathrm{Sp}_{\mathcal{A}}(a)$, one has

$$gh(\lambda) = g(\lambda)h(\lambda) = \begin{cases} 0 & \text{if } \lambda = \lambda_0 \text{ or } |\lambda - \lambda_0| \geqslant \frac{\delta}{2}; \\ g(\lambda)k(1 - \frac{2|\lambda - \lambda_0|}{\delta}) & \text{if } |\lambda - \lambda_0| < \frac{\delta}{2}. \end{cases}$$

If $|\lambda - \lambda_0| < \frac{\delta}{2}$, then $g(\lambda) < \frac{1}{k}$, and thus $gh(\lambda) < \left(1 - \frac{2|\lambda - \lambda_0|}{\delta}\right) \leqslant 1$. According to the Functional Calculus Theorem and the fact that $\Phi_a(g)$ is invertible, it then follows that

$$\begin{aligned} k &= \sup\{|h(\lambda)| : \lambda \in \mathrm{Sp}_{\mathcal{A}}(a)\} = \|\Phi_a(h)\| = \|\Phi_a(g)^{-1}(\Phi_a(g)\Phi_a(h))\| \\ &\leqslant \|\Phi_a(g)^{-1}\|\|\Phi_a(gh)\| = \|\Phi_a(g)^{-1}\| \sup\{|gh(\lambda)| : \lambda \in \mathrm{Sp}_{\mathcal{A}}(a)\} < k \cdot 1, \end{aligned}$$

which is a contradiction. □

Remark 1.42. Continuing on the line of thought in Remark 1.41, the Spectral Mapping Theorem provides yet another interpretation of (1.17), namely, $f(a)$ can be seen as the element one obtains when *transforming* the spectrum of a through the continuous function f. This point of view is most consistent with algebras of continuous functions or random variables, where the spectrum represents the image of functions.

Remark 1.43. An obvious Corollary to the Spectral Mapping Theorem (more precisely, an immediate consequence of part (2) of the Functional Calculus Theorem and the Spectral Mapping Theorem) is that Corollary 1.26 may be extended to any continuous function evaluated at a normal element, that is, for every normal element a in a C^*-algebra $(\mathcal{A}, *, \|\cdot\|)$ and continuous function $f : \mathrm{Sp}_{\mathcal{A}}(a) \to \mathbb{C}$, one has

$$\|f(a)\| = \sup\left\{\lambda : \lambda \in \mathrm{Sp}_{\mathcal{A}}\big(f(a)\big)\right\} = \rho\big(f(a)\big). \tag{1.19}$$

1.5. The Positive Cone

In the algebra of complex numbers, one calls an element $z \in \mathbb{C}$ **positive** if $z \in [0, \infty)$. Then, one can define an order on the self-adjoint elements of $\mathbb{C}$ (that is, $\mathbb{R}$) using positive elements as follows:

$$z \geqslant w \text{ if and only if } z - w \text{ is positive,}$$

and the positive elements can be characterized as

$$z \in \mathbb{C} \text{ is positive if and only if } z = w\overline{w} \text{ for some } w \in \mathbb{C}.$$

In this section, it is shown that a similar concept and characterization exists in every unital C^*-algebra.

Definition 1.44. Let $(\mathcal{A}, *)$ be a unital $*$-algebra. An element $a \in \mathcal{A}$ is called **positive** if it is self-adjoint and its spectrum is contained in $[0, \infty)$. The set of positive elements in $\mathcal{A}$ will be denoted by $\mathcal{A}^+$, and it will be called the **positive cone** of $\mathcal{A}$.

Lemma 1.45. *Let $(\mathcal{A}, *, \|\cdot\|)$ be a unital C^*-algebra, $a \in \mathcal{A}$ be self-adjoint, and $k > 0$ such that $\|a\| \leqslant k$. Then, a is positive if and only if $\|k1_{\mathcal{A}} - a\| \leqslant k$.*

Proof. According to part (1) of Theorem 1.21 and Lemma 1.27, one has $\mathrm{Sp}_{\mathcal{A}}(a) \subset [-k, k]$. Consider the function $f \in \mathrm{C}\big(\mathrm{Sp}_{\mathcal{A}}(a), \mathbb{C}\big)$ defined as

$$\begin{array}{rccc} f: & \mathrm{Sp}_{\mathcal{A}}(a) & \to & \mathbb{C} \\ & \lambda & \mapsto & k - \lambda \end{array}$$

(that is, $f(a) = k1_{\mathcal{A}} - a$). The Functional Calculus Theorem then implies that $\|k1_{\mathcal{A}} - a\| = \sup\big\{|k - \lambda| : \lambda \in \mathrm{Sp}_{\mathcal{A}}(a)\big\}$. Therefore, $\mathrm{Sp}_{\mathcal{A}}(a) \subset [0, k]$ (i.e., a is positive) if and only if $\|k1_{\mathcal{A}} - a\| \leqslant k$. □

Proposition 1.46. *Let $(\mathcal{A}, *, \|\cdot\|)$ be a unital C^*-algebra. Then, $\mathcal{A}^+$ is a* **cone** *in $\mathcal{A}$, that is,*

(1) If $k \geqslant 0$ and $a \in \mathcal{A}^+$, then $ka \in \mathcal{A}^+$; and

(2) If $a, b \in \mathcal{A}^+$, then $a + b \in \mathcal{A}^+$.

Proof. (1). Let $k \geqslant 0$ and $a \in \mathcal{A}^+$. Then, $(ka)^* = \overline{k}a^* = ka$, thus, ka is self-adjoint. If $k = 0$, then $ka = 0$, which implies that $\mathrm{Sp}_{\mathcal{A}}(a) = \{0\} \subset [0, \infty)$. If $k > 0$, then, according to Remark 1.19, $\mathrm{Sp}_{\mathcal{A}}(a) \subset [0, \infty)$ implies that $\mathrm{Sp}_{\mathcal{A}}(ka) = k\mathrm{Sp}_{\mathcal{A}}(a) \subset [0, \infty)$. In both cases, $ka \in \mathcal{A}^+$.

(2). Let $a, b \in \mathcal{A}^+$, and choose $k_1, k_2 > 0$ such that $\|a\| \leqslant k_1$ and $\|b\| \leqslant k_2$. Then, by Lemma 1.45, we know that $\|k_1 1_{\mathcal{A}} - a\| \leqslant k_1$ and $\|k_2 1_{\mathcal{A}} - a\| \leqslant k_2$. By the triangle inequality, we see that $\|a + b\| \leqslant \|a\| + \|b\| \leqslant k_1 + k_2$. Consequently, since

$$\|(k_1 + k_2)1_{\mathcal{A}} - (a + b)\| \leqslant \|k_1 1_{\mathcal{A}} - a\| + \|k_2 1_{\mathcal{A}} - b\| \leqslant k_1 + k_2,$$

we have by Lemma 1.45 that $a + b$ is positive. □

Lemma 1.47. *Let $(\mathcal{A}, *)$ be a unital $*$-algebra and $a \in \mathcal{A}$ be such that $a^*a \in -\mathcal{A}^+$.*[20] *Then, $a = 0$.*

Proof. It will first be necessary to establish that

$$\mathrm{Sp}_{\mathcal{A}}(a^*a) \setminus \{0\} = \mathrm{Sp}_{\mathcal{A}}(aa^*) \setminus \{0\}.$$

[20] We use $-\mathcal{A}^+$ to denote $\{-a : a \in \mathcal{A}^+\}$.

Let $b, c \in \mathcal{A}$ be arbitrary. Suppose that $(1_{\mathcal{A}} - bc)$ is invertible, and define $w = (1_{\mathcal{A}} - bc)^{-1}$ Then,

$$\begin{aligned}(1_{\mathcal{A}} + cwb)(1_{\mathcal{A}} - cb) &= 1_{\mathcal{A}} - cb + cwb - cwbcb \\ &= 1_{\mathcal{A}} - (cb - cwb + cwbcb) \\ &= 1_{\mathcal{A}} - c(1_{\mathcal{A}} - w + wbc)b \\ &= 1_{\mathcal{A}} - c(1_{\mathcal{A}} - w(1_{\mathcal{A}} - bc))b \\ &= 1_{\mathcal{A}} - c(1_{\mathcal{A}} - 1_{\mathcal{A}})b \qquad\qquad (\text{since } w = (1_{\mathcal{A}} - bc)^{-1}) \\ &= 1_{\mathcal{A}},\end{aligned}$$

and similarly, $(1_{\mathcal{A}} - cb)(1_{\mathcal{A}} + cwb) = 1_{\mathcal{A}}$. Therefore, $(1_{\mathcal{A}} - cb)$ is invertible and its inverse is given by $(1_{\mathcal{A}} - cb)^{-1} = (1_{\mathcal{A}} + cwb)$. Clearly, this implies that for every $b, c \in \mathcal{A}$, $(1_{\mathcal{A}} - bc)$ is invertible if and only if $(1_{\mathcal{A}} - cb)$ is invertible. Consequently,

$$\begin{aligned}\lambda \in \mathrm{Sp}_{\mathcal{A}}(a^*a) \setminus \{0\} &\iff (\lambda \cdot 1_{\mathcal{A}} - a^*a) \notin \mathrm{Inv}(\mathcal{A}) \text{ and } \lambda \neq 0 \\ &\iff \lambda(1_{\mathcal{A}} - \lambda^{-1}a^*a) \notin \mathrm{Inv}(\mathcal{A}) \\ &\iff (1_{\mathcal{A}} - \lambda^{-1}a^*a) \notin \mathrm{Inv}(\mathcal{A}) \\ &\iff (1_{\mathcal{A}} - a(\lambda^{-1}a^*)) \notin \mathrm{Inv}(\mathcal{A}) \\ &\iff (\lambda \cdot 1_{\mathcal{A}} - aa^*)) \notin \mathrm{Inv}(\mathcal{A}) \text{ and } \lambda \neq 0 \\ &\iff \lambda \in \mathrm{Sp}_{\mathcal{A}}(aa^*) \setminus \{0\}.\end{aligned}$$

Suppose that $a \in \mathcal{A}$ is such that $a^*a \in -\mathcal{A}^+$, that is, $-a^*a \in \mathcal{A}^+$. Then, Remark 1.19 implies that

$$\mathrm{Sp}_{\mathcal{A}}(a^*a) = -\mathrm{Sp}_{\mathcal{A}}(-a^*a) \subset (-\infty, 0]. \tag{1.20}$$

Since $\mathrm{Sp}_{\mathcal{A}}(a^*a) \setminus \{0\} = \mathrm{Sp}_{\mathcal{A}}(aa^*) \setminus \{0\}$, this implies that

$$\mathrm{Sp}_{\mathcal{A}}(aa^*) = \mathrm{Sp}_{\mathcal{A}}(a^*a) \subset (-\infty, 0]. \tag{1.21}$$

Since aa^* is self-adjoint, this implies that $-aa^* \in \mathcal{A}^+$. By Proposition 1.46, this implies that $-a^*a - aa^* \in \mathcal{A}^+$, hence

$$\mathrm{Sp}_{\mathcal{A}}(a^*a + aa^*) \subset (-\infty, 0]. \tag{1.22}$$

Write $a = \mathrm{Re}(a) + \mathrm{i} \cdot \mathrm{Im}(a)$, where $\mathrm{Re}(a) = \frac{a+a^*}{2}$ and $\mathrm{Im}(a) = \frac{(a-a^*)}{2\mathrm{i}}$ are self-adjoint. Notice that

$$\begin{aligned}2\mathrm{Re}(a)^2 + 2\mathrm{Im}(a)^2 &= \frac{a^2 + aa^* + a^*a + (a^*)^2}{2} - \frac{a^2 - aa^* - a^*a + (a^*)^2}{2} \\ &= aa^* + a^*a.\end{aligned}$$

Since $\mathrm{Re}(a)$ and $\mathrm{Im}(a)$ are self-adjoint (and thus normal), the functional calculus with continuous functions can be applied on them. Given that the function $f \in \mathrm{C}\big(\mathrm{Sp}_{\mathcal{A}}(a), \mathbb{C}\big)$ defined as $f(\lambda) = \lambda^2$ for every $\lambda \in \mathrm{Sp}_{\mathcal{A}}(a)$ has a nonnegative image, it follows from spectral mapping theorem that

$$\mathrm{Sp}_{\mathcal{A}}\big(\mathrm{Re}(a)^2\big) = \mathrm{Sp}_{\mathcal{A}}\Big(f\big(\mathrm{Re}(a)\big)\Big) = f\Big(\mathrm{Sp}_{\mathcal{A}}\big(\mathrm{Re}(a)\big)\Big) \subset [0, \infty)$$

and similarly, $\mathrm{Sp}_{\mathcal{A}}\big(\mathrm{Im}(a)^2\big) \subset [0,\infty)$. Therefore, according to Proposition 1.46, $2\mathrm{Re}(a)^2 + 2\mathrm{Im}(a)^2 = a^*a + aa^*$ is positive. Combining this fact with equation (1.22) yields $\mathrm{Sp}_{\mathcal{A}}(a^*a + aa^*) = \{0\}$, which, according to Corollary 1.26, implies that $\|a^*a + aa^*\| = 0$, and hence $a^*a + aa^* = 0$.

Now, $a^*a = -aa^*$ implies by Remark 1.19 and equation (1.21) that

$$\mathrm{Sp}_{\mathcal{A}}(a^*a) = \mathrm{Sp}_{\mathcal{A}}(-aa^*) = -\mathrm{Sp}_{\mathcal{A}}(aa^*) \subset [0,\infty).$$

Combining this with equation (1.20), one obtains $\mathrm{Sp}_{\mathcal{A}}(a^*a) = \{0\}$, and thus since a^*a is self-adjoint and $\mathcal{A}$ is a C^*-algebra,

$$\|a\|^2 = \|a^*a\| = \sup\{|\lambda| : \lambda \in \mathrm{Sp}_{\mathcal{A}}(a^*a)\} = 0,$$

which implies that $a = 0$. □

Theorem 1.48. *Let $(\mathcal{A}, *, \|\cdot\|)$ be a unital C^*-algebra. Then, the following statements are equivalent*

(1) $b \in \mathcal{A}^+$;

(2) $b = a^*a$ for some $a \in \mathcal{A}$.

Proof. (1) $\Longrightarrow$ (2). Let $b \in \mathcal{A}^+$. Since $\mathrm{Sp}_{\mathcal{A}}(b) \subset [0,\infty)$, the square root function $\sqrt{\cdot}$ is well defined on $\mathrm{Sp}_{\mathcal{A}}(b)$. Furthermore, this function is clearly continuous. Given that $\sqrt{\lambda} \in \mathbb{R}$ for every $\lambda \in \mathrm{Sp}_{\mathcal{A}}(b)$, it clearly is the case that $\overline{\sqrt{\cdot}} = \sqrt{\cdot}$. Letting Φ_b denote the functional calculus homomorphism on b (see Definition 1.39), one has

$$\Phi_b(\sqrt{\cdot})^*\Phi_b(\sqrt{\cdot}) = \Phi_b(\overline{\sqrt{\cdot}})\Phi_b(\sqrt{\cdot}) = \Phi_b(\sqrt{\cdot}^2) = \Phi_b(\mathrm{Id}) = b,$$

as desired.

(2) $\Longrightarrow$ (1). Let $b = a^*a$, where $a \in \mathcal{A}$. Then, b is clearly self-adjoint, and thus it only remains to prove that $\mathrm{Sp}_{\mathcal{A}}(b) \subset [0,\infty)$. Lemma 1.27 implies that $\mathrm{Sp}_{\mathcal{A}}(b)$ is a subset of $\mathbb{R}$. Thus, the functions $f, g : \mathrm{Sp}_{\mathcal{A}}(b) \to \mathbb{C}$ defined as $f(\lambda) = \max\{0,\lambda\}$ and $g(\lambda) = \max\{0,-\lambda\}$ for every $\lambda \in \mathrm{Sp}_{\mathcal{A}}(b)$ are well defined and continuous. Let Φ_b be the functional calculus homomorphism on b. According to the Spectral Mapping Theorem, $\mathrm{Sp}_{\mathcal{A}}\big(\Phi_b(f)\big) = f\big(\mathrm{Sp}_{\mathcal{A}}(b)\big)$ and $\mathrm{Sp}_{\mathcal{A}}\big(\Phi_b(g)\big) = g\big(\mathrm{Sp}_{\mathcal{A}}(b)\big)$, and thus the fact that $f(\lambda), g(\lambda) \geqslant 0$ for every $\lambda \in \mathrm{Sp}_{\mathcal{A}}(b)$ implies that $\mathrm{Sp}_{\mathcal{A}}(\Phi_b(f)), \mathrm{Sp}_{\mathcal{A}}(\Phi_b(g)) \subset [0,\infty)$. Moreover, since the images of f and g are contained in $\mathbb{R}$, f and g are self-adjoint, hence $\Phi_b(f)$ and $\Phi_b(g)$ are also self-adjoint. Consequently,

$$\Phi_b(f), \Phi_b(g) \in \mathcal{A}^+. \tag{1.23}$$

Given the definition of f and g, one clearly has

$$\Phi_b(f) - \Phi_b(g) = \Phi_b(f - g) = \Phi_b(\mathrm{Id}) = b \tag{1.24}$$

and, since $fg(\lambda) = gf(\lambda) = 0$ for every λ, it follows from Remark 1.6 that

$$\Phi_b(fg) = \Phi_b(gf) = 0_{\mathcal{A}}. \tag{1.25}$$

Combining equations (1.24) and (1.25) yields

$$\begin{aligned}(a\Phi_b(g))^*(a\Phi_b(g)) &= \Phi_b(g)a^*a\Phi_b(g) = \Phi_b(g)b\Phi_b(g) \\ &= \Phi_b(g)(\Phi_b(f) - \Phi_b(g))\Phi_b(g) = -\Phi_b(g)^3.\end{aligned} \tag{1.26}$$

According to (1.23) and the Spectral Mapping Theorem applied to the functional calculus on $\Phi_b(g)$ with the map $\lambda \mapsto \lambda^3$, it is clear that

$$\mathrm{Sp}_{\mathcal{A}}(\Phi_b(g)^3) = \mathrm{Sp}_{\mathcal{A}}(\Phi_b(g))^3 \subset [0, \infty),$$

and since $(\Phi_b(g)^3)^* = \Phi_b(g)^3$, then $\Phi_b(g)^3 \in \mathcal{A}^+$. According to equation (1.26), this implies that $(a\Phi_b(g))^*(a\Phi_b(g)) \in -\mathcal{A}^+$, and thus it follows from Lemma 1.47 that $a\Phi_b(g) = 0$.

$a\Phi_b(g) = 0$ implies that $(a\Phi_b(g))^*(a\Phi_b(g)) = -\Phi_b(g)^3 = 0$. Therefore, the Spectral Mapping Theorem implies that

$$\{0\} = \mathrm{Sp}_{\mathcal{A}}(\Phi_b(g)^3) = \mathrm{Sp}_{\mathcal{A}}(\Phi_b(g))^3,$$

which in turn implies that $\Phi_b(g) = 0$ since $\Phi_b(g)$ is self-adjoint (see Corollary 1.26). From equation (1.24), it can then be concluded that $b = \Phi_b(f)$, which implies that $b \in \mathcal{A}^+$ by (1.23). □

Remark 1.49. Notice that, a consequence of the first portion of the proof of Theorem 1.48 is that, much like it is the case in $\mathbb{R} \subset \mathbb{C}$, for every element a in the positive cone $\mathcal{A}^+$ of a unital C^*-algebra algebra $(\mathcal{A}, *, \|\cdot\|)$, there exists a **square root** $\sqrt{a} \in \mathcal{A}^+$ such that $\sqrt{a}^2 = a$.

Chapter 2

Fundamental Notions in Noncommutative Probability

The main sources that inspired the writing of this section are [**NS06**] Lectures 1 to 4, [**VDN92**] Chapter 2, [**Ta12**] Section 2.5.1, and [**AGZ09**] Section 5.2.

2.1. Noncommutative Probability Spaces

2.1.1. Basic Definitions and Examples. As explained in the introduction of this thesis, one of the most familiar models for an abstract algebra of random variables consists of the algebra of every complex-valued random variable with finite moments $L^{\infty-}$ equipped with the expected value $\mathbf{E}$. More precisely, given a probability space $(\Omega, \mathscr{A}, P)$, let

$$L^{\infty-} \coloneqq \bigcap_{1 \leqslant p < \infty} L^p(\Omega, \mathscr{A}, P),$$

and define the map $\mathbf{E} : L^{\infty-} \to \mathbb{C}$ as

$$\mathbf{E}[X] \coloneqq \int_\Omega X \, \mathrm{d}P, \qquad\qquad X \in L^{\infty-}.$$

When equipped with the usual operations of pointwise addition and multiplication, it is clear that $L^{\infty-}$ forms an algebra over the field of complex numbers. Furthermore, it can be noticed that $\mathbf{E} : L^{\infty-} \to \mathbb{C}$ is a linear functional, and since $P(\Omega) = 1$, $\mathbf{E}$ is *normalized*, meaning that $\mathbf{E}[\mathbf{1}_\Omega] = 1$. With these basic properties in mind, the general definition of a space of noncommutative random variables can be introduced by simply removing the assumption that the considered algebra is commutative:

Definition 2.1. Let $\mathcal{A}$ be a unital algebra over $\mathbb{C}$ (whose unit is denoted by $1_\mathcal{A}$) equipped with a function $\varphi : \mathcal{A} \to \mathbb{C}$. If φ is a linear functional and $\varphi(1_\mathcal{A}) = 1$, then the pair $(\mathcal{A}, \varphi)$ is called a **noncommutative probability space**, and the elements of $\mathcal{A}$ are called **noncommutative random variables** (or simply random variables when it is clear that the variables are not

necessarily commutative).[1] Furthermore, if $\varphi(ab) = \varphi(ba)$ for every $a, b \in \mathcal{A}$, then φ is called a **trace**.

Given that the expected value $\mathbf{E}$ on $L^{\infty-}$ is defined as an integral with respect to a probability measure P, it has many important properties other than those mentioned up to now, some of which have a nontrivial probabilistic interpretation. Among these properties are the facts that $\mathbf{E}$ is *positive* and *faithful*, which respectively mean that if a random variable $X \in L^{\infty-}$ takes values in $[0, \infty)$, then its expected value is nonnegative, and that the expected value of $|X|$ is exactly zero if and only if $X = 0$. At first glance, it may seem difficult to find analogs of these properties in general noncommutative $*$-algebras such that $\mathcal{M}_n(\mathbb{C})$ (the space of $n \times n$ matrices with complex entries), where it makes no sense to say that a matrix X *takes values* in $[0, \infty)$. However, recall that

(1) if one defines the operator $\bar{\cdot} : L^{\infty-} \to L^{\infty-}$ as $\overline{X}(\omega) = \overline{X(\omega)}$ for every random variable X and $\omega \in \Omega$ (where $\overline{X(\omega)}$ denotes the complex conjugate of $X(\omega)$), then $(L^{\infty-}, \bar{\cdot})$ is a $*$-algebra; and

(2) the spectrum $\mathrm{Sp}_{L^{\infty-}}(X) = \{\lambda \in \mathbb{C} : \lambda - X \notin \mathrm{Inv}(L^{\infty-})\}$ of a random variable $X \in L^{\infty-}$ is the image of X (that is, the set $\{\lambda \in \mathbb{C} : X(\omega) = \lambda \text{ for some } \omega \in \Omega\}$);

and thus the positivity and faithfulness of $\mathbf{E}$ can be expressed in the language of general $*$-algebras as follows: For every random variable X in the positive cone of $L^{\infty-}$ that is, $X = Y\overline{Y}$ for some $Y \in L^{\infty-}$ (or, equivalently, $X = \overline{X}$ and $\mathrm{Sp}_{L^{\infty-}}(X) = X(\Omega) \subset [0, \infty)$), one has that $\mathbf{E}[X] \geqslant 0$, and $\mathbf{E}[Y\overline{Y}] = 0$ if and only if $Y = 0$. Therefore, if a noncommutative probability space $(\mathcal{A}, \varphi)$ is such that $\mathcal{A}$ is a $*$-algebra, then more conditions may be imposed on the functional φ, leading us to the following definition:

Definition 2.2. Let $(\mathcal{A}, \varphi)$ be a noncommutative probability space. If $\mathcal{A}$ is a $*$-algebra and φ is positive (that is, $\varphi(a^*a) \geqslant 0$ for every $a \in \mathcal{A}$), then $(\mathcal{A}, *, \varphi)$ is called a **$*$-probability space**. Furthermore, the functional φ is said to be **faithful** if $\varphi(a^*a) = 0$ if and only if $a = 0$.

Now that noncommutative probability spaces and $*$-probability spaces have been defined, the following definition of a C^*-probability spaces is very natural:

Definition 2.3. Let $(\mathcal{A}, *, \varphi)$ be a $*$-probability space, and let $\|\cdot\| : \mathcal{A} \to \mathbb{C}$ be a norm on $\mathcal{A}$ that makes $(\mathcal{A}, *, \|\cdot\|)$ a C^*-algebra. Then, the quadruple $(\mathcal{A}, *, \|\cdot\|, \varphi)$ is called a ***C*-probability space***.

Example 2.4. Let G be a group with identity e. One defines the **group algebra** of G, denoted $\mathbb{C}G$, as the set of formal sums of the form

$$\sum_{i=1}^{n} \alpha_{g_i} g_i, \qquad \alpha_{g_1}, \ldots, \alpha_{g_n} \in \mathbb{C} \text{ and } g_1, \ldots, g_n \in G$$

[1] To avoid confusion, it should be noted that the algebra $\mathcal{A}$ in a noncommutative probability space $(\mathcal{A}, \varphi)$ is not necessarily noncommutative. The terminology *noncommutative probability space* is instead used to put emphasis on the fact that the algebras random variables we are *allowed* to consider are not necessarily commutative. Consequently, some authors propose the (perhaps more accurate) appellations **potentially noncommutative probability space** for $(\mathcal{A}, \varphi)$ and **potentially noncommutative random variables** for elements of $\mathcal{A}$.

together with the natural addition

$$\left(\sum_{i=1}^{n}\alpha_{g_i}g_i\right)+\left(\sum_{i=n+1}^{m}\alpha_{g_i}g_i\right)=\sum_{i=1}^{m}\alpha_{g_i}g_i$$

and multiplication

$$\left(\sum_{i=1}^{n}\alpha_{g_i}g_i\right)\cdot\left(\sum_{j=1}^{m}\beta_{h_j}h_j\right)=\sum_{i=1}^{n}\sum_{j=1}^{m}\alpha_{g_i}\beta_{h_j}g_ih_j.$$

When equipped with the $*$-operation defined as

$$\left(\sum_{i=1}^{n}\alpha_{g_i}g_i\right)^*=\sum_{i=1}^{n}\overline{\alpha_{g_i}}g_i^{-1},$$

$(\mathbb{C}G,*)$ is a unital $*$-algebra. Furthermore, if $(\mathbb{C}G,*)$ is equipped with the linear functional $\tau_e:\mathbb{C}G\to\mathbb{C}$ (which is called the **canonical trace**) defined as

$$\tau_e\left(\sum_{i=1}^{n}\alpha_{g_i}g_i\right)=\alpha_e,$$

then $(\mathbb{C}G,*,\tau_e)$ is a $*$-probability space, and it can easily be shown that τ_e is a faithful trace.

Remark 2.5. Notice that, in a group algebra $(\mathbb{C}G,*)$, every element of the group G is unitary, and an element of G is self-adjoint if and only if it is equal to the identity. Conversely, given a collection $(u_i:i\in I)$ of unitary random variables (that is, $u_iu_i^*=u_i^*u_i=1_{\mathcal{A}}$) in an arbitrary $*$-probability space $(\mathcal{A},*,\varphi)$, the set

$$H=\left\{u_{i(1)}^{n(1)}\cdots u_{i(t)}^{n(t)}:i(1),\ldots,i(t)\in I\text{ and }n(1),\ldots,n(t)\in\{0,1,*\}\right\}$$

together with the product operation it inherits from $\mathcal{A}$ forms a group (where the inverse is given by the $*$ operation), and the unital $*$-algebra generated by the u_i is trivially isomorphic to the group algebra $(\mathbb{C}H,*)$. It is not necessarily the case, however, that $(\mathcal{A},*,\varphi)$ and $(\mathbb{C}H,*,\tau_e)$ are isomorphic as $*$-probability spaces[2] for an arbitrary faithful trace φ.

Example 2.6. Let $n\in\mathbb{N}$ be fixed, and let $(\Omega,\mathscr{A},P)$ be a probability space.

(1) Consider the algebra $\mathcal{M}_n(\mathbb{C})$ of $n\times n$ matrices with entries in $\mathbb{C}$ equipped with the $*$-operation of conjugate transpose and the normalized trace $\frac{1}{n}\mathrm{tr}$. Then, the triple $\left(\mathcal{M}_n(\mathbb{C}),*,\frac{1}{n}\mathrm{tr}\right)$ is a $*$-probability space, and $\frac{1}{n}\mathrm{tr}$ is a faithful trace.

(2) Consider the algebra $\mathcal{M}_n(L^{\infty-})=\mathcal{M}_n(\mathbb{C})\otimes L^{\infty-}$ of $n\times n$ matrices whose entries are random variables in $L^{\infty-}$. If $\mathcal{M}_n(L^{\infty-})$ is equipped with the $*$-operation defined in such a way that, given a random matrix $X=(X_{ij}:1\leqslant i,j\leqslant n)$, one has

$$X^*=(Y_{ij}:1\leqslant i,j\leqslant n),\qquad\qquad(\text{where }Y_{ij}=\overline{X_{ji}})$$

and the functional $\frac{1}{n}\mathrm{tr}\otimes\mathbf{E}$ defined as

$$\left(\tfrac{1}{n}\mathrm{tr}\otimes\mathbf{E}\right)(X)=\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}[X_{ii}]$$

then $\left(\mathcal{M}_n(L^{\infty-}),*,\frac{1}{n}\mathrm{tr}\otimes\mathbf{E}\right)$ is a $*$-probability space, and $\frac{1}{n}\mathrm{tr}\otimes\mathbf{E}$ is a faithful trace.

[2] See Definition 2.19.

The procedure by which $\mathcal{M}_n(\mathbb{C})$ and $L^{\infty-}$ were combined to form the $*$-probability space of random matrices in Example 2.6 (2) can be extended to any pair of $*$-algebras. This gives the following proposition, whose proof in full generality will be provided in the next chapter (see Theorem 3.13).

Proposition 2.7. *Let $(\mathcal{A}_1, *_1, \varphi_1)$ and $(\mathcal{A}_2, *_2, \varphi_2)$ be $*$-probability spaces. Define the triple $(\mathcal{A}, *, \varphi)$ as follows: Let $\mathcal{A} = \mathcal{A}_1 \otimes \mathcal{A}_2$, let $*$ be defined as*

$$(a_1 \otimes a_2)^* = a_1^{*_1} \otimes a_2^{*_2}, \qquad a_1 \in \mathcal{A}_1 \text{ and } a_2 \in \mathcal{A}_2,$$

and let φ be defined as

$$\varphi(a_1 \otimes a_2) = \varphi_1(a_1)\varphi_2(a_2), \qquad a_1 \in \mathcal{A}_1 \text{ and } a_2 \in \mathcal{A}_2.$$

Then $(\mathcal{A}, *, \varphi)$ is a $*$-probability space, and if φ_1 and φ_2 are both traces (resp. faithful), then φ is also a trace (resp. faithful).

Example 2.8. The two *canonical*[3] examples of unital C^*-algebras can be equipped with *natural* functionals that make them C^*-probability spaces:

(1) Let Ω be a compact topological space, let $\mathscr{B}(\Omega)$ be the Borel σ-algebra on Ω, and let μ be a regular probability measure on Ω. Let $\mathrm{C}(\Omega, \mathbb{C})$ be the space of continuous functions from Ω to $\mathbb{C}$ (i.e., continuous $\mathbb{C}$-valued random variables defined on the probability space $\big(\Omega, \mathscr{B}(\mathbb{C}), \mu\big)$). If $\mathrm{C}(\Omega, \mathbb{C})$ is equipped with the usual pointwise addition and multiplication, the $*$-operation of complex conjugation $\bar{\cdot}$, the expected value functional $\mathbf{E}$ (with respect to μ), and the supremum norm

$$\|X\|_\infty = \sup_{\omega \in \Omega} |X(\omega)|,$$

then $\big(\mathrm{C}(\Omega, \mathbb{C}), \bar{\cdot}, \|\cdot\|_\infty, \mathbf{E}\big)$ is a C^*-probability space.

(2) Let $\big(H, \langle \cdot, \cdot \rangle\big)$ be a Hilbert space, and let $B(H)$ be the space of bounded linear operators on H. Then, as seen in Example 1.9, $\big(B(H), *, \|\cdot\|_{\mathrm{op}}\big)$ is a C^*-algebra. If $B(H)$ is equipped with the linear functional φ_{a_0} defined as

$$\varphi_{a_0}(T) = \langle T a_0, a_0 \rangle, \qquad T \in B(H)$$

where a_0 is a unit vector in H, then $\big(B(H), *, \|\cdot\|_{\mathrm{op}}, \varphi_{a_0}\big)$ is a C^*-probability space.

2.1.2. Generalized Moments. Now that the formal framework for noncommutative probability spaces has been introduced, it is worth taking some time to discuss the intuitive interpretation of this framework before exploring its mathematical properties.

Definition 2.9. Let $(\mathcal{A}, \varphi)$ be a noncommutative probability space and let $a \in \mathcal{A}$ be a random variable. The **centering** of a, denoted a°, is defined as the variable $a - \varphi(a) \cdot 1_{\mathcal{A}}$. The variable a is said to be **centered** if $a^\circ = a$, or, equivalently, if $\varphi(a) = 0$; and let $\mathcal{A}^\circ \coloneqq \big\{a \in \mathcal{A} : \varphi(a) = 0\big\}$ denote the collection of centered variables in $\mathcal{A}$.

According to the above definition, given a noncommutative probability space $(\mathcal{A}, \varphi)$, the functional φ provides a canonical way of decomposing $\mathcal{A}$ into a direct sum $\mathcal{A}_\varphi = \mathbb{C}1_{\mathcal{A}} \oplus \mathcal{A}^\circ$ in which every element $a \in \mathcal{A}$ is represented as $a = \varphi(a) \cdot 1_{\mathcal{A}} \oplus a^\circ$.

[3] Canonical in the sense that, in light of the Gelfand-Naimark Theorems, any unital C^*-algebra is isomorphic to one of these two examples.

Example 2.10. Let $(\Omega, \mathscr{A}, P)$ be a probability space and $(L^{\infty-}, \mathbf{E})$ be the noncommutative probability space of $\mathbb{C}$-valued random variables with finite moments of all orders. Then, $L^{\infty-}_{\mathbf{E}}$ yields the decomposition $\mathbb{C}\mathbf{1}_\Omega \oplus (L^{\infty-})^\circ$, where $\mathbb{C}\mathbf{1}_\Omega$ is the collection of almost-surely deterministic variables, and $(L^{\infty-})^\circ$ is the collection of random variables in $L^{\infty-}$ with mean zero. Given a random variable $X \in L^{\infty-}$, the decomposition $X = \mathbf{E}[X] \cdot \mathbf{1}_\Omega \oplus X^\circ$ *separates* X into

(1) a deterministic part $\mathbf{E}[X] \cdot \mathbf{1}_\Omega$ in which the expected value of X (i.e., the *information*[4] $\mathbf{E}$ *extracts* from X) is encoded; and

(2) a random part $X - \mathbf{E}[X] \cdot 1_\Omega$, which is a centered random variable about which $\mathbf{E}$ does not provide a priori information (indeed, given an arbitrary nondeterministic random variable X, the expected value alone gives no information about how X *fluctuates* about $\mathbf{E}[X]$).

Let

$$\mathcal{B} \coloneqq *\text{-alg}(\mathbf{1}_\Omega, X) = \text{Span}_{\mathbb{C}}\Big(\big\{X^m\overline{X}^n : m, n \in \mathbb{N} \cup \{0\}\big\}\Big) \subset L^{\infty-}$$

denote the unital $*$-algebra generated by the variable X. Then, being that $\mathbf{E}$ is linear, the total information that $\mathbf{E}$ can extract from $\mathcal{B}$ is completely determined by the moments of X, which can be obtained from the deterministic parts of the following variables:

$$X^m\overline{X}^n = \mathbf{E}\big[X^m\overline{X}^n\big] \cdot \mathbf{1}_\Omega - \big(X^m\overline{X}^n\big)^\circ, \qquad m, n \in \mathbb{N}.$$

Remark 2.11. In light of the above example, the following appears to be an adequate intuitive interpretation of the current framework for noncommutative probability: Let $(\mathcal{A}, \varphi)$ be a noncommutative probability space and $a \in \mathcal{A}$ be a noncommutative random variable. Then,

(1) the unit element $1_\mathcal{A} \in \mathcal{A}$ is regarded as the generic deterministic element of $\mathcal{A}$, that is, an element whose behaviour is completely known, and every other deterministic element of $\mathcal{A}$ is a $\mathbb{C}$-multiple of $1_\mathcal{A}$;

(2) the information that φ extracts from a can be encoded in the deterministic variable $\varphi(a) \cdot 1_\mathcal{A}$;

(3) the **moments** of a are defined as the quantities

$$\alpha_n \coloneqq \varphi(a^n), \qquad n \in \mathbb{N},$$

and being that φ is linear, the moments of a completely determine all of the information that φ extracts from

$$\text{alg}(1_\mathcal{A}, a) = \text{Span}_{\mathbb{C}}\Big(\big\{a^n : n \in \mathbb{N} \cup \{0\}\big\}\Big),$$

that is, the unital algebra generated by a; and similarly

(4) the **$*$-moments** of a are defined as the quantities

$$\varphi\left(a^{n(1)} \cdots a^{n(t)}\right), \qquad n(1), \ldots, n(t) \in \{1, *\}$$

and being that φ is linear, the $*$-moments of a completely determine all of the information that φ extracts from

$$*\text{-alg}(1_\mathcal{A}, a) = \text{Span}_{\mathbb{C}}\Big(\big\{a^{n(1)} \cdots a^{n(t)} : n(1), \ldots, n(t) \in \{0, 1, *\}\big\}\Big),$$

[4]Note that this usage of the term *information* is strictly for the purpose of aiding intuition and is not linked with other mathematical definitions of information that occur in information theory such as entropy.

that is, the unital $*$-algebra generated by a.

Example 2.12. Let $\big(\mathcal{M}_n(\mathbb{C}), *, \frac{1}{n}\mathrm{tr}\big)$ be the $*$-probability space of $n\times n$ matrices with $\mathbb{C}$-valued entries (as defined in Example 2.6), and let

$$A = (A_{ij} : 1 \leqslant i, j \leqslant n) \in \mathcal{M}_n(\mathbb{C})$$

be an arbitrary matrix. In this case, the information $\frac{1}{n}\mathrm{tr}$ extracts from A, i.e., the average of the diagonal entries $m_A = \frac{1}{n}(A_{11} + \cdots + A_{nn})$, is encoded in the following multiple of the identity matrix:

$$m_A \cdot \mathrm{Id}_n = \begin{bmatrix} m_A & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & m_A \end{bmatrix}.$$

Constant multiples of the identity matrix $\lambda\cdot\mathrm{Id}_n$ are considered deterministic in this $*$-probability space, as a knowledge of the normalized trace of such matrices is enough to uniquely determine the value of every entry. The centering

$$A^\circ = A - m_A \cdot \mathrm{Id}_n = \begin{bmatrix} A_{11} - m_A & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} - m_A & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} - m_A \end{bmatrix}$$

of A is considered to be *random* (or, perhaps more accurately in this specific case, uncertain or undetermined) in this $*$-algebra, as the entries above and below the diagonal as well as the deviations $A_{ii} - m_A$ between the diagonal entries and the average m_A are a priori not determined by the normalized trace.

Definition 2.13. Let $(\mathcal{A}, *, \varphi)$ be a $*$-probability space and $a \in \mathcal{A}$ be arbitrary. The quantity $\varphi\big((a - \varphi(a)\cdot 1_{\mathcal{A}})(a - \varphi(a)\cdot 1_{\mathcal{A}})^*\big)$ (denoted $\mathbf{Var}[a]$) is called the **variance** of a.

Remark 2.14. If the functional φ is faithful, then the variance retains some of its properties of probabilistic significance: For example, it immediately follows from the faithfulness of φ that $\mathbf{Var}[a] = 0$ if and only if $a = \varphi(a)\cdot 1_{\mathcal{A}}$, that is, if and only if a is a constant multiple of the unit $1_{\mathcal{A}}$ (i.e., deterministic).

2.1.3. Elementary Properties.

Proposition 2.15. *Let $(\mathcal{A}, *, \varphi)$ be a $*$-probability space. Then, φ is self-adjoint, that is, for every $a \in \mathcal{A}$, one has $\varphi(a^*) = \overline{\varphi(a)}$.*

Proof. Let $a \in \mathcal{A}$ be self-adjoint. Then, if one defines $a_0 \coloneqq \frac{a+1_{\mathcal{A}}}{2}$ and $a_1 \coloneqq \frac{a-1_{\mathcal{A}}}{2}$, it follows that $a = a_0^* a_0 - a_1^* a_1$. Therefore, the positivity of φ implies that $\varphi(a) \in \mathbb{R}$. Let $a \in \mathcal{A}$ be arbitrary. If one defines $\mathrm{Re}(a) \coloneqq \frac{a+a^*}{2}$ and $\mathrm{Im}(a) \coloneqq \frac{a-a^*}{2\mathrm{i}}$, then it is clear that $\mathrm{Re}(a)$ and $\mathrm{Im}(a)$ are self-adjoint and that $a = \mathrm{Re}(a) + \mathrm{i}\cdot\mathrm{Im}(a)$. Therefore, $\varphi\big(\mathrm{Re}(a)\big), \varphi\big(\mathrm{Im}(a)\big) \in \mathbb{R}$, hence

$$\varphi(a^*) = \varphi\Big(\big(\mathrm{Re}(a) + \mathrm{i}\cdot\mathrm{Im}(a)\big)^*\Big) = \varphi\big(\mathrm{Re}(a) - \mathrm{i}\cdot\mathrm{Im}(a)\big) = \overline{\varphi(a)},$$

which concludes the proof. □

Remark 2.16. Let $(\mathcal{A}, *, \varphi)$ be a $*$-probability space. Viewing $\mathcal{A}$ as a vector space over $\mathbb{C}$, the function

$$\begin{array}{ccc} \mathcal{A} \times \mathcal{A} & \to & \mathbb{C} \\ (a, b) & \mapsto & \varphi(b^* a) \end{array} \tag{2.1}$$

is a *semi-inner product*, meaning that $\langle a, b\rangle = \varphi(b^* a)$ satisfies all the axioms of an inner product on $\mathcal{A}$ except that it is positive-semidefinite instead of positive-definite (if φ is faithful, however, then (2.1) is an inner product). Thus, many results concerning semi-inner products can be formulated in the language of $*$-probability spaces, such as the **Cauchy-Schwarz Inequality** for semi-inner products:

$$|\varphi(b^* a)|^2 = |\langle a, b\rangle|^2 \leqslant \langle a, a\rangle\langle b, b\rangle = \varphi(a^* a)\varphi(b^* b), \qquad a, b \in \mathcal{A}. \tag{2.2}$$

Remark 2.17. According to the previous remark, if the functional on the $*$-probability space $(\mathcal{A}, *, \varphi)$ is faithful, then there is a natural choice for a norm on $\mathcal{A}$, namely:

$$\|a\|_\varphi \coloneqq \sqrt{\varphi(a^* a)}, \qquad a \in \mathcal{A}.$$

Thus, in this case, the variance can be seen as a measure of how much random variables deviate from their expected value, as

$$\mathbf{Var}[a] = \|a - \varphi(a) \cdot 1_{\mathcal{A}}\|_\varphi^2.$$

One of the advantages of imposing a C^*-algebra structure on $*$-probability spaces is that the linear functional is guaranteed to be continuous:

Theorem 2.18. *Let $(\mathcal{A}, *, \|\cdot\|, \varphi)$ be a C^*-probability space. Then, for every $a \in \mathcal{A}$, one has $|\varphi(a)| \leqslant \|a\|$.*

Proof. Suppose first that a is in the positive cone of $\mathcal{A}$, that is, there exists $b \in \mathcal{A}$ such that $a = b^* b$. Since a is self-adjoint, it follows from Proposition 1.27 and Theorems 1.48 and 1.21 that $\mathrm{Sp}_{\mathcal{A}}(a) \subset \big[0, \|a\|\big]$. Define the map $f : \mathrm{Sp}_{\mathcal{A}}(a) \to \mathbb{R}$ as $f(x) = \sqrt{\|a\| - x}$. Since f is continuous, it follows from the Functional Calculus Theorem that $f(a)$ defines an element in $\mathcal{A}$, and since a is self-adjoint, so is $f(a)$. Furthermore, $f(a)^2 = \|a\| - a$, and thus

$$\|a\| - \varphi(a) = \varphi\big(f(a)^2\big) = \varphi\big(f(a) f(a)^*\big) \geqslant 0.$$

Therefore, the theorem holds for every variable in the positive cone $\mathcal{A}^+$.

Let $a \in \mathcal{A}$ be arbitrary. According to the Cauchy-Schwarz Inequality (see Remark 2.16), it follows that $|\varphi(a)| = |\varphi(1_{\mathcal{A}}^* a)| \leqslant \sqrt{\varphi(a^* a)\varphi(1_{\mathcal{A}})}$. Since $a^* a$ is positive, then $\varphi(a^* a) \leqslant \|a^* a\|$, and thus, $|\varphi(a)| \leqslant \sqrt{\|a^* a\|} = \|a\|$. □

Definition 2.19. Let $(\mathcal{A}_1, \varphi_1)$ and $(\mathcal{A}_2, \varphi_2)$ be noncommutative probability spaces, and let $\Phi : \mathcal{A}_1 \to \mathcal{A}_2$ be an algebra homomorphism.

(1) If $\varphi_1(a) = \varphi_2\big(\Phi(a)\big)$ for every $a \in \mathcal{A}_1$, then Φ is called a **noncommutative probability spaces homomorphism**.

(2) Suppose that $(\mathcal{A}_1, *_1, \varphi_1)$ and $(\mathcal{A}_2, *_2, \varphi_2)$ are $*$-probability spaces. If Φ is a noncommutative probability spaces homomorphism and a $*$-homomorphism, then Φ is called a **$*$-probability space homomorphism**.

(3) Suppose that $(\mathcal{A}_1, *_1, \|\cdot\|_{(1)}, \varphi_1)$ and $(\mathcal{A}_2, *_2, \|\cdot\|_{(2)}, \varphi_2)$ are C^*-probability spaces. If Φ is a $*$-probability spaces homomorphism and norm preserving (i.e., $\|a\|_{(1)} = \|\Phi(a)\|_{(2)}$ for every $a \in \mathcal{A}_1$), then Φ is called a ***C**-probability space homomorphism**.

(4) If Φ^{-1} exists and is a noncommutative probability space homomorphism (resp. a $*$-probability space, or C^*-probability space homomorphism), Φ is called a **noncommutative probability space isomorphism** (resp. $*$-probability space, or C^*-probability space isomorphism), and $\cong$ is used to denote the existence of an isomorphism between probability spaces.

As shown in the next theorem, the Gelfand-Naimark Theorem III can be extended to the context of noncommutative probability spaces, showing that every commutative C^*-probability space can be realized as a space of $\mathbb{C}$-valued random variables equipped with the usual expectation $\mathbf{E}$.

Theorem 2.20. *Let $(\mathcal{A}, *, \|\cdot\|, \varphi)$ be a commutative C^*-probability space. Then, there exists a compact topological space Ω and a unique regular probability measure P on Ω such that*

$$(\mathcal{A}, *, \|\cdot\|, \varphi) \cong \big(\mathrm{C}(\Omega, \mathbb{C}), \bar{\cdot}, \|\cdot\|_\infty, \mathbf{E}\big).$$

Proof. According to Theorem 1.35 and the *Gelfand-Naimark Theorem III*,[5] there exists a compact topological space Ω (which is isomorphic to $\Delta(\mathcal{A})$, see Section 1.3.) such that

$$(\mathcal{A}, *, \|\cdot\|) \cong \big(\mathrm{C}(\Omega, \mathbb{C}), \bar{\cdot}, \|\cdot\|_\infty\big).$$

Let $\mathcal{G}$ denote the isomorphism from $\mathcal{A}$ to $\mathrm{C}(\Omega, \mathbb{C})$. Since $\mathcal{G}^{-1}$ is a $*$-algebra homomorphism, the function

$$\begin{array}{ccc} \mathrm{C}(\Omega, \mathbb{C}) & \to & \mathbb{C} \\ X & \mapsto & \varphi\big(\mathcal{G}^{-1}(X)\big) \end{array}$$

is a positive linear functional. Therefore, it follows from the Riesz Representation Theorem[6] that there exists a unique regular measure P on Ω such that

$$\int_\Omega X \, \mathrm{d}P = \varphi\big(\mathcal{G}^{-1}(X)\big), \qquad\qquad X \in \mathrm{C}(\Omega, \mathbb{C}).$$

Since $\varphi\big(\mathcal{G}^{-1}(\mathbf{1}_\Omega)\big) = \varphi(1_\mathcal{A}) = 1$, P must be a probability measure, hence

$$\mathbf{E}[X] = \int_\Omega X \, \mathrm{d}P = \varphi\big(\mathcal{G}^{-1}(X)\big)$$

and

$$\mathbf{E}[\mathcal{G}(X)] = \varphi\big(\mathcal{G}^{-1}(\mathcal{G}(X))\big) = \varphi(X),$$

which concludes the proof. □

[5] See Section 1.3.

[6] See Section B.6.

2.2. Distributions

In the measure-theoretic formulation of probability theory, complex-valued random variables are typically distinguished by their distributions only, meaning that the specific behaviour of a random variable as a function is considered somewhat irrelevant[7]. In noncommutative probability, there exists similar notions that are used to distinguish noncommutative random variables independently of the specific algebra to which they belong.

To give an abstract definition of the distribution of an arbitrary noncommutative random variable, it is necessary to introduce some terminology.

Definition 2.21. Let $(x_i : i \in I)$ be an arbitrary collection of abstract indeterminates.

(1) Let $\mathbb{C}\big\langle (x_i, x_i^* : i \in I)\big\rangle$ denote the free unital $*$-algebra over $\mathbb{C}$ generated by $(x_i : i \in I)$. (We adhere to the convention that x_i^0 is equal to the unit element in $\mathbb{C}\big\langle (x_i, x_i^* : i \in I)\big\rangle$ for every $i \in I$.)

(2) An expression of the form

$$M(x_i : i \in I) = x_{i(1)}^{n(1)} \cdots x_{i(t)}^{n(t)} \in \mathbb{C}\big\langle (x_i, x_i^* : i \in I)\big\rangle$$

(where $t \geqslant 1$, $i(1), \ldots, i(t) \in I$ and $n(1), \ldots, n(t) \in \{0, 1, *\}$) is called a **noncommutative $*$-word** in the indeterminates x_i.

(3) An expression of the form

$$P(x_i : i \in I) = \sum_{k=1}^{m} \alpha_k M_k(x_i : i \in I) \in \mathbb{C}\big\langle (x_i, x_i^* : i \in I)\big\rangle$$

(where $m \geqslant 1$, $\alpha_1, \ldots, \alpha_m \in \mathbb{C}$, and $M_1, \ldots, M_k$ are noncommutative $*$-words) is called a **noncommutative $*$-polynomial** in the indeterminates x_i.

Definition 2.22. Let $(\mathcal{A}, *, \varphi)$ be a $*$-probability space. The **$*$-distribution** of a random variable $a \in A$, which is denoted by σ_a, is defined as the linear functional

$$\begin{array}{rccc} \sigma_a : & \mathbb{C}\langle x, x^*\rangle & \to & \mathbb{C} \\ & P(x) & \mapsto & \varphi\big(P(a)\big). \end{array}$$

Two random variables a and b are said to be **identically distributed** if $\sigma_a = \sigma_b$.

Remark 2.23. Given that the functional φ with which a $*$-probability space is equipped is linear, it is clear that the $*$-distribution σ_a of any random variable a is completely determined by the values that it takes on $*$-words. Furthermore, in certain situations, the distribution can be completely determined by a specific subset of $*$-words. For instance,

(1) if a is normal, then every $*$-word M evaluated at a can be rewritten as an expression of the form $a^k(a^*)^l$ for some nonnegative integers k and l, and thus σ_a is determined by the values it takes on $*$-words of the form $M(x) = x^k(x^*)^l$;

(2) if a is unitary, then every $*$-word M evaluated at a can be reduced to an expression of the form a^k or $(a^*)^k$ for some $k \in \mathbb{N} \cup \{0\}$, and thus, since φ is self-adjoint, σ_a is determined by the values it takes on $*$-words of the form $M(x) = x^k$; and

[7] In fact, one may argue that knowing the precise outcome $X(\omega)$ of a random variable X for every point ω in the sample space makes X not so random after all.

(3) if a is self-adjoint, then every $*$-word M evaluated at a can be reduced to an expression of the form a^k, and thus σ_a is also determined by $*$-words of the form $M(x) = x^k$.

Let $X \in L^{\infty-}$ be a commutative (and hence normal) $\mathbb{C}$-valued random variable with distribution μ_X in the measure-theoretic sense. Then, the $*$-distribution σ_X of X can be recovered from μ_X through the relation

$$\sigma_X\big(x^k(x^*)^l\big) = \mathbf{E}\big[X^k(X^*)^l\big] = \int_{\mathbb{C}} z^k\overline{z}^l \, \mathrm{d}\mu_X(z), \qquad\qquad k,l \in \mathbb{N}\cup\{0\}.$$

Therefore, the $*$-distribution of a commutative $\mathbb{C}$-valued random variable X actually consists of the moments of the distribution of X in the measure theoretic sense. If the functional φ on an arbitrary $*$-probability space is interpreted as an analog of the expected value, then the distribution σ_a of a random variable a can be seen as a map from which every linear combination of moments of a can be obtained. In the measure-theoretic formulation of probability theory, linear combinations of moments of a random variable X (such as the mean, variance, skewness, kurtosis, etc.) are usually thought of as quantities that encode information about the *shape* of the distribution of the image (i.e., the spectrum) of the random variable X. In noncommutative probability, these quantities are the distribution. As shown in the two following examples, there are examples other than commutative random variables for which the $*$-distribution can be seen as way to encode information about the distribution of the spectrum.

Example 2.24. Let $X \in \mathcal{M}_n(\mathbb{C})$ be a normal matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$. Then, according to the Spectral Theorem for Normal Matrices,[8] there exists a unitary matrix $U \in \mathcal{M}_n(\mathbb{C})$ such that $X = UDU^*$, where $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ is a diagonal matrix with the eigenvalues of X on the diagonal. Therefore, for every $k,l \in \mathbb{N}\cup\{0\}$, one has

$$\tfrac{1}{n}\operatorname{tr}\big(X^k(X^*)^l\big) = \tfrac{1}{n}\operatorname{tr}\big(D^k(D^*)^l\big) = \frac{1}{n}\sum_{i=1}^{n} \lambda_i^k\overline{\lambda_i}^l = \int_{\mathbb{C}} z^k\overline{z}^l \, \mathrm{d}\mathscr{L}_X(z),$$

where $\mathscr{L}_X = \frac{1}{n}\sum_{i=1}^n \delta_{\lambda_i}$ is the **empirical eigenvalue distribution** of the matrix X.

Example 2.25. Let $(\Omega, \mathscr{A}, P)$ be a probability space, let $X \in \mathcal{M}_n(L^{\infty-})$ be a normal random matrix (i.e., $X(\omega)$ is normal for every $\omega \in \Omega$), and let $\lambda_1, \ldots, \lambda_n : \Omega \to \mathbb{C}$ be the (random) eigenvalues of X. Then, for every $\omega \in \Omega$, there exists a unitary matrix $U(\omega)$ such that $X(\omega) = U(\omega)D(\omega)U^*(\omega)$, where $D(\omega) = \operatorname{diag}\big(\lambda_1(\omega), \ldots, \lambda_n(\omega)\big)$. Therefore, for every $k,l \in \mathbb{N}\cup\{0\}$, one has

$$\begin{aligned}\big(\tfrac{1}{n}\operatorname{tr}\otimes\mathbf{E}\big)\big(X^k(X^*)^l\big) &= \mathbf{E}\left[\frac{1}{n}\sum_{i=1}^{n}\lambda_i^k\overline{\lambda_i}^l\right] \\ &= \mathbf{E}\left[\int_{\mathbb{C}} z^k\overline{z}^l \, \mathrm{d}\mathscr{L}_X(z)\right] \\ &= \int_{\mathbb{C}} z^k\overline{z}^l \, \mathrm{d}\overline{\mathscr{L}_X}(z),\end{aligned}$$

where $\overline{\mathscr{L}_X}$ is the expected empirical eigenvalue distribution[9] of the random matrix X.

[8] See section A.3.

[9] Let X be a random matrix whose spectrum is almost surely contained in a compact set K, and let $\mathscr{L}_X(\omega)$ be be the empirical eigenvalue distribution of $X(\omega)$ for each $\omega \in \Omega$. Then, the **expected empirical eigenvalue distribution** is defined as the unique measure $\overline{\mathscr{L}_X}$ such that $\int f(z) \, \mathrm{d}\overline{\mathscr{L}_X}(z) = \mathbf{E}\left[\int f(z) \, \mathrm{d}\mathscr{L}_X(z)\right]$ for every continuous function

In general, however, especially if a random variable is not normal, the $*$-distribution can have other interesting interpretations. Consider for instance the following example:

Example 2.26. Let G be a group with identity e generated by a collection $\{g_1, \ldots, g_n\}$. Let $\{W_n : n \in \mathbb{N} \cup \{0\}\}$ be a random walk on G that starts at e, that is, $W_0 = e$ and for every $n \in \mathbb{N}$, there exists a random *step* $s_n \in G$ such that $W_n = W_{n-1}s_n$. Suppose that the steps s_n are i.i.d. copies of a random element s such that $s = g_i$ with probability α_i and $s = g_i^{-1}$ with probability β_i, where $\alpha_1 + \cdots + \alpha_n + \beta_1 + \cdots + \beta_n = 1$. Then, for every $k \in \mathbb{N}$, the probability that $W_k = g$ for some $g \in G$ is given by the coefficient that multiplies g in

$$(\alpha_1 g_1 + \cdots + \alpha_n g_n + \beta_1 g_1^{-1} + \cdots + \beta_n g_n^{-1})^k \in \mathbb{C}G.$$

Therefore, if one defines the linear functional $\tau_g : \mathbb{C}G \to \mathbb{C}$ as

$$\tau_g\left(\sum_{i=1}^{m} \lambda_{h_i} h_i\right) = \lambda_g, \qquad\qquad h_1, \ldots, h_m \in G$$

then the $*$-distribution σ_S of the element

$$S \coloneqq \alpha_1 g_1 + \cdots + \alpha_n g_n + \beta_1 g_1^{-1} + \cdots + \beta_n g_n^{-1}$$

allows one to compute the probability $\mathbf{P}[W_k = g]$ for any $k \in \mathbb{N}$ through the relation $\mathbf{P}[W_k = g] = \tau_g(S^k)$.

Definition 2.27. Let a be a normal random variable in a $*$-probability space $(\mathcal{A}, *, \varphi)$. If there exists a measure μ_a on $\mathbb{C}$ from which the $*$-distribution σ_a of a can be determined by the relation

$$\sigma_a\big(x^k(x^*)^l\big) = \varphi\big(a^k(a^*)^l\big) = \int_{\mathbb{C}} z^k \overline{z}^l \, \mathrm{d}\mu_a(z), \qquad\qquad k, l \in \mathbb{N} \cup \{0\}$$

then μ_a is called an **analytic distribution** of a.

Note that, in the above definition, the measure μ_a is called *an* analytic distribution of a rather than *the* analytic distribution of a. The reason for this is that, in general, analytic distributions need not be unique. Indeed, for an analytic distribution μ_a to be unique, it must be *determined by its moments*, which means that no other measure on $\mathbb{C}$ has the same moments as μ_a. To ensure the uniqueness of analytic distributions, some authors require that every analytic distribution has a compact support.[10] If one does not wish to restrict analytic distributions to compactly supported measures, it may still be possible to prove the uniqueness of the analytic distributions in some cases using criterions such as Theorem B.6.

Definition 2.28. Let $(\mathcal{A}, *, \varphi)$ be a $*$-probability space.

(1) A self-adjoint variable $a \in \mathcal{A}$ is said to be **semicircular** if for every $n \in \mathbb{N} \cup \{0\}$, one has

$$\varphi(a^n) = \begin{cases} \frac{1}{n/2+1}\binom{n}{n/2} & \text{if } n \text{ is even; and} \\ 0 & \text{if } n \text{ is odd,} \end{cases}$$

---

f supported in K, whose existence and uniqueness is guaranteed by the Riesz Representation Theorem, since $f \mapsto \mathbf{E}\left[\int f(z) \, \mathrm{d}\mathscr{L}_X(z)\right]$ is a positive linear functional.

[10] One can easily prove, using the Stone-Weierstrass Approximation Theorem and Lebesque's Dominated Convergence Theorem, that two distinct compactly supported measures on $\mathbb{C}$ cannot have the same moments.

or, in other words, the $*$-distribution of a is given by the semicircular distribution $\frac{1}{2\pi}\sqrt{4-x^2}\,\mathrm{d}x$ on $[-2,2]$ through the relation

$$\sigma_a(x^k) = \int_{[-2,2]} x^k \frac{1}{2\pi}\sqrt{4-x^2}\,\mathrm{d}x, \qquad k \in \mathbb{N}\cup\{0\}.$$

(2) A self-adjoint variable $a \in \mathcal{A}$ is said to be **Gaussian** if for every $n \in \mathbb{N}\cup\{0\}$, one has

$$\varphi(a^n) = \begin{cases} (n-1)!! & \text{if } n \text{ is even; and} \\ 0 & \text{if } n \text{ is odd} \end{cases}$$

(where !! denotes the double factorial[11]), or, in other words, the $*$-distribution of a is given by the standard Gaussian distribution[12] $\frac{1}{\sqrt{2\pi}}\exp\left(-x^2/2\right)\,\mathrm{d}x$ on $\mathbb{R}$ through the relation

$$\sigma_a(x^k) = \int_{\mathbb{R}} x^k \frac{1}{\sqrt{2\pi}}\exp\left(-x^2/2\right)\,\mathrm{d}x, \qquad k \in \mathbb{N}\cup\{0\}.$$

(3) A unitary random variable $u \in \mathcal{A}$ is said to be **Haar unitary** if $\varphi(u^n) = 0$ whenever $n \neq 0$, or, in other words, the $*$-distribution of u is given by the normalized Haar measure μ on the unit circle $\mathbb{U}_1 = \{z \in \mathbb{C} : |z| = 1\}$ through the relation

$$\sigma_u\left(x^k(x^*)^l\right) = \int_{\mathbb{U}_1} z^k\overline{z}^l\,\mathrm{d}\mu(z) = \int_0^{2\pi} \frac{\mathrm{e}^{\mathrm{i}(k-l)x}}{2\pi}\,\mathrm{d}x, \qquad k,l \in \mathbb{N}\cup\{0\}.$$

(4) A unitary random variable $u \in \mathcal{A}$ is said to be ***p*-Haar unitary** (where $p \in \mathbb{N}$) if $u^n = 1_{\mathcal{A}}$ if p divides n and $\varphi(u^n) = 0$ otherwise, or, in other words, the $*$-distribution of u is given by the uniform measure on the roots of unity of order p

$$\mu_p = \frac{1}{p}\sum_{n=1}^{p} \delta_{\exp(2\cdot\mathrm{i}\cdot n\pi/p)}$$

through the relation

$$\sigma_u\left(x^k(x^*)^l\right) = \int_{\mathbb{C}} z^k\overline{z^l}\,\mathrm{d}\mu_p(z) = \frac{1}{p}\sum_{n=1}^{p} \mathrm{e}^{\frac{2\cdot\mathrm{i}\cdot n(k-l)\pi}{p}}, \qquad k,l \in \mathbb{N}\cup\{0\}.$$

Example 2.29. Let $(\mathbb{C}G, *, \tau_e)$ be defined as in Example 2.4. Then, every element of the group $g \in G$ is either Haar unitary or p-Haar unitary. More precisely, if g is of infinite order, then it is Haar unitary, and if it is of finite order p, then it is p-Haar unitary.

Example 2.30 (Haar Unitary Random Matrices and Their Analytic Distributions)**.** Let $(\Omega, \mathscr{A}, P)$ be a probability space, and for a fixed integer $n \in \mathbb{N}$, let $\mathbb{U}_n \subset \mathcal{M}_n(\mathbb{C})$ be the set of $n \times n$ unitary matrices. Since $\mathbb{U}_n$ is a compact topological group, it can be equipped with a normalized Haar measure μ. Let $(U_{ij} : 1 \leqslant i,j \leqslant n) = U : \Omega \to \mathbb{U}_n$ be a $n \times n$ random matrix whose distribution is given by the Haar measure μ, that is, for every Borel set $A \subset \mathbb{U}_n$, one has $\mathbf{P}[U \in A] = \mu(A)$ (such a matrix is usually called a **Haar unitary random matrix**). Then, for every $\omega \in \Omega$, the columns of $U(\omega)$ form an orthonormal basis of $\mathbb{C}^n$, which implies that every entry $U_{ij} : \Omega \to \mathbb{C}$ is a bounded random variable. Therefore, U is contained in the $*$-probability

[11] Recall that $-1!! = 0!! = 1$, $n!! = n(n-2)(n-4)\cdots 4\cdot 2$ if $n > 0$ is even, and $n!! = n(n-2)(n-4)\cdots 3\cdot 1$ if $n > 0$ is odd.

[12] Notice that the Gaussian distribution is determined by its moments, hence Gaussian variables have a unique analytic distribution.

space $\big(\mathcal{M}_n(L^{\infty-}), *, \frac{1}{n}\mathrm{tr}\otimes\mathbf{E}\big)$. Moreover, it can be shown that U is a Haar unitary variable: for any $k \in \mathbb{N}\setminus\{0\}$, every entry of the matrix U^k can be written as a linear combination of terms of the form

$$U_{i(1)j(1)}\cdots U_{i(k)j(k)}, \qquad 1 \leqslant i(1), j(1), \ldots, i(k), j(k) \leqslant n$$

and thus $(\frac{1}{n}\mathrm{tr}\otimes\mathbf{E})(U^k)$ is a linear combination of terms of the form

$$\mathbf{E}\left[U_{i(1)j(1)}\cdots U_{i(k)j(k)}\right] = \int_{\mathbb{U}_n} u_{i(1)j(1)}\cdots u_{i(k)j(k)}\ \mathrm{d}\mu(u_{ij}).$$

Let $1 \leqslant i(1), j(1), \ldots, i(k), j(k) \leqslant n$ be fixed, and define the map

$$\begin{array}{rccc} f: & \mathbb{U}_n & \mapsto & \mathbb{C} \\ & (u_{ij}) & \mapsto & u_{i(1)j(1)}\cdots u_{i(k)j(k)}. \end{array}$$

Let $d = \mathrm{e}^{\mathrm{i}\pi/k}\mathrm{Id}_n$, where Id_n is the $n\times n$ identity matrix. Then $d \in \mathbb{U}_n$, and since the Haar measure is translation invariant, it follows that

$$\begin{aligned} \int_{\mathbb{U}_n} u_{i(1)j(1)}\cdots u_{i(k)j(k)}\ \mathrm{d}\mu(u_{ij}) &= \int_{\mathbb{U}_n} f(u_{ij})\ \mathrm{d}\mu(u_{ij}) \\ &= \int_{\mathbb{U}_n} f(d\cdot u_{ij})\ \mathrm{d}\mu(u_{ij}) \\ &= \int_{\mathbb{U}_n} \mathrm{e}^{\mathrm{i}\pi/k}u_{i(1)j(1)}\cdots \mathrm{e}^{\mathrm{i}\pi/k}u_{i(k)j(k)}\ \mathrm{d}\mu(u_{ij}) \\ &= \mathrm{e}^{\mathrm{i}\pi}\int_{\mathbb{U}_n} u_{i(1)j(1)}\cdots u_{i(k)j(k)}\ \mathrm{d}\mu(u_{ij}). \end{aligned}$$

Given that $z = \mathrm{e}^{\mathrm{i}\pi}z$ if and only if $z = 0$, one has

$$\int_{\mathbb{U}_n} u_{i(1)j(1)}\cdots u_{i(k)j(k)}\ \mathrm{d}\mu(u_{ij}) = 0,$$

hence $\frac{1}{n}(\mathrm{tr}\otimes\mathbf{E})(U^k) = 0$.

Up to this point, we have seen simple examples of random variables with analytic distributions as well as examples of variables whose $*$-distributions are directly given by measures. When it comes to analytic distributions for arbitrary normal variables, there are two questions of special interest:

Question 2.31. Let $(\mathcal{A}, *, \varphi)$ be a $*$-probability space and $a \in \mathcal{A}$ be a normal variable.

(1) Under what conditions is a guaranteed to have an analytic distribution?

(2) Suppose that a has an analytic distribution μ_a. To what extent can μ_a be *described explicitly*?[13]

[13]For example, if μ_a is absolutely continuous with respect to the Lebesgue measure $\mathrm{d}z$ on $\mathbb{C}$, then it can be described explicitly using its density with respect to the Lebesgue measure, that is, a measurable function f such that $\mu_a(A) = \int_A f(z)\ \mathrm{d}z$ for every measurable set A.

One way to approach part (1) of Question 2.31 is to exploit the fact that the $*$-algebra generated by a normal element $a \in \mathcal{A}$ is commutative, and thus if a can be embedded in a C^*-probability space, then it can be associated to a commutative random variable (which obviously has an analytic distribution) by the Gelfand-Naimark Theorem. This method gives the following theorem.

Theorem 2.32. *Let a be a normal random variable in a C^*-probability space $(\mathcal{A}, *, \|\cdot\|, \varphi)$. Then, a has an analytic distribution μ_a such that*

(1) $\mathrm{supp}(\mu_a) \subset \mathrm{Sp}_{\mathcal{A}}(a)$ (which implies in particular that μ_a has a compact support); and

(2) for every continuous function $f : \mathrm{Sp}_{\mathcal{A}} \to \mathbb{C}$, one has

$$\varphi\big(f(a)\big) = \int_{\mathbb{C}} f(z)\ \mathrm{d}\mu_a(z),$$

where $f(a)$ is given by the functional calculus on $\mathcal{A}$.

Proof. The existence of μ_a is a direct consequence of Theorem 2.20: Since a is normal, it follows from Proposition 1.17 that $\mathcal{B} \coloneqq C^*(1_{\mathcal{A}}, a)$ (the unital C^*-algebra generated by a) is commutative, and thus $(\mathcal{B}, *, \|\cdot\|, \varphi)$ is isomorphic to $\big(\mathrm{C}(\Omega, \mathbb{C}), \bar{\cdot}, \|\cdot\|_\infty, \mathbf{E}\big)$ for some compact probability space $\big(\Omega, \mathscr{B}(\Omega), P\big)$. Let X be the random variable on Ω to which a is isomorphic, and let μ_X be the distribution of X in the measure-theoretic sense. Then, it follows that

$$\varphi\left(a^k (a^*)^l\right) = \mathbf{E}\left[X^k (X^*)^l\right] = \int_{\mathbb{C}} z^k \overline{z}^l\ \mathrm{d}\mu_X(z)$$

for every $k, l \in \mathbb{N} \cup \{0\}$, hence μ_a exists and is equal to μ_X.

(1). According to the proof of the Gelfand-Naimark Theorem III (see the end of section 1.3) the random variable X to which a is isomorphic in $\mathrm{C}(\Omega, \mathbb{C})$ is given by the Gelfand transform $\mathcal{G}(a)$ at a. According to Theorem 1.34, the image of $\mathcal{G}(a)$ is contained in $\mathrm{Sp}_{\mathcal{A}}(a)$. Therefore, the set of values that the random variable X can take is included in $\mathrm{Sp}_{\mathcal{A}}(a)$, which implies that the distribution of X must be supported in $\mathrm{Sp}_{\mathcal{A}}(a)$.

(2). Let $f : \mathrm{Sp}_{\mathcal{A}}(a) \to \mathbb{C}$ be a continuous function. Since $\mathrm{Sp}_{\mathcal{A}}(a)$ is compact, it follows from the Stone-Weierstrass Approximation Theorem that there exists a sequence of polynomial functions $\{p_n : n \in \mathbb{N}\}$ in $\mathbb{C}[z, \overline{z}]$ such that

$$\lim_{n\to\infty} \|p_n - f\|_\infty = \lim_{n\to\infty} \sup\big\{|p_n(\lambda) - f(\lambda)| : \lambda \in \mathrm{Sp}_{\mathcal{A}}(a)\big\} = 0.$$

It then follows from Lebesgue's Dominated Convergence Theorem that

$$\lim_{n\to\infty} \varphi\big(p_n(a)\big) = \lim_{n\to\infty} \int_{\mathbb{C}} p_n(z)\ \mathrm{d}\mu_a(z) = \int_{\mathbb{C}} f(z)\ \mathrm{d}\mu_a(z).$$

Furthermore, according to the Functional Calculus Theorem (see Section 1.4), one has $\|p_n(a) - f(a)\| = \|p_n - f\|_\infty$ for every n. Therefore, given that φ is continuous (see Theorem 2.18), it follows that $\varphi\big(p_n(a)\big) \to \varphi\big(f(a)\big)$. □

Given the strong link between spaces of bounded linear operators and C^*-algebras (see for instance the Gelfand-Naimark Theorem I), one easily obtains the following corollary, which slightly relaxes the hypotheses of Theorem 2.32.

Corollary 2.33. *Let $(\mathcal{A}, *, \varphi)$ be a $*$-algebra, let $a \in \mathcal{A}$ be a normal element, and let $\mathcal{B} := *\text{-alg}(1_{\mathcal{A}}, a)$. Suppose that there exists a $*$-probability space homomorphism Φ from $(\mathcal{B}, *, \varphi)$ to $\big(B(H), *, \varphi_{a_0}\big)$ (the latter being defined as in Example 2.8 (2)). Then, a has an analytic distribution μ_a that satisfies conditions (1) and (2) of Theorem 2.32.*

Proof. Let T_a be the image of a through Φ. Since Φ is a $*$-homomorphism, T_a is normal. Therefore, given that $\big(B(H), *, \|\cdot\|_{\text{op}}, \varphi_{a_0}\big)$ is a C^*-probability space, it follows from Theorem 2.32 that T_a has an analytic distribution μ_a supported on $\text{Sp}_{B(H)}(T_a)$. Let $k, l \in \mathbb{N} \cup \{0\}$ be arbitrary. Since Φ is a $*$-probability space homomorphism, it follows that

$$\varphi\left(a^k (a^*)^l\right) = \varphi_{a_0}\left(\Phi\left(a^k (a^*)^l\right)\right) = \varphi_{a_0}\left(T_a^k (T_a^*)^l\right) = \int_{\mathbb{C}} z^k \overline{z}^l \, \mathrm{d}\mu_a(z),$$

hence μ_a is an analytic distribution for a. Moreover, given that

$$\text{Sp}_{B(H)}(T_a) \subset \text{Sp}_{\mathcal{B}}(a) \subset \text{Sp}_{\mathcal{A}}(a)$$

(see Remark 1.20), it follows that the support of μ_a is contained in $\text{Sp}_{\mathcal{A}}(a)$. The proof that $\varphi\big(f(a)\big) = \int_{\mathbb{C}} f(z) \, \mathrm{d}\mu_a(z)$ for every continuous map f is identical to the proof of part (2) in Theorem 2.32. □

Another approach to part (1) Question 2.31, which requires a substantial knowledge of the moments of the normal random variable a in question, consists in working directly from the *array* or *sequence* of moments. (See the introduction of **[Bi89]**, **[At75]**, and chapter 6 of **[BCR89]** for a discussion/proof of the two following results.)

Hamburger's Moment Theorem. *Let $\big(\alpha_n : n \in \mathbb{N} \cup \{0\}\big)$ be a sequence of real numbers. There exists a measure μ on $\mathbb{R}$ such that*

$$\alpha_n = \int_{\mathbb{R}} x^n \, \mathrm{d}\mu(x), \qquad\qquad n \in \mathbb{N} \cup \{0\}$$

if and only if for every $k \in \mathbb{N} \cup \{0\}$, the matrix $A_k = (\alpha_{i+j} : 0 \leqslant i, j \leqslant k)$ is positive definite.

Complex Moment Theorem. *Let $\big(\alpha_{m,n} : m, n \in \mathbb{N} \cup \{0\}\big)$ be an array of complex numbers such that $\alpha_{m,n} = \overline{\alpha_{n,m}}$ for every m, n. There exists a measure μ on $\mathbb{C}$ supported on $\{z \in \mathbb{C} : |z| \leqslant 1\}$ such that*

$$\alpha_{m,n} = \int_{\mathbb{C}} z^m \overline{z}^n \, \mathrm{d}\mu(z), \qquad\qquad m, n \in \mathbb{N} \cup \{0\}$$

if and only if

(1) for every array $\big(c_{m,n} : m, n \in \mathbb{N} \cup \{0\}\big)$ of complex numbers with finitely many nonzero entries, one has

$$\sum_{m,n,j,k=0}^{\infty} \alpha_{(m+j),(n+k)} c_{n,j} \overline{c_{m,k}} \geqslant 0;$$

(2) for every sequence $\big(w_n : n \in \mathbb{N} \cup \{0\}\big)$ of complex numbers with finitely many nonzero terms, one has

$$\sum_{m,n=0}^{\infty} (\alpha_{m,n} - \alpha_{(m+1),(n+1)}) w_m \overline{w_n} \geqslant 0.$$

The solution to part (2) of Question 2.31 is in general much more difficult than part (1). If a random variable $a \in \mathcal{A}$ has a compactly supported analytic distribution μ_a and is self-adjoint (which implies that μ_a is supported on $\mathbb{R}$), then μ_a can be constructed through its moments by means of the *Stieltjes Inversion Formula*.[14]

Example 2.34 (Stieltjes Inversion Formula)**.** Let u be a Haar unitary element in a $*$-probability space $(\mathcal{A}, \varphi)$. Suppose we are interested in describing the analytic distribution of the self-adjoint element $u + u^*$. Given that u is unitary, that is, u commutes with u^* and $u^{-1} = u^*$, for every integer $n \geqslant 1$, one has

$$\varphi\big((u+u^*)^n\big) = \sum_{k=0}^{n} \binom{n}{k} \varphi\big(u^k (u^*)^{n-k}\big) = \sum_{k=0}^{n} \binom{n}{k} \varphi\big(u^{2k-n}\big).$$

Since u is Haar unitary, every term of the form $\varphi\big(u^{2k-n}\big)$ where $2k \neq n$ will vanish. Thus, we can conclude that

$$\varphi\big((u+u^*)^n\big) = \begin{cases} \binom{n}{n/2} & \text{if } n \text{ is even; and} \\ 0 & \text{if } n \text{ is odd.} \end{cases}$$

Suppose that there exists a compactly supported probability measure μ on $\mathbb{R}$ that is an analytic distribution for $u+u^*$, and let s_μ denote its Stieltjes transform. If $|z| > \sup\big\{|x| : x \in \operatorname{supp}(\mu)\big\}$, then it follows from Proposition B.9 that

$$s_\mu(z) = -\sum_{n=0}^{\infty} \frac{\varphi\big((u+u^*)^n\big)}{z^{n+1}} = -\sum_{n=0}^{\infty} \frac{\varphi\big((u+u^*)^{2n}\big)}{z^{2n+1}} = -\sum_{n=0}^{\infty} \frac{\binom{2n}{n}}{z^{2n+1}}.$$

Given that the product of two infinite series is given by the *Cauchy product*[15], one has

$$s_\mu(z)^2 = \sum_{n=0}^{\infty} \sum_{m=0}^{n} \frac{\binom{2m}{m}}{z^{2m+1}} \frac{\binom{2(n-m)}{n-m}}{z^{2(n-m)+1}} = \sum_{n=0}^{\infty} \frac{1}{(z^2)^{n+1}} \sum_{m=0}^{n} \binom{2m}{m} \binom{2(n-m)}{n-m}.$$

Using the well-known combinatorial identity (see for instance [**CG11**])

$$\sum_{m=0}^{n} \binom{2m}{m} \binom{2(n-m)}{n-m} = 4^n,$$

we finally obtain the following geometric series

$$s_\mu(z)^2 = \frac{1}{z^2} \sum_{n=0}^{\infty} \left(\frac{4}{z^2}\right)^n = \frac{1}{z^2} \frac{1}{1 - \frac{4}{z^2}} = \frac{1}{z^2 - 4}.$$

Therefore, one has

$$s_\mu(z) = \pm \frac{1}{\sqrt{z^2 - 4}} \tag{2.3}$$

for $|z| > \sup\big\{|x| : x \in \operatorname{supp}(\mu)\big\}$, and by the unicity of analytic continuations, it follows that (2.3) holds on all of $\mathbb{C} \setminus \mathbb{R}$. An application of the Stieltjes Inversion Formula then yields that μ

[14] See Section B.4.1 and Theorem B.10.

[15] The Cauchy Product for two infinite series $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ is given by $\sum_{n=0}^{\infty} c_n$, where, for every $n \geqslant 0$, $c_n = \sum_{m=0}^{n} a_m b_{n-m}$.

has a density given by

$$f(x) = \begin{cases} \frac{1}{\pi} \lim_{\varepsilon \to 0^+} \operatorname{Im}\big(s_\mu(x + \mathrm{i}\varepsilon)\big) & \text{whenever the limit exists; and} \\ 0 & \text{otherwise.} \end{cases}$$

which confirms that

$$s_\mu(z) = -\frac{1}{\sqrt{z^2 - 4}} = \frac{1}{\sqrt{4 - z^2}}$$

is the right choice, yielding

$$\begin{aligned} f(x) &= \begin{cases} \frac{1}{\pi} \lim_{\varepsilon \to 0^+} \operatorname{Im}\left(\frac{1}{\sqrt{4 - (x + \mathrm{i}\varepsilon)^2}}\right) & \text{whenever the limit exists; and} \\ 0 & \text{otherwise,} \end{cases} \\ &= \begin{cases} \frac{1}{\pi\sqrt{4 - x^2}} & \text{if } |x| < 2; \text{ and} \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

Definition 2.35. Let $(a_n : n \in \mathbb{N})$ be a sequence of noncommutative random variables such that for each n, a_n is an element of a $*$-probability space $(\mathcal{A}_n, *_n, \varphi_n)$, and let a be an element in a $*$-probability space $(\mathcal{A}, *, \varphi)$. The a_n are said to **converge in $*$-distribution** to a if the $*$-distribution of a_n converges to that of a as $n \to \infty$ pointwise, that is, for every $*$-polynomial P, one has

$$\lim_{n \to \infty} \varphi_n\big(P(a_n)\big) = \varphi\big(P(a)\big).$$

Remark 2.36. Let $(a_n : n \in \mathbb{N})$ be a sequence of normal random variables each having an analytic distribution μ_{a_n}. Suppose that a_n converges in $*$-distribution to a limit variable a having an analytic distribution μ_a. In this case, the convergence in $*$-distribution of the a_n is equivalent to the convergence of the moments of the distributions μ_{a_n}. Furthermore, if μ_a is determined by its moments, then it follows from the Method of Moments (see Section B.6) that μ_{a_n} converges weakly to μ_a.

Let $(X_i : i \in I)$ be a collection of $\mathbb{C}$-valued commutative random variables in $L^{\infty-}$ for some probability space $(\Omega, \mathscr{A}, P)$. Then, the concept of joint probability distribution gives a way of computing mixed $*$-moments

$$\mathbf{E}\Big[X_{i(1)}^{n(1)} \cdots X_{i(t)}^{n(t)}\Big], \qquad\qquad n(l) \in \mathbb{N},\ i(l) \in I \text{ distinct},$$

More precisely, recall that, given a collection $\mathbf{X} = (X_1, \ldots, X_t)$ of $\mathbb{C}$-valued random variables, the **joint probability distribution** of the collection $\mathbf{X}$, which is denoted $\mu_{\mathbf{X}}$, is defined as the unique probability measure on $\big(\mathbb{C}^t, \mathscr{B}(\mathbb{C}^t)\big)$ such that for every measurable rectangle $A = A_1 \times \cdots \times A_t \in \mathscr{B}(\mathbb{C}^t)$ (that is, A_i is Borel measurable for each $i \leqslant t$), one has

$$\mu_{\mathbf{X}}(A) = \mathbf{P}\left[\bigcap_{i=1}^{t} X_i \in A_i\right],$$

and for every $n(1), \ldots, n(t) \in \mathbb{N}$, one has

$$\mathbf{E}\Big[X_1^{n(1)} \cdots X_t^{n(t)}\Big] = \int_{\mathbb{C}^t} z_1^{n(1)} \cdots z_t^{n(t)} \, \mathrm{d}\mu_{\mathbf{X}}(z_1, \ldots, z_t).$$

Thus, in the moment-centric perspective of free probability theory, it seems appropriate to define the concept of joint $*$-distribution as follows:

Definition 2.37. Let $(\mathcal{A}, *, \varphi)$ be a $*$-probability space. The **joint $*$-distribution** of a collection $\mathbf{a} = (a_i : i \in I)$ of random variables in A, which is denoted by $\sigma_{\mathbf{a}}$, is defined as the linear functional

$$\begin{array}{rccc} \sigma_{\mathbf{a}} : & \mathbb{C}\big\langle (x_i, x_i^* :\in I)\big\rangle & \to & \mathbb{C} \\ & P(x_i : i \in I) & \mapsto & \varphi\big(P(a_i : i \in I)\big). \end{array}$$

Furthermore, quantities of the form

$$\varphi\Big(a_{i(1)}^{n(1)} \cdots a_{i(t)}^{n(t)}\Big), \qquad\qquad , n(1), \ldots, n(t) \in \{0, 1, *\}$$

are called **mixed $*$-moments** in the a_i.

Definition 2.38. For every $n \in \mathbb{N}$, let $\mathbf{a}^{(n)} = \Big(a_i^{(n)} : i \in I\Big)$ be a collection of noncommutative random variables in a $*$-probability space $(\mathcal{A}^{(n)}, *^{(n)}, \varphi^{(n)})$, and let $\mathbf{a} = (a_i : i \in I)$ be a collection in $(\mathcal{A}, *, \varphi)$. $\mathbf{a}^{(n)}$ is said to **converge in $*$-joint distribution** to $\mathbf{a}$ if $\sigma_{\mathbf{a}^{(n)}}$ converges to $\sigma_{\mathbf{a}}$ pointwise, that is, for every $*$-polynomial P, one has

$$\lim_{n\to\infty} \varphi^{(n)}\left(P\Big(a_i^{(n)} : i \in I\Big)\right) = \varphi\big(P(a_i : i \in I)\big).$$

Example 2.39. Let $(\Omega, \mathscr{A}, P)$ be a probability space, and let $L^{\infty-}$ be the algebra of $\mathbb{C}$-valued random variables on Ω with finite moments. Define the collection of random variables $(D_i : i \in \mathbb{N}), (Y_{ij} : 1 \leqslant i < j,\ i, j \in \mathbb{N}) \subset L^{\infty-}$ to be such that

(1) the D_i are i.i.d., $\mathbb{R}$-valued, and $\mathbf{E}[D_i] = 0$; and
(2) the Y_{ij} are i.i.d., $\mathbf{E}[Y_{ij}] = 0$, and $\mathbf{Var}[Y_{ij}] = \mathbf{E}[|Y_{ij}|^2] = 1$.

Given $n \in \mathbb{N}$, define the matrix $X^{(n)} = (X_{ij}^{(n)} : 1 \leqslant 1, j \leqslant n) \in \mathcal{M}_n(L^{\infty-})$ as

$$X_{ij}^{(n)} = \begin{cases} \frac{1}{\sqrt{n}} D_i & \text{if } i = j; \\ \frac{1}{\sqrt{n}} Y_{ij} & \text{if } i < j; \text{ and} \\ \frac{1}{\sqrt{n}} \overline{Y_{ji}} & \text{if } i > j. \end{cases}$$

($X^{(n)}$ is a special case of what is called a **Wigner Matrix**, see **[AGZ09]** equation (2.1.2).) Then, as $n \to \infty$, one has

(1) $X^{(n)}$ converges P-almost surely to a semicircular noncommutative random variable with respect to the normalized trace $\frac{1}{n}\mathrm{tr}$; and
(2) $X^{(n)}$ converges to a semicircular noncommutative random variable with respect to the expected normalized trace $\frac{1}{n}\mathrm{tr} \otimes \mathbf{E}$.

This result is known as **Wigner's Semicircle Law** (see the second chapter of **[AGZ09]**).

Part II. Free Probability

Chapter 3

Free Independence

The main sources for the concepts covered in this section are [**NS06**] Lectures 6 to 7, [**VDN92**] Chapter 2, [**Ta12**] Section 2.5.3, [**AGZ09**] Section 5.3, and [**Sp97**].

3.1. Independence of Commutative $\mathbb{C}$-valued Random Variables

The concept of independence for commutative $\mathbb{C}$-valued random variables is typically defined as follows:[1]

Definition 3.1. Let $\mathbf{X} = (X_i : i \in I)$ be a collection of $\mathbb{C}$-valued random variables on a probability space $(\Omega, \mathscr{A}, P)$ with respective distributions $(\mu_{X_i} : i \in I)$. The variables in $\mathbf{X}$ are said to be independent of each other if and only if the joint distribution $\mu_{\mathbf{X}}$ of the X_i is equal to the product $\times_{i\in I}\mu_{X_i}$ of the individual distributions of the X_i.

As evidenced by the above, the *classical* definition of independence is intimately linked with the measure-theoretic formulation of probability. In the present context of noncommutative probability (more specifically, the space $(L^{\infty-}, \bar{\tau}, \mathbf{E})$), where the only information about random variables one has access to are moments and mixed moments, it would be desirable to find a definition of independence that relies on moments alone. A good place to start is the following:[2]

Proposition 3.2. *Let $(\Omega, \mathscr{A}, P)$ be a standard*[3] *probability space and let $\mathbf{X} = (X_i : i \in I)$ be a collection of random variables on Ω. The X_i are independent of each other if and only if for every finite collection $i(1), \ldots, i(t) \in I$ of distinct indices and continuous functions $f_1, \ldots, f_t : \mathbb{C} \to \mathbb{C}$ such that $f_1(X_{i(1)}), \ldots, f_t(X_{i(t)})$ and $f_1(X_{i(1)}) \cdots f_t(X_{i(t)})$ are P-integrable, one has*

$$\mathbf{E}\big[f_1(X_{i(1)}) \cdots f_t(X_{i(t)})\big] = \mathbf{E}\big[f_1(X_{i(1)})\big] \cdots \mathbf{E}\big[f_t(X_{i(t)})\big]. \tag{3.1}$$

In the context of general noncommutative probability spaces, equation (3.1) is still potentially problematic. Indeed, given a noncommutative random variable a and a continuous function $f : \mathbb{C} \to \mathbb{C}$, unless a is normal and can be embedded in a C^*-probability space, there

[1] See Section B.1.2 for more details.

[2] See Proposition B.3 for the proof of Proposition 3.2.

[3] See [**It84**] Section 2.4.

is no guarantee that one can even make sense of the expression $f(a)$. Let $(\mathcal{A}, *, \|\cdot\|, \varphi)$ be a C^*-probability space of $\mathbb{C}$-valued random variables (for instance, $(\mathcal{A}, *, \|\cdot\|, \varphi) = \big(\mathrm{C}(\Omega, \mathbb{C}), \bar{\cdot}, \|\cdot\|_\infty, \mathbf{E}\big)$ for some compact probability space $(\Omega, \mathscr{A}, P)$). Let $\mathbf{a} = (a_i : i \in I)$ be a collection of variables in $\mathcal{A}$. In this case, it easily follows from the Stone-Weierstrass Approximation Theorem that equation (3.1) applied to $\mathbf{a}$ is equivalent to the following: For every finite collection $i(1), \ldots, i(t) \in I$ of distinct indices and $*$-polynomials $P_1, \ldots, P_t \in \mathbb{C}\langle x, x^*\rangle$, one has

$$\varphi\big(P_1(a_{i(1)}) \cdots P_t(a_{i(t)})\big) = \varphi\big(P_1(a_{i(1)})\big) \cdots \varphi\big(P_t(a_{i(t)})\big). \tag{3.2}$$

Therefore, whenever expressions of the form $f(a)$ (where a is a noncommutative random variable and f is a continuous function) can be defined consistently, the independence for $\mathbb{C}$-valued random variables is in fact equivalent to (3.2). Given that, for every $a \in \mathcal{A}$, one has

$$\begin{aligned} *\text{-alg}(1_\mathcal{A}, a) =& \mathrm{Span}_\mathbb{C}\Big(\big\{M(a) : M \in \mathbb{C}\langle x, x^*\rangle \text{ is a } *\text{-word}\big\}\Big) \\ =& \big\{P(a) : P \in \mathbb{C}\langle x, x^*\rangle \text{ is a } *\text{-polynomial}\big\} \end{aligned}$$

this leads to the following definition of independence for $\mathbb{C}$-valued random variables, in the context of noncommutative probability:

Definition 3.3. Let $(\mathcal{A}, \bar{\cdot}, \varphi)$ be a $*$-probability space of commutative $\mathbb{C}$-valued random variables, let $\mathbf{a} = (a_i : i \in I)$ be an arbitrary collection of random variables, and for every $i \in I$, let $\mathcal{A}_i \coloneqq *\text{-alg}(1_\mathcal{A}, a_i)$ be the $*$-algebra generated by a_i. The variables in $\mathbf{a}$ are independent of each other if and only if for every finite collection $i(1), \ldots, i(t) \in I$ of distinct indices and elements $b_1 \in \mathcal{A}_{i(1)}, \ldots, b_t \in \mathcal{A}_{i(t)}$, one has

$$\varphi(b_1 \cdots b_t) = \varphi(b_1) \cdots \varphi(b_t).$$

3.2. Tensor Independence

Inspired by the discussion in the previous section, the following definition extends the concept of classical independence to arbitrary noncommutative probability spaces.

Definition 3.4. Let $(\mathcal{A}, \varphi)$ be a noncommutative probability space and let I be an arbitrary indexing set. For every $i \in I$, let C_i be a collection of variables in $\mathcal{A}$, and let $\mathcal{A}_i \coloneqq \mathrm{alg}(1_\mathcal{A}, C_i)$ be the unital algebra generated by C_i. The collections in $\mathbf{C} = (C_i : i \in I)$ are said to be **tensor independent**[4] of each other if and only if they commute (that is, for every distinct $i, j \in I$ and $a_i \in C_i, a_j \in C_j$, one has $a_i a_j = a_j a_i$) and for every collection $i(1), \ldots, i(t) \in I$ of distinct indices and variables $a_1 \in \mathcal{A}_{i(1)}, \ldots, a_t \in \mathcal{A}_{i(t)}$, one has

$$\varphi(a_1 \cdots a_t) = \varphi(a_1) \cdots \varphi(a_t). \tag{3.3}$$

Similarly, if $(\mathcal{A}, *, \varphi)$ is a $*$-probability space and $\mathcal{A}_i = *\text{-alg}(1_\mathcal{A}, C_i)$ is defined as the unital $*$-algebra generated by C_i for every $i \in I$, then the collections in $\mathbf{C} = (C_i : i \in I)$ are said to be **$*$-tensor independent** of each other if and only if they commute and for every collection $i(1), \ldots, i(t) \in I$ of distinct indices and variables $a_1 \in \mathcal{A}_{i(1)}, \ldots, a_t \in \mathcal{A}_{i(t)}$, one has

$$\varphi(a_1 \cdots a_t) = \varphi(a_1) \cdots \varphi(a_t).$$

[4] This appellation will be explained later in this chapter.

Proposition 3.5. *Let $(\mathcal{A},\varphi)$ be a noncommutative probability space and let I be an indexing set. For every $i\in I$, let C_i be a collection of variables in $\mathcal{A}$, and let $\mathcal{A}_i \coloneqq \operatorname{alg}(1_{\mathcal{A}},C_i)$ be the unital algebra generated by C_i. Then, the collections in $\mathbf{C}=(C_i : i\in I)$ are tensor independent of each other if and only if for every collection $i(1),\dots,i(t)\in I$ of distinct indices and variables $a_1\in\mathcal{A}_{i(1)},\dots,a_t\in\mathcal{A}_{i(t)}$ such that $\varphi(a_1)=\cdots=\varphi(a_t)=0$, one has*

$$\varphi(a_1\cdots a_t)=0 \tag{3.4}$$

Proof. The fact that (3.3) implies (3.4) is trivial. Suppose that (3.4) holds. We proceed by induction on t:

Suppose that $t=2$, and let $a_1\in\mathcal{A}_{i(1)}, a_2\in\mathcal{A}_{i(2)}$ be arbitrary. Then, it clearly follows from (3.4) that

$$\varphi(a_1^\circ a_2^\circ)=\varphi\big((a_1-\varphi(a_1)\cdot 1_{\mathcal{A}})(a_2-\varphi(a_2)\cdot 1_{\mathcal{A}})\big)=0.$$

Thus, by linearity of φ, one has

$$0=\varphi(a_1a_2)-\varphi(a_1)\varphi(a_2)-\varphi(a_2)\varphi(a_1)+\varphi(a_1)\varphi(a_2),$$

from which it clearly follows that $\varphi(a_1a_2)=\varphi(a_1)\varphi(a_2)$. Thus, (3.3) holds for the base case $t=2$.

Suppose that, for some fixed $t\in\mathbb{N}$, (3.3) holds for every $s<t$. Let $i(1),\dots,i(t)\in I$ be a collection of distinct indices, and let $a_1\in\mathcal{A}_{i(1)},\dots,a_t\in\mathcal{A}_{i(t)}$ be arbitrary variables. Then (3.4) implies that

$$\varphi(a_1^\circ\cdots a_t^\circ)=\varphi\big((a_1-\varphi(a_1)\cdot 1_{\mathcal{A}})\cdots(a_t-\varphi(a_t)\cdot 1_{\mathcal{A}})\big)=0,$$

and hence

$$0=\varphi(a_1\cdots a_t)+\sum_{s=1}^{t}(-1)^s\sum_{1\leqslant j(1)<\cdots<j(s)\leqslant t}\varphi(a_{j(1)})\cdots\varphi(a_{j(s)})\varphi(a_{j'(1)}\cdots a_{j'(t-s)})$$

where $\{j'(1),\dots,j'(t-s)\}=\{1,\dots,t\}\setminus\{j(1),\dots,j(s)\}$. Given that (3.3) holds for every $s<t$,

$$\varphi(a_{j'(1)}\cdots a_{j'(t-s)})=\varphi(a_{j'(1)})\cdots\varphi(a_{j'(t-s)})$$

for every $1\leqslant j(1)<\cdots<j(s)\leqslant t$, and thus

$$\begin{aligned}
0=&\varphi(a_1\cdots a_t)+\sum_{s=1}^{t}(-1)^s\sum_{1\leqslant j(1)<\cdots<j(s)\leqslant t}\varphi(a_1)\cdots\varphi(a_t)\\
=&\varphi(a_1\cdots a_t)+\varphi(a_1)\cdots\varphi(a_t)\left(\sum_{s=1}^{t}(-1)^s\binom{t}{s}\right)\\
=&\varphi(a_1\cdots a_t)+\varphi(a_1)\cdots\varphi(a_t)\left(\sum_{s=0}^{t}(-1)^s\binom{t}{s}\right)-\varphi(a_1)\cdots\varphi(a_t).
\end{aligned}$$

According to the binomial theorem,

$$0=((-1)+1)^t=\sum_{s=0}^{t}(-1)^s(1)^{t-s}\binom{t}{s}=\sum_{s=0}^{t}(-1)^s\binom{t}{s},$$

and thus $\varphi(a_1\cdots a_t)=\varphi(a_1)\cdots\varphi(a_t)$, as desired. □

Remark 3.6. Equation (3.4) in the statement of the previous proposition is obviously equivalent to the following: For every collection $i(1), \ldots, i(t) \in I$ of distinct indices and variables $a_1 \in \mathcal{A}_{i(1)}, \ldots, a_t \in \mathcal{A}_{i(t)}$, one has

$$a_1^\circ \cdots a_t^\circ = \left(a_1^\circ \cdots a_t^\circ\right)^\circ \tag{3.5}$$

(i.e., $\varphi(a_1^\circ \cdots a_t^\circ) = 0$). Thus, one may argue that tensor independence is satisfying from an intuitive point of view, in that (3.5) can be interpreted as the assertion that the information that φ extracts from the individual algebras $\mathcal{A}_i = \operatorname{alg}(1_\mathcal{A}, C_i)$ completely determines the information that φ extracts from any product $a_1 \cdots a_t$ with $a_1 \in \mathcal{A}_{i(1)}, \ldots, a_t \in \mathcal{A}_{i(t)}$ and $i(1), \ldots, i(t) \in I$ distinct. Indeed, equation (3.5) states that no information (i.e., a deterministic quantity of the form $\lambda \cdot 1_\mathcal{A}$ with $\lambda \neq 0$) will emerge from the product $a_1^\circ \cdots a_t^\circ$ of the random parts a_j° of the a_j (provided the a_j are from different algebras $\mathcal{A}_{i(j)}$), and (3.3) provides an explicit formula to express the information φ extracts from $a_1 \cdots a_t$ in terms of the information it extracts from the individual a_j.[5]

In the following example, we see that the algebraic concept of tensor independence is actually closely linked to the concept of freeness for commuting groups:

Example 3.7. Let G be a group (with identity e) and let $(G_i : i \in I)$ be a collection of subgroups of G that commute—that is, for every two distinct indices $i, j \in I$ and $g_i \in G_i$ and $g_j \in G_j$, one has $g_i g_j = g_j g_i$. The G_i are said to be free in the abelian sense if there exists no nontrivial relations between elements from different subgroups G_i—that is, for every collection of distinct indices $i(1), \ldots, i(t) \in I$, elements $g_1 \in G_{i(1)} \setminus \{e\}, \ldots, g_t \in G_{i(t)} \setminus \{e\}$, one has $g_1 \cdots g_t \neq e$. One can prove without much difficulty that subgroups G_i are free in the abelian sense if and only if the $*$-algebras $\mathbb{C}G_i$ are $*$-tensor independent of each other in $(\mathbb{C}G, *, \tau_e)$.

3.2.1. Abstract Model for Tensor Independent Variables. One of the most important result concerning independence in the classical theory of probability is the existence and uniqueness (under favourable conditions) of product σ-algebra and product measures.[6] Indeed, this result guarantees that, given any collection of $\mathbb{C}$-valued random variables $(X_i : i \in I)$—each defined on a respective probability space $(\Omega_i, \mathscr{A}_i, P_i)$—there exists another probability space $(\Omega, \mathscr{A}, P)$ (namely, the product space $\left(\times_i \Omega_i, \times_i \mathscr{A}_i, \times_i P_i\right)$) in which

(1) events $E_i \in \mathscr{A}_i$ can be embedded in a natural way that preserves their probability, namely, $E_i \hookrightarrow \times_{j \in I} F_j$, where $F_i = E_i$ and $F_j = \Omega_j$ if $j \neq i$, and $P(E_i) = \times_j P_j(\times_j F_j)$;

(2) random variables $X_i : \Omega_i \to \mathbb{C}$ can be embedded in a natural way that preserves their distribution, namely, $X_i \hookrightarrow \times_{j \in I} Y_j$, where $Y_i = X_i$ and $Y_j = \mathbf{1}_{\Omega_j}$ if $j \neq i$, and X_i and $\times_j Y_j$ are identically distributed; and most importantly

(3) the embeddings of the X_i in the product space $\left(\times_i \Omega_i, \times_i \mathscr{A}_i, \times_i P_i\right)$ are independent of each other.

This, in particular, allows one, to make statements of the form "Let $(X_i : i \in I)$ be a collection of independent random variables with respective distributions μ_{X_i}" without fear of running into issues of existence, and thus enables one to study the properties of independence abstractly.

[5] See Remark 2.11 for a definition of the random part of a random variable and the information φ extracts from random variables.

[6] See for instance [**Ta11**] Section 2.4.

Thankfully, there exists an analogous construction in noncommutative probability from which the appellation *tensor independence* originates: Let $\big((\mathcal{A}_i, \varphi_i) : i \in I\big)$ be a collection of noncommutative probability spaces. Define the algebra $\mathcal{A}$ as the tensor product $\otimes_i \mathcal{A}_i$ equipped with the following product operation:

$$(\otimes_i a_i)(\otimes_i b_i) = \otimes_i a_i b_i, \qquad a_i, b_i \in \mathcal{A}_i,\ i \in I$$

and define the functional φ on $\mathcal{A}$ as $\otimes_i \varphi$, that is,

$$\otimes_i \varphi_i(\otimes_i a_i) = \prod_{i \in I} \varphi_i(a_i), \qquad a_i \in \mathcal{A}_i,\ i \in I.$$

(Note that the above product is well-defined even if I is infinite, since, in this case, only finitely many of the a_i in $\otimes_i a_i$ are permitted to be different from the unit $1_{\mathcal{A}_i}$.) For a fixed $i \in I$ and $a \in \mathcal{A}_i$, the element a can naturally be embedded in $\mathcal{A}$ as

$$\begin{array}{ccc} \mathcal{A}_i & \to & \mathcal{A} \\ a & \hookrightarrow & \otimes_j a_j, \end{array}$$

where $a_i = a$ and $a_j = 1_{\mathcal{A}_i}$ for all $j \neq i$. One easily checks that distributions are preserved by the above embedding, and that for any family of collections $(C_i \subset \mathcal{A}_i : i \in I)$, the embeddings of C_i into $\mathcal{A}$ are tensor independent of each other.

In the sequel, we will make the following notational shortcut:

Notation 3.8. Let $\big((\mathcal{A}_i, \varphi_i) : i \in I\big)$ be a collection of noncommutative probability spaces, $\mathcal{A} = \otimes_i \mathcal{A}_i$, and $\varphi = \otimes_i \varphi_i$. Given an element of the form $\otimes_i a_i \in \mathcal{A}$, suppose that $i(1), \ldots, i(t) \in I$ are the indices such that $a_i \neq 1_{\mathcal{A}_i}$ if and only if $i \in \{i(1), \ldots, i(t)\}$. Then, we will denote

$$\otimes_i a_i = (a_{i(1)} \otimes \cdots \otimes a_{i(t)})_I. \tag{3.6}$$

The following proposition will later be useful in technical proofs:

Proposition 3.9. *Let $\big((\mathcal{A}_i, \varphi_i) : i \in I\big)$ be a collection of noncommutative probability spaces, $\mathcal{A} = \otimes_i \mathcal{A}_i$, and $\varphi = \otimes_i \varphi_i$. Then, every $a \in \mathcal{A}$ may be written in the form (using the notation introduced in (3.6))*

$$a = \varphi(a) \cdot 1_{\mathcal{A}} + \sum_{k=1}^{K} \sum_{i(1),\ldots,i(k) \in I \text{ distinct}} A_{i(1),\ldots,i(k)}, \tag{3.7}$$

where for every $k \leqslant K$ and distinct indices $i(1), \ldots, i(k) \in I$, $A_{i(1),\ldots,i(k)}$ is either equal to the zero vector $0_{\mathcal{A}}$ or a linear combination of terms of the form $(a_{i(1)} \otimes \cdots \otimes a_{i(k)})_I$, where $a_{i(1)} \in \mathcal{A}^\circ_{i(1)}, \ldots, a_{i(k)} \in \mathcal{A}^\circ_{i(k)}$, that is, $\varphi_{i(j)}(a_{i(j)}) = 0$ for all $j \leqslant k$.

Proof Outline. Let $a \in \mathcal{A}$ be arbitrary. Then, by definition of $\mathcal{A}$, a can be written as a sum

$$a = \sum_{l=1}^{L} \otimes_i a_{l,i}, \tag{3.8}$$

where $L \in \mathbb{N}$ and for each $l \leqslant L$, $a_{l,i} \neq 1_{\mathcal{A}_i}$ for finitely many i. Thus, for every $l \leqslant L$ there exists $t_l \in \mathbb{N}$ and distinct indices $i_l(1), \ldots, i_l(t_l) \in I$ such that

$$\otimes_i a_{l,i} = (a_{i_l(1)} \otimes \cdots \otimes a_{i_l(t_l)})_I.$$

By writing every element $a_{i_l(1)}$ as $\varphi_{i_l(j)}(a_{i_l(1)}) + a^\circ_{i_l(1)}$, then using the bilinearity of tensor products, and then rearranging the elements in the sum (3.8), one obtains (3.7) (if no term of the form $(a_{i(1)} \otimes \cdots \otimes a_{i(k)})_I$ with $i(1), \ldots, i(k) \in I$ occur in (3.8), then $A_{i(1),\ldots,i(k)} = 0_{\mathcal{A}}$). $\square$

Example 3.10. Let $\big((\mathcal{A}_i, \varphi_i) : i \in I\big)$ (with $I = \{1, 2, 3\}$) be noncommutative probability spaces, let $\mathcal{A} = \otimes_i \mathcal{A}_i$ and $\varphi = \otimes_i \varphi_i$, and let $a_1, b_1 \in \mathcal{A}_1$, $a_2, b_2 \in \mathcal{A}_2$, and $a_3 \in \mathcal{A}_3$. Then,

$$\begin{aligned}
a =& (a_1 \otimes a_2)_I + (b_1 \otimes b_2)_I + (a_2 \otimes a_3)_I \\
=& \Big(\big(\varphi_1(a_1) \cdot 1_{\mathcal{A}_1} + a_1^\circ\big) \otimes \big(\varphi_2(a_2) \cdot 1_{\mathcal{A}_2} + a_2^\circ\big) \Big)_I \\
&\quad + \Big(\big(\varphi_1(b_1) \cdot 1_{\mathcal{A}_1} + b_1^\circ\big) \otimes \big(\varphi_2(b_2) \cdot 1_{\mathcal{A}_2} + b_2^\circ\big) \Big)_I \\
&\quad + \Big(\big(\varphi_2(a_2) \cdot 1_{\mathcal{A}_2} + a_2^\circ\big) \otimes \big(\varphi_3(a_3) \cdot 1_{\mathcal{A}_3} + a_3^\circ\big) \Big)_I \\
=& \big(\varphi_1(a_1)\varphi_2(a_2) + \varphi_1(b_1)\varphi_2(b_2) + \varphi_2(a_2)\varphi_3(a_3)\big) \cdot 1_{\mathcal{A}} \\
&\quad + \big(\varphi_1(a_1) \cdot a_2^\circ\big)_I + \big(\varphi_2(a_2) \cdot a_1^\circ\big)_I + (a_1^\circ \otimes a_2^\circ)_I \\
&\quad + \big(\varphi_1(b_1) \cdot b_2^\circ\big)_I + \big(\varphi_2(b_2) \cdot b_1^\circ\big)_I + (b_1^\circ \otimes b_2^\circ)_I \\
&\quad + \big(\varphi_2(a_2) \cdot a_3^\circ\big)_I + \big(\varphi_3(a_3) \cdot a_2^\circ\big)_I + (a_2^\circ \otimes a_3^\circ)_I \\
=& \varphi(a) \cdot 1_{\mathcal{A}} + \sum_{k=1}^{2} \sum_{i(1),\ldots,i(k) \text{ distinct}} A_{i(1),\ldots,i(k)},
\end{aligned}$$

where

(1) $\varphi(a) = \varphi_1(a_1)\varphi_2(a_2) + \varphi_1(b_1)\varphi_2(b_2) + \varphi_2(a_2)\varphi_3(a_3)$;

(2) $A_1 = \big(\varphi_2(a_2) \cdot a_1^\circ\big)_I + \big(\varphi_2(b_2) \cdot b_1^\circ\big)_I$;

(3) $A_2 = \big(\varphi_1(a_1) \cdot a_2^\circ\big)_I + \big(\varphi_1(b_1) \cdot b_2^\circ\big)_I + \big(\varphi_3(a_3) \cdot a_2^\circ\big)_I$;

(4) $A_3 = \big(\varphi_2(a_2) \cdot a_3^\circ\big)_I$;

(5) $A_{1,2} = (a_1^\circ \otimes a_2^\circ)_I + (b_1^\circ \otimes b_2^\circ)_I$;

(6) $A_{1,3} = 0_{\mathcal{A}}$; and

(7) $A_{2,3} = (a_2^\circ \otimes a_3^\circ)_I$.

Definition 3.11. Let $\big((\mathcal{A}_i, *_i, \varphi_i) : i \in I\big)$ be a collection of $*$-probability spaces, let $\mathcal{A} = \otimes_i \mathcal{A}_i$, and define the operator $\cdot^* : \mathcal{A} \to \mathcal{A}$ as follows:

$$(\otimes_i a_i)^* = \otimes_i a_i^{*_i}, \qquad a_i \in \mathcal{A}_i,\ i \in I. \tag{3.9}$$

Then, one easily checks that $(\mathcal{A}, *)$ is a $*$-algebra.

Example 3.12. Let $(G_i : i \in I)$ be a collection of groups (with respective neutral elements e_i), and let $G = \times_i G_i$ be the direct product of the G_i. Then, for every $j \in I$ and $g \in G_j$, the element g can be embedded in G as

$$\begin{array}{ccc} G_j & \to & G \\ g & \hookrightarrow & \times_i g_i, \end{array}$$

where $g_j = g$ and $g_i = e_i$ for all $i \neq j$. One easily checks that the order of every element is preserved by this embedding, and that the embeddings of the G_i into G are free in the abelian sense. In fact, $\big(\mathbb{C}(\times_i G_i), *, \tau_e\big)$ (where e denotes the neutral element in $\times_i G_i$) is isomorphic as a $*$-probability space to $(\otimes_i \mathbb{C}G_i, *, \otimes_i \tau_{e_i})$.

We end this subsection with the following theorem, which establishes that the tensor product preserves the main structure of interest in a $*$-probability space:

Theorem 3.13. *Let $\big((\mathcal{A}_i, *_i, \varphi_i) : i \in I\big)$ be a collection of noncommutative probability spaces, let $\mathcal{A} = \otimes_i \mathcal{A}_i$, let $\varphi = \otimes_i \varphi_i$, and let $*$ be defined as in (3.9). Then, $(\mathcal{A}, *, \varphi)$ is a $*$-probability space.*

Proof. In order to prove that $(\mathcal{A}, *, \varphi)$ is a $*$-probability space, we need only show that φ is positive. Let $a \in \mathcal{A}$ be arbitrary. Let us write a as in equation (3.7), that is

$$a = C \cdot 1_{\mathcal{A}} + \sum_{k=1}^{K} \sum_{i(1),\ldots,i(k) \in I \text{ distinct}} A_{i(1),\ldots,i(k)}.$$

In order to simplify notation, let us denote

$$A = \sum_{k=1}^{K} \sum_{i(1),\ldots,i(k) \in I \text{ distinct}} A_{i(1),\ldots,i(k)}.$$

Then, $aa^* = |C|^2 \cdot 1_{\mathcal{A}} + C \cdot A^* + \overline{C} \cdot A + AA^*$. Given that $|C|^2 \geqslant 0$ and $\varphi(A) = \varphi(A^*) = 0$, the proof that $\varphi(aa^*) \geqslant 0$ reduces to demonstrating that $\varphi(AA^*) \geqslant 0$.

$$AA^* = \sum_{1 \leqslant k,l \leqslant K} \sum_{\substack{i(1),\ldots,i(k) \in I \text{ distinct} \\ j(1),\ldots,j(l) \in I \text{ distinct}}} A_{i(1),\ldots,i(k)} A^*_{j(1),\ldots,j(l)}.$$

If $\{i(1), \ldots, i(k)\} \neq \{j(1), \ldots, j(l)\}$, then $\varphi(A_{i(1),\ldots,i(k)} A^*_{j(1),\ldots,j(l)}) = 0$, given that for every variables

$$a_{i(1)} \in \mathcal{A}^\circ_{i(1)}, \ldots, a_{i(k)} \in \mathcal{A}^\circ_{i(k)} \text{ and } a_{j(1)} \in \mathcal{A}^\circ_{j(1)}, \ldots, a_{j(t)} \in \mathcal{A}^\circ_{j(t)}$$

there will be at least one index (say, $i(m)$ or $j(n)$) such that

$$\varphi\big((a_{i(1)} \otimes \cdots \otimes a_{i(k)})_I (a^*_{j(1)} \otimes \cdots \otimes a^*_{j(l)})_I\big) = \varphi_{i(m)}(a_{i(m)}) \cdots$$

or

$$\varphi\big((a_{i(1)} \otimes \cdots \otimes a_{i(k)})_I (a^*_{j(1)} \otimes \cdots \otimes a^*_{j(l)})_I\big) = \varphi_{j(n)}(a_{j(n)}) \cdots,$$

and $\varphi_{i(m)}(a_{i(m)}) = \varphi_{j(n)}(a_{j(n)}) = 0$. Therefore,

$$\varphi(AA^*) = \sum_{k=1}^{K} \sum_{i(1),\ldots,i(k) \in I \text{ distinct}} \varphi(A_{i(1),\ldots,i(k)} A^*_{i(1),\ldots,i(k)})$$

Thus, we need only prove that $\varphi(A_{i(1),\ldots,i(k)}A^*_{i(1),\ldots,i(k)}) \geqslant 0$ for all distinct $i(1),\ldots,i(k) \in I$.

Let $i(1),\ldots,i(k) \in I$ be a fixed collection of distinct indices, and let

$$A_{i(1),\ldots,i(k)} = \sum_{l=1}^{L} (a_1^{(l)} \otimes \cdots \otimes a_k^{(l)})_I,$$

where, for each $l \leqslant L$, one has $a_1^{(l)} \in \mathcal{A}^\circ_{i(1)}, \ldots, a_k^{(l)} \in \mathcal{A}^\circ_{i(k)}$. Then,

$$\begin{aligned}\varphi(A_{i(1),\ldots,i(k)}A^*_{i(1),\ldots,i(k)}) &= \sum_{l,s=1}^{L} \varphi\big((a_1^{(l)} \otimes \cdots \otimes a_k^{(l)})_I((a_1^{(s)})^* \otimes \cdots \otimes (a_k^{(s)})^*)_I\big) \\ &= \sum_{l,s=1}^{L} \varphi_{i(1)}\big(a_1^{(l)}(a_1^{(s)})^*\big) \cdots \varphi_{i(k)}\big(a_k^{(l)}(a_k^{(s)})^*\big).\end{aligned}$$

For each $m \leqslant k$, define the matrix $B_m = \big(\varphi_{i(m)}\big(a_m^{(l)}(a_m^{(s)})^*\big) : 1 \leqslant l, s \leqslant L\big)$. According to Proposition A.10, every such matrix is positive.[7] Furthermore, if follows from Proposition A.9 that the Schur product[8]

$$S(B_1,\ldots,B_k) = \Big(\varphi_{i(1)}\big(a_1^{(l)}(a_1^{(s)})^*\big) \cdots \varphi_{i(k)}\big(a_k^{(l)}(a_k^{(s)})^*\big) : 1 \leqslant l, s \leqslant L\Big)$$

is positive. As $\varphi(A_{i(1),\ldots,i(k)}A^*_{i(1),\ldots,i(k)}) = \Big\langle S(B_1,\ldots,B_k)z, z\Big\rangle$, where $z = (1,\ldots,1) \in \mathbb{C}^L$, it follows from the definition of positivity for matrices that $\varphi(A_{i(1),\ldots,i(k)}A^*_{i(1),\ldots,i(k)}) \geqslant 0$, as desired. □

Remark 3.14. It is interesting to note that the tensor product of noncommutative probability spaces equipped with traces is also equipped with a trace, and the tensor product of $*$-probability spaces equipped with faithful functionals is also equipped with a faithful functional.

3.2.2. Tensor Independence as a Fundamentally Commutative Concept. Given that many interesting variables in $*$-probability spaces are not commutative, it is natural to wonder if it is possible to extend Definition 3.4 and equation (3.3) to the noncommutative case. The obvious place to start would be the following:

Tentative Definition 3.15. Let $(\mathcal{A},\varphi)$ be a noncommutative probability space and let I be an indexing set. For every $i \in I$, let C_i be a collection of variables in $\mathcal{A}$, and let $\mathcal{A}_i = \operatorname{alg}(1_\mathcal{A}, C_i)$ be the unital algebra generated by C_i. If the C_i are *tensor independent in the noncommutative sense*, then for every collection $i(1),\ldots,i(t) \in I$ of distinct indices and variables $a_1 \in \mathcal{A}_{i(1)},\ldots,a_t \in \mathcal{A}_{i(t)}$, as in the commutative case, one has $\varphi(a_1\cdots a_t) = \varphi(a_1)\cdots\varphi(a_t)$.

Notice that Tentative Definition 3.15 is phrased like a necessary condition for noncommutative tensor independence rather than an actual definition. This can be explained by the fact that, in the case where the collections of variables C_i commute, any mixed $*$-moment in the $\mathcal{A}_i = \operatorname{alg}(1_\mathcal{A}, C_i)$ can be written as an expression of the form $\varphi(a_1\cdots a_t)$ where the variables $a_1 \in \mathcal{A}_{i(1)},\ldots,a_t \in \mathcal{A}_{i(t)}$ are such that $i(1),\ldots,i(t) \in I$ are distinct (for example, $\varphi(a_1a_2a_1a_2)$ with $a_1 \in \mathcal{A}_{i(1)}$, $a_2 \in \mathcal{A}_{i(2)}$ and $i(1) \neq i(2)$ can be reduced to $\varphi(a_1^2a_2^2)$). This,

[7] See Definition A.8.

[8] See Proposition A.9.

however, is not necessarily the case when the collections C_i do not commute. Thus, if the notion of noncommutative tensor independence is to be as powerful as the commutative one, there should be a rule for decomposing mixed moments of the form $\varphi(a_1a_2a_1a_2)$ or, more generally, $\varphi(a_1\cdots a_t)$, where $a_1 \in \mathcal{A}_{i(1)}, \ldots, a_t \in \mathcal{A}_{i(t)}$ and $i(1), \ldots, i(t) \in I$ are such that $i(1) \neq i(2), i(2) \neq i(3), \ldots, i(t-1) \neq i(t)$ (with some repetitions in nonconsecutive indices allowed, such as $i(1) = i(3)$) in terms of moments in the individual a_i.

In light of the previous remarks, one may attempt to simply extend Tentative Definition 3.15 to the following:

Tentative Definition 3.16. Let $(\mathcal{A}, \varphi)$ be a noncommutative probability space and let I be an indexing set. For every $i \in I$, let C_i be a collection of variables in $\mathcal{A}$, and let $\mathcal{A}_i = \operatorname{alg}(1_\mathcal{A}, C_i)$ be the unital algebra generated by C_i. If the C_i are *tensor independent in the noncommutative sense*, then for every collection $i(1), \ldots, i(t) \in I$ such that $i(1) \neq i(2) \neq \cdots \neq i(t)$ and variables $a_1 \in \mathcal{A}_{i(1)}, \ldots, a_t \in \mathcal{A}_{i(t)}$, one has

$$\varphi(a_1 \cdots a_t) = \varphi(a_1) \cdots \varphi(a_t). \tag{3.10}$$

However, this definition is problematic: Suppose that two random variables $a_1, a_2 \in \mathcal{A}$ are independent in the sense of Tentative Definition 3.16, and let $\mathcal{A}_1 = \operatorname{alg}(1_\mathcal{A}, a_1)$ and $\mathcal{A}_2 = \operatorname{alg}(1_\mathcal{A}, a_2)$. Then, since $1_\mathcal{A}$ is in both $\mathcal{A}_1$ and $\mathcal{A}_2$, for every nontrivial $*$-word $M(x) = x^{n(1)} \cdots x^{n(t)}$ (with $n(1), \ldots, n(t) \in \{1, *\}$), one has

$$\begin{aligned}
\varphi\big(M(a_1)\big) =& \varphi\left(a_1^{n(1)} \cdots a_1^{n(t)}\right) = \varphi\left(a_1^{n(1)} \cdot 1_\mathcal{A} \cdot a_1^{n(2)} \cdots a_1^{n(t)}\right) = \varphi\left(a_1^{n(1)}\right) \varphi\left(a_1^{n(2)} \cdots a_1^{n(t)}\right) \\
=& \varphi\left(a_1^{n(1)}\right) \varphi\left(a_1^{n(2)} \cdot 1_\mathcal{A} \cdot a_1^{n(3)} \cdots a_1^{n(t)}\right) = \varphi\left(a_1^{n(1)}\right) \varphi\left(a_1^{n(2)}\right) \varphi\left(a_1^{n(3)} \cdots a_1^{n(t)}\right) \\
& \ldots \\
=& \varphi\left(a_1^{n(1)}\right) \cdots \varphi\left(a_t^{n(t)}\right),
\end{aligned} \tag{3.11}$$

and similarly, $\varphi\big(M(a_2)\big) = \varphi\left(a_2^{n(1)}\right) \cdots \varphi\left(a_2^{n(t)}\right)$. Thus, the concept of independence for noncommutative variables as it is currently defined imposes some very strong conditions on the $*$-distributions of the concerned variables. In fact, if $(\mathcal{A}, *, \varphi)$ is a $*$-probability space such that φ is faithful, given that $\mathbf{Var}[a] = \varphi(aa^*) - \varphi(a)\varphi(a^*)$, independence as defined in Tentative Definition 3.16 is in many situations restricted to deterministic (that is, constant multiples of the unit $1_\mathcal{A}$) variables. This is in stark contrast with tensor independence for commutative random variables, which imposes no condition on the distributions of the considered random variables. Indeed, as guaranteed by the tensor product of noncommutative probability spaces (see Subsection 3.2.1), collections of tensor independent random variables with any distribution can be generated at will.

One may attempt to salvage Tentative Definition 3.16 by making the following adjustment:

Tentative Definition 3.17. Let $(\mathcal{A}, \varphi)$ be a noncommutative probability space and let I be an indexing set. For every $i \in I$, let C_i be a collection of variables in $\mathcal{A}$, and let $\mathcal{A}_i = \operatorname{alg}(1_\mathcal{A}, C_i)$ be the unital algebra generated by C_i. If the C_i are *tensor independent in the noncommutative sense*, then for every collection $i(1), \ldots, i(t) \in I$ such that $i(1) \neq i(2) \neq \cdots \neq i(t)$ and variables $a_1 \in \mathcal{A}_{i(1)}, \ldots, a_t \in \mathcal{A}_{i(t)}$, one has

$$\varphi(a_1 \cdots a_t) = \varphi(a_{j(1)}) \cdots \varphi(a_{j(p)}) \varphi\big(a_{j'(1)} \cdots a_{j'(t-p)}\big),$$

where $1 \leqslant j(1), \ldots, j(p) \leqslant t$ are the indices such that $a_{j(s)}$ is deterministic, and $1 \leqslant j'(1) < \ldots < j'(t-p) \leqslant t$ are the remaining indices. Then, the remaining nondeterministic mixed $*$-moment $\varphi\big(a_{j'(1)} \cdots a_{j'(t-p)}\big)$ can be decomposed using equation (3.10).

Thus, for example, given two nondeterministic and noncommutative variables a_1 and a_2, Tentative Definition 3.17 would imply that if they are independent, then

$$\varphi\big(a_1 a_2 (2 \cdot 1_{\mathcal{A}}) a_2 a_1\big) = 2\varphi\big(a_1 a_2 a_2 a_1\big) = 2\varphi\big(a_1 a_2^2 a_1\big) = 2\varphi(a_1)^2 \varphi(a_2^2),$$

instead of $2\varphi(a_1)^2\varphi(a_2)^2$ for Tentative Definition 3.16. However, Tentative Definition 3.17 is still problematic: Suppose that $a_1, a_2 \in \mathcal{A}$ are independent of each other in the sense of Tentative Definition 3.17, and that a_2 is nondeterministic. Let P_1 and P_2 be arbitrary $*$-polynomials and let P be such that $\varphi\big(P(a_2)\big) \neq 0$ (it is always possible to find such a polynomial, provided a_2 is nondeterministic). On the one hand,

$$\varphi\Big(P_1(a_1)\Big(P(a_2) - \varphi\big(P(a_2)\big) \cdot 1_{\mathcal{A}_2}\Big) P_2(a_1)\Big) \\ = \varphi\big(P_1(a_1)\big)\varphi\Big(P(a_2) - \varphi\big(P(a_2)\big) \cdot 1_{\mathcal{A}_2}\Big)\varphi\big(P_2(a_1)\big) = 0,$$

and on the other hand,

$$\begin{aligned} &\varphi\Big(P_1(a_1)\Big(P(a_2) - \varphi\big(P(a_2)\big) \cdot 1_{\mathcal{A}_2}\Big) P_2(a_1)\Big) \\ =&\varphi\big(P_1(a_1)P(a_2)P_2(a_1)\big) - \varphi\big(P(a_2)\big)\varphi\big(P_1(a_1)P_2(a_1)\big) \\ =&\varphi\big(P_1(a_1)\big)\varphi\big(P(a_2)\big)\varphi\big(P_2(a_1)\big) - \varphi\big(P(a_2)\big)\varphi\big(P_1(a_1)P_2(a_1)\big) \\ =&\varphi\big(P(a_2)\big)\Big(\varphi\big(P_1(a_1)\big)\varphi\big(P_2(a_1)\big) - \varphi\big(P_1(a_1)P_2(a_1)\big)\Big). \end{aligned}$$

Thus, one could show that (3.11) holds for a_i for any $*$-word M, leading to the same shortcoming that Tentative Definition 3.17 has.

Consequently, it appears that, in parallel with Theorem 2.20 (which states that many of the commutative noncommutative probability spaces can be realized as an algebra of $\mathbb{C}$-valued random variables equipped with the expected value $\mathbf{E}$), tensor independence, i.e., the classical notion of independence for $\mathbb{C}$-valued random variables in the context of free probability, is a fundamentally commutative notion.

3.3. Definition and Basic Properties of Free Independence

As it turns out, the notion of tensor independence for commuting random variables does have a noncommutative analog, which is called free independence, or freeness. However, a naive extension of the rule for decomposing mixed moments in tensor independent random variables (i.e., equation (3.3)) to arbitrary mixed moments in noncommutative random variables (as it was attempted in Tentative Definitions 3.15, 3.16, and 3.17) is not the way to derive the rules of free independence.

In Remark 3.6, it was argued that the notion of tensor independence is satisfying from an intuitive point of view, as it implies that no new information emerges from the product of the centering (i.e., the random parts) $a_1^\circ \cdots a_t^\circ$ of distinct tensor independent variables $a_1, \ldots, a_t$

(since $(a_1^\circ \cdots a_t^\circ)^\circ = a_1^\circ \cdots a_t^\circ$). Furthermore, in Example 3.7, we highlighted the similarity between this concept and that of freeness for commuting subgroups, in that commuting subgroups are free in the abelian sense if no relation emerges from a product $g_1 \cdots g_t$ of nontrivial elements $g_1, \ldots, g_t$ from distinct groups that are free. In this context, it appears that equation (3.3) can be seen as consequence of tensor independence rather than its conceptual foundation. As we will see, the ideas of Remark 3.6 and the definition of freeness for noncommuting subgroups can be used as an inspiration to derive the definition of free independence:

Definition 3.18. Let $(\mathcal{A}, \varphi)$ be a noncommutative probability space and let I be an indexing set. For every $i \in I$, let C_i be a collection of variables in $\mathcal{A}$, and let $\mathcal{A}_i \coloneqq \mathrm{alg}(1_\mathcal{A}, C_i)$ be the unital algebra generated by C_i. The collections in $\mathbf{C} = (C_i : i \in I)$ are said to be **free** from each other if and only if for every collection $i(1), \ldots, i(t) \in I$ of indices such that $i(1) \neq i(2) \neq \cdots \neq i(t)$ and variables $a_1 \in \mathcal{A}_{i(1)}, \ldots, a_t \in \mathcal{A}_{i(t)}$, one has

$$(a_1^\circ \cdots a_t^\circ)^\circ = a_1^\circ \cdots a_t^\circ. \tag{3.12}$$

(i.e., $\varphi(a_1^\circ \cdots a_t^\circ) = 0$). Similarly, if $(\mathcal{A}, *, \varphi)$ is a $*$-probability space and $\mathcal{A}_i = *\text{-alg}(1_\mathcal{A}, C_i)$ is defined as the unital $*$-algebra generated by C_i for every $i \in I$, then the collections in $\mathbf{C} = (C_i : i \in I)$ are said to be **$*$-free** from each other if and only if for every collection $i(1), \ldots, i(t) \in I$ of indices such that $i(1) \neq i(2) \neq \cdots \neq i(t)$ and variables $a_1 \in \mathcal{A}_{i(1)}, \ldots, a_t \in \mathcal{A}_{i(t)}$, (3.12) holds.

Example 3.19. Let G be a group (with identity e) and let $(G_i : i \in I)$ be a collection of subgroups of G that do not commute. Recall that the G_i are said to be free from each other if there exists no nontrivial relations between elements from different subgroups G_i, that is, for every collection of indices $i(1), \ldots, i(t) \in I$ such that $i(1) \neq i(2) \neq \cdots \neq i(t)$ and elements $g_1 \in G_{i(1)} \setminus \{e\}, \ldots, g_t \in G_{i(t)} \setminus \{e\}$, one has $g_1 \cdots g_t \neq e$.[9] One easily sees that subgroups G_i are free if and only if the $*$-algebras $\mathbb{C}G_i$ are $*$-free from each other in $(\mathbb{C}G, *, \tau_e)$ (see for instance **[NS06]** Proposition 5.11).

As one would expect, much like tensor independence, free independence allows one to compute mixed moments in free variables $(a_i : i \in I)$ using the the distributions of the individual variables a_i only. However, as shown in the next proposition, the rule for computing such mixed moments is nowhere near as simple as the rule for tensor independence (i.e., equation (3.3)).

Proposition 3.20 (**[NS06]** Lemma 5.13)**.** *Let $(\mathcal{A}, \varphi)$ be a noncommutative probability space and let I be an indexing set. For every $i \in I$, let C_i be a collection of variables in $\mathcal{A}$, let $\mathcal{A}_i \coloneqq \mathrm{alg}(1_\mathcal{A}, C_i)$ be the unital algebra generated by C_i, and let $\mathcal{B} = \mathrm{alg}(\mathcal{A}_i : i \in I)$ be the unital algebra generated by the $\mathcal{A}_i$. If the C_i are free from each other, then the restriction $\varphi|_\mathcal{B}$ of φ to $\mathcal{B}$ is completely determined by the restrictions $\varphi|_{\mathcal{A}_i}$ of φ to $\mathcal{A}_i$, that is, for every word of the form $M = a_1 \cdots a_t \in \mathcal{B}$ with $a_1 \in \mathcal{A}_{i(1)}, \ldots, a_t \in \mathcal{A}_{i(t)}$ such that $i(1) \neq i(2) \neq \cdots \neq i(t)$, one can write*

$$\varphi(a_1 \cdots a_t) = F\big(M; \varphi(a) : a \in \mathcal{A}_i,\ i \in I\big), \tag{3.13}$$

where $F\big(M; \varphi(a) : a \in \mathcal{A}_i,\ i \in I\big)$ is an expression that depends only on the choice of the word M and the values that φ takes on elements of the individual algebras $\mathcal{A}_i$.

[9] See Section A.1.

Proof. Let $M = a_1 \cdots a_t \in \mathcal{B}$ be a word, where $a_1 \in \mathcal{A}_{i(1)}, \ldots, a_t \in \mathcal{A}_{i(t)}$ and $i(1) \neq i(2) \neq \cdots \neq i(t)$. We proceed by induction on t:

The case $t = 1$ is trivial.

Suppose that $t = 2$. Write a_1 and a_2 as their decompositions in $\mathbb{C}1_{\mathcal{A}} \oplus \mathcal{B}$, that is $a_1 = \varphi(a_1) \cdot 1_{\mathcal{A}} + a_1^\circ$ and $a_2 = \varphi(a_2) \cdot 1_{\mathcal{A}} + a_2^\circ$. Then, given that $\varphi(a_1^\circ) = \varphi(a_2^\circ) = 0$ and a_1 and a_2 are free from each other, one has

$$\begin{aligned}\varphi(a_1a_2) =& \varphi\Big(\big(\varphi(a_1) \cdot 1_{\mathcal{A}} + a_1^\circ\big)\big(\varphi(a_2) \cdot 1_{\mathcal{A}} + a_2^\circ\big)\Big)\\ =& \varphi(a_1)\varphi(a_2) + \varphi(a_1^\circ)\varphi(a_2) + \varphi(a_1)\varphi(a_2^\circ) + \varphi(a_1^\circ a_2^\circ)\\ =& \varphi(a_1)\varphi(a_2).\end{aligned}$$

Thus, in the case where $M = a_1a_2$ with $a_1 \in \mathcal{A}_{i(1)}, a_2 \in \mathcal{A}_{i(2)}$ and $i(1) \neq i(2)$, equation (3.13) holds with

$$F\big(M; \varphi(a) : a \in \mathcal{A}_i,\ i \in I\big) = \varphi(a_1)\varphi(a_2).$$

Let $t \geqslant 2$ be given, and suppose that equation (3.13) holds for every word N of length $t-1$. As in the case $t = 2$, write $a_i = \varphi(a_i) \cdot 1_{\mathcal{A}} + a_i^\circ$ for $i \leqslant t$. Then, one has

$$\begin{aligned}\varphi(a_1 \cdots a_t) =& \varphi\Big(\big(\varphi(a_1) \cdot 1_{\mathcal{A}} + a_1^\circ\big) \cdots \big(\varphi(a_t) \cdot 1_{\mathcal{A}} + a_t^\circ\big)\Big)\\ =& \varphi(a_1^\circ \cdots a_t^\circ) + \sum_{k=0}^{t-1} \sum_{1 \leqslant j(1), \ldots, j(k) \leqslant t} \varphi(a_{j(1)}^\circ \cdots a_{j(k)}^\circ)\varphi(a_{j'(1)}) \cdots \varphi(a_{j'(t-k)})\\ =& \sum_{k=0}^{t-1} \sum_{1 \leqslant j(1), \ldots, j(k) \leqslant t} \varphi(a_{j(1)}^\circ \cdots a_{j(k)}^\circ)\varphi(a_{j'(1)}) \cdots \varphi(a_{j'(t-k)}),\end{aligned} \tag{3.14}$$

where, for each $1 \leqslant j(1), \ldots, j(k) \leqslant t$, we define $\big\{j'(1), \ldots, j'(t-k)\big\} = \big\{1, \ldots, t\big\} \backslash \big\{j(1), \ldots, j(k)\big\}$. Given that every expression of the form $\varphi(a_{j(1)}^\circ \cdots a_{j(k)}^\circ)$ in (3.14) can be written as a linear combination of expressions of the form $\varphi(N)$, where N is a word in the $a_1, \ldots, a_t$ of length at most $t-1$, it follows that (3.13) holds in the case $M = a_1 \cdots a_t$ with $F\big(M; \varphi(a) : a \in \mathcal{A}_i,\ i \in I\big)$ equal to (3.14). □

Equation (3.14) provides, in principle, a way of recursively computing mixed moments in free variables from the distributions of the individual variables. However, if a mixed moment $\varphi(a_1 \cdots a_t)$ is such that t is very large, the computations involved can be very tedious. The following examples shows such computations for simple mixed moments.

Example 3.21. Let $(\mathcal{A}, \varphi)$ be a noncommutative probability space and let I be an indexing set. For every $i \in I$, let C_i be a collection of variables in $\mathcal{A}$, let $\mathcal{A}_i \coloneqq \mathrm{alg}(1_{\mathcal{A}}, C_i)$ be the unital algebra generated by C_i, and suppose that the C_i are free from each other. **(1).** Let $a_1 \in \mathcal{A}_{i(1)}$ and $a_2 \in \mathcal{A}_{i(2)}$, where $i(1) \neq i(2)$. As shown in the proof of Proposition 3.20, one has $\varphi(a_1a_2) = \varphi(a_1)\varphi(a_2)$.

(2). Let $a_1, a_1' \in \mathcal{A}_{i(1)}$ and $a_2 \in \mathcal{A}_{i(2)}$, where $i(1) \neq i(2)$. Then, given that

$$\varphi\big(a_1^\circ (a_1')^\circ\big) = \varphi(a_1a_1') - \varphi(a_1)\varphi(a_1')$$

and $\varphi(a_1^\circ)$, $\varphi(a_2^\circ)$, $\varphi\big((a_1')^\circ\big)$, $\varphi(a_1^\circ a_2^\circ)$, $\varphi\big(a_2^\circ(a_1')^\circ\big)$ and $\varphi\big(a_1^\circ a_2^\circ(a_1')^\circ\big)$ all vanish, one has

$$\begin{aligned}\varphi(a_1a_2a_1') =&\varphi\Big(\big(\varphi(a_1)\cdot 1_{\mathcal{A}}+a_1^\circ\big)\big(\varphi(a_2)\cdot 1_{\mathcal{A}}+a_2^\circ\big)\big(\varphi(a_1')\cdot 1_{\mathcal{A}}+(a_1')^\circ\big)\Big)\\ =&\varphi(a_1)\varphi(a_2)\varphi(a_1')+\varphi(a_2)\varphi\big(a_1^\circ(a_1')^\circ\big)\\ =&\varphi(a_1)\varphi(a_2)\varphi(a_1')+\varphi(a_2)\big(\varphi(a_1a_1')-\varphi(a_1)\varphi(a_1')\big)\\ =&\varphi(a_1a_1')\varphi(a_2).\end{aligned}$$

(3). Let $a_1, a_1' \in \mathcal{A}_{i(1)}$ and $a_2, a_2' \in \mathcal{A}_{i(2)}$, where $i(1) \neq i(2)$. Then, given that

$$\varphi\big(a_k^\circ(a_k')^\circ\big) = \varphi(a_ka_k') - \varphi(a_k)\varphi(a_k'), \qquad\qquad k = 1, 2,$$

and that $\varphi(a_1^\circ)$, $\varphi(a_2^\circ)$, $\varphi\big((a_1')^\circ\big)$, $\varphi\big((a_2')^\circ\big)$, $\varphi(a_1^\circ a_2^\circ)$, $\varphi\big(a_1^\circ(a_2')^\circ\big)$, $\varphi\big(a_2^\circ(a_1')^\circ\big)$, $\varphi\big((a_1')^\circ(a_2')^\circ\big)$, $\varphi\big(a_1^\circ a_2^\circ(a_1')^\circ\big)$, $\varphi\big(a_1^\circ a_2^\circ(a_2')^\circ\big)$, $\varphi\big(a_1^\circ(a_1')^\circ(a_2')^\circ\big)$, $\varphi\big(a_2^\circ(a_1')^\circ(a_2')^\circ\big)$ and $\varphi\big(a_1^\circ a_2^\circ(a_1')^\circ(a_2')^\circ\big)$ all vanish, one has

$$\begin{aligned}&\varphi(a_1a_2a_1'a_2')\\ =&\varphi\Big(\big(\varphi(a_1)\cdot 1_{\mathcal{A}}+a_1^\circ\big)\big(\varphi(a_2)\cdot 1_{\mathcal{A}}+a_2^\circ\big)\big(\varphi(a_1')\cdot 1_{\mathcal{A}}+(a_1')^\circ\big)\big(\varphi(a_2')\cdot 1_{\mathcal{A}}+(a_2')^\circ\big)\Big)\\ =&\varphi(a_1)\varphi(a_2)\varphi(a_1')\varphi(a_2')+\varphi(a_1)\varphi(a_1')\varphi\big(a_2^\circ(a_2')^\circ\big)+\varphi(a_2)\varphi(a_2')\varphi\big(a_1^\circ(a_1')^\circ\big)\\ =&\varphi(a_1)\varphi(a_2)\varphi(a_1')\varphi(a_2')+\varphi(a_1)\varphi(a_1')\big(\varphi(a_2a_2')-\varphi(a_2)\varphi(a_2')\big)\\ &\quad+\varphi(a_2)\varphi(a_2')\big(\varphi(a_1a_1')-\varphi(a_1)\varphi(a_1')\big)\\ =&\varphi(a_1)\varphi(a_1')\varphi(a_2a_2')+\varphi(a_2)\varphi(a_2')\varphi(a_1a_1')-\varphi(a_1)\varphi(a_2)\varphi(a_1')\varphi(a_2'),\end{aligned}$$

hence

$$\varphi(a_1a_2a_1'a_2') = \varphi(a_1)\varphi(a_1')\varphi(a_2a_2') + \varphi(a_2)\varphi(a_2')\varphi(a_1a_1') - \varphi(a_1)\varphi(a_2)\varphi(a_1')\varphi(a_2') \tag{3.15}$$

If a mixed moment has a very specific structure, it is sometimes possible to obtain a convenient formulation for it in terms of the individual moments. Consider for example the following proposition, which will be useful in later proofs.

Proposition 3.22 ([**NS06**] Lemma 5.18)**.** *Let $(\mathcal{A}, \varphi)$ be a noncommutative probability space and let I be an indexing set. For every $i \in I$, let C_i be a collection of variables in $\mathcal{A}$, let $\mathcal{A}_i \coloneqq \operatorname{alg}(1_{\mathcal{A}}, C_i)$ be the unital algebra generated by C_i, and suppose that the C_i are free from each other. Let $a_1, \ldots, a_s, b_1, \ldots, b_t \in \mathcal{A}$ be such that $a_1 \in \mathcal{A}_{i(1)}, \ldots, a_s \in \mathcal{A}_{i(s)}$ and $b_1 \in \mathcal{A}_{j(1)}, \ldots, b_t \in \mathcal{A}_{j(t)}$, where $i(1) \neq i(2) \neq \cdots \neq i(s)$ and $j(1) \neq j(2) \neq \cdots \neq j(t)$. Then, one has*

$$\varphi(a_1^\circ \cdots a_s^\circ b_t^\circ \cdots b_1^\circ) = \begin{cases} \varphi(a_1^\circ b_1^\circ)\cdots\varphi(a_t^\circ b_t^\circ) & \text{if } s = t \text{ and } i(l) = j(l) \text{ for } l \leqslant t; \text{ and} \\ 0 & \text{otherwise.}\end{cases}$$

Proof. If $i(s) \neq j(t)$, then it immediately follows from the definition of free independence that $\varphi(a_1^\circ \cdots a_s^\circ b_t^\circ \cdots b_1^\circ) = 0$. Suppose that $i(s) = j(t)$, and write $a_s^\circ b_t^\circ = \varphi(a_s^\circ b_t^\circ)\cdot 1_{\mathcal{A}_{i(s)}} + (a_s^\circ b_t^\circ)^\circ$.

Then, one has

$$\begin{aligned}&\varphi(a_1^\circ\cdots a_s^\circ b_t^\circ\cdots b_1^\circ)\\ =&\varphi\Big(a_1^\circ\cdots a_{s-1}^\circ\big(\varphi(a_s^\circ b_t^\circ)\cdot 1_{\mathcal{A}_{i(s)}}+(a_s^\circ b_t^\circ)^\circ\big)b_{t-1}^\circ\cdots b_1^\circ\Big)\\ =&\varphi(a_s^\circ b_t^\circ)\varphi\Big(a_1^\circ\cdots a_{s-1}^\circ b_{t-1}^\circ\cdots b_1^\circ\Big)+\varphi\Big(a_1^\circ\cdots a_{s-1}^\circ(a_s^\circ b_t^\circ)^\circ b_{t-1}^\circ\cdots b_1^\circ\Big).\end{aligned}$$

Given that $i(1)\neq i(2)\neq\cdots\neq i(s)=j(t)\neq j(t-1)\neq\cdots j(1)$, the mixed moment $\varphi\Big(a_1^\circ\cdots a_{s-1}^\circ(a_s^\circ b_t^\circ)^\circ b_{t-1}^\circ\cdots b_1^\circ\Big)$ vanishes, and thus

$$\varphi(a_1^\circ\cdots a_s^\circ b_t^\circ\cdots b_1^\circ)=\varphi(a_s^\circ b_t^\circ)\varphi\Big(a_1^\circ\cdots a_{s-1}^\circ b_{t-1}^\circ\cdots b_1^\circ\Big).$$

Applying the same reasoning as above to $\varphi\Big(a_1^\circ\cdots a_{s-1}^\circ b_{t-1}^\circ\cdots b_1^\circ\Big)$ yields that

$$\varphi\Big(a_1^\circ\cdots a_{s-1}^\circ b_{t-1}^\circ\cdots b_1^\circ\Big)=0$$

if $i(s-1)\neq j(t-1)$, and

$$\varphi\Big(a_1^\circ\cdots a_{s-1}^\circ b_{t-1}^\circ\cdots b_1^\circ\Big)=\varphi(a_{s-1}^\circ b_{t-1}^\circ)\varphi\Big(a_1^\circ\cdots a_{s-2}^\circ b_{t-2}^\circ\cdots b_1^\circ\Big)$$

otherwise. Thus, one may proceed by reverse induction and obtain that

$$\varphi(a_1^\circ\cdots a_s^\circ b_t^\circ\cdots b_1^\circ)=\varphi(a_1^\circ b_1^\circ)\cdots\varphi(a_t^\circ b_t^\circ)$$

if $s=t$ and $i(1)=j(1), i(2)=j(2),\ldots,i(t)=j(t)$, and

$$\varphi(a_1^\circ\cdots a_s^\circ b_t^\circ\cdots b_1^\circ)=\begin{cases}C_1\cdot\varphi(a_1^\circ\cdots a_{s-t}^\circ)=0 & \text{if } s>t\text{; and}\\ C_2\cdot\varphi(b_{t-s}^\circ\cdots b_1^\circ)=0 & \text{if } s<t\end{cases}$$

(where $C_1, C_2\in\mathbb{C}$), as desired. □

3.3.1. Abstract Model for Free Independent Variables. In order for the concept of free independence to be as powerful as tensor independence, it should also be possible to generate collections of free random variables with arbitrary distributions in order to be able to make statements such as "Let $(a_i : i\in I)$ be a collection of free random variables with respective distributions μ_{a_i}" without running into problems of existence.

As the appellation *free independence* might suggest, there exists such an abstract construction based on the free product of unital algebras. Let $\big((\mathcal{A}_i,\varphi_i) : i\in I\big)$ be a collection of noncommutative probability spaces. Define the algebra $\mathcal{A}$ as the free product[10]

$$*_{i\in I}\mathcal{A}_i=\mathbb{C}1_{\mathcal{A}}\oplus\left(\bigoplus_{t\geqslant 1}\bigoplus_{\substack{i(1),\ldots,i(t)\in I\\ i(1)\neq i(2)\neq\cdots\neq i(t)}}\mathcal{A}_{i(1)}^\circ\otimes\cdots\otimes\mathcal{A}_{i(t)}^\circ\right)$$

and define the functional $\varphi=*_{i\in I}\varphi_i$ on $\mathcal{A}$ as the linear extension of the map such that $\varphi(1_{\mathcal{A}})=1$ and $\varphi(a_1\otimes\cdots\otimes a_t)=0$ for every $t\in\mathbb{N}$ and $a_1\in\mathcal{A}_{i(1)}^\circ,\ldots,a_t\in\mathcal{A}_{i(t)}^\circ$ such that $i(1)\neq i(2)\neq$

[10] See Section A.2.

$\cdots \neq i(t)$. For a fixed $i \in I$ and $a \in \mathcal{A}_i$, the element a can naturally be embedded in $\mathcal{A}$ as

$$\begin{array}{ccc} \mathcal{A}_i & \to & \mathcal{A} \\ a & \hookrightarrow & \varphi_i(a) \cdot 1_{\mathcal{A}} + a^\circ. \end{array}$$

One easily checks that distributions are preserved by the above embedding, and that for any family of collections $(C_i \subset \mathcal{A}_i : i \in I)$, the embeddings of C_i into $\mathcal{A}$ are free from each other.

Example 3.23 ([**NS06**] Lecture 6, pages 83 and 84)**.** Let $(\mathcal{A}_1, \varphi)$ and $\mathcal{A}_2 = (\mathcal{A}_2, \varphi)$ be two noncommutative probability spaces, and let $a_1, b_1 \in \mathcal{A}_1^\circ$ and $a_2, b_2 \in \mathcal{A}_2^\circ$ be such that $a_i b_i \notin \mathcal{A}_i^\circ$ $(i = 1, 2)$. As it is done in Example A.6, we see that

$$\begin{aligned} &(a_1 \otimes a_2)(b_2 \otimes b_1) \\ =&\varphi_2(a_2 b_2) \cdot (a_1 \otimes b_1) + (a_1 \otimes (a_2 b_2)^\circ \otimes b_1) \\ =&\varphi_2(a_2 b_2)\varphi_1(a_1 b_1) \cdot 1_{\mathcal{A}} + \varphi_2(a_2 b_2) \cdot (a_1 b_1)^\circ + (a_1 \otimes (a_2 b_2)^\circ \otimes b_1), \end{aligned}$$

and thus,

$$(\varphi_1 * \varphi_2)\big((a_1 \otimes a_2)(b_2 \otimes b_1)\big) = \varphi_2(a_2 b_2)\varphi_1(a_1 b_1).$$

Notice that this is consistent with Example 3.21 (2).

Definition 3.24. Let $\big((\mathcal{A}_i, *, \varphi_i) : i \in I\big)$ be a collection of arbitrary $*$-probability spaces, let $\mathcal{A} = *_{i \in I} \mathcal{A}_i$, and define the operator $\cdot^* : \mathcal{A} \to \mathcal{A}$ as follows: $1_{\mathcal{A}}^* = 1_{\mathcal{A}}$,

$$(a_1 \otimes \cdots \otimes a_{t-1} \otimes a_t)^* = a_t^* \otimes a_{t-1}^* \otimes \cdots \otimes a_1^*$$

for every $t \in \mathbb{N}$ and $a_1 \in \mathcal{A}_{i(1)}^\circ, \ldots, a_t \in \mathcal{A}_{i(t)}^\circ$ such that $i(1) \neq i(2) \neq \cdots \neq i(t)$, and then extend $*$ antilinearly to all of $\mathcal{A}$ (that is, given $a, b \in \mathcal{A}$ and $\lambda, \mu \in \mathbb{C}$, we ensure that $(\lambda \cdot a + \mu \cdot b)^* = \overline{\lambda} \cdot a^* + \overline{\mu} \cdot b^*$). Then, one easily checks that $(\mathcal{A}, *)$ is a $*$-algebra.

Example 3.25. Let $(G_i : i \in I)$ be a collection of groups (with respective neutral elements e_i), and let $G = *_{i \in I} G_i$ be the free product of the G_i.[11] Then, for every $j \in I$ and $g \in G_j$, the element g can be embedded in G as

$$\begin{array}{ccc} G_j & \to & G \\ g & \hookrightarrow & g, \end{array}$$

One easily checks that the order of every element is preserved by this embedding, and that the embeddings of the G_i into G are free. In fact, $\big(\mathbb{C}(*_{i \in I} G_i), *, \tau_e\big)$ (where e denotes the neutral element in $*_i G_i$) is isomorphic as a $*$-probability space to $(*_{i \in I} \mathbb{C}G_i, *, *_{i \in I} \tau_{e_i})$.

The following proposition, whose proof is a matter of straightforward computations, further establishes the link between freeness in groups and freeness in $*$-probability spaces.

Proposition 3.26 ([**NS06**] Proposition 5.11)**.** *Let $(G_i : i \in I)$ be a collection of subgroups of a group G. Then, the groups G_i are free if and only if the $*$-algebras $\mathbb{C}G_i$ are free from each other with respect to the canonical trace τ_e on $\mathbb{C}G$.*

As in the case of tensor independence, we finish this subsection by demonstrating that the free product of $*$-probability spaces preserves positivity, and thus is also a $*$-probability space.

[11] See section A.1.

Theorem 3.27 ([**NS06**] Theorem 6.13)**.** *Let $\big((\mathcal{A}_i, *, \varphi_i) : i \in I\big)$ be a collection of $*$-probability spaces, let $\mathcal{A} = *_{i\in I}\mathcal{A}_i$, let $\varphi = *_{i\in I}\varphi_i$, and let $*$ be defined as in Definition 3.24. Then, $(\mathcal{A}, *, \varphi)$ is a $*$-probability space.*

Proof. We show that φ is positive. Let $a \in \mathcal{A}$ be arbitrary. Then, a can be written as a sum of the form

$$a = \varphi(a) \cdot 1_{\mathcal{A}} + \sum_{t=1}^{T} \sum_{\substack{i(1),\ldots,i(t)\in I \\ i(1)\neq i(2)\neq\cdots\neq i(t)}} A_{i(1),\ldots,i(t)}$$

where $T \in \mathbb{N}$, and for each $t \leqslant T$ and $i(1) \neq i(2) \neq \cdots \neq i(t)$, $A_{i(1),\ldots,i(t)}$ is a linear combination of terms of the form $a_{i(1)} \otimes \cdots \otimes a_{i(t)}$, where $a_{i(1)} \in \mathcal{A}^{\circ}_{i(1)}, \ldots, a_{i(t)} \in \mathcal{A}^{\circ}_{i(t)}$. According to Proposition 3.22, whenever the indices $i(1), \ldots, i(s), j(1), \ldots, j(t) \in I$ (where $i(1) \neq i(2) \neq \cdots \neq i(s)$, $j(1) \neq j(2) \neq \cdots \neq j(t)$) are such that $s \neq t$, or $s = t$ and $i(k) \neq j(k)$ for some $k \leqslant t$, then

$$\varphi\big((a_{i(1)} \otimes \cdots \otimes a_{i(s)})(a_{j(1)} \otimes \cdots \otimes a_{j(t)})^*\big) = \varphi\big((a_{i(1)} \otimes \cdots \otimes a_{i(s)})(a^*_{j(t)} \otimes \cdots \otimes a^*_{j(1)})\big) = 0.$$

Therefore, since $\varphi(A_{i(1),\ldots,i(t)}) = \varphi(A^*_{i(1),\ldots,i(t)}) = 0$, it follows that

$$\varphi(aa^*) = \big|\varphi(a)\big|^2 + \sum_{t=1}^{T} \sum_{\substack{i(1),\ldots,i(t)\in I \\ i(1)\neq i(2)\neq\cdots\neq i(t)}} \varphi(A_{i(1),\ldots,i(t)}A^*_{i(1),\ldots,i(t)}).$$

Thus, given that $\big|\varphi(a)\big|^2 \geqslant 0$, it suffices to prove that φ is positive when restricted to every summand of the form $A_{i(1),\ldots,i(t)}$.

Let $i(1), \ldots, i(t) \in I$ be a fixed collection of distinct indices, and let

$$A_{i(1),\ldots,i(t)} = \sum_{l=1}^{L} (a_1^{(l)} \otimes \cdots \otimes a_t^{(l)})_I,$$

where, for each $l \leqslant L$, one has $a_1^{(l)} \in \mathcal{A}^{\circ}_{i(1)}, \ldots, a_t^{(l)} \in \mathcal{A}^{\circ}_{i(t)}$. According to Proposition 3.22,

$$\begin{aligned}\varphi(A_{i(1),\ldots,i(t)}A^*_{i(1),\ldots,i(t)}) &= \sum_{l,s=1}^{L} \varphi\big((a_1^{(l)} \otimes \cdots \otimes a_t^{(l)})_I((a_1^{(s)})^* \otimes \cdots \otimes (a_t^{(s)})^*)_I\big) \\ &= \sum_{l,s=1}^{L} \varphi_{i(1)}\big(a_1^{(l)}(a_1^{(s)})^*\big) \cdots \varphi_{i(t)}\big(a_t^{(l)}(a_t^{(s)})^*\big).\end{aligned}$$

Then, using the same arguments as in part (1) Theorem 3.13, we conclude that φ is positive. □

Remark 3.28. As in the case of tensor independence, it is worth mentioning that the free product of noncommutative probability spaces equipped with traces is also equipped with a trace, and that the free product of $*$-probability spaces equipped with faithful functionals is also equipped with a faithful functional (see for instance [**NS06**] Propositions 6.8 and 6.14).

3.3.2. Free Independence as a Fundamentally Noncommutative Notion. Much like tensor independence could not be naively extended to noncommutative random variables, it can be argued that free independence is a fundamentally noncommutative notion.

Let $(\mathcal{A}, \varphi)$ be a noncommutative probability space, and let $a_1, a_2 \in \mathcal{A}$ be such that a_1 and a_2 commute and are free from each other. Then, it follows that $\varphi(a_1^2 a_2^2) = \varphi(a_1 a_2 a_1 a_2)$. According to Example 3.21, It then follows that

$$\varphi(a_1^2)\varphi(a_2^2) = \varphi(a_1^2 a_2^2) = \varphi(a_1 a_2 a_1 a_2) = \varphi(a_1^2)\varphi(a_2)^2 + \varphi(a_1)^2\varphi(a_2^2) - \varphi(a_1)^2\varphi(a_2)^2,$$

and thus, $0 = \big(\varphi(a_1^2) - \varphi(a_1)^2\big)\big(\varphi(a_2^2) - \varphi(a_2)^2\big)$, which imposes strong conditions on the distribution of a_1 or a_2. Similarly, one can prove that if a_1 and a_2 are in a $*$-probability space, free from each other, and commute, then $\varphi(a_1 a_1^* a_2 a_2^*) = \varphi(a_1 a_2 a_1^* a_2^*)$ implies that $0 = \mathbf{Var}[a_1]\mathbf{Var}[a_2]$, and thus, if φ is faithful, one of a_1 or a_2 must be deterministic.

3.4. Uniqueness of Tensor and Free Independence

Now that we have seen a concept of independence for noncommutative random variables that can be considered a true analog of the notion of classical independence for commutative random variables, it is natural to wonder if there exists other such concepts of independence. (In other words, how *special* are tensor and free independence?)

In order to answer this question, it is first necessary to rigorously define what is meant by a *concept of independence*, and then decide what properties this concept should satisfy.

As we will see in this section, given a very general definition of a concept of independence and under *mild* assumptions (i.e., assumptions that can easily be argued to be desirable of a concept of independence), one can prove that tensor and free independence are indeed special, in that they are the only possible concepts of independence.

3.4.1. Independence Rules. In the previous sections, we have seen two equivalent ways in which tensor and free independence were defined, namely,

(1) tensor and free independence means that no new information emerges from the product of independent variables (see equations (3.5) and (3.12)); and

(2) tensor and free independence are both rules for decomposing mixed moments in independent variables in terms of the individual variables (see equation (3.3) and Proposition 3.20).

We use the second item above to generalize the definition of a concept of independence:

Notation 3.29. Let $(\mathcal{A}, \varphi)$ be a noncommutative probability space. Let $n \in \mathbb{N}$ be fixed and let $\pi \in P(n)$ be a partition of $\{1, \ldots, n\}$ (see Section D.1 for the definitions, notations, and elementary proofs concerning partitions). For every block $B = \{i(1), \ldots, i(l)\} \in \pi$ (where $1 \leqslant i(1) < \ldots < i(l) \leqslant n$) of π, let $\varphi_B[a_1, \ldots, a_n] \coloneqq \varphi(a_{i(1)} \cdots a_{i(l)})$, and define

$$\varphi_\pi[a_1, \ldots, a_n] \coloneqq \prod_{B \in \pi} \varphi_B[a_1, \ldots, a_n].$$

Example 3.30. Consider the following noncrossing partitions of $\{1,2,3,4\}$:

$$\begin{aligned}\pi_1 =&\big\{\{1\},\{2,4\},\{3\}\big\}\\ \pi_2 =&\big\{\{1,3\},\{2\},\{4\}\big\}\\ \pi_3 =&0_4 = \big\{\{1\},\{2\},\{3\},\{4\}\big\}\end{aligned}$$

Then, given four variables a_1, a_2, a_3, a_4 in a noncommutative probability space $(\mathcal{A}, \varphi)$, one has

$$\varphi_{\pi_1}[a_1,a_2,a_3,a_4] = \varphi(a_1)\varphi(a_3)\varphi(a_2a_4),$$

$$\varphi_{\pi_2}[a_1,a_2,a_3,a_4] = \varphi(a_1a_3)\varphi(a_2)\varphi(a_4),$$

and

$$\varphi_{\pi_3}[a_1,a_2,a_3,a_4] = \varphi(a_1)\varphi(a_2)\varphi(a_3)\varphi(a_4).$$

In fact, as shown in Example 3.21 (3) (see equation (3.15)), if $a_1, a_3 \in C$ and $a_2, a_4 \in D$, where C and D are two collections that are free from each other, then

$$\varphi(a_1a_2a_3a_4) = \varphi_{\pi_1}[a_1,a_2,a_3,a_4] + \varphi_{\pi_2}[a_1,a_2,a_3,a_4] - \varphi_{\pi_3}[a_1,a_2,a_3,a_4].$$

Definition 3.31. Let $\big((\mathcal{A}_i,\varphi_i) : i \in I\big)$ be a collection of noncommutative probability spaces, and let $\mathcal{A} = *_{i\in I}\mathcal{A}_i$ be the free product of the $\mathcal{A}_i$. For each $n \in \mathbb{N}$ and partition $\pi \in P(n)$ let $\mathcal{A}_\pi \subset \mathcal{A}$ denote the set of products of the form $a_1 \cdots a_n$, where $a_1 \in \mathcal{A}_{i(1)} \setminus \{1_{\mathcal{A}_{i(1)}}\}, \ldots, a_n \in \mathcal{A}_{i(n)} \setminus \{1_{\mathcal{A}_{i(n)}}\}$ are such that $i(1) \neq i(2) \neq \cdots \neq i(n)$ and $i(k) = i(l)$ if and only if k and l are in the same block in π.

Remark 3.32. Several of the sets of the form $\mathcal{A}_\pi$ are empty. For example, if $\pi = \big\{\{1,2\},\{3\}\big\}$, then $\mathcal{A}_\pi = \varnothing$. Furthermore, one easily checks that, given $\pi \in P(m)$ and $\sigma \in P(n)$ (with m and n possibly equal), one has $\mathcal{A}_\pi \cap \mathcal{A}_\sigma = \varnothing$, and

$$\bigcup_{n\in\mathbb{N}} \bigcup_{\pi\in P(n)} \mathcal{A}_\pi = \mathcal{A}.$$

Example 3.33. Let $\pi = \big\{\{1,3\},\{2,4,6\},\{5\}\big\} \in P(6)$, and let $b_1 \in \mathcal{A}_{i(1)} \setminus \{1_{\mathcal{A}_{i(1)}}\}$, $b_2 \in \mathcal{A}_{i(2)} \setminus \{1_{\mathcal{A}_{i(2)}}\}$, and $b_3 \in \mathcal{A}_{i(3)} \setminus \{1_{\mathcal{A}_{i(3)}}\}$ be such that $i(1), i(2)$, and $i(3)$ are distinct. Define the word $a_1 \cdots a_6$ as follows: $a_1 = a_3 = b_1$, $a_2 = a_4 = a_6 = b_2$, and $a_5 = b_3$. Then, $a_1 \cdots a_6 \in \mathcal{A}_\pi$, but $a_1 \cdots a_6 \notin \mathcal{A}_\sigma$ for any other partition σ.

Definition 3.34 ([**Sp97**])**.** An **independence rule** consists of a sequence of functions

$$\mathscr{I} = \big\{\mathbf{t}_n : P(n)^{(2)} \to \mathbb{C} : n \in \mathbb{N}\big\}$$

(recall that $P(n)^{(2)} = \big\{(\sigma,\pi) \in P(n)\times P(n) : \sigma \leqslant \pi\big\}$, where $\leqslant$ is the reversed refinement order, see Notation D.8) such that for any collection $\big((\mathcal{A}_i,\varphi_i) : i \in I\big)$ noncommutative probability spaces, one can define a linear functional $*_{i\in I}^{(\mathscr{I})}\varphi_i : \mathcal{A} \to \mathbb{C}$ (where $\mathcal{A} = *_{i\in I}\mathcal{A}_i$ denotes the free product of the $\mathcal{A}_i$) satisfying the following conditions:

(1) for every $i \in I$, $*_{i\in I}^{(\mathscr{I})}\varphi_i|_{\mathcal{A}_i} = \varphi_i$, that is, for every $i \in I$ and $a \in \mathcal{A}_i$, one has $*_{i\in I}^{(\mathscr{I})}\varphi_i(a) = \varphi_i(a)$; and

(2) for every $n \in \mathbb{N}$ and $a_1 \cdots a_n \in \mathcal{A}_\pi$, one has

$$*_{i\in I}^{(\mathscr{I})}\varphi_i(a_1 \cdots a_n) = \sum_{\sigma\in P(n):\ \sigma\leqslant\pi} \mathbf{t}_n(\sigma,\pi)\varphi_\sigma[a_1,\ldots,a_n]. \tag{3.16}$$

Remark 3.35. Notice that every expression of the form $\varphi_\sigma[a_1,\ldots,a_n]$ in (3.16) can be computed using the functionals φ_i. Indeed, by requiring that $a_1\cdots a_n \in \mathcal{A}_\pi$, it follows that for every $\sigma \leqslant \pi$ and block $B = \{j(1),\ldots,j(s)\} \in \sigma$, the elements $a_{j(1)},\ldots,a_{j(s)}$ are contained in the same algebra $\mathcal{A}_i$, and thus $\varphi_B[a_1,\ldots,a_n] = \varphi_i(a_{j(1)}\cdots a_{j(s)})$.

Example 3.36 (**[Sp97]**)**.** Tensor independence is associated to the following independence rule: Let $\mathscr{T} = \{\mathbf{t}_n : P(n)^{(2)} \to \mathbb{C} : n \in \mathbb{N}\}$ be such that for every $n \in \mathbb{N}$ and $\sigma \leqslant \pi \in P(n)$, one has

$$\mathbf{t}_n(\sigma,\pi) = \begin{cases} 1 & \text{if } \sigma = \pi; \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

For instance, if $(\mathcal{A}_1,\varphi_1)$, $(\mathcal{A}_2,\varphi_2)$, and $(\mathcal{A}_3,\varphi_3)$ are tensor independent and $a_1 \in \mathcal{A}_1$, $a_2 \in \mathcal{A}_2$, and $a_3 \in \mathcal{A}_3$, one has (where $\pi = \{\{1,3\},\{2,4,6\},\{5,7\}\}$)

$$\begin{aligned}(\varphi_1\otimes\varphi_2\otimes\varphi_3)(a_1a_2a_1a_2a_3a_2a_3) &= *_{i\in I}^{(\mathscr{T})}\varphi_i(a_1a_2a_1a_2a_3a_2a_3)\\ &= 0 + \mathbf{t}_7(\pi,\pi)\varphi_\pi[a_1,a_2,a_1,a_2,a_3,a_2,a_3]\\ &= 0 + \varphi_\pi[a_1,a_2,a_1,a_2,a_3,a_2,a_3]\\ &= \varphi_1(a_1^2)\varphi_2(a_2^2)\varphi_3(a_3^3).\end{aligned}$$

Example 3.37. Let $\mathscr{F} = \{\mathbf{t}_n : P(n)^{(2)} \to \mathbb{C} : n \in \mathbb{N}\}$ be the independence rule associated to free independence, and let

$$\begin{array}{rclcrcl} \pi &=& \{\{1,3\},\{2,4\}\}, & & \pi_2 &=& \{\{1,3\},\{2\},\{4\}\},\\ \pi_1 &=& \{\{1\},\{2,4\},\{3\}\}, & & \pi_3 &=& 0_4 = \{\{1\},\{2\},\{3\},\{4\}\}.\end{array}$$

Then, it follows from Examples 3.21 and 3.30 that

$$\mathbf{t}_4(\sigma,\pi) = \begin{cases} 1 & \text{if } \sigma = \pi_1 \text{ or } \pi_2; \\ -1 & \text{if } \sigma = \pi_3; \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

3.4.2. Associativity of Tensor and Free Independence. Given the current definition of a concept of independence, it seems that any collection $\mathscr{I} = \{\mathbf{t}_n : P(n)^{(2)} \to \mathbb{C} : n \in \mathbb{N}\}$ of functions from which a functional $*_{i\in I}^{(\mathscr{I})}\varphi_i$ can be defined according to (3.16) is a concept of independence. However, there are properties that a concept of independence should have to be satisfying from an intuitive point of view. One such property is the **associativity of independence**,[12] which can briefly be stated as follows (using the notation of **[NS06]** Remark

[12] See Theorem B.4 in Section B.1.3 for the statement that classical independence for $\mathbb{C}$-valued random variables is associative.

5.20): Given three collections C_1, C_2, and C_3 of random variables

$$\begin{aligned} & C_1, C_2, C_3 \text{ are independent of each other} \\ \iff & C_1 \text{ is independent of } C_2 \cup C_3 \text{ and } C_2 \text{ is independent of } C_3 \\ \iff & C_2 \text{ is independent of } C_1 \cup C_3 \text{ and } C_1 \text{ is independent of } C_3 \\ \iff & C_3 \text{ is independent of } C_1 \cup C_2 \text{ and } C_1 \text{ is independent of } C_2. \end{aligned}$$

The general statement of the property of associativity and the proof that it holds for tensor and free independence is contained in the following Proposition:

Proposition 3.38. *Let $(\mathcal{A}, \varphi)$ be a noncommutative probability space and let I be an indexing set. For every $i \in I$, let C_i be a collection of variables in $\mathcal{A}$, and let $\mathcal{A}_i := \mathrm{alg}(1_{\mathcal{A}}, C_i)$ be the unital algebra generated by C_i. Let $(I_j : j \in J)$ be a partition of the indexing set I, and for each $j \in J$, let $\mathcal{B}_j = \mathrm{alg}(\mathcal{A}_i : i \in I_j)$ be the algebra generated by the $\mathcal{A}_i$ such that $i \in I_j$.*

(1) If the $\mathcal{A}_i$ are tensor independent of each other, then so are are the $\mathcal{B}_j$.

(2) Suppose that the algebras $(\mathcal{B}_j : j \in J)$ are tensor independent of each other, and that for every $j \in J$, the algebras in $(\mathcal{A}_i : i \in I_j)$ are tensor independent of each other. Then, the algebras in $(\mathcal{A}_i : i \in I_j)$ are tensor independent of each other.

(3) If the $\mathcal{A}_i$ are free from each other, then so are the $\mathcal{B}_j$.

(4) Suppose that the algebras $(\mathcal{B}_j : j \in J)$ are free from each other, and that for every $j \in J$, the algebras in $(\mathcal{A}_i : i \in I_j)$ are free from each other. Then, the algebras in $(\mathcal{A}_i : i \in I_j)$ are free from each other.

Proof. (1). Given that φ is linear, for a fixed $j \in J$ and an arbitrary element $b \in \mathcal{B}_j$, it is always possible to write

$$b = \alpha_0 \cdot 1_{\mathcal{A}} + \sum_{k=1}^{m} \alpha_k \cdot (a_{k,1}^{\circ} \cdots a_{k,n_k}^{\circ}),$$

where $m \in \mathbb{N}$, $\alpha_0, \ldots, \alpha_m \in \mathbb{C}$, and for every $k \leqslant m$, $n_k \in \mathbb{N}$ and $a_{k,1} \in \mathcal{A}_{i_k(1)}, \ldots, a_{k,n_k} \in \mathcal{A}_{i_k(n_k)}$, where $i_k(1), \ldots, i_k(n_k) \in I_j$ are distinct since the $\mathcal{A}_i$ commute. Therefore, since the $\mathcal{A}_i$ are tensor independent, it follows that

$$b^{\circ} = \sum_{k=1}^{m} \alpha_k \cdot (a_{k,1}^{\circ} \cdots a_{k,n_k}^{\circ})^{\circ} = \sum_{k=1}^{m} \alpha_k \cdot (a_{k,1}^{\circ} \cdots a_{k,n_k}^{\circ}).$$

Let $j(1), \ldots, j(t) \in J$ be distinct, and let $b_1 \in \mathcal{B}_{j(1)}, \ldots, b_t \in \mathcal{B}_{j(t)}$ be arbitrary. For each $l \leqslant t$, write

$$b_l^{\circ} = \sum_{k=1}^{m_l} \alpha_k^{(l)} \cdot \big((a_{k,1}^{(l)})^{\circ} \cdots (a_{k,n_k}^{(l)})^{\circ}\big).$$

Then, it is clear that the product $b_1^{\circ} \cdots b_t^{\circ}$ is a linear combination of terms of the form

$$\varphi\Big(\big(a_{k(1),1}^{(1)}\big)^{\circ} \cdots \big(a_{k(1),n_{k(1)}}^{(1)}\big)^{\circ} \cdots \big(a_{k(t),1}^{(t)}\big)^{\circ} \cdots \big(a_{k(t),n_{k(t)}}^{(t)}\big)^{\circ}\Big),$$

where $k(1) \leqslant m_1, \ldots, k(t) \leqslant m_t$. Given that $(I_j : j \in J)$ is a partition of I, no two elements among the collection

$$a_{k(1),1}^{(1)}, \ldots, a_{k(1),n_{k(1)}}^{(1)}, \ldots, a_{k(t),1}^{(t)}, \ldots, a_{k(t),n_{k(t)}}^{(t)}$$

are from the same algebra $\mathcal{A}_i$. Thus, since the $\mathcal{A}_i$ are tensor independent of each other, it follows that

$$\varphi\Big(\big(a^{(1)}_{k(1),1}\big)^\circ \cdots \big(a^{(1)}_{k(1),n_{k(1)}}\big)^\circ \cdots \big(a^{(t)}_{k(t),1}\big)^\circ \cdots \big(a^{(t)}_{k(t),n_{k(t)}}\big)^\circ\Big) = 0,$$

from which we conclude that $\varphi(b_1^\circ \cdots b_t^\circ) = 0$, as desired.

(2). Let $i(1), \ldots, i(s) \in I$ be distinct indices and let the variables $a_1 \in \mathcal{A}_{i(1)}, \ldots, a_s \in \mathcal{A}_{i(s)}$ be arbitrary. Let $I' = \big\{i(1), \ldots, i(s)\big\} \subset I$, and for every $j \in J$, let $I'_j = I' \cap I_j$. Then, there exists a finite collection of distinct indices $j(1), \ldots, j(t) \in J$ (with $t \leqslant s$) such that $I'_j \neq \varnothing$ if and only if $j \in \big\{j(1), \ldots, j(t)\big\}$. Thus, we can write $a_1^\circ \cdots a_s^\circ = b_1 \cdots b_t$, where, for each $l \leqslant t$, we have that

$$b_l = \prod_{i \in I'_{j(l)}} a_i^\circ.$$

Since the algebras in $(\mathcal{A}_i : i \in I_j)$ are tensor independent of each other for each j and the indices $i(1), \ldots, i(s)$ are distinct, it follows that

$$b_l^\circ = \left(\prod_{i \in I'_{j(l)}} a_i^\circ\right)^\circ = \prod_{i \in I'_{j(l)}} a_i^\circ = b_l$$

for each $l \leqslant t$, and since the $\mathcal{B}_j$ are independent of each other and the indices $j(1), \ldots, j(t)$ are distinct, it follows that

$$(b_1^\circ \cdots b_t^\circ)^\circ = b_1^\circ \cdots b_t^\circ = b_1 \cdots b_t,$$

and thus, since the $\mathcal{A}_i$ commute,

$$(a_1^\circ \cdots a_s^\circ)^\circ = (b_1 \cdots b_t)^\circ = (b_1^\circ \cdots b_t^\circ)^\circ = b_1 \cdots b_t = a_1^\circ \cdots a_s^\circ,$$

as desired.

(3). For a fixed $j \in J$ and an arbitrary element $b \in \mathcal{B}_j$, it is always possible to write

$$b = \alpha_0 \cdot 1_{\mathcal{A}} + \sum_{k=1}^{m} \alpha_k \cdot (a_{k,1}^\circ \cdots a_{k,n_k}^\circ),$$

where $m \in \mathbb{N}$, $\alpha_0, \ldots, \alpha_m \in \mathbb{C}$, and for every $k \leqslant m$, $n_k \in \mathbb{N}$ and $a_{k,1} \in \mathcal{A}_{i_k(1)}, \ldots, a_{k,n_k} \in \mathcal{A}_{i_k(n_k)}$, where $i_k(1), \ldots, i_k(n_k) \in I_j$ are such that $i_k(1) \neq i_k(2) \neq \cdots \neq i_k(n_k)$. Therefore, since the $\mathcal{A}_i$ are free from each other, it follows that

$$b^\circ = \sum_{k=1}^{m} \alpha_k \cdot (a_{k,1}^\circ \cdots a_{k,n_k}^\circ)^\circ = \sum_{k=1}^{m} \alpha_k \cdot (a_{k,1}^\circ \cdots a_{k,n_k}^\circ).$$

Let $j(1), \ldots, j(t) \in J$ be such that $j(1) \neq j(2) \neq \cdots \neq j(t)$, and let $b_1 \in \mathcal{B}_{j(1)}, \ldots, b_t \in \mathcal{B}_{j(t)}$ be arbitrary. For each $l \leqslant t$, write

$$b_l^\circ = \sum_{k=1}^{m_l} \alpha_k^{(l)} \cdot \big((a_{k,1}^{(l)})^\circ \cdots (a_{k,n_k}^{(l)})^\circ\big).$$

Then, it is clear that the product $b_1^\circ \cdots b_t^\circ$ is a linear combination of terms of the form

$$\varphi\Big(\big(a^{(1)}_{k(1),1}\big)^\circ \cdots \big(a^{(1)}_{k(1),n_{k(1)}}\big)^\circ \cdots \big(a^{(t)}_{k(t),1}\big)^\circ \cdots \big(a^{(t)}_{k(t),n_{k(t)}}\big)^\circ\Big),$$

where $k(1) \leqslant m_1, \ldots, k(t) \leqslant m_t$. Given that $(I_j : j \in J)$ is a partition of I and that $j(l) \neq j(l+1)$ for every $l \leqslant t-1$, it follows that $a^{(l)}_{k(l),n_{k(l)}}$ and $a^{(l+1)}_{k(l+1),1}$ are from different algebras $\mathcal{A}_i$ for every $l \leqslant t-1$. Thus, since the $\mathcal{A}_i$ are free from each other, it follows that

$$\varphi\Big(\big(a^{(1)}_{k(1),1}\big)^\circ \cdots \big(a^{(1)}_{k(1),n_{k(1)}}\big)^\circ \cdots \big(a^{(t)}_{k(t),1}\big)^\circ \cdots \big(a^{(t)}_{k(t),n_{k(t)}}\big)^\circ\Big) = 0,$$

from which we conclude that $\varphi(b_1^\circ \cdots b_t^\circ) = 0$, as desired.

(4). Let $i(1), \ldots, i(s) \in I$ be such that $i(1) \neq i(2) \neq \cdots \neq i(s)$, and let the variables $a_1 \in \mathcal{A}_{i(1)}, \ldots, a_s \in \mathcal{A}_{i(s)}$ be arbitrary. Let $j(1), \ldots, j(s) \in J$ be defined as follows: for every $l \leqslant s$, $j(l) \in J$ is the unique index such that $i(l) \in I_{j(l)}$. Then, define the elements $b_1, \ldots, b_t \in \mathcal{A}$ (with $t \leqslant s$) as follows:

$$b_1 = \begin{cases} a_1^\circ & \text{if } j(1) \neq j(2); \\ a_1^\circ a_2^\circ & \text{if } j(1) = j(2) \neq j(3); \\ \cdots & \\ a_1^\circ \cdots a_s^\circ & \text{if } j(1) = \cdots = j(s). \end{cases}$$

Then, in the event that $b_1 = a_1 \cdots a_{l_1}$ with $l_1 < s$, define

$$b_2 = \begin{cases} a^\circ_{l_1+1} & \text{if } j(l_1+1) \neq j(l_1+2); \\ a^\circ_{l_1+1} a^\circ_{l_1+2} & \text{if } j(l_1+1) = j(l_1+2) \neq j(l_1+3); \\ \cdots & \\ a^\circ_{l_1+1} \cdots a_s^\circ & \text{if } j(l_1+1) = \cdots = j(s), \end{cases}$$

and so on for $b_3, \ldots, b_t$ (in the event that $b_2 = a^\circ_{l_1+1} \cdots a^\circ_{l_2}$ with $l_2 < s$). Let $j'(1), \ldots, j'(t) \in J$ be defined as follows: for every $l \leqslant t$, $j'(l) \in J$ is the index such that $b_l \in \mathcal{B}_{j'(l)}$. Then, it clearly follows from the construction of the b_l that $j'(1) \neq j'(2) \neq \cdots \neq j'(t)$.

Given the definition of the b_l, we can write

$$a_1^\circ \cdots a_s^\circ = b_1 b_2 \cdots b_t = (a_1^\circ \cdots a^\circ_{l_1})(a^\circ_{l_1+1} \cdots a^\circ_{l_2}) \cdots (a^\circ_{l_{t-1}+1} \cdots a_s^\circ).$$

Since $i(1) \neq i(2) \neq \cdots \neq i(s)$ and the $(\mathcal{A}_i : i \in I_j)$ are free from each other for each $j \in J$, if follows that

$$\begin{aligned} b_1^\circ &= (a_1^\circ \cdots a^\circ_{l_1})^\circ = a_1^\circ \cdots a^\circ_{l_1}, \\ b_2^\circ &= (a^\circ_{l_1+1} \cdots a^\circ_{l_2})^\circ = a^\circ_{l_1+1} \cdots a^\circ_{l_2}, \\ &\cdots \\ b_t^\circ &= (a^\circ_{l_{t-1}+1} \cdots a_s^\circ)^\circ = a^\circ_{l_{t-1}+1} \cdots a_s^\circ. \end{aligned}$$

Furthermore, since $j'(1) \neq j'(2) \neq \cdots \neq j'(t)$ and the $\mathcal{B}_j$ are free from each other, it follows that

$$(b_1^\circ \cdots b_t^\circ)^\circ = b_1^\circ \cdots b_t^\circ.$$

Therefore,

$$\begin{aligned} a_1^\circ \cdots a_s^\circ &= (a_1^\circ \cdots a^\circ_{l_1})(a^\circ_{l_1+1} \cdots a^\circ_{l_2}) \cdots (a^\circ_{l_{t-1}+1} \cdots a_s^\circ) \\ &= (a_1^\circ \cdots a^\circ_{l_1})^\circ (a^\circ_{l_1+1} \cdots a^\circ_{l_2})^\circ \cdots (a^\circ_{l_{t-1}+1} \cdots a_s^\circ)^\circ \\ &= b_1^\circ \cdots b_t^\circ = (b_1^\circ \cdots b_t^\circ)^\circ = (a_1^\circ \cdots a_s^\circ)^\circ, \end{aligned}$$

as desired. □

3.4.3. Uniqueness. As mentioned in the beginning of this section, if one assumes that a concept of independence must be associative, then the only choices are tensor and free independence. More precisely:

Theorem 3.39. *Let $\mathscr{I} = \left\{\mathbf{t}_n : P(n)^{(2)} \to \mathbb{C} : n \in \mathbb{N}\right\}$ be an independence rule. Suppose that $\mathscr{I}$ is associative, that is, for every collection $\left((\mathcal{A}_i, \varphi_i) : i \in I\right)$ of noncommutative probability spaces and partition $(I_j : j \in J)$ of the indexing set I, one has*

$$*_{i\in I}^{(\mathscr{I})} \varphi_i = *_{j\in J}^{(\mathscr{I})} \left(*_{i\in I_j}^{(\mathscr{I})} \varphi_i\right).$$

Then, $\mathscr{I}$ is either the independence rule associated to tensor independence, or the independence rule associated to free independence.

Proof Outline (according to [Sp97]). Let $\mathscr{T} = \left\{\mathbf{t}_n^{(\mathscr{T})} : P(n)^{(2)} \to \mathbb{C} : n \in \mathbb{N}\right\}$ be the rule of tensor independence and $\mathscr{F} = \left\{\mathbf{t}_n^{(\mathscr{F})} : P(n)^{(2)} \to \mathbb{C} : n \in \mathbb{N}\right\}$ be the rule of free independence. The proof follows the following steps:

Step (1) Show that for every $n \in \mathbb{N}$, one has $\mathbf{t}_n^{(\mathscr{T})}(\pi, \pi) = 1$ for all $\pi \in P(n)$, and

$$\mathbf{t}_n^{(\mathscr{F})}(\pi, \pi) = \begin{cases} 1 & \text{if } \pi \text{ is crossing; and} \\ 0 & \text{otherwise.} \end{cases}$$

(See Definition D.2 for the definition of a noncrossing partition.)

Step (2) Show that, assuming $\mathscr{I}$ is associative, for every $n \in \mathbb{N}$, either $\mathbf{t}_n(\pi, \pi) = 1$ for al $\pi \in P(n)$, or

$$\mathbf{t}_n(\pi, \pi) = \begin{cases} 1 & \text{if } \pi \text{ is crossing; and} \\ 0 & \text{otherwise.} \end{cases}$$

Step (3) Show that for every $n \in \mathbb{N}$ and $\pi \in P(n)$, the values $\mathbf{t}_n$ takes on (π, π) uniquely determines the values that it takes on (σ, π) for all $\sigma \leqslant \pi$.

□

Chapter 4

Free Calculus

The main reference works for the concepts and proofs in this chapter are [**VDN92**] Chapter 3, [**Ta12**] Section 2.5.4, [**AGZ09**] Section 5.3, and [**NS06**] Lectures 11 and 12.

4.1. Introduction

In the previous chapter, we have seen that free independence can be seen as an analog of classical independence for noncommutative random variables and that, given free variables $a_1, \ldots, a_n$ and an arbitrary noncommutative polynomial $P(x_1, \ldots, x_n)$, the distribution of $P(a_1, \ldots, a_n)$ is completely determined by the distributions of the individual variables $a_1, \ldots, a_n$. However, as evidenced by the proof of Proposition 3.20 and the subsequent computations in Example 3.21, actually doing so in practice is a highly nontrivial task. Consequently, one of the most natural and important problem in the area of free probability is the following:

Problem 4.1. Let $(\mathcal{A}, *, \varphi)$ be a $*$-probability space, let $a_1, \ldots, a_n \in \mathcal{A}$ be $*$-free from each other, and let $P \in \mathbb{C}\langle x_1, \ldots, x_n \rangle$ be a $*$-polynomial. Devise an algorithm that enables one to compute the $*$-distribution of $P(a_1, \ldots, a_n)$ out of the $*$-distributions of the individual variables $a_1, \ldots, a_n$.

This chapter is devoted to exposing some of the techniques that have been devised to attack special cases of the above problem. In several cases, techniques used in classical probability will be used as an inspiration. Other notable works omitted from this section include the multiplicative free convolution and the S-transform algorithm, which allow one to compute the density of the analytic distribution of a product of two free self-adjoint variables (see for instance [**VDN92**] Section 3.6 or [**NS06**] Lectures 17 and 18); and the recent work of R. Speicher et al. on the operator-valued free additive and multiplicative convolutions, which provides a complete solution to Problem 4.1 in the case where the variables $a_1, \ldots, a_n$ and $P(a_1, \ldots, a_n)$ are self-adjoint (see [**BSTV15**] and [**BMS13**]).

4.2. Free Cumulants and Large Sum Approximations

In this section, we are interested in the following special case:

Problem 4.2. Let $(\mathcal{A}, *, \varphi)$ be a $*$-probability space and let $a_1, \ldots, a_n \in \mathcal{A}$ be $*$-free from each other (for a *large* integer n). *Approximate* the distribution of $a_1 + \cdots + a_n$. (In other words, find a limit distribution for $C(n) \cdot (a_1 + \cdots + a_n)$, where $C(n)$ is an appropriate normalization constant.)

4.2.1. Classical Cumulants and Central Limit Theorem. Let $(\Omega, \mathscr{A}, P)$ be a probability space and $X \in L^{\infty-}$ be a $\mathbb{R}$-valued random variable. The **moment generating function** $M_X : \mathbb{C} \to \mathbb{C} \cup \{\infty\}$ of X is defined as

$$M_X(z) = \mathbf{E}\big[\exp(zX)\big] = \sum_{n=0}^{\infty} \frac{z^n}{n!}\mathbf{E}[X^n], \qquad z \in \mathbb{C}.$$

Then, the **cumulant generating function** $C_X : \mathbb{C} \to \mathbb{C} \cup \{\infty\}$ is defined as $C_X(z) = \ln\big(M_X(z)\big)$. Since the formal power series for $\ln(1+z)$ is given by

$$\ln(1+z) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1} z^n}{n},$$

it follows that

$$C_X(z) = \ln\Big(1 + \big(M_X(z) - 1\big)\Big) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \left(\sum_{k=1}^{\infty} \frac{z^k}{k!}\mathbf{E}[X^k]\right)^n.$$

By expanding the above and performing a combinatorial analysis of the resulting coefficients, one obtains the following formal expansion (see **[Sp86]** equations (1.2) and (3.2)):

$$C_X(z) = \sum_{n=1}^{\infty} \frac{c_n[X] z^n}{n!},$$

where, for each $n \in \mathbb{N}$, the term

$$c_n[X] = \sum_{\pi \in P(n)} (-1)^{\#(\pi)-1}\big(\#(\pi)-1\big)! \prod_{B \in \pi} \mathbf{E}[X^{|B|}] = \sum_{\pi \in P(n)} \mu(\pi, 1_n) \prod_{B \in \pi} \mathbf{E}[X^{|B|}],$$

($\#(\pi)$ denotes the number of blocks in the partition π, and μ denotes the Möbius function on $P(n)^{(2)}$) is called the **classical cumulant of order n** of X. This leads to the following definition:

Definition 4.3. Let $(\mathcal{A}, *, \varphi)$ be a $*$-probability space and $a \in \mathcal{A}$ be a self-adjoint random variable. For each $n \in \mathbb{N}$, the term

$$c_n[a] = \sum_{\pi \in P(n)} \mu(\pi, 1_n) \prod_{B \in \pi} \varphi(a^{|B|}), \tag{4.1}$$

is called the **classical cumulant of order n** of a.

Example 4.4. Let a be an arbitrary self-adjoint variable. Then, $c_1[a] = \varphi(a)$, and

$$\begin{aligned} c_2[a] =& (-1)^{\#(0_2)-1}\big(\#(0_2)-1\big)!\varphi(a)^2 + (-1)^{\#(1_2)-1}\big(\#(1_2)-1\big)!\varphi(a^2) \\ =& -\varphi(a)^2 + \varphi(a^2) = \mathbf{Var}[a]. \end{aligned}$$

Remark 4.5. It clearly follows from (4.1) that the cumulants of a self-adjoint random variable are determined by its moments. Conversely, it can be shown **[Sp86]** that for every $n \in \mathbb{N}$,

$$\varphi(a^n) = \sum_{\pi \in P(n)} \prod_{B \in \pi} c_{|B|}[a], \tag{4.2}$$

and thus, the cumulants also determine the moments of a, and thus its distribution.

When one is interested in the distribution of sums of independent random variables, one of the advantage of using cumulants instead of moments comes from the following remark:

Remark 4.6. Let $X, Y \in L^{\infty-}$ be independent $\mathbb{R}$-valued random variables on a probability space $(\Omega, \mathscr{A}, P)$. Then, for every $t \in \mathbb{R}$, one has

$$\begin{aligned} C_{X+Y}(z) =& \ln\Big(\mathbf{E}\big[\exp(z(X+Y))\big]\Big) = \ln\Big(\mathbf{E}\big[\exp(zX)\exp(zY)\big]\Big) \\ =& \ln\Big(\mathbf{E}\big[\exp(zX)\big]\mathbf{E}\big[\exp(zY)\big]\Big) = \ln\Big(\mathbf{E}\big[\exp(zX)\big]\Big)\ln\Big(\mathbf{E}\big[\exp(zY)\big]\Big) \\ =& C_X(z) + C_Y(z), \end{aligned}$$

and thus, for every $n \in \mathbb{N}$, $c_n[X+Y] = c_n[X] + c_n[Y]$.

More generally, we have the following result:

Proposition 4.7 (**[Sp86]**). *Let $(\mathcal{A}, *, \varphi)$ be a $*$-probability space, and let $(a_i : i \in I) \subset \mathcal{A}$ be a collection of self-adjoint variables. Then, the a_i are $*$-tensor independent of each other if and only if for every $n \in \mathbb{N}$ and distinct indices $i(1), \ldots, i(k) \in I$ $(k \in \mathbb{N})$, one has*

$$c_n[a_{i(1)} + \cdots + a_{i(k)}] = c_n[a_{i(1)}] + \cdots + c_n[a_{i(k)}].$$

This leads, together with the following lemma, to a remarkably simple proof of the classical central limit theorem:

Lemma 4.8. *Let X be a standard Gaussian random variable (that is, of mean 0 and variance 1). Then, one has*

$$c_n[X] = \begin{cases} 1 & \text{if } n = 2\text{; and} \\ 0 & \text{otherwise.} \end{cases}$$

Proof. The fact that $c_1[X] = 0$ and $c_2[X] = 1$ clearly follows from Example 4.4. Suppose that $n \geqslant 3$. Given that the sum of two independent Gaussian random variables Y_1 and Y_2 is a Gaussian random variable of mean $\mathbf{E}[Y_1] + \mathbf{E}[Y_2]$ and variance $\mathbf{Var}[Y_1] + \mathbf{Var}[Y_2]$,[1] it follows that, given two i.i.d. copies X_1 and X_2 of X, the variables X and $\frac{1}{\sqrt{2}}X_1 + \frac{1}{\sqrt{2}}X_2$ are identically distributed. It then follows from Proposition 4.7 that

$$c_n[X] = c_n\left[\frac{1}{\sqrt{2}}X_1 + \frac{1}{\sqrt{2}}X_2\right] = c_n\left[\frac{1}{\sqrt{2}}X_1\right] + c_n\left[\frac{1}{\sqrt{2}}X_2\right].$$

Furthermore, one easily concludes from equation (4.1) that for every random variable Y and scalar $k \in \mathbb{R}$, $c_n[kY] = k^n c_n[Y]$. Thus,

$$c_n[X] = \frac{2}{\sqrt{2}^n} c_n[X],$$

[1] See Proposition 4.32.

from which it necessarily follows that $c_n[X] = 0$, as desired. □

Central Limit Theorem. *Let $(\mathcal{A}, *, \varphi)$ be a $*$-probability space, and let $(a_i : i \in I) \subset \mathcal{A}$ be self-adjoint, $*$-tensor independent of each other, and such that for every $i \in I$, $\varphi(a_i) = 0$ and $\varphi(a_i^2) = 1$. For every $n \in \mathbb{N}$, let us define $S_n = \frac{1}{\sqrt{n}}(a_1 + \cdots + a_n)$. If $\sup\{|\varphi(a_i^m)| : i \in I\} < \infty$ for all $m \in \mathbb{N}$, then, as $n \to \infty$, the variable S_n converges in distribution to a standard Gaussian variable.*

Proof. According to the Lemma 4.8 and Equation (4.2), the result will follow if we show that

$$\lim_{n\to\infty} c_m[S_n] = \begin{cases} 1 & \text{if } m = 2; \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

Given that the a_i are independent and of mean zero, it clearly follows that $c_1[S_n] = 0$ for every $n \in \mathbb{N}$. Furthermore, since the a_i are independent and of variance one, it follows that

$$c_2[S_n] = \mathbf{Var}[S_n] = \frac{1}{n}\sum_{i=1}^{n} \mathbf{Var}[a_i] = 1$$

for every $n \in \mathbb{N}$. Suppose that $m \geqslant 3$. According to Proposition 4.7 and the fact that the a_i are $*$-tensor independent of each other, it follows that

$$c_m[S_n] = \frac{1}{\sqrt{n}^m}\sum_{i=1}^{n} c_m[a_i] = \frac{1}{\sqrt{n}^m}\sum_{i=1}^{n}\left(\sum_{\pi\in P(m)} \mu(\pi, 1_m)\prod_{B\in\pi}\varphi(a_i^{|B|})\right).$$

Let us define the following constants (that only depend on m):

$$C_1(m) = \sup\Big\{\sup\big\{|\varphi(a_i^p)| : i \in I\big\} : p \leqslant m\Big\} < \infty$$
$$C_2(m) = \sup\Big\{\big|\mu(\pi, 1_m)\big| : \pi \in P(m)\Big\} < \infty$$

Then,

$$\begin{aligned}
\big|c_m[S_n]\big| &= \left|\frac{1}{\sqrt{n}^m}\sum_{i=1}^{n}\left(\sum_{\pi\in P(m)} \mu(\pi, 1_m)\prod_{B\in\pi}\varphi(a_i^{|B|})\right)\right| \\
&\leqslant \frac{1}{\sqrt{n}^m}\sum_{i=1}^{n}\left(\sum_{\pi\in P(m)} \big|\mu(\pi, 1_m)\big|\prod_{B\in\pi}\big|\varphi(a_i^{|B|})\big|\right) \\
&\leqslant \frac{1}{\sqrt{n}^m}\sum_{i=1}^{n}\left(\sum_{\pi\in P(m)} \big|\mu(\pi, 1_m)\big| C_1(m)^{\#(\pi)}\right) \\
&\leqslant \frac{1}{\sqrt{n}^m}\sum_{i=1}^{n}\left(\sum_{\pi\in P(m)} \big|\mu(\pi, 1_m)\big| C_1(m)^{m}\right) \\
&\leqslant |P(m)| C_1(m)^m C_2(m)\frac{n}{\sqrt{n}^m}
\end{aligned}$$

Since $m \geqslant 3$, the above converges to zero as $n \to \infty$, as desired. □

In similar fashion to how moments in a random variable a are generalized to mixed moments in multiple random variables $(a_i : i \in I)$ in order to explain the joint behaviour of said variables, cumulants in one variable can be extended to mixed cumulants, which describe the interactions between different variables.

Definition 4.9. Let $(\mathcal{A}, *, \varphi)$ be a $*$-probability space, and let $(a_i : i \in I) \subset \mathcal{A}$ be such that the algebra $*$-alg$(a_i : i \in I)$ is commutative (which implies in particular that the a_i are normal). For every $n \in \mathbb{N}$, the **mixed classical cumulants of order n** in the a_i are the expressions of the form

$$c_m[a_{i(1)}, \ldots, a_{i(m)}] = \sum_{\pi \in P(m)} \mu(\pi, 1_m)\varphi_\pi[a_{i(1)}, \ldots, a_{i(m)}], \tag{4.3}$$

where $i(1), \ldots, i(m) \in I$ are not all equal to each other, and φ_π is defined as in Notation 3.29.

Remark 4.10. If the requirement that $i(1), \ldots, i(n) \in I$ are not all equal to each other in the previous definition is violated, then it clearly follows that (4.3) reduces in this case to a *regular* classical cumulant. Indeed, for every $n \in \mathbb{N}$, self-adjoint random variable a, and partition $\pi \in P(n)$, one has

$$\varphi_\pi[\underbrace{a, \ldots, a}_{n \text{ times}}] = \prod_{B \in \pi} \varphi(a^{|B|}).$$

Thus, in the sequel, we will use the symbols $c_n[a]$ and $c_n[a, \ldots, a]$ interchangeably.

Notation 4.11. Let $(\mathcal{A}, *, \varphi)$ be a $*$-probability space, and let $(a_i : i \in I) \subset \mathcal{A}$ be such that $*$-alg$(a_i : i \in I)$ is commutative. For every $n \in \mathbb{N}$ and partition $\pi \in P(n)$, let

$$c_\pi[a_{i(1)}, \ldots, a_{i(n)}] = \prod_{B \in \pi} c_B[a_{i(1)}, \ldots, a_{i(n)}],$$

where, given a block $B = \{j(1), \ldots, j(s)\} \in \pi$, $c_B = c_{|B|}[a_{i(j(1))}, \ldots, a_{i(j(s))}]$.

Example 4.12. Let $\pi = \big\{\{1, 3\}, \{2, 6\}, \{4, 5, 7\}\big\} \in P(7)$. Then,

$$c_\pi[a_1, \ldots, a_7] = c_2[a_1, a_3]c_2[a_2, a_6]c_3[a_4, a_5, a_7].$$

Equation (4.2) can also be extended to mixed cumulants, which yields the following relation between mixed moments and mixed cumulants:

Proposition 4.13 (**[Sp86]**). *Let $(\mathcal{A}, *, \varphi)$ be a $*$-probability space, let $n \in \mathbb{N}$, and let $a_1, \ldots, a_n \in \mathcal{A}$ be such that $*$-alg$(a_i : i \leqslant n)$ is commutative. Then,*

$$\varphi(a_1 \cdots a_n) = \sum_{\pi \in P(n)} c_\pi[a_1, \ldots, a_n]. \tag{4.4}$$

Remark 4.14. In the previous section, it was shown that classical cumulants uniquely determine the distribution of self-adjoint elements. According to (4.4), one notices that the $*$-distribution of any normal random variable a is uniquely determined by mixed cumulants in a and a^*. Indeed, for any $n \in \mathbb{N}$ and $e(1), \ldots, e(n) \in \{1, *\}$, one has

$$\varphi(a^{e(1)} \cdots a^{e(n)}) = \sum_{\pi \in P(n)} c_\pi[a^{e(1)}, \ldots, a^{e(n)}].$$

The main properties of interest of mixed classical cumulants are summarized in the following theorem.

Theorem 4.15 ([**Sp86**] Section 4)**.** *Let $(\mathcal{A}, *, \varphi)$ be a $*$-probability space and let $(a_i : i \in I)$ be a collection of variables such that $*$-alg$(a_i : i \in I)$ is commutative. Then,*

(1) classical cumulants are multilinear, that is, for every $n, k_1, \ldots, k_n \in \mathbb{N}$,

$$c_n\left[\sum_{j_1=1}^{k_1} \alpha_{1j_1} a_{i_1(j_1)}, \ldots, \sum_{j_n=1}^{k_n} \alpha_{nj_n} a_{i_n(j_n)}\right] = \sum_{j_1=1}^{k_1} \cdots \sum_{j_n=1}^{k_n} (\alpha_{1j_1} \cdots \alpha_{nj_n}) c_n[a_{i_1(j_1)}, \ldots, a_{i_n(j_n)}],$$

where the α_{mj_m} are complex scalars and the $i_m(j_m) \in I$; and

(2) classical mixed cumulants characterize $*$-tensor independence, that is, the $(a_i : i \in I)$ are $*$-tensor independent of each other if and only if for every $n \geqslant 2$, $e(1), \ldots, e(n) \in \{1, *\}$, and collection of indices $i(1), \ldots, i(n) \in I$ not all equal to each other, $c_n\left[a_{i(1)}^{e(1)}, \ldots, a_{i(n)}^{e(1)}\right] = 0$.

From this, one can derive the following extension of the Central Limit Theorem:

Theorem 4.16. *Let $(\mathcal{A}, *, \varphi)$ be a $*$-probability space, and let $(a_i : i \in I) \subset \mathcal{A}$ be $*$-tensor independent of each other and such that $*$-alg$(a_i : i \in I)$ is commutative. Suppose that*

(1) for every $i \in I$, $\varphi(a_i) = 0$ and $\mathbf{Var}[a_i] = \varphi(a_i a_i^*) = 1$;

(2) the limit

$$C = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \varphi(a_i^2)$$

exists and is such that $C \in [-1, 1]$; and

(3) for every $m \in \mathbb{N}$ and $e(1), \ldots, e(m) \in \{1, *\}$,

$$\sup\left\{|\varphi(a_i^{e(1)} \cdots a_i^{e(m)})| : i \in I\right\} < \infty.$$

Then, the variable $S_n = \frac{1}{\sqrt{n}}(a_1 + \cdots + a_n)$ converges in $*$-distribution towards a variable $s = \frac{1}{\sqrt{2}}(\lambda_1 \cdot g_1 + \mathrm{i}\lambda_2 \cdot g_2)$, where g_1 and g_2 are $*$-tensor independent standard Gaussian variables, $\lambda_1 = \sqrt{1 + C}$, and $\lambda_2 = \sqrt{1 - C}$.

Proof. First, we compute the cumulants of the candidate limit s. According to Theorem 4.15, the $*$-tensor independence of the a_i and the multilinearity of classical cumulants implies that for every and $m \in \mathbb{N}$ and $e(1), \ldots, e(m) \in \{1, *\}$, one has

$$c_m[s^{e(1)}, \ldots, s^{e(m)}] = \frac{1}{\sqrt{2}^m}\left(\lambda_1^m c_m[g_1^{e(1)}, \ldots, g_1^{e(m)}] + (-1)^{k_{e(1),\ldots,e(m)}} (\mathrm{i}\lambda_2)^m c_m[g_2^{e(1)}, \ldots, g_2^{e(m)}]\right),$$

where $k_{e(1),\ldots,e(m)}$ is the number of exponents $e(1),\ldots,e(m)$ that are equal to $*$. Since g_1 and g_2 are standard Gaussian variables, the above cumulant vanishes for all m except $m=2$. As for $m=2$, we see that

$$c_2[s] = \frac{\lambda_1^2}{2}c_2[g_1] - \frac{\lambda_2^2}{2}c_2[g_1] = \frac{1+C}{2} - \frac{1-C}{2} = C,$$

and similarly $c_2[s^*] = C$. Furthermore,

$$c_2[s,s^*] = \frac{|\lambda_1|^2}{2}c_2[g_1,g_1^*] + \frac{|\lambda_2|^2}{2}c_2[g_1,g_1^*] = \frac{1+C}{2} + \frac{1-C}{2} = 1,$$

and similarly, $c_2[s^*,s] = 1$.

We now prove that the cumulants of S_n converge to that of s as $n \to \infty$. According to Theorem 4.15, the $*$-tensor independence of the a_i and the multilinearity of classical cumulants implies that for every $n \in \mathbb{N}$, $m \in \mathbb{N}$ and $e(1),\ldots,e(m) \in \{1,*\}$, one has

$$c_m[S_n^{e(1)},\ldots,S_n^{e(m)}] = \frac{1}{\sqrt{n}^m}\sum_{i(1),\ldots,i(m)=1}^{n} c_m[a_{i(1)}^{e(1)},\ldots,a_{i(m)}^{e(m)}] = \frac{1}{\sqrt{n}^m}\sum_{i=1}^{n} c_m[a_i^{e(1)},\ldots,a_i^{e(m)}].$$

Given that $c_1[a] = \varphi(a)$ and that the a_i are centered (i.e., $\varphi(a_i)=0$), it clearly follows that $c_1[S_n] = c_1[S_n^*] = 0$ for all n.

Given that $c_2[a,b] = \varphi(ab) - \varphi(a)\varphi(b)$, it follows that

$$\lim_{n\to\infty} c_2[S_n] = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} c_2[a_i] = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\varphi(a_i^2) = C,$$

and similarly for $c_2[S_n^*]$. Furthermore,

$$\lim_{n\to\infty} c_2[S_n,S_n^*] = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} c_2[a_i,a_i^*] = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} 1 = 1,$$

and similarly, $c_2[S_n^*,S_n] \to 1$.

Let $m \geqslant 3$ and $e(1),\ldots,e(m) \in \{1,*\}$ be arbitrary. Then,

$$c_m[S_n^{e(1)},\ldots,S_n^{e(m)}] = \frac{1}{\sqrt{n}^m}\sum_{i=1}^{n}\left(\sum_{\pi\in P(m)}\mu(\pi,1_m)\varphi_\pi[a_i^{e(1)},\ldots,a_i^{e(m)}]\right).$$

Let

$$C_1(m) = \sup\Big\{\sup\big\{|\varphi(a_i^{f(1)},\ldots,a_i^{f(p)})| : i\in I\big\} : p\leqslant m,\ f(1),\ldots,f(p)\in\{1,*\}\Big\} < \infty$$

$$C_2(m) = \sup\Big\{\big|\mu(\pi,1_m)\big| : \pi\in P(m)\Big\} < \infty.$$

Then, since $m \geqslant 3$, it follows that

$$\lim_{n\to\infty}\big|c_m[S_n^{e(1)},\ldots,S_n^{e(m)}]\big| \leqslant |P(m)|C_1(m)^m C_2(m)\lim_{n\to\infty}\frac{n}{\sqrt{n}^m} = 0,$$

as desired. □

4.2.2. Free Cumulants and Central Limit Theorem.

Definition 4.17. Let $(\mathcal{A}, *, \varphi)$ be a $*$-probability space. We call a **free cumulant** any expression of the form

$$\kappa_\pi[a_1, \ldots, a_n] = \sum_{\sigma \in NC(n):\ \sigma \leqslant \pi} \mu(\sigma, 1_n)\varphi_\pi[a_1, \ldots, a_n], \tag{4.5}$$

where $n \in \mathbb{N}$, $\pi \in NC(n)$, and $a_1, \ldots, a_n \in \mathcal{A}$.

Notation 4.18. Let $(\mathcal{A}, *, \varphi)$ be a $*$-probability space and $n \in \mathbb{N}$ be fixed.

(1) In the special case where $\pi = 1_n$, we denote $\kappa_n[a_1, \ldots, a_n] = \kappa_{1_n}[a_1, \ldots, a_n]$ for any $a_1, \ldots, a_n \in \mathcal{A}$;

(2) if $a_1, \ldots, a_n = a \in \mathcal{A}$, then we denote $\kappa_\pi[a] = \kappa_\pi[a_1, \ldots, a_n]$;

(3) if $a_1, \ldots, a_n \in \mathcal{A}$ are not all equal to each other, then $\kappa_n[a_1, \ldots, a_n]$ is called a **mixed free cumulant**; and

(4) given a collection $(a_i : i \in I) \subset \mathcal{A}$, we call a free cumulant in the variables a_i any expression of the form $\kappa_n\left[a_{i(1)}^{e(1)}, \ldots, a_{i(n)}^{e(n)}\right]$, where $i(1), \ldots, i(n) \in I$ and $e(1), \ldots, e(n) \in \{1, *\}$.

Remark 4.19. One can notice that the definition of free cumulants (4.5) is essentially a copy of the definition of classical mixed cumulants (4.3) with the lattice of partitions $P(n)$ replaced by the lattice of noncrossing partition $NC(n)$. This is consistent with the difference between the independence rules for tensor and free independence uncovered in the proof sketch of Theorem 3.39.

Example 4.20. Let $a, b \in (\mathcal{A}, *, \varphi)$ be noncommutative random variables. Then, $\kappa_1[a] = \varphi(a)$, and it follows from Example D.16 that $\kappa_2[a, b] = \mu(0_2, 1_2)\varphi(a)\varphi(b) + \mu(1_2, 1_2)\varphi(ab) = -\varphi(a)\varphi(b) + \varphi(ab)$, as in the classical case.

Example 4.21. Let us define $\pi_i \in NC(4)$ $(i \leqslant 14)$ and $\Delta(E)$ as in Example D.17, that is,

$$\begin{array}{rclrcl}
\pi_1 &=& 0_4 = \{\{1\},\{2\},\{3\},\{4\}\} & \pi_8 &=& \{\{1,4\},\{2,3\}\} \\
\pi_2 &=& \{\{1\},\{2\},\{3,4\}\} & \pi_9 &=& \{\{1,2\},\{3,4\}\} \\
\pi_3 &=& \{\{1\},\{2,3\},\{4\}\} & \pi_{10} &=& \{\{1,2,3\},\{4\}\} \\
\pi_4 &=& \{\{1,2\},\{3\},\{4\}\} & \pi_{11} &=& \{\{1\},\{2,3,4\}\} \\
\pi_5 &=& \{\{1,4\},\{2\},\{3\}\} & \pi_{12} &=& \{\{1,3,4\},\{2\}\} \\
\pi_6 &=& \{\{1,3\},\{2\},\{4\}\} & \pi_{13} &=& \{\{1,2,4\},\{3\}\} \\
\pi_7 &=& \{\{1\},\{2,4\},\{3\}\} & \pi_{14} &=& 1_4 = \{\{1,2,3,4\}\}
\end{array}.$$

Then, given noncommutative random variables a_1, a_2, a_3 and a_4 in a $*$-probability space $(\mathcal{A}, *, \varphi)$ such that $\varphi(a_l) = 0$ $(l \leqslant 4)$, one has

$$\kappa_n[a_1, a_2, a_3, a_4] = \sum_{i=1}^{14} \varphi_{\pi_i}[a_1, a_2, a_3, a_4]\mu(\pi_i, 1_4).$$

Since $\varphi(a_l) = 0$, the only noncrossing partitions for which $\varphi_{\pi_i}[a_1, a_2, a_3, a_4]$ does not necessarily vanish are π_8,π_9 and $\pi_{14} = 1_4$. Moreover, given the entries of the matrix $\Delta(E)^{-1}$ in Example

D.17, it follows that $\mu(\pi_8, 1_4) = \mu(\pi_9, 1_4) = -1$ and $\mu(1_4, 1_4) = 1$.[2] Thus,

$$\kappa_4[a_1, a_2, a_3, a_4] = \varphi(a_1a_2a_3a_4) - \varphi(a_1a_2)\varphi(a_3a_4) - \varphi(a_1a_4)\varphi(a_2a_3).$$

One similarly finds that

$$\kappa_{\pi_8}[a_1, a_2, a_3, a_4] = \varphi(a_1a_4)\varphi(a_2a_3) \text{ and } \kappa_{\pi_9}[a_1, a_2, a_3, a_4] = \varphi(a_1a_2)\varphi(a_3a_4).$$

Equation (4.4) further established the link between classical cumulants and moments. In the following proposition, we see that a very similar relationship holds between free cumulants and moments.

Proposition 4.22 ([**NS06**] Proposition 11.4)**.** *Let $(\mathcal{A}, *, \varphi)$ be a $*$-probability space, For every $n \in \mathbb{N}$, variables $a_1, \ldots, a_n \in \mathcal{A}$, and partition $\pi \in NC(n)$ one has*

$$\varphi_\pi[a_1, \ldots, a_n] = \sum_{\sigma \in NC(n):\ \sigma \leqslant \pi} \kappa_\pi[a_1, \ldots, a_n]. \tag{4.6}$$

Proof. Let $n \in \mathbb{N}$, and $a_1, \ldots, a_n \in \mathcal{A}$ be fixed. Define the functions $f, g : NC(n) \to \mathbb{C}$ as follows: for every $\pi \in NC(n)$, let $f(\pi) = \varphi_\pi[a_1, \ldots, a_n]$ and $g(\pi) = \kappa_\pi[a_1, \ldots, a_n]$. Then, the result is a direct application of the Möbius Inversion Formula (see Proposition D.15), which states that

$$f(\pi) = \sum_{\sigma \in NC(n):\ \sigma \leqslant \pi} g(\sigma), \qquad \pi \in NC(n)$$

if and only if

$$g(\pi) = \sum_{\sigma \in NC(n):\ \sigma \leqslant \pi} f(\sigma)\mu(\sigma, \pi), \qquad \pi \in NC(n).$$

□

Remark 4.23. As a consequence of the above proposition, one concludes that the free cumulants also uniquely determine $*$-distributions and joint $*$-distributions in $*$-probability spaces. Indeed, given a collection $(a_i : i \in I)$ of variables in a $*$-probability space $(\mathcal{A}, *, \varphi)$, Equation (4.6) implies that for every $n \in \mathbb{N}$, indices $i(1), \ldots, i(n) \in I$, and exponents $e(1), \ldots, e(n) \in \{1, *\}$, one has

$$\varphi\Big(a_{i(1)}^{e(1)} \cdots a_{i(n)}^{e(n)}\Big) = \sum_{\pi \in NC(n)} \kappa_\pi\Big[a_{i(1)}^{e(1)}, \ldots, a_{i(n)}^{e(1)}\Big].$$

Proposition 4.24 ([**NS06**] Proposition 11.4)**.** *Let $(\mathcal{A}, *, \varphi)$ be a $*$-probability space. For every $n \in \mathbb{N}$, variables $a_1, \ldots, a_n \in \mathcal{A}$, and partition $\pi \in NC(n)$, one has*

$$\kappa_\pi[a_1, \ldots, a_n] = \prod_{B \in \pi} \kappa_B[a_1, \ldots, a_n],$$

where, given a block $B = \{j(1), \ldots, j(s)\} \in \pi$ with $j(1) < j(2) < \cdots < j(s)$, we define $\kappa_B[a_1, \ldots, a_n] = \kappa_{|B|}[a_{j(1)}, \ldots, a_{j(s)}]$.

Theorem 4.25. *Let $(\mathcal{A}, *, \varphi)$ be a $*$-probability space and let $(a_i : i \in I)$ be a collection of random variables. Then,*

[2] see Proposition D.11 and Remark D.14

(1) free cumulants are multilinear, that is, for every $n, k_1, \ldots, k_n \in \mathbb{N}$,

$$\kappa_n\left[\sum_{j_1=1}^{k_1} \alpha_{1j_1} a_{i_1(j_1)}, \ldots, \sum_{j_n=1}^{k_n} \alpha_{nj_n} a_{i_n(j_n)}\right] = \sum_{j_1=1}^{k_1} \cdots \sum_{j_n=1}^{k_n} (\alpha_{1j_1} \cdots \alpha_{nj_n}) \kappa_n[a_{i_1(j_1)}, \ldots, a_{i_n(j_n)}],$$

where the α_{mj_m} *are complex scalars and the* $i_m(j_m) \in I$*; and*

(2) mixed free cumulants characterize $*$*-free independence, that is, the* $(a_i : i \in I)$ *are* $*$*-free from each other if and only if mixed cumulants in the* a_i *vanish, that is, for every* $n \geqslant 2$, $e(1), \ldots, e(n) \in \{1, *\}$*, and collection of indices* $i(1), \ldots, i(n) \in I$ *not all equal to each other,* $\kappa_n\left[a_{i(1)}^{e(1)}, \ldots, a_{i(n)}^{e(n)}\right] = 0$.

The fact that free cumulants are multilinear (i.e., part (1) in the above theorem) follows directly from (4.5): indeed, every term of the form $\varphi_\pi[a_1, \ldots, a_n]$ is multilinear. The proof that mixed free cumulants characterize $*$-freeness is more difficult and requires the following preliminary results:

Lemma 4.26 ([**NS06**] Theorem 11.12)**.** *Let* $(\mathcal{A}, *, \varphi)$ *be a* $*$*-probability space and let* $(a_i : i \in I)$ *be a collection of random variables. Let* $m \in \mathbb{N}$ *be arbitrary, and let* $i(1), \ldots, i(m) \in I$*. Let* $k_1, \ldots, k_m \in \mathbb{N}$*, and, for each* $j \leqslant m$*, let* $A_j = a_{i(j)}^{e_j(1)} \cdots a_{i(j)}^{e_j(k_j)}$*, where* $e_j(1), \ldots, e_j(k_j) \in \{1, *\}$*. Then,*

$$\kappa_m[A_1, \ldots, A_m] = \sum_{\pi \vee \sigma = 1_n} \kappa_\pi\left[a_{i(1)}^{e_1(1)}, \ldots, a_{i(1)}^{e_1(k_1)}, \ldots, a_{i(m)}^{e_m(1)}, \ldots, a_{i(m)}^{e_m(k_m)}\right],$$

where $n = k_1 + \cdots + k_m$*,* $\vee$ *denotes the join of two partitions (see Notation D.5), and* $\sigma = \big\{\{1, \ldots, k_1\}, \{k_1 + 1, \ldots, k_1 + k_2\}, \ldots, \{k_1 + \cdots + k_{m-1} + 1, \ldots, n\}\big\} \in NC(n)$.

Lemma 4.27 ([**NS06**] Proposition 11.15)**.** *Let* $(\mathcal{A}, *, \varphi)$ *be a* $*$*-probability space, let* $n \geqslant 2$*, and let* $a_1, \ldots, a_n \in \mathcal{A}$*. If there exists* $i \leqslant n$ *such that* a_i *is a constant multiple of* $1_\mathcal{A}$*, then* $\kappa_n[a_1, \ldots, a_n] = 0$.

Lemma 4.28 ([**NS06**] Theorem 11.16)**.** *Let* $(\mathcal{A}, *, \varphi)$ *be a* $*$*-probability space, and let* $(\mathcal{A}_i : i \in I)$ *be* $*$*-subalgebras of* $\mathcal{A}$*. The* $\mathcal{A}_i$ *are* $*$*-free independent from each other if and only if mixed cumulants in variables in the* $\mathcal{A}_i$ *vanish, that is, for every* $n \geqslant 2$*, every collection of indices* $i(1), \ldots, i(n) \in I$ *not all equal to each other, and every* $a_1 \in \mathcal{A}_{i(1)}, \ldots, a_n \in \mathcal{A}_{i(n)}$*, one has* $\kappa_n[a_1, \ldots, a_n] = 0$.

Proof. We first show that the vanishing of mixed cumulants implies $*$-freeness of the $*$-algebras $\mathcal{A}_i$. Let $i(1), \ldots, i(n) \in I$ be such that $i(1) \neq i(2) \neq \cdots \neq i(n)$, and let $a_1 \in \mathcal{A}_{i(1)}, \ldots, a_n \in \mathcal{A}_{i(n)}$ be such that $\varphi(a_1) = \cdots = \varphi(a_n) = 0$. According to (4.6), one has

$$\varphi(a_1 \cdots a_n) = \sum_{\pi \in NC(n)} \kappa_\pi[a_1, \ldots, a_n].$$

Let $\pi \in NC(n)$ be a noncrossing partition. Then, either π contains a singleton block $B = \{i\}$, in which case $\varphi(a_i) = 0$ implies that $\kappa_B[a_1, \ldots, a_n] = 0$, or π contains a block B such that $s, t \in B$ and $i(s) \neq i(t)$ (since $i(1) \neq i(2) \neq \cdots \neq i(n)$ and π is noncrossing), in which

case $\kappa_B[a_1,\ldots,a_n] = 0$, since mixed cumulants in elements of the $\mathcal{A}_i$ vanish. Therefore, $\kappa_\pi[a_1,\ldots,a_n] = 0$, and thus $\varphi(a_1\cdots a_n)$ must be equal to zero.

We now prove that $*$-free independence implies the vanishing of mixed cumulants in the $\mathcal{A}_i$. Let $i(1),\ldots,i(n) \in I$ be such that $i(1) \neq i(2) \neq \cdots \neq i(n)$, and let $a_1 \in \mathcal{A}_{i(1)},\ldots,a_n \in \mathcal{A}_{i(n)}$ be such that $\varphi(a_1) = \cdots = \varphi(a_n) = 0$. According to (4.5), one has

$$\kappa_n[a_{i(1)},\ldots,a_{i(n)}] = \sum_{\pi\in NC(n)} \mu(\pi,1_n)\varphi_\pi[a_{i(1)},\ldots,a_{i(n)}].$$

Let π be a noncrossing partition. Then, either π contains a singleton $B = \{i\}$, in which case $\varphi_B[a_1,\ldots,a_n] = 0$ since $\varphi(a_i) = 0$, or π contains a block of the form $B = \{j, j+1,\ldots,j+s\}$, in which case $\varphi_B[a_1,\ldots,a_n] = 0$, since the $\mathcal{A}_i$ are $*$-free, $i(j) \neq i(j+1) \neq \cdots \neq i(j+s)$, and $\varphi(a_j) = \cdots = \varphi(a_{j+s}) = 0$. Thus, $\varphi_\pi[a_1,\ldots,a_n] = 0$, which implies that the mixed cumulant $\kappa_n[a_1,\ldots,a_n]$ vanishes. To prove the lemma in full generality, we must now get rid of the assumption that $\varphi(a_i) = 0$ for all $i \leqslant n$, and we must relax the condition that $i(1) \neq i(2) \neq \cdots \neq i(n)$ to simply assuming that $i(1),\ldots,i(n)$ are not all equal to each other.

We first deal with the assumption that $\varphi(a_i) = 0$, thus, let $i(1),\ldots,i(n) \in I$ be such that $i(1) \neq i(2) \neq \cdots \neq i(n)$, and let $a_1 \in \mathcal{A}_{i(1)},\ldots,a_n \in \mathcal{A}_{i(n)}$ be arbitrary. Then, since free cumulants are multilinear and cumulants with constant (i.e., $\mathbb{C}$-multiples $1_\mathcal{A}$) coefficients vanish (Lemma 4.27), it follows that

$$\kappa_\pi[a_1,\ldots,a_n] = \kappa_\pi[a_1^\circ + \varphi(a_1)\cdot 1_\mathcal{A},\ldots,a_n^\circ + \varphi(a_n)\cdot 1_\mathcal{A}] = \kappa_\pi[a_1^\circ,\ldots,a_n^\circ] + 0.$$

Thus, if the result holds for centered variables (i.e., $\varphi(a_i) = 0$), it also holds for arbitrary variables.

We now deal with the assumption that $i(1) \neq i(2) \neq \cdots \neq i(n)$. Let $i(1),\ldots,i(n) \in I$ be not all equal to each other, and let $a_1 \in \mathcal{A}_{i(1)},\ldots,a_n \in \mathcal{A}_{i(n)}$ be arbitrary. We proceed by induction on n. If $n = 2$, then it must be the case that $i(1) \neq i(2)$, and thus $\kappa_2[a_1,a_2] = 0$. Suppose that $n \geqslant 3$, and that for every $m < n$, one has $\kappa_m[b_1,\ldots,b_m] = 0$ whenever $b_1 \in \mathcal{A}_{j(1)},\ldots,b_m \in \mathcal{A}_{j(m)}$ with $j(1),\ldots,j(m) \in I$ not all equal to each other.

Define the elements $A_1,\ldots,A_m \in \mathcal{A}$ (with $m \leqslant n$) as follows:

$$A_1 = \begin{cases} a_1 & \text{if } i(1) \neq i(2); \\ a_1a_2 & \text{if } i(1) = i(2) \neq i(3); \\ \cdots & \\ a_1\cdots a_{n-1} & \text{if } i(1) = \cdots = i(n-1) \neq i(n). \end{cases}$$

Then, letting $A_1 = a_1\cdots a_{k_1}$, define

$$A_2 = \begin{cases} a_{k_1+1} & \text{if } i(k_1+1) \neq i(k_1+2); \\ a_{k_1+1}a_{k_1+2} & \text{if } i(k_1+1) = i(k_1+2) \neq i(k_1+3); \\ \cdots & \\ a_{k_1+1}\cdots a_n & \text{if } i(k_1+1) = \cdots = i(n), \end{cases}$$

and so on for $A_3,\ldots,A_m$ (in the event that $b_2 = a_{k_1+1}\cdots a_{k_1+k_2}$ with $k_1+k_2 < n$). Thus, we have that

$$a_1\cdots a_n = (a_1\cdots a_{k_1})(a_{k_1+1}\cdots a_{k_1+k_2})\cdots(a_{k_1+\cdots+k_{m-1}+1}\cdots a_n) = A_1\cdots A_m,$$

where $A_1 \in \mathcal{A}_{j(1)}, \ldots, A_m \in \mathcal{A}_{j(m)}$ with $j(1) \neq j(2) \neq \cdots \neq j(m)$. Thus, according to what we have established so far, $\kappa_m[A_1, \ldots, A_m] = 0$. According to Lemma 4.26, we have that

$$0 = \kappa_m[A_1, \ldots, A_m] = \sum_{\pi \vee \sigma = 1_n} \kappa_\pi[a_1, \ldots, a_n],$$

where

$$\sigma = \big\{\{1, \ldots, k_1\}, \{k_1 + 1, \ldots, k_1 + k_2\}, \ldots, \{k_1 + \cdots + k_{m-1} + 1, \ldots, n\}\big\} \in NC(n).$$

Since $1_n \vee \pi = 1_n$ for every partition π, we can rephrase the above as

$$0 = \kappa_n[a_1, \ldots, a_n] + \sum_{\pi \vee \sigma = 1_n,\ \pi \neq 1_n} \kappa_\pi[a_1, \ldots, a_n].$$

Let $\pi \in NC(n)$ be a noncrossing partition other that 1_n. Since $\pi \vee \sigma = 1_m$, there exists two distinct blocks $B_1, B_2 \in \sigma$ such that $l_1 \in B_1$, $l_2 \in B_2$, and l_1, l_2 are in the same block B of π. Since $i(l_1) \neq i(l_2)$ (given how the blocks of σ were defined) and $|B| < n$ (since $\pi \neq 1_n$), it follows from our induction hypothesis that $\kappa_B[a_1, \ldots, a_n] = 0$. Thus, $\kappa_\pi[a_1, \ldots, a_n] = 0$ for every π other than 1_n, which implies that $0 = \kappa_n[a_1, \ldots, a_n] + 0$, as desired. □

Proof of Theorem 4.25 Part (2). For every $i \in I$, let $\mathcal{A}_i = *\text{-alg}(1_{\mathcal{A}}, a_i)$. Then, it follows from Lemma 4.28 that for every $n \geqslant 2$, $*$-polynomials $P_1, \ldots, P_n \in \mathbb{C}\langle x, x^* \rangle$, and collection of indices $i(1), \ldots, i(n) \in I$ not all equal to each other,

$$\kappa_n\Big[P_1(a_{i(1)}), \ldots, P_n(a_{i(n)})\Big] = 0. \tag{4.7}$$

Thus, it only remains to show that the vanishing of mixed free cumulants in the a_i and a_i^* implies the vanishing of mixed free cumulants of the form (4.7).

Suppose that mixed cumulants in the a_i and a_i^* vanish. Let $m \geqslant 2$, let $M_1, \ldots, M_m \in \mathbb{C}\langle x, x^* \rangle$ be $*$-words, and let $i(1), \ldots, i(m) \in I$ be not all equal to each other. Given the definition of $*$-words, for every $t \leqslant m$, there exists $k_t \in \mathbb{N}$ and $e_t(1), \ldots, e_t(k_t) \in \{1, *\}$ such that

$$M_t(x) = x^{e_t(1)} \cdots x^{e_t(k_t)},$$

and thus, if one lets $n = k_1 + \cdots + k_m$ and

$$\sigma = \big\{\{1, \ldots, k_1\}, \{k_1 + 1, \ldots, k_1 + k_2\}, \ldots, \{k_1 + \cdots + k_{m-1} + 1, \ldots, n\}\big\} \in NC(n),$$

it follows from Lemma 4.26 that

$$\kappa_m\big[M_1(a_{i(1)}), \ldots, M_m(a_{i(m)})\big] = \sum_{\pi \vee \sigma = 1_n} \kappa_\pi\Big[a_{i(1)}^{e_1(1)}, \ldots, a_{i(1)}^{e_1(k_1)}, \ldots, a_{i(m)}^{e_m(1)}, \ldots, a_{i(m)}^{e_m(k_m)}\Big].$$

Let $\pi \in NC(n)$ be such that $\pi \vee \sigma = 1_n$. Since $i(1), \ldots, i(m)$ are not all equal to each other, then there exists $s < t \leqslant m$ such that $i(s) \neq i(t)$. Let

$$B_s = \{k_1 + \cdots + k_{s-1} + 1, \ldots, k_1 + \cdots + k_{s-1} + k_s\}$$

and

$$B_t = \{k_1 + \cdots + k_{t-1} + 1, \ldots, k_1 + \cdots + k_{t-1} + k_t\}$$

be the two blocks of σ associated to s and t. Since $\pi \vee \sigma = 1_n$, there must exist two elements $j_s \in B_s$ and $j_t \in B_t$ and a block $B \in \pi$ such that $j_s, j_t \in B$. Therefore,

$$\kappa_B\Big[a_{i(1)}^{e_1(1)}, \ldots, a_{i(1)}^{e_1(k_1)}, \ldots, a_{i(m)}^{e_m(1)}, \ldots, a_{i(m)}^{e_m(k_m)}\Big] = 0,$$

which in turn implies that

$$\kappa_\pi\Big[a_{i(1)}^{e_1(1)},\ldots,a_{i(1)}^{e_1(k_1)},\ldots,a_{i(m)}^{e_m(1)},\ldots,a_{i(m)}^{e_m(k_m)}\Big]=0.$$

Thus, we can conclude that $\kappa_m\big[M_1(a_{i(1)}),\ldots,M_m(a_{i(m)})\big]=0$, which then implies that $\kappa_m\big[P_1(a_{i(1)}),\ldots,P_m(a_{i($ 0 for every $*$-polynomials $P_1,\ldots,P_m\in\mathbb{C}\langle x,x^*\rangle$ by multilinearity of free cumulants. □

Having established these properties, we are now able to provide the following free analog of the central limit theorem, which allows one to approximate the distribution of some large sums of free random variables:

Lemma 4.29. *Let $(\mathcal{A},*,\varphi)$ be a $*$-probability space, and let s be a semicircular random variable (see Definition 2.28). Then, one has*

$$c_n[s]=\begin{cases}1 & \text{if } n=2\text{; and}\\ 0 & \text{otherwise.}\end{cases}$$

Proof. First, according to Example 4.20, it is clear that $\kappa_1[s]=\varphi(s)=0$ and $\kappa_1[s]=\varphi(s^2)-\varphi(s)^2=1$. Thus, we must now prove that $\kappa_n[s]=0$ for all $n\geqslant 3$.

Let $n\geqslant 3$ be odd. Then, according to equation (4.5), one has

$$\kappa_n[s]=\sum_{\pi\in NC(n)}\mu(\pi,1_n)\varphi_\pi[s,\ldots,s]=\sum_{\pi\in NC(n)}\mu(\pi,1_n)\prod_{B\in\pi}\varphi(s^{|B|}).$$

Given that $\varphi(s^m)=0$ whenever m is odd, and that any partition of $\{1,\ldots,n\}$ when n is odd is guaranteed to contain a block B such that $|B|$ is odd, it must be the case that $\kappa_n[a,\ldots,a]=0$, as desired.

We now prove that $\kappa_{2n}[s]=0$ for all $n\geqslant 2$. We proceed by induction: According to equation (4.5), one has

$$\kappa_4[s]=\sum_{\pi\in NC(4)}\mu(\pi,1_4)\prod_{B\in\pi}\varphi(s^{|B|}).$$

Again, since $\varphi(s)=0$, the only partitions of $NC(4)$ for which $\prod_{B\in\pi}\varphi(s^{|B|})$ does not vanish are those that contain only blocks with an even number of elements, namely (using the same notation as in Example 4.21), $\pi_8=\big\{\{1,4\},\{2,3\}\big\}$, $\pi_9=\big\{\{1,2\},\{3,4\}\big\}$, and $1_4=\big\{\{1,2,3,4\}\big\}$. Thus,

$$\kappa_4[s]=\mu(\pi_8,1_4)\varphi(s^2)^2+\mu(\pi_9,1_4)\varphi(s^2)^2+\mu(1_4,1_4)\varphi(s^4).$$

According to the definition of semicircular elements and the computations done in Example D.17, this yields

$$\kappa_4[s]=-2\varphi(s^2)^2+\varphi(s^4)=-2+\frac{1}{3}\binom{4}{2}=0.$$

Now for the induction hypothesis, suppose that $\kappa_{2k}[s]=0$ for all $1\leqslant k<n$. According to equation (4.6), it follows that

$$\frac{1}{n+1}\binom{2n}{n}=\varphi(s^{2n})=\sum_{\pi\in NC(2n)}\kappa_\pi[s].$$

Let $\pi \in NC(2n)$ be a *pairing*, that is, every block of π contains exactly two elements. Then, $\kappa_\pi[s] = \kappa_2[s]^n = 1$. Given that the number of noncrossing pairings of $\{1, \ldots, 2n\}$ is given by $\frac{1}{n+1}\binom{2n}{n}$,[3] it then follows that

$$0 = \sum_{\pi \in NC'(2n)} \kappa_\pi[s], \tag{4.8}$$

where $NC'(2n)$ is the set of noncrossing partitions of $\{1, \ldots, 2n\}$ that are not pairings. Let $\pi \in NC'(2n)$. Then, exactly one of the following cases must occur

(1) π contains a block B such that $|B|$ is odd, in which case $\kappa_{|B|}[s] = 0$, and thus $\kappa_\pi[s] = 0$;

(2) π contains a block B such that $4 \leqslant |B| < 2n - 2$, in which case $\kappa_{|B|}[s] = 0$ by our induction hypothesis, and hence $\kappa_\pi[s] = 0$; or

(3) $\pi = 1_{2n}$.

Therefore, it follows from (4.8) that

$$0 = \sum_{\pi \in NC'(2n)} \kappa_\pi[s] = \kappa_{1_{2n}}[s] + \sum_{\pi \in NC'(2n) \setminus \{1_{2n}\}} \kappa_\pi[s] = \kappa_{2n}[s] + 0,$$

which concludes the proof. □

Free Central Limit Theorem. *Let $(\mathcal{A}, *, \varphi)$ be a $*$-probability space, and let $(a_i : i \in I) \subset \mathcal{A}$ be $*$-free from each other. Suppose that*

(1) for every $i \in I$, $\varphi(a_i) = 0$ and $\mathbf{Var}[a_i] = \varphi(a_i a_i^*) = 1$;

(2) the limit

$$C = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \varphi(a_i^2)$$

exists and is such that $C \in [-1, 1]$; and

(3) for every $m \in \mathbb{N}$ and $e(1), \ldots, e(m) \in \{1, *\}$,

$$\sup \left\{ |\varphi(a_i^{e(1)} \cdots a_i^{e(m)})| : i \in I \right\} < \infty.$$

Then, the variable $S_n = \frac{1}{\sqrt{n}}(a_1 + \cdots + a_n)$ converges in distribution towards a variable $s = \frac{1}{\sqrt{2}}(\lambda_1 \cdot s_1 + \mathrm{i}\lambda_2 \cdot s_2)$, where s_1 and s_2 are $*$-free semicircular variables, $\lambda_1 = \sqrt{1 + C}$, and $\lambda_2 = \sqrt{1 - C}$.

Proof. Given Theorem 4.25 and Lemma 4.29, the proof of the Free Central Limit Theorem follows the same steps as that of Theorem 4.16. □

Remark 4.30. The Free Central Limit Theorem and Lemma 4.29 (together with their classical or tensor independent counterparts) both hint at the fact that the semicircular distribution can be seen as the *free analog* of the Gaussian distribution.

[3] See for instance **[NS06]** Lemma 8.9.

4.3. Free Additive Convolution

In this section, we are concerned with the following special case of Problem 4.1:

Problem 4.31. Let $(\mathcal{A}, *, \varphi)$ be a $*$-probability space, and let $a, b \in \mathcal{A}$ be $*$-free from each other and self-adjoint. Assume that a, b and $a+b$ have respective analytic distributions μ_a, μ_b and μ_{a+b} on $\big(\mathbb{R}, \mathscr{B}(\mathbb{R})\big)$, and that the analytic distributions have respective density functions f_a, f_b and f_{a+b}. Devise an algorithm to compute f_{a+b} from f_a and f_b.

Given two independent random variables $X, Y \in L^{\infty-}$ (with respective densities μ_X and μ_Y) defined on a probability space $(\Omega, \mathscr{A}, P)$, the distribution of $X+Y$, which we denote by $\mu_X * \mu_Y$ and call the **classical convolution** of μ_X and μ_Y, is given by the following formula:

$$\mu_X * \mu_Y(A) = \int_{\mathbb{C}} \int_{\mathbb{C}} \mathbf{1}_A(x+y)\ \mathrm{d}\mu_X(x)\mathrm{d}\mu_Y(y), \qquad A \in \mathscr{B}(\mathbb{C}) \tag{4.9}$$

and if X and Y have respective density functions f_X and f_Y on $\mathbb{R}$, then the density of $\mu_X * \mu_Y$, denoted $f_X * f_Y$, is given by

$$f_X * f_Y(w) = \int_{\mathbb{R}} f_X(x) f_Y(w-x)\ \mathrm{d}x, \qquad w \in \mathbb{R}. \tag{4.10}$$

Given two $*$-free variables a and b in a $*$-probability space $(\mathcal{A}, *, \varphi)$ with respective analytic distributions μ_a and μ_b, the analytic distribution of $a+b$ will be denoted $\mu_a \boxplus \mu_b$ and will be called the **free convolution** of μ_a and μ_b. Thus, the objective of this section will be to present algorithms that allows one to *decribe* $\mu_a \boxplus \mu_b$ using a density function from the density functions of μ_a and μ_b in the special case where a and b are self-adjoint (and thus μ_a, μ_b, and $\mu_a \boxplus \mu_b$ are measures on the real line $\mathbb{R}$).

4.3.1. Fourier Transform Algorithm for Tensor Independence. While equations (4.9) and (4.10) provide a solution to the classical counterpart of Problem 4.31, it is difficult to imagine how they could be generalized to the free/noncommutative case, as their proof is typically intimately linked with the measure-theoretic conceptualization of random variables.

Consequently, we consider for inspiration the following algorithm (called the **Fourier transform algorithm**), which is more consistent with the moment-centric conceptualization of random variables in free probability: Let $a, b \in (\mathcal{A}, *, \varphi)$ be self-adjoint variables with respective analytic distributions μ_a and μ_b, and suppose that μ_a and μ_b have respective densities f_a and f_b, and that a and b are tensor independent of each other. Then, we can obtain the density f_{a+b} of the analytic distribution μ_{a+b} of $a+b$ as follows:

Step 1. Compute the Fourier transforms $\widehat{f_a}$ and $\widehat{f_b}$ of the densities of a and b (see Section B.7).
Step 2. By using the relations $M_a(it) = \widehat{f_a}(t)$, $M_b(it) = \widehat{f_b}(t)$ (for every $t \in \mathbb{R}$), $C_a(z) = \ln\big(M_a(z)\big)$, and $C_b(z) = \ln\big(M_b(z)\big)$ (for all $z \in \mathbb{C}$), where M_a and M_b denote the moment generating functions of a and b respectively, and C_a and C_b denote the cumulant generating functions, compute

$$C_a(it) = \ln\big(M_a(it)\big) = \ln\big(\widehat{f_a}(t)\big)$$

and

$$C_b(it) = \ln\big(M_b(it)\big) = \ln\big(\widehat{f_b}(t)\big).$$

Step 3. By using the relations $M_{a+b}(it) = \widehat{f_{a+b}}(t)$ (for all $t \in \mathbb{R}$) and $M_{a+b}(z) = \exp\big(C_{a+b}(z)\big)$, as well as the fact that $C_{a+b} = C_a + C_b$ (since a and b are tensor independent), compute

$$\widehat{f_{a+b}}(t) = M_{a+b}(it) = \exp\big(C_a(it) + C_b(it)\big) = M_a(it)M_b(it) = \widehat{f_a}(t)\widehat{f_b}(t).$$

Step 4. Using the Inverse Fourier Transform (see Theorem B.15), compute the density of the analytic distribution of $a + b$ as follows:

$$f_{a+b}(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}tx}\widehat{f_{a+b}}(t)\ \mathrm{d}t, \qquad x \in \mathbb{R}.$$

For example, this algorithm can be used to establish the fact that a sum of two tensor independent Gaussian variables is again a Gaussian variable:

Proposition 4.32. *Let $a, b \in (\mathcal{A}, *, \varphi)$ be tensor independent standard Gaussian variables, that is, a and b are self-adjoint, and their distribution is given by the moments of the standard Gaussian density $f(x) = \frac{1}{\sqrt{2\pi}}\exp\big(-x^2/2\big)$ ($x \in \mathbb{R}$). Then, $\frac{1}{\sqrt{2}}(a+b)$ is also a standard Gaussian variable.*

Proof. We apply the above algorithm. Let f be the standard Gaussian density on $\mathbb{R}$. Then,

$$\widehat{f}(t) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}tx - x^2/2}\ \mathrm{d}x = \mathrm{e}^{-t^2/2},$$

and thus $\widehat{f_a}(t) = \widehat{f_b}(t) = \mathrm{e}^{-t^2/2}$. Therefore, $\widehat{f_{a+b}}(t) = \widehat{f_a}(t)\widehat{f_b}(t) = \mathrm{e}^{-t^2}$, from which we obtain that

$$f_{a+b}(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}tx - t^2}\ \mathrm{d}t = \frac{\exp(-x^2/4)}{2\sqrt{\pi}},$$

which is the density of a Gaussian variable of mean 0 and variance 2. Consequently, $\frac{1}{\sqrt{2}}(a+b)$ has a standard Gaussian distribution, as desired. □

4.3.2. R-Transform Algorithm for Free Independence. If one analyses the Fourier transform algorithm, it seems that the essential ingredients are as follows:

(1) A method to compute the density of $a + b$ (with a and b tensor independent) from a power series that depends entirely on the moments of $a+b$, that is, the inverse Fourier transform, which can be computed using the moment generating function $M_{a+b}(z)$; and

(2) a method to compute the the moment generating function M_{a+b} of $a + b$ from the moment generating functions M_a and M_b of a and b respectively, namely, the relation $M_{a+b}(z) = \exp\big(C_a(z) + C_b(z)\big) = M_a(z)M_b(z)$.

Thus, a viable strategy to design an analog of the Fourier transform algorithm in the free/noncommutative setting seems to be to proceed as follows:

(1) Choose a power series $G_a(z)$ whose definition depends entirely on the moments of a random variable a and from which the density of the analytic distribution μ_a of a can be recovered; and

(2) find a way to compute $G_{a+b}(z)$ from G_a and G_b when a and b are free from each other.

The algorithm that we will cover in this section, which we call the **R-transform algorithm**, was first proposed by D. Voiculescu (see **[Vo86]**).

Let $a \in (\mathcal{A}, *, \varphi)$ be a self-adjoint noncommutative random variable with an analytic distribution μ_a given by a density f_a, and consider the power series

$$G_a(z) = \frac{1}{z} + \sum_{n=1}^{\infty} \frac{\varphi(a^n)}{z^{n+1}}.$$

If $|z| > \rho(a)$ (recall that ρ denotes the spectral radius of an operator a), then $-G_a(z)$ coincides with the Stieltjes transform $s_{\mu_a}(z)$ of the measure μ_a, from which we have already seen that f_a can be recovered (See Theorem B.10 and Example 2.34). Thus, the power series G_a appears to be a good candidate for the foundation of the R-transform algorithm. We now attempt to find a method of computing G_{a+b} from G_a and G_b when a and b are free from each other.

Proposition 4.33. *Let $a \in (\mathcal{A}, *, \varphi)$ be a self-adjoint random variable, and define the power series*

$$F_a(z) = z^{-1} G_a(z^{-1}) = 1 + \sum_{n=1}^{\infty} \varphi(a^n) z^n$$

(or, equivalently, $G_a(z) = z^{-1} F_a(z^{-1})$) and

$$K_a(z) = 1 + \sum_{n=1}^{\infty} \kappa_n[a] z^n.$$

Then, $K_a\big(zF_a(z)\big) = F_a(z)$ and $F_a(z/K_a(z)) = K_a(z)$ for all $z \in \mathbb{C}$.

Proof. For every $n \in \mathbb{N}$, define the functions $f_n, g_n : NC(n) \to \mathbb{C}$ as $f_n(\pi) = \varphi_\pi[a, \ldots, a]$ and $g_n(\pi) = \kappa_\pi[a, \ldots, a]$. According to Equations (4.5) and (4.6), it follows that for every $n \in \mathbb{N}$, one has

$$\varphi_\pi[a, \ldots, a] = \sum_{\sigma \in NC(n):\ \sigma \leqslant \pi} \kappa_\sigma[a, \ldots, a] \text{ and }, \qquad \pi \in NC(n)$$

and

$$\kappa_\pi[a, \ldots, a] = \sum_{\sigma \in NC(n):\ \sigma \leqslant \pi} \mu(\sigma, \pi) \varphi_\sigma[a, \ldots, a], \qquad \pi \in NC(n)$$

According to the definition of φ_π, it follows that each f_n is multiplicative, and Proposition 4.24 implies that each g_n is multiplicative (see Definition D.18 for the definition of a multiplicative function). Therefore, the present result is a direct application of Theorem D.19, since the coefficients of F_a are given by $f_n(1_n)$, and the coefficients of K_a are given by $g_n(1_n)$. □

Definition 4.34. Let $a \in (\mathcal{A}, *, \varphi)$ be a self-adjoint random variable. Then, the **R-transform** of a, denoted $\mathcal{R}_a$, is defined as the following power series:

$$\mathcal{R}_a(z) = \sum_{n=1}^{\infty} \kappa_n[a] z^{n-1} = \frac{1}{z} K_a(z) - \frac{1}{z}.$$

Remark 4.35. Given the definition of the R-transform and that mixed cumulants in free variables vanish, it follows that if $a, b \in (\mathcal{A}, *, \varphi)$ are free from each other, then $\mathcal{R}_{a+b} = \mathcal{R}_a + \mathcal{R}_b$.

Proposition 4.36. *Let $a \in (\mathcal{A}, *, \varphi)$ be a self-adjoint random variable. Then, the R-transform of a and the power series G_a are related to each other as follows:*

$$\frac{1}{G_a(z)} + \mathcal{R}_a\big(G_a(z)\big) = z \tag{4.11}$$

and

$$G_a\big(\mathcal{R}_a(z) + 1/z\big) = z \tag{4.12}$$

Proof. According to Proposition 4.33, it follows that

$$\begin{aligned}
\frac{1}{G_a(z)} + \mathcal{R}_a\big(G_a(z)\big) &= \frac{1}{G_a(z)} + \frac{1}{G_a(z)} K_a\big(G_a(z)\big) - \frac{1}{G_a(z)} \\
&= \frac{z}{F_a(z^{-1})} K_a\big(z^{-1} F_a(z^{-1})\big) \\
&= \frac{z}{F_a(z^{-1})} F_a(z^{-1}) \\
&= z,
\end{aligned}$$

as desired. Furthermore,

$$G_a\big(\mathcal{R}_a(z) + 1/z\big) = G_a\big(K_a(z)/z\big) = \frac{z}{K_a(z)} F_a(z/K_a(z)) = \frac{z}{K_a(z)} K_a(z) = z,$$

concluding the proof of the proposition. □

A consequence of the previous proposition is that, given two self-adjoint variables $a, b \in (\mathcal{A}, *, \varphi)$ that are $*$-free from each other, whenever G_a and G_b are such that $G_a(w + 1/z) = z$ and $G_b(w + 1/z) = z$ can easily be solved for $w \in \mathbb{C}$, and $\mathcal{R}_{a+b} = \mathcal{R}_a + \mathcal{R}_b$ is such that $1/w + \mathcal{R}_{a+b}(w) = z$ can easily be solved for $w \in \mathbb{C}$, then $\mathcal{R}_a$ and $\mathcal{R}_b$ can be computed from G_a and G_b, and then G_{a+b} can be computed from $\mathcal{R}_{a+b}$. This observation, combined with the fact that the density of $a + b$ can, in principle, be obtained from G_{a+b}, seems to indicate that we now have two *similar ingredients* to those mentioned in the beginning of this subsection (the two essential ingredients to the Fourier transform algorithm), namely,

(1) A method to compute the density of $a + b$ (with a and b tensor independent) from a power series that depends entirely on the moments of $a + b$, namely, the Stieltjes Inversion Formula, which can be computed using the series $G_{a+b}(z)$; and

(2) a method to compute G_{a+b} from G_a and G_b namely, computing the R-transforms $\mathcal{R}_a$ and $\mathcal{R}_b$ from G_a and G_b (the latter being completely determined by the individual distributions of a and b respectively) using (4.12), and then computing G_{a+b} from $\mathcal{R}_{a+b} = \mathcal{R}_a + \mathcal{R}_b$ using (4.11).

We can now present Voiculescu's R-transform algorithm: Let $a, b \in (\mathcal{A}, *, \varphi)$ be self-adjoint variables with respective analytic distributions μ_a and μ_b, and suppose that μ_a and μ_b have respective densities f_a and f_b, and that a and b are free from each other. Then, we can obtain the density f_{a+b} of the analytic distribution μ_{a+b} of $a+b$ as follows:

Step 1. Compute the Stieltjes transforms of μ_a and μ_b to obtain closed form expressions for the power series G_a and G_b.
Step 2. Solve for $w_a, w_b \in \mathbb{C}$ in the expressions $G_a(w_a + 1/z) = z$ and $G_b(w_b + 1/z) = z$, and set $\mathcal{R}_a(z) = w_a$ and $\mathcal{R}_b(z) = w_b$.
Step 3. Since a and b are free from each other, compute $\mathcal{R}_{a+b} = \mathcal{R}_a + \mathcal{R}_b$.
Step 4. Solve for $w \in \mathbb{C}$ in the expression $1/w + \mathcal{R}_{a+b}(w) = z$ and set $G_{a+b}(z) = w$.
Step 5. Using the Stieltjes inversion formula, recover the density f_{a+b} from G_{a+b} as:

$$f_{a+b}(x) = \frac{-1}{\pi} \lim_{\varepsilon \to 0^+} \operatorname{Im}\big(G_{a+b}(x + \mathrm{i}\varepsilon)\big)$$

whenever the above limit exists and $f_{a+b}(x) = 0$ otherwise.

We now see examples of the R-transform algorithm being applied to compute the density of a sum of two free self-adjoint variables.

Proposition 4.37. *Let $a, b \in (\mathcal{A}, *, \varphi)$ be free semicircular variables, that is, a and b are self-adjoint, and their distribution is given by the moments of the semicircular distribution $f(x) = \frac{1}{2\pi}\sqrt{4 - x^2}$ (for $x \in [-2, 2]$), that is,*

$$\varphi(a^n) = \varphi(b^n) = \begin{cases} \frac{1}{n/2+1}\binom{n}{n/2} & \text{if } n \text{ is even; and} \\ 0 & \text{if } n \text{ is odd.} \end{cases}$$

Then, $\frac{1}{\sqrt{2}}(a + b)$ is also a semicircular variable.

Proof. We apply the R-transform algorithm. For every $n \in \mathbb{N}$, we notice that $\varphi(a^{2n}) = \varphi(b^{2n}) = C_n$, where C_n is the well-known Catalan number of order n (see Section D.3). Knowing that whenever $|z|$ is outside the support of a real measure μ, then the Stieltjes transform at that point is given by the power series

$$s_\mu(z) = -\sum_{n=0}^{\infty} \frac{1}{z^{n+1}} \int_{\mathbb{R}} x^n \, \mathrm{d}\mu(x)$$

(see Proposition B.9), it follows that

$$G_a(z) = G_b(z) = G(z) = \sum_{n=0}^{\infty} \frac{C_n}{z^{2n+1}}.$$

Using the Cauchy product formula for the product of two series (as it was done in Example 2.34), we then obtain that

$$G(z)^2 = \sum_{n=0}^{\infty} \sum_{m=0}^{n} \frac{C_m}{z^{2m+1}} \frac{C_{n-m}}{z^{2(n-m)+1}} = \sum_{n=0}^{\infty} \frac{1}{z^{2(n+1)}} \sum_{m=0}^{n} C_m C_{n-m}.$$

By substituting $p = n + 1$ and then $q = m + 1$ in the above, we see that

$$G(z)^2 = \sum_{p=1}^{\infty} \frac{1}{z^{2p}} \sum_{m=0}^{p-1} C_m C_{p-(m+1)} = \sum_{p=1}^{\infty} \frac{1}{z^{2p}} \sum_{q=1}^{p} C_{q-1} C_{p-q}.$$

Then, an application of Segner's recursive formula for the Catalan numbers (see section D.3) yields

$$G(z)^2 = \sum_{p=1}^{\infty} \frac{C_p}{z^{2p}} = \sum_{p=0}^{\infty} \frac{C_p}{z^{2p}} - 1 = z \sum_{p=0}^{\infty} \frac{C_p}{z^{2p+1}} - 1 = zG(z) - 1.$$

Therefore, we see that $G(z)$ satisfies $G(z)^2 - zG(z) + 1 = 0$, which implies that

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

Now,

$$G(w + 1/z) = \frac{w + 1/z \pm \sqrt{(w + 1/z)^2 - 4}}{2} = z,$$

which implies that

$$(w + 1/z)^2 - 4 = (2z - w - 1/z)^2,$$

and hence

$$w^2 + \frac{2w}{z} + \frac{1}{z^2} - 4 = w^2 - 4wz + \frac{2w}{z} + 4z^2 + \frac{1}{z^2} - 4.$$

From this, we obtain that

$$0 = -4wz + 4z^2 = 4z(z - w),$$

and thus $w = \mathcal{R}_a(z) = \mathcal{R}_b(z) = z$.

Given that free cumulants are multiplicative, it follows that $\mathcal{R}_{\frac{a}{\sqrt{2}}}(z) = \frac{1}{2}\mathcal{R}_a(z)$ and likewise for $\mathcal{R}_{\frac{b}{\sqrt{2}}}$, and thus

$$\mathcal{R}_{\frac{a+b}{\sqrt{2}}}(z) = \frac{z}{2} + \frac{z}{2} = z.$$

It then immediately follows that $G_{\frac{a+b}{\sqrt{2}}}(z)$ is the same as $G_a(z)$ and $G_b(z)$, from which we conclude the result. □

Proposition 4.38 ([**NS06**] Exercise 12.9)**.** *Let $a, b \in \mathcal{A}$ be self-ajoint and free from each other. Suppose that the distribution of a is given by a density f_a, and that the distribution of b is given by the dirac mass δ_α at some point $\alpha \in \mathbb{R}$. Then, $f_{a+b}(x) = f_a(x - \alpha)$, that is, the distribution of the sum $a + b$ is the distribution of a shifted to the right by an amount of α.*

Proof. We apply the R-transform algorithm. For every $z \in \mathbb{C}$, we have that

$$G_b(z) = -s_{\delta_\alpha}(z) = \int_{\mathbb{R}} \frac{1}{z - x} \, \mathrm{d}\delta_\alpha(x) = \frac{1}{z - \alpha}.$$

To obtain $\mathcal{R}_b$, we solve for $w_b \in \mathbb{C}$ in

$$z = G_b(w_b + 1/z) = \frac{1}{w_b + 1/z - \alpha},$$

which yields

$$\mathcal{R}_b(z) = w_b = \frac{1}{z} - \frac{1}{z} + \alpha = \alpha.$$

Therefore, $\mathcal{R}_{a+b}(z) = \mathcal{R}_a(z) + \alpha$. To find G_{a+b}, we solve for $w \in \mathbb{C}$ in

$$\frac{1}{w} + \mathcal{R}_{a+b}(w) = \frac{1}{w} + \mathcal{R}_a(w) + \alpha = z,$$

or

$$\frac{1}{w} + \mathcal{R}_a(w) = z - \alpha = z',$$

where $z' = z - \alpha$. It is then clear that $G_{a+b}(z) = w = G_a(z') = G_a(z - \alpha)$, which implies that

$$f_{a+b}(x) = \frac{-1}{\pi} \lim_{\varepsilon \to 0^+} \operatorname{Im}\big(G_a(x + \mathrm{i}\varepsilon - \alpha)\big) = f_a(x - \alpha),$$

as desired. □

Remark 4.39. Using similar arguments as in the above proposition, one could prove that, for any probability measure μ on $\big(\mathbb{R}, \mathscr{B}(\mathbb{R})\big)$ and $\alpha \in \mathbb{R}$, then $\nu = \mu \boxplus \delta_\alpha$ is the measure such that $\nu(A) = \mu(A - \alpha)$ for each $A \in \mathscr{B}(\mathbb{R})$.

Part III. Applications to Random Matrix Theory

Chapter 5

Random Matrix Models of Freeness

For more detailed accounts of the concepts covered in this chapter, see [**Gu14**], [**AGZ09**] Sections 5.4 and 5.5, [**NS06**] Lectures 22 and 23, and [**Sp09**].

5.1. Introduction

The usefulness of free probability in the study of random matrices is mainly due to the following fact: Let $(\Omega, \mathscr{A}, P)$ be a probability space, let $L^{\infty-}$ be the space of $\mathbb{C}$-valued random variables with finite moments of all orders on $(\Omega, \mathscr{A}, P)$, and for every $n \in \mathbb{N}$, let $\big(\mathcal{M}_n(\mathbb{C}) \otimes L^{\infty-}, *\big)$ be the $*$-algebra of random matrices whose entries are random variables in $L^{\infty-}$. For every $n \in \mathbb{N}$, let $\big(X_i^{(n)} : i \in I\big) \subset \mathcal{M}_n(\mathbb{C}) \otimes L^{\infty-}$ be a collection of random matrices that are independent from each other (in the classical measure-theoretic sense). Suppose that for every $i \in I$, the sequence $Y_i^{(n)} := C_i(n) X_i^{(n)}$ converges in $*$-distribution (with respect to $\frac{1}{n}\mathrm{tr} \otimes \mathbf{E}$, or with respect to $\frac{1}{n}\mathrm{tr}$ almost surely) to a limit variable x_i in a $*$-probability space $(\mathcal{A}, *, \varphi)$ as $n \to \infty$, where $C_i(n)$ is some normalization sequence. Then, in many cases, the $Y_i^{(n)}$ are **asymptotically $*$-free** from each other, meaning that the limiting variables $(x_i : i \in I)$ are $*$-free from each other in the limit $*$-probability space $(\mathcal{A}, *, \varphi)$.

This implies in particular that if one can compute the distribution of a $*$-polynomial $P(x_i : i \in I)$ evaluated in the x_i from the individual distributions of the x_i (that is, answer the general free probability problem introduced in Chapter 4), then it is possible to obtain the limiting distribution of the polynomial $P\big(Y_i^{(n)} : i \in I\big)$ from the limiting distributions of the individual $Y_i^{(n)}$. Given that this asymptotic $*$-freeness phenomenon occurs with respect to the functionals $\frac{1}{n}\mathrm{tr} \otimes \mathbf{E}$ and $\frac{1}{n}\mathrm{tr}$ (almost surely) and that these functionals provide the moments of the eigenvalue distribution of a given matrix (see Examples 2.24 and 2.25), we then see that free probability provides a way of answering the following general random matrix theory problem:

Problem 5.1. Let $\big(X_i^{(n)} : i \in I\big) \subset \mathcal{M}_n(\mathbb{C}) \otimes L^{\infty-}$ be a collection of random matrices that are independent from each other and such that for every $i \in I$, the empirical eigenvalue distribution

of $Y_i^{(n)} \coloneqq C_i(n)X_i^{(n)}$ converges (in expectation or almost surely) to a limiting distribution μ_i. Given an arbitrary $*$-polynomial $P \in \langle (x_i, x_i^* : i \in I) \rangle$, devise an algorithm that allows one to compute the limiting eigenvalue distribution (in expectation or almost surely) of $P\Big(\big(Y_i^{(n)} : i \in I\big)\Big)$ from the individual μ_i.

In this context, it seems clear that a question of interest in the theory of random matrices is to understand when asymptotic freeness occurs.

5.2. Asymptotic Freeness of Random Matrices

Notation 5.2. To avoid confusion with the multiple indices, given a collection of $n \times n$ matrices $\big(X_i^{(n)} : i \in I\big)$, we will use $X_i^{(n)}(j,k)$ to denote the entry on row j and column k (where $1 \leqslant j,k \leqslant n$) of the matrix $X_i^{(n)}$.

The following theorem establishes the asymptotic freeness of a very general class of random matrices.

Theorem 5.3 ([**Gu14**] Section 2)**.** *For every $n \in \mathbb{N}$, define the families of random matrices $\mathbf{X}_n = \big(X_i^{(n)} : i \in I\big)$, $\mathbf{U}_n = \big(U_j^{(n)} : j \in J\big)$, and $\mathbf{D}_n = \big(D_k^{(n)} : k \in K\big)$ in $\mathcal{M}_n(\mathbb{C}) \otimes L^{\infty-}$ as follows:*

(1) For every $n \in \mathbb{N}$, the $X_i^{(n)}$ are independent $n \times n$ Wigner matrices (see Example 2.39) such that for every $m \in \mathbb{N}$, one has

$$\sup_{n\in\mathbb{N}} \sup_{i\in I} \sup_{1\leqslant j,k\leqslant n} \mathbf{E}\Big[\big|X_i^{(n)}(j,k)\big|^m\Big] < \infty;$$

(2) for every $n \in \mathbb{N}$, the $U_j^{(n)}$ are i.i.d. $n \times n$ Haar unitary random matrices (see Example 2.30);

(3) the $D_k^{(n)}$ are self-adjoint deterministic (i.e., elements of $\mathcal{M}_n(\mathbb{C}) \otimes \mathbf{1}_{L^{\infty-}}$) and such that $\mathbf{D}_n$ converges in joint $*$-distribution to a limit family $\mathbf{d} = (d_k : k \in K)$, and

$$\sup_{m\in\mathbb{N}} \max_{k\in K} \sup_{n\in\mathbb{N}} \frac{1}{n}\mathrm{tr}\Big(\big(D_k^{(n)}\big)^m\Big)^{1/m} < \infty;$$

for all $m \in \mathbb{N}$; and

(4) for every $n \in \mathbb{N}$, the matrices in $\mathbf{X}_n$ are independent (in the measure-theoretic sense) of the matrices in $\mathbf{U}_n$.

Then, the family $(\mathbf{X}_n, \mathbf{U}_n, \mathbf{D}_n)$ converges in joint $*$-distribution (with respect to $\frac{1}{n}\mathrm{tr}$ almost surely, and with respect to $\frac{1}{n}\mathrm{tr} \otimes \mathbf{E}$) to a limit family $(\mathbf{x}, \mathbf{u}, \mathbf{d})$ in a limit $*$-probability space $(\mathcal{A}, *, \varphi)$, where

(1) $\mathbf{x} = (x_i : i \in I)$ contains $*$-free semicircular variables;

(2) $\mathbf{u} = (u_j : j \in J)$ contains $*$-free Haar unitary variables; and

(3) $\mathbf{x}$, $\mathbf{u}$, and $\mathbf{d}$ are $*$-free from each other.

Proposition 5.4. *Let $(\mathcal{A}, *, \varphi)$ be a $*$-probability space, let $\mathbf{a} = (a_i : i \in I) \subset \mathcal{A}$ be a collection of variables, and let $u \in \mathcal{A}$ be a Haar unitary variable. If the collection $\mathbf{a}$ is $*$-free from u. Then, the collection $\mathbf{a}$ is $*$-free from the collection $u\mathbf{a}u^* = (ua_iu^* : i \in I)$.*

Proof. For every $i \in I$, let us denote $b_i = ua_iu^*$. Then, we need to prove that the b_i are $*$-free from the a_i.

Let $i \in I$ be arbitrary and $M \in \mathbb{C}\langle x, x^*\rangle$ be a nontrivial $*$-word. According to the definition of $*$-words, there exists $t \in \mathbb{N}$ and exponents $n(1), \ldots, n(t) \in \{1, *\}$ such that $M(x) = x^{n(1)}x^{n(2)} \cdots x^{n(t)}$. Consequently,

$$M(b_i) = (ua_iu^*)^{n(1)}(ua_iu^*)^{n(2)} \cdots (ua_iu^*)^{n(t)} = ua_i^{n(1)}u^*ua_i^{n(2)}u^* \cdots ua_i^{n(t)}u^* = uM(a_i)u^*$$

(indeed, $(uau^*)^* = ua^*u^*$ for all $a \in \mathcal{A}$). Moreover, since a_i is $*$-free from u and $uu^* = 1_{\mathcal{A}}$ (as u is unitary), it follows from Example 3.21 (2) that

$$\begin{aligned} M(b_i)^\circ &= \big(uM(a_i)u^*\big)^\circ = uM(a_i)u^* - \varphi\big(uM(a_i)u^*\big) \cdot 1_{\mathcal{A}} \\ &= uM(a_i)u^* - \varphi(uu^*)\varphi\big(M(a_i)\big) \cdot 1_{\mathcal{A}} \\ &= uM(a_i)u^* - \varphi\big(M(a_i)\big) \cdot 1_{\mathcal{A}} \\ &= uM(a_i)u^* - \varphi\big(M(a_i)\big) \cdot uu^* \\ &= uM(a_i)u^* - u\Big(\varphi\big(M(a_i)\big) \cdot 1_{\mathcal{A}}\Big)u^* \\ &= uM(a_i)^\circ u^*. \end{aligned}$$

Let $t \in \mathbb{N}$ be fixed, and let $M_1, N_1 \ldots, M_t, N_t \in \mathbb{C}\langle x, x^*\rangle$ be $*$-words such that only M_1 and N_t are possibly trivial (i.e., equal to x^0). Then, for all $i(1), \ldots, i(t), j(1), \ldots, j(t) \in I$,

$$\begin{aligned} &\varphi\Big(M_1(a_{i(1)})^\circ N_1(b_{j(1)})^\circ M_2(a_{i(2)})^\circ N_2(b_{j(2)})^\circ \cdots M_t(a_{i(t)})^\circ N_t(b_{j(t)})^\circ\Big) \\ =&\varphi\Big(M_1(a_{i(1)})^\circ uN_1(a_{j(1)})^\circ u^*M_2(a_{i(2)})^\circ uN_2(a_{j(2)})^\circ u^* \cdots M_t(a_{i(t)})^\circ uN_t(a_{j(t)})^\circ u^*\Big). \end{aligned}$$

Since $\varphi(u) = \varphi(u^*) = 0$, and $\mathbf{a}$ and u are $*$-free from each other, it then follows that

$$\varphi\Big(M_1(a_{i(1)})^\circ N_1(b_{j(1)})^\circ M_2(a_{i(2)})^\circ N_2(b_{j(2)})^\circ \cdots M_t(a_{i(t)})^\circ N_t(b_{j(t)})^\circ\Big) = 0,$$

which implies that $\mathbf{a}$ and $u\mathbf{a}u^*$ are $*$-free (see Proposition 6.6). □

From the two previous results, one easily obtains the following corollary:

Corollary 5.5. *For every $n \in \mathbb{N}$, let U_n be a Haar unitary random matrix, and let $\mathbf{A}_n = (A_i^{(n)} : i \in I)$ be a sequence of self-adjoint deterministic matrices such that $\mathbf{A}_n$ converges in joint $*$-distribution to a limit family $\mathbf{a} = (a_i : i \in I)$, and for all $m \in \mathbb{N}$,*

$$\sup_{m\in\mathbb{N}} \max_{i\in I} \sup_{n\in\mathbb{N}} \frac{1}{n}\mathrm{tr}\Big(\big(A_i^{(n)}\big)^m\Big)^{1/m} < \infty.$$

Then, the families $\mathbf{A}_n$ and $U_n\mathbf{A}_nU_n^* = (U_nA_i^{(n)}U_n^* : i \in J)$ are asymptotically $*$-free from each other with respect to $\frac{1}{n}\mathrm{tr} \otimes \mathbf{E}$.

Remark 5.6. Given a collection $(A_i : i \in I)$ of $n \times n$ (random or deterministic) matrices, the eigenvalues of a $*$-polynomial P evaluated in the A_i typically depends on the eigenvectors of the A_i, and there is no guarantee that one can obtain the eigenvalue distribution of $P(A_i : i \in I)$ by simply knowing the eigenvalue distributions of the individual A_i. For example, if A_1 and A_2 have the same eigenvectors $v_1, \ldots, v_n$, then we can expect that the eigenvalues of $A_1 + A_2$ will be $\lambda_1(A_1) + \lambda_1(A_2), \ldots, \lambda_n(A_1) + \lambda_n(A_2)$, where, for each $i \leqslant n$, $\lambda_i(A_1)$ is the eigencalue of A_1 associated to the eigenvector v_i, and likewise for $\lambda_i(A_2)$. However, if no such *simple* relation exist between the eigenvectors of A_1 and A_2, then the relationship between the eigenvalues of $A_1 + A_2$, A_1, and A_2 becomes much more complicated (see for instance [**Ta12**] Section 1.3).

In this context, it appears that Corollary 5.5 can be intuitively interpreted as follows: Let A_1 and A_2 be $n \times n$ deterministic matrices, where n is *very large*. If it is assumed that the eigenvectors of A_2 are in a *generic position* with respect to the eigenvectors of A_1, that is, we uniformly rotate the eigenvectors of A_2 at random by conjugating A_2 with a $n \times n$ Haar unitary random matrix ($A_2 \hookrightarrow UA_2U^*$), then the averaged eigenvalue distribution of any $*$-polynomial $P(A_1, A_2)$ evaluated in A_1 and A_2 can, in principle, be approximated using our knowledge of the eigenvalue distributions of A_1 and A_2.

5.3. Strong Asymptotic Freeness of Random Matrices

Definition 5.7. For every $n \in \mathbb{N}$, let $\mathbf{a}_n = (a_i^{(n)} : i \in I)$ be a collection of variables in a C^*-probability space $(\mathcal{A}_n, *, \|\cdot\|, \varphi_n)$; and let $\mathbf{a} = (a_i : i \in I) \in (\mathcal{A}, *, \|\cdot\|, \varphi)$. Suppose that $\mathbf{a}_n$ converges in joint $*$-distribution to $\mathbf{a}$. If it is also the case that

$$\lim_{n\to\infty} \|P(a_i^{(n)} : i \in I)\| = \|P(a_i : i \in I)\|$$

for every $*$-polynomial $P \in \mathbb{C}\langle (x_i, x_i^* : i \in I)\rangle$, then $\mathbf{a}$ is the **strong limit in distribution** of the sequence $\mathbf{a}_n$.

The potential usefulness of strong limits in distribution for the study of the spectrum of large random matrices comes from the following proposition:

Proposition 5.8 ([**CM14**] Proposition 2.1)**.** *For every $n \in \mathbb{N}$, let $\mathbf{a}_n = (a_i^{(n)} : i \in I)$ be a collection of self-adjoint variables in a C^*-probability space $(\mathcal{A}_n, *, \|\cdot\|, \varphi_n)$, and let $\mathbf{a} = (a_i : i \in I) \subset (\mathcal{A}, *, \|\cdot\|, \varphi)$ be the strong limit in distribution of the $\mathbf{a}_n$ and such that the a_i are self-adjoint. Let $P \in \mathbb{C}\langle (x_i, x_i^* : i \in I)\rangle$ be a $*$-polynomial P such that $P(a_i^{(n)} : i \in I)$ and $P(a_i : i \in I)$ are self-adjoint; for each $n \in \mathbb{N}$, let μ_n be the analytic distribution of $P(a_i^{(n)} : n \in \mathbb{N})$ (this real probability measure is guaranteed to exist thanks to Theorem 2.32); and let μ be the analytic distribution of $P(a_i : i \in I)$. For every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ large enough such that if $n \geqslant N$, then*

$$\operatorname{supp}(\mu_n) \subset \operatorname{supp}(\mu) + (-\varepsilon, \varepsilon),$$

where, given a set $A \subset \mathbb{R}$,

$$A + (-\varepsilon, \varepsilon) = A \cup \left(\bigcup_{t\in(-\varepsilon,\varepsilon)} A + t \right).$$

Definition 5.9. For every $n \in \mathbb{N}$, let $\mathbf{a}_n = (a_i^{(n)} : i \in I)$ be a collection of self-adjoint variables in a C^*-probability space $(\mathcal{A}_n, *, \|\cdot\|, \varphi_n)$; and let $\mathbf{a} = (a_i : i \in I) \in (\mathcal{A}, *, \|\cdot\|, \varphi)$ be the strong limit in distribution of the $\mathbf{a}_n$. Suppose that $\mathbf{a}_n$ converges in joint $*$-distribution to $\mathbf{a}$. If the variables in $\mathbf{a}$ are also $*$-free from each other, then the $\mathbf{a}_n$ are said to be **strongly asymptotically $*$-free** from each other.

Let $\left(X_i^{(n)} : i \in I\right) \subset \mathcal{M}_n(\mathbb{C}) \otimes L^{\infty-}$ be a collection of self-adjoint random matrices such that, for every $\omega \in \Omega$, $\mathbf{X}_n = \left(X_i^{(n)}(\omega) : i \in I\right)$ are strongly asymptotically $*$-free and converge towards the strong limit $\mathbf{x} = (x_i : i \in I)$ (which does not depend on ω) in some C^*-probability space. Then, an immediate consequence of the previous proposition is that, given a $*$-polynomial P such that $P(X_i^{(n)} : i \in I)$ and $P(x_i : i \in I)$ are self-adjoint, then not only does the empirical eigenvalue distribution of $P(X_i^{(n)} : i \in I)$ converges weakly to the analytic distribution μ of $P(x_i : i \in I)$ (which is completely determined by the distribution of the individual x_i), but the extremal eigenvalues of $P(X_i^{(n)} : i \in I)$ almost-surely converge inside of the support of μ. Thus, strong asymptotic freeness provides additional information on the eigenvalue behaviour at the edge of the limiting spectrum for polynomials of large random matrices. In similar fasion to Theorem 5.10, the following result establishes the occurence of strong asymptotic freeness for several types of random matrices.

Theorem 5.10 ([**Gu14**] Section 2.1)**.** *For every $n \in \mathbb{N}$, define the families of random matrices $\mathbf{X}_n = \left(X_i^{(n)} : i \in I\right)$, $\mathbf{U}_n = \left(U_j^{(n)} : j \in J\right)$, and $\mathbf{D}_n = \left(D_k^{(n)} : k \in K\right)$ in $\mathcal{M}_n(\mathbb{C}) \otimes L^{\infty-}$ as follows:*

(1) For every $n \in \mathbb{N}$, the $X_i^{(n)}$ are i.i.d. $n \times n$ Gaussian Wigner matrices, that is, for every $n \in \mathbb{N}$ and $i \in I$, the diagonal entries of $X_i^{(n)}$ are real standard Gaussian variables, and the above diagonal entries of $X_i^{(n)}$ are complex standard gaussian variables;

(2) for every $n \in \mathbb{N}$, the $U_j^{(n)}$ are i.i.d. $n \times n$ Haar unitary random matrices;

(3) the D_k^n are self-adjoint, deterministic, and such that $\mathbf{D}_n$ converges in joint $*$-distribution to a limit family $\mathbf{d} = (d_k : k \in K)$, and there exists a constant $C < \infty$ such that

$$\max_{k \in K} \sup_{n \in \mathbb{N}} \frac{1}{n} \mathrm{tr}\left(\left(D_k^{(n)}\right)^{2m}\right) < C^{2m}$$

for every $m \in \mathbb{N}$; and

(4) for every $n \in \mathbb{N}$, the matrices in $\mathbf{X}_n$ are independent (in the measure-theoretic sense) of the matrices in $\mathbf{U}_n$.

Then, the family $(\mathbf{X}_n, \mathbf{U}_n, \mathbf{D}_n)$ is strongly asymptotically $*$-free, and its strong limit in distribution is $(\mathbf{x}, \mathbf{u}, \mathbf{d})$, where

(1) $\mathbf{x} = (x_i : i \in I)$ contains $*$-free semicircular variables;

(2) $\mathbf{u} = (u_j : j \in J)$ contains $*$-free Haar unitary variables; and

(3) $\mathbf{x}$, $\mathbf{u}$, and $\mathbf{d}$ are $*$-free from each other.

Chapter 6

$*$-freeness in Tensor Products

Unless cited from an outside source, all the work featured in this chapter constitutes original research done in collaboration with/under the supervision of Professor Benoit Collins (thesis supervisor) and Camille Male (CNRS, Université Paris V).

6.1. Introduction

The starting point of the original research project exposed in this chapter is the following question, which was first asked by Gilles Pisier to Benoit Collins:

Question 6.1. For each $n \in \mathbb{N}$, let $\mathbf{U}_n = \left(U_i^{(n)} : i \in I\right)$ be a collection of i.i.d. $n \times n$ Haar unitary random matrices, where I is an arbitrary indexing set (that is, the $U_i^{(n)}$ are i.i.d. random variables distributed according to the normalized Haar measure on the compact group $\mathbb{U}_n$ of $n \times n$ unitary matrices). Consider the collection

$$\mathbf{U}_n \otimes \overline{\mathbf{U}_n} = \left(U_i^{(n)} \otimes \overline{U_i^{(n)}} : i \in I\right).$$

(Note that, given a $n \times n$ matrix A with $\mathbb{C}$-valued entries, $\overline{A}$ denotes the entry-wise complex conjugate of A, not to be confused with the conjugate transpose, which we denote by A^*.) Are the variables in $\mathbf{U}_n \otimes \overline{\mathbf{U}_n}$ asymptotically $*$-free with respect to the expected normalized trace $\mathbf{E}\left[\frac{1}{n}\mathrm{tr} \otimes \frac{1}{n}\mathrm{tr}\right]$ and/or almost surely asymptotically $*$-free with respect to the normalized trace $\frac{1}{n}\mathrm{tr} \otimes \frac{1}{n}\mathrm{tr}$?

In order to answer this question, we first look at a more general situation: Fix an integer $K \in \mathbb{N}$, and for every $1 \leqslant k \leqslant K$, let

$$\mathbf{a}_k^{(n)} = \left(a_{k;i}^{(n)} : i \in I\right) \subset \left(\mathcal{A}_k^{(n)}, *, \varphi_k^{(n)}\right), \qquad n \in \mathbb{N}$$

be a sequence of collections of noncommutative random variables. Suppose further that for each $k \leqslant K$, the collection $\mathbf{a}_k^{(n)}$ converges in joint $*$-distribution towards a limit family

$$\mathbf{a}_k = (a_{k;i} : i \in I) \subset (\mathcal{A}_k, *, \varphi_k)$$

as $n \to \infty$. Then, it can easily be shown (see Proposition 6.8) that

$$\otimes_k \mathbf{a}_k^{(n)} = \Big(\otimes_k a_{k;i}^{(n)} : i \in I \Big) \subset \Big(\otimes_k \mathcal{A}_k^{(n)}, *, \otimes_k \varphi_k^{(n)} \Big)$$

converges in $*$-distribution to

$$\otimes_k \mathbf{a}_k = (\otimes_k a_{k;i} : i \in I) \subset (\otimes_k \mathcal{A}_k, *, \otimes_k \varphi_k)$$

as $n \to \infty$.

Consequently, if one wishes to investigate the asymptotic $*$-freeness of a family of tensor products $\otimes_k \mathbf{a}_k^{(n)}$ whose individual component families $\mathbf{a}_k^{(n)}$ are known to converge to respective limits $\mathbf{a}_k$,[1] one need only study the $*$-freeness relations in the limit family $\otimes_k \mathbf{a}_k$. Thus, Question 6.1 is actually a special case of the following problem in free probability:

Problem 6.2. Given K families

$$\mathbf{a}_k = (a_{k;i} : i \in I) \subset (\mathcal{A}_k, *, \varphi_k), \qquad 1 \leqslant k \leqslant K$$

of noncommutative random variables, characterize the $*$-freeness of the elements in the tensor product collection

$$\otimes_k \mathbf{a}_k = (\otimes_k a_{k;i} : i \in I) \subset (\otimes_k \mathcal{A}_k, *, \otimes_k \varphi_k)$$

in terms of our knowledge of the behaviour of the individual factor collections $\mathbf{a}_k$.

It can be noticed that, in Question 6.1, the families $\mathbf{U}_n$ and $\overline{\mathbf{U}_n}$ are themselves both almost surely asymptotically $*$-free with respect to $\frac{1}{n}\mathrm{tr}$ and asymptotically $*$-free with respect to $\mathbf{E}\big[\frac{1}{n}\mathrm{tr}\big]$ (see Theorem 5.10). Additionally, if one lets $(u_i : i \in I) \subset (\mathcal{A}, *, \varphi)$ denote the family of noncommutative random variables towards which either $\mathbf{U}_n$ or $\overline{\mathbf{U}_n}$ converges in $*$-distribution, it is clear that the linear functional φ is a faithful trace when restricted to the $*$-algebra generated by the u_i.[2] Thus, a special case of Problem 6.2 that is more approachable while still being interesting (as it would solve Question 6.1) is the following:

Problem 6.3. Solve Problem 6.2 with the added assumptions that some $*$-freeness occurs in the factor families $\mathbf{a}_k = (a_{k;i} : i \in I)$, and that the functionals φ_k are faithful.

Remark 6.4. If one of the the collections $\mathbf{a}_k$ contains the zero vector, or if there exists $i \in I$ such that $a_{k;i} = 1_{\mathcal{A}_k}$ for all $k \leqslant K$, then $0_{\otimes_k \mathcal{A}_k}$ or $1_{\otimes_k \mathcal{A}_k}$ will be contained $\otimes_k \mathbf{a}_k$. Given that the zero vector and the unit vector are trivially $*$-free from any element in a $*$-probability space, we will assume that the following conditions hold throughout this chapter without fear that it will result in missing out on any interesting example:

(1) for every $i \in I$ and $k \leqslant K$, $a_{k;i} \neq 0_{\mathcal{A}_k}$; and

[1] That is, answer questions similar to Question 6.1, as it is well-known that $\mathbf{U}_n$ and $\overline{\mathbf{U}_n}$ both converge in joint $*$-distribution (with respect to $\mathbf{E}\big[\frac{1}{n}\mathrm{tr}\big]$ and $\frac{1}{n}\mathrm{tr}$ almost surely) towards families of Haar unitary variables, see Example 2.30.

[2] More precisely, since the u_i are Haar unitary, they are invertible, and thus the $*$-algebra $\mathbb{C}\big\langle (u_i, u_i^* : i \in I) \big\rangle$ generated by them is a group algebra, which in turn implies that φ restricted to $\mathbb{C}\big\langle (u_i, u_i^* : i \in I) \big\rangle$ is in fact the canonical trace since the u_i are $*$-free.

(2) for every $i \in I$, there exists $k \leqslant K$ such that $a_{k;i} \neq 1_{\mathcal{A}_k}$.

The organisation of this chapter is as follows: In Section 6.2, we introduce a few elementary results in free probability which will be useful throughout the chapter. In section 6.3, we provide sufficient conditions for the $*$-freeness in tensor products in the context of Problems 6.2 and 6.3, which, in particular, settle Question 6.1 in the affirmative. In section 6.4, we achieve a complete characterization of the following special case of Problem 6.3 (see Theorem 6.20):

Problem 6.5. Given two families $\mathbf{a}_1 = (a_{1;i} : i \in I) \subset (\mathcal{A}_1, *, \varphi_1)$ and $\mathbf{a}_2 = (a_{2;i} : i \in I) \subset (\mathcal{A}_2, *, \varphi_2)$ of noncommutative random variables such that

(1) the variables in $\mathbf{a}_1$ are $*$-free from each other;

(2) the variables in $\mathbf{a}_2$ are $*$-free from each other; and

(3) φ_1 and φ_2 are faithful,

characterize the $*$-freeness of the elements in the tensor product collection

$$\mathbf{a}_1 \otimes \mathbf{a}_2 = (a_{1;i} \otimes a_{2;i} : i \in I) \subset (\mathcal{A}_1 \otimes \mathcal{A}_2, *, \varphi_1 \otimes \varphi_2)$$

in terms of our knowledge of the behaviour of the individual factor collections $\mathbf{a}_1$ and $\mathbf{a}_2$.

In Section 6.5, we prove that, in certain cases, it is necessary that the variables in a factor collection $\mathbf{a}_1$ be $*$-free from each other in order for the variables in the tensor product collection $\mathbf{a}_1 \otimes \mathbf{a}_2$ to be $*$-free from each other. In doing so, we derive results in the theory of free groups that we believe may be of independent interest (see for instance Proposition 6.41 and Lemma 6.40). In Section 6.6, we explore further questions raised by our results and possible directions for future investigations.

6.2. Preliminary Results in Free Probability

Given that the functionals with which $*$-probability spaces are endowed are linear, one can easily prove the following result, which simplifies the verification that a given collection of noncommutative random variables are $*$-free.

Proposition 6.6. *Let $\mathbf{a} = (a_i : i \in I)$ be a collection of random variables in a $*$-probability space $(\mathcal{A}, *, \varphi)$. The elements in $\mathbf{a}$ are $*$-free from each other if and only if for every $i(1), \ldots, i(t) \in I$ such that $i(1) \neq i(2) \neq \cdots \neq i(t)$ and for every $*$-words $M_1, \ldots, M_t \in \mathbb{C}\langle x, x^* \rangle$, one has*

$$\varphi\big(M_1(a_{i(1)})^\circ \cdots M_t(a_{i(t)})^\circ\big) = 0. \tag{6.1}$$

Proof. If the elements in $\mathbf{a}$ are $*$-free, then (6.1) follows from the definition of $*$-freeness. Suppose now that (6.1) holds. Let $i(1), \ldots, i(t) \in I$ be such that $i(1) \neq i(2) \neq \cdots \neq i(t)$, and let $P_1, \ldots, P_t \in \mathbb{C}\langle x, x^* \rangle$ be $*$-polynomials such that $\varphi\big(P_l(a_{i(l)})\big) = 0$ for each $l \leqslant t$. Since the P_l are $*$-polynomials, one can write

$$P_l(x) = \alpha_{l0} x^0 + \sum_{k=1}^{m_l} \alpha_{lk} M_{lk}(x),$$

where for each $l \leqslant t$, $\alpha_{l0}, \ldots, \alpha_{lt} \in \mathbb{C}$, $m_l \geqslant 1$ is an integer, and $M_{l1}, \ldots, M_{lk}$ are $*$-words. Thus, if one defines the sets

$$K_l \coloneqq \big\{k \leqslant m_l : \varphi\big(M_{lk}(a_{i(l)})\big) = 0\big\}, \qquad\qquad l \leqslant t$$

then the fact that $\varphi\big(P_l(a_{i(l)})\big) = 0$ implies that

$$\alpha_{l0} = -\sum_{k\notin K_l} \varphi\big(M_{lk}(a_{i(l)})\big),$$

hence

$$P_l(a_{i(l)}) = \sum_{k\in K_l} \alpha_{lk} M_{lk}(a_{i(l)}) + \sum_{k\notin K_l} \alpha_{lk}\Big(M_{lk}(a_{i(l)}) - \varphi\big(M_{lk}(a_{i(l)})\big)\cdot 1_A\Big)$$
$$= \sum_{k=1}^{m_l} \alpha_{lk} M_{lk}(a_{i(l)})^\circ.$$

Consequently, it is clear that

$$\varphi\big(P_1(a_{i(1)})\cdots P_t(a_{i(t)})\big) = \sum_{k_1\leqslant m_1,\ldots,k_t\leqslant m_t} \alpha_{1k_1}\cdots\alpha_{tk_t}\varphi\big(M_{1k_1}(a_{i(1)})^\circ\cdots M_{tk_t}(a_{i(t)})^\circ\big) = 0,$$

as desired. □

In this chapter, the usefulness of the free cumulants will mostly come from the following theorem:

Theorem 6.7 ([**NS06**] Theorem 14.4)**.** *Let $(\mathcal{A},*,\varphi)$ be a $*$-probability space. Suppose that the set $\{a_1,\ldots,a_n\}$ is $*$-free from $\{b_1,\ldots,b_n\}$ in $(\mathcal{A},*,\varphi)$. Then, one has*

$$\varphi(a_1b_1a_2b_2\cdots a_nb_n) = \sum_{\pi\in NC(n)} \kappa_\pi[a_1,\ldots,a_n]\varphi_{K(\pi)}[b_1,\ldots,b_n]. \tag{6.2}$$

(Where $K(\pi)$ denotes the Kreweras complement.)

Proposition 6.8. *Fix an integer $K\in\mathbb{N}$, and for every $1\leqslant k\leqslant K$, let*

$$\mathbf{a}_k^{(n)} = \Big(a_{k;i}^{(n)} : i\in I\Big) \subset \Big(\mathcal{A}_k^{(n)},*,\varphi_k^{(n)}\Big), \qquad n\in\mathbb{N}$$

be a sequence of collections of noncommutative random variables. If

$$\mathbf{a}_k^{(n)} \xrightarrow{*\text{-d}} \mathbf{a}_k = (a_{k;i} : i\in I) \subset (\mathcal{A}_k,*,\varphi_k)$$

for every $k\leqslant K$, then

$$\otimes_k \mathbf{a}_k^{(n)} \xrightarrow{*\text{-d}} \otimes_k \mathbf{a}_k = (\otimes_k a_{k;i} : i\in I) \subset (\otimes_k\mathcal{A}_k,*,\otimes_k\varphi_k)$$

(as $n\to\infty$).

Proof. For every $*$-word $M\in\mathbb{C}\big\langle (x_i,x_i^* : i\in I)\big\rangle$, one has

$$\lim_{n\to\infty} \otimes_k\varphi_k^{(n)}\Big(M\Big(\otimes_k a_{k;i}^{(n)} : i\in I\Big)\Big) = \lim_{n\to\infty} \otimes_k\varphi_k^{(n)}\Big(\otimes_k M\Big(a_{k;i}^{(n)} : i\in I\Big)\Big)$$
$$= \lim_{n\to\infty} \prod_{k\leqslant K} \varphi_k^{(n)}\Big(M\Big(a_{k;i}^{(n)} : i\in I\Big)\Big)$$
$$= \prod_{k\leqslant K} \varphi_k\big(M(a_{k;i} : i\in I)\big)$$
$$= \otimes_k\,\varphi_k\big(M(\otimes_k a_{k;i} : i\in I)\big).$$

Being that the φ_k and $\otimes_k\varphi_k$ are linear, the above concludes the proof of the result. □

6.3. Sufficient Conditions

6.3.1. Centering Sufficient Conditions. Suppose we take for granted that at least one of the families $\mathbf{a}_k$ contains variables that are $*$-free from each other. Under this assumption, we obtain the following equivalent condition for the $*$-freeness of the variables in $\otimes_k\mathbf{a}_k$, which simply amounts to a reformulation of the definition of $*$-freeness in the statement of Proposition 6.6 (equation (6.1)):

Proposition 6.9. *Let $(\mathcal{A}_k, *, \varphi_k)$ (where $k \leqslant K$) and the collections $\mathbf{a}_k$ and $\otimes_k\mathbf{a}_k$ be defined as in Problem 6.2, and suppose that there exists $l \leqslant K$ such that the variables in $\mathbf{a}_l$ are $*$-free from each other. Then, the variables in $\otimes_k\mathbf{a}_k$ are $*$-free from each other if and only if for every collection $i(1), \ldots, i(t) \in I$ of indices such that $i(1) \neq i(2) \neq \cdots \neq i(t)$ and $*$-words $M_1, \ldots, M_t \in \mathbb{C}\langle x, x^*\rangle$, there exists a constant $C \in \mathbb{C}$ (which may or may not depend on the choice of indices $i(j)$ or $*$-words M_j) such that*

$$\otimes_k\varphi_k\Big(M_1\big(\otimes_k a_{k;i(1)}\big)^\circ \cdots M_t\big(\otimes_k a_{k;i(t)}\big)^\circ\Big) = \varphi_l\Big(M_1\big(a_{l;i(1)}\big)^\circ \cdots M_t(a_{l;i(t)})^\circ\Big) \cdot C. \tag{6.3}$$

Proof. According to the definition of $*$-freeness (Proposition 6.6), the variables in $\otimes_k\mathbf{a}_k$ are $*$-free from each other if and only if every expression of the form

$$\otimes_k\varphi_k\Big(M_1\big(\otimes_k a_{k;i(1)}\big)^\circ \cdots M_t\big(\otimes_k a_{k;i(t)}\big)^\circ\Big)$$

vanishes. As we have assumed that the variables in $\mathbf{a}_l$ are $*$-free from each other, it follows that

$$\varphi_l\Big(M_1\big(a_{l;i(1)}\big)^\circ \cdots M_t(a_{l;i(t)})^\circ\Big) = 0$$

for every indices $i(1) \neq i(2) \neq \cdots \neq i(t)$ and $*$-words $M_1, \ldots, M_t \in \mathbb{C}\langle x, x^*\rangle$. □

Although the above proposition does not offer much insight into the occurence of $*$-freeness of tensor products, as it is a nearly trivial reformulation of the definition for the $*$-freeness of the elements in $\otimes_k\mathbf{a}_k$ under the assumption that the variables in $\mathbf{a}_l$ are $*$-free, thinking of the $*$-freeness in those terms allows one to easily extract the following corollary, which provides a first sufficient condition for the $*$-freeness that we call the **centering sufficient conditions**:

Definition 6.10 (Centering Sufficient Conditions)**.** Let $(\mathcal{A}_k, *, \varphi_k)$ (where $k \leqslant K$) and the collections $\mathbf{a}_k$ and $\otimes_k\mathbf{a}_k$ be defined as in Problem 6.2. If there exists $l \leqslant K$ such that the variables in $\mathbf{a}_l$ are $*$-free from each other, and for every $i \in I$ and $*$-word $M \in \mathbb{C}\langle x, x^*\rangle$, one has

$$M(\otimes_k a_{k;i})^\circ = \otimes_k N(a_{k;i}), \tag{6.4}$$

where

$$N(a_{k;i}) = \begin{cases} M(a_{k;i})^\circ & \text{if } k = l; \\ M(a_{k;i}) & \text{otherwise,} \end{cases} \tag{6.5}$$

then $\otimes_k\mathbf{a}_k$ is said to satisfy the centering sufficient conditions, and $\mathbf{a}_l$ is called a **dominating family** of the tensor product $\otimes_k\mathbf{a}_k$.

Corollary 6.11. *Let $(\mathcal{A}_k, *, \varphi_k)$ (where $k \leqslant K$) and the collections $\mathbf{a}_k$ and $\otimes_k \mathbf{a}_k$ be defined as in Problem 6.2. If $\otimes_k \mathbf{a}_k$ satisfies the centering sufficient conditions, then the variables in $\otimes_k \mathbf{a}_k$ are $*$-free from each other.*

Proof. Let $l \leqslant K$ be such that the variables in $\mathbf{a}_l$ are $*$-free from each other, and that tensor products of $*$-words in the $\otimes_k \mathbf{a}_k$ factor as in (6.4) (at least one such l is guaranteed to exist if the centering sufficient conditions are satisfied—$\mathbf{a}_l$ is a dominating family). Let $i(1), \ldots, i(t) \in I$ be a collection of indices such that $i(1) \neq i(2) \neq \cdots \neq i(t)$, and let $M_1, \ldots, M_t \in \mathbb{C}\langle x, x^* \rangle$ be arbitrary $*$-words. Then,

$$\begin{aligned}
&\otimes_k \varphi_k\Big(M_1\big(\otimes_k a_{k;i(1)}\big)^\circ \cdots M_t\big(\otimes_k a_{k;i(t)}\big)^\circ\Big) \\
=&\otimes_k \varphi_k\Big(\Big(\otimes_k N_1\big(a_{k;i(1)}\big)\Big) \cdots \Big(\otimes_k N_t\big(a_{k;i(t)}\big)\Big)\Big) \\
=&\varphi_l\Big(M_1\big(a_{l;i(1)}\big)^\circ \cdots M_t\big(a_{l;i(t)}\big)^\circ\Big) \prod_{k \neq l} \varphi_k\Big(M_1\big(a_{k;i(1)}\big) \cdots M_t\big(a_{k;i(t)}\big)\Big),
\end{aligned}$$

and thus (6.3) is satisfied with $C = \prod_{k \neq l} \varphi_k\Big(M_1\big(a_{k;i(1)}\big) \cdots M_t\big(a_{k;i(t)}\big)\Big)$. □

As claimed in the introduction to this chapter, Corollary 6.11 settles Question 6.1 in the affirmative. Indeed, if u and v are Haar unitary variables in respective $*$-probability spaces $(\mathcal{A}_1, *, \varphi_1)$ and $(\mathcal{A}_2, *, \varphi_2)$, then

(1) every nontrivial $*$-word $M \in \mathbb{C}\langle x, x^* \rangle$ evaluated in u, v, or $u \otimes v$ can be reduced to u^n, v^n, or $(u \otimes v)^n$ for some $n \in \mathbb{Z} \setminus \{0\}$; and

(2) for every $n \in \mathbb{Z} \setminus \{0\}$, one has $\varphi_1(u^n) = \varphi_2(v^n) = (\varphi_1 \otimes \varphi_2)\big((u \otimes v)^n\big) = 0$, and hence

$$\big((u \otimes v)^n\big)^\circ = (u \otimes v)^n = u^n \otimes v^n = \big(u^n\big)^\circ \otimes v^n = u^n \otimes \big(v^n\big)^\circ,$$

which implies that the limit family $\mathbf{u} \otimes \mathbf{v} = (u_i \otimes v_i : i \in I)$ towards which $\mathbf{U}_n \otimes \overline{\mathbf{U}_n}$ converges in $*$-distribution satisfies the centering sufficient conditions, and thus contains variables that are $*$-free from each other. (Note that, in this particular case, $\mathbf{u}$ and $\mathbf{v}$ are both dominating families.) In fact, by applying the above argument, one easily obtains the following proposition, of which Question 6.1 is a special case:

Proposition 6.12. *Let $\mathbf{u}^{(n)} = \Big(u_i^{(n)} : i \in I\Big)$ and $\mathbf{v}^{(n)} = \Big(v_i^{(n)} : i \in I\Big)$ be sequences of families of random variables (for example, random matrices) such that*

(1) $\mathbf{u}^{(n)}$ converges in joint $*$-distribution towards a family $\mathbf{u} = (u_i : i \in I)$ of unitary variables that are $*$-free from each other;

(2) $\mathbf{v}^{(n)}$ converges in joint $*$-distribution towards a family $\mathbf{v} = (v_i : i \in I)$ of unitary variables that may or may not be $*$-free from each other;

(3) for every $i \in I$, either u_i is a Haar unitary variable and v_i is an arbitrary unitary variable; or u_i is a p-Haar unitary variable and v_i is a q-Haar unitary variable, where q divides p.

Then, it follows from the centering sufficient conditions that the variables in $\mathbf{u} \otimes \mathbf{v}$ are $*$-free from each other, which implies in particular that the variables in $\mathbf{u}^{(n)} \otimes \mathbf{v}^{(n)}$ are asymptotically $*$-free.

6.3.2. Reformulated Centering Conditions. Although the centering sufficient conditions enable us to answer Question 6.1, it suffers from a few shortcomings. Firstly, the centering sufficient conditions may seem abstruse, in that they arguably fail to provide an intuitive understanding for when the $*$-freeness should occur in $\mathbf{a} \otimes \mathbf{b}$ when it is already present in $\mathbf{a}$. Secondly, although the centering sufficient conditions are easy to apply to unitary variables, it is not immediately obvious if they can be used to settle questions similar to the following (i.e., asymptotic freeness in tensor products when the factor families do not converge towards unitary variables):

Question 6.13. For every $n \in \mathbb{N}$, let $\mathbf{X}_n = \left(X_i^{(n)} : i \in I\right)$ and $\mathbf{Y}_n = \left(Y_i^{(n)} : i \in I\right)$ be collections of i.i.d. $n \times n$ Wigner matrices. Are the variables in the collection

$$\mathbf{X}_n \otimes \mathbf{Y}_n = \left(X_i^{(n)} \otimes Y_i^{(n)} : i \in I\right)$$

asymptotically $*$-free with respect to the expected normalized trace $\mathbf{E}\left[\frac{1}{n}\mathrm{tr} \otimes \frac{1}{n}\mathrm{tr}\right]$ and/or almost surely asymptotically $*$-free with respect to the normalized trace $\frac{1}{n}\mathrm{tr} \otimes \frac{1}{n}\mathrm{tr}$?

Fortunately, it is possible to reformulate the centering sufficient condition in a way that solves these defects. We will call this reformulation the **reformulated centering conditions**.

Definition 6.14 (Reformulated Centering Conditions)**.** Let $(\mathcal{A}_k, *, \varphi_k)$ (where $k \leqslant K$) and the collections $\mathbf{a}_k$ and $\otimes_k \mathbf{a}_k$ be defined as in Problem 6.2. If there exists $l \leqslant K$ such that the variables in $\mathbf{a}_l$ are $*$-free from each other, $i \in I$ and $*$-word $M \in \mathbb{C}\langle x, x^*\rangle$, one has

(1) if $\otimes_k \varphi_k\big(M(\otimes_k a_{k;i})\big) = 0$, then $\varphi_l\big(M(a_{l;i})\big) = 0$; and
(2) if $\otimes_k \varphi_k\big(M(\otimes_k a_{k;i})\big) \neq 0$, then $M(a_{k;i}) = \varphi_k\big(M(a_{k;i})\big) \cdot 1_{\mathcal{A}_k}$ for every $k \neq l$,

then $\otimes_k \mathbf{a}_k$ is said to satisfy the reformulated centering conditions, and $\mathbf{a}_l$ is said to be a **dominating family** of the tensor product $\otimes_k \mathbf{a}_k$.

Proposition 6.15. *The reformulated centering conditions imply the centering sufficient conditions.*

Proof. Suppose that the reformulated centering conditions hold. If $\otimes_k \varphi_k\big(M(\otimes_k a_{k;i})\big) = 0$, then $M(\otimes_k a_{k;i})^\circ = M(\otimes_k a_{k;i})$. Furthermore, since $\varphi_l\big(M(a_{l;i})\big) = 0$ must also be true in this case, it follows that $M(a_{l;i}) = M(a_{l;i})^\circ$, hence

$$M(\otimes a_{k;i})^\circ = M(\otimes_k a_{k;l}) = \otimes_k M(a_{k;i}) = \otimes_k N(a_{k;i}),$$

where N is defined as in equation (6.5). If $\otimes_k \varphi_k\big(M(\otimes_k a_{k;i})\big) \neq 0$, then $M(a_{k;i}) = \varphi_k\big(M(a_{k;i})\big) \cdot 1_{\mathcal{A}_k}$ for all $k \neq l$. Therefore,

$$\begin{aligned} M(\otimes_k a_{k;i})^\circ =& M(\otimes_k a_{k;i}) - \otimes_k \varphi_k\big(M(\otimes_k a_{k;i})\big) \cdot 1_{\otimes_k \mathcal{A}_k} \\ =& \otimes_k M(a_{k;i}) - \otimes_k \Big(\varphi_k\big(M(a_{k;i})\big) \cdot 1_{\mathcal{A}_k}\Big) \\ =& \otimes_k M(a_{k;i}) - \otimes_k \widetilde{N}(a_{k;i}) \\ =& \otimes_k N(a_{k;i}), \end{aligned}$$

where N is defined as in equation (6.5), and $\widetilde{N}$ as follows:

$$\widetilde{N}(a_{k;i}) = \begin{cases} \varphi_k\big(M(a_{k;i})\big) \cdot 1_{\mathcal{A}_k} & \text{if } k = l; \\ M(a_{k;i}) & \text{otherwise.} \end{cases}$$

Thus, the centering sufficient conditions also hold. □

In light of the reformulated centering conditions, it seems that we can think of the sufficient conditions obtained so far in the following intuitive terms: Suppose that we are intially given a collection $\mathbf{a}_1 = (a_{1;i} : i \in I) \subset (\mathcal{A}_1, *, \varphi_1)$ whose constituent variables are $*$-free from each other. Then, suppose that for each $i \in I$, we *modify* a_i by *affixing an extra part onto it* (represented by the variable $\otimes_{k \geqslant 2} a_{k;i}$)

$$\begin{array}{ccc} \mathcal{A}_1 & \to & \otimes_k \mathcal{A}_k \\ a_{1;i} & \hookrightarrow & a_{1;i} \otimes \Big(\otimes_{k \geqslant 2} a_{k;i} \Big), \end{array} \tag{6.6}$$

and that the joint $*$-distribution of the $a_{1;i}$ is modified in a way that is consistent with the definition of $\otimes_k \varphi_k$, that is, for every $*$-polynomial $P \in \mathbb{C}\big\langle (x_i, x_i^* : i \in I) \big\rangle$, one has

$$\varphi_1\big(P(a_{1;i} : i \in I)\big) \hookrightarrow \otimes_k \varphi_k\big(P(\otimes_k a_{k;i} : i \in I)\big).$$

Then, thanks to Proposition 6.15, the reformulated centering conditions provide sufficient conditions for the modified versions of the a_i to retain their $*$-freeness, namely, for every $*$-word M, affixing $\otimes_{k \geqslant 2} a_{k;i}$ to $a_{1;i}$ cannot mean that $\otimes_k \varphi_k\big(M(\otimes_k a_{k;i})\big) = 0$ when it was not already the case that $\varphi_1\big(M(a_{1;i})\big) = 0$, and if $\varphi_1\big(M(a_{1;i})\big) \neq 0$, then the word $M\Big(\otimes_{k \geqslant 2} a_{k;i} \Big)$ is deterministic (i.e., a constant multiple of the unit vector $\otimes_{k \geqslant 2} 1_{\mathcal{A}_k}$). Thus, our sufficient conditions show that if, for every $i \in I$, the $*$-distribution of $a_{1;i}$ is not modified too significantly by joining $\otimes_{k \geqslant 2} a_{k;i}$ onto it as is done in (6.6), then the $*$-freeness of the $a_{1;i}$ is preserved through the modification, hence the $a_{1;i} \otimes \Big(\otimes_{k \geqslant 2} a_{k;i} \Big) = \otimes_k a_{k;i}$ are $*$-free from each other.

6.3.3. Further Questions. In light of the discussion following the statement of the reformulated centering conditions, two questions seem to be especially interesting:

Firstly, it is natural to wonder to what extent the reformulated centering conditions are necessary. Indeed, thanks to Proposition 6.15, we know that if we *modify* the $a_{1;i}$ to no greater extent than that which is allowed by the reformulated centering conditions, then the $*$-freeness of the $a_{1;i}$ is retained in the $\otimes_k a_{k;i}$. Is this also the *full* extent to which the $a_{1;i}$ can be *modified*? Moreover, is it necessary to assume that one family contains $*$-free variables? For example, is it possible to start with two families $\mathbf{a}_1$ and $\mathbf{a}_2$, neither of which contain $*$-free variables, and then join them together to obtain $\mathbf{a}_1 \otimes \mathbf{a}_2$, where the $a_{1;i} \otimes a_{2;i}$ are $*$-free from each other? Answers to these questions will be pursued in Sections 6.4 and 6.5 of this chapter.

Secondly, it is also natural to wonder about how restrictive the reformulated centering conditions are on the distributions of the $a_{k;i}$. If the $a_{1;i}$ are $*$-free and the $a_{k;i}$ (with $k \geqslant 2$) are all chosen such that $a_{k;i} = \varphi_k(a_{k;i}) \cdot 1_{\mathcal{A}_k}$, then the $*$-freeness of the $\otimes_k a_{k;i}$ is trivial. Although Proposition 6.12 provides many examples where the the $*$-freeness of the $\otimes_k a_{k;i}$ holds without the $a_{k;i}$ ($k \geqslant 2$) necessarily being constant multiples of the identity, it is unclear if there are some $*$-distributions on the $a_{1;i}$ for which the $a_{k;i}$ ($k \geqslant 2$) have no other choice than to be a

constant multiple of $1_{\mathcal{A}_k}$ for the reformulated centering conditions to hold. This question will be addressed in the remainder of this subsection.

Definition 6.16. Consider a $*$-word $M(x) = x^{n(1)} \cdots x^{n(t)}$, where $n(1), \ldots, n(t) \in \{0, 1, *\}$. Then, one can define a finite sequence of 1 and -1 associated to M as follows: Let $s_M := s_M(1), \ldots, s_M(t)$, where for each $i \leqslant t$, one has

$$s_M(i) = \begin{cases} 1 & \text{if } n(i) = 1; \\ 0 & \text{if } n(i) = 0; \text{ and} \\ -1 & \text{if } n(i) = *. \end{cases}$$

Then, we define the **order** of M, which we denote $\operatorname{ord}(M)$, as the sum $s_M(1) + \cdots + s_M(t)$ of the elements in s_M.

Remark 6.17. If u is unitary, then for every $*$-word M, one has $M(u) = u^{\operatorname{ord}(M)}$.

Proposition 6.18. *Let $(\mathcal{A}_k, *, \varphi_k)$ (where $k \leqslant K$) and the collections $\mathbf{a}_k$ and $\otimes_k \mathbf{a}_k$ be defined as in Problem 6.2, and suppose that φ_1 is faithful. If $\otimes_k \mathbf{a}_k$ satisfies the reformulated centering conditions with $\mathbf{a}_1$ as a dominating family, then, for every $i \in I$, one has that*

(1) for every $k \geqslant 2$, $a_{k;i}$ is a constant multiple of a unitary variable; and

(2) if there exists $*$-words M and N such that $\varphi_1\big(M(a_{1;i})\big), \varphi_1\big(N(a_{1;i})\big) \neq 0$ and $|\operatorname{ord}(M)|$ and $|\operatorname{ord}(N)|$ are relative primes, then $a_{k;i}$ is a $\mathbb{C}$-multiple of $1_{\mathcal{A}_k}$ for all $k \geqslant 2$.

Proof. (1). Since φ_1 is faithful, it follows from our assumption that $a_{1;i} \neq 0_{\mathcal{A}_1}$ (see Remark 6.4) that $\varphi_1(a_{1;i}a_{1;i}^*) \neq 0$ for all $i \in I$. Thus, it follows from the reformulated centering conditions that for every $i \in I$ and $k \geqslant 2$, the variable $a_{k;i}a_{k;i}^*$ is a constant multiple of $1_{\mathcal{A}_k}$.

(2). Since $\varphi_1\big(M(a_{1;i})\big), \varphi_1\big(N(a_{1;i})\big) \neq 0$, it follows from the reformulated centering conditions that $M(a_{k;i})$ and $N(a_{k;i})$ are nonzero constant multiples of $1_{\mathcal{A}_k}$ for every $i \in I$ and $k \geqslant 2$. Furthermore, according to part (1) of this proposition, we know that $a_{k;i}$ is a constant multiple of a unitary random variable for all $i \in I$ and $k \geqslant 2$, and thus $M(a_{k;i}) = c_{k;i}^{(M)} \cdot a_{k;i}^{\operatorname{ord}(M)}$ and $N(a_{k;i}) = c_{k;i}^{(N)} \cdot a_{k;i}^{\operatorname{ord}(N)}$ for some constants $\Big(c_{k;i}^{(M)}, c_{k;i}^{(N)} \in \mathbb{C} : i \in I \text{ and } k \geqslant 2\Big)$. Since $|\operatorname{ord}(M)|$ and $|\operatorname{ord}(N)|$ are relative primes, it follows from Bézout's identity that $\alpha \cdot \operatorname{ord}(M) + \beta \cdot \operatorname{ord}(N) = 1$ for some $\alpha, \beta \in \mathbb{Z}$, and thus $a_{k;i} = a_{k;i}^{\alpha \cdot \operatorname{ord}(M) + \beta \cdot \operatorname{ord}(N)}$ is a constant multiple of $1_{\mathcal{A}_k}$ for all $i \in I$ and $k \geqslant 2$. □

6.4. Tensor Product of Two $*$-free Families

In this section, we solve Problem 6.3 in the case where $K = 2$ and each factor family contrains $*$-free variables, that is (we pick the notation $\mathbf{a} = (a_i : i \in I)$ and $\mathbf{b} = (b_i : i \in I)$ instead of $\mathbf{a}_1 = (a_{1;i} : i \in I)$ and $\mathbf{a}_2 = (a_{2;i} : i \in I)$ in order to diminish the amount of indices and thus improve readability in the technical proofs to come in this section)

Problem 6.19. Let $\mathbf{a} = (a_i : i \in I)$ and $\mathbf{b} = (b_i : i \in I)$ be two families of noncommutative random variables in respective $*$-algebras $(\mathcal{A}, *, \varphi)$ and $(\mathcal{B}, *, \psi)$. Characterize the $*$-freeness of the elements in $\mathbf{a} \otimes \mathbf{b} = (a_i \otimes b_i : i \in I)$ in $(\mathcal{A} \otimes \mathcal{B}, *, \varphi \otimes \psi)$ under the assumption that

(1) the a_i are $*$-free from each other;

(2) the b_i are $*$-free from each other; and

(3) φ and ψ are faithful.

The solution to the above problem is the subject of the following theorem:

Theorem 6.20. *Let $(\mathcal{A}, *, \varphi)$, $(\mathcal{B}, *, \psi)$, and the collections $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{a} \otimes \mathbf{b}$ be defined as in Problem 6.19. The variables in $\mathbf{a} \otimes \mathbf{b}$ are $*$-free from each other in $(\mathcal{A} \otimes \mathcal{B}, *, \varphi \otimes \psi)$ if and only if $\mathbf{a} \otimes \mathbf{b}$ satisfies the reformulated centering conditions. Furthermore, if the variables in $\mathbf{a} \otimes \mathbf{b}$ are $*$-free from each other, then at least one of the collection $\mathbf{a}$ or $\mathbf{b}$ contains only multiples of unitary variables, and if one collection contains an element that is not a multiple of a unitary variable, then that family is a dominating family of $\mathbf{a} \otimes \mathbf{b}$ (see Definitions 6.10 and 6.14). For example, if $\mathbf{a}$ contains an element that is not a multiple of a unitary variable, then for every $i \in I$ and $*$-word $M \in \mathbb{C}\langle x, x^* \rangle$, one has*

(1) if $\varphi \otimes \psi\big(M(a_i \otimes b_i)\big) = 0$, then $\varphi\big(M(a_i)\big) = 0$; and

(2) if $\varphi \otimes \psi\big(M(a_i \otimes b_i)\big) \neq 0$, then $M(b_i) = \psi\big(M(b_i)\big) \cdot 1_{\mathcal{B}}$.

Thus, throughout this section, we let $(\mathcal{A}, *, \varphi)$ and $(\mathcal{B}, *, \psi)$ be $*$-probability spaces such that φ and ψ are faithful, and we let the collections $\mathbf{a} = (a_i : i \in I) \subset \mathcal{A}$ and $\mathbf{b} = (b_i : i \in I) \subset \mathcal{B}$ be such that the a_i are $*$-free from each other and likewise for the b_i.

Remark 6.21. In relation to Question 6.13, we notice that, while Proposition 6.18 simply asserted that the reformulated centering condition cannot be applied to prove that $\mathbf{X}_n \otimes \mathbf{Y}_n$ is $*$-asymptotically free (where $\mathbf{X}_n$ and $\mathbf{Y}_n$ are families of i.i.d. Wigner matrices), Theorem 6.20 implies that $\mathbf{X}_n \otimes \mathbf{Y}_n$ is definitely *not* asymptotically $*$-free.

6.4.1. Method of Proof. The method of proof used in this section is fairly straightforward. According to Corollary 6.11 and Proposition 6.15, it clearly is the case that the reformulated centering conditions imply the $*$-freeness of the elements in $\mathbf{a} \otimes \mathbf{b}$. Thus, we need only prove that if the variables in $\mathbf{a} \otimes \mathbf{b}$ are $*$-free from each other, then one of $\mathbf{a}$ or $\mathbf{b}$ contains multiples of unitary random variables, $\mathbf{a} \otimes \mathbf{b}$ satisfies the reformulated centering conditions, and if one of $\mathbf{a}$ or $\mathbf{b}$ contains a variable that is not unitary, then that family is a dominating family.

We will find necessary conditions for the $*$-freeness of the elements in $\mathbf{a} \otimes \mathbf{b}$ as follows: By assuming that the a_i are $*$-free from each other in $(\mathcal{A}, *, \varphi)$ and likewise for the b_i in $(\mathcal{B}, *, \psi)$, it follows from equation (6.2) that for every distinct indices $p, q \in I$, $n \in \mathbb{N}$ and $*$-words $M_1, \ldots, M_n, N_1, \ldots, N_n \in \mathbb{C}\langle x, x^* \rangle$, one has

$$\varphi\big(M_1(a_p)N_1(a_q)\cdots M_n(a_p)N_n(a_q)\big) = \sum_{\pi \in NC(n)} \kappa_\pi\big[M_1(a_p), \ldots, M_n(a_p)\big]\varphi_{K(\pi)}\big[N_1(a_q)\cdots N_n(a_q)\big]$$

and

$$\psi\big(M_1(b_p)N_1(b_q)\cdots M_n(b_p)N_n(b_q)\big) = \sum_{\pi \in NC(n)} \kappa_\pi\big[M_1(b_p), \ldots, M_n(b_p)\big]\psi_{K(\pi)}\big[N_1(b_q)\cdots N_n(b_q)\big],$$

and hence

$$\begin{aligned}
&(\varphi\otimes\psi)\big(M_1(a_p\otimes b_p)N_1(a_q\otimes b_q)\cdots M_n(a_p\otimes b_p)N_n(a_q\otimes b_q)\big)\\
=&\varphi\big(M_1(a_p)N_1(a_q)\cdots M_n(a_p)N_n(a_q)\big)\psi\big(M_1(b_p)N_1(b_q)\cdots M_n(b_p)N_n(b_q)\big)\\
=&\left(\sum_{\pi\in NC(n)}\kappa_\pi\big[M_1(a_p),\ldots,M_n(a_p)\big]\varphi_{K(\pi)}\big[N_1(a_q)\cdots N_n(a_q)\big]\right)\\
&\qquad\times\left(\sum_{\pi\in NC(n)}\kappa_\pi\big[M_1(b_p),\ldots,M_n(b_p)\big]\psi_{K(\pi)}\big[N_1(b_q)\cdots N_n(b_q)\big]\right)\\
=&\Bigg(\sum_{\sigma,\pi\in NC(n)}\kappa_\sigma\big[M_1(a_p),\ldots,M_n(a_p)\big]\kappa_\pi\big[M_1(b_p),\ldots,M_n(b_p)\big]\\
&\qquad\times\varphi_{K(\sigma)}\big[N_1(a_q),\cdots,N_n(a_q)\big]\psi_{K(\pi)}\big[N_1(b_q),\cdots,N_n(b_q)\big]\Bigg).
\end{aligned}\tag{6.7}$$

If it is also assumed that the $a_i\otimes b_i$ are $*$-free from each other, we obtain from equation (6.2) that

$$\begin{aligned}
&(\varphi\otimes\psi)\big(M_1(a_p\otimes b_p)N_1(a_q\otimes b_q)\cdots M_n(a_p\otimes b_p)N_n(a_q\otimes b_q)\big)\\
=&\sum_{\pi\in NC(n)}\kappa_\pi\big[M_1(a_p\otimes b_p),\ldots,M_n(a_p\otimes b_p)\big](\varphi\otimes\psi)_{K(\pi)}\big[N_1(a_q\otimes b_q)\cdots N_n(a_q\otimes b_q)\big]\\
=&\sum_{\pi\in NC(n)}\kappa_\pi\big[M_1(a_p\otimes b_p),\ldots,M_n(a_p\otimes b_p)\big]\varphi_{K(\pi)}\big[N_1(a_q)\cdots N_n(a_q)\big]\\
&\qquad\times\psi_{K(\pi)}\big[N_1(b_q)\cdots N_n(b_q)\big].
\end{aligned}\tag{6.8}$$

Then, subtracting (6.8) from (6.7) yields

$$\begin{aligned}
0=&\Bigg(\sum_{\pi\in NC(n)}\Big(\kappa_\pi\big[M_1(a_p),\ldots,M_n(a_p)\big]\kappa_\pi\big[M_1(b_p),\ldots,M_n(b_p)\big]\\
&\quad-\kappa_\pi\big[M_1(a_p\otimes b_p),\ldots,M_n(a_p\otimes b_p)\big]\Big)\\
&\qquad\times\varphi_{K(\pi)}\big[N_1(a_q),\ldots,N_n(a_q)\big]\psi_{K(\pi)}\big[N_1(b_q),\ldots,N_n(b_q)\big]\Bigg)\\
&+\Bigg(\sum_{\sigma,\pi\in NC(n):\ \sigma\neq\pi}\kappa_\sigma\big[M_1(a_p),\ldots,M_n(a_p)\big]\kappa_\pi\big[M_1(b_p),\ldots,M_n(b_p)\big]\\
&\qquad\times\varphi_{K(\sigma)}\big[N_1(a_q),\ldots,N_n(a_q)\big]\psi_{K(\pi)}\big[N_1(b_q),\ldots,N_n(b_q)\big]\Bigg)
\end{aligned}\tag{6.9}$$

Thus, equation (6.9) offers, in principle, infinitely many necessary conditions on the $*$-distributions of the a_i and the b_i for the $*$-freeness of the $a_i\otimes b_i$. The *kind* of information on the distributions of the a_i and b_i we wish to extract (for example, if they are unitary and the $*$-words M for

which $\varphi\big(M(a_i)\big)$ or $\psi\big(M(b_i)\big)$ are zero) will inform the choice of $*$-words M_k and N_k ($k \leqslant n$) used in equation (6.9).

In the following propositions, we specialize equation (6.9) to two special cases that will be especially useful for our purposes:

Proposition 6.22. *Let $i, j \in I$ be distinct indices and let $M, N \in \mathbb{C}\langle x, x^*\rangle$ be nontrivial $*$-words. If $a_i \otimes b_i$ and $a_j \otimes b_j$ are $*$-free from each other, then*

$$\begin{aligned}
0 =& \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(N(b_j)\big)\Big|^2\varphi\big(N(a_j)N(a_j)^*\big)\psi\big(M(b_i)M(b_i)^*\big) \\
&+ \Big|\varphi\big(N(a_j)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2\varphi\big(M(a_i)M(a_i)^*\big)\psi\big(N(b_j)N(b_j)^*\big) \\
&- \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2\Big|\psi\big(N(b_j)\big)\Big|^2\varphi\big(N(a_j)N(a_j)^*\big) \\
&- \Big|\varphi\big(N(a_j)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2\Big|\psi\big(N(b_j)\big)\Big|^2\varphi\big(M(a_i)M(a_i)^*\big) \\
&- \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\varphi\big(N(a_j)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2\psi\big(N(b_j)N(b_j)^*\big) \\
&- \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\varphi\big(N(a_j)\big)\Big|^2\Big|\psi\big(N(b_j)\big)\Big|^2\psi\big(M(b_i)M(b_i)^*\big) \\
&+ 2\Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\varphi\big(N(a_j)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2\Big|\psi\big(N(b_j)\big)\Big|^2. \qquad (6.10)
\end{aligned}$$

Proof. By applying equation (6.9) with $p = i$, $q = j$, $n = 2$ and

$$M_1(x) = M(x),\ N_1(x) = N(x),\ M_2(x) = M(x)^*, \text{ and } N_2(x) = N(x)^*,$$

we obtain that

$$\begin{aligned}
0 =& \Bigg(\sum_{\pi\in NC(2)}\Big(\kappa_\pi\big[M(a_i), M(a_i)^*\big]\kappa_\pi\big[M(b_i), M(b_i)^*\big] - \kappa_\pi\big[M(a_i\otimes b_i), M(a_i\otimes b_i)^*\big]\Big) \\
&\qquad \times \varphi_{K(\pi)}\big[N(a_j), N(a_j)^*\big]\psi_{K(\pi)}\big[N(b_j), N(b_j)^*\big]\Bigg) \\
&+ \Bigg(\sum_{\sigma,\pi\in NC(2):\ \sigma\neq\pi} \kappa_\sigma\big[M(a_i), M(a_i)^*\big]\kappa_\pi\big[M_1(b_i), M(b_i)^*\big] \\
&\qquad \times \varphi_{K(\sigma)}\big[N(a_j), N(a_j)^*\big]\psi_{K(\pi)}\big[N(b_j), N(b_j)^*\big]\Bigg) \qquad (6.11)
\end{aligned}$$

According to Example D.16, it is clear that for any two noncommutative random variables c and d in a $*$-probability space $(\mathcal{C}, *, \tau)$, one has $\kappa_{0_2}[c, d] = \tau(c)\tau(d)$ and $\kappa_{1_2}[c, d] = \tau(cd) - \tau(c)\tau(d)$, hence

$$\kappa_{0_2}\big[M(c), M(c)^*\big] = \Big|\tau\big(M(c)\big)\Big|^2 \text{ and } \kappa_{1_2}\big[M(c), M(c)^*\big] = \tau\big(M(c)M(c)^*\big) - \Big|\tau\big(M(c)\big)\Big|^2.$$

Similarly,

$$\kappa_{0_2}\big[M(a_i\otimes b_i), M(a_i\otimes b_i)^*\big] = \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2,$$

and

$$\kappa_{1_2}\big[M(a_i\otimes b_i), M(a_i\otimes b_i)^*\big] \\ = \varphi\big(M(a_i)M(a_i)^*\big)\psi\big(M(b_i)M(b_i)^*\big) - \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2.$$

We now carry out the two sums in the right-hand side of (6.11) one summand at a time:
First sum in (6.11), summand $\pi = 0_2$. Since

$$\begin{aligned}\kappa_{0_2}\big[M(a_i), M(a_i)^*\big]\kappa_{0_2}\big[M(b_i), M(b_i)^*\big] =& \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2 \\ =& \kappa_{0_2}\big[M(a_i\otimes b_i), M(a_i\otimes b_i)^*\big],\end{aligned}$$

then

$$0 = \kappa_{0_2}\big[M(a_i), M(a_i)^*\big]\kappa_{0_2}\big[M(b_i), M(b_i)^*\big] - \kappa_{0_2}\big[M(a_i\otimes b_i), M(a_i\otimes b_i)^*\big],$$

and thus this summand vanishes.
First sum in (6.11), summand $\pi = 1_2$. Given that

$$\begin{aligned}&\kappa_{1_2}\big[M(a_i), M(a_i)^*\big]\kappa_{1_2}\big[M(b_i), M(b_i)^*\big] \\ =&\Big(\varphi\big(M(a_i)M(a_i)^*\big) - \Big|\varphi\big(M(a_i)\big)\Big|^2\Big)\Big(\psi\big(M(b_i)M(b_i)^*\big) - \Big|\psi\big(M(b_i)\big)\Big|^2\Big) \\ =&\varphi\big(M(a_i)M(a_i)^*\big)\psi\big(M(b_i)M(b_i)^*\big) - \varphi\big(M(a_i)M(a_i)^*\big)\Big|\psi\big(M(b_i)\big)\Big|^2 \\ &- \Big|\varphi\big(M(a_i)\big)\Big|^2\psi\big(M(b_i)M(b_i)^*\big) + \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2\end{aligned}$$

and

$$\kappa_{1_2}\big[M(a_i\otimes b_i), M(a_i\otimes b_i)^*\big] \\ = \varphi\big(M(a_i)M(a_i)^*\big)\psi\big(M(b_i)M(b_i)^*\big) - \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2,$$

then

$$\begin{aligned}&\kappa_{1_2}\big[M(a_i), M(a_i)^*\big]\kappa_{1_2}\big[M(b_i), M(b_i)^*\big] - \kappa_{1_2}\big[M(a_i\otimes b_i), M(a_i\otimes b_i)^*\big] \\ =& -\varphi\big(M(a_i)M(a_i)^*\big)\Big|\psi\big(M(b_i)\big)\Big|^2 - \Big|\varphi\big(M(a_i)\big)\Big|^2\psi\big(M(b_i)M(b_i)^*\big) \\ &+ 2\Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2.\end{aligned}$$

Therefore, since $K(1_2) = 0_2$ and $\varphi_{0_2}[a, b] = \varphi(a)\varphi(b)$, one has

$$\begin{aligned}
&\Big(\kappa_{1_2}\big[M(a_i), M(a_i)^*\big]\kappa_{1_2}\big[M(b_i), M(b_i)^*\big] - \kappa_{0_2}\big[M(a_i \otimes b_i), M(a_i \otimes b_i)^*\big]\Big) \\
&\qquad \times \varphi_{K(1_2)}\big[N(a_j), N(a_j)^*\big]\psi_{K(1_2)}\big[N(b_j), N(b_j)^*\big] \\
=&\Big(- \varphi\big(M(a_i)M(a_i)^*\big)\Big|\psi\big(M(b_i)\big)\Big|^2 - \Big|\varphi\big(M(a_i)\big)\Big|^2\psi\big(M(b_i)M(b_i)^*\big) \\
&\quad + 2\Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2\Big)\Big|\varphi\big(N(a_j)\big)\Big|^2\Big|\psi\big(N(b_j)\big)\Big|^2 \\
=& \ - \Big|\varphi\big(N(a_j)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2\Big|\psi\big(N(b_j)\big)\Big|^2\varphi\big(M(a_i)M(a_i)^*\big) \\
&\quad - \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\varphi\big(N(a_j)\big)\Big|^2\Big|\psi\big(N(b_j)\big)\Big|^2\psi\big(M(b_i)M(b_i)^*\big) \\
&\quad + 2\Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\varphi\big(N(a_j)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2\Big|\psi\big(N(b_j)\big)\Big|^2. \qquad (6.12)
\end{aligned}$$

Second sum in (6.11), summand $\sigma = 0_2$ and $\pi = 1_2$. In this case,

$$\begin{aligned}
&\kappa_{0_2}\big[M(a_i), M(a_i)^*\big]\kappa_{1_2}\big[M(b_i), M(b_i)^*\big]\varphi_{1_2}\big[N(a_j), N(a_j)^*\big]\psi_{0_2}\big[N(b_j), N(b_j)^*\big] \\
=&\Big|\varphi\big(M(a_i)\big)\Big|^2\Big(\psi\big(M(b_i)M(b_i)^*\big) - \Big|\psi\big(M(b_i)\big)\Big|^2\Big)\varphi\big(N(a_j)N(a_j)^*\big)\Big|\psi\big(N(b_j)\big)\Big|^2 \\
=&\Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(N(b_j)\big)\Big|^2\varphi\big(N(a_j)N(a_j)^*\big)\psi\big(M(b_i)M(b_i)^*\big) \\
&\quad - \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2\Big|\psi\big(N(b_j)\big)\Big|^2\varphi\big(N(a_j)N(a_j)^*\big). \qquad (6.13)
\end{aligned}$$

Second sum in (6.11), summand $\sigma = 1_2$ and $\pi = 0_2$. In this case,

$$\begin{aligned}
&\kappa_{1_2}\big[M(a_i), M(a_i)^*\big]\kappa_{0_2}\big[M(b_i), M(b_i)^*\big]\varphi_{0_2}\big[N(a_j), N(a_j)^*\big]\psi_{1_2}\big[N(b_j), N(b_j)^*\big] \\
=&\Big(\varphi\big(M(a_i)M(a_i)^*\big) - \Big|\varphi\big(M(a_i)\big)\Big|^2\Big)\Big|\psi\big(M(b_i)\big)\Big|^2\Big|\varphi\big(N(a_j)\big)\Big|^2\psi\big(N(b_j)N(b_j)^*\big) \\
=&\Big|\varphi\big(N(a_j)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2\varphi\big(M(a_i)M(a_i)^*\big)\psi\big(N(b_j)N(b_j)^*\big) \\
&\quad - \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\varphi\big(N(a_j)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2\psi\big(N(b_j)N(b_j)^*\big) \qquad (6.14)
\end{aligned}$$

Combining all of the above summands (that is, equations (6.12), (6.13), and (6.14)) shows that (6.11) does indeed reduce to (6.10), as desired. □

Proposition 6.23. *Let $i, j \in I$ be distinct indices. If a_j and b_j are unitary, such that $\varphi(a_j) = \psi(b_j) = 0$, and $a_i \otimes b_i$ is $*$-free from $a_j \otimes b_j$, then for every $*$-words $N, M \in \mathbb{C}\langle x, x^*\rangle$, one has*

$$\begin{aligned}
0 =&\varphi\Big(N(a_i)N(a_i)^*\Big)\psi\Big(M(b_i)M(b_i)^*\Big)\Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(N(b_i)\big)\Big|^2 \\
&+ \varphi\Big(M(a_i)M(a_i)^*\Big)\psi\Big(N(b_i)N(b_i)^*\Big)\Big|\varphi\big(N(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2 \\
&- \psi\Big(M(b_i)M(b_i)^*\Big)\Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\varphi\big(N(a_i)\big)\Big|^2\Big|\psi\big(N(b_i)\big)\Big|^2 \\
&- \psi\Big(N(b_i)N(b_i)^*\Big)\Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\varphi\big(N(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2 \\
&- \varphi\Big(M(a_i)M(a_i)^*\Big)\Big|\varphi\big(N(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2\Big|\psi\big(N(b_i)\big)\Big|^2 \\
&- \varphi\Big(N(a_i)N(a_i)^*\Big)\Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2\Big|\psi\big(N(b_i)\big)\Big|^2 \\
&+ 2\Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\varphi\big(N(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2\Big|\psi\big(N(b_i)\big)\Big|^2. \qquad (6.15)
\end{aligned}$$

In particular, if $N = M$, then

$$\begin{aligned}
0 =&\varphi\Big(M(a_i)M(a_i)^*\Big)\psi\Big(M(b_i)M(b_i)^*\Big)\Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2 \\
&- \psi\Big(M(b_i)M(b_i)^*\Big)\Big|\varphi\big(M(a_i)\big)\Big|^4\Big|\psi\big(M(b_i)\big)\Big|^2 \\
&- \varphi\Big(M(a_i)M(a_i)^*\Big)\Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^4 \\
&+ \Big|\varphi\big(M(a_i)\big)\Big|^4\Big|\psi\big(M(b_i)\big)\Big|^4. \qquad (6.16)
\end{aligned}$$

Proof. By applying equation (6.9) with $p = j$, $q = i$, $n = 4$ and the following choice of $*$-words:

$$\begin{array}{cccc} M_1(x) = x & M_2(x) = x^* & M_3(x) = x & M_4(x) = x^* \\ N_1(x) = M(x) & N_2(x) = N(x) & N_3(x) = M(x)^* & N_4(x) = N(x)^*, \end{array}$$

we obtain that

$$\begin{aligned}
0 =&\Bigg(\sum_{\pi\in NC(4)}\Big(\kappa_\pi\big[a_j, a_j^*, a_j, a_j^*\big]\kappa_\pi\big[b_j, b_j^*, b_j, b_j^*\big] \\
&\quad - \kappa_\pi\big[a_j\otimes b_j, (a_j\otimes b_j)^*, a_j\otimes b_j, (a_j\otimes b_j)^*\big]\Big) \\
&\quad \times \varphi_{K(\pi)}\big[M(a_i), N(a_i), M(a_i)^*, N(a_i)^*\big]\psi_{K(\pi)}\big[M(b_i), N(b_i), M(b_i)^*, N(b_i)^*\big]\Bigg) \\
&+\Bigg(\sum_{\sigma,\pi\in NC(4):\ \sigma\neq\pi}\kappa_\sigma\big[a_j, a_j^*, a_j, a_j^*\big]\kappa_\pi\big[b_j, b_j^*, b_j, b_j^*\big] \\
&\quad \times \varphi_{K(\sigma)}\big[M(a_i), N(a_i), M(a_i)^*, N(a_i)^*\big]\psi_{K(\pi)}\big[M(b_i), N(b_i), M(b_i)^*, N(b_i)^*\big]\Bigg). \qquad (6.17)
\end{aligned}$$

Given that

$$\varphi(a_j) = \psi(b_j) = (\varphi\otimes\psi)(a_j\otimes b_j) = 0,$$

the only partitions $\pi \in NC(4)$ for which the free cumulants

$$\kappa_\pi[a_j, a_j^*, a_j, a_j^*],\ \kappa_\pi[b_j, b_j^*, b_j, b_j^*]$$

and

$$\kappa_\pi[a_j \otimes b_j, (a_j \otimes b_j)^*, a_j \otimes b_j, (a_j \otimes b_j)^*]$$

are possibly nonzero are those that do not contain any singleton, namely (using the notation introduced in Example D.17),

$$\begin{array}{rcl} \pi_8 & = & \big\{\{1,4\},\{2,3\}\big\} \\ \pi_9 & = & \big\{\{1,2\},\{3,4\}\big\} \\ \pi_{14} & = & 1_4 = \big\{\{1,2,3,4\}\big\} \end{array},$$

whose Kreweras complements are given by

$$\begin{array}{rcl} K(\pi_8) & = & \pi_6 = \big\{\{1,3\},\{2\},\{4\}\big\} \\ K(\pi_9) & = & \pi_7 = \big\{\{1\},\{2,4\},\{3\}\big\} \\ K(\pi_{14}) & = & 0_4 = \big\{\{1\},\{2\},\{3\},\{4\}\big\} \end{array}.$$

Using the computations made in Example 4.21 and the fact that a_j and b_j are unitary, we obtain that

$$\begin{aligned} &\kappa_{1_4}\big[a_j \otimes b_j, (a_j \otimes b_j)^*, a_j \otimes b_j, (a_j \otimes b_j)^*\big] \\ &\qquad = \kappa_{1_4}\big[a_j, a_j^*, a_j, a_j^*\big] = \kappa_{1_4}\big[b_j, b_j^*, b_j, b_j^*\big] = -1 \end{aligned} \tag{6.18}$$

and

$$\begin{aligned} &\kappa_{\pi}\big[a_j \otimes b_j, (a_j \otimes b_j)^*, a_j \otimes b_j, (a_j \otimes b_j)^*\big] \\ &\qquad = \kappa_{\pi}\big[a_j, a_j^*, a_j, a_j^*\big] = \kappa_{\pi}\big[b_j, b_j^*, b_j, b_j^*\big] = 1 \end{aligned} \tag{6.19}$$

for $\pi = \pi_8$ and $\pi = \pi_9$.

We now carry out the two sums in the right-hand side of (6.17) one summand at a time:

First sum in (6.17), summands $\pi = \pi_8$ and $\pi = \pi_9$. According to (6.19),

$$0 = \kappa_\pi\big[a_j \otimes b_j, (a_j \otimes b_j)^*, a_j \otimes b_j, (a_j \otimes b_j)^*\big] - \kappa_\pi\big[a_j, a_j^*, a_j, a_j^*\big]\kappa_\pi\big[b_j, b_j^*, b_j, b_j^*\big],$$

for $\pi = \pi_8, \pi_9$, hence the summand vanishes for $\pi = \pi_8, \pi_9$.

First sum in (6.17), summand $\pi = 1_4$. According to (6.18), one has

$$2 = \kappa_{1_4}\big[a_j, a_j^*, a_j, a_j^*\big]\kappa_{1_4}\big[b_j, b_j^*, b_j, b_j^*\big] - \kappa_{1_4}\big[(a_j \otimes b_j), (a_j \otimes b_j), (a_j \otimes b_j)^*, (a_j \otimes b_j)^*\big]$$

Therefore,

$$\begin{aligned} &\Big(\kappa_{1_4}\big[a_j, a_j^*, a_j, a_j^*\big]\kappa_{1_4}\big[b_j, b_j^*, b_j, b_j^*\big] - \kappa_{1_4}\big[(a_j \otimes b_j), (a_j \otimes b_j), (a_j \otimes b_j)^*, (a_j \otimes b_j)^*\big]\Big) \\ &\qquad \times \varphi_{0_4}\big[M(a_i), N(a_i), M(a_i)^*, N(a_i)^*\big]\psi_{0_4}\big[M(b_i), N(b_i), M(b_i)^*, N(b_i)^*\big] \\ =&2\Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\varphi\big(N(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2\Big|\psi\big(N(b_i)\big)\Big|^2. \end{aligned} \tag{6.20}$$

Second sum in (6.17), summand $\sigma = \pi_8$ and $\pi = \pi_9$. According to equation (6.19) and the fact that $K(\pi_8) = \pi_6$ and $K(\pi_9) = \pi_7$, one has

$$\begin{aligned}&\kappa_{\pi_8}\big[a_j, a_j^*, a_j, a_j^*\big]\kappa_{\pi_9}\big[b_j, b_j^*, b_j, b_j^*\big]\varphi_{\pi_6}\big[M(a_i), N(a_i), M(a_i)^*, N(a_i)^*\big]\\ &\qquad\times \psi_{\pi_7}\big[M(b_i), N(b_i), M(b_i)^*, N(b_i)^*\big]\\ =&\varphi\Big(M(a_i)M(a_i)^*\Big)\psi\Big(N(b_i)N(b_i)^*\Big)\Big|\varphi\big(N(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2. \end{aligned} \tag{6.21}$$

Second sum in (6.17), summand $\sigma = \pi_9$ and $\pi = \pi_8$. According to equation (6.19) and the fact that $K(\pi_9) = \pi_7$ and $K(\pi_8) = \pi_6$, one has

$$\begin{aligned}&\kappa_{\pi_9}\big[a_j, a_j^*, a_j, a_j^*\big]\kappa_{\pi_8}\big[b_j, b_j^*, b_j, b_j^*\big]\varphi_{\pi_7}\big[M(a_i), N(a_i), M(a_i)^*, N(a_i)^*\big]\\ &\qquad\times \psi_{\pi_6}\big[M(b_i), N(b_i), M(b_i)^*, N(b_i)^*\big]\\ =&\varphi\Big(N(a_i)N(a_i)^*\Big)\psi\Big(M(b_i)M(b_i)^*\Big)\Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(N(b_i)\big)\Big|^2. \end{aligned} \tag{6.22}$$

Second sum in (6.17), summand $\sigma = 1_4$ and $\pi = \pi_8$. According to equations (6.19) and (6.18), and the fact that $K(1_4) = 0_4$ and $K(\pi_8) = \pi_6$, one has

$$\begin{aligned}&\kappa_{1_4}\big[a_j, a_j^*, a_j, a_j^*\big]\kappa_{\pi_8}\big[b_j, b_j^*, b_j, b_j^*\big]\varphi_{0_4}\big[M(a_i), N(a_i), M(a_i)^*, N(a_i)^*\big]\\ &\qquad\times \psi_{\pi_6}\big[M(b_i), N(b_i), M(b_i)^*, N(b_i)^*\big]\\ =&-\psi\Big(M(b_i)M(b_i)^*\Big)\Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\varphi\big(N(a_i)\big)\Big|^2\Big|\psi\big(N(b_i)\big)\Big|^2. \end{aligned} \tag{6.23}$$

Second sum in (6.17), summand $\sigma = \pi_8$ and $\pi = 1_4$. According to equations (6.19) and (6.18), and the fact that $K(1_4) = 0_4$ and $K(\pi_8) = \pi_6$, one has

$$\begin{aligned}&\kappa_{\pi_8}\big[a_j, a_j^*, a_j, a_j^*\big]\kappa_{1_4}\big[b_j, b_j^*, b_j, b_j^*\big]\varphi_{\pi_6}\big[M(a_i), N(a_i), M(a_i)^*, N(a_i)^*\big]\\ &\qquad\times \psi_{0_4}\big[M(b_i), N(b_i), M(b_i)^*, N(b_i)^*\big]\\ =&-\varphi\Big(M(a_i)M(a_i)^*\Big)\Big|\varphi\big(N(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2\Big|\psi\big(N(b_i)\big)\Big|^2. \end{aligned} \tag{6.24}$$

Second sum in (6.17), summand $\sigma = 1_4$ and $\pi = \pi_9$. According to equations (6.19) and (6.18), and the fact that $K(1_4) = 0_4$ and $K(\pi_9) = \pi_7$, one has

$$\begin{aligned}&\kappa_{1_4}\big[a_j, a_j^*, a_j, a_j^*\big]\kappa_{\pi_9}\big[b_j, b_j^*, b_j, b_j^*\big]\varphi_{0_4}\big[M(a_i), N(a_i), M(a_i)^*, N(a_i)^*\big]\\ &\qquad\times \psi_{\pi_7}\big[M(b_i), N(b_i), M(b_i)^*, N(b_i)^*\big]\\ =&-\psi\Big(N(b_i)N(b_i)^*\Big)\Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\varphi\big(N(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2. \end{aligned} \tag{6.25}$$

Second sum in (6.17), summand $\sigma = \pi_9$ and $\pi = 1_4$. According to equations (6.19) and (6.18), and the fact that $K(1_4) = 0_4$ and $K(\pi_9) = \pi_7$, one has

$$\begin{aligned}&\kappa_{\pi_9}\big[a_j, a_j^*, a_j, a_j^*\big]\kappa_{1_4}\big[b_j, b_j^*, b_j, b_j^*\big]\varphi_{\pi_7}\big[M(a_i), N(a_i), M(a_i)^*, N(a_i)^*\big]\\ &\qquad\times \psi_{0_4}\big[M(b_i), N(b_i), M(b_i)^*, N(b_i)^*\big]\\ =&-\varphi\Big(N(a_i)N(a_i)^*\Big)\Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2\Big|\psi\big(N(b_i)\big)\Big|^2. \end{aligned} \tag{6.26}$$

Combining all the above summands (that is, equations (6.20) to (6.26)) shows that (6.17) does indeed reduce to (6.15), as desired. Then, by replacing every instance of N by M in (6.15), we obtain (6.16). □

6.4.2. Technical Assumptions/Results. In order to simplify computations, we will begin with a few technical remarks and propositions.

Remark 6.24. In the context of Problem 6.3, the assumption that the a_i and b_i are nonzero (see Remark 6.4) can be taken one step further: Since φ and ψ are faithful, it follows that $\varphi(a_i a_i^*)$ and $\psi(b_i b_i^*)$ are nonzero for every $i \in I$. Therefore, given that the $*$-freeness of the elements in a collection $(c_j : j \in J) \subset (\mathcal{C}, *, \tau)$ is in no way affected by scaling the variables c_j using nonzero constants $\lambda_j \in \mathbb{C}$ (that is, the c_j are $*$-free from each other if and only if the $\lambda_j c_j$ are $*$-free from each other, provided the λ_j are nonzero), we will henceforth assume that $\varphi(a_i a_i^*) = \psi(b_i b_i^*) = 1$ for every $i \in I$, which amounts to scaling the a_i by $\big(\varphi(a_i a_i^*)\big)^{-1/2}$ and the b_i by $\big(\psi(b_i b_i^*)\big)^{-1/2}$.

Proposition 6.25. *Let $(\mathcal{A}, *, \varphi)$ be a $*$-probability space such that φ is faithful, and let $a \in (\mathcal{A}, *, \varphi)$ be such that $\varphi(aa^*) = 1$. If $|\varphi(a)| = 1$, then $a = \varphi(a) \cdot 1_{\mathcal{A}}$. Furthermore, if $a = \lambda \cdot 1_A$ for some $\lambda \in \mathbb{C}$, then $|\lambda| = 1$.*

Proof. Let $\lambda = \varphi(a)$, and suppose that $|\lambda| = 1$. Then,

$$\begin{aligned}\mathbf{Var}[a] =& \varphi\big((a - \lambda \cdot 1_A)(a - \lambda \cdot 1_A)^*\big) = \varphi(aa^* - \overline{\lambda}a - \lambda a^* + \lambda\overline{\lambda}) \\ =& 1 - |\lambda|^2 - |\lambda|^2 + |\lambda|^2 = 0.\end{aligned}$$

Since φ is faithful, this implies that $a = \lambda \cdot 1_A$.

Suppose that $a = \lambda \cdot 1_A$ for some $\lambda \in \mathbb{C}$. Then, $\varphi(aa^*) = 1$ obviously implies that $|\lambda|^2 = 1$, as desired. □

Remark 6.26. As shown in Proposition 6.25, under the assumption we have just made that $\varphi(a_i a_i^*) = \psi(b_i b_i^*) = 1$, proving that a_i or b_i is a constant multiple of $1_{\mathcal{A}}$ amounts to showing that $|\varphi(a_i)|$ or $|\psi(b_i)|$ is equal to 1. Thus, in this case, the reformulated centering conditions with $\mathbf{a}$ as a dominating family would be equivalent to the following: for every $i \in I$ and $*$-word $M \in \mathbb{C}\langle x, x^*\rangle$,

(1) if $(\varphi \otimes \psi)\big(M(a_i \otimes b_i)\big) = 0$, then $\varphi\big(M(a_i)\big) = 0$; and

(2) if $(\varphi \otimes \psi)\big(M(a_i \otimes b_i)\big) \neq 0$, then $\Big|\psi\big(M(b_i)\big)\Big| = 1$.

Under the assumption that $\varphi(a_i a_i^*) = \psi(b_i b_i^*) = 1$, we also obtain the following inequality, which will be useful in some proofs:

Proposition 6.27. *Let a be a noncommutative random variable in a $*$-probability space $(\mathcal{A}, *, \varphi)$. If $\varphi(aa^*) = 1$, then*

$$|\varphi(a)| \leqslant 1 \leqslant \varphi\big((aa^*)^2\big). \tag{6.27}$$

Proof. Since φ is positive, it follows that for every $\theta \in [0, 2\pi)$, one has

$$0 \leqslant \varphi\big((\mathrm{e}^{\mathrm{i}\theta} \cdot 1_{\mathcal{A}} + a)(\mathrm{e}^{-\mathrm{i}\theta} \cdot 1_{\mathcal{A}} + a^*)\big) = 2 + 2\mathrm{Re}\big(\mathrm{e}^{-\mathrm{i}\theta}\varphi(a)\big).$$

If we choose θ in such a way that $\mathrm{Re}(\mathrm{e}^{-\mathrm{i}\theta}\varphi(a)) = -|\varphi(a)|$, the above yields $|\varphi(a)| \leqslant 1$. Furthermore,

$$0 \leqslant \varphi\big((aa^* - 1)(aa^* - 1)\big) = \varphi\big((aa^*)^2\big) - 1,$$

which implies that $\varphi\big((aa^*)^2\big) \geqslant 1$. □

6.4.3. At Least One Unitary Family. We now prove the following claim in Theorem 6.20: If the variables in $\mathbf{a} \otimes \mathbf{b}$ are $*$-free from each other, then at least one of the collection $\mathbf{a}$ or $\mathbf{b}$ contains only unitary variables (under the assumption that $\varphi(a_i a_i^*) = \psi(b_i b_i^*) = 1$ for all $i \in I$).

Lemma 6.28. *Let $i, j \in I$ be distinct indices and suppose that $a_i \otimes b_i$ and $a_j \otimes b_j$ are $*$-free from each other. If both a_j and b_j are unitary, then at least one of a_i or b_i must be unitary as well.*

Proof. Suppose that both a_j and b_j are unitary. We consider the two possible cases with respect to the expectation of a_j and b_j, namely,

(1) at least one of $\varphi(a_j)$ or $\psi(b_j)$ is not equal to zero; and

(2) $\varphi(a_j)$ and $\psi(b_j)$ are both equal to zero.

(1). Suppose that at least one of $\varphi(a_j)$ or $\psi(b_j)$ is not equal to zero. Define the $*$-words $M, N \in \mathbb{C}\langle x, x^*\rangle$ as $M(x) = xx^*$ and $N(x) = x$. Since a_j and b_j are unitary, $N(a_j)N(a_j)^* = 1_{\mathcal{A}}$ and $N(b_j)N(b_j)^* = 1_{\mathcal{B}}$. Combining this with the assumption that $\varphi\big(M(a_i)\big) = \psi\big(M(b_i)\big) = 1$, equation (6.10) reduces in this case to

$$\begin{aligned} 0 = \Big|\psi\big(b_j\big)\Big|^2\psi\big((b_i b_i^*)^2\big) + \Big|\varphi\big(a_j\big)\Big|^2\varphi\big((a_i a_i^*)^2\big) - \Big|\psi\big(b_j\big)\Big|^2 - \Big|\varphi\big(a_j\big)\Big|^2\Big|\psi\big(b_j\big)\Big|^2\varphi\big((a_i a_i^*)^2\big) \\ - \Big|\varphi\big(a_j\big)\Big|^2 - \Big|\varphi\big(a_j\big)\Big|^2\Big|\psi\big(b_j\big)\Big|^2\psi\big((b_i b_i^*)^2\big) + 2\Big|\varphi\big(a_j\big)\Big|^2\Big|\psi\big(b_j\big)\Big|^2. \end{aligned}$$

To simplify computations, let us pose

$$k \coloneqq \Big|\varphi\big(a_j\big)\Big|^2 \text{ and } t \coloneqq \Big|\psi\big(b_j\big)\Big|^2.$$

Thus, we have that

$$0 = t \cdot \psi\big((b_i b_i^*)^2\big) + k \cdot \varphi\big((a_i a_i^*)^2\big) - t - kt \cdot \varphi\big((a_i a_i^*)^2\big) - k - kt \cdot \psi\big((b_i b_i^*)^2\big) + 2kt,$$

which can be rearranged into

$$k + t - 2kt = k(1-t) \cdot \varphi\big((a_i a_i^*)^2\big) + t(1-k) \cdot \psi\big((b_i b_i^*)^2\big).$$

Thus,

$$\Big(\varphi\big((a_i a_i^*)^2\big), \psi\big((b_i b_i^*)^2\big)\Big) \in (0, \infty)^2$$

must be among the couples $(x, y) \in (0, \infty)^2$ that satisfy

$$k + t - 2kt = k(1-t)x + t(1-k)y. \tag{6.28}$$

Notice that

- the couple $(x, y) = (1, 1)$ satisfies (6.28) no matter what the value of t and k are;
- acording to equation (6.27), $0 \leqslant k, t \leqslant 1$;

- given our assumption that at least one of a_j and b_j is not a multiple of the unit (see Remark 6.4), it follows from Proposition 6.25 that at least one of k and t is not equal to 1; and
- since at least one of $\varphi(a_j)$ or $\psi(b_j)$ is nonzero, at least one of k and t is nonzero.

Therefore, (6.28) can always be written as a linear equation of the form

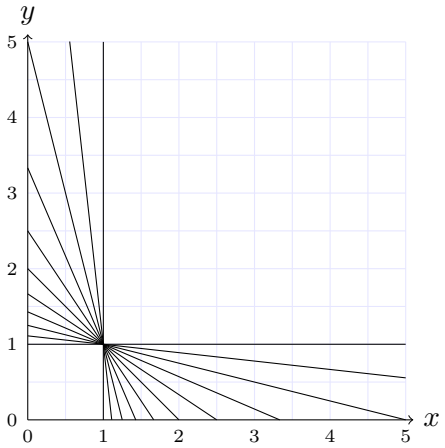
$$y = \frac{1}{t(1-k)}\big(-k(1-t)x + (t+k-2tk)\big)$$

or

$$x = \frac{1}{k(1-t)}\big(-t(1-k)y + (t+k-2tk)\big),$$

which goes through the point $(1,1)$ and has a slope $\frac{\mathrm{d}x}{\mathrm{d}y}$ contained in $(-\infty, 0]$ (indeed, $t(1-k)$ and $k(1-t)$ are both nonnegative, see Figure 1 below for examples of possible set of solutions of (6.28)). Therefore, it necessarily is the case that $\varphi\big((a_i a_i^*)^2\big) \leqslant 1$ or $\psi\big((b_i b_i^*)^2\big) \leqslant 1$, which,

Figure 1.



when combined with (6.27), implies that $\varphi\big((a_i a_i^*)^2\big) = 1$ or $\psi\big((b_i b_i^*)^2\big) = 1$. As $\mathbf{Var}[a_i a_i^*] = \varphi\big((a_i a_i^*)^2\big) - 1$ and similarly for b_i, we conclude that $a_i a_i^* = 1_{\mathcal{A}}$ or $b_i b_i^* = 1_{\mathcal{B}}$, as desired.

(2). Suppose that $\varphi(a_j) = \psi(b_j) = 0$, and define the $*$-word $M \in \mathbb{C}\langle x, x^*\rangle$ as $M(x) = xx^*$. Then, since $\varphi(a_i a_i^*) = \psi(b_i b_i^*) = 1$, equation (6.16) reduces in this case to

$$0 = \varphi\big((a_i a_i^*)^2\big)\psi\big((b_i b_i^*)^2\big) - \varphi\big((a_i a_i^*)^2\big) - \psi\big((b_i b_i^*)^2\big) + 1 = \Big(\varphi\big((a_i a_i^*)^2\big) - 1\Big)\Big(\psi\big((b_i b_i^*)^2\big) - 1\Big),$$

from which we conclude that one of $\mathbf{Var}[a_i a_i^*]$ and $\mathbf{Var}[b_i b_i^*]$ must be zero, as desired.

□

Lemma 6.29. *Let $i, j \in I$ be distinct indices and suppose that $a_i \otimes b_i$ is $*$-free from $a_j \otimes b_j$.*

(1) It cannot be the case that a_i and b_i are both not unitary; and

(2) it cannot be the case that a_i and b_j are both not unitary.

Proof. Define the $*$-words $M, N \in \mathbb{C}\langle x, x^*\rangle$ as $M(x) = N(x) = xx^*$. Since $\varphi\big(M(a_\iota)\big) = \psi\big(M(b_\iota)\big) = 1$ for $\iota = i, j$, equation (6.10) reduces in this case to

$$\begin{aligned}
0 =&\varphi\big((a_j a_j^*)^2\big)\psi\big((b_i b_i^*)^2\big) + \varphi\big((a_i a_i^*)^2\big)\psi\big((b_j b_j^*)^2\big) - \varphi\big((a_j a_j^*)^2\big) - \varphi\big((a_i a_i^*)^2\big) - \psi\big((b_j b_j^*)^2\big) \\
&- \psi\big((b_i b_i^*)^2\big) + 2 \\
=&\varphi\big((a_i a_i^*)^2\big)\psi\big((b_j b_j^*)^2\big) - \varphi\big((a_i a_i^*)^2\big) - \psi\big((b_j b_j^*)^2\big) + 1 + \varphi\big((a_j a_j^*)^2\big)\psi\big((b_i b_i^*)^2\big) \\
&- \varphi\big((a_j a_j^*)^2\big) - \psi\big((b_i b_i^*)^2\big) + 1 \\
=&\Big(\varphi\big((a_i a_i^*)^2\big) - 1\Big)\Big(\psi\big((b_j b_j^*)^2\big) - 1\Big) + \Big(\varphi\big((a_j a_j^*)^2\big) - 1\Big)\Big(\psi\big((b_i b_i^*)^2\big) - 1\Big).
\end{aligned}$$

According to equation (6.27), $\varphi\big((a_\iota a_\iota^*)^2\big), \psi\big((b_\iota b_\iota^*)^2\big) \geqslant 1$ for $\iota = i, j$. Therefore, the above equation implies that

$$\Big(\varphi\big((a_i a_i^*)^2\big) - 1\Big)\Big(\psi\big((b_j b_j^*)^2\big) - 1\Big) = 0, \tag{6.29}$$

and

$$\Big(\varphi\big((a_j a_j^*)^2\big) - 1\Big)\Big(\psi\big((b_i b_i^*)^2\big) - 1\Big) = 0. \tag{6.30}$$

(1). Suppose by contradiction that a_i and b_i are both not unitary, i.e., $\varphi\big((a_i a_i^*)^2\big)$ and $\psi\big((b_i b_i)^2\big)$ are both different from 1. Then, equations (6.29) and (6.30) imply that both a_j and b_j are unitary. This, however, contradicts Lemma 6.28.

(2). Suppose by contradiction that a_i and b_j are both not unitary, i.e., $\varphi\big((a_i a_i^*)^2\big)$ and $\psi\big((b_j b_j)^2\big)$ are both different from 1. This contradicts equation (6.29). □

Proposition 6.30. *Let $(\mathcal{A}, *, \varphi)$, $(\mathcal{B}, *, \psi)$, and the collections $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{a} \otimes \mathbf{b}$ be defined as in Problem 6.19. If the variables in $\mathbf{a} \otimes \mathbf{b}$ are $*$-free from each other, then at least one of the collection $\mathbf{a}$ or $\mathbf{b}$ contains only unitary variables (under the assumption that $\varphi(a_i a_i^*) = \varphi(b_i b_i^*) = 1$ for all $i \in I$).*

Proof. If a_i and b_i are both unitary for every $i \in I$, then the result trivially holds. Thus, suppose that there exists $i \in I$ such that a_i and b_i are not both unitary. Then, according to part (1) of Lemma 6.29, exactly one of a_i and b_i must be unitary. Suppose without loss of generality that b_i is unitary and that a_i is not. Since a_i is not unitary, it then follows from part (2) of Lemma 6.29 that b_j is unitary for every $j \neq i$, and thus every element in $\mathbf{b}$ is unitary, as desired. □

6.4.4. Reformulated Centering Conditions. We now prove the remaining claims in Theorem 6.20, namely, if the variables in $\mathbf{a} \otimes \mathbf{b}$ are $*$-free from each other, then $\mathbf{a} \otimes \mathbf{b}$ satisfies the reformulated centering conditions, and if one of the families $\mathbf{a}$ and $\mathbf{b}$ contains a variable that is not unitary, then that family is a dominating family in $\mathbf{a} \otimes \mathbf{b}$.

Lemma 6.31. *Let $i, j \in I$ be distinct, suppose that $a_i \otimes b_i$ is $*$-free from $a_j \otimes b_j$, and suppose that a_i and a_j are both not unitary, and that b_i and b_j are both unitary. Then, the reformulated centering conditions hold for the couple $(a_i \otimes b_i, a_j \otimes b_j)$, and (a_i, a_j) is a dominating family.*

Proof. Define $M \in \mathbb{C}\langle x, x^*\rangle$ as $M(x) = xx^*$, and let $N \in \mathbb{C}\langle x, x^*\rangle$ be arbitrary. Then, since b_i and b_j are unitary and $\varphi\big(M(a_i)\big) = \psi\big(M(b_i)\big) = 1$, (6.10) reduces in this case to

$$\begin{aligned}
0 =& \Big|\psi\big(N(b_j)\big)\Big|^2 \varphi\big(N(a_j)N(a_j)^*\big) + \Big|\varphi\big(N(a_j)\big)\Big|^2 \varphi\big((a_i a_i^*)^2\big) - \Big|\psi\big(N(b_j)\big)\Big|^2 \varphi\big(N(a_j)N(a_j)^*\big) \\
&- \Big|\varphi\big(N(a_j)\big)\Big|^2 \Big|\psi\big(N(b_j)\big)\Big|^2 \varphi\big((a_i a_i^*)^2\big) - \Big|\varphi\big(N(a_j)\big)\Big|^2 \\
&- \Big|\varphi\big(N(a_j)\big)\Big|^2 \Big|\psi\big(N(b_j)\big)\Big|^2 + 2\Big|\varphi\big(N(a_j)\big)\Big|^2 \Big|\psi\big(N(b_j)\big)\Big|^2 \\
=& \Big|\varphi\big(N(a_j)\big)\Big|^2 \varphi\big((a_i a_i^*)^2\big) - \Big|\varphi\big(N(a_j)\big)\Big|^2 \Big|\psi\big(N(b_j)\big)\Big|^2 \varphi\big((a_i a_i^*)^2\big) - \Big|\varphi\big(N(a_j)\big)\Big|^2 \\
&+ \Big|\varphi\big(N(a_j)\big)\Big|^2 \Big|\psi\big(N(b_j)\big)\Big|^2 \\
=& \Big|\varphi\big(N(a_j)\big)\Big|^2 \Big(\varphi\big((a_i a_i^*)^2\big) - 1\Big)\Big(1 - \Big|\psi\big(N(b_j)\big)\Big|^2\Big).
\end{aligned}$$

Since a_i is not unitary, this yields

$$0 = \Big|\varphi\big(N(a_j)\big)\Big|^2 \Big(1 - \Big|\psi\big(N(b_j)\big)\Big|^2\Big),$$

and since a_j is not unitary, by reversing the role of i and j, we similarly obtain that

$$0 = \Big|\varphi\big(N(a_i)\big)\Big|^2 \Big(1 - \Big|\psi\big(N(b_i)\big)\Big|^2\Big).$$

Consequently, we see that for $\iota = i, j$, it cannot be the case that both $\varphi\big(N(a_\iota)\big) \neq 0$ and $\varphi\big(N(b_\iota)\big) = 0$, thus

$$\text{if } (\varphi \otimes \psi)\big(N(a_\iota \otimes b_\iota)\big) = 0, \text{ then } \varphi\big(N(a_\iota)\big) = 0.$$

Furthermore, it clearly follows that

$$\text{if } (\varphi \otimes \psi)\big(N(a_\iota \otimes b_\iota)\big) \neq 0, \text{ then } \Big|\psi\big(N(b_\iota)\big)\Big| = 1,$$

which concludes the proof of the result by Remark 6.26. □

Lemma 6.32. *Let $i, j \in I$ be distinct, suppose that $a_i \otimes b_i$ is $*$-free from $a_j \otimes b_j$, that a_i is not unitary, and a_j, b_i, and b_j are all unitary. Then, the reformulated centering conditions hold for the couple $(a_i \otimes b_i, a_j \otimes b_j)$, and (a_i, a_j) is a dominating family.*

Proof. Using the same reasoning as in Lemma 6.31, we obtain that for every $*$-word $N \in \mathbb{C}\langle x, x^*\rangle$, one has

$$0 = \Big|\varphi\big(N(a_j)\big)\Big|^2 \Big(1 - \Big|\psi\big(N(b_j)\big)\Big|^2\Big).$$

(Note, however, that the above does not necessarily also hold for i instead of j, as a_j is unitary.) Therefore, as it was done in Lemma 6.31, we conclude that

$$\text{if } (\varphi \otimes \psi)\big(N(a_j \otimes b_j)\big) = 0, \text{ then } \varphi\big(N(a_j)\big) = 0, \tag{6.31}$$

and

$$\text{if } (\varphi \otimes \psi)\big(N(a_j \otimes b_j)\big) \neq 0, \text{ then } \Big|\psi\big(N(b_j)\big)\Big| = 1. \tag{6.32}$$

It now remains to prove that the same holds for $a_i \otimes b_i$.

We consider the two possible cases with respect to the expectation of a_j and b_j, namely,

(1) at least one of $\varphi(a_j)$ or $\psi(b_j)$ is not equal to zero; and

(2) $\varphi(a_j)$ and $\psi(b_j)$ are both equal to zero.

(1). Suppose that $\varphi(a_j) \neq 0$ or $\psi(b_j) \neq 0$. Define $N(x) = x$, and let M be an arbitrary $*$-word. Since a_j, b_i and b_j are unitary, then (6.10) reduces in this case to

$$\begin{aligned}
0 =& \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(b_j\big)\Big|^2 + \Big|\varphi\big(a_j\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2\varphi\big(M(a_i)M(a_i)^*\big) \\
&- \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2\Big|\psi\big(b_j\big)\Big|^2 \\
&- \Big|\varphi\big(a_j\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2\Big|\psi\big(b_j\big)\Big|^2\varphi\big(M(a_i)M(a_i)^*\big) \\
&- \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\varphi\big(a_j\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2 - \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\varphi\big(a_j\big)\Big|^2\Big|\psi\big(b_j\big)\Big|^2 \\
&+ 2\Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\varphi\big(a_j\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2\Big|\psi\big(b_j\big)\Big|^2. \qquad (6.33)
\end{aligned}$$

Suppose by contradiction that $\varphi\big(M(a_i)\big) \neq 0$ and $\psi\big(M(b_i)\big) = 0$. Then, the above reduces to

$$\begin{aligned}
0 =& \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(b_j\big)\Big|^2 - \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\varphi\big(a_j\big)\Big|^2\Big|\psi\big(b_j\big)\Big|^2 \\
=& \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(b_j\big)\Big|^2\Big(1 - \varphi(a_j)\Big).
\end{aligned}$$

Thus, it must be the case that $\Big|\psi\big(b_j\big)\Big|^2 = 0$ or $\Big(1 - \varphi(a_j)\Big) = 0$. If $\Big|\psi\big(b_j\big)\Big|^2 = 0$, then our assumption that $\varphi(a_j) \neq 0$ or $\psi(b_j) \neq 0$ implies that $\varphi(a_j) \neq 0$ and $\psi(b_j) = 0$, which contradicts equation (6.31) with $N(x) = x$. If $1 - \big|\varphi(a_j)\big| = 0$, then Proposition 6.25 implies that $a_j = \varphi(a_j) \cdot 1_{\mathcal{A}}$, and according to equation (6.32) it must be the case that b_j is a constant multiple of $1_{\mathcal{B}}$. However, this contradicts our assumption that $a_j \otimes b_j$ is not a constant multiple of the unit (see Remark 6.4). Therefore, $\varphi\big(M(a_i)\big) \neq 0$ and $\psi\big(M(b_i)\big) = 0$ leads to a contradiction, which implies that

$$\text{if } (\varphi \otimes \psi)\big(M(a_i \otimes b_i)\big) = 0, \text{ then } \varphi\big(M(a_i)\big) = 0. \qquad (6.34)$$

Suppose that $\varphi\big(M(a_i)\big), \psi\big(M(b_i)\big) \neq 0$. We consider the following two sub-cases of $\varphi(a_j) \neq 0$ or $\psi(b_j) \neq 0$, which we know are exhaustive thanks to equations (6.31) and (6.32):

(1.1) $\varphi(a_j) = 0$ and $\psi(b_j) \neq 0$; and

(1.2) $\varphi(a_j) \neq 0$ and $\big|\psi(b_j)\big| = 1$

(1.1). Suppose that $\varphi(a_j) = 0$ and $\psi(b_j) \neq 0$. Then, equation (6.33) reduces in this case to

$$\begin{aligned}
0 =& \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(b_j\big)\Big|^2 - \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2\Big|\psi\big(b_j\big)\Big|^2 \\
=& \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(b_j\big)\Big|^2\Big(1 - \Big|\psi\big(M(b_i)\big)\Big|^2\Big),
\end{aligned}$$

hence $\Big|\psi\big(M(b_i)\big)\Big| = 1$.

(1.2). Suppose that $\varphi(a_j) \neq 0$ and $\big|\psi\big(b_j\big)\big| = 1$. Then, equation (6.33) yields in this case

$$\begin{aligned}
0 =& \Big|\varphi\big(M(a_i)\big)\Big|^2 + \Big|\varphi\big(a_j\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2\varphi\big(M(a_i)M(a_i)^*\big) - \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2 \\
&- \Big|\varphi\big(a_j\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2\varphi\big(M(a_i)M(a_i)^*\big) - \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\varphi\big(a_j\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2 \\
&- \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\varphi\big(a_j\big)\Big|^2 + 2\Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\varphi\big(a_j\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2 \\
=& \Big|\varphi\big(M(a_i)\big)\Big|^2 - \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2 - \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\varphi\big(a_j\big)\Big|^2 \\
&+ \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\varphi\big(a_j\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2 \\
=& \Big|\varphi\big(M(a_i)\big)\Big|^2\Big(1 - \Big|\psi\big(M(b_i)\big)\Big|^2\Big)\Big(1 - \Big|\varphi\big(a_j\big)\Big|^2\Big),
\end{aligned}$$

which necessarily implies that $\Big|\psi\big(M(b_i)\big)\Big| = 1$, since $|\varphi(a_j)|^2 = 1$ would mean that $a_j \otimes b_j$ is a constant multiple of $1_{\mathcal{A}\otimes\mathcal{B}}$.

Combining the sub-cases (1.1) and (1.2), we see that

$$\text{if } (\varphi\otimes\psi)\big(M(a_i\otimes b_i)\big) \neq 0, \text{ then } \Big|\psi\big(M(b_i)\big)\Big| = 1,$$

which, when combined with (6.34), implies that the reformulated centering conditions hold for $(a_i\otimes b_i, a_j\otimes b_j)$ and that (a_i, a_j) is a dominating family in the case where $\varphi(a_j) \neq 0$ or $\psi(b_j) \neq 0$.

(2). Suppose that $\varphi(a_j) = \psi(b_j) = 0$. Since b_i is unitary and $\varphi(a_i a_i^*) = 1$, equation (6.15) with $N(x) = xx^*$ and $M \in \mathbb{C}\langle x, x^*\rangle$ arbitrary reduces in this case to

$$\begin{aligned}
0 =& \varphi\Big((a_i a_i^*)^2\Big)\Big|\varphi\big(M(a_i)\big)\Big|^2 + \varphi\Big(M(a_i)M(a_i)^*\Big)\Big|\psi\big(M(b_i)\big)\Big|^2 - \Big|\varphi\big(M(a_i)\big)\Big|^2 \\
&- \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2 - \varphi\Big(M(a_i)M(a_i)^*\Big)\Big|\psi\big(M(b_i)\big)\Big|^2 \\
&- \varphi\Big((a_i a_i^*)^2\Big)\Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2 + 2\Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2 \\
=& \varphi\Big((a_i a_i^*)^2\Big)\Big|\varphi\big(M(a_i)\big)\Big|^2 - \Big|\varphi\big(M(a_i)\big)\Big|^2 - \varphi\Big((a_i a_i^*)^2\Big)\Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2 \\
&+ \Big|\varphi\big(M(a_i)\big)\Big|^2\Big|\psi\big(M(b_i)\big)\Big|^2. \\
=& \Big|\varphi\big(M(a_i)\big)\Big|^2\Big(1 - \varphi\Big((a_i a_i^*)^2\Big)\Big)\Big(\Big|\psi\big(M(b_i)\big)\Big|^2 - 1\Big).
\end{aligned} \tag{6.35}$$

Since a_i is not unitary, Proposition 6.25 implies that $\varphi\Big((a_i a_i^*)^2\Big) \neq 1$. Consequently, we can divide both sides of (6.35) by $1 - \varphi\Big((a_i a_i^*)^2\Big)$ and obtain that

$$0 = \Big|\varphi\big(M(a_i)\big)\Big|^2\Big(\Big|\psi\big(M(b_i)\big)\Big|^2 - 1\Big).$$

From the above equation it clearly follows that

$$\text{if } (\varphi\otimes\psi)\big(M(a_i\otimes b_i)\big) = 0, \text{ then } \varphi\big(M(a_i)\big) = 0,$$

and

$$\text{if } (\varphi\otimes\psi)\big(M(a_i\otimes b_i)\big)\neq 0, \text{ then } \Big|\varphi\big(M(b_i)\big)\Big| = 1,$$

and hence the reformulated centering conditions hold for $(a_i\otimes b_i, a_j\otimes b_j)$ and (a_i, a_j) is a dominating family in the case where $\varphi(a_j)=\psi(b_j)=0$. □

Proposition 6.33. *Let $(\mathcal{A},*,\varphi)$, $(\mathcal{B},*,\psi)$, and the collections $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{a}\otimes\mathbf{b}$ be defined as in Problem 6.19, and suppose that one of $\mathbf{a}$ of $\mathbf{b}$ contains at least one variable that is not unitary. If the variables in $\mathbf{a}\otimes\mathbf{b}$ are $*$-free from each other, then $\mathbf{a}\otimes\mathbf{b}$ satisfies the reformulated centering conditions, and the family that contains a variable that is not unitary is a dominating family.*

Proof. Suppose without loss of generality that it is $\mathbf{a}$ that contains a variable that is not unitary, say a_i. Then, it follows from Proposition 6.30 that every variable in $\mathbf{b}$ is unitary.

Let $j\in I\setminus\{i\}$ be arbitrary. If a_j is not unitary, then it follows from Lemma 6.31 that for every $*$-word M and $\iota=i,j$,

(1) if $\varphi\otimes\psi\big(M(a_\iota\otimes b_\iota)\big)=0$, then $\varphi\big(M(a_\iota)\big)=0$; and

(2) if $\varphi\otimes\psi\big(M(a_\iota\otimes b_\iota)\big)\neq 0$, then $M(b_\iota)=\psi\big(M(b_\iota)\big)\cdot 1_{\mathcal{B}}$.

If a_j is unitary, then Lemma 6.32 implies the same. Since $j\neq i$ was arbitrary, we conclude that the above holds for all $\iota\in I$, as desired. □

Lemma 6.34. *Let $i,j\in I$ be distinct, suppose that $a_i\otimes b_i$ is $*$-free from $a_j\otimes b_j$, and that a_i, a_j, b_i, and b_j are all unitary. For $\iota=i,j$, define the sets*

$$A_\iota=\big\{n\in\mathbb{N}:\varphi(a_\iota^n)=0\big\} \text{ and } B_\iota=\big\{n\in\mathbb{N}:\psi(b_\iota^n)=0\big\}.$$

Then, one of the following holds:

(1) $A_i\supset B_i$ and $A_j\supset B_j$; or

(2) $A_i\subset B_i$ and $A_j\subset B_j$.

(That is, item (1) of the reformulated centering conditions holds for $(a_i\otimes b_i, a_j\otimes b_j)$.)

Proof. If $A_i=B_i$ and $A_j=B_j$, then the result trivially holds. Thus, we need only consider the case where at least one of $A_i\neq B_i$ or $A_j\neq B_j$ is true.

Assume without loss of generality that $A_i\not\subset B_i$, that is, there exists $p\in\mathbb{N}$ such that $\varphi(a_i^p)=0\neq\psi(b_i^p)$. Since every variable considered is unitary, equation (6.10) implies that

$$\begin{aligned}0=&|\varphi(a_i^m)|^2|\psi(b_j^n)|^2+|\varphi(a_j^n)|^2|\psi(b_i^m)|^2-|\varphi(a_i^m)|^2|\psi(b_i^m)|^2|\psi(b_j^n)|^2\\&-|\varphi(a_j^n)|^2|\psi(b_i^m)|^2|\psi(b_j^n)|^2-|\varphi(a_i^m)|^2|\varphi(a_j^n)|^2|\psi(b_i^m)|^2\\&-|\varphi(a_i^m)|^2|\varphi(a_j^n)|^2|\psi(b_j^n)|^2+2|\varphi(a_i^m)|^2|\varphi(a_j^n)|^2|\psi(b_i^m)|^2|\psi(b_j^n)|^2.\end{aligned}\tag{6.36}$$

for every $n,m\in\mathbb{N}$ (with $M,N\in\mathbb{C}\langle x,x^*\rangle$ defined as $M(x)=x^m$ and $N(x)=x^n$). Thus, applying equation (6.36) with $m=p$ implies that

$$0=|\varphi(a_j^n)|^2|\psi(b_i^p)|^2-|\varphi(a_j^n)|^2|\psi(b_i^p)|^2|\psi(b_j^n)|^2=|\varphi(a_j^n)|^2|\psi(b_i^p)|^2\big(1-|\psi(b_j^n)|^2\big)\tag{6.37}$$

holds for every $n\in\mathbb{N}$. Since $|\psi(b_i^p)|^2\neq 0$, then $0=|\varphi(a_j^n)|^2\big(1-|\psi(b_j^n)|^2\big)$, from which we conclude that $A_j\supset B_j$.

We now prove that $A_i\supset B_i$ as well. To do so, we separate the proof into the following cases:

(1) $A_j \neq B_j$;

(2) $A_j = B_j \neq \varnothing$; and

(3) $A_j = B_j = \varnothing$.

(1). Suppose that $A_j \neq B_j$. Then, since $A_j \supset B_j$, it must be the case that $A_j \not\subset B_j$, that is, there exists $p \in \mathbb{N}$ such that $\varphi(a_j^p) = 0 \neq \psi(b_j^p)$. By applying (6.36) and (6.37) similarly to how it was done in the preceding paragraphs of this proof (but interchanging i and j), we obtain that $A_i \supset B_i$.

(2). Suppose that $A_j = B_j \neq \varnothing$, which implies in particular that there exists $p \in \mathbb{Z}$ such that $\varphi(a_j^p) = \psi(b_j^p) = 0$. Assume by contradiction that $A_i \not\supset B_i$. Since it is also the case that $A_i \not\subset B_i$, then neither A_i nor B_i is the empty set. Therefore, there exists $n, m \in \mathbb{Z} \setminus \{0\}$ such that $\varphi(a_i^n) = 0 \neq \psi(b_i^n)$ and $\varphi(a_i^m) \neq 0 = \psi(b_i^m)$. Since a_j^p and b_j^p are unitary and centered, equation (6.15) with $M(x) = x^m$ and $N(x) = x^n$ reduces in this case to

$$\begin{aligned}
0 =& \left|\varphi\left(a_i^m\right)\right|^2\left|\psi\left(b_i^n\right)\right|^2 + \left|\varphi\left(a_i^n\right)\right|^2\left|\psi\left(b_i^m\right)\right|^2 - \left|\varphi\left(a_i^m\right)\right|^2\left|\varphi\left(a_i^n\right)\right|^2\left|\psi\left(b_i^n\right)\right|^2 \\
&- \left|\varphi\left(a_i^m\right)\right|^2\left|\varphi\left(a_i^n\right)\right|^2\left|\psi\left(b_i^m\right)\right|^2 - \left|\varphi\left(a_i^n\right)\right|^2\left|\psi\left(b_i^m\right)\right|^2\left|\psi\left(b_i^n\right)\right|^2 \\
&- \left|\varphi\left(a_i^m\right)\right|^2\left|\psi\left(b_i^m\right)\right|^2\left|\psi\left(b_i^n\right)\right|^2 + 2\left|\varphi\left(a_i^m\right)\right|^2\left|\varphi\left(a_i^n\right)\right|^2\left|\psi\left(b_i^m\right)\right|^2\left|\psi\left(b_i^n\right)\right|^2 \\
=& \left|\varphi\left(a_i^m\right)\right|^2\left|\psi\left(b_i^n\right)\right|^2,
\end{aligned}$$

from which we conclude that either $\left|\varphi\left(a_i^m\right)\right|^2 = 0$ or $\left|\psi\left(b_i^n\right)\right|^2 = 0$, which is a contradiction. Thus, it must be the case that $A_i \supset B_i$.

(3). Suppose that $A_j = B_j = \varnothing$. Suppose by contradiction that $A_i \not\supset B_i$, which, as argued in the previous case, implies that there exists $p, q \in \mathbb{Z} \setminus \{0\}$ such that $\varphi(a_i^p) = 0 \neq \psi(b_i^p)$ and $\varphi(a_i^q) \neq 0 = \psi(b_i^q)$. Applying equation (6.36) with $m = p$ and $n = 1$ reduces to

$$0 = |\varphi(a_j)|^2|\psi(b_i^p)|^2 - |\varphi(a_j)|^2|\psi(b_i^p)|^2|\psi(b_j)|^2 = |\varphi(a_j)|^2|\psi(b_i^p)|^2\left(1 - |\psi(b_j)|^2\right),$$

which, since $|\varphi(a_j)|^2, |\psi(b_i^p)|^2 \neq 0$, implies that $|\psi(b_j)|^2 = 1$, hence b_j is a constant multiple of $1_{\mathcal{B}}$ by Proposition 6.25. By applying equation (6.36) with $m = q$ and $n = 1$, we similarly obtain that a_j is a constant multiple of $1_{\mathcal{A}}$. Thus, $a_j \otimes b_j$ is a contant multiple of the unit, contradicting our assumption that such is not the case (see Remark 6.4). Therefore, $A_i \supset B_i$.

Using the same arguments and case by case analysis (but interchanging i for j, or a_i for b_i and a_j for b_j), it can be shown that

(1) if $A_j \not\subset B_j$, then $A_i \supset B_i$ and $A_j \supset B_j$;

(2) if $A_i \not\supset B_i$, then $A_i \subset B_i$ and $A_j \subset B_j$; and

(3) if $A_j \not\supset B_j$, then $A_i \subset B_i$ and $A_j \subset B_j$,

which concludes the proof of the lemma. □

Lemma 6.35. *Let $i, j \in I$ be distinct, suppose that $a_i \otimes b_i$ is $*$-free from $a_j \otimes b_j$, and that a_i, a_j, b_i, and b_j are all unitary. Then, the couple $(a_i \otimes b_i, a_j \otimes b_j)$ satisfies the reformulated centering conditions.*

Proof. For $\iota = i, j$, define the sets

$$A_\iota = \big\{n \in \mathbb{N} : \varphi(a_\iota^n) = 0\big\} \text{ and } B_\iota = \big\{n \in \mathbb{N} : \psi(b_\iota^n) = 0\big\}.$$

We separate this proof in the three following cases (which we know are exhaustive thanks to Lemma 6.34):

(1) $A_i = B_i$ and $A_j = B_j$;
(2) $A_\iota \supset B_\iota$ for $\iota = i, j$; and $A_i \neq B_i$ or $A_j \neq B_j$; and
(3) $A_\iota \subset B_\iota$ for $\iota = i, j$; and $A_i \neq B_i$ or $A_j \neq B_j$.

(1). Suppose that $A_i = B_i$ and $A_j = B_j$. We will separate the proof of this case in the following sub-cases:

(1.1) $A_i = B_i \neq \varnothing$ and $A_j = B_j \neq \varnothing$;
(1.2) $A_i = B_i = \varnothing$ and $A_j = B_j \neq \varnothing$;
(1.3) $A_i = B_i \neq \varnothing$ and $A_j = B_j = \varnothing$; and
(1.4) $A_i = B_i = A_j = B_j = \varnothing$.

(1.1). Suppose that $A_i = B_i \neq \varnothing$ and $A_j = B_j \neq \varnothing$, that is, there exists $p, q \in \mathbb{N}$ such that $\varphi(a_i^p) = \psi(b_i^p) = \varphi(a_j^q) = \psi(b_j^q) = 0$.

Suppose that $n \in \mathbb{N}$ is such that $n \notin A_i, B_i$, that is, $(\varphi \otimes \psi)\big((a_i \otimes b_i)^n\big) \neq 0$. Then, applying equation (6.16) with $N(x) = M(x) = x^n$ and a_j^q and b_j^q as the centered unitary elements yields

$$\begin{aligned} 0 =& \big|\varphi(a_i^n)\big|^2\big|\psi(b_i^n)\big|^2 - \big|\varphi(a_i^n)\big|^4\big|\psi(b_i^n)\big|^2 - \big|\varphi(a_i^n)\big|^2\big|\psi(b_i^n)\big|^4 + \big|\varphi(a_i^n)\big|^4\big|\psi(b_i^n)\big|^4 \\ =& \Big(\big|\varphi(a_i^n)\big|^4 - \big|\varphi(a_i^n)\big|^2\Big)\Big(\big|\psi(b_i^n)\big|^4 - \big|\psi(b_i^n)\big|^2\Big). \end{aligned} \tag{6.38}$$

Since $\varphi(a_i^n), \psi(b_i^n) \neq 0$, then $\big|\varphi(a_i^n)\big| = 1$ or $\big|\psi(b_i^n)\big| = 1$.

Suppose by contradiction that there exists $m, n \in \mathbb{N}$ such that $m \neq n$, $\big|\varphi(a_i^m)\big| = 1$ and $\big|\psi(b_i^m)\big| \neq 1$, and $\big|\varphi(a_i^n)\big| \neq 1$ and $\big|\psi(b_i^n)\big| = 1$. Then, applying equation (6.15) with $M(x) = x^m$, $N(x) = x^n$, and a_j^q and b_j^q as the centered unitary element yields

$$\begin{aligned} 0 =& 1 + \big|\varphi(a_i^n)\big|^2\big|\psi(b_i^m)\big|^2 - \big|\varphi(a_i^n)\big|^2 - \big|\varphi(a_i^n)\big|^2\big|\psi(b_i^m)\big|^2 - \big|\varphi(a_i^n)\big|^2\big|\psi(b_i^m)\big|^2 \\ & - \big|\psi(b_i^m)\big|^2 + 2\big|\varphi(a_i^n)\big|^2\big|\psi(b_i^m)\big| \\ =& 1 - \big|\varphi(a_i^n)\big|^2 - \big|\psi(b_i^m)\big|^2 + \big|\varphi(a_i^n)\big|^2\big|\psi(b_i^m)\big|^2 \\ =& \Big(\big|\varphi(a_i^n)\big|^2 - 1\Big)\Big(\big|\psi(b_i^m)\big|^2 - 1\Big), \end{aligned} \tag{6.39}$$

which is a contradiction. Therefore, we conclude that for every $n \in \mathbb{N}$, $n \notin A_i, B_i$ implies that $\big|\varphi(a_i^n)\big| = 1$; or that for every $n \in \mathbb{N}$, $n \notin A_i, B_i$ implies that $\big|\psi(b_i^n)\big| = 1$.

By using the same arguments but replacing i for j, (i.e., apply equations (6.16) and (6.15) with a_i^p and b_i^p as the centered unitary elements), we can conclude that for every $n \in \mathbb{N}$, $n \notin A_j, B_j$ implies that $\big|\varphi(a_j^n)\big| = 1$; or that for every $n \in \mathbb{N}$, $n \notin A_j, B_j$ implies that $\big|\psi(b_j^n)\big| = 1$. Thus, to prove that the desired result holds in this sub-case, it only remains to prove that it cannot happen that $\big|\varphi(a_i^m)\big| = 1$, $\big|\psi(b_i^m)\big| \neq 1$, $\big|\varphi(a_j^n)\big| \neq 1$, and $\big|\psi(b_j^n)\big| = 1$ for some $m, n \in \mathbb{N}$.

Suppose by contradiction that there exists $m, n \in \mathbb{N}$ such that $\big|\varphi(a_i^m)\big| = 1$, $\big|\psi(b_i^m)\big| \neq 1$, $\big|\varphi(a_j^n)\big| \neq 1$, and $\big|\psi(b_j^n)\big| = 1$. Then, equation (6.10) with $M(x) = x^m$ and $N(x) = x^n$ implies

that

$$\begin{aligned}0 =&1 + \left|\varphi(a_j^n)\right|^2\left|\psi(b_i^m)\right|^2 - \left|\psi(b_i^m)\right|^2 - \left|\varphi(a_j^n)\right|^2\left|\psi(b_i^m)\right|^2 - \left|\varphi(a_j^n)\right|^2\left|\psi(b_i^m)\right|^2\\ &- \left|\varphi(a_j^n)\right|^2 + 2\left|\varphi(a_j^n)\right|^2\left|\psi(b_i^m)\right|^2\\ =&1 - \left|\psi(b_i^m)\right|^2 - \left|\varphi(a_j^n)\right|^2 + \left|\varphi(a_j^n)\right|^2\left|\psi(b_i^m)\right|^2\\ =&\Big(\left|\varphi(a_j^n)\right|^2 - 1\Big)\Big(\left|\psi(b_i^m)\right|^2 - 1\Big),\end{aligned}$$

which is a contradiction.

Therefore, it must be the case that either

$$\text{if } \varphi\otimes\psi\big(a_\iota^n\otimes b_\iota^n\big) \neq 0, \text{ then } a_\iota^n = \varphi(a_\iota^n)\cdot 1_{\mathcal{A}} \text{ for } \iota = i, j; \tag{6.40}$$

or

$$\text{if } \varphi\otimes\psi\big(a_\iota^n\otimes b_\iota^n\big) \neq 0, \text{ then } b_\iota^n = \psi(b_\iota^n)\cdot 1_{\mathcal{B}} \text{ for } \iota = i, j, \tag{6.41}$$

which, when combined with the fact that $A_\iota = B_\iota$ for $\iota = 1, 2$, implies that $(a_i\otimes b_i, a_j\otimes b_j)$ satisfies the reformulated centering conditions.

(1.2). Suppose that $A_i = B_i = \varnothing$ and $A_j = B_j \neq \varnothing$. Using the same arguments as in the previous case (i.e., equations (6.38) and (6.39) in case (1.1)), the fact that $A_j = B_j \neq \varnothing$ can be used to show that for every $n\in\mathbb{N}$, $n\notin A_i, B_i$ implies that $\left|\varphi(a_i^n)\right| = 1$; or that for every $n\in\mathbb{N}$, $n\notin A_i, B_i$ implies that $\left|\psi(b_i^n)\right| = 1$. We now show that the same holds for j.

Let $n\notin A_j, B_j$, that is, $\varphi(a_j^n), \psi(b_j^n)\neq 0$. Then, applying equation (6.10) with $M(x) = x$ and $N(x) = x^n$ implies that

$$\begin{aligned}0 =&\left|\varphi(a_i)\right|^2\left|\psi(b_j^n)\right|^2 + \left|\varphi(a_j^n)\right|^2\left|\psi(b_i)\right|^2 - \left|\varphi(a_i)\right|^2\left|\psi(b_i)\right|^2\left|\psi(b_j^n)\right|^2\\ &- \left|\varphi(a_j^n)\right|^2\left|\psi(b_i)\right|^2\left|\psi(b_j^n)\right|^2 - \left|\varphi(a_i)\right|^2\left|\varphi(a_j^n)\right|^2\left|\psi(b_i)\right|^2\\ &- \left|\varphi(a_i)\right|^2\left|\varphi(a_j^n)\right|^2\left|\psi(b_j^n)\right|^2 + 2\left|\varphi(a_i)\right|^2\left|\varphi(a_j^n)\right|^2\left|\psi(b_i)\right|^2\left|\psi(b_j^n)\right|^2.\end{aligned}$$

Given that $\varphi(a_i), \psi(b_i)\neq 0$, it must be the case that $\left|\varphi(a_i)\right| = 1$ or $\left|\psi(b_i)\right| = 1$. Suppose that $\left|\varphi(a_i)\right| = 1$. Then, the above reduces to

$$\begin{aligned}0 =&\left|\psi(b_j^n)\right|^2 + \left|\varphi(a_j^n)\right|^2\left|\psi(b_i)\right|^2 - \left|\psi(b_i)\right|^2\left|\psi(b_j^n)\right|^2 - \left|\varphi(a_j^n)\right|^2\left|\psi(b_i)\right|^2\left|\psi(b_j^n)\right|^2\\ &- \left|\varphi(a_j^n)\right|^2\left|\psi(b_i)\right|^2 - \left|\varphi(a_j^n)\right|^2\left|\psi(b_j^n)\right|^2 + 2\left|\varphi(a_j^n)\right|^2\left|\psi(b_i)\right|^2\left|\psi(b_j^n)\right|^2\\ =&\left|\psi(b_j^n)\right|^2 - \left|\psi(b_i)\right|^2\left|\psi(b_j^n)\right|^2 - \left|\varphi(a_j^n)\right|^2\left|\psi(b_j^n)\right|^2 + \left|\varphi(a_j^n)\right|^2\left|\psi(b_i)\right|^2\left|\psi(b_j^n)\right|^2\\ =&\left|\psi(b_j^n)\right|^2\Big(1 - \left|\psi(b_i)\right|^2 - \left|\varphi(a_j^n)\right|^2 + \left|\varphi(a_j^n)\right|^2\left|\psi(b_i)\right|^2\Big)\\ =&\left|\psi(b_j^n)\right|^2\Big(1 - \left|\psi(b_i)\right|^2\Big)\Big(1 - \left|\varphi(a_j^n)\right|^2\Big).\end{aligned}$$

Given that $\left|\psi(b_i)\right| \neq 1$ (as the opposite would imply that $a_i\otimes b_i$ is a constant multiple of the unit vector, contradicting Remark 6.4), we conclude that $\left|\varphi(a_j^n)\right| = 1$. Similarly, if it is assumed that $\left|\psi(b_i)\right| = 1$, we can conclude that $\left|\varphi(b_j^n)\right| = 1$ for every $n\notin A_j, B_j$. Therefore, we conclude that either (6.40) or (6.41) also holds in this sub-case.

(1.3). Suppose that $A_i = B_i \neq \varnothing$ and $A_j = B_j = \varnothing$. The proof of this sub-case follows from that of sub-case (1.2) by interchanging i and j.

(1.4). Suppose that $A_i = B_i = A_j = B_j = \varnothing$. According to (6.10) with $M(x) = N(x) = x$,

$$\begin{aligned}0 =&\big|\varphi(a_i)\big|^2\big|\psi(b_j)\big|^2 + \big|\varphi(a_j)\big|^2\big|\psi(b_i)\big|^2 - \big|\varphi(a_i)\big|^2\big|\psi(b_i)\big|^2\big|\psi(b_j)\big|^2\\ &- \big|\varphi(a_j)\big|^2\big|\psi(b_i)\big|^2\big|\psi(b_j)\big|^2 - \big|\varphi(a_i)\big|^2\big|\varphi(a_j)\big|^2\big|\psi(b_i)\big|^2\\ &- \big|\varphi(a_i)\big|^2\big|\varphi(a_j)\big|^2\big|\psi(b_j)\big|^2 + 2\big|\varphi(a_i)\big|^2\big|\varphi(a_j)\big|^2\big|\psi(b_i)\big|^2\big|\psi(b_j)\big|^2.\end{aligned}$$

In order to simplify computations, let us denote $x = \big|\varphi(a_i)\big|^2$, $y = \big|\psi(b_i)\big|^2$, $w = \big|\varphi(a_j)\big|^2$, and $z = \big|\psi(b_j)\big|^2$, which yields

$$0 = xz + wy - xyz - wyz - xwy - xwz + 2xwyz.$$

This can be rearranged into

$$\begin{aligned}0 =&(xz - xyz - xwz + xwyz) + (wy - wyz - xwy + xwyz)\\ =&xz(1 - y - w + yw) + wy(1 - z - x + xz)\\ =&xz(1 - y)(1 - w) + wy(1 - z)(1 - x).\end{aligned} \tag{6.42}$$

According to equation (6.27), we know that $0 < x, y, w, z \leqslant 1$. Therefore,

$$xz(1 - y)(1 - w), wy(1 - z)(1 - x) \geqslant 0,$$

which, when combined with (6.42), implies that

$$(1 - y)(1 - w) = 0 \text{ and } (1 - z)(1 - x) = 0.$$

Given that neither $a_i \otimes b_i$ nor $a_j \otimes b_j$ can be a constant multiple of the unit vector (Remark 6.4), neither $x = y = 1$ nor $w = z = 1$ can be true. Therefore, it must be the case that $x = w = 1$ or $y = z = 1$. This implies that $a_i = \varphi(a_i) \cdot 1_{\mathcal{A}}$ and $a_j = \varphi(a_j) \cdot 1_{\mathcal{A}}$; or $b_i = \psi(b_i) \cdot 1_{\mathcal{B}}$ and $b_j = \psi(b_j) \cdot 1_{\mathcal{B}}$, from which we conclude that (6.40) and (6.41) both hold in this case.

(2). Suppose that $A_\iota \supset B_\iota$ for $\iota = i, j$; and that $A_i \neq B_i$ or $A_j \neq B_j$. We separate the proof of this case in the following sub-cases:

(2.1) $A_i \not\subset B_i$ and $A_j \not\subset B_j$;

(2.2) $A_i \not\subset B_i$ and $A_j = B_j$; and

(2.3) $A_i = B_i$ and $A_j \not\subset B_j$.

(2.1). Suppose that $A_i \not\subset B_i$ and $A_j \not\subset B_j$, that is, there exists $p, q \in \mathbb{N}$ such that $\varphi(a_i^p) = 0$, $\psi(b_i^p) \neq 0$, $\varphi(a_j^q) = 0$, and $\psi(b_j^q) \neq 0$. Then, as it was done in Lemma 6.34, equations (6.36) and (6.37) imply that $0 = |\varphi(a_j^n)|^2\big(1 - |\psi(b_j^n)|^2\big)$ for every $n \in \mathbb{N}$. Thus, if $n \notin A_j, B_j$, it must be the case that $|\psi(b_j^n)| = 1$. Using the same argument but interchanging i for j, it can be shown that $n \notin A_i, B_i$ implies that $|\psi(b_i^n)| = 1$, and thus the reformulated centering conditions hold with (a_i, a_j) as a dominating family.

(2.2). Suppose that $A_i \not\subset B_i$ and $A_j = B_j$, that is, there exists $p \in \mathbb{N}$ such that $\varphi(a_i^p) = 0$ and $\psi(b_i^p) \neq 0$. Using the same argument as in case (2.1) above (i.e., equations (6.36) and (6.37)), we can show that $n \notin A_j, B_j$ implies that $|\psi(b_j^n)| = 1$.

We must now prove that the same is true for i. Consider the following sub-sub-cases:

(2.2.1) Suppose that $A_j = B_j = \varnothing$. Given that, as we have just shown, $n \notin A_j, B_j$ implies that $|\psi(b_j^n)| = 1$, it then follows that $|\psi(b_j)| = 1$. If it were also the case that $|\varphi(a_j)| = 1$, then $a_j \otimes b_j$ would be a constant multiple of the unit vector, which we have assumed is

not the case (Remark 6.4). Therefore, $|\varphi(a_j)| \neq 1$. Let $m \in \mathbb{N}$ be such that $m \notin A_i, B_i$. Then, equation (6.10) with $M(x) = x^m$ and $N(x) = x$ implies that

$$\begin{aligned}0 =&\big|\varphi(a_i^m)\big|^2\big|\psi(b_j)\big|^2 + \big|\varphi(a_j)\big|^2\big|\psi(b_i^m)\big|^2 - \big|\varphi(a_i^m)\big|^2\big|\psi(b_i^m)\big|^2\big|\psi(b_j)\big|^2\\ &- \big|\varphi(a_j)\big|^2\big|\psi(b_i^m)\big|^2\big|\psi(b_j)\big|^2 - \big|\varphi(a_i^m)\big|^2\big|\varphi(a_j)\big|^2\big|\psi(b_i^m)\big|^2\\ &- \big|\varphi(a_i^m)\big|^2\big|\varphi(a_j)\big|^2\big|\psi(b_j)\big|^2 + 2\big|\varphi(a_i^m)\big|^2\big|\varphi(a_j)\big|^2\big|\psi(b_i^m)\big|^2\big|\psi(b_j)\big|^2.\end{aligned}$$

Since $|\psi(b_j)| = 1$, the above reduces to

$$\begin{aligned}0 =&\big|\varphi(a_i^m)\big|^2 + \big|\varphi(a_j)\big|^2\big|\psi(b_i^m)\big|^2 - \big|\varphi(a_i^m)\big|^2\big|\psi(b_i^m)\big|^2 - \big|\varphi(a_j)\big|^2\big|\psi(b_i^m)\big|^2\\ &- \big|\varphi(a_i^m)\big|^2\big|\varphi(a_j)\big|^2\big|\psi(b_i^m)\big|^2 - \big|\varphi(a_i^m)\big|^2\big|\varphi(a_j)\big|^2\\ &+ 2\big|\varphi(a_i^m)\big|^2\big|\varphi(a_j)\big|^2\big|\psi(b_i^m)\big|^2\\ =&\big|\varphi(a_i^m)\big|^2 - \big|\varphi(a_i^m)\big|^2\big|\psi(b_i^m)\big|^2 - \big|\varphi(a_i^m)\big|^2\big|\varphi(a_j)\big|^2\\ &+ \big|\varphi(a_i^m)\big|^2\big|\varphi(a_j)\big|^2\big|\psi(b_i^m)\big|^2\\ =&\big|\varphi(a_i^m)\big|^2\Big(1 - \big|\psi(b_i^m)\big|^2 - \big|\varphi(a_j)\big|^2 + \big|\varphi(a_j)\big|^2\big|\psi(b_i^m)\big|^2\Big)\\ =&\big|\varphi(a_i^m)\big|^2\Big(1 - \big|\psi(b_i^m)\big|^2\Big)\Big(1 - \big|\varphi(a_j)\big|^2\Big).\end{aligned}$$

Given that $\varphi(a_i^m) \neq 0$ and $\big|\varphi(a_j)\big| \neq 1$, it must be the case that $\big|\psi(b_i^m)\big| = 1$. Therefore, $n \notin A_i, B_i$ implies that $\big|\psi(b_i^n)\big| = 1$.

(2.2.2) Suppose that $A_j = B_j \neq \varnothing$, that is, there exists $q \in \mathbb{N}$ such that $\varphi(a_j^q) = \psi(b_j^q) = 0$. According to equation (6.16) with $M(x) = N(x) = x^n$ and a_j^q and b_j^q as the centered unitary elements, one has

$$\begin{aligned}0 =&\big|\varphi(a_i^n)\big|^2\big|\psi(b_i^n)\big|^2 - \big|\varphi(a_i^n)\big|^4\big|\psi(b_i^n)\big|^2 - \big|\varphi(a_i^n)\big|^2\big|\psi(b_i^n)\big|^4\\ &+ \big|\varphi(a_i^n)\big|^4\big|\psi(b_i^n)\big|^4\\ =&\Big(\big|\varphi(a_i^n)\big|^4 - \big|\varphi(a_i^n)\big|^2\Big)\Big(\big|\psi(b_i^n)\big|^4 - \big|\psi(b_i^n)\big|^2\Big).\end{aligned}$$

Thus, if $n \notin A_i, B_i$, then $|\varphi(a_i^n)| = 1$ or $|\psi(b_i^n)| = 1$. Suppose by contradiction that there exists $m \notin A_i, B_i$ such that $|\varphi(a_i^m)| = 1$, and $|\psi(b_i^m)| \neq 1$. Then, applying equation (6.15) with $M(x) = x^m$ and $N(x) = x^p$ (recall that $\varphi(a_i^p) = 0$ and $\psi(b_i^p) \neq 0$) and a_j^q and b_j^q as the centered unitary elements, one has

$$\begin{aligned}0 =&\big|\varphi(a_i^m)\big|^2\big|\psi(b_i^p)\big|^2 + \big|\varphi(a_i^p)\big|^2\big|\psi(b_i^m)\big|^2 - \big|\varphi(a_i^m)\big|^2\big|\varphi(a_i^p)\big|^2\big|\psi(b_i^p)\big|^2\\ &- \big|\varphi(a_i^m)\big|^2\big|\varphi(a_i^p)\big|^2\big|\psi(b_i^m)\big|^2 - \big|\varphi(a_i^p)\big|^2\big|\psi(b_i^m)\big|^2\big|\psi(b_i^p)\big|^2\\ &- \big|\varphi(a_i^m)\big|^2\big|\psi(b_i^m)\big|^2\big|\psi(b_i^p)\big|^2 + 2\big|\varphi(a_i^m)\big|^2\big|\varphi(a_i^p)\big|^2\big|\psi(b_i^m)\big|^2\big|\psi(b_i^p)\big|^2.\end{aligned}$$

Given that $|\varphi(a_i^m)| = 1$ and $\varphi(a_i^p) = 0$, the above reduces to

$$0 = \big|\psi(b_i^p)\big|^2 - \big|\psi(b_i^m)\big|^2\big|\psi(b_i^p)\big|^2 = \big|\psi(b_i^p)\big|^2\big(1 - \big|\psi(b_i^m)\big|^2\big).$$

This contradicts that $\psi(b_i^p) \neq 0$ and $\big|\psi(b_i^m)\big| \neq 1$. Therefore, we conclude that if $n \notin A_i, B_i$, then $\big|\psi(b_i^n)\big| = 1$.

Given that the above two sub-sub-cases are exhaustive, we conclude that $n \notin A_i, B_i$ implies that $\left|\psi(b_i^n)\right| = 1$, hence the reformulated centering conditions hold with (a_i, a_j) as a dominating family.

(2.3). Suppose that $A_i = B_i$ and $A_j \not\subset B_j$. By using the same arguments as in case (2.2) but interchanging i and j, it can be shown that $n \notin A_i, B_i$ implies that $|\psi(b_i^n)| = 1$, and that $n \notin A_j, B_j$ implies that $|\psi(b_j^n)| = 1$, and thus the reformulated centering conditions hold with (a_i, a_j) as a dominating family.

(3). Suppose that $A_\iota \subset B_\iota$; and $A_i \neq B_i$ or $A_j \neq B_j$ for $\iota = i, j$. The proof of this case follows from that of case (2) by interchanging a_i for b_i, and a_j for b_j, which will yield that the reformulated centering conditions hold with (b_i, b_j) as a dominating family. □

Thus, we obtain the following Proposition:

Proposition 6.36. *Let $(\mathcal{A}, *, \varphi)$, $(\mathcal{B}, *, \psi)$, and the collections $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{a} \otimes \mathbf{b}$ be defined as in Problem 6.19, and suppose that $\mathbf{a}$ and $\mathbf{b}$ both contain unitary variables only. If the variables in $\mathbf{a} \otimes \mathbf{b}$ are $*$-free from each other, then $\mathbf{a} \otimes \mathbf{b}$ satisfies the reformulated centering conditions.*

Proof. For every $i \in I$, define the sets

$$A_i = \left\{n \in \mathbb{N} : \varphi(a_i^n) = 0\right\} \text{ and } B_i = \left\{n \in \mathbb{N} : \psi(b_i^n) = 0\right\}.$$

Thanks to Lemma 6.34, we know that one of the three following cases must occur:

(1) $A_i = B_i$ for all $i \in I$;

(2) $A_i \supset B_i$ for all $i \in I$, and $A_j \neq B_j$ for some $j \in J$; and

(3) $A_i \subset B_i$ for all $i \in I$, and $A_j \neq B_j$ for some $j \in J$.

(Indeed, $A_i \not\supset B_i$ and $A_j \not\subset B_j$ for some $i, j \in I$ would clearly contradict Lemma 6.34.)

(1). Suppose that $A_i = B_i$ for all $i \in I$. Let $i \in I$ be arbitrary. Then, according to Lemma 6.35, it follows that for every $j \in I \setminus \{i\}$, either

$$\text{for } \iota = i, j \text{ and for all } n \in \mathbb{N}, \text{ if } \varphi \otimes \psi\left(a_\iota^n \otimes b_\iota^n\right) \neq 0, \text{ then } a_\iota^n = \varphi(a_\iota^n) \cdot 1_{\mathcal{A}};$$

or

$$\text{or } \iota = i, j \text{ and for all } n \in \mathbb{N}, \text{ if } \varphi \otimes \psi\left(a_\iota^n \otimes b_\iota^n\right) \neq 0, \text{ then } b_\iota^n = \psi(b_\iota^n) \cdot 1_{\mathcal{B}}.$$

Suppose by contradiction that there exists $n \in \mathbb{N}$ and two distinct indices $j(1), j(2) \in I \setminus \{i\}$ such that $\left|\varphi\left(a_{j(1)}^n\right)\right| = 1$ and $\left|\varphi\left(b_{j(1)}^n\right)\right| \neq 1$; and $\left|\varphi\left(a_{j(2)}^n\right)\right| \neq 1$ and $\left|\varphi\left(b_{j(2)}^n\right)\right| = 1$. Then,

according to equation (6.10) with $M(x) = N(x) = x^n$, one has

$$\begin{aligned}
0 =&\left|\varphi\left(a_{j(1)}^n\right)\right|^2\left|\psi\left(b_{j(2)}^n\right)\right|^2 + \left|\varphi\left(a_{j(2)}^n\right)\right|^2\left|\psi\left(b_{j(1)}^n\right)\right|^2 - \left|\varphi\left(a_{j(1)}^n\right)\right|^2\left|\psi\left(b_{j(1)}^n\right)\right|^2\left|\psi\left(b_{j(2)}^n\right)\right|^2 \\
&- \left|\varphi\left(a_{j(2)}^n\right)\right|^2\left|\psi\left(b_{j(1)}^n\right)\right|^2\left|\psi\left(b_{j(2)}^n\right)\right|^2 - \left|\varphi\left(a_{j(1)}^n\right)\right|^2\left|\varphi\left(a_{j(2)}^n\right)\right|^2\left|\psi\left(b_{j(1)}^n\right)\right|^2 \\
&- \left|\varphi\left(a_{j(1)}^n\right)\right|^2\left|\varphi\left(a_{j(2)}^n\right)\right|^2\left|\psi\left(b_{j(2)}^n\right)\right|^2 + 2\left|\varphi\left(a_{j(1)}^n\right)\right|^2\left|\varphi\left(a_{j(2)}^n\right)\right|^2\left|\psi\left(b_{j(1)}^n\right)\right|^2\left|\psi\left(b_{j(2)}^n\right)\right|^2 \\
=&1 + \left|\varphi\left(a_{j(2)}^n\right)\right|^2\left|\psi\left(b_{j(1)}^n\right)\right|^2 - \left|\psi\left(b_{j(1)}^n\right)\right|^2 - \left|\varphi\left(a_{j(2)}^n\right)\right|^2\left|\psi\left(b_{j(1)}^n\right)\right|^2 - \left|\varphi\left(a_{j(2)}^n\right)\right|^2\left|\psi\left(b_{j(1)}^n\right)\right|^2 \\
&- \left|\varphi\left(a_{j(2)}^n\right)\right|^2 + 2\left|\varphi\left(a_{j(2)}^n\right)\right|^2\left|\psi\left(b_{j(1)}^n\right)\right|^2 \\
=&1 - \left|\psi\left(b_{j(1)}^n\right)\right|^2 - \left|\varphi\left(a_{j(2)}^n\right)\right|^2 + \left|\varphi\left(a_{j(2)}^n\right)\right|^2\left|\psi\left(b_{j(1)}^n\right)\right|^2 \\
=&\left(1 - \left|\psi\left(b_{j(1)}^n\right)\right|^2\right)\left(1 - \left|\varphi\left(a_{j(2)}^n\right)\right|^2\right),
\end{aligned}$$

which is a contradiction. Therefore, it follows that either

$$\text{for all } n \in \mathbb{N} \text{ and } i \in I, \text{ if } \varphi \otimes \psi\left(a_i^n \otimes b_i^n\right) \neq 0, \text{ then } a_i^n = \varphi(a_i^n) \cdot 1_{\mathcal{A}};$$

or

$$\text{for all } n \in \mathbb{N} \text{ and } i \in I, \text{ if } \varphi \otimes \psi\left(a_i^n \otimes b_i^n\right) \neq 0, \text{ then } b_i^n = \psi(b_i^n) \cdot 1_{\mathcal{B}},$$

which, when combined with the fact that $A_i = B_i$ for every $i \in I$, implies the desired conclusion.

(2). Suppose that $A_i \supset B_i$ for all $i \in I$, and that there is an index $j \in I$ such that $A_j \not\subset B_j$. Then, for every $i \neq j$, case (2) of Lemma 6.35 implies that

$$\text{for all } n \in \mathbb{N}, \text{ if } \varphi \otimes \psi(a_\iota \otimes b_\iota) \neq 0, \text{ then } b_\iota = \psi(b_\iota) \cdot 1_{\mathcal{B}} \text{ for } \iota = 1, 2,$$

as desired.

(3). Suppose that $A_i \subset B_i$ for all $i \in I$, and that there is an index $j \in I$ such that $A_j \not\supset B_j$. Then, for every $i \neq j$, case (3) of Lemma 6.35 implies that

$$\text{for all } n \in \mathbb{N}, \text{ if } \varphi \otimes \psi(a_\iota \otimes b_\iota) \neq 0, \text{ then } a_\iota = \varphi(a_\iota) \cdot 1_{\mathcal{A}} \text{ for } \iota = 1, 2,$$

as desired. □

By combining Propositions 6.30, 6.33, and 6.36, we conclude that Theorem 6.20 holds.

6.5. Freeness in at Least One Factor

In this section, we approach the second part of the first question raised in Subsection 6.3.3 (namely, is it necessary that at least one of the families $\mathbf{a}_k$ (where $k \leqslant K$) be $*$-free for the tensor product collection $\otimes_k \mathbf{a}_k$ to be $*$-free) in the special case of group algebras equipped with the canonical trace.

Notation 6.37. By a slight abuse of notation, we will denote every neutral element in a group by e, and recall that $\langle (g_i : i \in I) \rangle$ denotes the group generated by the collection $(g_i : i \in I)$.

Lemma 6.38 ([**MKS66**] Corollary 4.1.6)**.** *Let $(G_i : i \in I)$ be a collection of groups and $G = *_{i \in I} G_i$ be their free product. Suppose that $g, h \in G$ commute. Then,*

(1) $g, h \in \langle w \rangle$ for some $w \in G$; or

(2) $g, h \in zG_iz^{-1}$*, for some* $z \in G$ *and* $i \in I$*.*

Lemma 6.39 ([**MKS66**] Corollary 4.1.4)**.** *Let* $(G_i : i \in I)$ *be a collection of groups and* $G = *_{i\in I}G_i$ *be their free product. Let* $g \in G$ *be of finite order. Then, there exists* $z \in G$*,* $i \in I$ *and an element* $g_i \in G_i$ *of finite order such that* $g = zg_iz^{-1}$*.*

Lemma 6.40. *Let* G *and* H *be groups, and let* $g = (g_1, g_2), h = (h_1, h_2) \in G \times H$ *be such that* $g, h \neq e$*. Suppose that* h *is of finite order, that the order of* h_1 *does not divide the order of* h_2*, and that the order of* h_2 *does not divide the order of* h_1*. Then,* g *and* h *are not free in* G*.*

Proof. Since the orders of h_1 and h_2 do not divide each other, there exists $n, m \in \mathbb{N} \setminus \{0\}$ such that $h_1^n = e \neq h_1^m$ and $h_2^n \neq e = h_2^m$. Since g is not the identity element, if g and h were free, it should be the case that $h^ngh^mg^{-1}h^{-n}gh^{-m}g^{-1} \neq e$. However,

$$\begin{aligned}
&h^ngh^mg^{-1}h^{-n}gh^{-m}g^{-1}\\
=&(e, h_2^n)(g_1, g_2)(h_1^m, e)(g_1^{-1}, g_2^{-1})(e, h_2^{-n})(g_1, g_2)(h_1^{-m}, e)(g_1^{-1}, g_2^{-1})\\
=&(g_1h_1^mg_1^{-1}g_1h_1^{-m}g_1^{-1}, h_2^ng_2g_2^{-1}h_2^{-n}g_2g_2^{-1})\\
=&e.
\end{aligned}$$

Thus, g and h are not free in $G \times H$. □

Proposition 6.41. *Let* G *and* H *be groups, and let* $\mathbf{g} = (g_i : i \in I) \subset G$ *and* $\mathbf{h} = (h_i : i \in I) \subset H$ *be collections of group elements. Define the collection* $\mathbf{g} \times \mathbf{h} = \big((g_i, h_i) : i \in I\big) \subset G \times H$*. If the elements in the collection* $\mathbf{g} \times \mathbf{h}$ *are free in* $G \times H$*, then the elements in* $\mathbf{g}$ *are free in* G*, or the elements in* $\mathbf{h}$ *are free in* H*.*

Proof. Let $F = \big\langle (g_i, h_i) : i \in I \big\rangle \subset G \times H$. Since the (g_i, h_i) are free from each other, then F is trivially isomorphic to the free product $*_{i\in I}\big\langle (g_i, h_i)\big\rangle$.

Define π_1 and π_2 as the projection homomorphisms from F to G and H respectively, that is, for every $(w_1, w_2) \in F$, let $\pi_1\big((w_1, w_2)\big) = w_1$ and $\pi_2\big((w_1, w_2)\big) = w_2$. Then, $\ker(\pi_1)$ and $\ker(\pi_2)$ are normal subgroups of F. Furthermore, given that $\ker(\pi_1) \subset \{e\} \times H$ and $\ker(\pi_2) \subset G \times \{e\}$, it follows that $\ker(\pi_1)$ and $\ker(\pi_2)$ commute and that $\ker(\pi_1) \cap \ker(\pi_2) = \{e\}$.

Suppose that $\ker(\pi_1) = \{e\}$, Then, it must be the case that the elements in $\mathbf{g}$ are free in G. Otherwise, there would exist $i(1), \ldots, i(t) \in I$ and $n(1), \ldots, n(t) \in \mathbb{Z} \setminus \{0\}$ such that

- $i(1) \neq i(2) \neq \ldots \neq i(t)$;
- $g_{i(1)}^{n(1)}, \ldots, g_{i(t)}^{n(t)} \neq e$, and hence $(g_{i(1)}, h_{i(1)})^{n(1)}, \ldots, (g_{i(t)}, h_{i(t)})^{n(t)} \neq e$; and
- $g_{i(1)}^{n(1)} \cdots g_{i(t)}^{n(t)} = e$, that is, $(g_{i(1)}, h_{i(1)})^{n(1)} \cdots (g_{i(t)}, h_{i(t)})^{n(t)} \in \ker(\pi_1)$.

Since $\ker(\pi_1) = \{e\}$, this implies that $(g_{i(1)}, h_{i(1)})^{n(1)} \cdots (g_{i(t)}, h_{i(t)})^{n(t)} = e$, contradicting the freeness of the (g_i, h_i) in $G \times H$. We could similarly prove that $\ker(\pi_2) = \{e\}$ implies that the elements in $\mathbf{h}$ are free in H. Therefore, if it is shown that $\ker(\pi_k) = \{e\}$ is always true for one of $k = 1$ or $k = 2$, then the proof of the proposition will be complete.

Assume by contradiction that $\ker(\pi_1), \ker(\pi_2) \neq \{e\}$. Then, $\ker(\pi_1) \not\subset \ker(\pi_2)$ and $\ker(\pi_2) \not\subset \ker(\pi_1)$, otherwise one would have that $\ker(\pi_1) \cap \ker(\pi_2) \neq \{e\}$. Therefore, there exists elements $g, h \neq e$ such that $g \in \ker(\pi_1) \setminus \ker(\pi_2)$ and $h \in \ker(\pi_2) \setminus \ker(\pi_1)$. Since $\ker(\pi_1)$ and $\ker(\pi_2)$

commute, g and h commute, and thus it follows from Lemma 6.38 that $g, h \in \langle w \rangle$ for some $w \in F$, or that $g, h \in z\big\langle (g_i, h_i) \big\rangle z^{-1}$ for some $z \in F$ and $i \in I$.

Suppose that $g, h \in \langle w \rangle$ for some $w \in G$. Then, there exists $n, m \in \mathbb{N}$ such that $g = w^n$ and $h = w^m$. Since $g \in \ker(\pi_1)$ and $h \in \ker(\pi_2)$, it follows that $w^{nm} = e$. Therefore, Lemma 6.39 implies that $w \in z\big\langle (g_i, h_i) \big\rangle z^{-1}$ for some $z \in G$ and $i \in I$. Consequently, $g \in \ker(\pi_1) \setminus \ker(\pi_2)$ and $h \in \ker(\pi_2) \setminus \ker(\pi_1)$ necessarily implies that $g, h \in z\big\langle (g_i, h_i) \big\rangle z^{-1}$ for some $z \in G$ and $i \in I$.

Let $m, n \in \mathbb{Z} \setminus \{0\}$ be such that $g = z(g_i, h_i)^n z^{-1}$ and $h = z(g_i, h_i)^m z^{-1}$, that is, $(g_i, h_i)^n = z^{-1} g z$ and $(g_i, h_i)^m = z^{-1} h z$. Then, given that $\ker(\pi_1)$ and $\ker(\pi_2)$ are normal subgroups, $(g_i, h_i)^n \in \ker(\pi_1) \setminus \ker(\pi_2)$ and $(g_i, h_i)^m \in \ker(\pi_2) \setminus \ker(\pi_1)$. In other words, (g_i, h_i) is an element of finite order such that the order of g_i does not divide the order of h_i and the order of h_i does not divide the order of g_i. According to Lemma 6.40, this contradicts that the collection $\big((g_i, h_i) : i \in I\big)$ is free. Therefore, the assumption that $\ker(\pi_1), \ker(\pi_2) \neq \{e\}$ must be false, which concludes the proof of the proposition. □

Corollary 6.42. *Let G and H be groups, and consider the $*$-probability spaces $(\mathbb{C}G, *, \tau_e)$ and $(\mathbb{C}H, *, \tau_e)$. Let $\mathbf{g} = (g_i : i \in I) \subset G$ and $\mathbf{h} = (h_i : i \in I) \subset H$ be arbitrary collections of variables, and define $\mathbf{g} \otimes \mathbf{h} = (g_i \otimes h_i : i \in I) \subset \mathbb{C}G \otimes \mathbb{C}H$. If the variables in $\mathbf{u} \otimes \mathbf{v}$ are $*$-free from each other with respect to $\tau_e \otimes \tau_e$, then the variables in $\mathbf{u}$ are $*$-free from each other with respect to τ_e, or the variables in $\mathbf{v}$ are $*$-free from each other with respect to τ_e.*

Proof. Given that the freeness of a collection of groups $(G_i : i \in I)$ is equivalent to the $*$-freeness of their associated group algebras $(\mathbb{C}G_i : i \in I)$ with respect to τ_e (see Proposition 3.26), the result follows directly from the previous proposition. □

The following counterexample shows that Corollary 6.42 does not hold for arbitrary $*$-probability spaces, and that, even if we simply consider group algebras with functionals other than the canonical trace.

Example 6.43. Let $G = \mathbb{F}_2 = \langle g, h \rangle$ be the free group with two generators g and h, and let $H = (\mathbb{Z}, +)$ be the additive group on $\mathbb{Z}$. Let $\mathcal{A}_1 = \mathbb{C}\mathbb{F}_2$ and $\mathcal{A}_2 = \mathbb{C}\mathbb{Z}$, equip $\mathcal{A}_2$ with the canonical trace $\varphi_2 = \tau_e$, and equip $\mathcal{A}_1$ with the linear extension of the map φ_1 defined on $\mathbb{F}_2$ as follows:

$$\varphi_1(m) = \begin{cases} 1 & \text{if } m = e; \\ \beta & \text{if } m = gh, (gh)^{-1}; \text{ and} \\ 0 & \text{otherwise} \end{cases} \tag{6.43}$$

(where $0 < \beta < 1$).

Clearly, $(\mathcal{A}_2, \varphi_2)$ is a $*$-probability space. It can also be shown that φ_1 is positive on $\mathcal{A}_1$: Let

$$w = \sum_{m \in \mathbb{F}_2} \alpha_m m \qquad \text{(with only finitely many } \alpha_m \in \mathbb{C} \text{ nonzero)}$$

be an arbitrary element of $\mathbb{C}\mathbb{F}_2$. Then,

$$\begin{aligned}\varphi_1(ww^*) &= \varphi_1\left(\sum_{m,n\in\mathbb{F}_2}\alpha_m\overline{\alpha_n}mn^{-1}\right)\\ &= \sum_{m\in\mathbb{F}_2}|\alpha_m|^2+\sum_{mn^{-1}=gh}\alpha_m\overline{\alpha_n}\varphi_1(mn^{-1})+\sum_{nm^{-1}=(gh)^{-1}}\alpha_n\overline{\alpha_m}\varphi_1(nm^{-1})\\ &= \sum_{m\in\mathbb{F}_2}|\alpha_m|^2+\beta\left(\sum_{mn^{-1}=gh}\alpha_m\overline{\alpha_n}+\sum_{nm^{-1}=(gh)^{-1}}\alpha_n\overline{\alpha_m}\right).\end{aligned}$$

Given that $nm^{-1}=(gh)^{-1}$ if and only if $mn^{-1}=gh$, the above reduces to

$$\begin{aligned}\varphi_1(ww^*) &= \sum_{m\in\mathbb{F}_2}|\alpha_m|^2+\beta\left(\sum_{mn^{-1}=gh}\alpha_m\overline{\alpha_n}+\alpha_n\overline{\alpha_m}\right)\\ &= \sum_{m\in\mathbb{F}_2}|\alpha_m|^2+2\beta\sum_{mn^{-1}=gh}\mathrm{Re}(\alpha_m\overline{\alpha_n})\\ &\geqslant \sum_{m\in\mathbb{F}_2}|\alpha_m|^2-2\beta\sum_{mn^{-1}=gh}|\alpha_m||\alpha_n|\\ &\geqslant \sum_{mn^{-1}=gh}|\alpha_m|^2+|\alpha_n|^2-2\beta|\alpha_m||\alpha_n|.\end{aligned}$$

So long as $0\leqslant\beta\leqslant 1$, the above implies that $\varphi_1(ww*)\geqslant 0$, and if $\beta\neq 0,1$, then $\varphi_1(ww^*)=0$ if and only if $w=0$, and thus φ_1 as it is defined in (6.43) is positive and faithful on $\mathcal{A}_1$.

Since g and h are free in $\mathbb{F}_2$ and of infinite order, it follows that $(g,\overline{1})$ and $(h,\overline{2})$ are free in $\mathbb{F}_2\times\mathbb{Z}$, and thus $g\otimes\overline{1}$ and $h\otimes\overline{2}$ are $*$-free as members of the $*$-probability space $(\mathbb{C}\mathbb{F}_2\otimes\mathbb{C}\mathbb{Z},\tau_e\otimes\tau_e)$ (we use $\overline{n}$ to denote integers in the additive group $(\mathbb{Z},+)$ to avoid confusing them with scalars in $\mathbb{C}$). Consider a nontrivial (i.e., not the identity) reduced word m in $g\otimes\overline{1}$ and $h\otimes\overline{2}$, that is,

$$m=(g\otimes\overline{1})^{k(1)}(h\otimes\overline{2})^{l(1)}\cdots(g\otimes\overline{1})^{k(t)}(h\otimes\overline{2})^{l(t)},$$

where $k(1),l(1),\ldots,k(t),l(t)\in\mathbb{Z}$ are nonzero except possibly $k(1)$ and $l(t)$. If

$$g^{k(1)}h^{l(1)}\cdots g^{k(t)}h^{l(t)}\notin\{gh,(gh)^{-1}\},$$

then

$$\varphi_1(g^{k(1)}h^{l(1)}\cdots g^{k(t)}h^{l(t)})=0,$$

and thus $(\varphi_1\otimes\varphi_2)(m)=0$. If $g^{k(1)}h^{l(1)}\cdots g^{k(t)}h^{l(t)}\in\{gh,(gh)^{-1}\}$, then it must be the case that $k(1)+\cdots+k(t)=1=l(1)+\cdots+l(t)$ or $k(1)+\cdots+k(t)=-1=l(1)+\cdots+l(t)$, which implies that $m=gh\otimes\overline{3}$ or $m=(gh)^{-1}\otimes\overline{-3}$. Since both $\varphi_2(\overline{3})$ and $\varphi_2(\overline{-3})$ are equal to zero, one has $(\varphi_1\otimes\varphi_2)(m)=0$. Therefore, when restricted to the $*$-algebra generated by $g\otimes\overline{1}$ and $h\otimes\overline{2}$, the functional $\varphi_1\otimes\varphi_2$ is equal to the canonical trace, hence $g\otimes\overline{1}$ and $h\otimes\overline{2}$ are $*$-free as members of the $*$-probability space $(\mathcal{A}_1\otimes\mathcal{A}_2,\varphi_1\otimes\varphi_2)$.

Therefore, if we define $\mathbf{a}_1=(a_{1;1},a_{1;2})=(g,h)$ and $\mathbf{a}_2=(a_{2;1},a_{2;2})=(\overline{1},\overline{2})$, then $\mathbf{a}_1\otimes\mathbf{a}_2=(g\otimes\overline{1},h\otimes\overline{2})$ is $*$-free as a collection in $(\mathcal{A}_1\otimes\mathcal{A}_2,\varphi_1\otimes\varphi_2)$. However, $\mathbf{a}_1$ is not $*$-free, as $\varphi_1(gh)=\beta\neq 0$ while $\varphi_1(g)=\varphi_1(h)=0$; and $\mathbf{a}_2$ is not $*$-free, since $\varphi_2(\overline{1}\cdot\overline{2}\cdot\overline{1}^*\cdot\overline{2}^*)=\varphi_2(\overline{1}+\overline{2}-\overline{1}-\overline{2})=\varphi_2(\overline{0})=1$, yet $\varphi_2(\overline{1})=\varphi_2(\overline{2})=0$.

6.6. Further Questions and Future Investigations

6.6.1. Tensor Product with Only One $*$-free Family. Thanks to Theorem 6.20, we understand the occurence of $*$-freeness in the tensor product $\mathbf{a} \otimes \mathbf{b}$ of two families $\mathbf{a}$ and $\mathbf{b}$ that themselves contain $*$-free variables (assuming the functionals on the considered $*$-probability spaces are faithful). A next step towards solving Problem 6.2 or its special case Problem 6.3 could be to try to answer the following:

Problem 6.44. Let $\mathbf{a} = (a_i : i \in I)$ and $\mathbf{b} = (b_i : i \in I)$ be two families of noncommutative random variables in respective $*$-algebras $(\mathcal{A}, *, \varphi)$ and $(\mathcal{B}, *, \psi)$. Characterize the $*$-freeness of the elements in $\mathbf{a} \otimes \mathbf{b} = (a_i \otimes b_i : i \in I)$ in $(\mathcal{A} \otimes \mathcal{B}, *, \varphi \otimes \psi)$ under the assumption that

(1) the a_i are $*$-free from each other;

(2) the b_i are *not* $*$-free from each other; and

(3) φ and ψ are faithful.

As was done in Section 6.4, one may attempt to find necessary conditions for the $*$-freeness of the elements in $\mathbf{a} \otimes \mathbf{b}$ by expressing $\varphi \otimes \psi\big(M(a_i \otimes b_i : i \in I)\big)$ in terms of the distribution of the individual $a_i \otimes b_i$ (where M is a $*$-word), and then relate this expression to $\varphi\big(M(a_i : i \in I)\big)\psi\big(M(b_i : i \in I)\big)$. However, given that the b_i are not $*$-free from each other, one expects that the expression $\psi\big(M(b_i : i \in I)\big)$ will depend heavily on the joint distribution of the b_i. Thus, it seems that approaching Problem 6.44 either requires a new proof technique, or a better understanding of *dependence* for noncommutative random variables, the latter being a very difficult problem.

6.6.2. Tensor Products of Three or More $*$-free Families. A possible direction for further inquiry could be to attempt to extend the results of Section 6.4 to tensor products of three or more families. A *naive* or *natural* extension of Problems 6.5/6.19 could be the following:

Problem 6.45. Given $K \in \mathbb{N}$ families $\mathbf{a}_k = (a_{k;i} : i \in I) \subset (\mathcal{A}_k, *, \varphi_k)$ (where $1 \leqslant k \leqslant K$) of noncommutative random variables such that for every $k \leqslant K$

(1) the variables in $\mathbf{a}_k$ are $*$-free from each other; and

(2) φ_k is faithful,

characterize the $*$-freeness of the elements in the tensor product collection

$$\otimes_k \mathbf{a}_k = (\otimes_k a_{k;i} : i \in I) \subset (\otimes_k \mathcal{A}_k, *, \otimes_k \varphi_k)$$

in terms of our knowledge of the behaviour of the individual factor collections $\mathbf{a}_k$.

A potential approach to this Problem would be to apply the methods of Section 6.4 directly: By assuming that the $a_{k;i}$ are $*$-free from each other in $(\mathcal{A}_k, *, \varphi_k)$ for all $k \leqslant K$, it follows from equation (6.2) that for every distinct indices $p, q \in I$, $n \in \mathbb{N}$ and $*$-words $M_1, \ldots, M_n, N_1, \ldots, N_n \in \mathbb{C}\langle x, x^* \rangle$, one has

$$\begin{aligned}\varphi_k\big(M_1(a_{k;p})N_1(a_{k;q}) \cdots M_n(a_{k;p})N_n(a_{k;q})\big) \\ = \sum_{\pi \in NC(n)} \kappa_\pi\big[M_1(a_{k;p}), \ldots, M_n(a_{k;p})\big]\varphi_{K(\pi)}\big[N_1(a_{k;q}) \cdots N_n(a_{k;q})\big]\end{aligned}$$

for each $k \leqslant K$, and hence

$$\begin{aligned}
&\otimes_k \varphi_k\big(M_1(\otimes_k a_{k;p})N_1(\otimes_k a_{k;q})\cdots M_n(\otimes_k a_{k;p})N_n(\otimes_k a_{k;q})\big)\\
&= \prod_{k\leqslant K} \varphi_k\big(M_1(a_{k;p})N_1(a_{k;q})\cdots M_n(a_{k;p})N_n(a_{k;q})\big)\\
&= \prod_{k\leqslant K} \sum_{\pi\in NC(n)} \kappa_\pi\big[M_1(a_{k;p}),\ldots,M_n(a_{k;p})\big](\varphi_k)_{K(\pi)}\big[N_1(a_{k;q})\cdots N_n(a_{k;q})\big]\\
&= \sum_{\pi_1,\ldots,\pi_K\in NC(n)} \prod_{k\leqslant K} \kappa_{\pi_k}\big[M_1(a_{k;p}),\ldots,M_n(a_{k;p})\big](\varphi_k)_{K(\pi_k)}\big[N_1(a_{k;q})\cdots N_n(a_{k;q})\big]. \qquad (6.44)
\end{aligned}$$

If it is also assumed that the $\otimes_k a_{k;i}$ are $*$-free from each other, then we obtain from equation (6.2) that

$$\begin{aligned}
&\otimes_k \varphi_k\big(M_1(\otimes_k a_{k;p})N_1(\otimes_k a_{k;q})\cdots M_n(\otimes_k a_{k;p})N_n(\otimes_k a_{k;q})\big)\\
&\quad= \sum_{\pi\in NC(n)} \kappa_\pi\big[M_1(\otimes_k a_{k;p}),\ldots,M_n(\otimes_k a_{k;p})\big]\otimes_k (\varphi_k)_{K(\pi)}\big[N_1(\otimes_k a_{k;q})\cdots N_n(\otimes_k a_{k;q})\big]\\
&\quad= \sum_{\pi\in NC(n)} \kappa_\pi\big[M_1(\otimes_k a_{k;p}),\ldots,M_n(\otimes_k a_{k;p})\big]\prod_{k\leqslant K}(\varphi_k)_{K(\pi)}\big[N_1(\otimes_k a_{k;q})\cdots N_n(\otimes_k a_{k;q})\big].
\end{aligned}$$

Then, subtracting the above from 6.44 would yield a necessary condition similar to equation (6.9), which we could attempt to specialize as was done in Propositions 6.22 and 6.23.

However, it is possible that replicating Propositions 6.22 and 6.23 for the above case would be very impractical. Indeed, given that the sum in (6.44) involves K partitions $\pi_1, \pi_2, \ldots, \pi_K$ instead of two, replicating the summand-by-summand analysis done in Propositions 6.22 and 6.23 could involve an overwhelming amount of cases and/or the analogs of equations (6.10), (6.15) and (6.16) would be very impractical. Thus, it is possible that a rethinking of the current methods or finding completely new techniques would be necessary.

6.6.3. Full Characterization of $*$-freeness for Tensor Products of Group Algebras with the Canonical Trace. In Section 6.5, we have have found a first necessary condition for the $*$-freeness of a tensor product collection $\mathbf{g}\otimes\mathbf{h} = (g_i \otimes h_i : i \in I)$ in $(\mathbb{C}G\otimes\mathbb{C}H, *, \tau_e\otimes\tau_e)$, where $\mathbf{g} \subset G$ and $\mathbf{h} \subset H$, namely, we have shown that it is necessary that one of $\mathbf{g}$ or $\mathbf{h}$ be $*$-free.

If $\mathbf{g}$ and $\mathbf{h}$ both contain $*$-free elements, then the results in Section 6.4 show that the Reformulated Centering Conditions are necessary for the $*$-freeness of $\mathbf{g}\otimes\mathbf{h}$. Thus, all that remains to be done to achieve a full characterization of this special case is to find strong enough necessary conditions for the $*$-freeness of $\mathbf{g}\otimes\mathbf{h}$ when one of $\mathbf{g}$ and $\mathbf{h}$ contains $*$-free variables and the other does not. (Note that this would provide a solution to a special case of Problem 6.44.)

If we conjecture that the Reformulated Centering Conditions are also necessary in the case when only one of $\mathbf{g}$ and $\mathbf{h}$ contains $*$-free variables, then the problem we are attempting to solve is the following:

Problem 6.46. Let G and H be groups, and let $\mathbf{g} = (g_i : i \in I) \subset G$ and $\mathbf{h} = (h_i : i \in I) \subset H$ be collections such that $\mathbf{g}$ is free in G and $\mathbf{h}$ is not free in H. Show that for every $i \in I$, if $\mathbf{g}\times\mathbf{h}$ is free in $G\times H$, then $\mathbf{g}\otimes\mathbf{h}$ satisfies the Reformulated Centering Conditions in $(\mathbb{C}G\otimes\mathbb{C}H, *, \tau_e\otimes\tau_e)$

with $\mathbf{g}$ as a dominating family, that is, if $(g_i, h_i)^n \neq e$ for some $n \in \mathbb{N}$, then $g_i^n \neq e$ (the condition "if $(g_i, h_i)^n = e$, then $h_i^n = e$" is trivially satisfied in this case).

While our main interest in the above problem stems from applications in free probability, one may argue that it could be of independent interest in the context of the theory of free groups. In this context, a more thorough review of the current literature on results concerning the occurence of freeness in direct products of groups would be in order.

Part IV. Appendix

Appendix A

Algebra

A.1. Free Groups

See [**MKS66**] Chapter 4.

Definition A.1. Let G be a group (with identity e) and $(G_i : i \in I)$ be a collection of subgroups of G. The G_i are said to be free from each other if and only if for every collection $i(1), \ldots, i(t) \in I$ of indices such that $i(1) \neq i(2) \neq \cdots \neq i(t)$ and elements $g_1 \in G_{i(1)} \setminus \{e\}, \ldots, g_t \in G_{i(t)} \setminus \{e\}$, one has $g_1 \cdots g_t \neq e$, that is, there exists no nontrivial relations between elements from different subgroups G_i.

Definition A.2. Let $(G_i : i \in I)$ be a collection of groups with respective identities e_i. A **reduced nontrivial word** in the G_i is an expression of the form $g_1 \cdots g_t$, where $g_1 \in G_{i(1)} \setminus \{e_{i(1)}\}, \ldots, g_t \in G_{i(t)} \setminus \{e_{i(t)}\}$ and $i(1) \neq i(2) \neq \cdots i(t)$.

Let $(G_i : i \in I)$ be a collection of groups with respective identities e_i. Then, the *free product* of the G_i, denoted $*_{i \in I} G_i$, is defined as the set of every reduced nontrivial word in the G_i together with a generic identity element labeled e, which is equipped with the following product operation: For every reduced nontrivial word $g \in *_{i \in I} G$, one has $e \cdot g = g \cdot e = g$; and for every two reduced nontrivial words $g = g_1 \cdots g_s$ and $h = h_1 \cdots h_t$, where

$$g_1 \in G_{i(1)} \setminus \{e_{i(1)}\}, \ldots, g_s \in G_{i(s)} \setminus \{e_{i(s)}\} \text{ and } i(1) \neq i(2) \neq \cdots i(s)$$

and

$$h_1 \in G_{j(1)} \setminus \{e_{j(1)}\}, \ldots, h_t \in G_{j(t)} \setminus \{e_{j(t)}\} \text{ and } j(1) \neq i(2) \neq \cdots j(t),$$

one defines the product $g \cdot h$ using the following recursive algorithm:

Step 1. If $i(s) \neq j(1)$, then the concatenated word $g_1 \cdots g_s h_1 \cdots h_t$ is a nontrivial reduced word, and we output $g \cdot h = g_1 \cdots g_s h_1 \cdots h_t$. If $i(s) = j(1)$, then go to Step 2.

Step 2. Let $i = i(s) = j(1)$. If $g_s h_1 \neq e_i$, then the concatenated word $g_1 \cdots (g_s h_1) \cdots h_t$ is a nontrivial reduced word (since $i(1) \neq \cdots \neq i(s) = j(1) \neq \cdots j(t)$), and we output $g \cdot h = g_1 \cdots (g_s h_1) \cdots h_t$. If $g_s h_1 = e_i$, then go to Step 3.

Step 3. If $g = g_s$ and $h = h_1$ (i.e., $s = t = 1$), then we output $g \cdot h = e$. If $g = g_s$ and $h \neq h_1$ (i.e., $s = 1 \neq t$), then $h_2 \cdots h_t$ is a reduced nontrivial word, and we output $g \cdot h = h_2 \cdots h_t$. If $g \neq g_s$ and $h = h_1$ (i.e., $s \neq 1 = t$), then $g_1 \cdots g_{s-1}$ is a reduced nontrivial word, and we output $g \cdot h = g_1 \cdots g_{s-1}$. If $g \neq g_s$ and $h \neq h_1$ (i.e., $s, t \neq 1$), then redefine $g = g_1 \cdots g_{s-1}$ and $h = h_2 \cdots h_t$, and go back to Step 1.

Proposition A.3. *The algorithm outlined above always outputs either the formal identity e or a reduced word, and $*_{i\in I}G_i$ with this operation and e as an identity element forms a group. Furthermore, if one embeds every group G_i into $*_{i\in I}G_i$ as follows: $e_i \hookrightarrow e$, and $g \hookrightarrow g$ for every $g \in G_i \setminus \{e_i\}$, then the embeddings of the G_i are free in $*_{i\in I}G_i$.*

Notation A.4. Let G be a group with identity e. Given an arbitrary $g \in G$, let $\langle g \rangle$ denote the subgroup of G generated by g, that is, if g is of finite order n, then

$$\langle g \rangle = \{e, g, g^2, \ldots, g^{n-1}\} \cong (\mathbb{Z}/n\mathbb{Z}, +),$$

and if g is of infinite order, then

$$\langle g \rangle = \left\{e, g^n : n \in \mathbb{Z} \setminus \{0\}\right\} \cong (\mathbb{Z}, +).$$

Notation A.5. Let $(G_i : i \in I)$ be a collection of groups each generated by a single element g_i of infinite order, that is, for every $i \in I$,

$$G_i = \langle g_i \rangle \cong (\mathbb{Z}, +),$$

and for every $m, n \in \mathbb{Z}$, $g_i^m g_i^n = g_i^{m+n}$ (where $g_i^0 = e_i$). Then, the free product $*_{i\in I}G_i$ is denoted $\mathbb{F}_I$ if I is infinite, and $\mathbb{F}_{|I|}$ if I is finite, where $|I|$ denotes the cardinality of I. In the case where $|I| = n \in \mathbb{N}$, the group $\mathbb{F}_n$ will be called the **free group with n generators**.

A.2. Free Unital Algebras

See **[VDN92]** Chapter 1 and **[NS06]** Lecture 6.

Let $(\mathcal{A}_i : i \in I)$ be a collection of unital algebras over $\mathbb{C}$ with respective unit elements $1_{\mathcal{A}_i}$, and suppose that for every $i \in I$, there exists a subspace $\mathcal{A}_i^\dagger \subset \mathcal{A}_i$ of codimension 1 such that $\mathcal{A}_i = \mathbb{C}1_{\mathcal{A}_i} \oplus \mathcal{A}_i^\dagger$. Then, the **free product** of the algebras $\mathcal{A}_i$, denoted $*_{i\in I}\mathcal{A}_i$, is defined as the vector space

$$\mathbb{C}1_{\mathcal{A}} \oplus \left(\bigoplus_{t \geqslant 1} \bigoplus_{\substack{i(1),\ldots,i(t)\in I \\ i(1)\neq i(2)\neq\cdots\neq i(t)}} \mathcal{A}_{i(1)}^\dagger \otimes \cdots \otimes \mathcal{A}_{i(t)}^\dagger \right)$$

(where $1_{\mathcal{A}}$ is a formal unit element) together with the following product operation: For every $a \in *_{i\in I}\mathcal{A}_i$, $1_{\mathcal{A}}a = a1_{\mathcal{A}} = a$, for every two elements of the form $a = a_1 \otimes \cdots \otimes a_s$ and $b = b_1 \otimes \cdots \otimes b_t$, where

$$a_1 \in \mathcal{A}_{i(1)}^\dagger, \ldots, a_s \in \mathcal{A}_{i(s)}^\dagger \text{ and } i(1) \neq i(2) \neq \cdots i(s)$$

and

$$b_1 \in \mathcal{A}_{j(1)}^\dagger, \ldots, b_t \in \mathcal{A}_{j(t)}^\dagger \text{ and } j(1) \neq i(2) \neq \cdots j(t),$$

one defines the product $a \cdot b$ using the following recursive algorithm:

Step 1. If $i(s) \neq j(1)$, then output

$$a \cdot b = a_1 \otimes \cdots \otimes a_s \otimes b_1 \otimes \cdots \otimes b_t,$$

which is an element of $*_{i\in I}\mathcal{A}_i$, as it inside

$$\mathcal{A}_{i(1)}^{\dagger} \otimes \cdots \otimes \mathcal{A}_{i(s)}^{\dagger} \otimes \mathcal{A}_{j(1)}^{\dagger} \otimes \cdots \otimes \mathcal{A}_{j(t)}^{\dagger},$$

where

$$i(1) \neq i(2) \neq \cdots \neq i(s) \neq j(1) \neq j(2) \neq \cdots \neq j(t).$$

If $i(s) = j(1)$, then go to Step 2.

Step 2. Let $i = i(s) = j(1)$. If $a_s b_1 \in \mathcal{A}_i^{\dagger}$, then output

$$a \cdot b = a_1 \otimes \cdots \otimes a_{s-1} \otimes a_s b_1 \otimes b_2 \otimes \cdots \otimes b_t,$$

which is an element of $*_{i\in I}\mathcal{A}_i$, as it inside

$$\mathcal{A}_{i(1)}^{\dagger} \otimes \cdots \otimes \mathcal{A}_{i(s-1)}^{\dagger} \otimes \mathcal{A}_{i}^{\dagger} \otimes \mathcal{A}_{j(2)}^{\dagger} \otimes \cdots \otimes \mathcal{A}_{j(t)}^{\dagger},$$

where

$$i(1) \neq i(2) \neq \cdots \neq i(s-1) \neq i \neq j(2) \neq j(2) \neq \cdots \neq j(t).$$

If $a_s b_1 \notin \mathcal{A}_i^{\dagger}$, then go to Step 3.

Step 3. Let $c = a_s b_1 \in \mathcal{A}_i$. Since $\mathcal{A}_i = \mathbb{C}1_{\mathcal{A}_i} \oplus \mathcal{A}_i^{\dagger}$, there exists $\lambda \in \mathbb{C}$ and $c^{\dagger} \in \mathcal{A}_i^{\dagger}$ such that $c = \lambda \cdot 1_{\mathcal{A}_i} + c^{\dagger}$. Clearly,

$$a_1 \otimes \cdots \otimes a_{s-1} \otimes c^{\dagger} \otimes b_2 \otimes \cdots \otimes b_t$$

is an element of $*_{i\in I}\mathcal{A}_i$, as it inside

$$\mathcal{A}_{i(1)}^{\dagger} \otimes \cdots \otimes \mathcal{A}_{i(s-1)}^{\dagger} \otimes \mathcal{A}_{i}^{\dagger} \otimes \mathcal{A}_{j(2)}^{\dagger} \otimes \cdots \otimes \mathcal{A}_{j(t)}^{\dagger},$$

where

$$i(1) \neq i(2) \neq \cdots \neq i(s-1) \neq i \neq j(2) \neq j(2) \neq \cdots \neq j(t).$$

If $s = 1$ or $t = 1$ or $i(s-1) \neq j(2)$, then the element

$$a_1 \otimes \cdots \otimes a_{s-1} \otimes b_2 \otimes \cdots \otimes b_t$$

is also in $*_{i\in I}\mathcal{A}_i$, as it is inside

$$\mathcal{A}_{i(1)}^{\dagger} \otimes \cdots \otimes \mathcal{A}_{i(s-1)}^{\dagger} \otimes \mathcal{A}_{j(2)}^{\dagger} \otimes \cdots \otimes \mathcal{A}_{j(t)}^{\dagger},$$

where

$$i(1) \neq i(2) \neq \cdots \neq i(s-1) \neq j(2) \neq j(2) \neq \cdots \neq j(t).$$

In this case, output

$$a \cdot b = \lambda \cdot (a_1 \otimes \cdots \otimes a_{s-1} \otimes b_2 \otimes \cdots \otimes b_t) + a_1 \otimes \cdots \otimes a_{s-1} \otimes c^{\dagger} \otimes b_2 \otimes \cdots \otimes b_t.$$

If $s \neq 1$, $t \neq 1$, and $i(s-1) = j(2)$, then let $a' = a_1 \otimes \cdots \otimes a_{s-1}$ and $b' = b_2 \otimes \cdots \otimes b_t$, return a' and b' in Step 2, and the output will be

$$a \cdot b = \lambda \cdot (a' \cdot b') + a_1 \otimes \cdots \otimes a_{s-1} \otimes c^{\dagger} \otimes b_2 \otimes \cdots \otimes b_t.$$

Then, extend the product operation to all of $*_{i\in I}\mathcal{A}_i$ to make it distributive over the direct sums.

Example A.6 ([**NS06**] Lecture 6, pages 83 and 84)**.** Let $\mathcal{A}_1 = \mathbb{C}1_{\mathcal{A}_1} \oplus \mathcal{A}_1^\dagger$ and $\mathcal{A}_2 = \mathbb{C}1_{\mathcal{A}_2} \oplus \mathcal{A}_2^\dagger$ be two unital algebras, and let $a_1, b_1 \in \mathcal{A}_1^\dagger$ and $a_2, b_2 \in \mathcal{A}_2^\dagger$ be such that $a_i b_i \notin \mathcal{A}_i^\dagger$ $(i = 1, 2)$.

Consider the product $(a_1 \otimes a_2) \cdot (b_2 \otimes b_1)$. We apply the above algorithm. Since $a_2 b_2 \notin \mathcal{A}_2^\dagger$, we may proceed directly to Step 3. Let $c = a_2 b_2 = \lambda \cdot 1_{\mathcal{A}_2} + c^\dagger$, where $\lambda \in \mathbb{C}$ and $c^\dagger \in \mathcal{A}_2^\dagger$. Then,

$$(a_1 \otimes a_2) \cdot (b_2 \otimes b_1) = \lambda \cdot (a_1 \otimes b_1) + a_1 \otimes c^\dagger \otimes b_1.$$

Since $a_1 b_1 \notin \mathcal{A}_1^\dagger$, we may proceed directly to Step 3 with $a_1 \otimes b_1$. Letting $d = a_1 b_1 = \mu \cdot 1_{\mathcal{A}} + d^\dagger$, where $\mu \in \mathbb{C}$ and $d^\dagger \in \mathcal{A}_1^\dagger$, we finally obtain

$$(a_1 \otimes a_2) \cdot (b_2 \otimes b_1) = \lambda\mu \cdot 1_{\mathcal{A}} + \lambda \cdot d^\dagger + a_1 \otimes c^\dagger \oplus b_1,$$

which is an element of the free product $\mathcal{A}_1 * \mathcal{A}_2$.

Proposition A.7. *The algorithm outlined above always outputs either the formal unit element* $1_{\mathcal{A}}$ *or a sum in*

$$\bigoplus_{t \geqslant 1} \bigoplus_{\substack{i(1),\ldots,i(t) \in I \\ i(1) \neq i(2) \neq \cdots \neq i(t)}} \mathcal{A}_{i(1)}^\dagger \otimes \cdots \otimes \mathcal{A}_{i(t)}^\dagger$$

and $*_{i \in I} \mathcal{A}_i$ *with this operation and* $1_{\mathcal{A}}$ *as a unit element forms an algebra over* $\mathbb{C}$*.*

A.3. Linear Algebra

Spectral Theorem for Normal Matrices. *Let* $n \in \mathbb{N}$*, let* $A \in \mathcal{M}_n(\mathbb{C})$ *be a* $n \times n$ *matrix, and let* $\lambda_1, \ldots, \lambda_n$ *be the eigenvalues of* A*. Then,* A *is normal if and only if there exists a unitary matrix* U *such that* $A = UDU^*$*, where* $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ *is a diagonal matrix with the eigenvalues of* A *on its diagonal.*

Definition A.8. A $n \times n$ square matrix $A \in \mathcal{M}_n(\mathbb{C})$ is said to be **positive** if one of the following (equivalent) conditions hold:

(1) $A = A^*$ and $\operatorname{Sp}_{\mathcal{M}_n(\mathbb{C})}(A) \subset [0, \infty)$;

(2) $A = XX^*$ for some $X \in \mathcal{M}_n(\mathbb{C})$; and

(3) $\langle Az, z \rangle \geqslant 0$ for every vector $z \in \mathbb{C}^n$, where $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product on $\mathbb{C}^n$.

Proposition A.9 ([**NS06**] Lemma 6.11)**.** *Let* $A = (A_{ij} : 1 \leqslant i, j \leqslant n)$ *and* $B = (B_{ij} : 1 \leqslant i, j \leqslant n)$ *be* $n \times n$ *matrices in* $\mathcal{M}_n(\mathbb{C})$*, and define their* **Schur product***, denoted* $S(A, B)$*, as*

$$S(A, B) = (A_{ij} B_{ij} : 1 \leqslant i, j \leqslant n) \in \mathcal{M}_n(\mathbb{C}).$$

If A *and* B *are positive, then so is* $S(A, B)$*.*

Proposition A.10 ([**NS06**] Lemma 6.12)**.** *Let* $(\mathcal{A}, *, \varphi)$ *be a* $*$*-algebra. Then,* φ *is positive if and only if for every* $n \in \mathbb{N}$ *and* $a_1, \ldots, a_n \in \mathcal{A}$*, the matrix* $\big(\varphi(a_i^* a_j) : 1 \leqslant i, j \leqslant n\big)$ *is positive.*

Appendix B

Analysis/Probability

B.1. Classical Probability Theory

B.1.1. Axioms of Classical Probability. The axiomatic basis of the modern mathematical theory of probability was introduced in the 1930s by Andrey Kolmogorov.[1] In this formalism, the fundamental mathematical framework used to study the behaviour of a random system consists of a triple $(\Omega, \mathscr{A}, P)$ called a probability space, which is defined in the following manner: Consider an experiment E whose outcome is uncertain.

(1) Firstly, one defines Ω (the sample space) as a set that *contains* every possible outcome of the experiment E, that is, every conceivable result of E is represented as a mathematical object $\omega \in \Omega$.

(2) Secondly, one defines $\mathscr{A}$ (the event space) as a collection of subsets $A \subset \Omega$. Each subset $A \in \mathscr{A}$ is called an event. The event space is required to satisfy certain axioms, which can be summarized as requiring that $\mathscr{A}$ be a σ-algebra. An event $A \subset \Omega$ is said to be a success (or to occur) if it contains the outcome of the experiment.

(3) Thirdly, one defines P as a function that assigns to every event A its *chance of success*

$$P(A) \equiv \mathbf{P}[\text{“}A \text{ occurs"}] \in [0, 1].$$

P is also required to fulfill some conditions, which can be summarized as requiring that P be a probability measure on $\mathscr{A}$.

When analyzing a random experiment, it is often the case that one is interested in a function of the outcome of the experiment rather than the outcome itself. For example, given a game of chance G (upon which one or several players wager money) modelled by the probability space $(\Omega, \mathscr{A}, P)$, one might be interested in the function $W : \Omega \to \mathbb{R}$ that assigns to any outcome ω of the game the amount of money $W(\omega) \in \mathbb{R}$ a given player wins (or loses if $W(\omega) < 0$) for that particular outcome. This gives rise to the notion of a random variable. In general, given a probability space $(\Omega, \mathscr{A}, P)$, a random variable X is defined as a function $X : \Omega \to \mathbb{C}$ that

[1] See **[Ko56]** for an english translation of Kolmogorov's original work.

is measurable with respect to $\mathscr{A}$ and the Borel σ-algebra $\mathscr{B}(\mathbb{C})$ on $\mathbb{C}$. Assuming that random variables be measurable ensures that every subset of Ω of the form

$$X^{-1}(A) = [X \in A] = \big\{\omega \in \Omega : X(\omega) \in A\big\}, \qquad A \text{ Borel}$$

is an event, and thus has a probability of succes assigned to it by P. Moreover, it enables one to assign to any variable X a distribution μ_X, which is defined as a measure on $\big(\mathbb{C}, \mathscr{B}(\mathbb{C})\big)$ that satisfies

$$\mu_X(A) = \mathbf{P}[X \in A] = P\big(X^{-1}(A)\big) = P\big(\{\omega \in \Omega : X(\omega) \in A\}\big)$$

for every Borel set A. Combining this with the fact that for every Borel probability measure μ on $\mathbb{C}$ and for every *standard*[2] probability space $(\Omega, \mathscr{A}, P)$, there exists a random variable $X : \Omega \to \mathbb{C}$ whose distribution is μ_X, we are enabled to study random variables independently of the sample space from which they originate using the theory of measure and integration by viewing the sample space as an abstract *source of randomness* rather than as a model for a specific random experiment (that is, through abstract statements such as: "let X be a random variable with distribution μ_X"). In this abstract context, random variables are usually distinguished using their distributions, and any difference between two measurable functions $X : \Omega \to \mathbb{C}$ and $Y : \Omega \to \mathbb{C}$ that does not manifest itself in the distributions μ_X and μ_Y are seen as irrelevant from a probabilistic point of view.

B.1.2. Independence. In informal terms, two events A and B are said to be independent if the occurrence of A has no impact on the probability that B occurs, and vice versa. In mathematical terms (i.e., in the measure-theoretic formulation of probability), the independence of two events is typically defined as follows: Let $(\Omega, \mathscr{A}, P)$ be a probability space. Two events $A, B \in \mathscr{A}$ are said to be **independent** if and only if

$$\mathbf{P}[A \text{ and } B \text{ both occur}] = P(A \cap B) = P(A)P(B).^{3}$$

An equivalent (although more abstract) way of defining the independence of events is the following: Let $(\Omega, \mathscr{A}, P)$ be a probability space and $A, B \in \mathscr{A}$ be events. Let $\sigma_A = \{\Omega, A, A^c, \varnothing\}$ and $\sigma_B = \{\Omega, B, B^c, \varnothing\}$ be the σ-algebras generated by A and B respectively. Define the probability measure $P_{A,B}$ on the product space $(\Omega \times \Omega, \sigma_A \times \sigma_B)$ as

$$P_{A,B}(X \times Y) \coloneqq \mathbf{P}[X \text{ and } Y \text{ both occur}], \qquad (X \in \sigma_A \text{ and } Y \in \sigma_B).$$

Then A and B are **independent** if and only if $P_{A,B}$ is equal to the product measure $P \times P$ on $\mathscr{A} \times \mathscr{A}$ restricted to $\sigma_A \times \sigma_B$. The advantage of this abstract formulation is that it can easily be extended to the case where we consider the independence of collections of events rather than the independence of single events, or the independence of infinitely many collections of events, giving the following general definition:

---

[2] See **[It84]** Section 2.4 for a definition of a standard probability space.

[3] It is interesting to note that this definition has its origins in the *frequentist* approach to probability: Suppose that the same experiment E is conducted n times, and that, out of those n, the events A, B and $[A \text{ and } B]$ occur n_A, n_B and $n_{A\cap B}$ times respectively. Then, if A and B are independent, it is to be expected that the frequency of occurrence of A (i.e. the ratio n_A/n) will be in no way affected by the occurrence of B and vice versa. Thus, the ratio $n_{A\cap B}/n_A$, which represents the frequency of occurrence of B when A has occurred, should be roughly equal to to n_B/n, the probability that B occurs, hence $n_{A\cap B}/n = (n_{A\cap B}/n_A)(n_A/n) \approx (n_B/n)(n_A/n)$.

Definition B.1. Let $(\Omega, \mathscr{A}, P)$ be a standard probability space[4] and let I be an arbitrary (possibly infinite) indexing set. For every $i \in I$, let $C_i \subset \mathscr{A}$ be a collection of of events, and let σ_i be the σ-algebra generated by C_i. Define the probability measure $P_{\mathbf{C}}$ on the product space $(\times_{i\in I}\Omega, \times_{i\in I}\sigma_i)$ as

$$P_{\mathbf{C}}(\times_{i\in I}A_i) \coloneqq \mathbf{P}[A_i \text{ occurs for every } i \in I],$$

where $A_i \in \sigma_i$ for every $i \in I$, and $A_i \neq \Omega$ for only finitely many i. Then, the C_i are said to be **independent** of each other if and only if $P_{\mathbf{C}}$ is equal to the product measure $\times_{i\in I}P$ on $\times_{i\in I}\mathscr{A}$ restricted to $\times_{i\in I}\sigma_i$, Equivalently, the C_i are said to be independent of each other if and only if for every finite collection $i(1), \ldots, i(t) \in I$ of distinct indices and events $A_1 \in C_{i(1)}, \ldots, A_n \in C_{i(t)}$, one has

$$\mathbf{P}[A_i \text{ occurs for every } i \leqslant t] = \prod_{i=1}^{t} \mathbf{P}[A_i \text{ occurs}].$$

With this general definition of independence for events, we can introduce the definition of independence for random variables:

Definition B.2. Let $(\Omega, \mathscr{A}, P)$ be a standard probability space and let I be an arbitrary indexing set. For every $i \in I$, let $\mathcal{X}_i$ be a collection of random variables on Ω, and let C_i be the collection of every event of the form

$$[X_i \text{ is in } A] = \big\{\omega \in \Omega : X_i(\omega) \in A\big\}, \qquad X_i \in \mathcal{X}_i \text{ and } A \in \mathscr{B}(\mathbb{C}).$$

Then, the collections of random variables in $\boldsymbol{\mathcal{X}} = (\mathcal{X}_i : i \in I)$ are **independent** of each other if and only if the collections of events in $\mathbf{C} = (C_i : i \in I)$ are independent of each other.

Notice that, if the collections $\mathcal{X}_i$ each contain only one random variable X_i, then the independence of the X_i (i.e., the independence of the collections of singletons $\mathcal{X}_i$) is equivalent to the assertion that the joint distribution $\mu_{\mathbf{X}}$ of the collection $\mathbf{X} = (X_i : i \in I)$ is equal to the product $\times_{i\in I}\mu_{X_i}$ of the individual distributions of the X_i on $\big(\times_{i\in I}\mathbb{C}, \times_{i\in I}\mathscr{B}(\mathbb{C})\big)$.

Proposition B.3. *Let $(\Omega, \mathscr{A}, P)$ be a standard probability space and let $\mathbf{X} = (X_i : i \in I)$ be a collection of random variables on Ω. The X_i are independent of each other if and only if for every finite collection $i(1), \ldots, i(t) \in I$ of distinct indices and continuous functions $f_1, \ldots, f_t : \mathbb{C} \to \mathbb{C}$ such that $f_1(X_{i(1)}), \ldots, f_t(X_{i(t)})$ and $f_1(X_{i(1)}) \cdots f_t(X_{i(t)})$ are P-integrable, one has*

$$\mathbf{E}\big[f_1(X_{i(1)}) \cdots f_t(X_{i(t)})\big] = \mathbf{E}\big[f_1(X_{i(1)})\big] \cdots \mathbf{E}\big[f_t(X_{i(t)})\big]. \tag{B.1}$$

Proof. Suppose that the X_i are independent. Then, since the joint distribution of the X_i is the product of the individual distributions, (B.1) follows directly from Fubini's Theorem.

Suppose that (B.1) holds. Let $\times_{i\in I}A_i \in \times_{i\in I}\mathscr{A}$ be such that $A_i \neq \Omega$ for only finitely many i, say, $i(1), \ldots, i(t)$. For each $j \leqslant t$ and $n \in \mathbb{N}$, let $f_n^{(j)} : \mathbb{C} \to \mathbb{C}$ be defined as

$$f_n^{(j)}(z) = \begin{cases} 1 & \text{if } z \in A_{i(j)}; \\ 1 - n \cdot d & \text{if } \inf\big\{|z - w| : w \in A_{i(j)}\big\} = d \leqslant \frac{1}{n}; \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

[4]The necessity of assuming that $(\Omega, \mathscr{A}, P)$ is standard can be explained by the fact that defining a product measure for uncountably many measure spaces can otherwise be problematic (see **[Ta11]** Section 2.4.). However, it may be argued that the vast majority of probability spaces that arise in practical applications are standard (see **[It84]** Section 3.1).

Then, every $f_n^{(j)}$ is bounded and continuous, and thus, every $f_n^{(j)}$ is integrable, as well as the product $f_n^{(1)} \cdots f_n^{(t)}$. Furthermore, $f_n^{(j)}$ and converges pointwise to the indicator function $\mathbf{1}_{A_{i(j)}}$ for each $l \leqslant t$. Thus, it follows from Lebesgue's Dominated Convergence Theorem that both

$$\lim_{n\to\infty} \int f_n^{(1)}(z_{i(1)}) \cdots f_n^{(t)}(z_{i(t)}) \, \mathrm{d}\mu_{\mathbf{X}}(z_i : i \in I) = \mu_{\mathbf{X}}(\times_{i\in I} A_i)$$

(where $\mu_{\mathbf{X}}$ denotes the joint distribution of the X_i) and

$$\lim_{n\to\infty} \int f_n^{(1)}(z) d\mu_{X_{i(1)}}(z) \cdots \int f_n^{(t)}(z) \, \mathrm{d}\mu_{X_{i(t)}}(z) = \mu_{X_{i(1)}}(A_{i(1)}) \cdots \mu_{X_{i(t)}}(A_{i(t)}).$$

According to (B.1), this implies that

$$\mu_{\mathbf{X}}(\times_{i\in I} A_i) = \mu_{X_{i(1)}}(A_{i(1)}) \cdots \mu_{X_{i(t)}}(A_{i(t)}),$$

that is, the joint distribution $\mu_{\mathbf{X}}$ is equal to the product $\times_{i\in I}\mu_{X_i}$, hence the X_i are independent. □

B.1.3. Associativity of Independent Events.

Theorem B.4 ([**Bi95**] Corollaries 1 and 2 of Theorem 4.2)**.** *Let $(X_i : i \in I)$ be a family of independent random variables on a probability space $(\Omega, \mathscr{A}, P)$, and let $(I_j : j \in J)$ be a partition of I. Let $(Y_j : j \in J)$ be a family of random variables on a probability space $(\Omega', \mathscr{A}', P')$ such that for every $j \in J$, there exists a measurable map function $f_j : \times_{j\in I_j}\Omega \to \Omega'$ such that $Y_j = f_j(X_i : i \in I_j)$. Then, the Y_j are independent.*

B.1.4. Method of Moments.

Definition B.5. Let μ be a measure on $\big(\mathbb{R}, \mathscr{B}(\mathbb{R})\big)$. μ is said to be **determined by its moments** if, given a measure ν on $\big(\mathbb{R}, \mathscr{B}(\mathbb{R})\big)$,

$$\int x^n \, \mathrm{d}\mu(x) = \int x^n \, \mathrm{d}\nu(x)$$

for every $n \in \mathbb{N}$ implies that $\mu = \nu$. Similarly, a measure μ on $\big(\mathbb{C}, \mathscr{B}(\mathbb{C})\big)$ is determined by its moments if, given a measure ν on $\big(\mathbb{C}, \mathscr{B}(\mathbb{C})\big)$,

$$\int z^m \overline{z}^n \, \mathrm{d}\mu(z) = \int z^m \overline{z}^n \, \mathrm{d}\nu(z)$$

for every $m, n \in \mathbb{N}$ implies that $\mu = \nu$.

Theorem B.6 ([**Bi95**] Theorem 30.1)**.** *Let μ be a measure on $\big(\mathbb{R}, \mathscr{B}(\mathbb{R})\big)$ whose moments are given by*

$$\alpha_n = \int_{\mathbb{R}} x^n \, \mathrm{d}\mu(x), \qquad\qquad n \in \mathbb{N} \cup \{0\}.$$

If the power series

$$\sum_{n=0}^{\infty} \frac{\alpha_n x^n}{n!}$$

converges for some $x > 0$, then μ is determined by its moments.

Skorokhod's Representation Theorem ([**Bi99**] Theorem 6.7)**.** *Let $(\mu_n : n \in \mathbb{N})$ and μ be probability measures on $\big(\mathbb{C}, \mathscr{B}(\mathbb{C})\big)$ such that μ_n converges in distribution to μ. Then, there exists $\mathbb{C}$-valued random variables $\{X_n : n \in \mathbb{N}\}$ and X all defined on the same probability space $(\Omega, \mathscr{A}, P)$ such that*

(1) for each $n \in \mathbb{N}$, X_n is distributed according to μ_n, and X is distributed according to μ; and

(2) X_n converges to X almost surely, that is, for every $\omega \in \Omega'$ (where $P(\Omega') = 1$), one has

$$\lim_{n\to\infty} X_n(\omega) = X(\omega).$$

Theorem B.7 ([**Bi99**] Corollary to Theorem 5.1)**.** *Let $\{\mu_n : n \in \mathbb{N}\}$ be a sequence of tight probability measures on $\mathbb{C}$. If every subsequence of μ_n that converges weakly does so towards the same measure μ, then then entire sequence μ_n converges weakly to μ.*

Method of Moments. *Let $(X_n : n \in \mathbb{N})$ be a sequence of $\mathbb{C}$-valued random variables with moments of every order. Suppose that*

$$\lim_{n\to\infty} \mathbf{E}\left[X_n^k \overline{X_n}^l\right] = \mathbf{E}\left[X^k \overline{X}^l\right], \qquad k, l \in \mathbb{N} \cup \{0\}$$

for some limiting random variable X. If the distribution of X is determined by its moments, then X_n converges weakly to X.

Proof. Let μ be the distribution of X, and for every $n \in \mathbb{N}$, let μ_n be the distribution of X_n. Given that the sequence of real numbers

$$\mathbf{E}\left[X_n^m \overline{X_n}^m\right] = \int_{\mathbb{C}} z^m \overline{z}^m \, \mathrm{d}\mu_n(z) = \int_{\mathbb{C}} |z|^{2m} \, \mathrm{d}\mu_n(z)$$

converges as $n \to \infty$ for every m, it is bounded by some $K(m) > 0$. It therefore follows from *Markov's Inequality* that

$$\mathbf{E}\left[|X_n|^2 \geqslant \varepsilon^2\right] = \mu_n\big(\{z \in \mathbb{C} : |z|^2 \geqslant \varepsilon^2\}\big) \leqslant \frac{1}{\varepsilon^2} \int_{\mathbb{C}} |z|^2 \, \mathrm{d}\mu_n(z) \leqslant \frac{K(1)}{\varepsilon^2}$$

for every $\varepsilon > 0$, hence the sequence of probability measures (μ_n) is tight. Let $(X_{n_i} : i \in \mathbb{N})$ be a subsequence of (X_n) that converges weakly to some limiting random variable Y. According to the *Skorokhod Representation Theorem*, it may be assumed without loss of generality that X_{n_i} converges to Y almost surely. Let $k, l \in \mathbb{N} \cup \{0\}$ be arbitrary. Define the random variables $W_n = \mathrm{Re}\left(X_n^k \overline{X_n}^l\right)$ and $Z_n = \mathrm{Im}\left(X_n^k \overline{X_n}^l\right)$. Then, W_n and Z_n converge almost surely to $W = \mathrm{Re}\left(Y^k \overline{Y}^l\right)$ and $Z = \mathrm{Im}\left(Y^k \overline{Y}^l\right)$ respectively. Let $m \in \mathbb{N} \cup \{0\}$ be such that $2m \geqslant k + l$. Then, one has

$$|W_n|, |Z_n| \leqslant \left|X_n^k \overline{X_n}^l\right| = |X_n|^{k+l} \leqslant |X_n|^{2m} + 1 = X_n^m \overline{X_n}^m + 1 \leqslant K(m) + 1$$

for every $n \in \mathbb{N}$. Since the probability space on which the X_n are defined is a finite measure space, the constant function $K(m) + 1$ is certainly integrable on that space. It then follows from *Lebesgue's Dominated Convergence Theorem* that $\mathbf{E}[W_n]$ and $\mathbf{E}[Z_n]$ converge to $\mathbf{E}[W]$ and $\mathbf{E}[Z]$ respectively, hence $\mathbf{E}\left[X^k \overline{X}^l\right] = \mathbf{E}\left[Y^k Y^l\right]$, which implies that X and Y have the same distribution since X is determined by its moments. Since the distributions of X_n are tight, this implies that X_n converges weakly to X by Theorem B.7. □

B.2. Real Analysis

Uniform Limit Theorem ([**Mu99**] theorem 21.6)**.** *Let (X, τ) be a topological space and (Y, d) be a metric space. If $\{f_n\}_{n\in\mathbb{N}}$ is a sequence of continuous functions from X to Y that converges uniformly to a function $f : X \to Y$, then f is continuous.*

B.3. Stone-Weierstrass Approximation Theorem

Stone-Weierstrass Approximation Theorem ([**CP11**] theorem 1.1.6)**.** *Let (X, d) be a metric space and $K \subset X$ be compact. Let $\mathcal{A} \subset \mathrm{C}(K, \mathbb{C})$ be an algebra of functions which contains the constant functions; separates points (i.e. for every $x, y \in K$, there exists $f \in \mathcal{A}$ such that $f(x) \neq f(y)$); and which is closed under complex complementation. Then, $\mathcal{A}$ is dense in $\mathrm{C}(K, \mathbb{C})$.*[5]

Corollary. *Let (X, d) be a metric space and $K \subset X$ be compact. Let $\mathcal{A} \subset \mathrm{C}(K, \mathbb{C})$ be an algebra of functions which contains the constant functions; separates points (i.e. for every $x, y \in K$, there exists $f \in \mathcal{A}$ such that $f(x) \neq f(y)$); and which is closed under complex complementation. If $\mathcal{A}$ is closed, then $\mathcal{A} = \mathrm{C}(K, \mathbb{C})$.*

B.4. Complex Analysis

Let $(V, \|\cdot\|)$ be a Banach space (where V is a vector space over $\mathbb{C}$). A function $f : U \to V$ (where $U \subset \mathbb{C}$ is open) is said to be **analytic** if for every $\lambda_0 \in U$, there exists a sequence $\{x_n\}_{n=0}^{\infty} \subset V$ and a neighbourhood U_0 of λ_0 such that for every $\lambda \in U_0$, we have that

$$f(\lambda) = \sum_{n=0}^{\infty} (\lambda - \lambda_0)^n x_n.$$

If $U = \mathbb{C}$, then we say that f is **entire**.

Proposition ([**Ma88**] lemma 5.6)**.** *Let $(V, \|\cdot\|)$ be a Banach space, $U \subset \mathbb{C}$ be open and $f : U \to V$ be analytic. Then, f is continuous.*

Cauchy-Hadamard Theorem ([**Ma88**] theorem 4.3)**.** *Let $(V, \|\cdot\|)$ be a Banach space, $\{x_n\}_{n\in\mathbb{N}} \subset V$ be a sequence and $\lambda_0 \in \mathbb{C}$ be fixed. If we let*

$$R = \left(\limsup_{n\to\infty} \|x_n\|^{1/n} \right)^{-1},$$

then the power series

$$\sum_{n=0}^{\infty} (\lambda - \lambda_0)^n x_n$$

converges if $|\lambda - \lambda_0| < R$ and diverges if $|\lambda - \lambda_0| > R$.[6]

Liouville Theorem ([**EMT04**] theorem 10.3.6)**.** *Let $(\mathcal{A}, \|\cdot\|)$ be a Banach algebra and $f : \mathbb{C} \to \mathcal{A}$ be a function. If f is entire and uniformly bounded (that is, there exists $C \geqslant 0$ such that $\|f(\lambda)\| \leqslant C$ for every $\lambda \in \mathbb{C}$), then f is constant.*

[5] $\mathrm{C}(K, \mathbb{C})$ is equipped with the metric topology induced by the *uniform norm* $\|f - g\|_\infty$.

[6] If $|\lambda - \lambda_0| = R$, the series may or may not converge.

Cauchy's Integral Theorem ([**MH99**] Theorem 2.2.3)**.** *Let $f : U \to \mathbb{C}$ be an analytic function, where U is simply connected. For every closed and piecewise C^1 curve C in U, one has*

$$\oint_C f(z)\ \mathrm{d}z = 0.$$

Morera's Theorem ([**MH99**] Theorem 2.4.10)**.** *Let $f : U \to \mathbb{C}$ be a continuous function, where U is a simply connected open set. If*

$$\oint_C f(z)\ \mathrm{d}z = 0$$

holds for every closed and piecewise C^1 curve C in U, then f is analytic.

B.4.1. Stieltjes Inversion Formula. ([**NS06**] pages 30 to 33, [**HOA07**] Section 1.9)

Let μ be a probability measure on $\big(\mathbb{R}, \mathscr{B}(\mathbb{R})\big)$. The **Stieltjes transform** of μ is defined as the map

$$\begin{array}{rccc} s_\mu : & \mathbb{C} \setminus \mathbb{R}(\mu) & \to & \mathbb{C} \\ & z & \mapsto & \displaystyle\int_{\mathbb{R}} \frac{1}{x - z}\ \mathrm{d}\mu(x). \end{array}$$

The main properties of Stieltjes transforms are summarized in the following propositions.

Proposition B.8. *Let μ be a probability measure on $\big(\mathbb{R}, \mathscr{B}(\mathbb{R})\big)$ and s_μ its Stieltjes transform. Then,*

(1) $|s_\mu(z)| \leqslant \big|\mathrm{Im}(z)\big|^{-1}$ whenever $z \notin \mathbb{R}$;

(2) s_μ is continuous;

(3) s_μ is analytic on every connected open subset of its domain;

(4) $\overline{s_\mu(z)} = s_\mu(\overline{z})$ for every $z \in \mathbb{C} \setminus \mathbb{R}$;[7] *and*

(5) if $\mathrm{Im}(z) > 0$, then $\mathrm{Im}\big(s_\mu(z)\big) > 0$ and vice versa.

Proof. (1). This property follows immediately from the fact that $|x - z| \geqslant |\mathrm{Im}(z)|$ whenever $x \in \mathbb{R}$.

(2). Let $z \in \mathbb{C} \setminus \mathbb{R}$ and $\{z_n : n \in \mathbb{N}\} \subset \mathbb{C} \setminus \mathbb{R}$ be such that $z_n \to z$. Then, since the map $z \mapsto (x - z)^{-1}$ is continuous for every fixed $x \in \mathbb{R}$, if follows that $(x - z_n)^{-1} \to (x - z)^{-1}$. Thus, if one define the maps $f : \mathbb{R} \to \mathbb{C}$ and $f_n : \mathbb{R} \to \mathbb{C}$ as $f(x) = (x - z)^{-1}$ and $f_n(x) = (x - z_n)^{-1}$ for $n \in \mathbb{N}$, it follows that $f_n \to f$ pointwise. Given that f and each f_n is bounded by the constant (and hence integrable) map $x \mapsto \sup\{|\mathrm{Im}(z_n)|, |\mathrm{Im}(z)| : n \in \mathbb{N}\}$, it follows from Lebesgue's Dominated Convergence Theorem that $s_\mu(z_n) \to s_\mu(z)$, as desired.

(3). Let $U \subset \mathbb{C} \setminus \mathbb{R}$ be a connected open set. For any fixed $x \in \mathbb{R}$, the map $f_x(z) = (x - z)^{-1}$ is analytic on $\mathbb{C} \setminus \{x\}$. Thus, it follows from Cauchy's Integral Theorem that for any closed and piecewise C^1 curve $C \coloneqq \gamma : [0, 1] \to U$, one has

$$\oint_C f_x(z)\ \mathrm{d}z = 0.$$

[7] This implies in particular that the Stieltjes transform of a probability measure is completely determined by the values that it takes on $\{a + b\mathrm{i} \in \mathbb{C} : b > 0\}$.

Thus, according to Fubini's Theorem, it follows that

$$\begin{aligned}\int_C s_\mu(z)\ \mathrm{d}z &= \int_{[0,1]} s_\mu\big(\gamma(t)\big)\gamma'(t)\ \mathrm{d}t = \int_{[0,1]}\int_{\mathbb{R}} \frac{\gamma'(t)}{x-\gamma(t)}\ \mathrm{d}x\ \mathrm{d}t \\ &= \int_{\mathbb{R}}\int_{[0,1]} \frac{\gamma'(t)}{x-\gamma(t)}\ \mathrm{d}t\ \mathrm{d}x = \int_{\mathbb{R}} \oint_C f_x(z)\ \mathrm{d}z\ \mathrm{d}x = 0,\end{aligned}$$

which concludes the proof by Morera's Theorem. and the continuity of s_μ.

(4). Given that for any measure μ on $\mathbb{R}$ and integrable function $f : \mathbb{R} \to \mathbb{C}$, one has

$$\overline{\int_{\mathbb{R}} f(x)\ \mathrm{d}\mu(x)} = \int_{\mathbb{R}} \overline{f(x)}\ \mathrm{d}\mu(x),$$

it is clear that $\overline{s_\mu(z)} = s_\mu(\overline{z})$ for every $z \in \mathbb{C} \setminus \mathbb{R}$.

(5). Let $z = a + b\mathrm{i}$. Then,

$$s_\mu(z) = \int_{\mathbb{R}} \frac{1}{x-z}\ \mathrm{d}\mu(x) = \int_{\mathbb{R}} \frac{x-a}{(x-a)^2+b^2}\ \mathrm{d}\mu(x) + \mathrm{i}\int_{\mathbb{R}} \frac{b}{(x-a)^2+b^2}\ \mathrm{d}\mu(x).$$

Thus, if $b > 0$, it clearly is the case that

$$\int_{\mathbb{R}} \frac{b}{(x-a)^2+b^2}\ \mathrm{d}\mu(x) > 0$$

as well. □

Proposition B.9. *Let X be a $\mathbb{R}$-valued random variable whose distribution μ_X has a compact support. If we define $K = \sup\{|x| : x \in \operatorname{supp}(\mu_X)\}$, then,*

$$s_{\mu_X}(z) = -\sum_{n=0}^{\infty} \frac{\mathbf{E}[X^n]}{z^{n+1}}$$

whenever $|z| > K$.

Proof. Knowing that we have the series representation

$$\frac{1}{z-x} = -\sum_{n=0}^{\infty} \frac{x^n}{z^{n+1}}$$

whenever $x \in \mathbb{R}$ and $z \in \mathbb{C}$ are such that $|z| > |x|$, it follows that if $|z| > K$, then

$$s_{\mu_X}(z) = -\int_{\mathbb{R}} \sum_{n=0}^{\infty} \frac{x^n}{z^{n+1}}\ \mathrm{d}\mu_X(x). \tag{B.2}$$

Given that, for a fixed z, the convergence of

$$\sum_{n=0}^{\infty} \frac{x^n}{z^{n+1}}$$

is uniform for x in the support of μ (as it is a compact set), we may permute the sum and the integral in (B.2), which gives the desired result. □

The following theorem, sometimes called the **Stieltjes Inversion Formula**, provides, in principle, a method of *explicitly describing* μ from its Stieltjes transform. If μ has a compact support, it then follows from Proposition B.9 that a measure can be described from its moments.

Theorem B.10 ([**HOA07**] Theorems 1.99, 1.100 and 1.102; and Corollary 1.103)**.** *Let μ be a probability measure on $\big(\mathbb{R}, \mathscr{B}(\mathbb{R})\big)$ and let s_μ be its Stieltjes transform. Then, for every real numbers $a < b$, one has*

$$\frac{2}{\pi} \lim_{\varepsilon \to 0^+} \int_a^b \operatorname{Im}\big(s_\mu(x + \mathrm{i}\varepsilon)\big)\, \mathrm{d}x = \mu\big(\{a\}\big) + \mu\big(\{b\}\big) + 2\mu\big((a,b)\big).$$

Letting F be the distribution function associated to the measure μ (that is, $F(x) = \mu\big((-\infty, x]\big)$), one has

$$\frac{2}{\pi} \lim_{\varepsilon \to 0^+} \int_a^b \operatorname{Im}\big(s_\mu(x + \mathrm{i}\varepsilon)\big)\, \mathrm{d}x = F(b) + \lim_{\varepsilon \to 0^+} F(b - \varepsilon) - F(a) - \lim_{\varepsilon \to 0^+} F(a - \varepsilon).$$

Furthermore, for every point x at which F is differentiable,

$$F'(x) = \frac{1}{\pi} \lim_{\varepsilon \to 0^+} \operatorname{Im}\big(s_\mu(x + \mathrm{i}\varepsilon)\big).$$

Let $f(x)\, \mathrm{d}x$ be the absolutely continuous part of μ.[8] *Then, the limit*

$$\frac{1}{\pi} \lim_{\varepsilon \to 0^+} \operatorname{Im}\big(s_\mu(x + \mathrm{i}\varepsilon)\big)$$

is defined almost everywhere for x, and one has

$$f(x) = \begin{cases} \dfrac{1}{\pi} \lim\limits_{\varepsilon \to 0^+} \operatorname{Im}\big(s_\mu(x + \mathrm{i}\varepsilon)\big) & \text{whenever the limit exists; and} \\ 0 & \text{otherwise.} \end{cases}$$

B.5. Functional Analysis

Proposition B.11 ([**Co13**] proposition 3.2.5)**.** *Let $(V, \|\cdot\|)$ be a normed vector space over a field F. Then V (equipped with the metric topology) is complete if and only if every absolutely convergent series with terms in V is convergent.*

Banach-Alaoglu Theorem ([**EMT04**] theorem 9.7.9 and remak 9.7.10)**.** *Let $(V, \|\cdot\|)$ be a Banach space over $\mathbb{C}$. Then, the set $\{\varphi \in V^* : \|\varphi\|_{\mathrm{op}} \leqslant 1\}$ is compact in the w^*-topology.*

B.5.1. Ideals of Continuous Functions. See [**La93**] pages 55 to 57.

Let X be a compact topological space, and let $\mathrm{C}(X, \mathbb{C})$ be the algebra of continuous functions from X to $\mathbb{C}$. Equip $\mathrm{C}(X, \mathbb{C})$ with the norm topology it inherits from the supremum norm $\|\cdot\|_\infty$.

Lemma B.12. *Let $I \subset \mathrm{C}(X, \mathbb{C})$ be an ideal. Then, either $I = \mathrm{C}(X, \mathbb{C})$, or*

$$\text{there exists } x \in X \text{ such that } f(x) = 0 \text{ for every } f \in I. \tag{B.3}$$

Proof. Let $I \subset \mathrm{C}(X, \mathbb{C})$ be an ideal different from $\mathrm{C}(X, \mathbb{C})$. Suppose that for every $x \in X$, there exists $f_x \in I$ such that $f_x(x) \neq 0$. Then, since f_x is continuous, there exists an open set $U_x \subset X$ such that $|f(y)| > 0$ for every $y \in U_x$. Given that the collection $(U_x : x \in X)$ is a cover for X and that X is compact, there exists $x_1, \ldots, x_n \in X$ such that $X \subset \bigcup_i U_{x_i}$. Let

$$g = f_{x_1}\overline{f_{x_1}} + \cdots + f_{x_n}\overline{f_{x_n}} = |f_{x_1}|^2 + \cdots + |f_{x_n}|^2.$$

[8] That is, for every Borel set $A \subset \mathbb{R}$, one has $\mu(A) = \int_A f(x)\, \mathrm{d}x + \nu(A)$, where μ and ν are singular (see [**Br71**]). In the case where μ is absolutely continuous, then $\nu = 0$.

Since I is an ideal, $g \in I$. Furthermore, since $g(X) \subset (0, \infty]$, it follows that the inverse function g^{-1} exists and is continuous. Therefore, $gg^{-1} = \mathbf{1}_X \in I$, which contradicts that $I \neq \mathrm{C}(X, \mathbb{C})$. □

Theorem B.13. *There is a one to one correspondance between open sets of $U \subset X$ and the closed ideals of $\mathrm{C}(X, \mathbb{C})$, which are all of the form*

$$\big\{ f \in \mathrm{C}(X, \mathbb{C}) : f \text{ vanishes outside of } U \big\}. \tag{B.4}$$

Proof. Clearly, every set of the form (B.4) is an ideal. Moreover, if a uniformly convergent sequence of functions $\{f_n : n \in \mathbb{N}\} \subset \mathrm{C}(X, \mathbb{C})$ is such that f_n vanishes outside of an open set U for every n, then the same must be true for the limit function (as uniform convergence implies pointwise convergence). Consequently, every ideal of the form (B.4) is a closed ideal.

Let $I \subset \mathrm{C}(X, \mathbb{C})$ be a closed ideal. Define the set

$$F \coloneqq \bigcap_{f \in I} \{x \in X : f(x) = 0\} = \bigcap_{f \in I} f^{-1}\big(\{0\}\big).$$

Since every function in I is continuous, F is closed, and thus its complement $U \coloneqq F^c$ is open. Clearly, every function in I vanishes outside of U. We claim that I is in fact equal to (B.4). Let $f \in \mathrm{C}(X, \mathbb{C})$ be such that f vanishes outside of U. For a fixed $0 < \varepsilon < 1$, let

$$V \coloneqq \big\{ x \in X : |f(x)| < \varepsilon \big\}.$$

Clearly, V is open and contains F, and its complement $G = V^c$ is closed and is contained in U. According to the definition of the set U, for every $x \in U$, there exists $f_x \in I$ such that $f_x(x) \neq 0$. Then, since f_x is continuous, there exists an open set $U_x \subset X$ such that $|f(y)| > 0$ for every $y \in U_x$. Given that the collection $(U_x : x \in X)$ is a cover for $G \subset U$ and that G is compact (as it is a closed subset of a compact set), there exists $x_1, \ldots, x_n \in U$ such that $G \subset \bigcup_i U_{x_i}$. Let

$$g = f_{x_1}\overline{f_{x_1}} + \cdots + f_{x_n}\overline{f_{x_n}} = |f_{x_1}|^2 + \cdots + |f_{x_n}|^2.$$

Since I is an ideal, $g \in I$. Furthermore, $g(x) > 0$ for every $x \in G$, and $1 + g(x) > 1$ for every $x \in X$. Therefore, for every $n \in \mathbb{N}$, the function $h_n = n \cdot g(1 + n \cdot g)^{-1}$ is well defined and in the ideal I, and h_n converges uniformly to 1 on G, hence $h_n f$ converges uniformly to f on G. Given that $0 \leqslant h_n(x) \leqslant 1$ on all of X, by definition of V, it must be the case that $|h_n(x) f(x)| \leqslant \varepsilon$ for every $x \in V$. Thus, by taking $\varepsilon \to 0$, $h_n f$ converges uniformly to f on all of X, hence f is in the closure of I. Since it was assumed that I is closed, it follows that $f \in I$, which concludes the proof of the theorem. □

B.6. Measure Theory

Riesz Representation Theorem (**[Co13]** Theorem 7.2.8)**.** *Let X be a topological space, and let $\mathrm{C}_c(X, \mathbb{C})$ be the set of continuous functions with compact support from X to $\mathbb{C}$. Let $I : \mathrm{C}_c(X, \mathbb{C}) \to \mathbb{C}$ be a positive linear functional, that is, $I(f\overline{f}) \geqslant 0$ for every $f \in \mathrm{C}_c(X, \mathbb{C})$. Then, there exists a unique regular measure μ on $\mathbb{C}$ such that*

$$I(f) = \int_{\mathbb{C}} f(z) \, \mathrm{d}\mu(z), \qquad\qquad f \in \mathrm{C}_c(X, \mathbb{C}).$$

B.7. Harmonic Analysis

Let $f : \mathbb{C} \to \mathbb{R}$ be integrable with respect to the Lebesgue measure. The (radial) **Fourier transform** of f, which we denote $\widehat{f}$, is defined as

$$\widehat{f}(t) = \int_{\mathbb{R}} f(x) \mathrm{e}^{\mathrm{i}tx} \, \mathrm{d}x, \qquad t \in \mathbb{R}.$$

Remark B.14. If f is the density function of a random variable X and $M_X : \mathbb{C} \to \mathbb{C}$ is the moment generating function of X (see Subsection 4.2.1), then $\widehat{f}(t) = M_X(it)$ for all $t \in \mathbb{R}$.

Theorem B.15 (Inverse Fourier Transform, **[Bi95]** Theorem 26.2 and equation (26.20))**.** *Let $f : \mathbb{R} \to \mathbb{C}$ be integrable and $\widehat{f}$ be its Fourier transform. Then,*

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}tx} \widehat{f}(t) \, \mathrm{d}t, \qquad x \in \mathbb{R}.$$

Appendix C

Topology

C.1. Bases and Sub-bases

See [**Mu99**] section 13.

Let X be a set. A **basis** on X is a collection $(A_i)_{i\in I}$ of subsets of X such that

- For every $x \in X$, there exists $i \in I$ such that $x \in A_i$; and
- If $x \in B_i \cap B_j$ for some $i, j \in I$, then there exists $k \in I$ such that $x \in A_k$ and $A_k \subset A_i \cap A_j$.

The topology τ generated by the basis $(A_i)_{i\in I}$ is defined as follows: A subset U of X is open in τ if and only if for every $x \in U$, there exists $i \in I$ such that $x \in A_i$ and $A_i \subset U$. A **sub-basis** on X is a collection $(B_j)_{j\in J}$ of subsets of X such that

$$\bigcup_{j\in J} B_j = X.$$

The topology τ' generated by the sub-basis $(B_j)_{j\in J}$ is that which admits the following collection as a basis:

$$\tau' = \{B_{j_1} \cap \cdots \cap B_{j_t} : t \in \mathbb{N} \text{ and } j_1, ..., j_t \in J\} \cup \{X\} \cup \{\varnothing\}.$$

Proposition ([**Mu99**] section 13 exercise 5)**.** *Let X be a set and $(A_i)_{i\in I}$ be a basis on X. Then, the topology generated by $(A_i)_{i\in I}$ is the intersection of every topology that contains $(A_i)_{i\in I}$. The same is true for sub-bases.*

Proposition ([**Mu99**] page 103)**.** *Let (X, τ) and (Y, τ') be topological spaces, where τ' is generated by the basis $(A_i)_{i\in I}$ on Y. If $f^{-1}(A_i)$ is open in τ for every $i \in I$, then f is continuous. The same is true for sub-bases.*

C.2. Compact Sets

Heine-Borel Theorem ([**Mu99**] theorem 27.3)**.** *A set is compact in Euclidean space (i.e., $\mathbb{R}^n$ or $\mathbb{C}^n$ for some $n \in \mathbb{N}$ equipped with the Euclidean topology) if and only if it is closed and bounded.*

Proposition ([**Ho95**] theorem 1.8)**.** *Let (X, d) be a metric space. If $A \subset X$ is compact in the metric topology, then it is closed and bounded.*

Proposition ([**Mu99**] theorem 26.5)**.** *Let (X, τ) and (Y, τ') be topological spaces. If $f : X \to Y$ is continuous and $A \subset X$ is compact in τ, then $f(A) \subset Y$ is compact in τ'.*

C.3. Product Topology

See [**Mu99**] section 19.

Let $(X_i, \tau_i)_{i\in I}$ be a collection of topological spaces. For every $j \in I$, define the j^{th}-coordinate projection map π_j as follows:

$$\begin{array}{rccc} \pi_j : & \prod_{i\in I} X_i & \to & X_j \\ & (x_i)_{i\in I} & \mapsto & x_j. \end{array}$$

If, for every $i \in I$, we define

$$B_i = \{\pi_i^{-1}(U) : U \text{ is open in } \tau_i\},$$

then $\{B_i\}_{i\in I}$ is a sub-basis on the cartesian product $\prod_i X_i$, and the topology generated by this sub-basis is called the **product topology** on the cartesian product.

Tychonoff's Theorem (see [**Wr94**])**.** *Let $(X_i, \tau_i)_{i\in I}$ be a collection of compact topological spaces. Then, $\prod_i X_i$ equipped with the product topology is a compact topological space.*

C.4. w^*-topology

See [**EMT04**] section 9.7.

Let V be a vector space over $\mathbb{C}$ and V^* be the dual space of V. For every $\varphi \in V^*$, $x \in V$ and $\varepsilon > 0$, define the set

$$U(\varphi, x, \varepsilon) = \{\psi \in V^* : |\psi(x) - \varphi(x)| < \varepsilon\}. \tag{C.1}$$

We define the ***w**-topology** on V^* as the topology which admits the following collection as a sub-basis:

$$\{U(\varphi, x, \varepsilon) : \varphi \in V^*,\ x \in V,\ \varepsilon > 0\}.$$

Remark. Notice that the w^*-topology may also be generated by the collection of all sets of the form

$$U(\lambda, x, \varepsilon) = \{\psi \in V^* : |\psi(x) - \lambda| < \varepsilon\}, \tag{C.2}$$

where $\lambda \in \mathbb{C}$, $x \in V$ and $\varepsilon > 0$. Clearly, every set of the form (C.1) is a set of the form (C.2) with $\lambda = \varphi(x)$. The fact that every set of the form (C.2) is open in the topology generated by (C.1) is also easy to see: Let $\lambda \in \mathbb{C}$, $x \in V$ and $\varepsilon > 0$ be arbitrary. If $\varphi(x) = 0$ for every $\varphi \in V^*$,

then $U(\lambda, x, \varepsilon)$ is either V^* or the empty set, depending on whether $|\lambda| < \varepsilon$ or $|\lambda| \geqslant \varepsilon$. If there exists $\varphi \in V^*$ such that $\varphi(x) = \lambda \neq 0$, then $\frac{\lambda}{\varphi(x)}\varphi(x) = \lambda$, and thus $U(\lambda, x, \varepsilon) = U(\frac{\lambda}{\varphi(x)}\varphi, x, \varepsilon)$.

Appendix D

Discrete Mathematics

D.1. Partitions

See [**NS06**] Lecture 9.

Definition D.1. Let A be a set. The collection of sets $\pi = (A_i : i \in I)$ is called a **partition** of A if and only if the following conditions hold:

(1) $A_i \subset A$ for every $i \in I$;

(2) the A_i are pairwise disjoint, that is, for every distinct $i, j \in I$, one has $A_i \cap A_j = \varnothing$; and

(3) the A_i cover A, that is, $A = \bigcup_{i \in I} A_i$.

The elements A_i are called the **blocks** of the partition π.

In this thesis, we are mostly interested in partitions of the set $\{1, \ldots, n\}$ for some $n \in \mathbb{N}$. In this context, we adopt the following notations: Given a fixed integer $n \in \mathbb{N}$, let $P(n)$ denote the set of partitions of $\{1, \ldots, n\}$, and let us denote $1_n \coloneqq \big\{\{1, \ldots, n\}\big\}$ and $0_n \coloneqq \big\{\{1\}, \ldots, \{n\}\big\}$.

Definition D.2. Let $n \in \mathbb{N}$ be fixed, and let $\pi = \{B_1, \ldots, B_t\} \in P(n)$ be a partition of $\{1, \ldots, n\}$ (where the B_i are the blocks of π). π is said to be a **noncrossing partition** if and only if for every quadruple $n_1, m_1, n_2, m_2 \in \{1, \ldots, n\}$ such that $n_1 < m_1 < n_2 < m_2$, it cannot be the case that there exists $i, j \leqslant t$ such that $n_1, n_2 \in B_i$ and $m_1, m_2 \in B_j$. The set of noncrossing partitions of $\{1, \ldots, n\}$ is denoted $NC(n)$.

Notation D.3. Let $n \in \mathbb{N}$ be fixed. Given partitions $\pi = \{B_1, \ldots, B_t\}$ and $\sigma = \{C_1, \ldots, C_s\}$ in $P(n)$, we write $\pi \leqslant \sigma$ whenever for every $i \leqslant t$, there exists $j \leqslant s$ such that $B_i \subset C_j$. This relation defines a partial order on $NC(n)$ called the **reversed refined order** (note that this order makes $P(n)$ and $NC(n)$ a lattice, see [**NS06**] Lecture 9).

Proposition D.4. *Let $n \in \mathbb{N}$ be fixed and $\pi, \sigma \in NC(n)$, and define the sets*

$$U(\pi, \sigma) = \{\tau \in NC(n) : \tau \geqslant \pi \text{ and } \tau \geqslant \sigma\}$$

and

$$L(\pi, \sigma) = \{\tau \in NC(n) : \tau \leqslant \pi \text{ and } \tau \leqslant \sigma\}.$$

Then, the sets $U(\pi,\sigma)$ and $L(\pi,\sigma)$ are nonempty, the set $U(\pi,\sigma)$ has a minimal element τ_0 (that is, τ_0 is such that $\tau_0 \leqslant \tau$ for each $\tau \in U(\pi,\sigma)$), and the set $L(\pi,\sigma)$ has a maximal element τ_0' (that is, τ_0 is such that $\tau_0 \geqslant \tau$ for each $\tau \in L(\pi,\sigma)$).

Notation D.5. Let $n \in \mathbb{N}$ be fixed and $\pi, \sigma \in NC(n)$, let $U(\pi,\sigma)$ and $L(\pi,\sigma)$ be defined as in the previous proposition.

(1) The minimal element of $U(\pi,\sigma)$ will henceforth be denoted $\pi \vee \sigma$, and will be called the **join** of π and σ.

(2) The maximal element of $L(\pi,\sigma)$ will henceforth be denoted $\pi \wedge \sigma$, and will be called the **meet** of π and σ.

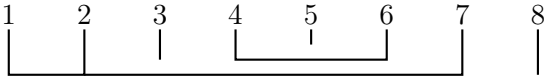
Definition D.6. Let $n \in \mathbb{N}$ and $\pi \in NC(n)$ be arbitrary. The **Kreweras complement** of π, denoted $K(\pi)$ is the largest noncrossing partition of $\{\overline{1}, \overline{2}, \ldots, \overline{n}\}$ (and thus, $K(\pi) \in NC(n)$) such that $\pi \cup K(\pi)$ is a noncrossing partition of $\{1, \overline{1}, 2, \overline{2}, \ldots, n, \overline{n}\}$.

The following example shows how one can obtain the Kreweras complement of a noncrossing partition in practice:

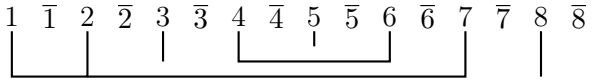
Example D.7 ([**NS06**] Example 9.22)**.** Consider the partition

$$\pi = \big\{\{1,2,7\},\{3\},\{4,6\},\{5\},\{8\}\big\} \subset NC(8).$$

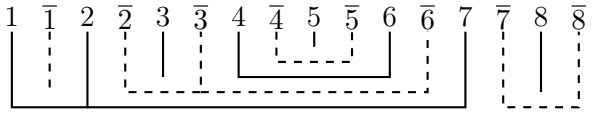
This partition can be graphically represented as follows:



Then, one adds the symbol $\overline{1}$ to the right of 1, the symbol $\overline{2}$ to the right of 2, and so on, which yields:



and then complete the above graphical representation way that gives the biggest possible noncrossing partition of $\{1, \overline{1}, 2, \overline{2}, \ldots, 8, \overline{8}\}$:



The Kreweras complement of π is then given by the resulting partition of $\{\overline{1}, \ldots, \overline{n}\}$, that is,

$$K(\pi) = \big\{\{1\},\{2,3,6\},\{4,5\},\{7,8\}\big\}.$$

D.2. Möbius Function

See [**NS06**] Lecture 10.

Notation D.8. Let $(P, \leqslant)$ be a partially ordered set (such as $P(n)$ or $NC(n)$ with the reversed refinement order). Let us denote

$$P^{(2)} = \big\{(\pi,\sigma) \in P \times P : \pi \leqslant \sigma\big\}.$$

Definition D.9. Given two functions $f, g : P^{(2)} \to \mathbb{C}$, define their convolution (denoted $f * g$) as the map

$$(f * g)(\pi, \sigma) = \sum_{\tau \in P:\ \pi \leqslant \tau \leqslant \sigma} f(\pi, \tau) g(\tau, \sigma).$$

Definition D.10. Let $(P, \leqslant)$ be a finite partially ordered set (with n elements). Let $E = (\pi_1, \ldots, \pi_n)$ be an enumeration of the elements in P such that for every $1 \leqslant i < j \leqslant n$, either π_i and π_j are incomparable or $\pi_i < \pi_j$.[1] For any function $f : P^{(2)} \to \mathbb{C}$, define the upper triangular matrix $\Delta(f, E) = \{\Delta_{ij}(f, E) : 1 \leqslant i, j \leqslant C_n\}$ as

$$\Delta_{ij}(f, E) = \begin{cases} 0 & \text{if } i > j; \\ f(\pi_i, \pi_j) & \text{if } i = j, \text{ or } i < j \text{ and } \pi_i < \pi_j; \text{ and} \\ 0 & \text{if } i < j \text{ and } \pi_i \text{ and } \pi_j \text{ are not comparable.} \end{cases}$$

Proposition D.11. *Let $(P, \leqslant)$ be a finite partially ordered set (with n elements), and let $E = (\pi_1, \ldots, \pi_n)$ be an enumeration of the elements in P such that for every $1 \leqslant i < j \leqslant n$, either π_i and π_j are incomparable or $\pi_i < \pi_j$. Given two functions $f, g : P^{(2)} \to \mathbb{C}$, one has*

$$\Delta(f * g, E) = \Delta(f, E)\Delta(g, E).$$

Notation D.12. Let $(P, \leqslant)$ be a finite partially ordered set (with n elements), and let $E = (\pi_1, \ldots, \pi_n)$ be an enumeration of the elements in P such that for every $1 \leqslant i < j \leqslant n$, either π_i and π_j are incomparable or $\pi_i < \pi_j$. Let $\delta : P^{(2)} \to \mathbb{C}$ be defined as

$$\delta(\pi, \sigma) = \begin{cases} 1 & \text{if } \pi = \sigma; \text{ and} \\ 0 & \text{otherwise,} \end{cases}$$

and let $\zeta : P^{(2)} \to \mathbb{C}$ be defined as $\zeta(\pi, \sigma) = 1$ for every $(\pi, \sigma) \in P^{(2)}$. We will henceforth denote $\Delta(E) \coloneqq \Delta(\zeta, E)$.

Definition D.13. Let $(P, \leqslant)$ be a finite partially ordered set (with n elements), and let $E = (\pi_1, \ldots, \pi_n)$ be an enumeration of the elements in P such that for every $1 \leqslant i < j \leqslant n$, either π_i and π_j are incomparable or $\pi_i < \pi_j$. The inverse under convolution of the function $\zeta : P^{(n)} \to \mathbb{C}$ introduced in Notation D.12 is called the *Möbius function* of P, and it is denoted μ (that is, $\Delta(\zeta * \mu, E) = \Delta(\mu * \zeta, E) = \Delta(\delta, E)$).

Remark D.14. It is clear from Proposition D.11 that $\Delta(\mu, E) = \Delta(E)^{-1}$ (where $\Delta(E) = \Delta(\zeta, E)$, as per Notation D.12). That is, the entries of $\Delta(E)$ are given by

$$\Delta_{ij}(E) = \begin{cases} 0 & \text{if } i > j; \\ 1 & \text{if } i \leqslant j \text{ and } \pi_i \leqslant \pi_j; \text{ and} \\ 0 & \text{if } i < j \text{ and } \pi_i, \pi_j \text{ are not comparable,} \end{cases}$$

and for every $i \leqslant j$ such that $\pi_i \leqslant \pi_j$, $\mu(\pi_i, \pi_j)$ is given by the entry on the row i and column j of $\Delta(E)^{-1}$.

[1] Note that such an enumeration is possible in every partially ordered set. See **[NS06]** Exercise 10.25 for an outline of an algorithm to produce such an order.

Proposition D.15 (Möbius Inversion Formula, [**NS06**] Proposition 10.6)**.** *Let $(P, \leqslant)$ be a finite partially ordered set (with n elements), and let $E = (\pi_1, \ldots, \pi_n)$ be an enumeration of the elements in P such that for every $1 \leqslant i < j \leqslant n$, either π_i and π_j are incomparable or $\pi_i < \pi_j$. Let $f, g : P \to \mathbb{C}$ be functions. Then,*

$$f(\pi) = \sum_{\sigma \in P:\ \sigma \leqslant \pi} g(\sigma), \qquad \pi \in P$$

holds if and only if the following also holds

$$g(\pi) = \sum_{\sigma \in P:\ \sigma \leqslant \pi} f(\sigma)\mu(\sigma, \pi), \qquad \pi \in P.$$

We now consider two examples of computations with the Möbius function in the partially ordered set of noncrossing partitions.

Example D.16. The only partitions in $NC(2)$ are $0_2 = \big\{\{1\},\{2\}\big\}$ and $1_2 = \big\{\{1,2\}\big\}$. If $E = (0_2, 1_2)$, then

$$\Delta(E) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix},$$

and thus

$$\Delta(E)^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}.$$

Therefore, $\mu(0_2, 0_2) = \mu(1_2, 1_2) = 1$, and $\mu(0_2, 1_2) = -1$.

Example D.17. Consider $NC(4)$. In this case, there are 14 noncrossing partitions. Define $E = (\pi_1, \ldots, \pi_{14})$, where

$$\begin{array}{rclrcl}
\pi_1 &=& 0_4 & \pi_8 &=& \big\{\{1,4\},\{2,3\}\big\} \\
\pi_2 &=& \big\{\{1\},\{2\},\{3,4\}\big\} & \pi_9 &=& \big\{\{1,2\},\{3,4\}\big\} \\
\pi_3 &=& \big\{\{1\},\{2,3\},\{4\}\big\} & \pi_{10} &=& \big\{\{1,2,3\},\{4\}\big\} \\
\pi_4 &=& \big\{\{1,2\},\{3\},\{4\}\big\} & \pi_{11} &=& \big\{\{1\},\{2,3,4\}\big\} \\
\pi_5 &=& \big\{\{1,4\},\{2\},\{3\}\big\} & \pi_{12} &=& \big\{\{1,3,4\},\{2\}\big\} \\
\pi_6 &=& \big\{\{1,3\},\{2\},\{4\}\big\} & \pi_{13} &=& \big\{\{1,2,4\},\{3\}\big\} \\
\pi_7 &=& \big\{\{1\},\{2,4\},\{3\}\big\} & \pi_{14} &=& 1_4
\end{array}.$$

One easily checks that E satisfies the condition in Definition D.10. Furthermore, by observing which pairs π_i, π_j are incomparable or such that $\pi_i < \pi_j$ (with $i < j$), we obtain

$$\Delta(E) = \begin{array}{c} \\ \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \\ \pi_5 \\ \pi_6 \\ \pi_7 \\ \pi_8 \\ \pi_9 \\ \pi_{10} \\ \pi_{11} \\ \pi_{12} \\ \pi_{13} \\ \pi_{14} \end{array} \begin{array}{c} \begin{array}{cccccccccccccc} \pi_1 & \pi_2 & \pi_3 & \pi_4 & \pi_5 & \pi_6 & \pi_7 & \pi_8 & \pi_9 & \pi_{10} & \pi_{11} & \pi_{12} & \pi_{13} & \pi_{14} \end{array} \\ \left[\begin{array}{cccccccccccccc} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \end{array}.$$

With the aid of computer software capable of matrix calculations (such as *Maple* or *Matlab*), one easily obtains

$$\Delta(E)^{-1} = \begin{array}{c} \\ \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \\ \pi_5 \\ \pi_6 \\ \pi_7 \\ \pi_8 \\ \pi_9 \\ \pi_{10} \\ \pi_{11} \\ \pi_{12} \\ \pi_{13} \\ \pi_{14} \end{array} \begin{array}{c} \begin{array}{cccccccccccccc} \pi_1 & \pi_2 & \pi_3 & \pi_4 & \pi_5 & \pi_6 & \pi_7 & \pi_8 & \pi_9 & \pi_{10} & \pi_{11} & \pi_{12} & \pi_{13} & \pi_{14} \end{array} \\ \left[\begin{array}{cccccccccccccc} 1 & -1 & -1 & -1 & -1 & -1 & -1 & 1 & 1 & 2 & 2 & 2 & 2 & -5 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 & -1 & 0 & 2 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & -1 & -1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & -1 & 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & -1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \end{array}.$$

Definition D.18. Let $n \in \mathbb{N}$ be fixed. A function $f : NC(n) \to \mathbb{C}$ is said to be **multiplicative** if there exists a collection of complex numbers $\{\alpha_B : B \subset \{1, \ldots, n\}\}$ such that for every partition $\pi \in NC(n)$, one has

$$f(\pi) = \prod_{B \in \pi} \alpha_B.$$

Theorem D.19 ([**NS06**] Theorem 10.23)**.** *Let $(f_n : n \in \mathbb{N})$ and $(g_n : n \in \mathbb{N})$ be two sequences of functions such that for every $n \in \mathbb{N}$, g_n and f_n are defined on $NC(n)$ and take values in $\mathbb{C}$,*

g_n and f_n are multiplicative functions, and

$$f_n(\pi) = \sum_{\sigma \in NC(n):\ \sigma \leqslant \pi} g(\sigma) \text{ and }, \qquad \pi \in NC(n)$$

and

$$g_n(\pi) = \sum_{\sigma \in NC(n):\ \sigma \leqslant \pi} f(\sigma)\mu(\sigma, \pi), \qquad \pi \in NC(n)$$

(that is, f_n and g_n are related as in Proposition D.15). Define the sequences of complex numbers $(\alpha_n : n \in \mathbb{N})$ and $(\beta_n : n \in \mathbb{N})$ as $\alpha_n = f_n(1_n)$ and $\beta_n = g_n(1_n)$ for all $n \in \mathbb{N}$. Then, the formal power series

$$F(z) = 1 + \sum_{n=1}^{\infty} \alpha_n z^n$$

and

$$G(z) = 1 + \sum_{n=1}^{\infty} \beta_n z^n$$

are such that $G\big(zF(z)\big) = F(z)$, and $F\big(z/G(z)\big) = G(z)$.

D.3. Catalan Numbers

See [**Ko09**]. For every $n \in \mathbb{N}$, the **Catalan number** of order n is defined as

$$C_n = \frac{1}{n+1}\binom{2n}{n}.$$

Segner's Recursive Formula. *Let $n \geqslant 1$. Then,*

$$C_n = \sum_{m=1}^{n} C_{m-1} C_{n-m}.$$

Bibliography

[AGZ09] G. W. Anderson, A. Guionnet, O. Zeitouni, **An introduction to random matrices.** Cambridge Studies in Advanced Mathematics, 118. *Cambridge University Press, Cambridge,* 2010.

[At75] A, Atzmon, **A moment problem for positive measures on the unit disc.** Pacific J. Math. **59** (1975), no. 2, 317–325.

[BCR89] C. Berg, J. P. R. Christensen, P. Ressel, **Harmonic analysis on semigroups. Theory of positive definite and related functions.** Graduate Texts in Mathematics, 100. *Springer-Verlag, New York,* 1984.

[Be99] J. A. Beachy, **Introductory lectures on rings and modules.** London Mathematical Society Student Texts, 47. *Cambridge University Press, Cambridge,* 1999.

[Bi89] T. M. Bisgaard, **The two-sided complex moment problem.** Ark. Mat. 27 (1989), no. 1, 23–28.

[Bi95] P. Billingsley, **Probability and measure.** Third edition. Wiley Series in Probability and Mathematical Statistics. A Wiley-Interscience Publication. *John Wiley & Sons, Inc., New York,* 1995.

[Bi99] P. Billingsley, **Convergence of probability measures.** Second edition. Wiley Series in Probability and Statistics: Probability and Statistics. A Wiley-Interscience Publication. *John Wiley & Sons, Inc., New York,* 1999.

[BMS13] S. Belinschi, T. Mai, R. Speicher, **Analytic subordination theory of operator-valued free additive convolution and the solution of a general random matrix problem.** arXiv preprint 1303.3196 version 3, 2013.

[Br71] J. K. Brooks, **Classroom Notes: The Lebesgue Decomposition Theorem for Measures.** Amer. Math. Monthly 78 (1971), no. 6, 660–662.

[BSTV15] S. Belinschi, R. Speicher, J. Treilhard, C. Vargas, **Operator-Valued Free Multiplicative Convolution: Analytic Subordination Theory and Applications to Random Matrix Theory.** To appear in Int. Math. Res. Not. (2015).

[CG11] G. Chang, C. Xu, **Generalization and probabilistic proof of a combinatorial identity.** *Amer. Math. Monthly* 118 (2011), no. 2, 175–177.

[CM14] B. Collins, C. Male, **The strong asymptotic freeness of Haar and deterministic matrices.** *Ann. Sci. Čc. Norm. Supér. (4)* 47 (2014), no. 1, 147–163.

[Co13] D. L. Cohn, **Measure theory.** Second edition. BirkhŁuser Advanced Texts: Basler Lehrb§cher. *BirkhŁuser/Springer, New York,* 2013.

[Co94] A. Connes, **Noncommutative geometry.** *Academic Press, Inc., San Diego, CA,* 1994.

[CP11] I. Chalendar, J. R. Partington, **Modern approaches to the invariant-subspace problem.** Cambridge Tracts in Mathematics, 188. *Cambridge University Press, Cambridge,* 2011.

[Da96] K. R. Davidson, **C^*-algebras by example.** Fields Institute Monographs, 6. *American Mathematical Society, Providence, RI,* 1996.

[DB86] R. S. Doran, V. A. Belfi, **Characterizations of C^*-algebras.** The Gel'fand-Naimark theorems. Monographs and Textbooks in Pure and Applied Mathematics, 101. *Marcel Dekker, Inc., New York,* 1986.

[Di77] J. Dixmier, ***C**-algebras**, Translated from the French by Francis Jellett. North-Holland Mathematical Library, Vol. 15. *North-Holland Publishing Co., Amsterdam-New York-Oxford,* 1977.

[EMT04] Y. Eidelman, V. D. Milman, A. Tsolomitis, **Functional Analysis. An Introduction.** Graduate Studies in Mathematics, 66. *American Mathematical Society, Providence, RI,* 2004.

[Gu14] A. Guionnet, **Free probability and random matrices**, second paper in **Modern Aspects of Random Matrix Theory.** Papers from the 2013 American Mathematical Society (AMS) Short Course held in San Diego, CA, January 6–7, 2013. Edited by Van H. Vu. Proceedings of Symposia in Applied Mathematics, 72. *American Mathematical Society, Providence, RI,* 2014.

[He86] I. N. Herstein, **Abstract algebra.** *Macmillan Publishing Company, New York,* 1986.

[Ho95] N. R. Howes, **Modern analysis and topology.** Universitext. *Springer-Verlag, New York,* 1995.

[HOA07] A. Hora, N. Obata, L. Accardi, **Quantum probability and spectral analysis of graphs.** Theoretical and Mathematical Physics. *Springer, Berlin,* 2007.

[It84] K. Ito, **Introduction to probability theory.** Translated from the Japanese by the author. *Cambridge University Press, Cambridge,* 1984.

[Ka09] E. Kaniuth, **A course in commutative Banach algebras.** Graduate Texts in Mathematics, 246. *Springer, New York,* 2009.

[Ko56] A. N. Kolmogorov, **Foundations of the theory of probability.** Translation edited by Nathan Morrison, with an added bibliography by A. T. Bharucha-Reid. *Chelsea Publishing Co., New York,* 1956.

[Ko09] T. Koshy, **Catalan numbers with applications.** *Oxford University Press,* Oxford, 2009.

[La93] S. Lang, **Real and functional analysis.** Third edition. Graduate Texts in Mathematics, 142. *Springer-Verlag, New York,* 1993.

[Ma88] I. J. Maddox, **Elements of functional analysis.** Second edition. *Cambridge University Press, Cambridge,* 1988.

[MH99] J. E. Mardsen, M. J. Hoffman, **Basic complex analysis.** Second edition. *W. H. Freeman and Company, New York,* 1987.

[MKS66] W. Magnus, A. Karras, D. Solitar, **Combinatorial group theory: Presentations of groups in terms of generators and relations.** Interscience Publishers [John Wiley & Sons, Inc.], *New York-London-Sydney,* 1966.

[Mu99] J. Munkres, **Topology: a first course.** Second edition. *Prentice-Hall, Inc., Upper Saddle River, N.J.* 2000.

[NS06] A. Nica, R. Speicher, **Lectures on the combinatorics of free probability.** London Mathematical Society Lecture Note Series, 335. *Cambridge University Press, Cambridge,* 2006.

[Sp86] T. Speed, **Cumulants and partition lattices.** Austral. J. Statist. 25 (1983), no. 2, 378–388.

[Sp97] R. Speicher, **On Universal Products**, *Free probability theory* (Waterloo, ON, 1995), 257–266, Fields Inst. Commun., 12, *Amer. Math. Soc., Providence, RI,* 1997.

[Sp09] R. Speicher **Free probability theory.** *The Oxford handbook of random matrix theory,* 452–470, *Oxford Univ. Press, Oxford,* 2011.

[Ta11] T. Tao, **An introduction to measure theory.**| Graduate Studies in Mathematics, 126. *American Mathematical Society, Providence, RI,* 2011.

[Ta12] T. Tao, **Topics in random matrix theory.** Graduate Studies in Mathematics, 132. *American Mathematical Society, Providence, RI,* 2012.

[Va06] J. C. Varilly, **An introduction to noncommutative geometry.** EMS Series of Lectures in Mathematics. *European Mathematical Society (EMS), Zurich,* 2006.

[VDN92] D. Voiculescu, K. Dykema, A. Nica, **Free random variables.** CRM Monograph Series, 1. *American Mathematical Society, Providence, RI,* 1992.

[Vo86] D. Voiculescu **Addition of certain noncommuting random variables.** *J. Funct. Anal.* 66 (1986), no. 3, 323–346.

[Wr94] D. G. Wright, **Tychonoff's Theorem**, *Proc. Amer. Math. Soc.* 120 (1994), no. 3, 985–987.

Index