Universidade Federal Rural de Pernambuco
Departamento de Estatística e Informática

Pós Graduação em Biometria e Estatística Aplicada

$*$-DISTRIBUIÇÕES E APLICAÇÃO À COMPUTAÇÃO QUÂNTICA

Edneide Florivalda Ramos Ramalho

DISSERTAÇÃO DE MESTRADO

Recife - PE
2015

Universidade Federal Rural de Pernambuco
Departamento de Estatística e Informática

Edneide Florivalda Ramos Ramalho

$*$-DISTRIBUIÇÕES E APLICAÇÃO À COMPUTAÇÃO QUÂNTICA

Trabalho apresentado ao Programa de Pós Graduação em Biometria e Estatística Aplicada do Departamento de Estatística e Informática da Universidade Federal Rural de Pernambuco como requisito parcial para obtenção do grau de Mestre em Biometria e Estatística Aplicada.

Orientador: *Cláudio Tadeu Cristino*

Recife - PE
2015

Ficha catalográfica

R165d Ramalho, Edneide Florivalda Ramos
* - distribuição e aplicação à computação quântica / Edneide Florivalda Ramos Ramalho. -- 2015.
99 f.: il.

Orientador: Cláudio Tadeu Cristino.
Dissertação (Mestrado em Biometria e Estatística Aplicada) – Universidade Federal Rural de Pernambuco, Departamento de Estatística e Informática, Recife, 2015.
Referências.

1. Algoritmo de Grover 2. Probabilidade não comutativa 3. Distribuições não cumulativas 4. Medidas I. Cristino, Cláudio Tadeu, orientador II. Título

CDD 574.018

UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA

$*$-Distribuições e Aplicação à Computação Quântica

Edneide Florivalda Ramos Ramalho

Dissertação julgada adequada para obtenção do título de mestre em Biometria e Estatística Aplicada, defendida e aprovada por unanimidade em 26/02/2015 pela Comissão Examinadora.

Orientador:

Prof. Dr. Cláudio Tadeu Cristino
Universidade Federal Rural de Pernambuco

Banca Examinadora:

Prof. Dr. Anderson Luiz da Rocha e Barbosa
Universidade Federal Rural de Pernambuco
DF/PPGFA/UFPRE

Prof. Dr. Tiago Alessandro Espínola Ferreira
Universidade Federal Rural de Pernambuco
DEINFO/PPGBEA/PPGIA/UFRPE

Prof. Dr. Wilson Rosa de Oliveira Júnior
Universidade Federal Rural de Pernambuco
DEINFO/PPGBEA/PPGIA/UFRPE

RESUMO

A probabilidade quântica (livre, ou Não Comutativa) traz um conceito generalizado de variáveis aleatórias com a permissão de que as álgebras sejam Não Comutativas, diferentemente da concepção usual de variáveis aleatórias. Há uma modificação na maneira de se conceber os Espaços de Probabilidades Não Comutativos, bem como na obtenção das suas medidas de distribuição. No contexto dos algoritmos quânticos, sabe-se que estes possuem uma melhora significativa se comparados aos algoritmos clássicos, e que o algoritmo de busca de Grover, em especial, possui uma melhora polinomial. Os algoritmos quânticos podem ser pensados como aplicações sucessivas de operadores lineares, de uma forma geral. Neste sentido, este trabalho busca uma associação entre as medidas de distribuição Não Comutativa, chamadas de $*$-distribuições, dos operadores utilizados no algoritmo de Grover, e a probabilidade usual de se encontrar um dado elemento de interesse numa lista não ordenada.

SUMÁRIO

CAPÍTULO 1

INTRODUÇÃO

Na probabilidade clássica ou axiomática, um modelo (ou espaço de probabilidade) é determinado dando-se um conjunto Ω de resultados w, especificando-se quais subconjuntos $S \in \Omega$ são considerados como *eventos*, e associando uma *probabilidade* $P(S)$ para cada um desses eventos.

Em probabilidade quântica, perde-se um pouco desse esquema. Deve-se abdicar da ideia de Ω ser uma amostra de pontos: um ponto $w \in \Omega$ num modelo clássico, decide sobre a ocorrência ou não de todos os eventos simultaneamente, e abandona-se essa ideia. Na verdade, toma-se como *eventos* certos *subespaços fechados* de um *espaço de Hilbert*, ou de forma equivalente, um conjunto de projeções. Para todas essas projeções associa-se probabilidades [Kuperberg, 2005]. A probabilidade quântica é frequentemente chamada de "probabilidade não comutativa". Isso ocorre porque, um sistema probabilístico clássico (ou espaço mensurável) é um álgebra de variáveis aleatórias que satisfaz axiomas relevantes. Uma das restrições da álgebra é a comutatividade: Se x e y são duas variáveis aleatórias de valor real ou complexo, então xy e yx são a mesma variável aleatória. Na probabilidade quântica, esta álgebra comutativa é substituída por uma álgebra não comutativa chamada de álgebra de Von Neumann. As outras definições permanecem, na medida do possível [Griffiths, 2005].

Alguns aspectos onde pode-se observar a teoria quântica de probabilidade são a informação quântica e a computação quântica. Na *informação quântica* surge uma nova unidade de informação chamada de *quibit*, substituindo a unidade clássica de informação, o bit. Na teoria clássica de informação, esta unidade serve para quantificar todas as formas de informação, pois podem ser convertidas umas nas outras através de cópias. Já em sistemas quânticos, as informações não podem ser copiadas, mas podem ser convertidas umas nas outras. O tratamento deste novo tipo de informação é feito a partir de seus portadores, chamados de 'canais' [Massen, 2004]. A utilização de algoritmos aleatorizados pode resultar em respostas mais rápidas do que certos algoritmos determinísticos para alguns problemas computacionais; e os algoritmos quânticos podem ser ainda mais

rápidos, do que suas alternativas clássicas ou aleatorizadas [Kuperberg, 2005], como os algoritmos de fatoração, por exemplo. A ideia da computação quântica vem surgindo desde os anos de 1970 com a criação do código conjugado por Stephen Wiesner, e vem se expandindo com contribuições de muitos outros cientistas como Richard Feynman, que em 1981, em uma de suas palestras, observou que parecia ser improvável, em geral, simular a evolução de um sistema quântico em um computador clássico de forma eficiente; David Deutsch, que em 1985 descreveu o primeiro computador quântico universal na Universidade de Oxford; Peter Shor, que em 1994 criou um importante algoritmo que permitia que um computador quântico fatorasse números inteiros de grande ordem rapidamente; entre tantos outros.

1.1 OBJETIVO

O objetivo deste trabalho foi o de estudar e descrever a probabilidade não comutativa (ou quântica, ou livre), fornecendo um embasamento matemático para o seu entendimento, bem como analisar, do ponto de vista probabilístico, o algoritmo quântico de busca de Grover.

1.2 ORGANIZAÇÃO DO TRABALHO

Neste capítulo introdutório, apresenta-se uma visão geral do trabalho, mostrando as teorias importantes para a construção do mesmo, bem como os objetivos a serem alcançados.

No capítulo 2, é feita uma introdução teórica de conceitos matemáticos importantes para o entendimento da probabilidade não comutativa. Tais conceitos envolvem a Teoria dos Conjuntos juntamente com suas operações; a noção de supremo e ínfimo de um conjunto numérico; o conceito de álgebra; a definição de espaços vetoriais e topológicos, com destaque para o Espaço de Hilbert, que é o espaço no qual a mecânica quântica atua.

No capítulo 3, traz-se uma introdução à mecânica quântica, que dará embasamento ao entendimento do algoritmo de busca quântico de Grover. É mostrada noções da álgebra linear como bases e independência linear, operadores lineares e matrizes, produtos internos, entre outros, utilizando a notação de Dirac (bra-ket). Também são mostrados os postulados que fundamentam a mecânica quântica.

No capítulo 4, os conceitos de Espaço de probabilidade e distribuições não comutativas ganham forma. A estrutura de um $*$-espaço de probabilidade é construída e exemplificada. E as $*$-distribuições para elementos normais são mostradas. Este capítulo é encerrado com os C^*-espaços de probabilidade , que fornecem um ambiente natural onde as ideias da probabilidade não comutativa podem ser trabalhadas.

No capítulo 5, descreve-se o algoritmo quântico de busca de Grover, de uma maneira simplificada, mostrando a ideia por trás do mesmo, além dos passos por ele executados. A partir disso, fornece-se uma descrição, via probabilidade não comutativa, dos operadores que atuam na aplicação do algoritmo.

Por fim, faz-se as considerações finais acerca do trabalho e comentários sobre trabalhos futuros.

CAPÍTULO 2

CONJUNTOS, ÁLGEBRAS E ESPAÇOS

Neste capítulo é feita uma introdução matemática de temas relevantes para a compreensão dos conceitos apresentados nos capítulos subsequentes.

2.1 TEORIA DOS CONJUNTOS

Esta Seção tem por objetivo expor alguns conceitos e noções importantes da Teoria dos Conjuntos, principalmente com o intuito de se fixar a notação (mesmo que seja a usual) e de servir como fonte de referências para este tópico.

2.1.1 Conjuntos e operações

A noção primitiva de conjunto é a de que este é constituído por uma coleção de objetos chamados *elementos*.

Dado um conjunto A, diz-se que x é um *elemento* de A e escreve-se esta relação entre elemento e conjunto como $x \in A$ (x pertence à A), ou $A \ni x$. Para negar esta afirmação, ou seja, dizer que um elemento não pertence a um conjunto, usa-se a notação $x \notin A$. Se A é um *subconjunto* de S, escreve-se esta relação entre conjuntos como $A \subset S$ (A está contido em S). Ou $S \supset A$ (S contém A). Para negar essa afirmação, pode-se escrever $A \not\subset S$.

Duas relações que permitem ordenar e igualar conjuntos são, para dois conjuntos A e B:

Relação de contenção: $A \subset B \Leftrightarrow x \in A \Rightarrow x \in B$.

Relação de igualdade: $A = B \Leftrightarrow A \subset B$ e $B \subset A$.

Dados dois conjuntos A e B, têm-se as seguintes operações elementares entre conjuntos:

União: A união de A e B, escrita como $A \cup B$, é o conjunto de elementos que pertencem à A, à B ou à ambos:

$$A \cup B = \{x : x \in A \text{ ou } x \in B\}.$$

Intersecção: A intersecção de A e B, escrita como $A \cap B$, é o conjunto de elementos que pertencem à ambos, A e B:

$$A \cap B = \{x : x \in A \text{ e } x \in B\}.$$

Complemento: o complementar de A, escrito como A^C, é o conjunto de todos os elementos que não estão em A:

$$A^C = \{x : x \notin A\}.$$

Diferença: A diferença entre A e B, é o conjunto formado por todos os elementos de A, exceto os que também estejam em B:

$$A - B \text{ ou } A \cap B^C = \{x : x \in A, x \notin B\}.$$

Diferença Simétrica: A diferença simétrica entre A e B é dada por todos os elementos de $A \cup B$, exceto os que também estejam em $A \cap B$. Podemos representar a diferença simétrica por: $A \triangle B$ ou $(A \cap B^C) \cup (A^C \cap B)$.

As operações elementares podem ser combinadas. Abaixo segue algumas propriedades úteis das operações de conjuntos.

Teorema 2.1.1. *Para quaisquer três conjuntos A, B e C, sendo estes subconjuntos de um conjunto S, têm-se:*

a.	***Comutatividade***	$A \cup B = B \cup A,$
		$A \cap B = B \cap A;$
b.	***Associatividade***	$A \cup (B \cup C) = (A \cup B) \cup C,$
		$A \cap (B \cap C) = (A \cap B) \cap C;$
c.	***Leis Distributivas***	$A \cap (B \cup C) = (A \cap B) \cup (A \cap C),$
		$A \cup (B \cap C) = (A \cup B) \cap (A \cup C);$
d.	***Leis de Morgan***	$(A \cup B)^C = A^C \cap B^C,$
		$(A \cap B)^C = A^C \cup B^C;.$

Para demonstração dos fatos acima ver [Casella and Berger, 2002] ou [Magalhães, 2006].

O *conjunto vazio*, simbolizado por $\emptyset$, é o conjunto que não contém nenhum elemento.

Considere qualquer conjunto arbitrário A. Visto que o conjunto vazio $\emptyset$ não contém elementos, é logicamente correto dizer que qualquer elemento pertencente à $\emptyset$, também pertence à A, ou $\emptyset \subset A$. Em outras palavras, para qualquer conjunto A, é verdade que $\emptyset \subset A$ [DeGroot, 1989].

Dois conjuntos são ditos *disjuntos* ou *mutuamente exclusivos* se sua intersecção é o conjunto vazio.

$$A \text{ e } B \text{ são disjuntos } \Leftrightarrow A \cap B = \emptyset.$$

Diz-se que $A_1, A_2, \ldots, A_n$ formam uma *partição* de um conjunto S, quando são disjuntos e sua união é S, ou seja:

$$\bigcup_{i=1}^{n} A_i = S, \text{ com } A_i \cap A_j = \emptyset, \forall i \neq j.$$

O *conjunto das partes* de S é formado por todos os subconjuntos de S. Esse conjunto é denotado por S_p.

Se um número infinito de subconjuntos estiver envolvido nas operações acima mencionadas, elas são definidas de maneira análoga [Magalhães, 2006].

2.1.2 Supremo e ínfimo de um conjunto numérico

Dado um subconjunto C, de números reais, diz-se que ele é *limitado à direita* ou *limitado superiormente* se existe um número K tal que $c \leq K$ para todo $c \in C$. De maneira análoga, C é *limitado à esquerda* ou *limitado inferiormente* se existe um número k tal que $k \leq c$ para todo $c \in C$. Os números K e k são chamados *cota superior* e *cota inferior*, respectivamente, do conjunto C. Como exemplos podemos citar:

- O conjunto dos números naturais é limitado inferiormente, mas não superiormente;
- O conjunto dos números racionais menores que 8 é limitado superiormente, mas não inferiormente;
- O conjunto dos números reais x tais que $x^2 \leq 10$ é limitado tanto à direita quanto à esquerda, pois $-\sqrt{10} \leq x \leq \sqrt{10}$.

Um conjunto, como o citado anteriormente, que é limitado à esquerda e à direita, é dito, simplesmente, *conjunto limitado*. É também limitado qualquer intervalo de extremos finitos a e b. [Ávila, 1999].

Quando um conjunto é limitado superiormente, ele pode ter um elemento que seja o maior de todos, chamado de *máximo* do conjunto. Por exemplo, o conjunto dos números racionais que x tais que $x \leq 10$ tem 10 como seu máximo.

Entretanto, o conjunto

$$A = \left\{ \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots, \frac{n}{n+1}, \ldots \right\}$$

não tem máximo, embora seja limitado superiormente. Os elementos desse conjunto, são frações dispostas de maneira crescente:

$$\frac{1}{2} < \frac{2}{3} < \frac{3}{4} < \ldots < \frac{n}{n+1} < \ldots$$

e nenhuma dessas frações é maior que todas as outras. Pelo contrário, qualquer delas é superada pela que vem logo a seguir, isto é,

$$\frac{n}{n+1} < \frac{n+1}{n+2}.$$

Além disso, qualquer elemento do conjunto é menor que 1, o qual é, portanto, uma de suas *cotas superiores*. Aliás, 1 é a menor dessas cotas, pois, dado qualquer número $c < 1$, é sempre possível encontrar n tal que $c < \frac{n}{(n+1)}$, o que quer dizer que c não é cota superior.

Este último exemplo, ilustra uma situação interessante: o conjunto é limitado superiormente, não tem máximo, mas tem uma cota superior que é a menor de todas. Isso sugere uma definição de *supremo* de um conjunto.

Definição 2.1.1. Chama-se supremo de um conjunto C à menor de suas cotas superiores. Ou seja, chama-se supremo de um conjunto numérico C ao número S que satisfaz as duas condições seguintes: $a)\, c \leq S$ para todo $c \in C$; $b)$ dado qualquer número $\varepsilon > 0$, existe um elemento $c \in C$ tal que $S - \varepsilon < c$.

Proposição 2.1.2 (Propriedade do supremo)**.** *Todo conjunto não vazio de números reais, que seja limitado superiormente, possui supremo.*

A noção de *ínfimo* é introduzida de maneira análoga à de supremo.

Definição 2.1.2. Chama-se ínfimo de um conjunto C à maior de suas cotas inferiores; ou ainda chama-se ínfimo de um conjunto C ao número s que satisfaz as duas condições seguintes: $a)\, s \leq c$ para todo $c \in C$; $b)$ dado qualquer número $\varepsilon > 0$, existe um elemento $c \in C$ tal que $c < s + \varepsilon$.

Com a propriedade do supremo prova-se que *todo conjunto não vazio de números reais, que seja limitado inferiormente possui ínfimo.*

Conjuntos não limitados à direita certamente não possuem supremos finitos. Convenciona-se considerar $+\infty$ como supremo desses conjuntos. Analogamente, $-\infty$ é considerado o ínfimo dos conjuntos não limitados inferiormente.

Observe que se os conjuntos dos números racionais forem considerados, então não será verdade que todo conjunto limitado superiormente tenha supremo ou que todo conjunto limitado inferiormente tenha ínfimo.

2.2 ÁLGEBRA DE CONJUNTOS

Definição 2.2.1. Seja Ω um conjunto e $\mathcal{F}$ uma família de subconjuntos de Ω. Diz-se que $\mathcal{F}$ é uma *álgebra de conjuntos* se satisfaz as seguintes propriedades:

i. $\Omega \in \mathcal{F}$;

ii. $A \in \mathcal{F} \Rightarrow A^C \in \mathcal{F}$;

iii. Se $A_1, A_2, \ldots, A_n \in \mathcal{F}$, então $\bigcup_{i=1}^{n} A_i \in \mathcal{F}$

Se, além dessas propriedades, for verificada a propriedade abaixo, a álgebra de conjuntos passa a ser chamada de *σ-álgebra de conjuntos*.

iii.a Se $A_1, A_2, \ldots \in \mathcal{F}$, então $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

Assim, qualquer σ-álgebra de conjuntos é também uma álgebra de conjuntos, mas o recíproco não é verdadeiro [Guerra, 2014].

Outro conceito importante aqui é o da *σ-álgebra (ou álgebra) de Borel*, que é a menor σ-álgebra gerada por um subconjunto $A \in \Omega$. A σ-álgebra de Borel em $\mathbb{R}$ ($\mathcal{B}(\mathbb{R})$), por exemplo, é a menor σ-álgebra que contém todos os intervalos abertos dos reais. É possível verificar que ela pode ser gerada pelos intervalos $(-\infty, x)$ com $x \in \mathbb{R}$. Existem outras escolhas para o intervalo gerador, mas o que é importante é que, qualquer tipo de intervalo dos reais pode ser obtido através de um número enumerável, finito ou infinito, de operações de uniões e intersecções com o intervalo acima [Magalhães, 2006].

2.3 MEDIDA

Seja Ω um conjunto fixo. Se $\mathcal{A}$ é a família de subconjuntos de Ω tal que:

i. $\emptyset \in \mathcal{A}$;

ii. Se $A \in \mathcal{A}$, então $A^C \in \mathcal{A}$;

iii. Se $A_1, A_2, \cdots, A_n \subset \Omega$ então $\bigcup_{i=1}^{n} A_i \in \mathcal{A}$.

Neste caso, $\mathcal{A}$ é chamado de *álgebra* de conjuntos sobre Ω.

Dada uma propriedade adicional

iii'. $A_1, A_2, \cdots, A_n \subset \Omega$, então $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$, então $\mathcal{A}$ é chamado de σ-álgebra sobre Ω.

Pode-se ainda, definir uma função sobre o par $(\Omega, \mathcal{A})$:

$$\begin{aligned} \mu :&\mathcal{A} \to \mathbb{R}^+ \\ &A \mapsto \mu(A) \end{aligned}$$

Tal que:

1. $\mu(\emptyset) = 0$.
2. Se $A, B \in \mathcal{A}$ e $A \subset B$ então $\mu(A) \leq \mu(B)$.
3. Se $A_1, A_2, \ldots \in \mathcal{A}$ e $A_i \cap A_j = \emptyset$ para $i \neq j$, então

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$$

4. No caso mais geral,onde de a interseção não é necessariamente disjunta, ou seja, podendo ocorrer $A_i \cap A_j \neq \emptyset$, tem-se que

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu(A_i).$$

Neste caso, μ é chamada *medida* sobre $(\Omega, \mathcal{A})$ e a tripla $(\Omega, \mathcal{A}, \mu)$ é um *espaço de medida*. Quando $\mu(\Omega) = 1$, este espaço passa a ser chamado de *espaço de probabilidade* [Leadbetter et al., 2014].

O conceito de medida $\mu(A)$ de um conjunto A é uma generalização natural de conceitos como:

- O comprimento $l(r)$ de um segmento de reta r.
- A área $S(P)$ de uma figura plana P.
- O volume $V(T)$ de uma figura T no espaço.
- O incremento $f(b) - f(a)$ de uma função não decrescente $f(t)$ em um intervalo aberto da reta $[a, b)$.

- A integral de uma função não negativa tomada sobre alguma linha, plano ou região no espaço.

O conceito de medida de um conjunto teve origem na teoria de funções de variáveis reais, e tem diversas aplicações em teoria da probabilidade, teoria de sistemas dinâmicos, análise funcional e diversos outros campos da matemática [Kolmogorov and Fomin, 1960].

2.4 ESPAÇOS VETORIAIS E TOPOLÓGICOS

2.4.1 Espaço Vetorial

Um *espaço vetorial* sobre um corpo $\mathbb{K}$ [1], é um conjunto não vazio V, cujos elementos são chamados de vetores, munidos de duas operações, chamadas de adição e produto por escalar. A adição $(+)$ associa a cada par (x, y) do conjunto $V \times V$ a um novo elemento em V, indicado por $x+y$. O produto por escalar $(\cdot)$, associa a cada par (λ, x) do conjunto $\mathbb{K} \times V$ a um novo elemento em V, indicado por $\lambda \cdot V$ [Barroso, 2014].

Tais operações satisfazem as seguintes propriedades, para quaisquer $x, y \in V$ e $\lambda, \mu \in \mathbb{K}$.

- **Adição**:

  (a_1) (Comutatividade): $x + y = y + x$;

  (a_2) (Associatividade): $x + (y + z) = (x + y) + z$;

  (a_3) (Elemento neutro): Existe um elemento $e \in V$ tal que $x + e = x\,, \quad \forall x \in V$;

  (a_4) (Elemento inverso): Para cada $x \in V$ existe $\widehat{x} \in V$ tal que $x + \widehat{x} = e$.

  Para quaisquer $x\,, y\,, z \in V$.

[1] Dizer que um conjunto numérico é um corpo significa que estão definidas, neste conjunto, duas operações: *adição* e *multiplicação* que cumprem certas condições. A adição faz corresponder a cada par de elementos, por exemplo, $x\,, y \in \mathbb{R}$ (ou $\mathbb{C}$), sua *soma* $x + y \in \mathbb{R}$ (ou $\mathbb{C}$), enquanto a *multiplicação* associa a esses elementos o seu produto $x \cdot y \in \mathbb{R}$ (ou $\mathbb{C}$).
Os axiomas que essas operações obedecem são: *Associatividade*, *Comutatividade*, *Existência de elementos neutros* (tanto na adição quanto na multiplicação), *Existência de um Inverso Aditivo e um Inverso Multiplicativo* e a *Distributividade*. Para mais detalhes, consultar [Lima, 2006].

- **Produto por escalar**:

(p_1) (Distributividade): $\lambda \cdot (x+y) = \lambda \cdot x + \lambda \cdot y \quad$ e $\quad (\lambda+\mu)\cdot x = \lambda \cdot x + \mu \cdot x$;

(p_2) (Associatividade): $\lambda \cdot (\mu \cdot x) = (\lambda\mu) \cdot x$;

(p_3) (Elemento neutro): Para cada $x \in V$, tem-se $1 \cdot x = x,$ onde $1 \in \mathbb{K}$.

2.4.2 Espaço Topológico

Definição 2.4.1. Um conjunto X com uma família T de subconjuntos de X é chamado um *espaço topológico* se satisfaz as condições:

(i) $\emptyset$ e X pertencem a T;

(ii) A união de qualquer subfamília de T pertence à T;

(iii) A intersecção de qualquer subfamília finita de T está em T.

A família T é chamada uma *topologia* em X e os elementos de T são chamados *conjuntos abertos* de X nessa topologia.

Exemplo 2.4.1. Exemplos de topologia em X são:

a) Sendo $X \neq \emptyset$, $T = \{X\,, \emptyset\}$ é uma topologia de X (chamada de topologia indiscreta).

b) Sendo $X \neq \emptyset$ e $T = \mathfrak{p}(X)$ = conjunto das partes de X, T é uma topologia em X chamada *topologia discreta* de X. [Nowosad, 1967]

2.4.3 Espaço Métrico

Definição 2.4.2. (**Distâncias**)
Seja X um conjunto. Uma *distância sobre* X é um a aplicação d de $X \times X$ no conjunto dos números reais $\mathbb{R}$, obedecendo as seguintes propriedades:

(i) $d(x,y) \geq 0$ para todo $x, y \in X$, sendo que $d(x,y) = 0$ se, e somente se, $x = y$;

(ii) $d(x,y) = d(y,x)$ para $x, y \in X$;

(iii) $d(x,y) \leq d(x,y) + d(z,y)$ para quaisquer $x, y, z \in X$. (Desigualdade Triangular.)

Espaço Métrico é o conjunto X juntamente com uma distância sobre X. [Nowosad, 1967]

Exemplo 2.4.2. A função $d(x,y) = |x - y|$, no conjunto dos números reais, satisfaz às condições impostas acima. O conjunto dos números reais, juntamente com essa métrica chama-se *reta real* $\mathbb{R}$.

Exemplo 2.4.3. Qualquer uma das funções abaixo é uma distância no plano euclidiano $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$:

$$d_1(x,y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$$

$$d_2(x,y) = |x_1 - x_2| + |y_1 - y_2|$$

$$d_3(x,y) = max\{|x_1 - x_2|, |y_1 - y_2|\}$$

com $x = (x_1, x_2)$ e $y = (y_1, y_2)$.

Proposição 2.4.4. *Em* $\mathbb{K}^n$ *toda norma ou distância são equivalente, ou seja,* $\alpha d_2 \leq d_1 \leq \beta d_3$.

Exemplo 2.4.5. Seja $C[a,b]$ o conjunto das funções complexas contínuas, definidas no intervalo finito $[a,b]$, munido da distância

$$d(x,y) = \max_{a \leq t \leq b} |x(t) - y(t)|.$$

$C[a,b]$ é um espaço métrico.

Definição 2.4.3. Sequência de Cauchy
Uma sequência de elementos x_n de um espaço métrico com métrica $d(x\,,y)$ é chamada *sequência de Cauchy* se para todo $\varepsilon > 0$ existe um n_0 tal que para todo $k\,, m \geq n_0 \rightarrow d(x_k\,, x_m) < \varepsilon$. [Bierens, 2014]

A noção sobre sequência de Cauchy tem papel crucial na definição de *espaços de Hilbert*.

Teorema 2.4.6. *Toda sequência de Cauchy em* $\mathbb{R}^d$ *ou em* $\mathbb{C}^d$ *com* $d < \infty$*, tem um limite no espaço envolvido.*

Prova. Ver [Bierens, 2014].

2.4.4 Espaço com produto interno

Definição 2.4.4. Espaço com produto interno: Tome V como um espaço vetorial sobre $\mathbb{C}$. Um produto interno em V significa um mapa[2] $f : V \times V \to \mathbb{C}$, descrito por $(x\,, y) \mapsto \langle x|\, y\rangle$, tal que, para todo $x\,, x'\,, y \in V$ e todo $\alpha \in \mathbb{C}$, as seguintes identidades são satisfeitas:

(1) $\langle x + x'|\, y\rangle = \langle x|\, y\rangle + \langle x'|\, y\rangle$;

(2) $\langle \alpha x|\, y\rangle = \alpha\langle x|\, y\rangle$;

(3) $\overline{\langle x|\, y\rangle} = \langle y|\, x\rangle$, então, em particular, $\langle x\,|x\rangle \in \mathbb{R}$;

(4) $\langle x|\, x\rangle \geq 0$, com a igualdade ocorrendo, se, e somente se, $x = 0_V$ (vetor nulo).

Por um *espaço complexo com produto interno* entende-se um espaço vetorial V sobre $\mathbb{C}$ junto com o produto interno em V. Por um *espaço real com produto interno* entende-se um espaço vetorial V sobre $\mathbb{R}$ junto com o produto interno em V [Blyth and Robertson, 2006].

Há outras identidades úteis que seguem imediatamente às vistas acima:

(5) $\langle x\,|y + y'\rangle = \langle x|y\rangle + \langle x|y'\rangle$;

(6) $\langle x\,|\alpha y\rangle = \overline{\alpha}\langle x|y\rangle$;

(7) $\langle x\,|0_V\rangle = 0 = \langle 0_V\,|x\rangle$

Exemplo 2.4.7. No espaço vetorial $\mathbb{R}^n$ de n-uplas de números reais, tome

$$\langle (x_1, \ldots, x_n)\,|\,(y_1, \ldots, y_n)\rangle = \sum_{i=1}^{n} x_i y_i.$$

Então, verifica-se que isto define um produto interno em $\mathbb{R}^n$, chamado de *produto interno padrão em* $\mathbb{R}^n$.

Nos casos onde $n = 2\,, 3$, este produto interno é frenquentemente chamado de *produto escalar*. Esta terminologia é popular quando trata-se de aplicações geométricas de vetores [Blyth and Robertson, 2006].

[2]Refere-se à uma função ou relação matemática que trata de domínios e/ou contradomínios não numéricos. Também pode ser chamado de aplicação matemática ou transformação.

Exemplo 2.4.8. No espaço vetorial $\mathbb{C}^n$ de n-uplas de números complexos tome:

$$\langle (z_1, \ldots, z_n) \,|\, (w_1, \ldots, w_n) \rangle = \sum_{i=1}^{n} z_i \overline{w_i}.$$

Então, verifica-se que isto define um produto interno em $\mathbb{C}$, chamado de *produto interno padrão em* $\mathbb{C}$.

Exemplo 2.4.9. Tome $a\,, b \in \mathbb{R}$ com $a < b$ e tome V como o espaço vetorial real de funções contínuas $f : [a\,, b] \to \mathbb{R}$. Defini-se um mapa de $V \times V$ em $\mathbb{R}$ por

$$(f\,, g) \mapsto \langle f \,|g\rangle = \int_a^b fg.$$

Então, pelas propriedades elementares da integral, isto define um produto interno em V.

2.4.5 Espaço vetorial normado

Definição 2.4.5. Uma *norma* em um espaço vetorial V é uma aplicação $||\cdot|| : V \to \mathbb{R}$, que satisfaz as seguintes condições:

(i) $||x|| \geq 0$ e $||x|| = 0$ se, e somente se, $x = 0$;

(ii) $||\lambda x|| = |\lambda| \cdot ||x||$;

(iii) $||x + y|| \leq ||x|| + ||y||$.

Daí resulta que a função definida como $d(x, y) = ||x - y||$ é uma distância.

Um espaço vetorial também considerado como espaço métrico, com a métrica induzida por esta norma, é dito *espaço vetorial normado* [Nowosad, 1967].

Exemplo 2.4.10. Nos exemplos que seguem serão considerados espaços constituídos por funções reais ou complexas definidas em um certo conjunto T. Estes espaços se tornam vetoriais quando a soma e o produto por escalar são definidos por

$$(x + y)(t) = x(t) + y(t), \quad (\lambda x)(t) = \lambda x(t), \quad t \in T.$$

Em particular, se $T = \{1, 2, 3, \ldots, n\}$ obtêm-se o espaço $V_n(\mathbb{R})$ ou $V_n(\mathbb{C})$, das n-uplas reais ou complexas com produto escalar definido pelas operações correspondentes sobre as componentes.

Exemplo 2.4.11. Pode-se introduzir várias normas em $V_n(\mathbb{R})$ ou $V_n(\mathbb{C})$. Para $p \geq 1$ e $x = (\xi_1, \xi_2, \ldots, \xi_n) \in V_n(\mathbb{C})$, fazendo

$$||x|| = (|\xi_1|^p + |\xi_2|^p + \ldots + |\xi_n|^p)^{1/p} \tag{2.1}$$

é imediato que os axiomas $i)$ e $ii)$ da definição de norma são satisfeitos, quanto à desigualdade $iii)$ sua expressão em termos de (2.1) é:

$$\left(\sum_{i=1}^{n} |\xi_i + \eta_i|^p\right)^{1/p} \leq \left(\sum_{i=1}^{n} |\xi_i|^p\right)^{1/p} + \left(\sum_{i=1}^{n} |\eta_i|^p\right)^{1/p} \tag{2.2}$$

que se chama *desigualdade de Minkowski.*

Para $p = 1$ a validade dessa igualdade é imediata. Para $p > 1$, ela é provada a seguir.

A desigualdade de Hölder é dada por:

$$\sum_{i=1}^{n} |\xi_i \cdot \eta_i| \leq \left(\sum_{i=1}^{n} |\xi_i|^p\right)^{1/p} \cdot \left(\sum_{i=1}^{n} |\eta_i|^q\right)^{1/q} \tag{2.3}$$

onde q é definido por $q = \dfrac{p}{p-1}$ e portanto satisfaz:

$$\frac{1}{p} + \frac{1}{q} = 1 \tag{2.4}$$

Como a (2.3) é homogênea, isto é, contínua e válida ao substituirmos x por λx e y por λy, onde λ é um escalar, basta prová-la no caso em que:

$$\sum_{i=1}^{n} |\xi_i|^p = \sum_{i=1}^{n} |\eta_i|^q = 1 \tag{2.5}$$

Daí, o segundo termo da (2.3) é igual a 1. Observando-se que se $\eta = f(\xi)$, com $\eta, \xi \in \mathbb{R} \geq 0$, for uma função contínua monótona não decrescente tal que, $f(0) = 0$ e $f(\xi) \to \infty$ quando $\xi \to \infty$, então dados dois números positivos a e b quaisquer, a soma das áreas A, B indicadas na figura 2.1 , é maior que ab.

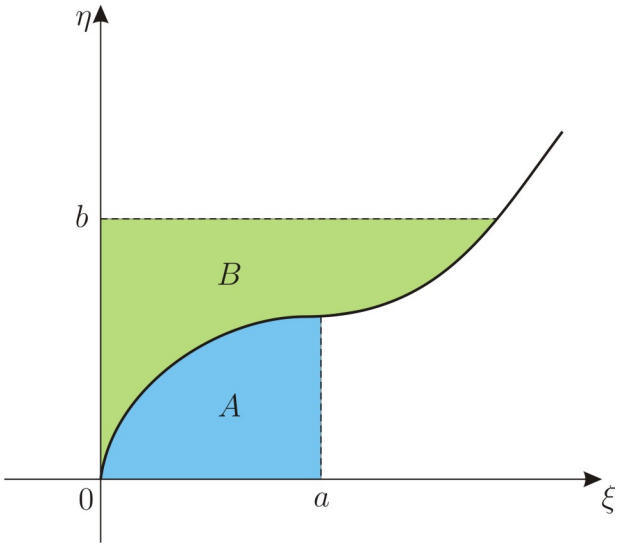

Figura 2.1 Gráfico ilustrativo da função $\eta = f(\xi)$.

No caso em que $\eta = f(\xi) = \xi^{p-1}$ e aplicando-se este resultado:

$$\xi = \eta^{\frac{1}{p-1}} = \eta^{q-1}$$

Daí, obtém-se:

$$A = \int_0^a \xi^{p-1} d\xi = \frac{a^p}{p}$$

$$B = \int_0^a \eta^{q-1} d\eta = \frac{b^q}{q}$$

Daí, $ab \leq \dfrac{a^p}{p} + \dfrac{b^q}{q}$. Fazendo $a = |\xi_i|$, $b = |\eta_i|$ e somando sobre i de 1 a n tem-se:

$$\sum_{i=1}^{n} |\xi_i \eta_i| \leq 1,$$

levando em conta (2.5) e (2.4), prova-se a afirmação.

Para provar a desigualdade de Minkowisk substitui-se $a = |\xi_i|$, $b = |\eta_i|$ na identidade:

$$(a+b)^p = (a+b)^{p-1} \cdot a + (a+b)^{p-1} \cdot b;$$

válida para $a\,, b \geq 0$, e somando-se sobre i de 1 a n obtém-se:

$$\sum_{i=1}^{n} (|\xi_i| + |\eta_i|)^p = \sum_{i=1}^{n} (|\xi_i| + |\eta_i|)^{p-1} \cdot |\xi_i| + \sum_{i=1}^{n} (|\xi_i| + |\eta_i|)^{p-1} \cdot |\eta_i|.$$

Aplicando-se a desigualdade de Hölder a cada um dos somatórios do segundo membro dessa igualdade, separadamente, e levando em conta que $(p-1)q = p$ obtém-se:

$$\sum_{i=1}^{n}(|\xi_i| + |\eta_i|)^p \leq \left[\sum_{i=1}^{n}(|\xi_i| + |\eta_i|)^p\right]^{1/q} \cdot \left[(\sum_{i=1}^{n}|\xi_i|^p)^{1/p} + (\sum_{i=1}^{n}|\eta_i|^p)^{1/p}\right].$$

Dividindo os dois membros dessa desigualdade pelo primeiro fator do segundo obtém-se a desigualdade de Minkowski (2.2).

Observação 2.4.12. (**Espaço** ℓ^p) O espaço vetorial obtido dessa maneira é denotado por $\ell^p(n)$. Ao espaço vetorial normado obtido com essa norma definida sobre $V_n(\mathbb{R})$ chama-se $\ell^p(n)$ sobre os reais.

Uma outra norma dada em $V_n(\mathbb{C})$ é dada por

$$||x|| = \max\{|\xi_1|, \ldots, |\xi_n|\}.$$

Neste caso, denota-se o correspondente espaço vetorial normado por $\ell^\infty(n)$, que é motivado pelo fato de que

$$\max_{1\leq i\leq n}|\xi_i| = \lim_{p\to\infty}(|\xi_1|^p, \ldots, |\xi_n|^p)^{1/p}.$$

Definição 2.4.6. Define-se o espaço ℓ^p, $p \geq 1$, como o espaço vetorial das sequências $x = \{\xi\}_{i=1}^{\infty}$ para as quais vale $\sum_{i=1}^{\infty}|\xi_i|^p < \infty$, sendo a norma definida por

$$||x_p|| = \left(\sum_{i=1}^{\infty}|\xi_i|^p\right)^{1/p}.$$

O espaço ℓ^∞ é definido como espaço vetorial das sequências $x = \{\xi\}_{i=1}^{\infty}$ limitadas, com a norma

$$||x||_\infty = \sup_i |\xi_i|$$

Observação 2.4.13. No caso em que $p = 2$, o espaço ℓ^2, que é a generalização dos espaços $\ell^2(n) =$ espaço unitário de dimensão n, chama-se *espaço de Hilbert* das sequências quadrado somáveis.

Neste caso, q também é igual a 2, e as desigualdades de Hölder e de Minkowisk tomam, respectivamente, as formas

$$\sum_i |\xi_i \eta_i| \leq \left(\sum_i |\xi_i|^2 \cdot \sum_i |\eta_i|^2 \right)^{1/2}$$

$$\left(\sum_i |\xi_i + \eta_i|^2 \right)^{1/2} \leq \left(\sum_i |\xi_i|^2 \right)^{1/2} + \left(\sum_i |\eta_i|^2 \right)^{1/2}$$

A primeira é a *desigualdade de Cauchy-Schwarz*, e a segunda simplesmente expressa o fato de que o comprimento do lado de um triângulo é menor ou igual do que a soma dos comprimentos dos outros dois.

Definição 2.4.7. Desigualdade de Cauchy-Schwarz: Tome V como sendo um *espaço com produto interno* e $x\,, y \in X$. Então

$$|\langle x|y\rangle| \leq ||x|| \cdot ||y||;$$

a igualdade ocorre, se, e somente se, x e y são linearmente dependentes [3].

Exemplo 2.4.14. Seja $p \geq 1$; no intervalo finito $[a, b]$ considera-se o espaço vetorial de todas as funções complexas contínuas, e define-se sua norma por

$$||x|| = \left(\int_a^b |x(t)|^p dt \right)^{1/p}.$$

Pode-se provar que esta é de fato uma norma, para isso usa-se a desigualdade de Minkowski para integrais

$$\left(\int_a^b |x + y|^p dt \right)^{1/p} \leq \left(\int_a^b |x(t)|^p dt \right)^{1/p} + \left(\int_a^b |y(t)|^p dt \right)^{1/p}$$

[3] Dado um espaço vetorial V e um conjunto de vetores do mesmo $A = v_1, \ldots, v_n \in V$, e considerando-se a equação:

$$a_1 v_1 + \ldots + a_n v_n = 0$$

Sabe-se que esta equação admite, pelo menos uma solução $a_1 = a_2 = \ldots = a_n = 0$, chamada solução trivial. O conjunto A diz-se *linearmente dependente* (LI), ou os vetores $v_1, \ldots, v_n$ são LI, caso a equação acima admita apenas a solução trivial. Se existirem soluções $a_i \neq 0$, diz-se que o conjunto A é *linearmente dependente* LD, ou que os vetores $v_1, \ldots, v_n$ são LD.

cuja prova se obtém substituindo-se o símbolo de somatório pelo o de integral na prova do exemplo 2.4.11 . Este espaço vetorial normado é denotado por $L^p[a, b]$.

2.4.6 Funcional Linear

Definição 2.4.8. (Funcional Linear): Um *funcional linear* é definido como toda função em que o domínio é um espaço vetorial e a imagem é o seu corpo de escalares.

De maneira formal: Um *funcional linear* num espaço vetorial V é uma função: $f \to \mathbb{R}(\text{ou}\mathbb{C})$ que satisfaz as propriedades seguintes:

(i) $f(x + y) = f(x) + f(y)$;

(ii) $f(\lambda x) = \lambda f(x)$,

para quaisquer vetores $x\,, y \in V$ e um escalar $\lambda \in \mathbb{R}(\text{ou } \mathbb{C})$.

2.4.7 Espaço de Hilbert

O espaço Euclidiano $\mathbb{R}^n$ é um espaço vetorial dotado de um produto interno $\langle x|y\rangle = x^T y$, de uma norma $||x|| = \sqrt{x^T x} = \sqrt{\langle x|x\rangle}$ e uma métrica associada $||x - y||$, tal que toda sequência de Cauchy toma um limite em $\mathbb{R}^n$. Isto torna $\mathbb{R}^n$ um *espaço de Hilbert* [Bierens, 2014].

Definição 2.4.9. Espaço de Hilbert
Um *espaço de Hilbert* $\mathcal{H}$ é um espaço vetorial com produto interno, tal que toda sequência de Cauchy em $\mathcal{H}$ tenha um limite em $\mathcal{H}$.

Um espaço de Hilbert também é um *espaço de Banach*.

Definição 2.4.10. Espaço de Banach
Um espaço de Banach $\mathcal{B}$ é um espaço normado com uma métrica associada $d(x\,, y) = ||x - y||$ tal que toda sequência de Cauchy em $\mathcal{B}$ tem um limite em $\mathcal{B}$.

Se $\mathcal{H}$ é um espaço de Hilbert e $\langle\cdot|\cdot\rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{K}$, então define-se (naturalmente), $||\cdot|| : \mathcal{H} \to \mathbb{R}$ como:

$$||x|| := (\langle x|x\rangle)^{1/2}, \quad \forall x \in \mathcal{H}$$

e, daí,

$$d(x,\, y) = ||x - y||, x, y \in \mathcal{H}$$

como a distância entre os vetores x e y.

A diferença entre o espaço de Banach e o espaço de Hilbert é a forma como a norma é obtida. No caso do espaço de Hilbert a norma é definida via produto interno, $||x|| = \sqrt{\langle x|x\rangle}$, enquanto que no caso do espaço de Banach a norma é definida diretamente pela métrica $d(x, y) = ||x - y||$. Então, o espaço de Hilbert é um espaço de Banach, mas a recíproca não é verdadeira, pois em alguns casos a norma não pode ser associada à um produto interno.

Como exemplos de espaços de Hilbert, pode-se citar: $\mathbb{R}^n, C^n, \mathcal{B}(\mathbb{C}^n), l^2(\mathbb{C}), L^2(a, b)$, etc.

O espaço $L^2(a, b)$ é a coleção de funções quadrado integráveis e Borel mensuráveis de valores complexos ou reais f em $(a,\, b)$, isto é

$$\int_a^b |f(t)| < \infty,$$

dotado de um produto interno $\langle f|g\rangle = \int f(t)\overline{g(t)}dt$, e normas e métricas associadas

$$||f|| = \sqrt{\int_a^b |f(t)|^2 dt}$$

$$d(f, g) = ||f - g|| = \sqrt{\int_a^b |f(t) - g(t)|^2 dt},$$

respectivamente, onde as integrais envolvidas são as integrais de Lebesgue.

Teorema 2.4.15. *O espaço $L^2(\mu)$ de funções reais Borel mensuráveis em $\mathbb{R}$, satisfazendo $\int f(x)d\mu(x) < \infty$, dotado com produto interno $\langle f|g\rangle = \int f(x)g(x)d\mu(x)$, norma associada $||f|| = \sqrt{\langle f|f\rangle}$ e métrica $||f - g||$, é um espaço de Hilbert.*

O mesmo acontece para os outros espaços citados acima [Bierens, 2014].

Neste capítulo, foi feita uma breve descrição dos elementos matemáticos que serão largamente utilizados no decorrer do presente trabalho. Dentre esses elementos, dá-se destaque ao espaço de Hilbert, que é o ambiente natural para o desenvolvimento da probabilidade quântica (ou não comutativa); teoria que será vista no capítulo 4.

CAPÍTULO 3

INTRODUÇÃO À MECÂNICA QUÂNTICA

Alguns conceitos oriundos da mecânica quântica, como sua estrutura matemática e seus postulados, constituem um embasamento importante para se entender a computação e a informação quântica. Sendo assim, esse capítulo traz, de forma resumida, algumas ferramentas necessárias para a construção do conhecimento dessas áreas. O entendimento dessas ferramentas ajudará a consolidar as noções básicas da mecânica quântica elementar. Este capítulo é inteiramente baseado em [Nielsen and Chuang, 2010].

3.1 NOÇÕES DE ÁLGEBRA LINEAR

A *álgebra linear* é o estudo de espaços vetoriais e de operações lineares nestes espaços. Nesta seção, alguns conceitos básicos da álgebra linear serão revisados, e algumas notações básicas que serão usadas no estudo da mecânica quântica serão descritas.

Os objetos básicos da álgebra linear são os *espaços vetoriais*. O espaço vetorial de maior interesse para nós é o $\mathbb{C}^n$, o espaço de todas n-uplas de números complexos, $(z_1, \ldots, z_n)$. Os elementos de um espaço vetorial são chamados de *vetores*, e às vezes, uma matriz coluna é usada para representá-lo

$$\begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}.$$

Há uma operação de *adição* definida, que transforma um par de vetores em outro vetor. Em $\mathbb{C}^n$ a adição é definida como:

$$\begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} + \begin{bmatrix} z'_1 \\ \vdots \\ z'_n \end{bmatrix} = \begin{bmatrix} z_1 + z'_1 \\ \vdots \\ z_n + z'_n \end{bmatrix},$$

onde as operações de adição do lado direito são apenas adições ordinárias de números complexos.

Além disso, em um espaço vetorial há uma operação de *multiplicação por escalar*. Em $\mathbb{C}^n$ essa operação é definida como:

$$z \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} = \begin{bmatrix} zz_1 \\ \vdots \\ zz_n \end{bmatrix},$$

onde z é um *escalar*, isto é, um número complexo, e as multiplicações do lado direito são multiplicações usuais de números complexos.

Na mecânica quântica, há uma notação padrão para se representar um vetor em um espaço vetorial:

$$|\psi\rangle.$$

A notação $|\cdot\rangle$ é utilizada para indicar que um objeto é um vetor. Todo o objeto $|\psi\rangle$ é chamado de *ket*.

Um espaço vetorial também contém um vetor especial, o *vetor zero*, denotado por 0. Ele satisfaz a propriedade de que para qualquer outro vetor $|v\rangle$, $|v\rangle + 0 = |v\rangle$. Note que não se usa a notação *ket* para o vetor zero, pois $|0\rangle$ representa o vetor (coluna)

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

A multiplicação por escalar é tal que $z0 = 0$ para qualquer número complexo z. Por conveniência usa-se a notação $(z_1, z_2, \ldots, z_n)$ para representar uma matriz coluna com entradas $z_1, z_2, \ldots, z_n$. Em $\mathbb{C}^n$ o elemento zero é $(0, 0, \ldots, 0)$.

Um *subespaço vetorial* de um espaço vetorial V é um subconjunto W de V tal que W também é um espaço vetorial, isto é, W precisa ser fechado quanto às operações de adição e multiplicação por escalar.

Tabela 3.1 Resumo de algumas notações básicas em mecânica quântica para noções de álgebra linear. Este tipo de notação é conhecido como notação de *Dirac* [Nielsen and Chuang, 2010].

Notação	Descrição
z^*	Complexo conjugado do número complexo z.
	$(1+i)^* = 1-i$
$\|\psi\rangle$	Vetor. Também conhecido como *ket*.
$\langle\psi\|$	Vetor dual de $\|\psi\rangle$. Também conhecido como *bra*.
$\langle\varphi\|\psi\rangle$	Produto interno entre os vetores $\|\varphi\rangle$ e $\|\psi\rangle$.
$\|\varphi\rangle \otimes \|\psi\rangle$	Produto tensorial de $\|\varphi\rangle$ e $\|\psi\rangle$.
$\|\varphi\rangle\|\psi\rangle$	Notação abreviada para o produto tensorial de $\|\varphi\rangle$ e $\|\psi\rangle$.
A^*	Complexo conjugado da matriz A.
A^T	Transposta da matriz A.
$A^\dagger$	Conjugada Hermitiana ou adjunta da matriz A. $A^\dagger = (A^T)^*$.
	$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^\dagger = \begin{bmatrix} a^* & c^* \\ b^* & d^* \end{bmatrix}$
$\langle\varphi\|A\|\psi\rangle$	Produto interno entre $\|\varphi\rangle$ e $A\|\psi\rangle$.
	Ou, de forma equivalente, produto interno entre $A^\dagger\|\varphi\rangle$ e $\|\psi\rangle$.

3.1.1 Bases e independência linear

Um *conjunto gerador* para um espaço vetorial é um conjunto de vetores $|v_1\rangle, \ldots, |v_n\rangle$ tais que qualquer vetor $|v\rangle$ neste espaço vetorial possa ser escrito como combinação linear $|v\rangle = \sum_i a_i|v_i\rangle$ de vetores deste conjunto. Por exemplo, um conjunto gerador para o $\mathbb{C}^2$ é o conjunto:

$$|v_1\rangle \equiv \begin{bmatrix} 1 \\ 0 \end{bmatrix}; \;\; |v_2\rangle \equiv \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

visto que qualquer vetor

$$|v\rangle = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

em $\mathbb{C}^2$ pode ser escrito como combinação linear $|v\rangle = a_1|v_1\rangle + a_2|v_2\rangle$ dos vetores $|v_1\rangle$ e $|v_2\rangle$. Diz-se que os vetores $|v_1\rangle$ e $|v_2\rangle$ geram o $\mathbb{C}^2$. De forma geral, um espaço vetorial

pode ter muitos conjuntos geradores diferentes.

Um conjunto de vetores não-nulos $|v_1\rangle, \ldots, |v_n\rangle$ são *linearmente dependentes* se existe um conjunto de números complexos $a_1, \ldots, a_n$ com $a_i \neq 0$ para, pelo menos, um valor de i, tal que

$$a_1|v_1\rangle + a_2|v_2\rangle + \ldots + a_n|v_n\rangle = 0.$$

Um conjunto de vetores é *linearmente independentes* se não é linearmente dependente. Pode ser mostrado que quaisquer dois conjuntos de vetores linearmente independentes que geram um espaço vetorial V contém o mesmo número de elementos. A este conjunto damos o nome de *base* para V. O número de elementos da base é definido como *dimensão* de V.

3.1.2 Operadores lineares e matrizes

Um *operador linear* entre os espaços vetoriais V e W é definido como sendo qualquer função $A : V \to W$ que é linear em suas entradas,

$$A\left(\sum_i a_i|v_i\rangle\right) = \sum_i a_i A(|v_i\rangle).$$

Quando se diz que um operador linear é definido *em* um espaço vetorial, V, significa que A é um operador linear de V em V. Um operador linear importante em qualquer espaço vetorial V é o *operador identidade*, I_V, definido pela equação $I_V|v\rangle \equiv |v\rangle$. Outro importante operador linear é o *operador zero*, denotado por 0. O operador zero mapeia todos os vetores ao vetor zero, $0|v\rangle \equiv 0$.

Suponha que V, W e X sejam espaços vetoriais, e $A : V \to W$ e $B : W \to X$ são operadores lineares. A notação BA denota a *composição* de B com A, definida por $(BA)(|v\rangle) \equiv B(A|v\rangle)$. A maneira mais conveniente de entender operadores lineares é a partir da *representação matricial*. Operadores lineares e representações matriciais são completamente equivalentes.

Tome $A : V \to W$ como um operador linear entre os espaços vetoriais V e W. Suponha que $|v_1\rangle, \ldots, |v_m\rangle$ seja uma base para V e $|w_1\rangle, \ldots, |w_n\rangle$ seja uma base para W. Então, para cada j em $1, \ldots, m$ existem números complexos de A_{1j} à A_{nj} tal que

$$A|v_j\rangle = \sum_i A_{ij}|w_i\rangle.$$

A matriz cujas entradas são os valores A_{ij} forma a *representação matricial* do operador A.

3.1.3 As matrizes de Pauli

As *matrizes de Pauli* são quatro matrizes extremamente úteis no estudo da computação e informação quântica, e são usadas com frequência. Essa importância se dá pelo fato de elas formarem uma base dos operadores lógicos AND, NOT, OR, etc. São matrizes 2 por 2 com uma variedade de notações.

$$\sigma_0 \equiv I \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \sigma_1 \equiv \sigma_x \equiv X \equiv \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\sigma_2 \equiv \sigma_y \equiv Y \equiv \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \qquad \sigma_3 \equiv \sigma_z \equiv Z \equiv \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Note que as matrizes de Pauli representam transformações unitárias (aquelas que conservam a norma dos vetores) sobre espaços 2×2. Cabe a observação das transformações sobre os estados-base $|0\rangle$ e $|1\rangle$:

$$X|0\rangle = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = |1\rangle, \qquad X|1\rangle = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = |0\rangle.$$

$$Y|0\rangle = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ i \end{bmatrix} = i \begin{bmatrix} 0 \\ 1 \end{bmatrix} = i|1\rangle,$$

$$Y|1\rangle = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -i \\ 0 \end{bmatrix} = -i \begin{bmatrix} 1 \\ 0 \end{bmatrix} = -i|0\rangle.$$

$$Z|0\rangle = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = |0\rangle, \qquad Z|1\rangle = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix} = -|1\rangle.$$

3.1.4 Produto Interno

Na Seção 2.4.4 foi definido o espaço com produto interno. Na mecânica quântica, o produto interno entre os vetores $|v\rangle$ e $|w\rangle$ é representado por $\langle v|w\rangle$. A notação $\langle v|$ é usada para o vetor dual do vetor $|v\rangle$; o dual é um funcional linear definido por $|v\rangle(|w\rangle) \equiv \langle v|w\rangle$. O vetor dual pode ser representado matricialmente como um vetor linha.

3.1.4.1 Autovalores e autovetores. Um *autovetor* de um operador linear A num espaço vetorial, é, em geral, um vetor não-nulo $|v\rangle$ tal que $A|v\rangle = v|v\rangle$, onde v é um número complexo conhecido como *autovalor* de A correspondente a $|v\rangle$. Para calcular os autovalores e autovetores, necessita-se da *função característica*[1]. A *função característica* é definida como sendo $c(\lambda) \equiv \det|A - \lambda I|$. As raízes da *função característica* $c(\lambda) = 0$ são os autovalores do operador A. Pelo Teorema fundamental da álgebra, todo polinômio tem, pelo menos, uma raiz complexa, então todo operador A tem, ao menos, um autovalor, e um autovetor correspondente. O *autoespaço* correspondente a um autovalor v é o conjunto de vetores que têm autovalor v. É um subespaço vetorial de um espaço vetorial onde A atua.

Uma representação diagonal para um operador A em um espaço vetorial V é uma representação $A = \sum_i \lambda_i |i\rangle\langle i|$, onde os vetores $|i\rangle$ formam um conjunto ortonormal de autovetores para A, com autovalores correspondentes λ_i. Um operador é dito ser *diagonalizável* se tem uma representação diagonal. Como um exemplo de representação diagonal, note que a matiz de Pauli Z pode ser escrita como

[1] A função característica é um polinômio de grau igual ao tamanho da matriz, ou seja, número de linhas e/ou número de colunas.

$$Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = |0\rangle\langle 0| - |1\rangle\langle 1|,$$

onde a representação matricial é feita com respeito aos vetores ortonormais $|0\rangle$ e $|1\rangle$, respectivamente. Às vezes, as representações diagonais são conhecidas como *decomposição ortonormal*.

Quando um autoespaço tem dimensão maior que um, diz-se que ele é *degenerado*. Por exemplo, a matriz A definida por

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

tem um autoespaço bidimensional correspondendo ao autovalor 2. Os auto vetores $(1, 0, 0)$ e $(0, 1, 0)$ são ditos *degenerados* por serem autovetores linearmente independentes de A com o mesmo autovalor.

3.1.5 Operadores Hermitianos e adjuntos

Suponha que A é qualquer operador linear no espaço de Hilbert V. Pode ser mostrado que existe um único operador linear $A^\dagger$ em V tal que para todos os vetores $|v\rangle$, $|w\rangle \in V$,

$$(|v\rangle, A|w\rangle) = (A^\dagger|v\rangle, |w\rangle).$$

Este operador linear é conhecido com *adjunto* ou *Hermitiano conjugado* do operador A. A partir da definição é fácil ver que $(AB)^\dagger = B^\dagger A^\dagger$. Por convenção, se $|v\rangle$ é um vetor, então defini-se $|v\rangle^\dagger \equiv \langle v|$. Com isso, não é difícil ver que $(A|v\rangle)^\dagger \equiv \langle v|A^\dagger$.

3.1.6 Produto Tensorial

O *produto tensorial* é uma forma de juntar espaços vetoriais, a fim de formar espaços vetoriais maiores. Esta construção é crucial para entender a mecânica quântica de múltiplas partículas.

Suponha que V e W são espaços vetoriais de dimensão m e n, respectivamente. Suponha também que V e W são espaços de Hilbert. Então, o produto tensorial destes espaços vetoriais, $V \otimes W$, é um espaço vetorial $m \cdot n$ dimensional. Os elementos de $V \otimes W$ são combinações lineares dos 'produtos tensoriais' $|v\rangle \otimes |w\rangle$ de elementos $|v\rangle$ de V e $|w\rangle$ de W. Em particular, se $|i\rangle$ e $|j\rangle$ são bases ortonormais do espaço V e W, então $|i\rangle \otimes |j\rangle$ é uma base para $V \otimes W$.

Por exemplo, se V é uma espaço vetorial bidimensional com vetores da base $|0\rangle$ e $|1\rangle$, então $|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle$ é um elemento de $V \otimes V$.

Por definição o produto tensorial satisfaz algumas propriedades básicas:

1. Para um escalar arbitrário z e elementos $|v\rangle$ de V e $|w\rangle$ de W,

$$z(|v\rangle \otimes |w\rangle) = (z|v\rangle) \otimes |w\rangle = |v\rangle \otimes (z|w\rangle)$$

2. Para vetores arbitrários $|v_1\rangle$ e $|v_2\rangle$ em V e $|w\rangle$ em W,

$$(|v_1\rangle + |v_2\rangle) \otimes |w\rangle = |v_1\rangle \otimes |w\rangle + |v_2\rangle \otimes |w\rangle$$

3. Para um $|v\rangle$ arbitrário em V e $|w_1\rangle$ e $|w_2\rangle$ em W,

$$|v\rangle \otimes (|w_1\rangle + |w_2\rangle) = |v\rangle \otimes |w_1\rangle + |v\rangle \otimes |w_2\rangle$$

Suponha que $|v\rangle$ e $|w\rangle$ são vetores em V e W, respectivamente. Então, pode-se definir o operador linear $A \otimes B$ em $V \otimes W$ pela equação

$$(A \otimes B)(|v\rangle \otimes |w\rangle) \equiv A|v\rangle \otimes B|w\rangle.$$

A definição de $A \otimes B$ é estendida para todos os elementos de $V \otimes W$ de maneira natural, para assegurar a linearidade de $A \otimes B$, isto é

$$(A \otimes B) \left(\sum_i a_i |v_i\rangle \otimes |w_i\rangle \right) \equiv \sum_i a_i A|v_i\rangle \otimes B|w_i\rangle. \tag{3.1}$$

Pode ser mostrado que $A \otimes B$ definido desta maneira é um operador linear bem definido em $V \otimes W$.

A representação matricial é conhecida como *Produto de Kronecker*. Suponha que A é uma matriz $m \times n$, e B uma matriz $p \times q$. Então tem-se a representação matricial:

$$A \otimes B \equiv \begin{bmatrix} A_{11}B & A_{12}B & \dots & A_{1n}B \\ A_{21}B & A_{22}B & \dots & A_{2n}B \\ \vdots & \vdots & \vdots & \vdots \\ A_{m1}B & A_{m2}B & \dots & A_{mn}B \end{bmatrix}$$

Os termos como $A_{11}B$ denotam submatrizes p por q cujas entradas são proporcionais a B, com constante de proporcionalidade global A_{11}.

Por exemplo, o produto tensorial dos vetores (1, 2) e (2, 3) é o vetor:

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} \otimes \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \times 2 \\ 1 \times 3 \\ 2 \times 2 \\ 2 \times 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 4 \\ 6 \end{bmatrix}$$

O produto tensorial das matrizes de Pauli X e Y é:

$$X \otimes Y = \begin{bmatrix} 0 \cdot Y & 1 \cdot Y \\ 1 \cdot Y & 0 \cdot Y \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{bmatrix}$$

Outra notação útil é $|\psi\rangle^{\otimes k}$, que significa o produto tensorial de $|\psi\rangle$ com ele mesmo k vezes. Uma notação análoga é usada também para operadores em espaços de produto tensorial.

A *decomposição espectral* é um teorema de representação extremamente útil para operadores normais.

Teorema 3.1.1. ***(Decomposição Espectral)***
Qualquer operador normal M num espaço vetorial V é diagonal com respeito à alguma base ortonormal para V. Reciprocamente, qualquer operador normal é diagonalizável.

3.1.7 Funções de operadores

Há muitas funções importantes que podem ser definidas por operadores e matrizes. De forma geral, dada uma função f dos números complexos para os números complexos, é possível definir uma função matricial correspondente em matrizes normais (ou alguma subclasse como matrizes Hermitianas) pela seguinte construção. Tome $A = \sum_a a|a\rangle\langle a|$ como uma decomposição espectral para um operador normal A. Defina $f(A) \equiv \sum_a f(a)|a\rangle\langle a|$. Note que $f(a)$ é unicamente definida. Este procedimento pode ser usado, por exemplo, para definir a raiz quadrada de um operador definido-positivo, ou a exponencial de um operador normal, como por exemplo:

$$exp(\theta Z) = \begin{bmatrix} e^{\theta} & 0 \\ 0 & e^{-\theta} \end{bmatrix}$$

dado que Z tem autovalores $|0\rangle$ e $|1\rangle$.

3.1.8 O comutador e o anti-comutador

O *comutador* entre dois operadores A e B é definido como

$$[A,\, B] \equiv AB - BA.$$

Se $[A,\, B] = 0$, isto é, $AB = BA$, então diz-se que A *comuta* com B.

De modo similar, o *anti-comutador* de dois operadores A e B é definido por

$$\{A,\, B\} \equiv AB + BA;$$

diz-se que A *anti-comuta* com B se $\{A,\, B\} = 0$.

Disto resulta que muitas propriedades de pares de operadores podem ser deduzidas a partir de seus comutadores e anti-comutadores. Talvez, a relação mais útil seja a conexão entre o comutador e a propriedade de os operadores A e B serem Hermitianos

e diagonalizáveis simultaneamente, isto é, escrever $A = \sum_i a_i |i\rangle\langle i|$, $B = \sum_i b_i |i\rangle\langle i|$, onde $|i\rangle$ é algum conjunto comum ortonormal de autovetores para A e B.

Teorema 3.1.2. *(**Diagonalização Simultânea**): Suponha que A e B são operadores Hermitianos. Então $[A, B] = 0$ se, e somente se existe uma base ortonormal tal que ambos A e B são diagonais com respeito à esta base. Diz-se que A e B são* simultaneamente diagonalizáveis *neste caso.*

Este resultado conecta o comutador de dois operadores, que é frequentemente fácil de computar, à propriedade de serem simultaneamente diagonalizáveis, que é *a priori*, particularmente difícil de se determinar.

Como exemplo considere:

$$[X, Y] = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} - \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = 2i \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = 2iZ.$$

então, X e Y não comutam, e não têm autovetores comuns, como se esperava pelo Teorema da diagonalização simultânea.

3.1.9 Valores polar e singular de decomposição

As decomposições em valores *polar* e *singular* são estratégias úteis para separar operadores lineares em partes mais simples. Em particular, esta decomposição nos permite 'quebrar' operadores lineares gerais em produtos de operadores unitários e operadores positivos.

Teorema 3.1.3. *(**Decomposição polar**) Tome A como um operador linear num espaço vetorial V. Então, existem operadores unitários U e operadores positivos J e K tais que*

$$A = UJ = KU,$$

onde os únicos operadores positivos J e K satisfazendo estas equações, são definidos por $J \equiv \sqrt{A^\dagger A}$ e $K \equiv \sqrt{AA^\dagger}$. Além disso, se A é inversível, então U é único. A expressão $A = UJ$ é chamada decomposição polar à esquerda *de A, e $A = KU$ de* decomposição polar à direita *de A. Na maioria das vezes, a nomenclatura 'direita' e 'esquerda' é omitida e o termo 'decomposição polar' é usado em ambos os casos, com o contexto indicando o significado.*

O valor singular de decomposição combina a decomposição polar e o teorema da espectral.

Corolário 3.1.4. *(**Valor singular de decomposição**): Tome A como sendo uma matriz quadrada. Então existem matrizes unitárias U e V, e uma matriz diagonal D com entradas não-negativas tais que*

$$A = UDV.$$

Os elementos da matriz diagonal D são chamados valores singulares *de A.*

3.2 OS POSTULADOS DA MECÂNICA QUÂNTICA

A mecânica quântica é uma estrutura matemática para o desenvolvimento de teorias físicas. Por si só, a mecânica quântica não informa quais leis um sistema físico deve obedecer, mas fornece uma estrutura matemática e conceitual para o desenvolvimento dessas leis.

Os postulados fornecem uma conexão entre o mundo físico e o formalismo matemático da mecânica quântica. Muitas vezes as motivações para os postulados não são muito claras. Mas o objetivo aqui é conhecer os postulados, e, quando e como aplicá-los.

3.2.1 Espaço de estado

O primeiro postulado da mecânica quântica estabelece o cenário no qual ela se desenvolve. Este cenário é o espaço de Hilbert.

Postulado 1: Associado a qualquer sistema físico isolado há um espaço vetorial complexo com produto interno (isto é, um espaço de Hilbert) conhecido como *espaço de estado* do sistema. O sistema é completamente descrito pelo seu *vetor de estado*, que é um vetor unitário no espaço de estado do sistema.

O sistema mais simples da mecânica quântica, e que é de interesse aqui, é o *qubit*. Um qubit tem um espaço de estados bidimensional. Suponha que $|0\rangle$ e $|1\rangle$ formam uma base ortonormal para este espaço de estado. Então, um vetor de estado arbitrário neste

espaço de estado pode ser escrito como:

$$|\psi\rangle = a|0\rangle + b|1\rangle,$$

onde a e b são números complexos. A condição de que $|\psi\rangle$ seja um vetor unitário, $\langle\psi|\psi\rangle = 1$, é então, equivalente a $|a|^2 + |b|^2 = 1$. A condição $\langle\psi|\psi\rangle = 1$ é frequentemente chamada de *condição de normalização* para vetores de estado.

Intuitivamente, os estados $|0\rangle$ e $|1\rangle$ são análogos aos valores 0 e 1 que um bit pode tomar. A maneira na qual um qubit difere de um bit é que a *superposição* de seus dois estados, na forma, $a|0\rangle + b|1\rangle$, também pode existir, e, não é possível dizer que o qubit está definitivamente no estado $|0\rangle$, ou definitivamente no estado $|1\rangle$.

Diz-se que qualquer combinação linear $\sum_i \alpha_i|\psi_i\rangle$ é uma superposição dos estados $|\psi_i\rangle$ com *amplitude* α_i para o estado $|\psi_i\rangle$. Então, por exemplo, o estado

$$\frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

é uma superposição dos estados $|0\rangle$ e $|1\rangle$ com amplitude $1/\sqrt{2}$ para o estado $|0\rangle$ e $-1/\sqrt{2}$ para o estado $|1\rangle$.

3.2.2 Evolução

Este postulado nos diz como o estado $|\psi\rangle$ de um sistema quântico muda com o tempo.

Postulado 2: A evolução de um sistema quântico *fechado* é descrito por uma *transformação unitária*. Isto é, o estado $|\psi\rangle$ de um sistema no tempo t_1 é relacionado ao estado $|\psi'\rangle$ do sistema no tempo t_2 por um operador unitário U. Tal operador depende apenas dos tempos t_1 e t_2,

$$|\psi'\rangle = U|\psi\rangle.$$

Como a mecânica quântica não informa o espaço de estado ou o estado quântico de um sistema quântico *particular*, não há informação sobre qual operador unitário U descreve a

dinâmica quântica do mundo real. A mecânica quântica apenas assegura que a evolução de qualquer sistema quântico fechado pode ser descrita de alguma maneira. No caso de um único qubit, isto implica que *qualquer* operador unitário pode ser considerado em sistemas reais.

Alguns exemplos de operadores unitários em um único qubit são importantes em computação e informação quântica, como as matrizes de Pauli. A matriz X é muitas vezes chamada de porta lógica quântica NOT. As matrizes X e Z de Pauli também são conhecidas como matrizes "*bit flip*" (inversão) e "*phase flip*" (mudança de fase): a matriz X transforma $|0\rangle$ em $|1\rangle$, e $|1\rangle$ em $|0\rangle$, daí o nome "bit flip"; e a matriz Z deixa o $|0\rangle$ invariante e transforma $|1\rangle$ em $-|1\rangle$, com o fator extra, -1 conhecido como *fator de fase*, justificando o termo usado.

Outro operador unitário interessante é o operador de *porta lógica Hadamard*, denotado por H. Sua ação é $H|0\rangle \equiv (|0\rangle + |1\rangle)/\sqrt{2}$, $H|1\rangle \equiv (|0\rangle - |1\rangle)/\sqrt{2}$. Sua representação matricial é:

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

3.2.3 Medição quântica

O Postulado 3, fornece um meio para descrever os efeitos das medições em sistemas quânticos.

Postulado 3: Medições quânticas são descritas por uma coleção $\{M_m\}$ de *operadores de medida*. Estes operadores atuam no espaço de estado do sistema a ser medido. O índice m se refere aos resultados medidos que podem ocorrer no experimento. Se o estado do sistema quântico é $|\psi\rangle$ imediatamente antes da medida então, a probabilidade de m ocorrer é dada por

$$p(m) = \langle\psi|M_m^\dagger M_m|\psi\rangle,$$

e o estado do sistema após a medição é

$$\frac{M_m|\psi\rangle}{\sqrt{\langle\psi|M_m^\dagger M_m|\psi\rangle}}.$$

Os operadores de medida satisfazem a *equação de completude*,

$$\sum_m M_m^\dagger M_m = I.$$

A equação de completude expressa o fato de que as probabilidades somam 1:

$$1 = \sum_m p(m) = \sum_m \langle\psi|M_m^\dagger M_m|\psi\rangle.$$

Esta equação, que é satisfeita por todos os $|\psi\rangle$, é equivalente à equação de completude. Porém, a equação de completude é mais fácil de ser checada diretamente, então este é o motivo de ela aparecer na declaração deste postulado.

Um exemplo simples, porém importante de medição é a *medição de um qubit na base computacional*. Esta é uma medição num único qubit com dois resultados definidos por dois operadores de medição $M_0 = |0\rangle\langle 0|$, $M_1 = |1\rangle\langle 1|$. Observe que cada operador de medição é Hermitiano, e que $M_0^2 = M_0$, $M_1^2 = M_1$. Como a relação de completude é obedecida, $I = M_0^\dagger M_0 + M_1\dagger M_1 = M_0 + M_1$. Suponha que o estado a ser medido é $|\psi\rangle = a|0\rangle + b|1\rangle$. Então, a probabilidade de se obter o resultado 0 é

$$p(0) = \langle\psi|M_0^\dagger M_0|\psi\rangle = \langle\psi|M_0|\psi\rangle = |a|^2.$$

De modo similar, a probabilidade de se obter a medição do resultado 1 é $p(1) = |b|^2$. O estado depois da medição nos dois casos é então:

$$\frac{M_0|\psi\rangle}{|a|} = \frac{a}{|a|}|0\rangle$$

$$\frac{M_1|\psi\rangle}{|b|} = \frac{b}{|b|}|1\rangle.$$

A condição do Postulado 3 como um postulado fundamental intriga muitas pessoas. Instrumentos de medição são sistemas da mecânica quântica, tais que o sistema quântico a ser medido e o instrumento de medição juntos, fazem parte de uma sistema quântico maior e isolado.

De acordo com o Postulado 2, a evolução deste sistema isolado maior pode ser descrito por uma evolução unitária. É possível derivar o Postulado 3 como uma consequência deste fato? Apesar das investigações consideráveis em torno deste assunto, ainda há uma divergência entre físicos sobre se isto é ou não é possível. Aqui, a abordagem será mais pragmática, que, na prática, é elucidar quando aplicar o Postulado 2 e quando aplicar o Postulado 3, sem a preocupação de derivar um postulado do outro.

3.2.4 Distinguindo estados quânticos

Uma aplicação importante do Postulado 3 é o problema de *distinguir estados quânticos*. No mundo clássico, estados distintos de um objeto são usualmente distinguíveis, pelo menos em princípio. Por exemplo, pode-se sempre identificar quando o lançamento de uma moeda resultou em cara ou coroa, pelo menos no limite ideal. Na mecânica quântica, a situação é um pouco mais complicada.

O Postulado 3 fornece uma demonstração convincente do fato de que estados quânticos não ortogonais não podem ser distinguidos.

Distinguibilidade, como muitas outras ideias da mecânica quântica e da informação quântica, é mais facilmente entendida usando metáforas de jogos envolvendo duas partes, vamos chamá-las de Alice e Bob. Alice escolhe um estado $|\psi_i\rangle (1 \leq i \leq n)$ a partir de algum conjunto fixo de estados conhecidos por ambas as partes. Ela fornece o estado $|\psi_i\rangle$ para Bob, cuja tarefa é identificar o índice i do estado que Alice deu à ele.

Suponha que os estados $|\psi_i\rangle$ são ortonormais. Então, Bob pode fazer uma medição quântica para *distinguir* estes estados, usando o seguinte procedimento. Definir operadores de medição $M_i = |\psi_i\rangle\langle\psi_i|$, um para cada possibilidade de índices i, e um operador de medição adicional M_0, definido como a raiz quadrada positiva do operador positivo $I - \sum_{i \neq 0} |\psi_i\rangle\langle\psi_i|$. Estes operadores satisfazem a relação de completude, e se o estado $|\psi_i\rangle$ é preparado, então $p(i) = \langle\psi_i|M_i|\psi_i\rangle = 1$, então o resultado i ocorre com certeza.

Portanto, é possível distinguir, de forma confiável, o estado ortonormal $|\psi_i\rangle$.

Em contrapartida, se os estados $|\psi_i\rangle$ não são ortonormais, então pode-se provar que *não há medição quântica capaz de distinguir os estados*. A ideia é que Bob fará a medição descrita por operadores de medição M_j, com resultado j. Dependendo do resultado da medição, Bob tenta adivinhar qual foi o índice i usando alguma regra, $i = f(j)$, onde $f(\cdot)$ representa a regra que ele usa para dar o palpite. O ponto chave pelo qual Bob não pode distinguir estados não-ortogonais $|\psi_1\rangle$ e $|\psi_2\rangle$ é a observação de que $|\psi_2\rangle$ pode ser decomposto em um componente (não-nulo) paralelo à $|\psi_1\rangle$, e uma componente ortogonal à $|\psi_1\rangle$. Suponha que j seja um resultado da medição tal que $f(j) = 1$, isto é, Bob acha que o estado foi $|\psi_1\rangle$ quando ele observou j. Mas, por conta da componente de $|\psi_2\rangle$ paralela à $|\psi_1\rangle$, há uma probabilidade não nula de se ter o resultado j quando $|\psi_2\rangle$ é preparado, então, às vezes, Bob cometerá um erro ao identificar qual estado foi preparado.

3.2.5 Medição projetiva

Nesta seção será explicado um caso especial do postulado geral de medição, o Postulado 3, conhecido como *medição projetiva*.

Medição Projetiva: A medição projetiva é descrita por um *observável*, M, um operador Hermitiano no espaço de estado do sistema a ser observado. O observável tem uma decomposição espectral,

$$M = \sum_m m P_m,$$

onde P_m é o projetor no autoespaço de M com autovalor m. Os resultados possíveis da medição correspondem aos autovalores, m, do observável. Depois de medir o estado $|\psi\rangle$, a probabilidade de obter o resultado m é dada por

$$p(m) = \langle\psi|P_m|\psi\rangle.$$

Dado que o resultado m ocorreu, o estado do sistema quântico imediatamente após a medição é

$$\frac{P_m|\psi\rangle}{\sqrt{p(m)}}.$$

As medições projetivas podem ser entendidas como casos especiais do Postulado 3. Suponha que os operadores de medição do Postulado 3, além de obedecerem à relação de completude $\sum_m {M_m}^\dagger M_m = I$, também satisfaçam as condições de que M_m são projetores ortogonais, isto é, M_m são Hermitianos, e $M_m M'_m = \delta_{m,m'} M_m$. Com essas restrições adicionais, o Postulado 3 se reduz à medição projetiva como definida anteriormente.

Medições projetivas tem muitas propriedades sutis. Em particular, é fácil calcular valores médios para medições projetivas. Por definição, o valor médio da medição é

$$\begin{aligned}
\mathbf{E}(M) &= \sum_m mp(m) \\
&= \sum_m m\langle\psi|P_m|\psi\rangle \\
&= \langle\psi|\left(\sum_m mp(m)\right)|\psi\rangle \\
&= \langle\psi|M|\psi\rangle.
\end{aligned}$$

Esta é uma fórmula útil, que simplifica muitos cálculos. O valor esperado de um observável M é frequentemente escrito como $\langle M\rangle \equiv \langle\psi|M|\psi\rangle$. A partir desta fórmula para o valor médio segue-se uma fórmula para o desvio padrão associado às observações de M,

$$\begin{aligned}
[\Delta(M)]^2 &= \langle(M - \langle M\rangle)^2\rangle \\
&= \langle M^2\rangle - \langle M\rangle^2.
\end{aligned}$$

O desvio padrão é uma medida de dispersão típica dos valores observados sobre medidas de M. Em particular, se forem feitos um grande número de experimentos no qual o estado $|\psi\rangle$ é preparado e o observável M é medido, então o desvio padrão $\Delta(M)$ dos valores observados é determinado pela fórmula $\Delta(M) = \sqrt{\langle M^2\rangle - \langle M\rangle^2}$. Esta formulação de medição e desvio padrão em termos de observáveis fornece um aspecto elegante para resultados como o *princípio da incerteza de Heinsenberg*.

3.2.6 Medição POVM

O Postulado da medição quântica, Postulado 3, envolve dois elementos. Primeiro, é dada uma regra descrevendo medições estatísticas, isto é, as respectivas probabilidades das medições dos diferentes resultados possíveis. Segundo, é dada uma regra descrevendo o estado do sistema após a medição. No entanto, para algumas aplicações, este estado pós-medição é de pouco interesse, sendo que o aspecto de maior interesse está nas probabilidades das medições dos respectivos resultados. Neste caso, por exemplo, em um experimento onde o sistema é medido apenas uma vez, sob conclusão do experimento. Nestes casos há uma ferramenta matemática conhecida como *POVM formalismo* que é especialmente bem adaptado à análises de medições. (O acrônimo POVM vem do inglês 'Positive Operator-Valued Measure' - 'Medida com Operador de Valor Positivo').

Suponha que a medição escrita pelos operadores de medição M_m é preparado sobre um sistema quântico no estado $|\psi\rangle$. Então, a probabilidade de um resultado m é dado por $p(m) = \langle\psi|M_m^\dagger M_m|\psi\rangle$. Suponha que seja definido

$$E_m \equiv M_m^\dagger M_m.$$

Então, do Postulado 3 e da álgebra linear, E_m é um operador positivo tal que $\sum_m E_m = I$ e $p(m) = \langle\psi|E_m|\psi\rangle$. Então, um grupo de operadores E_m é suficiente para determinar as probabilidades de diferentes medições de resultados (saídas). Os operadores E_m são conhecidos como $elementos POVM$ associados à medição. O conjunto completo $\{E_m\}$ é conhecido como um POVM.

Como um exemplo de POVM, considere uma medição projetiva descrita pelos operadores de medição P_m, onde os P_m são projetores tais que, $P_m P_{m'} = \delta_{mm'} P_m$ e $\sum_m P_m = I$. Neste caso (e apenas neste caso) todo os elementos POVM são os mesmos, como os próprios operadores de medição, visto que $E_m \equiv P_m^\dagger P_m = P_m$.

3.2.7 Fase

'Fase' é um termo comumente usado em mecânica quântica, com muitos significados diferentes dependendo do contexto. Aqui, é conveniente revisar dois desses significados.

Considere, por exemplo, o estado $e^{i\theta}|\psi\rangle$, onde $|\psi\rangle$ é um vetor de estado, e θ é um número real. Diz-se que o estado $e^{i\theta}|\psi\rangle$ é igual a $|\psi\rangle$, a menos de um *fator de fase global* $e^{i\theta}$. É interessante notar que a *estatística de medição* preditas para esses dois estados é a mesma. Suponha que M_m é um operador de medição associado à alguma medição quântica, e note que as respectivas probabilidades de os resultados m ocorrerem são $\langle\psi|M_m^\dagger M_m|\psi\rangle$ e $\langle\psi|e^{-i\theta}M_m^\dagger M_m e^{i\theta}|\psi\rangle = \langle\psi|M_m^\dagger M_m|\psi\rangle$. Por conta disso, a partir de um ponto de vista observacional estes dois estados são idênticos. Por essa razão, pode-se ignorar fatores de fase global por serem irrelevantes para as propriedades observadas do sistema físico.

Há outro tipo de fase, conhecido como *fase relativa*, cujo significado é um pouco diferente. Considere os estados

$$\frac{|0\rangle + |1\rangle}{\sqrt{2}} \quad \text{e} \quad \frac{|0\rangle - |1\rangle}{\sqrt{2}}.$$

No primeiro estado, a amplitude de $|1\rangle$ é $1/\sqrt{2}$. Para o segundo estado, a amplitude é $-1/\sqrt{2}$. Em cada caso, a *magnitude* das amplitudes é a mesma, mas elas diferem quanto ao sinal. De forma geral, diz-se que as duas amplitudes, a e b, *diferem por uma fase relativa* se há um número real θ tal que $a = \exp(i\theta)b$. De forma ainda mais geral, dois estados são ditos *diferir por uma fase relativa* em alguma base se cada uma das amplitudes nestas bases é relacionada por algum fator de fase. Por exemplo, os dois estados mostrados anteriormente são os mesmos, a menos de um deslocamento relativo de fase, porque as amplitudes do $|0\rangle$ são idênticas (um fator de fase relativo de 1), e as amplitudes do $|1\rangle$ diferem apenas por um fator de fase relativo de -1. A diferença entre fatores de fase relativo e global é que para a fase relativa os fatores de fase podem variar de amplitude para amplitude. Isto faz da fase relativa um conceito que depende da base, diferente da fase global. Como resultado, estados que diferem apenas na fase relativa em algumas bases dão origem à diferenças fisicamente observáveis em medições estatísticas, e não é possível considerar estes estados como fisicamente equivalentes, como se faz com estados diferindo por um fator de fase global.

3.2.8 Sistemas compostos

Suponha que seja de interesse estudar um sistema quântico composto, formado por dois (ou mais) sistemas físicos distintos. Como os estados desse sistema composto podem

ser descritos? O postulado seguinte descreve como o espaço de estado de um sistema composto é constituído a partir do espaço de estado dos sistemas que o compõe.

Postulado 4: O espaço de estado de um sistema físico composto é o produto tensorial do espaço de estados dos sistema físicos que o compõe. Além disso, se há sistemas numerados de 1 a n, e o sistema de número i é preparado no estado $|\psi_i\rangle$, então o estado conjunto do sistema total é $|\psi_1\rangle \otimes |\psi_2\rangle \otimes \ldots \otimes |\psi_n\rangle$.

Porque o produto tensorial é a estrutura matemática usada para descrever o espaço de estado de um sistema composto? Em certa nível, pode-se simplesmente aceitar isto como um postulado básico, não redutível à algo mais elementar, e prosseguir. Afinal, certamente espera-se que exista alguma *forma canônica* de descrever sistemas compostos em mecânica quântica. Será que há outro caminho pelo qual possamos chegar à este postulado? Há uma heurística que às vezes é usada. Físicos, às vezes, gostam de falar em *princípio de superposição da mecânica quântica*, que estabelece que se $|x\rangle$ e $|y\rangle$ são dois estados de um sistema quântico, então qualquer superposição $\alpha|x\rangle + \beta|y\rangle$ também poderia ser um estado possível de um sistema quântico, onde $|a|^2 + |b|^2 = 1$. Para sistemas compostos, parece natural que se $|A\rangle$ é um estado do sistema A e $|B\rangle$ é um estado sistema B, então, pode haver algum estado correspondente, denotado por $|A\rangle|B\rangle$, do sistema conjunto AB. Aplicando o princípio da superposição à produtos de estados dessa forma, chega-se ao postulado do produto tensorial visto anteriormente. Não é uma derivação, visto que não se fala do princípio da superposição como uma parte fundamental da descrição da mecânica quântica, mas isto fornece uma gama das várias maneiras nas quais essas ideias são, às vezes, reformuladas.

Uma variedade de diferentes notações para diferentes para sistemas compostos aparecem na literatura. Parte da razão para esta proliferação é que diferentes notações são melhores adaptadas para diferentes aplicações, e pode-se achar conveniente introduzir algumas notações especializadas em cada ocasião. Neste ponto, é suficiente mencionar uma notação de subscrito útil para denotar estados e operadores em diferentes sistemas, quando não fica claro pelo contexto. Por exemplo, em um sistema contendo três qubits, X_2 é o operador de Pauli σ_x agindo no segundo qubit.

Observação 3.2.1. As provas de teoremas e corolários presentes neste capítulo podem ser encontradas em: [Nielsen and Chuang, 2010].

CAPÍTULO 4

ESPAÇOS DE PROBABILIDADE E DISTRIBUIÇÕES NÃO COMUTATIVAS

Neste capítulo analisa-se a estrutura analítica da probabilidade não comutativa. Sua principal característica reside no fato de que, aqui, permite-se que as álgebras de variáveis aleatórias sejam não comutativas. Isto significa que um conceito generalizado de variáveis aleatórias deve ser considerado, visto que na sua concepção usual, as álgebras de variáveis aleatórias precisariam ser comutativas.

Ao longo de todo este capítulo utiliza-se como fonte principal [Nica and Speicher, 2006], [García et al., 2007] e [Accardi, 1991].

4.1 ESPAÇOS DE PROBABILIDADE NÃO COMUTATIVOS

Definição 4.1.1. (1) Um *espaço de probabilidade não comutativo* $(\mathcal{A}, \varphi)$ consiste de uma álgebra unital [1] sobre $\mathbb{C}$ e um funcional linear unital

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

Os elementos $a \in \mathcal{A}$ são chamados de *variáveis aleatórias não comutativas* em $(\mathcal{A}, \varphi)$.

Uma propriedade adicional que, às vezes, será imposta ao funcional linear φ é a de que ele é um *traço*, isto é, possui a seguinte propriedade:

$$\varphi(ab) = \varphi(ba), \quad \forall a, b \in \mathcal{A}.$$

Quando isso acontece, diz-se que o espaço de probabilidade não comutativo $(\mathcal{A}, \varphi)$ é *tracial*.

[1] Que possui um elemento unitário. Por exemplo, no caso dos números reais com a multiplicação usual, o elemento unitário é o 1; no caso de matrizes quadradas com produto à direita, o elemento unitário é a matriz identidade.

(2) No esquema da parte (1) da definição, suponha que $\mathcal{A}$ é uma $*$ - álgebra, isto é, $\mathcal{A}$ também é dotada de uma operação antilinear, chamada de $*$-operação: $\mathcal{A} \ni a \mapsto a^* \in \mathcal{A}$, tal que $(a^*)^* = a$ e $(ab)^* = b^*\, a^*$ para todo $a\,, b \in \mathcal{A}$.

No caso em que se tem:

$$\varphi(a^*\, a) \geq 0, \qquad \forall a \in \mathcal{A},$$

então diz-se que o funcional φ é positivo.

(3) Na estrutura de um $*$-espaço de probabilidade encontram-se:

- variáveis aleatórias **autoadjuntas**: são os elementos $a \in \mathcal{A}$ com a propriedade $a = a^*$;
- variáveis aleatórias **unitárias**: são os elementos $u \in \mathcal{A}$ com a propriedade $u^*u = uu^* = 1_{\mathcal{A}}$;
- variáveis aleatórias **normais**: são os elementos $a \in \mathcal{A}$ com a propriedade $a^*a = aa^*$.

O objeto de interesse aqui é o $*$-espaço de probabilidade, que fornecerá ferramentas úteis em aplicações futuras, como na obtenção das $*$-distribuições de probabilidade.

Observação 4.1.1. Tome $(\mathcal{A}\,, \varphi)$ como um $*$-espaço de probabilidade.

(1) O funcional φ é autoadjunto, isto é, tem a propriedade

$$\varphi(a^*) \,=\, \varphi(a), \qquad \forall a \in \mathcal{A}.$$

De fato, visto que todo $a \in \mathcal{A}$ pode ser escrito unicamente na forma $a = x+iy$, onde $x\,, y \in A$ são autoadjuntos, a equação anterior é imediatamente vista como sendo equivalente ao fato de que $\varphi(x) \in \mathbb{R}$ para cada elemento autoadjunto $x \in \mathcal{A}$. Isto, por sua vez se deve à positividade de φ e ao fato de que todo elemento autoadjunto de $\mathcal{A}$ pode ser escrito na forma $x \,=\, a^*a - b^*b$ para algum $a\,, b \,\in\, \mathcal{A}$ (pode-se considerar como exemplo $a \,=\, (x + 1_{\mathcal{A}})/2, b \,=\, (x - 1_{\mathcal{A}})/2$).

(2) Outra consequência da positividade de φ é:

$$|\varphi(b^*a)|^2 \leq \varphi(a^*a)\varphi(b^*b), \qquad \forall a\,, b \in \mathcal{A}. \tag{4.1}$$

A desigualdade (4.1) é comumente chamada de *desigualdade de Cauchy-Schwarz* para o funcional φ.

(3) Se um elemento $a \in \mathcal{A}$ é tal que $\varphi(a^*a) \,=\, 0$, então a desigualdade de Cauchy-Schwarz (4.1) implica que $\varphi(ba) \,=\, 0$ para todo $b \in \mathcal{A}$ (então, a é, de certa forma, um elemento degenerado [2] do funcional φ). Usa-se o termo "fiel" para a situação na

[2] Significa dizer que $\varphi(a^*a)$ faz parte do núcleo de φ, ou seja, essa operação resulta no valor zero.

qual não existem tais elementos degenerados, exceto no caso em que $a = 0$, como é dito na definição seguinte.

Definição 4.1.2. Tome $(\mathcal{A}, \varphi)$ como um $*$-espaço de probabilidade . Se há a implicação:

$$a \in \mathcal{A}, \quad \varphi(a^*a) = 0 \Rightarrow a = 0,$$

Então, diz-se que o funcional φ é *fiel*.

Exemplo 4.1.2. Tome $(\Omega, \mathcal{Q}, P)$ como um espaço de probabilidade no sentido clássico, isto é, Ω é um conjunto, $\mathcal{Q}$ é uma σ-álgebra de subconjuntos de Ω e $P : \mathcal{Q} \to [0, 1]$ é uma medida de probabilidade. Fazendo $\mathcal{A} = L^\infty(\Omega, P)$ [3], e definindo φ por

$$\varphi(a) = \int_\Omega a(\omega) dP(\omega), \qquad a \in \mathcal{A}.$$

Então, $(\mathcal{A}, \varphi)$ é um $*$-espaço de probabilidade (a $*$-operação em $\mathcal{A}$ é a operação de complexo conjugado de uma função complexa). As variáveis aleatórias neste exemplo são, então, variáveis aleatórias genuínas no sentido "usual" da teoria de probabilidade.

Pode-se pensar que este exemplo só funciona com variáveis aleatórias que são limitadas, e se esquece, por exemplo, de uma das mais importantes variáveis aleatórias da probabilidade usual - àquelas que possuem distribuição Gaussiana. Este problema pode ser contornado, substituindo-se a álgebra $L^\infty(\Omega, P)$ por:

$$L^{\infty-}(\Omega, P) := \bigcap_{1 \leq p < \infty} L^p(\Omega, P).$$

Isto é, transforma-se a álgebra $\mathcal{A}$ em uma álgebra de variáveis aleatórias genuínas que possuem momentos finitos de todas as ordens. Neste caso, a álgebra considerada conteria as variáveis aleatórias gaussianas.

É lógico que há casos, na probabilidade clássica, de variáveis aleatórias que não possuem momentos de todas as ordens. Estas, entretanto, fogem ao escopo da teoria tratada neste texto, pois não se encaixam na definição dos espaços de probabilidade não comutativos.

Exemplo 4.1.3. Tome d como um inteiro positivo e $M_d(\mathbb{C})$ como a álgebra de matrizes complexas $d \times d$ com multiplicação usual de matrizes; $tr : M_d(\mathbb{C}) \to \mathbb{C}$ é o traço normalizado,

[3] Espaço das funções mensuráveis limitadas.

$$tr(a) = \frac{1}{d} \cdot \sum_{i=1}^{d} \alpha_{ii} \qquad \text{para} \qquad a = (\alpha_{ij})_{i,j=1}^{d} \in M_d(\mathbb{C}). \tag{4.2}$$

Então $(M_d(\mathbb{C}), tr)$ é uma $*$-espaço de probabilidade (onde a $*$-operação é dada a partir da transposta da matriz e do complexo conjugado de suas entradas).

Este é o exemplo de interesse no presente trabalho, pois ao tratar-se dos fenômenos relacionados à computação quântica, surge a descrição dos operadores lineares que, por sua vez, são representados por matrizes com entradas complexas.

Exemplo 4.1.4. Tome $\mathcal{H}$ como um espaço de Hilbert e $B(\mathcal{H})$ a álgebra de todos os operadores lineares limitados [4] em $\mathcal{H}$. Esta é uma $*$-álgebra, onde o adjunto a^* de um operador $a \in B(\mathcal{H})$ é unicamente determinado pelo fato de que:

$$\langle a\xi \, , \eta \rangle = \langle \xi \, , a^* \eta \rangle \, , \qquad \forall \xi \, , \eta \in \mathcal{H}.$$

Suponha que $\mathcal{A}$ é uma $*$-subálgebra de $B(\mathcal{H})$ e que $\xi_0 \in \mathcal{H}$ é um vetor de norma um ($||\xi_0|| := \langle \xi_0 \, , \xi_0 \rangle^{1/2} = 1$). Então, tem-se um exemplo de $*$-espaço de probabilidade $(\mathcal{A} \, , \varphi)$, onde $\varphi : \mathcal{A} \to \mathbb{C}$ é definido por:

$$\varphi(a) := \langle a\xi_0 \, , \xi_0 \rangle; \qquad a \in \mathcal{A} \tag{4.3}$$

Um funcional linear como o definido em (4.3) é usualmente chamado de *vetor de estado* (na álgebra de operadores $\mathcal{A}$).

Definição 4.1.3. (1) Um *morfismo* entre dois $*$-espaços de probabilidade $(\mathcal{A} \, , \varphi)$ e $(\mathcal{B} \, , \psi)$ é um homomorfismo unital da $*$-álgebra $\Phi : \mathcal{A} \to \mathcal{B}$ com a propriedade $\psi \circ \Phi = \varphi$.

(2) No caso em que $(\mathcal{B}, \psi)$ é um $*$-espaço de probabilidade do tipo especial discutido no exemplo 4.1.4, refere-se a um morfismo Φ de $(\mathcal{A} \, , \varphi)$ em $(\mathcal{B}, \psi)$ como uma *representação* de $(\mathcal{A} \, , \varphi)$. Então, para ser preciso, dada uma representação $(\mathcal{A} \, , \varphi)$ é o mesmo que uma tripla $(\mathcal{H}, \Phi, \xi_0)$ onde $\mathcal{H}$ é um espaço de Hilbert, $\Phi : \mathcal{A} \to B(\mathcal{H})$ é um $*$-homomorfismo unital, e $\xi_0 \in \mathcal{H}$ é um vetor de norma um, tal que $\varphi(a) = \langle \Phi(a)\xi_0, \xi_0 \rangle$ para todo $a \in \mathcal{A}$.

Observação 4.1.5. O $*$-espaços de probabilidade que aparecem nos exemplos 4.1.2 e 4.1.3 têm representações naturais nos espaços de Hilbert a partir de como as álgebras de variáveis aleatórias foram construídas.

[4] Ou seja, $B(\mathcal{H}) = \{\varphi : \mathcal{H} \to \mathcal{H} : ||\varphi|| \leq M, \, M \in \mathbb{R}\}$; e $\varphi(ax + y) = a\varphi(x) + \varphi(y)$.

4.2 $*$-DISTRIBUIÇÕES (CASO DOS ELEMENTOS NORMAIS)

Um conceito fundamental no estudo estatístico de variáveis aleatórias é o de distribuição de uma variável aleatória . Na estrutura do $*$-espaço de probabilidade $(\mathcal{A}, \varphi)$, o conceito apropriado é o de $*$-distribuição de um elemento $a \in \mathcal{A}$. À grosso modo, a $*$-distribuição de a tem de ser normalizada, de modo a ler o os valores de φ na $*$-subálgebra unital gerada por a.

O caso mais simples de $*$-distribuições ocorre quando a é normal, isto é, $a^*a = aa^*$. Neste caso, a $*$-álgebra unital gerada por a é:

$$\mathcal{A} := span\{a^k(a^*)^l \,|\, k, l \geq 0\}; \tag{4.4}$$

(onde *span* significa: "conjunto gerado por") a tarefa da $*$-distribuição de a deve, então, ser a de acompanhar os valores de $\varphi(a^k(a^*)^l)$, onde k e l assumem valores em $\mathbb{N} \cup \{0\}$. O tipo de objeto que cumpre essa tarefa e que é de interesse se obter sempre que possível, é uma medida de probabilidade com suporte compacto em $\mathbb{C}$.

Definição 4.2.1. Tome $(\mathcal{A}, \varphi)$ como $*$-espaço de probabilidade e a como um elemento normal de $\mathcal{A}$. Se existe uma medida de suporte compacto [5] μ em $\mathbb{C}$ tal que

$$\int z^k \overline{z}^l d\mu(z) \;=\; \varphi(a^k(a^*)^l), \qquad \text{para todo } k, l \in \mathbb{N} \tag{4.5}$$

então esta medida de probabilidade μ é unicamente determinada e será chamada de $*$-*distribuição* de a.

Observação 4.2.1. (1) O fato de uma medida de probabilidade μ com suporte compacto em $\mathbb{C}$ ser unicamente determinada por como ela integra funções da forma $z \mapsto z^k\overline{z}^l$ com $k, l \in \mathbb{N}$ é uma consequência imediata do teorema de Stone-Weierstrass [6]. Ou, mais precisamente; devido à Stone-Weierstrass , μ é determinado como um funcional linear no espaço $C(K)$ de funções contínuas complexas em K, onde K é o suporte de μ; é bem conhecido, por sua vez, que isto determina μ unicamente.

(2) Não é dito que todo elemento normal em um $*$-espaço de probabilidade tem, obrigatoriamente, uma $*$-distribuição no sentido definido anteriormente. Mas, isso se

[5] O suporte compacto de uma função é o conjunto fora do qual a função se anula. Um conjunto é compacto quando toda cobertura finita deste conjunto tem uma subcobertura finita.

[6] O Teorema de Stone-Weierstrass afirma que toda função real contínua cujo domínio é um intervalo compacto, ou seja, fechado e limitado pode ser aproximado uniformemente por polinômios. (Para mais detalhes ver [Rudin, 1976])

torna verdade para um bom número de exemplos importantes. De fato, isto é sempre verdade quando se olha para um $*$-espaço de probabilidade que tem uma representação no espaço de Hilbert, no sentindo da definição 4.1.3.

Observação 4.2.2. *(O caso do elemento auto-adjunto)*
Tome $(\mathcal{A}, \varphi)$ como um $*$-espaço de probabilidade e a como um elemento auto-adjunto de $\mathcal{A}$ (isto é, $a = a^*$; o que implica, em particular, que a é normal). Suponha que a tem $*$-distribuição μ, no sentido da definição 4.2.1. Então μ é suportada em $\mathbb{R}$. De fato, tem-se

$$\begin{aligned}\int_{\mathbb{C}} |z-\overline{z}|^2 d\mu(z) &= \int_{\mathbb{C}} (z-\overline{z})(\overline{z}-z) d\mu(z) \\ &= \int_{\mathbb{C}} (2z\overline{z} - z^2 - \overline{z}^2) d\mu(z) \\ &= 2\varphi(aa^*) - \varphi(a^2) - \varphi((a^*)^2) = 0.\end{aligned}$$

Visto que $z \mapsto |z-\overline{z}|^2$ é uma função contínua não negativa, deve-se ter que $z - \overline{z}$ desaparece no suporte $supp(\mu)$ da medida dada, e então:

$$supp(\mu) \subset \{z \in \mathbb{C} | z = \overline{z}\} = \mathbb{R}.$$

Então, nesse caso, μ é realmente uma medida em $\mathbb{R}$, e a equação (4.5) é melhor escrita nesse caso como segue-se abaixo, com $t \in \mathbb{R}$

$$\int t^p d\mu(t) = \varphi(a^p), \qquad \forall p \in \mathbb{N}. \tag{4.6}$$

De outra forma, supondo que exista uma medida μ com suporte compacto em $\mathbb{R}$ tal que (4.6) se mantém, então, claramente, μ é a $*$-distribuição de a no sentido da definição 4.2.1.

A conclusão retirada dessa discussão é que para um elemento auto-adjunto $a \in \mathcal{A}$ é mais apropriado falar na *distribuição* de a (ao invés de falar da sua $*$-distribuição); isto é definido como uma medida de suporte compacto em $\mathbb{R}$ tal que (4.6) se mantém.

Exemplo 4.2.3. (1) Considere o esquema do exemplo 4.1.2, onde a álgebra de variáveis aleatórias é $L^\infty(\Omega, P)$. Tome a como um elemento em $\mathcal{A}$; em outras palavras a é

uma função mensurável limitada, $a : \Omega \to \mathbb{C}$. Considere a medida de probabilidade ν em $\mathbb{C}$ que é chamada a "distribuição de a" na probabilidade usual; esta é definida por

$$\nu(E) = P(\{\omega \in \Omega : a(\omega) \in E\}), \qquad E \subset \mathbb{C} \quad \text{é um conjunto de Borel.} \tag{4.7}$$

Note que ν é compactamente suportado. Mais precisamente, se um r positivo é escolhido de forma que $|a(\omega)| \leq r, \quad \forall \omega \in \Omega$, então é claro que ν é suportado no disco fechado centrado em zero de raio r.

Agora, a é um elemento normal de $\mathcal{A}$ (todo elemento de $\mathcal{A}$ é normal visto que $\mathcal{A}$ é comutativa). Então, faz sentido colocar a na estrutura da definição 4.2.1. Será mostrado que a medida acima ν é exatamente a $*$-distribuição de a neste caso.

De fato, a equação (4.7) pode ser lida como

$$\int_{\mathbb{C}} f(z)d\nu(z) = \int_{\Omega} f(a(\omega))dP(\omega), \tag{4.8}$$

onde f é a função característica do conjunto E. Através do processo usual de tomar combinações lineares de funções características, e então fazer aproximações de uma função mensurável limitada por funções degrau, vê-se que a equação (4.8) de fato funciona para toda função mensurável $f : \mathbb{C} \to \mathbb{C}$. Finalmente, faça $k\,, l$ serem inteiros arbitrários não-negativos, e $r > 0$ tal que $|a(\omega)| \leq r$ para todo $\omega \in \Omega$. Considere uma função mensurável limitada $f : \mathbb{C} \to \mathbb{C}$ tal que $f(z) = z^k\overline{z}^l$ para todo $z \in \mathbb{C}$ tendo $|z| \leq r$. Dado que ν é suportado no disco fechado de raio r centrado em zero, segue que

$$\int_{\mathbb{C}} f(z)d\nu(z) = \int_{\mathbb{C}} z^k\overline{z}^l d\nu(z),$$

e, consequentemente que

$$\int_{\Omega} f(a(\omega))dP(\omega) = \int_{\Omega} a(\omega)^k\overline{a(\omega)}^l dP(\omega) = \varphi(a^k(a^*)^l).$$

Então para uma particular escolha de f, a equação (4.8) resulta em

$$\int_{\mathbb{C}} z^k\overline{z}^l d\nu(z) = \varphi(a^k(a^*)^l),$$

e isto é precisamente (4.5), implicando que ν é a $*$-distribuição de a no sentido da definição 4.2.1.

(2) Considere a estrutura do exemplo 4.1.3, e tome $a \in M_d(\mathbb{C})$ como uma matriz normal. Tome $\lambda_1, \ldots, \lambda_d$ como os autovalores de a, contando com as multiplicidades. Diagonalizando a temos que

$$tr(a^k(a^*)^l) = \frac{1}{d}\sum_{i=1}^{d} \lambda_i^k \overline{\lambda}_i^l, \qquad k\,, l \in \mathbb{N}$$

Esta última quantidade pode ser escrita como $\int z^k \overline{z}^l d\mu(z)$, onde

$$\mu := \frac{1}{d}\sum_{i=1}^{d} \delta_{\lambda_i} \tag{4.9}$$

(o δ_λ funciona, aqui, como o Delta de Dirac (*Dirac mass*) em $\lambda \in \mathbb{C}$). Segue-se que a tem uma $*$-distribuição μ, que é descrita pela Equação (4.9). Usualmente μ é chamada de *distribuição de autovalores* da matriz a.

Pode-se considerar a questão de como generalizar o fato acima no caso de matrizes aleatórias. Pode ser mostrado que a fórmula que apareceria no lugar da (4.9) seria

$$\mu := \frac{1}{d}\sum_{i=1}^{d} \int_\Omega \delta_{\lambda_i} dP(\omega), \tag{4.10}$$

onde $a = a^* \in M_d(L^{\infty-}(\Omega\,, P))$, e onde $\lambda_1(\omega) \leq \ldots \leq \lambda_d(\omega)$ são os autovalores de $a(\omega)\,, \omega \in \Omega$. Estritamente falando, a equação (4.10) requer uma extensão da estrutura usada na definição 4.2.1, visto que a distribuição de autovalores ponderados resultante, μ, geralmente não teria suporte compacto.

4.3 C^*-ESPAÇOS DE PROBABILIDADE

C^*-álgebras fornecem um ambiente natural onde as ideias de probabilidade não-comutativa podem ser trabalhadas. A ênfase aqui será a de conceituar o C^*-espaço de probabilidade e as relações entre *espectro* e $*$-distribuição para um elemento normal no C^*-espaço de probabilidade.

4.3.1 Cálculo funcional na C^*-álgebra

Um **C^*-espaço de probabilidade** é um $*$-espaço de probabilidade $(\mathcal{A}\,, \varphi)$ onde a $*$-álgebra $\mathcal{A}$ é tomada como sendo uma C^*-álgebra unital. Ser uma **C^*-álgebra** unital

significa que, além de ser uma $*$-álgebra unital, $\mathcal{A}$ é dotada de uma norma $||\cdot|| : \mathcal{A} \to [0\,, \infty)$, fazendo com que ela seja um *espaço vetorial normado e completo*, e tem-se:

$$||ab|| \leq ||a|| \cdot ||b||, \quad \forall a\,, b \in A; \tag{4.11}$$

$$||a^*a|| = ||a||^2, \quad \forall a \in \mathcal{A}. \tag{4.12}$$

Se $\mathcal{A}$ é uma C^*-álgebra unital e se $a \in \mathcal{A}$, então o espectro de a é o conjunto

$$\mathrm{Sp}(a) = \{z \in \mathbb{C} \,|\, z1_{\mathcal{A}} - a \text{ não é inversível } \}.$$

No Teorema a seguir, estão apresentadas algumas propriedades e cuja demonstração pode ser encontrada em [García et al., 2007].

Teorema 4.3.1. *Seja $\mathcal{A}$ uma C^*-álgebra unital.*

(1) *Para cada $a \in \mathcal{A}\,, \mathrm{Sp}(a)$ é um subconjunto compacto não-vazio de $\mathbb{C}$, contido no disco $\{z \in \mathbb{C} || z| \leq ||a||\}$.*

(2) *Seja a um elemento normal de $\mathcal{A}$, isto é, $a^*a = aa^*$, e considere a álgebra $C(\mathrm{Sp}(a))$ de funções contínuas de valor complexo em $\mathrm{Sp}(a)$. Existe um mapa $\Phi : C(\mathrm{Sp}(a)) \to \mathcal{A}$ com as seguintes propriedades:*

- **(i)** *Φ é um $*$-álgebra homomorfismo unital.*
- **(ii)** *$||\Phi(f)|| = ||f||_\infty, \forall f \in C(\mathrm{Sp}(a))$ (onde para $f \in C(\mathrm{Sp}(a))$ defini-se $||f||_\infty := \sup\{|f(z)||z \in \mathrm{Sp}(a)\}$).*
- **(iii)** *Se $id : \mathrm{Sp}(a) \to \mathbb{C}$ é a função identidade $id(z) = z$, tem-se $\Phi(id) = a$.*

Observação 4.3.2. Seja $\mathcal{A}$ uma C^*-álgebra unital, seja a um elemento normal de $\mathcal{A}$, e $\Phi : C(\mathrm{Sp}(a)) \to \mathcal{A}$ tem as propriedades **(i)**, **(ii)** e **(iii)** listadas no Teorema 4.3.1, item **(2)**.

(1) A condição **(ii)** (juntamente com a linearidade de **(i)**) implica que Φ é um-a-um. Consequentemente, em um certo sentido, Φ fornece uma cópia da álgebra $C(\mathrm{Sp}(a))$ que se encontra dentro de $\mathcal{A}$.

(2) Suponha que $p : \mathrm{Sp}(a) \to \mathbb{C}$ é um polinômio em z e $\overline{z}$, isto é, é da forma

$$p(z) = \sum_{j,\,k=0}^{n} \alpha_{j,\,k}\, z^j\, \overline{z}^k, \quad z \in \mathrm{Sp}(a). \tag{4.13}$$

$$\Phi(p) = \sum_{j,\,k=0}^{n} \alpha_{j,\,k}\, a^j (a^*)^k. \tag{4.14}$$

(3) A observação anterior mostra que os valores de Φ em polinômios em z e $\overline{z}$ são unicamente determinados. Visto que estes polinômios são densos em $C(\mathrm{Sp}(a))$ com respeito à convergência uniforme, e dado (por **(i)** + **(ii)**) Φ é contínuo com respeito à convergência uniforme, disto, segue que as propriedades **(i)**, **(ii)** e **(iii)** *determinam* Φ *unicamente*.

(4) O nome comumente usado para Φ é **cálculo funcional com funções contínuas** para o elemento a. Uma justificativa para esse nome pode ser vista olhando para polinômios p tais como aparecem na equação 4.13. De fato, para um tal p, o elemento correspondente $\Phi(p) \in \mathcal{A}$ (que aparece na equação 4.14) é o que naturalmente denota-se por "$p(a)$". É, de fato, costumeiro usar a notação

$$\text{“}f(a)\text{”} \text{ ao invés de } \text{“}\Phi(f)\text{”} \tag{4.15}$$

quando $f(a)$ é uma função arbitrária contínua em $\mathrm{Sp}(a)$ (não necessariamente um polinômio em z e $\overline{z}$).

Observação 4.3.3. Seja $\mathcal{A}$ uma C^*-álgebra unital. O Teorema 4.3.1 no item (2) contém, de uma maneira concentrada, uma grande quantidade de informação sobre o espectro de elementos normais de $\mathcal{A}$. Aqui, alguns fatos que são decorrentes deste teorema serão mostrados.

(1) Se a é um elemento normal de $\mathcal{A}$, então

$$||a|| = ||a^*|| = \sup\{|z| | z \in \mathrm{Sp}(a)\}. \tag{4.16}$$

Isto é visto através do uso de **(ii)** do Teorema 4.3.1 - parte (2) para as funções id e $\overline{id}$ no $\mathrm{Sp}(a)$.

(2) Se x é um elemento auto-adjunto de $\mathcal{A}$ então $\mathrm{Sp}(x) \subset \mathbb{R}$. De fato, quando se aplica **(ii)** do Teorema 4.3.1 - parte (2) à função $id - \overline{id}$ em $\mathrm{Sp}(x)$, obtém-se

$$||x - x^*|| = \sup\{|z - \overline{z}||z \in \mathrm{Sp}(x)\}. \tag{4.17}$$

O lado direito da equação 4.17 é 0; consequentemente o lado esquerdo também o é, e isto implica que $\mathrm{Sp}(x) \subset \{z \in \mathbb{C} : z - \overline{z} = 0\} = \mathbb{R}$.

De modo contrário, se $x \in \mathcal{A}$ é normal e tem-se que $\mathrm{Sp}(x) \subset \mathbb{R}$, então segue-se que $x = x^*$; novamente por 4.17, agora sabe-se que o lado direito desaparece.

(3) Se u é um elemento unitário de $\mathcal{A}$, então $\mathrm{Sp}(u) \subset \mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$. E, de modo contrário, se $u \in \mathcal{A}$ é normal e tem $\mathrm{Sp} \subset \mathbb{T}$, então u tem de ser unitário. O argumento é o mesmo da parte (2) dessa observação, onde agora usa-se a equação:

$$||1 - u^*u|| = \sup\{|1 - |z|^2||z \in \mathrm{Sp}(u)\}. \tag{4.18}$$

Teorema 4.3.4 (Teorema da aplicação espectral)**.** *Seja $\mathcal{A}$ uma C^*-álgebra unital, tome a como um elemento normal de $\mathcal{A}$, e tome $f : \mathrm{Sp}(a) \to \mathbb{C}$ como uma função contínua. Então, o elemento $f(a) \in \mathcal{A}$ (definido pelo cálculo funcional) tem*

$$\mathrm{Sp}(f(a)) = f(\mathrm{Sp}(a)). \tag{4.19}$$

Prova: Ver [Nica and Speicher, 2006].

Observação 4.3.5. Seja $\mathcal{A}$ uma C^*-álgebra unital. Costuma-se definir o conjunto dos *elementos positivos* de $\mathcal{A}$ como

$$\mathcal{A}^+ := \{p \in \mathcal{A}|p = p^* \text{ e } \mathrm{Sp}(p) \subset [0, \infty)\}. \tag{4.20}$$

Pode-se mostrar que

$$p, q \in \mathcal{A}^+, \alpha, \beta \in [0, \infty) \Rightarrow \alpha p + \beta q \in \mathcal{A}^+, \tag{4.21}$$

isto é, $\mathcal{A}^+$ é um cone convexo no espaço vetorial real de elementos auto-adjuntos de $\mathcal{A}$. Além disso, o cone $\mathcal{A}^+$ é "pontiagudo", no sentido de que $\mathcal{A}^+ \cap (-\mathcal{A}^+) = \{0\}$. (Ou, em outras palavras: se um elemento auto-adjunto $x \in \mathcal{A}$) é tal que se ambos, x e $-x$ estão em $\mathcal{A}^+$, então $x = 0$. Isto é, de fato, por que $x, -x \in \mathcal{A} \to \mathrm{Sp}(x) \subset [0, \infty) \cap (-\infty, 0] = \{0\} \to ||x|| = \sup\{|z||z \in \mathrm{Sp}(x) = 0.\}$

Nota-se também, que o teorema do mapeamento espectral fornece um rico suprimento de elementos positivos em $\mathcal{A}$. De fato, se a é um elemento arbitrário de $\mathcal{A}$ e se $f : \mathrm{Sp}(a) \to [0, \infty)$ é uma função contínua, então o elemento $f(a)$ está em $\mathcal{A}^+$ (é auto-adjunto porque $f = \overline{f}$, e tem um espectro em $[0, \infty)$ pelo Teorema 4.3.4).

É importante lembrar que um funcional linear $\varphi : \mathcal{A} \to \mathbb{C}$ é dito ser positivo quando satisfaz a condição $\varphi(a^*a) \geq 0$, $\forall a \in \mathcal{A}$. Isto levanta a questão de se há alguma relação entre $\mathcal{A}^+$ e o conjunto $\{a^*a | a \in \mathcal{A}\}$. É bastante conveniente que se esses dois conjuntos coincidam.

Proposição 4.3.6. *Seja $\mathcal{A}$ uma C^*-álgebra unital, e considere o conjunto $\mathcal{A}^+$ de elementos positivos de $\mathcal{A}$ (definido na equação 4.20 da observação anterior). Então*

$$\mathcal{A}^+ = \{a^*a \,|\, a \in \mathcal{A}\}. \tag{4.22}$$

Provar: Ver [Nica and Speicher, 2006].

4.3.2 C^*-espaços de probabilidade

Definição 4.3.1. Um C^*-espaço de probabilidade é um $*$-espaço de probabilidade $(\mathcal{A}, \varphi)$ onde $\mathcal{A}$ é uma C^*-álgebra unital.

Na estrutura de C^*, o funcional esperança é automaticamente contínuo. De forma precisa, tem-se o seguinte.

Proposição 4.3.7. *Tome $(\mathcal{A}, \varphi)$ como sendo um C^*-espaço de probabilidade.. Então*

$$|\varphi(a)| \leq ||a||, \quad \forall a \in \mathcal{A}. \tag{4.23}$$

Prova: Ver [Nica and Speicher, 2006].

Observação 4.3.8. O inverso da Proposição 4.3.7 também é verdadeiro. Tome $\mathcal{A}$ como uma C^*-espaço de probabilidade unital. Tome $\varphi : \mathcal{A} \to \mathbb{C}$ como um funcional linear tal que $|\varphi(a)| \leq ||a||, \quad \forall a \in \mathcal{A}$, e tal que $\varphi(1_{\mathcal{A}}) = 1$ (onde $1_{\mathcal{A}}$ é a unidade de $\mathcal{A}$). Então, φ é positivo e então $(\mathcal{A}, \varphi)$ é um C^*-espaço de probabilidade.

Exemplo 4.3.9. Tome Ω como um espaço topológico compacto de Hausdorff, e μ como uma medida aleatória de probabilidade na σ-álgebra de Borel de Ω. Perguntar se a medida de probabilidade μ é uma "medida aleatória" resulta em requerer que para todo conjunto de Borel $\mathcal{A} \in \Omega$ tem-se:

$$\mu(A) = \sup\{\mu(K)|K \subset \mathcal{A},\ \text{compacto }\} = \inf\{\mu(D)|D \supset A,\ \text{aberto }\}.$$

Em muitas situações naturais - quando Ω é um espaço métrico compacto por exemplo - tem-se que toda medida de probabilidade na σ-álgebra de Borel de Ω é, de fato, uma medida aleatória.

Considere a álgebra $\mathcal{A} = \mathbb{C}(\Omega)$, de funções complexas contínuas em Ω, e tome $\varphi : \mathcal{A} \to \mathbb{C}$ definido por:

$$\varphi(f) = \int_\Omega f d\mu,\ f \in \mathcal{A}. \tag{4.24}$$

Então $(\mathcal{A}, \varphi)$ é um C^*-espaço de probabilidade , e todos os elementos de $\mathcal{A}$ são normais. O cálculo funcional com funções contínuas para um elemento $a \in \mathcal{A}$ é reduzido neste caso para representar uma composição funcional.

Há dois importantes teoremas na análise funcional que valem a pena ser relembrados na conexão com esse exemplo. Primeiro, o teorema básico de Riesz determina que todo funcional linear positivo em $C(\Omega)$ pode ser colocado na forma da equação 4.24 para uma medida aleatória de probabilidade μ.

Segundo, o Teorema de Gelfand estabelece que toda C^*-espaço de probabilidade unital comutativa $\mathcal{A}$ pode ser identificada como $C(\Omega)$ para um espaço compacto de Hausdorff apropriado Ω. Portanto, o exemplo apresentado aqui é "genérico", tanto quanto os C^*-espaço de probabilidade comutativos estão interessados.

Em exemplos não comutativos, a C^*-espaço de probabilidade aparece mais frenquentemente como $*$-subálgebras $\mathcal{A} \subset B(\mathcal{H})$ ($\mathcal{H}$ é o espaço de Hilbert), tal que $\mathcal{A}$ é fechado na norma da topologia de $B(\mathcal{H})$.

4.3.3 $*$-distribuição norma e espectro para um elemento normal

As demonstrações dos próximos resultados podem ser encontradas em [Nica and Speicher, 2006].

Proposição 4.3.10. *Tome $(\mathcal{A}, \varphi)$ com um C^*-espaço de probabilidade e a como um*

elemento normal de $\mathcal{A}$. Então, a tem uma $*$-distribuição μ no sentido analítico (como descrito na definição 4.2.1). Além disso:

(1) *O suporte de μ está contido no espectro de a.*

(2) *Para $f \in \mathbb{C}(\mathrm{Sp}(a))$ tem-se a equação*

$$\int f d\mu = \varphi(f(a)), \tag{4.25}$$

onde, no lado direito $f(a) \in \mathcal{A}$ é obtido pelo cálculo funcional, e no lado esquerdo, μ é visto como uma medida de probabilidade em $\mathrm{Sp}(a)$.

Corolário 4.3.11. *Tome $(\mathcal{A}, \varphi)$ como um $*$-espaço. Se $(\mathcal{A}, \varphi)$ admite uma representação no espaço de Hilbert (no sentido na definição 4.1.3), então todo elemento normal de $\mathcal{A}$ tem uma $*$-distribuição no sentido analítico.*

No restante da seção vê-se alguns fatos adicionais que podem ser derivados de um C^*-espaço de probabilidade onde a esperança é fiel.

Proposição 4.3.12. *Tome $(\mathcal{A}, \varphi)$ como um C^*-espaço de probabilidade , onde φ é fiel. Tome a como sendo um elemento normal de $\mathcal{A}$, e μ como a $*$-distribuição de a no sentido analítico. Então, o suporte de μ é igual ao* $\mathrm{Sp}(a)$.

Observação 4.3.13. A proposição anterior pode ser interpretada como se segue: se $(\mathcal{A}, \varphi)$ é um C^*-espaço de probabilidade tal que φ é fiel, e se a é um elemento normal de $\mathcal{A}$, então, o conhecimento da $*$-distribuição μ de a permite que compute-se o espectro de a através da fórmula:

$$\mathrm{Sp}(a) = supp(\mu). \tag{4.26}$$

Note que ao conhecer μ, obtém-se a norma de a - de fato, por 4.26 e 4.16 da observação 4.3.3 segue-se que

$$||a|| = \sup\{|z||z \in supp(\mu).\} \tag{4.27}$$

A próxima proposição indica outro caminho (mais direto) de computar a norma de a a partir informação combinatorial dos $*$-momentos.

Proposição 4.3.14. *Tome $(\mathcal{A}, \varphi)$ como um C^*-espaço de probabilidade , onde φ é fiel. Para cada $a \in \mathcal{A}$ (normal ou não) tem-se que:*

$$||a|| = \lim_{n\to\infty} \varphi((a^*a)^n)^{1/2n}. \tag{4.28}$$

4.4 DISTRIBUIÇÕES CONJUNTAS NÃO COMUTATIVAS

Será denotada por $\mathbb{C}\langle X_1, \ldots, X_s\rangle$ a álgebra com unidade gerada livremente por s variáveis indeterminadas e não comutativas $X_1, \ldots, X_s$. Ou seja, os monômios da forma $X_{r_1}X_{r_2}\cdots X_{r_n}$, em que $n \geq 0$ e $1 \leq r_1, \ldots, r_n \leq s$ fornece uma base linear para $\mathbb{C}\langle X_1, \ldots, X_s\rangle$, e a multiplicação de dois de tais monômios é feita por justaposição.

Seja $\mathcal{A}$ uma álgebra com unidade, e $a_1, \ldots, a_s \in \mathcal{A}$. Para todo $P \in \mathbb{C}\langle X_1, \ldots, X_s\rangle$ denotaremos por $P(a_1, \ldots, a_s) \in \mathcal{A}$ que é obtido pela substituição de $X_1, \ldots, X_s$ por $a_1, \ldots, a_s$, respectivamente, na forma escrita explícita de P. Da mesma forma,

$$\mathbb{C}\langle X_1, \ldots, X_s\rangle \ni P \mapsto P(a_1, \ldots, a_s) \in \mathcal{A} \tag{4.29}$$

é o homomorfismo de álgebras com unidade determinada pelo fato que ela mapeia X_r em a_r, para $1 \leq r \leq s$.

Definição 4.4.1. Seja $(\mathcal{A}, \varphi)$ um espaço de probabilidade não comutativo, e sejam $a_1, \ldots, a_s$ elementos de $\mathcal{A}$.

1. A família
$$\{\varphi(a_{r_1}, \ldots, a_{r_n}) \,|\, n \geq 1,\, 1 \leq r_1, \ldots, r_s \leq s\} \tag{4.30}$$
é chamada a *família de momentos conjuntos* de $a_1, ..., a_s$.

2. O funcional linear $\mu : \mathbb{C}\langle X_1, \ldots, X_s\rangle \to \mathbb{C}$ definido por
$$\mu(P) := \varphi\big(P(a_1, \ldots, a_s)\big), \quad P \in \mathbb{C}\langle X_1, \ldots, X_s\rangle \tag{4.31}$$
é chamado a distribuição conjunta de $a_1, ..., a_s$ em $(\mathcal{A}, \varphi)$.

A distribuição conjunta de $a_1, ..., a_s$ é assim determinada pelo fato que ela mapeia todo monômio $X_{r_1}\cdots X_{r_n}$ no correspondente momento conjunto, $\varphi(a_{r_1}\cdots a_{r_n})$, de $a_1, ..., a_s$.

Observação 4.4.1. As definições acima podem ser estendidas para ocaso de uma família arbitrária $(a_i)_{i\in I}$ de variáveis aleatórias (I aqui é um conjunto de índices que poderia ser infinito, até não enumerável). A distribuição conjunta de $(a_i)_{i\in I}$ é então um funcional linear sobre a álgebra com unidade $\mathbb{C}\langle X_i \,|\, i \in I\rangle$, que é livremente gerada pelas variáveis indeterminadas não comutativas X_i ($i \in I$).

Exemplo 4.4.2. Seja $(\Omega, \mathcal{Q}, P)$ um espaço de probabilidade, e sejam $f_1, \ldots, f_s : \Omega \to \mathbb{R}$ variáveis aleatórias limitadas. Então $f_1, \ldots, f_s$ são ao mesmo tempo elementos do espaço de probabilidade não comutativo $L^\infty(\Omega, P)$ (com $\varphi(a) = \int_\Omega a(\omega)dP(\omega)$ para $a \in L^\infty(\Omega, P)$). A distribuição conjunta μ of $f_1, \ldots, f_s$ in $L^\infty(\Omega, P)$ é determinada pela fórmula:

$$\mu(X_{r_1} \cdots X_{r_n}) = \int_\Omega f_{r_1}(\omega) \cdots f_{r_n}(\omega)\, dP(\omega), \tag{4.32}$$

valendo para todo $n \geq 1$ e $1 \leq r_1, \ldots, r_n \leq s$.

Neste exemplo particular, existe um conceito paralelo distribuição conjunta de $f_1, \ldots, f_s$ vindo a partir da probabilidade clássica: isto é a medida ν sobre a σ-álgebra de Borel de $\mathbb{R}^s$ que tem para todo conjunto de Borel $E \subset \mathbb{R}^s$,

$$\nu(E) = P\big(\{\omega \in \Omega \,|\, (f_1(\omega), \ldots, f_2(\omega)) \in E\}\big). \tag{4.33}$$

Exemplo 4.4.3. Seja d um inteiro positivo e considere o $*$-espaço de probabilidade $(M_d(\mathbb{C}), \mathrm{tr})$ (o traço normalizado sobre matrizes complexas $d \times d$). Sejam $A_1, A_2 \in M_d(\mathbb{C})$ matrizes hermitianas. Sua distribuição conjunta $\mu : \mathbb{C}\langle X_1, X_2\rangle \to \mathbb{C}$ é determinada por

$$\mu(X_{r_1} \cdots X_{r_n}) = \mathrm{tr}(A_{r_1} \cdots A_{r_n}),\ \forall\, n \geq 1,\ \forall\, 1 \geq r_1, \ldots, r_n \leq 2. \tag{4.34}$$

A menos que A_1 e A_2 comutem, o funcional μ não pode ser trocado por um objeto mais simples (como uma medida de probabilidade sobre $\mathbb{R}^2$) que lembra a mesma informação.

Exemplo 4.4.4. Seja $(\mathcal{A}, \varphi)$ um $*$-espaço de probabilidade, e sejam x e y elementos de $\mathcal{A}$. Para todo $n \geq 1$ pode-se expandir $(x + y)^n$ como uma soma de $2n$ monômios não comutativos em x e y (obviamente apesar da fórmula binomial não se aplicar geralmente) como uma consequência, os momentos $\varphi((x + y)^n)$, $n \geq 1$ (e, portanto, a distribuição de $x + y$) são determinados pelo conhecimento da distribuição conjunta de x e y.

Por outro lado, está claro que, para x e y como acima, apenas sabendo quais são as distribuições individuais de x e de y não seria suficiente em geral para se determinar a distribuição de $x + y$.

Exemplo 4.4.5. Seja G um grupo e $g, h \in G$ dois elementos de ordem infinita. Considere o $*$-espaço de probabilidade $(\mathbb{C}G, \tau_G)$, como antes. Relembre que $\mathbb{C}G$ tem uma base canônica indexada por G; os elementos desta base são denotados pelas mesmas letras que os próprios elementos do grupo, e eles são unidades em $\mathbb{C}G$. Assim temos em particular que $g, h \in \mathbb{C}G$, e que $g^* = g^{-1}, h^* = h^{-1}$. Tais g e h tornam-se um unitário de Haar em $(\mathbb{C}G, \tau_G)$; como uma consequência, cada um dos elementos autoadjuntos $x := g + g^{-1}$ e

$y := h + h^{-1}$ tem uma distribuição arcsin. Assim, se no contexto do último parágrafo olharmos o elemento

$$\Delta := x + y = g + g^{-1} + h + h^{-1} \in \mathbb{C}G, \tag{4.35}$$

então Δ será sempre a soma de dois elementos autoadjuntos com distribuições arcsin. A fim de esclarecer as ideias, podemos considerar as seguintes situações:

(1) $G = \mathbf{Z}^2$, com $g = (1, 0)$ e $h = (0, 1)$. Nesta situação, o correspondente *grafo de Cayley*[7] é o reticulado $\mathbf{Z}^2$, e a contagem de caminhos fechados que produzem os momentos Δ é bastante simples. A fórmula obtida é

$$\tau_{\mathbf{Z}^2}(\Delta^n) = \begin{cases} 0, & \text{se } n \text{ é ímpar} \\ \binom{2p}{p}^2, & \text{se } n \text{ é par,} n = 2p. \end{cases} \tag{4.36}$$

(2) G é o grupo livre não comutativos sobre dois geradores, $G = \mathbb{F}_2$, e g, h são dois geradores livres de $\mathbb{F}_2$. Nesta situação, o grafo de Cayley que aparece é uma árvore e a contagem de caminhos fechados dá os momentos de Δ é devido a Kesten[8]. Obtém-se uma relação recorrente entre os momentos, que pode ser expressa concisamente como uma fórmula dando as séries geradoras de momentos:

$$\sum_{n=0}^{\infty} \tau_{\mathbb{F}_2}(\Delta^n) z^n = \frac{2\sqrt{1 - 12z^2} - 1}{1 - 16z^2} = 1 + 4z^2 + 28z^4 + 232z^6 + \cdots \tag{4.37}$$

Existe uma possível derivação da fórmula (4.37) que ilustra os métodos de probabilidade livre - isso porque na situação (2) os elementos $x = u + u^*$ e $y = v + v^*$ de $\mathbb{C}\mathbb{F}_2$ são *livremente independentes*, e consequentemente pode-se usar a técnica para computar a distribuição de uma soma dos elementos livremente independentes.

4.4.1 $*$-distribuições conjuntas

Seja $(\mathcal{A}, \varphi)$ um $*$-espaço de probabilidade e seja a um elemento de $\mathcal{A}$. Ao olhar para o que seria a $*$-distribuição de a no sentido algébrico, vemos que é isso realmente é a

[7] Suponha que G seja um grupo e S seja um conjunto de geradores. O grafo de Cayley $\Gamma = \Gamma(G, S)$ é um grafo direcionado colorido construído como se segue: (1) A cada elemento g de G é atribuído um vértice: o conjunto de vértices $V(\Gamma)$ de Γ é identificado com G. (2) A cada gerador s de S é atribuída uma cor c_s. (3) Para qualquer $g \in G$, $s \in S$, os vértices correspondentes aos elementos g e gs são unidos por uma aresta de cor c_s. Assim, o conjunto de arestas $E(\Gamma)$ consiste em pares da forma (g, gs), com $s \in S$ proporcionando a cor.

[8] H. Kesten: Symmetric random walks on groups, *Trans of the American Math Society* **92** (1959), 336-359.

mesma coisa que a distribuição conjunta de a e a^*, com a única diferença que trocamos a variável X_2 de $\mathbb{C}\langle X_1, X_2\rangle$ por X_1^*, e usamos esta notação para introduzir uma $*$-operação sobre $\mathbb{C}\langle X_1, X_2\rangle$. Será conveniente ter este formalismo posto também para n-uplas de elementos. Assim, introduz-se as notações a seguir.

Seja s um inteiro positivo.

1. Denota-se por $\mathbb{C}\langle X_1, X_1^*, \ldots, X_s, X_s^*\rangle$ a álgebra com unidade livremente gerada por $2s$ variáveis não comutativas $X_1, X_1^*, \ldots, X_s, X_s^*$, que tem uma $*$-operação natural, que é determinada fazendo com que a $*$-operação aplicada a X_r dá X_r^*, para $1 \leq r \leq s$.

2. Seja $\mathcal{A}$ uma $*$-álgebra com unidade e sejam $a_1, \ldots, a_s$ elementos de $\mathcal{A}$. Para todo $Q \in \mathbb{C}\langle X_1, X_1^*, \ldots, X_s, X_s^*\rangle$ denotaremos por $Q(a_1, \ldots, a_s)$ o elemento de $\mathcal{A}$ que é obtido substituindo-se X_1 por a_1, X_1^* por a_1^*, ..., X_s por a_s, X_s^* por a_s^* na escrita explicita de Q. Equivalentemente,

$$\mathbb{C}\langle X_1, X_1^*, \ldots, X_s, X_s^*\rangle \ni Q \mapsto Q(a_1, \ldots, a_s) \in \mathcal{A} \tag{4.38}$$

é um $*$-homomorfismo com unidade unicamente determinado pelo fato que ele mapeia X_r em a_r, para $1 \leq r \leq s$.

Definição 4.4.2. Seja $(\mathcal{A}, \varphi)$ um $*$-espaço de probabilidade, e sejam $a_1, \ldots, a_s$ elementos de $\mathcal{A}$.

1. A família
$$\left\{ \varphi(a_{r_1}^{\varepsilon_1}, \ldots, a_{r_n}^{\varepsilon_n}) \;\middle|\; \begin{array}{c} n \geq 1,\ 1 \leq r_1, \ldots, r_n \leq s \\ \varepsilon_1, \ldots, \varepsilon_n \in \{1, *\} \end{array} \right\} \tag{4.39}$$
é chamada a *família de $*$-momentos conjuntos* de $a_1, \ldots, a_s$.

2. O funcional linear $\mu : \mathbb{C}\langle X_1, X_1^*, \ldots, X_s, X_s^*\rangle \to \mathbb{C}$ definido por
$$\mu(Q) := \varphi\big(Q(a_1, \ldots, a_s)\big),\ Q \in \mathbb{C}\langle X_1, X_1^*, \ldots, X_s, X_s^*\rangle \tag{4.40}$$
é chamado a *$*$-distribuição conjunta* de $a_1, \ldots, a_s$ em $(\mathcal{A}, \varphi)$.

Exemplo 4.4.6. Seja θ um número em $[0, 2\pi]$. Suponha que $(\mathcal{A}, \varphi)$ é $*$-espaço de probabilidade em que a $*$-álgebra $\mathcal{A}$ é gerada por dois unitários u_1, u_2 que satisfazem

$$u_1 u_2 = e^{i\theta} u_2 u_1 \tag{4.41}$$

e em que $\varphi : \mathcal{A} \to \mathbb{C}$ é um funcional fiel e positivo tal que

$$\varphi(u_1^m u_2^n) = \begin{cases} 1, & \text{if } m = n = 0, \\ 0, & \text{caso contrário,} \end{cases} \quad \text{for } m, n \in \mathbf{Z}. \tag{4.42}$$

Observe que a partir de (4.41) temos

$$\begin{cases} (u_1^m u_2^n) \cdot (u_1^p u_2^q) = e^{-inp\theta}(u_1^{m+p} u_2^{n+q}), \\ (u_1^m u_2^n)^* = e^{-imn}(u_1^{-m} u_2^{-n}), \end{cases} \quad m, n \in \mathbf{Z}. \tag{4.43}$$

Isso implica que

$$\mathcal{A} = span\{u_1^m u_2^n \,|\, m, n \in \mathbf{Z}\} \tag{4.44}$$

Em particular isso mostra que o funcional linear φ é completamente descrito pela Equação (4.42). Outro fato que segue facilmente é que φ é um traço. De fato, verificar esta afirmação reduz a verificar que para todo $m, n, p, q \in \mathbf{Z}$ temos

$$\varphi\big((u_1^m u_2^n) \cdot (u_1^p u_2^q)\big)\varphi\big((u_1^p u_2^q) \cdot (u_1^m u_2^n)\big)$$

mas (de (4.42) e (4.43)) ambos os lados desta equação são iguais a $e^{-imn\theta}$, quando $(p, q) = -(m, n)$, e são iguais a 0 em todos os outros casos.

Seja $\mu : \mathbb{C}\langle X_1, X_1^*, X^2, X_2^*\rangle \to \mathbb{C}$ a $*$-distribuição conjunta dos unitários u_1 e u_2. Então para todo $n \geq 1$ e $r_1, \ldots, r_n \in \{1, 2\}$, $\varepsilon_1, \ldots, \varepsilon_n \in \{1, *\}$ o valor de μ no monômio $X_{r_1}^{\varepsilon_1} \cdots X_{r_n}^{\varepsilon_n}$ ou é 0, ou da forma $e^{ik\theta}$, para algum $k \in \mathbf{Z}$.

4.4.2 $*$-distribuições conjuntas e isomorfismo

Teorema 4.4.7. *Sejam $(\mathcal{A}, \varphi)$ e $(\mathcal{B}, \psi)$ $*$-espaços de probabilidade tais que φ e ψ sejam fieis. Denote as unidades de $\mathcal{A}$ e de $\mathcal{B}$ por $1_\mathcal{A}$ e $1_\mathcal{B}$, respectivamente. Suponha que $a_1, \ldots, a_s \in \mathcal{A}$ e $b_1, \ldots, b_s \in \mathcal{B}$ sejam tais que:*

(i) *$a_1, \ldots, a_s$ e $1_\mathcal{A}$ geram $\mathcal{A}$ como uma $*$-álgebra;*

(ii) *$b_1, \ldots, b_s$ e $1_\mathcal{B}$ geram $\mathcal{B}$ como uma $*$-álgebra;*

(iii) *A $*$-distribuição conjunta de $a_1, \ldots, a_s$ in $(\mathcal{A}, \varphi)$ é igual à $*$-distribuição conjunta de $b_1, \ldots, b_s$ em $(\mathcal{B}, \psi)$.*

Então existe um $*$-isomorfismo $\Phi : A \to B$, unicamente determinado, tal que $\Phi(a_1) = b_1, \ldots, \Phi(a_s) = b_s$. Tal Φ é também um isomorfismo entre $(\mathcal{A}, \varphi)$ e $(\mathcal{B}, \psi)$, i.e. ele tem a propriedade que $\psi \circ \Phi = \varphi$.

Observe que o tipo de isomorfismo que aparece no Teorema 4.4.7 é adequado para a categoria das C^*-álgebras com unidade, i.e. ele inclui a propriedade métrica apropriada de ser isométrica ($\|\Phi(a)\|_{\mathcal{B}} = \|a\|_{\mathcal{A}}$, para todo $a \in \mathcal{A}$). Vale ressaltar aqui que de fato um $*$-homomorfismo bijetivo com unidade entre C^*-álgebras com unidade é sempre isométrico (a propriedade métrica é uma consequência automática das propriedades algébricas).

O Teorema 4.4.7 tem uma versão para a qual as famílias $a_1, \ldots, a_s$ e $b_1, \ldots, b_s$ consistem de elementos autoadjuntos (de $\mathcal{A}$ e de $\mathcal{B}$, respectivamente), e em que a hipótese (iii) no teorema é ajustada para requerer que a distribuição conjunta de $a_1, \ldots, a_s$ em $(\mathcal{A}, \varphi)$ é igual à distribuição conjunta de $b_1, \ldots, b_s$ em $(\mathcal{B}, \psi)$.

Outra possível generalização o Teorema 4.4.7 é no sentido que as famílias de geradores considerados para $\mathcal{A}$ e $\mathcal{B}$ são infinitas.

Exemplo 4.4.8. Olhemos novamente para a situação no Exemplo 4.4.6, mas considerando agora no contexto de C^*. Neste caso, seja θ um número fixo em $[0, 2\pi]$. Suponha que $(\mathcal{A}, \varphi)$ é um C^*-espaço de probabilidade, em que a C^*-álgebra $\mathcal{A}$ é gerada pelos dois unitários u_1, u_2 que satisfazem a Equação (4.41), e onde $\varphi : \mathcal{A} \to \mathbb{C}$ é um funcional fiel e positivo satisfazendo a Equação (4.42). Então exatamente com no Exemplo 4.37, vemos que as relações (4.43) são válidas e implicam que

1. $\mathcal{A} = cl\ span\{u_1^m u_2^n \,|\, m, n \in \mathbb{Z}\}$, e
2. φ é um traço.

Agora, o Teorema 4.4.7 implica que um C^*-espaço de probabilidade $(\mathcal{A}, \varphi)$ como descrito no último *slide* é unicamente determinado a menos de isomorfismo. Em particular, a classe de isomorfismo da C^*-álgebra $\mathcal{A}$ envolvida no exemplo é unicamente determinada; por isso faz sentido (e isso é costumeiro) se referir a tal $\mathcal{A}$ denominado-a uma C^*-álgebra de rotação por θ.

É claro que a fim de falar sobre a C^*-álgebra ode rotação por θ deve-se também mostrar que ela existe - isto é, deve-se construir um exemplo de C^*-espaço de probabilidade $(\mathcal{A}, \varphi)$ em que φ é fiel e para qual as Equações (4.41) e (4.42) são satisfeitas.

Considere O espaço de Hilbert $\ell^2(\mathbb{Z}^2)$, e denote sua base ortonormal canônica por $\{\xi_{(m,n)} \,|\, m, n \in \mathbb{Z}\}$. É imediato que podemos definir dois operadores unitários U_1, U_2 sobre $\ell^2(\mathbb{Z}^2)$ descrevendo suas ações sobre a base ortonormal canônica como se segue:

$$\begin{cases} U_1(\xi_{(m,n)} & = \xi_{(m+1,n)} \\ U_2(\xi_{(m,n)} & = e^{-im\theta}\xi_{(m,n+1)} \end{cases} \quad m, n \in \mathbb{Z} \tag{4.45}$$

Seja $\mathcal{A}$ a C^*-subálgebra de $B(\ell^2(\mathbb{Z}^2))$ que é gerada por U_1 e U_2, e seja $\varphi : \mathcal{A}to\mathbb{C}$ o vetor-estado definido pelo vetor $xi_{(0,0)}$;

$$\varphi(T) = \langle T\xi_{(0,0)}, \xi_{(0,0)}\rangle, \quad T \in \mathcal{A}. \tag{4.46}$$

De (4.45) é imediato que $U_1U_2 = e^{i\theta}U_2U_1$ (de fato, ambos U_1U_2 e $e^{i\theta}U_2U_1$ levam $\xi_{(m,n)}$ para $e^{-im\theta}\xi_{(m+1,n+1)}$, par todos $m, n \in \mathbf{Z}$). Assim a fim de que o C^*-espaço de probabilidade $(\mathcal{A}, \varphi)$ tenha as propriedades requeridas, somente precisamos verificar que φ é fiel.

Observe que mesmo não sabendo que φ é fiel, podemos ver que é um traço. Isto é verificado exatamente como no Exemplo 4.4.6, em que a Equação (4.44) é agora substituída pelo fato que $\mathcal{A}$ é igual a $cl\, span\{U_1^m U_2^n \,|\, m, n \in \mathbf{Z}\}$; os detalhes desse fato não serão expostos aqui.

Exemplo 4.4.9. Agora suponha que $T \in \mathcal{A}$ é tal que $\varphi(T^*T) = 0$. Como $\varphi(T^*T) = \|T\xi_{(0,0)}\|^2$, temos assim que $T\xi_{(0,0)} = 0$. Mas então, para todo $m, n, p, q \in \mathbf{Z}$, podemos escrever:

$$\begin{aligned}
\langle T\xi_{(m,n)}, \xi_{(p,q)}\rangle &= \langle T(U_1^m U)2^n)\xi_{(0,0)}, (U_1^pU_2^q)\xi_{(0,0)}\rangle \\
&= \langle (U_1^pU_2^q)^*T(U_1^mU_2^n)\xi_{(0,0)}, \xi_{(0,0)}\rangle \\
&= \varphi\Big((U_1^pU_2^q)^*T(U_1^mU_2^n)\Big) \\
&= \varphi\Big((U_1^mU_2^n)(U_1^pU_2^q)^*T\Big) && \text{(pois } \varphi \text{ é traço)} \\
&= \langle (U_1^mU_2^n)(U_1^pU_2^q)^*T\xi_{(0,0)}, \xi_{(0,0)}\rangle \\
&= 0 && \text{(porque } T\xi_{(0,0)} = 0\text{)}.
\end{aligned}$$

Portanto $\langle T\xi_{(m,n)}, \xi_{(p,q)}\rangle = 0$ para todo $m, n, p, q \in \mathbf{Z}$, e está claro que implica que $T = 0$ (assim completando a verificação de que φ é fiel).

Sem entrar em detalhes, mencionamos aqui que as propriedades de universalidade e unicidade da C^*-álgebra $\mathcal{A}$ pela rotação de um ângulo θ pode ser obtida sem tomar o traço canônico $\varphi : \mathcal{A} \to \mathbb{C}$ como parte de nossos dados iniciais as (mas então os argumentos não estão muito longe daqueles apresentados acima).

4.5 DEFINIÇÃO E PROPRIEDADES DE INDEPENDÊNCIA LIVRE

4.5.1 A situação clássica: Independência tensorial

Definição 4.5.1. Seja $(\mathcal{A}, \varphi)$ um espaço de probabilidade não-comutativo e seja I um conjunto fixo de índices.

(1) Sub-álgebras unitais $(\mathcal{A}_i)_{i\in I}$ são ditas ser *tensor independentes* se as sub-álgebras $\mathcal{A}_i$ comutam (i.e. $ab = ba$ para todo $a \in \mathcal{A}_i$ e todo $b \in \mathcal{A}_j$ e todo $i, j \in I$ com $i \neq j$) e se φ é fatorado da seguinte maneira:

$$\varphi\left(\prod_{j\in J} a_j\right) = \prod_{j\in J} \varphi(a_j) \tag{4.47}$$

para todos os subconjuntos $J \subset I$ e todo $a_j \in \mathcal{A}_j$ $(j \in J)$.

(2) Independência tensorial (ou clássica) de variáveis aleatórias é definida pela independência tensorial das álgebras unitais geradas; portanto "a e b independentes tensoriais" significa nada mais que a e b comutam e têm os momentos mistos fatoráveis, ou seja,

$$ab = ba \quad \text{e} \quad \varphi(a^n b^m) = \varphi(a^n)\varphi(b^m) \ \ \forall n, m \geq 0. \tag{4.48}$$

Do ponto de vista combinatorial podemos considerar a independência tensorial como uma regra especial, a saber (4.48), para calcular momentos mistos de variáveis aleatórias independentes a partir dos momentos individuais de cada uma delas. Independência livre será apenas uma outra regra específica.

Note que no contexto não-comutativo, deve-se que especificar uma quantidade maior de momentos que no caso comutativo. Se a e b comutam, então todo momento misto em a e b podem ser reduzidos ao momento da forma $\varphi(a^n b^m)$ e, assim, a regra da fatoração em (4.48) para estas variáveis contém a completa informação sobre a distribuição conjunta de a e b, desde que sejam conhecidas as distribuições de a de b. Se, por outro lado, a e b não comutam, então $\varphi(a^n b^m)$ é somente uma pequena parte da distribuição conjunta de a e b, pois temos que considerar momentos tais com $\varphi(a^{n_1} b^{m_1} a^{n_2} b^{m_2} \cdots a^{n_k} b^{m_k})$, e estes não podem ser reduzidos em geral apenas para $\varphi(a^n b^m)$. Como uma primeira aproximação para uma regra de fatoração para situações não-comutativas, pode-se pensar numa extensão direta da regra clássica, a saber,

$$\varphi(a^{n_1} b^{m_1} \cdots a^{n_k} b^{m_k}) = \varphi(a^{n_1}) \cdot \varphi(b^{m_1}) \cdots \varphi(a^{n_k}) \cdot \varphi(b^{m_k}) \tag{4.49}$$

Isso, entretanto, não é a regra de independência livre. Pode-se ver facilmente que (4.49) não é consistente em geral se pusermos, por exemplo, algum dos m_i ou algum dos n_i igual a 0. Se estiremos dispostos a aceitar esta deficiência então a regra (4.49) pode ser usada para definir a assim chamada "independência booleana". Pode-se desenvolver elementos de uma teoria de probabilidade booleana, entretanto, sua estrutura é um pouco mais trivial quando comparada à profundidade da teoria da probabilidade livre. Para uma leitura mais aprofundada da fatoração booleana ver [Mishchenko et al., 2008].

Definição 4.5.2 (Independência Livre)**.** Seja $(\mathcal{A}, \varphi)$ um espaço de probabilidade não-comutativo e seja I um conjunto de índices fixo.

(1) Par cada $i \in I$, $\mathcal{A}_i \subset \mathcal{A}$, considere as sub-álgebras unitais $\mathcal{A}_i$. As sub-álgebras $(\mathcal{A}_i)_{i\in I}$ são ditas serem livremente independentes se

$$\varphi(a_1 \cdots a_k) = 0$$

sempre que se tenha o que se segue:

- k é um inteiro positivo;
- $a_{i(j)} \in \mathcal{A}_{i(j)}$ $(i(j) \in I)$ para todo $j = 1, \ldots, k$;
- $\varphi(a_j) = 0$ para todo $j = 1, \ldots, k$;
- e os elementos vizinhos vem de diferentes sub-álgebras, i.e. $i(1) \neq i(2)$, $i(2) \neq i(3), \ldots, i(k-1) \neq i(k)$.

(2) Sejam $\mathcal{X}_i \subset \mathcal{A}$ $(i \in I)$ subconjuntos de $\mathcal{A}$. Então $(\mathcal{X}_i)_{i\in I}$ são ditas livremente independentes se $(\mathcal{A}_i)_{i\in I}$ são livremente independentes, em que para $i \in I$, $\mathcal{A}_i := \text{alg}(1, \mathcal{X}_i)$ é a álgebra unital gerada por $\mathcal{X}_i$.

(3) Em particular, se as álgebras unitais $\mathcal{A}_i := (1, a_i)$ geradas pelos elementos $a_i \in \mathcal{A}_i$ $(i \in I)$ são livremente independentes, então $(a_i)_{i\in I}$ são chamadas *variáveis aleatórias livremente independentes*.

(4) Se no contexto de um $*$-espaço de probabilidade, as $*$-álgebras unitais $\mathcal{A}_i := \text{alg}(1, a_i, a_i^*)$ geradas pelas variáveis aleatórias a_i $(i \in I)$ são livremente independentes, então dizemos que $(a_i)_{i\in I}$ são$*$-*livremente independentes*.

Deve-se observar que o conceito de independência livre relaciona-se com o funcional linear φ. As variáveis aleatórias que são livremente independentes com respeito a algum funcional φ não são geralmente livremente independentes a outro funcional ψ. Ou seja, deve-se nomear esta propriedade como "livremente independente com respeito a φ".

Uma regra usual para que variáveis aleatórias sejam livremente independentes por ser explicitada por: $(a_i)_{i\in I}$ são livremente independentes se $\varphi(P_1(a_{i(1)})\dots P_k(a_{i(k)}) = 0$ para todos os polinômios $P_1,\dots,P_k \in \mathbb{C}\langle X\rangle$ em uma variável X e todos $i(1) \neq i(2) \neq \cdots \neq i(k)$, tais que $\varphi(P_j(a_{i(j)})) = 0$ para todo $j = 1,\dots,k$.

A independência livre é definida em termos as sub-álgebras geradas, mas este conceito também poderia se extender às C^*-álgebras geradas. Por simplicidade, diz-se que álgebras, conjuntos ou variáveis aleatórias são "livre" ou "$*$-livres" ao invés de "livremente independentes", ou $*$-livremente independentes"

4.5.2 Independência Livre e Momentos Conjuntos

Embora não seja tão óbvio quanto no caso de independência tensorial, a independência livre do ponto de vista combinatorial não é nada mais que uma regra especial de cálculo de momentos conjuntos de variáveis simples. O que explicita isto é a seguinte afirmação: *se uma família de variáveis aleatórias é livremente independente, então a distribuição conjunta da família é completamente determinada pelo conhecimento das distribuições individuais das variáveis.* Segue a seguinte Proposição.

Lema 4.5.1. *Seja (A,φ) um espaço de probabilidade não-comutativo e considere as sub-álgebras unitais A_i $(i \in I)$ sendo livres. Denote por $\mathcal{B}$ a álgebra que é gerada por todo A_i, $B := alg(A_i \,|\, i \in I)$. Então $\varphi_{|\mathcal{B}}$ é unicamente deterninada por $\varphi_{|A_i}$ para todo $i \in I$ e pela condição de independência livre.*

Prova: Ver [Nica and Speicher, 2006].

Exemplo 4.5.2. Seja $(\mathcal{A},\varphi)$ um espaço de probabilidade não-comutativo fixo e considere duas sub-álgebras livres $\tilde{\mathcal{A}}$ e $\tilde{\mathcal{B}}$. Para elementos $a, a_1, a_2 \in \tilde{\mathcal{A}}$ e $b, b_1, b_2 \in \tilde{\mathcal{B}}$, deseja-se calcular concretamente alguns momentos mistos de pequena ordem. O principal *truque* é se reduzir um momento misto geral para casos especiais, considerados na definição de independência livre, centralizando-se as variáveis envolvidas. Descrevendo em passos:

1. De acordo com a definição de independência livre, tem-se diretamente que $\varphi(ab) = 0$ se $\varphi(a) = 0$ e $\varphi(b) = 0$. Para se calcular $\varphi(ab)$ em geral, centraliza-se as variáveis como feito anteriormente:
$$\begin{aligned} 0 &= \varphi\Big((a-\varphi(a)1)(b-\varphi(b)1)\Big) \\ &= \varphi(ab) - \varphi(a_1)\varphi(b) - \varphi(a)\varphi(1b) + \varphi(a)\varphi(b)\varphi(1) \\ &= \varphi(ab) - \varphi(a)\varphi(b) \end{aligned}$$

o que implica que

$$\varphi(ab) = \varphi(a)\varphi(b) \qquad \text{se } a \text{ e } b \text{ são livres.} \tag{4.50}$$

2. Da mesma maneira, escreve-se

$$\varphi\Big((a_1 - \varphi(a_1)1)(b - \varphi(b)1)(a_2 - \varphi(a_2)1)\Big) = 0$$

implicando que

$$\varphi(a_1 b a_2) = \varphi(a_1 a_2)\varphi(b) \qquad \text{se } \{a_1, a_2\} \text{ e } b \text{ são livres.} \tag{4.51}$$

3. Os exemplos até agora renderam o mesmo resultado que teríamos para variáveis aleatórias tensor independentes. Para ver a diferença entre "independência livre" e "tensor independência", deve-se considerar agora $\varphi(a_1 b_1 a_2 b_2)$. Começando com

$$\varphi\Big((a_1 - \varphi(a_1)1)(b_1 - \varphi(b_1)1)(a_2 - \varphi(a_2)1)(b_2 - \varphi(b_2)1)\Big) = 0$$

obtém-se após alguns cálculos:

$$\begin{aligned}\varphi(a_1 b_1 a_2 b_2) = \varphi(a_1 a_2)\varphi(b_1)\varphi(b_2) + \varphi(a_1)\varphi(a_2)\varphi(b_1 b_2)\\ -\varphi(a_1)\varphi(b_1)\varphi(a_2)\varphi(b_2),\end{aligned} \tag{4.52}$$

se $\{a_1, a_2\}$ e $\{b_1, b_2\}$ são livres.

4.5.3 Algumas Propriedades Básicas de Independência Livre

Embora os exemplos anteriores sejam a "ponta do iceberg", eles permitem inferir algumas afirmações gerais sobre variáveis aleatórias livremente independentes. Em particular, pode-se ver que o conceito de independência livre é um conceito genuinamente não comutativo e somente alguns vestígios podem ser observados no "mundo comutativo".

Uma questão natural que surge é: quando variáveis aleatórias que comutam entre si são livremente independentes? Pode-se afirmar que isto ocorre se pelo menos uma delas tem variância nula, ou seja, se

$$\varphi\Big((a - \varphi(a)1)^2\Big) = 0 \quad \text{ou} \quad \varphi\Big((b - \varphi(b)1)^2\Big) = 0.$$

De fato, sejam a e b livres e tais que $ab = ba$. Então, pela combinação das Equações (4.50) (para a_2 e b_2 ao invés de a e b) e (4.52) (para o caso que $a_1 = a_2 = a$ e $b_1 = b_2 = b$),

tem-se

$$\begin{aligned}\varphi(a^2)\varphi(b^2) &= \varphi(a^2b^2)\\ &= \varphi(abab)\\ &= \varphi(a^2)\varphi(b)^2 + \varphi(a)^2\varphi(b^2) - \varphi(a)^2\varphi(b)^2,\end{aligned}$$

e, portanto,

$$\begin{aligned}0 &= \Big(\varphi(a2) - \varphi(a)^2\Big)\Big(\varphi(b2) - \varphi(b)2\Big)\\ &= \varphi\Big((a - \varphi(a)1)2\Big)\cdot\varphi\Big((b - \varphi(b)1)2\Big),\end{aligned}$$

o que implica que pelo menos um dos dois fatores tenha que ser nulo.

Em particular, se a e b são variáveis aleatórias clássicas, então elas podem somente ser livres se pelo menos uma delas e constante q.t.p.. Isto mostra que a independência livre é de fato um conceito não-comutativo e não pode ser considerado um caso especial de dependência entre variáveis aleatórias clássicas.

Um caso especial do que foi dito é se a é livremente independente de si própria, então $\varphi(a^2) = \varphi(a)^2$. Se estiver sendo considerado um $*$-espaço de probabilidade $(\mathcal{A}, \varphi)$ emm que φ é fiel, e se $a = a^*$, então isto implica que a é uma constante: $a = \varphi(a)1$. De outra forma, se as álgebras $\mathcal{A}_1$ e $\mathcal{A}_2$ são $*$-livres no $*$-espaço de probabilidade $(\mathcal{A}, \varphi)$ e se φ é fiel, então:

$$\mathcal{A}_1 \cap \mathcal{A}_2 = \mathbb{C}1.$$

Outra afirmação geral sobre variáveis aleatórias livremente geradas que pode ser inferidas diretamente a partir da definição é que variáveis aleatórias constantes são livres de qualquer outra. Segue o seguinte Lema.

Lema 4.5.3. *Seja $(\mathcal{A}, \varphi)$ um espaço de probabilidade não-comutativo e $\mathcal{B} \subset \mathcal{A}$ uma sub-álgebra unital. Então as sub-álgebras $\mathbb{C}1$ e $\mathcal{B}$ são livremente independentes.*

Prova: Ver [Nica and Speicher, 2006].

Na Proposição a seguir, será observado o fato de que independência livre tem um bom comportamento com respeito à propriedade tracial. Para demonstração deste resultado faz-se necessário o seguinte Lema:

Lema 4.5.4. *Seja $(\mathcal{A}, \varphi)$ um espaço de probabilidade não-comutativo, e seja $(\mathcal{A}_i)_{i\in I}$ uma família livremente independente de sub-álgebras unitais de $\mathcal{A}$. Sejam $a_1, \ldots, a_k$ elementos das álgebras $\mathcal{A}_{i(1)}, \ldots, A_{i(k)}$, respectivamente, para as quais os índices $i(1), \ldots, i(k) \in I$ são tais que $i(1) \neq i(2), \ldots, i(k-1) \neq i(k)$, e para as quais $\varphi(a1) = \cdots = \varphi(ak) = 0$. Da mesma forma, sejam $b_1, \ldots, b_\ell$ elementos de $A_{j(1)}, \ldots, A_{j(\ell)}$, respectivamente, tais que $j(1) \neq j(2), \ldots, j(\ell-1) \neq j(\ell)$, e tais que $\varphi(b_1) = \cdots = \varphi(b_\ell) = 0$. Então*

$$\varphi(a_1 \cdots a_k b_\ell \cdots b_1) = \begin{cases} \varphi(a_1 b_1) \cdots \varphi(a_k b_k), & \text{se } k = \ell, i(1) = j(1), ..., i(k) = j(k) \\ 0, & \text{caso contrário.} \end{cases} \tag{4·53}$$

Prova: Ver [Nica and Speicher, 2006].

Proposição 4.5.5. *Seja $(\mathcal{A}, \varphi)$ um espaço de probabilidade não-comutativo, seja $(\mathcal{A}_i)_{i\in I}$ uma família livremente independente de sub-álgebras unitais de $\mathcal{A}$, e seja $\mathcal{B}$ a sub-álgebra de $\mathcal{A}$ gerada por $\bigcup_{i\in I} \mathcal{A}_i$. Se $\varphi_{|\mathcal{A}_i}$ é um traço para todo $i \in I$, então $\varphi_{|\mathcal{B}}$ é um traço.*

Prova: Ver [Nica and Speicher, 2006].

4.6 MODELAGEM QUÂNTICA DE MODELOS CLÁSSICOS

Considere um dos mais simples modelos probabilísticos: o lançamento de uma moeda. Sobre tal experimento define-se naturalmente a variável aleatória (clássica) X pelo modelo de Bernoulli, cuja função de probabilidade é dado como

$$P(X = x) = \begin{cases} \frac{1}{2}, & \text{se } = +1, \text{ (vitória!)} \\ \\ \frac{1}{2}, & \text{se } = -1, \text{ (derrota!)} \end{cases} \tag{4·54}$$

Pode-se reescrever a Função (4·54) acima, a partir da medida (essencial),

$$\mu = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_{+1} \tag{4·55}$$

Sabe-se que a sequência de momentos é fundamental na caracterização de uma medida (e, consequentemente, da variável por ela modelada). Para o modelo ora em discussão, tem-se que os momentos de ordem m, denotados por $M_m(\mu)$, da medida μ é dado a partir

de (4.55) simplesmente por

$$M_m(\mu) = \int_{-\infty}^{+\infty} x^m \mu(dx) = \begin{cases} 1, & \text{se } m \text{ é par} \\ 0, & \text{se } m \text{ é ímpar.} \end{cases} \tag{4.56}$$

O chamado *problema de determinação de momentos* é aquele que a partir da sequência de momentos, deseja-se determinar a medida associada. No passado os quatro primeiros momentos eram usados para se caracterizar uma medida. Com o avanço dos modelos, percebeu-se que são necessários os momentos de todas as ordens para esta caracterização e mesmo assim uma tarefa árdua. Obviamente para a sequência $(0, 1, 0, 1, \ldots, 0, 1, \ldots)$ dada em 4.56 não é problema de se reconhecer como a medida de Bernoulli.

Considere, agora, os seguintes objetos:

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \qquad e_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad e_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{4.57}$$

Como visto neste capítulo, $\{e_0, e_1\}$ é considerado como uma base ortonormal de espaço bidimensional de Hilbert $\mathbb{C}^2$ e A é um operador hermitiano agindo sobre este espaço. Pode-se ver que

$$\langle e_1 \,|\, A^m e_1 \rangle == \begin{cases} 1, & \text{se } m \text{ é par} \\ 0, & \text{se } m \text{ é ímpar.} \end{cases} \tag{4.58}$$

que coincide com (4.56). Ou seja, o lançamento de uma moeda é também modelado usando o espaço de Hilbert C^2 e o operador A. No que foi descrito aqui, seja A a $*$-álgebra gerada por A, então o lançamento de moedas é modelado pela variável aleatória não-comutativa A em um $*$-espaço de probabilidade $(\mathcal{A}, e_1)$. Diz-se que A é uma realização algébrica da variável aleatória (clássica) X.

Neste Capítulo foi apresentada uma boa introdução à teoria da probabilidade não-comutativa. Apesar de que tais resultados estejam bem descritos na literatura, o capítulo serve de um guia para esta área. Como indicado nesta última seção, no Capítulo a seguir, será feita a descrição de um exemplo concreto de álgebra e de sua aplicação e as relações que podem ser obtidas com os resultados aqui apresentados.

CAPÍTULO 5

ALGORITMO DE BUSCA DE GROVER

Um *algoritmo* é um procedimento computacional bem definido que toma algum valor, ou conjunto de valores, como *entrada* e produz algum valor, ou conjunto de valores, como *saída*. Trata-se de uma sequência de passos computacionais que transformam uma entrada em uma saída. Um algoritmo também pode ser visto como uma ferramenta para resolver um *problema computacional* específico. A descrição do problema especifica em termos gerais a relação desejada entre entrada/saída. O algoritmo descreve um procedimento computacional específico para alcançar essa relação entre entrada/saída [Cormen et al., 1990].

Procurar um item em uma lista não ordenada de tamanho N, custa um tempo $O(N)$ para um computador clássico. Se N é muito grande, é como 'procurar uma agulha no palheiro'. Será que um computador quântico de busca conseguiria executar esta tarefa de forma mais eficiente do que um computador clássico? Em 1995, Grover respondeu à essa questão de maneira afirmativa, propondo um algoritmo de busca que consulta esta lista apenas $O(\sqrt{N})$ vezes [Grover, 1997]. Em contraste com algoritmos como o de fatoração quântica que fornece uma melhora ('aceleração') exponencial, o algoritmo de busca fornece apenas uma melhora quadrática. Porém, o algoritmo é bastante importante, pois tem uma ampla aplicação e porque a mesma técnica pode ser usada para acelerar algoritmos em problemas NP-completos [Varizani, 2013].

Pode-se imaginar se há algoritmos de busca ainda mais rápidos. De qualquer maneira, pode-se mostrar que uma melhora quadrática é ótima. Isto foi provado ainda antes do algoritmo de busca de Grover, em 1994. Qualquer algoritmo de busca deve consultar a lista, pelo menos, $\sqrt{N}$ vezes. Há duas maneiras de se pensar no algoritmo de busca de Grover, ambas serão discutidas em seguida.

5.1 BUSCA QUÂNTICA

5.1.1 Ideia do algoritmo

O algoritmo de busca de Grover se empenha em resolver o exato problema: é dada uma função booleana $f : \{1, \ldots, N\} \to \{0, 1\}$, e espera-se que, para exatamente um $a \in \{1, \ldots, N\}$, $f(a) = 1$. Logicamente a é o elemento pelo qual se procura.

A ideia básica do algoritmo de Grover é melhor descrita geometricamente. Como nossa função 'caixa preta' possui apenas duas saídas, pode-se identificar dois estados importantes: $|a\rangle$, o estado pelo qual se procura; e o restante, chamado de $|e\rangle = \sum_{x \neq a} \frac{1}{\sqrt{N-1}}|x\rangle$. Esses dois vetores geram um subespaço bidimensional, que contém a superposição uniforme $|\psi_0\rangle = \sum_x \frac{1}{\sqrt{N}}|x\rangle$. Além disso, $|a\rangle$ e $|e\rangle$ são ortogonais. Este espaço bidimensional pode ser representado geometricamente como pode ser representado na Figura 5.1.

Como $|\psi_0\rangle$ é constituído de $N - 1$ partes de $|e\rangle$ e apenas uma parte de $|a\rangle$, ele se encontra muito próximo de $|e\rangle$. O algoritmo de Grover começa trabalhando com o estado $|\psi_0\rangle$ e aumenta sucessivamente o ângulo entre ele e o vetor $|e\rangle$, para eventualmente chegar cada vez mais próximo de $|a\rangle$. Isto é feito por uma sequência de reflexões: primeiro uma reflexão sobre $|e\rangle$, e então uma reflexão sobre $|\psi_0\rangle$. O efeito líquido dessas duas reflexões é aumentar o ângulo entre o estado e $|e\rangle$. Repetindo este par de reflexões, o estado de se move para cada vez mais longe de $|e\rangle$, e então cada vez mais perto de $|a\rangle$. A primeira dessas reflexões pode ser vista na Figura 5.2. Uma vez que o vetor está perto o suficiente, a medição dos estados resulta na saída a com uma boa probabilidade.

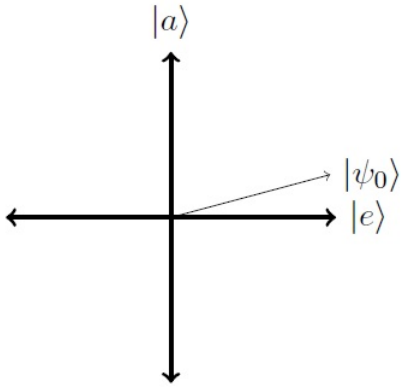

Figura 5.1 Espaço bidimensional gerado por $|a\rangle$ e $|e\rangle$ e a superposição uniforme $|\psi_0\rangle$.

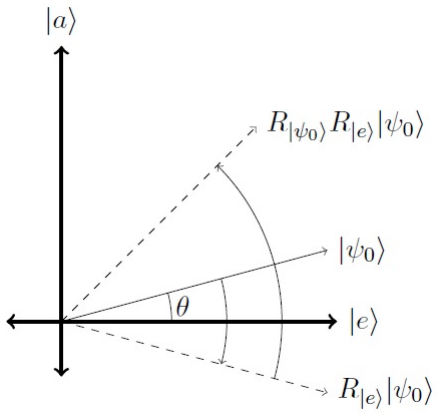

Figura 5.2 Primeira reflexão sobre $|e\rangle$, e então sobre $|\psi_0\rangle$.

Agora a questão é como executar exatamente estas duas reflexões.

5.2 O ORÁCULO QUÂNTICO

Da função booleana $f : \{1, \ldots, N\} \to \{0, 1\}$, pode-se construir um circuito quântico U_f para executar esta computação. Dado que sabemos que a função f pode ser calculada classicamente em um tempo polinomial, pode-se também computá-la em superposição:

$$\sum_x \alpha_x |x\rangle|0\rangle \to \sum_x \alpha_x |x\rangle|f(x)\rangle.$$

Há também uma maneira complicada de colocar este resultado numa forma que contenha igualmente toda informação relevante para o problema. Pode-se registrar a resposta na superposição $|-\rangle$, tal que quando f é implementada, a informação é armazenada na fase ou sinal do resultado:

$$\sum_x \alpha_x |x\rangle|-\rangle \to (-1)^{f(x)} \alpha_x |x\rangle|-\rangle.$$

De forma mais detalhada

$$\begin{aligned} U_f : \quad \sum_x \alpha_x |x\rangle \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) \quad \mapsto \quad & \sum_x \alpha_x \left(\frac{|x\rangle|f(x)\rangle - |x\rangle\overline{|f(x)\rangle}}{\sqrt{2}}\right) \\ & = \sum_x \alpha_x |x\rangle \left(\frac{|f(x)\rangle - \overline{|f(x)\rangle}}{\sqrt{2}}\right) \\ & = \sum_x \alpha_x |x\rangle (-1)^{f(x)} \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right). \end{aligned}$$

U_f tem a propriedade de que quando $x = a$, a fase do estado será multiplicada por -1. Será visto que esta implementação do circuito é equivalente à uma reflexão sobre o vetor $|e\rangle$.

5.3 ALGORITMO DE GROVER

O algoritmo de Grover encontra a em $O(\sqrt{N})$ passos. Como antes, considere o subespaço bidimensional gerado pelos dois estados $|a\rangle$ e $|e\rangle$, onde $|e\rangle$ é como citado anteriormente. Tome θ como o ângulo entre $|e\rangle$ e $|\psi_0\rangle = \sum_x \frac{1}{\sqrt{N}}|x\rangle$ (ver Figura 5.1) .

Dado que a é o alvo, tem-se que aumentar θ: isto é, rotacionar a entrada. Uma maneira de rotacionar um vetor é fazer duas reflexões. Em particular, pode-se rotacionar um vetor $|v\rangle$ de 2θ através da reflexão sobre $|e\rangle$ e depois sobre $|\psi_0\rangle$. Esta transformação está ilustrada na Figura 5.2.

A cada vez que estas reflexões são implementadas, ocorre uma rotação de 2θ. Isto significa que são necessárias $\frac{\pi}{2}/2\theta = \pi/\theta$ iterações para o algoritmo estar completo. Mas, qual é o valor de θ? Tem-se que

$$\langle\psi_0 \,|\, a\rangle = \cos(\pi/2 - \theta) = \text{sen}(\theta), \qquad \langle\psi_0 \,|\, a\rangle = \sum_x \frac{1}{\sqrt{N}}\langle x \,|\, a\rangle = \frac{1}{\sqrt{N}}$$

então,

$$\text{sen}(\theta) = \frac{1}{\sqrt{N}}.$$

Visto que $1/\sqrt{N}$ é muito pequeno, $\text{sen}\,\theta \approx \theta$, e $\theta \approx 1/\sqrt{N}$. Então, são necessárias $O(\sqrt{N})$ iterações para o algoritmo ficar completo. No final, chega-se bem próximo de $|a\rangle$, e então, com alta probabilidade, uma medida de estado fornece a.

Isso resolveria todo problema, exceto para o mecanismo exato de cada reflexão. Como a reflexão sobre $|\psi_0\rangle$ não depende do conhecimento de a, deveria-se ter a habilidade de construir isso puramente a partir de portas lógicas regulares e unitárias. A reflexão sobre $|e\rangle$ entretanto, requer o conhecimento de $|a\rangle$, então é necessário usar o oráculo U_f para construir esta reflexão. A utilização do oráculo resulta na facilidade de reflexão de $|e\rangle$. Tudo o que se precisa fazer é trocar a fase da componente na direção de $|a\rangle$. Já foi visto como se consegue fazer isso a partir de uma segunda implementação de f, como visto anteriormente.

Para a reflexão sobre $|\psi_0\rangle$, usa-se o operador de difusão D (assume-se $N = 2^n$), que opera como se segue. Primeiro, aplica-se H_{2^n}, que mapeia $|\psi_0\rangle \mapsto |00\ldots0\rangle$. Então, reflete-se sobre $|00\ldots0\rangle$ (isto é realizado pelo circuito U_g, onde g é uma função tal que $g(00\ldots0) = 0$ e $g(x) = 1$ para $x \neq 00\ldots0$). Finalmente, aplica-se H_{2^n} para retornar à base original. Note que esta é simplesmente uma reflexão sobre o vetor zero na base de Hadamard. Um melhor entendimento dessa operação pode ser obtido a partir de uma segunda descrição do algoritmo, dada abaixo.

5.4 UMA OUTRA APROXIMAÇÃO

Nesta seção será apresentado uma visão diferente sobre o algoritmo de busca, com todas as superposições.

Afirmação 1. Seja D o operador difusão como dado no último parágrafo na seção anterior. Então D tem as seguintes propriedades:

1. É unitário e pode ser eficientemente percebido.
2. Pode ser visto como 'uma inversão sobre a média'.

Prova. 1. Para $N = 2^n$, D pode ser decomposto e reescrito como:

$$D = H_N \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 \end{pmatrix} H_N = H_N \left(\begin{pmatrix} 2 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} - I \right) H_N$$

$$= H_N \begin{pmatrix} 2 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} H_N - I = \begin{pmatrix} 2/N & 2/N & \cdots & 2/N \\ 2/N & 2/N & \cdots & 2/N \\ \vdots & \vdots & \ddots & \vdots \\ 2/N & 2/N & \cdots & 2/N \end{pmatrix} - I$$

$$= \begin{pmatrix} 2/N - 1 & 2/N & \cdots & 2/N \\ 2/N & 2/N - 1 & \cdots & 2/N \\ \vdots & \vdots & \ddots & \vdots \\ 2/N & 2/N & \cdots & 2/N - 1 \end{pmatrix}$$

Observe que D é expresso como produto de três matrizes unitárias (duas matrizes de Hadamard separadas por uma matriz de mudança de fase condicional). Então,

D é também unitário. Com respeito à implementação, ambas as transformações, a de Hadamard e mudança de fase condicional, podem ser eficientemente realizadas com $O(n)$ portas lógicas.

2. Considere D operando em um vetor $|\alpha\rangle$ gerando outro vetor $|\beta\rangle$:

$$D \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_i \\ \vdots \\ \alpha_N \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_i \\ \vdots \\ \beta_N \end{pmatrix}$$

Tomando $\mu = N^{-1} \sum_j \alpha_j$ como a amplitude média, então a expressão $2\mu - \alpha_i$ descreve uma reflexão de α_i sobre a média. Isto pode ser mais facilmente visualizado escrevendo-se $\mu + (\mu - \alpha_i)$. Dessa forma, a amplitude de $\beta_i = -\frac{2}{N} \sum_j \alpha_j + \alpha_i = -2\mu + \alpha_i$ pode ser considerada uma "inversão sobre a média" com respeito à α_i. ∎

O algoritmo de busca quântico melhora iterativamente a probabilidade de medição de uma solução. Isto ocorre da seguinte maneira:

1. O estado inicial é $|\psi_0\rangle = \sum_x \frac{1}{\sqrt{N}} |x\rangle$.
2. Inverte-se a fase de a usando f.
3. Então faz-se uma inversão sobre a média usando D.
4. Repete-se os passos 2 e 3 $O(\sqrt{N})$ vezes, então em cada iteração, α_a cresce de $\frac{2}{\sqrt{N}}$.

Para saber o número de vezes que o algoritmo deve ser implementado utiliza-se a equação abaixo:

$$k = \frac{\arccos(1/\sqrt{N})}{\arccos\left(\frac{N-2}{N}\right)}. \tag{5.1}$$

Este processo é ilustrado nas Figuras 5.3 (a), (b) e (c) a seguir.

Note que em qualquer ponto do algoritmo, o estado pode ser descrito por dois números, a amplitude α_a de a, e a amplitude α' de qualquer $x \neq a$. Inicialmente $\alpha' = \alpha_x = 1/\sqrt{N}$.

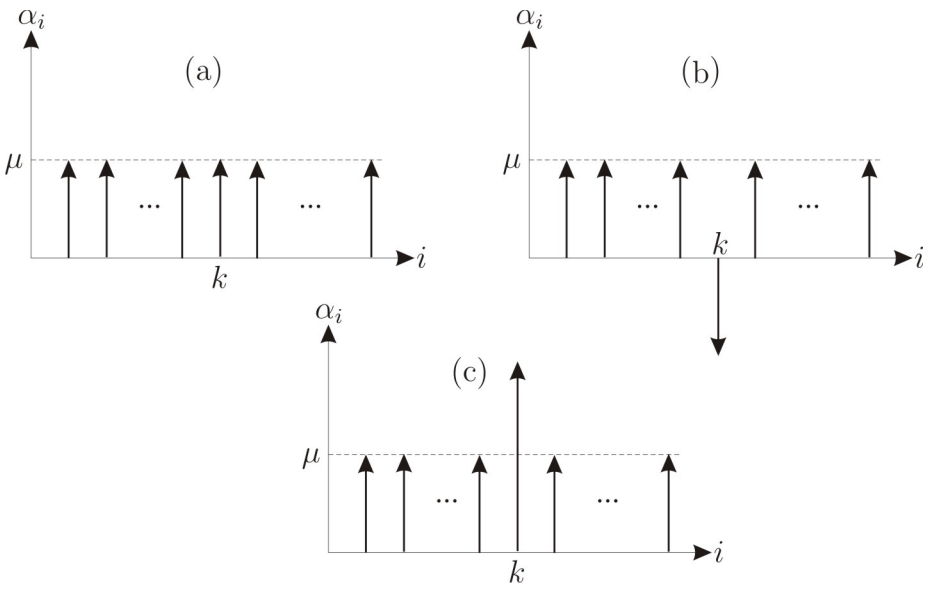

Figura 5.3 Os primeiros três passos do algoritmo de Grover. (a) Superposição uniforme de todos os vetores da base. (b) Uso da função f para inverter a fase de α_k. (c) Depois de executar o operador de difusão D, amplifica-se α_k enquanto diminui-se todas as outras amplitudes. Adaptado de [Varizani, 2013].

Quando o algoritmo é executado, α_a cresce e α' decresce. Suponha que queira-se achar a com probabilidade de apenas $\frac{1}{2}$. Então, tem-se que executar o algoritmo até $\alpha_a \approx 1/\sqrt{2}$. Neste ponto, $\alpha' \approx \frac{1}{\sqrt{2N}}$. Isto significa $\alpha' \geq \frac{1}{\sqrt{2N}}$ durante toda a execução do algoritmo. Visto que em cada iteração do algoritmo, α_a cresce de, pelo menos, $2\alpha'$ segue que o incremento é de, pelo menos, $\frac{2}{\sqrt{2N}} = \sqrt{\frac{2}{N}}$. Visto que o alvo é $\alpha_a = 1/\sqrt{2}$, o número de iterações $\leq 1/\sqrt{2}/\sqrt{\frac{2}{N}} = \sqrt{N}$.

5.5 PROBABILIDADE NÃO COMUTATIVA E O ALGORITMO DE GROVER

A teoria quântica é baseada em dois pilares: funções de onda (estado ou superposição quântica) e operadores. O estado de um sistema é representado por sua função de onda, e os observáveis são representados por operadores. Matematicamente, funções de onda satisfazem as condições definidas para vetores abstratos, e operadores agem neles como transformações lineares. Então, a linguagem natural da mecânica quântica é a álgebra linear.

Os vetores encontrados em mecânica quântica são, em sua maioria, funções que se

encontram em espaços de dimensão finita. A coleção de todas as funções de x, por exemplo, constituem um espaço vetorial. Contudo, a dimensão desse espaço é muito grande para certos propósitos. Para representar um possível estado físico, a função de onda ψ deve ser normalizada

$$\int |\psi|^2 = 1.$$

O conjunto de todas as funções quadrado integráveis em um intervalo específico

$$f(x) \text{ tal que } \int_a^b |f(x)|^2 dx < \infty$$

constituem um espaço vetorial mais facilmente tratável. Esse conjunto é chamado de $L^2(a, b)$ e é um espaço de Hilbert. Em mecânica quântica, as funções residem no espaço de Hilbert[Griffiths, 2005].

Então, usualmente, a mecânica quântica é definida em termos de operadores atuando no espaço de Hilbert. Os vetores de estado evoluem por meio de operadores unitários, os observáveis são medidos por operadores hermitianos, e os valores medidos têm distribuição de probabilidade [Kuperberg, 2005].

Os operadores utilizados na computação quântica (ou nos algoritmos quânticos) podem ser representados por matrizes com entradas complexas.

Foi visto no Exemplo 4.1.3 do capítulo 4, que a álgebra de matrizes complexas $d \times d$ com multiplicação usual de matrizes, juntamente com o traço normalizado, se encaixam na definição de um $*$-espaço de probabilidade. A sua $*$-operação é dada a partir da transposta da matriz e do complexo conjugado de suas entradas.

$$\mathrm{tr}(a) = \frac{1}{d} \cdot \sum_{i=1}^{d} \alpha_{ii} \qquad \text{para} \qquad a = (\alpha_{ij})_{i,j=1}^{d} \in M_d(\mathbb{C}).$$

Sendo assim, os operadores presentes no algoritmo de busca de Grover, estão presentes neste $*$-espaço de probabilidade e podem ser descritos pela probabilidade não comutativa. E à cada um desses operadores pode-se falar em uma $*$-distribuição.

Outro fato importante de se estudar a probabilidade não comutativa para algoritmos quânticos é o fato de que a diagonalização de matrizes com entradas complexas é sempre possível, pois sempre haverá raízes para os polinômios que surgem no cálculo da equação característica.

É importante lembrar da Definição 4.2.1 que a $*$-distribuição μ de um elemento a da álgebra é unicamente determinado resolvendo-se:

$$\int z^k \overline{z}^l d\mu(z) \;=\; \varphi(a^k(a^*)^l), \qquad \text{para todo } k\,, l \in \mathbb{N}.$$

Porém, vimos que dado que $a \in M_d(\mathbb{C})$, diagonalizando-se a e acha-se $\lambda_1, \ldots, \lambda_d$ como seus autovalores. Tem-se que

$$tr(a^k(a^*)^l) = \frac{1}{d}\sum_{i=1}^{d} \lambda_i^k \overline{\lambda}_i^l, \qquad k\,, l \in \mathbb{N}$$

Esta última quantidade pode ser escrita como $\int z^k \overline{z}^l d\mu(z)$, onde

$$\mu := \frac{1}{d}\sum_{i=1}^{d} \delta_{\lambda_i} \tag{5.2}$$

(o δ_λ funciona, aqui, como o Delta de Dirac (*Dirac mass*) em $\lambda \in \mathbb{C}$). Segue-se que a tem uma $*$-distribuição μ, que é descrita pela Equação (5.2). Usualmente μ é chamada de *distribuição de autovalores* da matriz a.

Como os operadores definidos no algoritmo de busca de Grover podem ser representados por matrizes, tem-se uma maneira direta para calcular as $*$-distribuições .

O algoritmo de Grover pode ser interpretado como aplicações sucessivas de operadores lineares. Uma responsável pela rotação de fase, usando o operador R e outra pela inversão sobre a média, usando o operador D. A composição desses dois operadores, dada pela multiplicação usual de matrizes, é chamada de operador de Grover, representado por $G = D \cdot R$. O operador R pode ser escrito como uma matriz diagonal preenchida com $1's$ e com um único valor -1 ocupando uma posição, tal que inverterá a fase do elemento procurado. Assim, $R_{ij} = 0$ se $i \neq j$, $R_{ii} = -1$ na i-ésima coluna referente ao elemento

procurado e $R_{ii} = 1$ para as demais colunas. A matriz de inversão sobre a média, também chamada de transformação de difusão D, é definida por $D_{ij} = \frac{2}{N}$ se $i \neq j$ e $D_{ii} = -1 + \frac{2}{N}$ [Grover, 1997].

De modo geral, o operador de Grover pode ser escrito matricialmente como

$$G = D \cdot R = \begin{bmatrix} \left(-1+\frac{2}{N}\right) R_{11} & \left(\frac{2}{N}\right) R_{22} & \cdots & \left(\frac{2}{N}\right) R_{NN} \\ \left(\frac{2}{N}\right) R_{11} & \left(-1+\frac{2}{N}\right) R_{22} & \cdots & \left(\frac{2}{N}\right) R_{NN} \\ \vdots & \vdots & \ddots & \vdots \\ \left(\frac{2}{N}\right) R_{11} & \left(\frac{2}{N}\right) R_{22} & \cdots & \left(-1+\frac{2}{N}\right) R_{NN} \end{bmatrix} \tag{5.3}$$

Onde, N é o número de elementos da lista, $R_{i_0 i_0} = -1$ quando o i_0-ésimo elemento da lista for o procurado e $R_{ii} = 1$, quando $i \neq i_0$.

Em posse do operador de Grover, os passos para se calcular as medidas associadas são:

i Obtenção dos autovalores e autovetores associados ao operador G;

ii normalização dos autovetores para garantir a unitariedade do operador de Grover;

iii escrever cada vetor de entrada, representando os elementos da lista, como combinação linear dos autovetores associados ao operador G;

iv Calcular a medida não comutativa, $*$-distribuição, associada a cada elemento da lista.

Por exemplo, um vetor genérico $|k\rangle$, representando o k-ésimo elemento da lista, pode ser escrito como:

$$|k\rangle = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_N v_N. \tag{5.4}$$

Onde $v_1, v_2, \ldots, v_N$ são os autovetores normalizados do operador G e $\alpha_1, \alpha_2, \cdots, \alpha_N$ são os projeções dos autovetores.

Calculando-se a medida associada à esse vetor, tem-se:

$$\begin{aligned}\mu(G_{|k\rangle}) &= \mu(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_N v_N) \\ &\leq |\alpha_1|\mu(v_1) + |\alpha_2|\mu(v_2) + \cdots + |\alpha_N|\mu(v_N)\end{aligned} \tag{5.5}$$

A medida $\mu(v_i)$ com $i = 1, \ldots, N$ é dada pela Equação 4.9:

$$\mu(v_i) = \frac{1}{N}\sum_{j=1}^{N} \delta_{\lambda_j}(v_i). \tag{5.6}$$

Substituindo 5.6 em 5.5, tem-se que a medida $\mu(G_{|k\rangle})$ será da forma:

$$\mu(G_{|k\rangle}) = \frac{1}{N}\sum_{i=1}^{N} |\alpha_i|. \tag{5.7}$$

Espera-se que para o elemento procurado $|i_0\rangle$ essa medida seja, de certa forma, diferenciada das medidas associadas aos demais elementos da lista.

Agora, pode-se observar os passos da aplicação do algoritmo de Grover e como obtém-se as medidas ($*$-distribuições) para alguns casos.

A princípio, será exemplificado a aplicação do algoritmo de Grover para uma lista de quatro elementos e posteriormente há a extensão dos resultados para uma lista de oito e dezesseis elementos a fim de calcular-se as probabilidades clássicas de se medir o elemento procurado e as medidas não comutativas, $*$-distribuições, associadas.

Para uma lista de 4 elementos tem-se $N = 4$, e sendo $N = 2^n$, precisa-se de dois qubits no primeiro registrador. De acordo com a Equação 5.1, tem-se $k = 1$, o algoritmo será implementado apenas uma vez. Os passos do algoritmo seriam, então,

(i) Inicializa-se o primeiro registrador com o estado:

$$|\psi\rangle = |\mathbf{0}\rangle = |00\rangle.$$

Em notação matricial:

$$|\psi\rangle = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

(ii) Aplicando $H^{\otimes n}$ gera-se uma superposição de estados.

$$\begin{aligned} |\psi_0\rangle &= H|0\rangle \otimes H|0\rangle \\ &= \left(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\right) \otimes \left(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\right) \\ &= \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle) \end{aligned}$$

ou em notação decimal,

$$|\psi_0\rangle = \frac{1}{2}(|0\rangle + |1\rangle + |2\rangle + |3\rangle)$$

Em notação matricial

$$|\psi_0\rangle = \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix}.$$

(iii) Suponha que a função f escolha o elemento $|11\rangle = |3\rangle$. Aplicando o operador inversão de fase $U_f(I \otimes H)$, passo $(iii - a)$, o único elemento a ter a sua fase invertida é o elemento $|11\rangle$. Assim, o novo estado passa a ser:

$$|\psi_1\rangle = \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle - |11\rangle).$$

Em notação matricial:

$$|\psi_1\rangle = \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ -1/2 \end{bmatrix}$$

A matriz associada ao operador inversão de fase, R, é dada por:

$$R = U_f(I \otimes H) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Aplicando-se o operador inversão sobre a média, também chamado de operador de difusão, $D = -I + 2A$, passo (iii - b), tem-se que o novo estado é:

$$|\psi_2\rangle = 1 \cdot |11\rangle.$$

Em notação matricial:

$$|\psi_2\rangle = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

A matriz associada ao operador inversão sobre a média, nesse caso, é dada por:

$$D = \begin{bmatrix} -1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 & 1/2 \\ 1/2 & 1/2 & 1/2 & -1/2 \end{bmatrix}$$

(iv) Medindo os qubits tem-se que o elemento procurado é encontrado com uma probabilidade de $(1)^2 = 100\%$. Para os outros elementos esta probabilidade se anula.

Pode-se sintetizar a aplicação do algoritmo de Grover em notação matricial como:

$$\underbrace{\begin{bmatrix} -1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 & 1/2 \\ 1/2 & 1/2 & 1/2 & -1/2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}}_{\text{Operador G}} \cdot \underbrace{\begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix}}_{\text{Superposição de estados}} = \underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}}_{\text{Medição dos qubits}}.$$

Quando elemento procurado é o $|11\rangle = |3\rangle$, o operador G é representado matricialmente por:

$$G = \begin{bmatrix} -1/2 & 1/2 & 1/2 & -1/2 \\ 1/2 & -1/2 & 1/2 & -1/2 \\ 1/2 & 1/2 & -1/2 & -1/2 \\ 1/2 & 1/2 & 1/2 & 1/2 \end{bmatrix}$$

As $*$-distribuições associadas ao operador G para cada elemento da lista, dadas pela Equação 5.7 foram calculadas e são mostradas na Tabela 5.1:

	Medidas
Estado $\|k\rangle$	Medida $\mu(G_{\|k\rangle})$
$\|0\rangle$	$0,439826$
$\|1\rangle$	$0,557678$
$\|2\rangle$	$0,557678$
$\|3\rangle$	$0,353553$

Tabela 5.1 Tabela mostrando as medidas associados a cada elemento da lista quando o elemento procurado é representado pelo vetor $|3\rangle$.

Os mesmos passos foram realizados para os casos em que os outros valores da lista eram procurados, estes resultados são mostrados na Tabela 5.2. Observa-se que o elemento procurado é o que possui a menor medida associada.

	Medidas		
Estados	$\mu(G_{\|0\rangle})$	$\mu(G_{\|1\rangle})$	$\mu(G_{\|2\rangle})$
$\|0\rangle$	$0,353553$	$0,439826$	$0,439826$
$\|1\rangle$	$0,439826$	$0,353553$	$0,557678$
$\|2\rangle$	$0,557678$	$0,557678$	$0,353553$
$\|3\rangle$	$0,557678$	$0,557678$	$0,557678$

Tabela 5.2 Tabela mostrando as medidas associados a cada elemento da lista para diferentes escolhas do elemento procurado.

No caso de uma lista com oito elementos, a Equação 5.1 fornece $k \cong 1,67$, assim, arredondando-se esse valor, o algoritmo deve ser implementado duas vezes. Suponha que o estado $|5\rangle$ seja o procurado. Aplicando-se o operador G uma vez, obtêm-se que o elemento procurado é medido com probabilidade de 78.12%. Na segunda aplicação, este valor cresce para 94.53%. Já as medidas, $*$-distribuições, mantém seus valores inalterados após a segunda aplicação, pois os autovetores não se alteram mesmo depois de uma nova aplicação do operador G. Na Tabela 5.3 observa-se as medidas associadas para a alguns elementos procurados. Note que os estados procurados continuam possuindo a menor medida.

	Medidas		
Estados	$\mu(G_{\|0\rangle})$	$\mu(G_{\|5\rangle})$	$\mu(G_{\|7\rangle})$
$\|0\rangle$	$0,176777$	$0,218338$	$0,218338$
$\|1\rangle$	$0,218338$	$0,344607$	$0,344607$
$\|2\rangle$	$0,344607$	$0,344607$	$0,344607$
$\|3\rangle$	$0,344607$	$0,344607$	$0,344607$
$\|4\rangle$	$0,344607$	$0,344607$	$0,344607$
$\|5\rangle$	$0,344607$	$0,176777$	$0,344607$
$\|6\rangle$	$0,344607$	$0,344607$	$0,344607$
$\|7\rangle$	$0,344607$	$0,344607$	$0,176777$

Tabela 5.3 Tabela mostrando as medidas associados a cada elemento da lista para diferentes escolhas do elemento procurado.

Para uma lista contendo dezesseis elementos a Equação 5.1 fornece $k \cong 2,61$, logo o algoritmo será implementado três vezes. Para um dado elemento a ser procurado na lista, este será medido com probabilidade $47,27\%$ na primeira aplicação do operador G, $90,84\%$ na segunda aplicação e $96,13\%$ na terceira. Novamente, as medidas, $*$-distribuições, não são alteradas pelas sucessivas aplicações do operador G e continuam apresentando menor valor para o dado elemento procurado. Na Tabela 5.4 é possível visualizar as medidas associadas para alguns elementos procurados.

A partir do exposto é possível enunciar duas proposições, a saber:

Proposição 5.5.1. *Dada uma lista não ordenada contendo N elementos, se o i_0-ésimo elemento é o elemento procurado, via alagoritmo de Grover, então ele possui a menor medida, ou $*$-distribuição.*

Proposição 5.5.2. *O número de implementações necessárias para se medir o elemento procurado no algoritmo de Grover não altera as medidas, ou $*$-distribuições, associadas aos elementos da lista.*

Com isso, pode-se notar que existe uma relação entre a implementação do algoritmo quântico de busca de Grover e as $*$-distribuições provenientes da probabilidade não comutativa. Uma relação matemática explícita entre a medida probabilidade de se medir um elemento procurado via algoritmo de Grover e as $*$-distribuições precisa ser investigada mais a fundo. A interpretação das $*$-distribuições, que possuem o menor valor para o elemento procurado, pode ser extraída da visão geométrica da aplicação do operador linear e unitário de Grover G, que faz com que o estado de superposição inicial se aproxime sucessivamente, à partir de reflexões, do estado procurado. Estas $*$-distribuições podem,

	Medidas		
Estados	$\mu(G_{\|0\rangle})$	$\mu(G_{\|5\rangle})$	$\mu(G_{\|15\rangle})$
$\|0\rangle$	$0,0883883$	$0,105318$	$0,105318$
$\|1\rangle$	$0,105318$	$0,181921$	$0,181921$
$\|2\rangle$	$0,181921$	$0,181921$	$0,181921$
$\|3\rangle$	$0,181921$	$0,181921$	$0,181921$
$\|4\rangle$	$0,181921$	$0,181921$	$0,181921$
$\|5\rangle$	$0,181921$	$0,0883883$	$0,181921$
$\|6\rangle$	$0,181921$	$0,181921$	$0,181921$
$\|7\rangle$	$0,181921$	$0,181921$	$0,181921$
$\|8\rangle$	$0,181921$	$0,181921$	$0,181921$
$\|9\rangle$	$0,181921$	$0,181921$	$0,181921$
$\|10\rangle$	$0,181921$	$0,181921$	$0,181921$
$\|11\rangle$	$0,181921$	$0,181921$	$0,181921$
$\|12\rangle$	$0,181921$	$0,181921$	$0,181921$
$\|13\rangle$	$0,181921$	$0,181921$	$0,181921$
$\|14\rangle$	$0,181921$	$0,181921$	$0,181921$
$\|15\rangle$	$0,181921$	$0,181921$	$0,0883883$

Tabela 5.4 Tabela mostrando as medidas associados a cada elemento da lista para diferentes escolhas do elemento procurado.

então, estar relacionadas à "distância" entre o vetor de superposição e o estado que se quer medir.

Uma investigação mais profunda da relação entre probabilidade não comutativa e resultados do algoritmo de Grover será objeto de trabalhos futuros.

5.6 A UTILIZAÇÃO DA $*$-DISTRIBUIÇÃO AOS OPERADORES QUÂNTICOS

Como dito na Seção 4.6 do capítulo anterior, a modelagem de experimentos aleatórios via $*$-espaços de probabilidade podem representar um importante fonte de recursos na formulação e demonstração de propriedades, tais como leis (quânticas) de grandes números, teoremas (quânticos) de limites centrais, etc ([Hora and Obata, 2007, Woronowicz, 1970, Shohat and Tamarkin, 1943, Accardi and Bożejko, 1998]).

Neste capítulo, apresentou-se os princípios da análise à luz da probabilidade não-comutativa de operadores usuais de algoritmos quânticos. A determinação de momentos e sua utilização em avaliações estatística (no sentido quântico, via teoria inferencial oriunda

das definições apresentadas no Capítulo 4), indicam um ótimo caminho de estudos sobre os algoritmos quânticos e sua utilização tanto em problemas clássicos, quanto em "não-clássicos".

A determinação de propriedades de algoritmos como sua complexidade, por exemplo, podem ser feitas a partir do estabelecimento mais concreto dos significados dos resultados apresentados neste trabalho.

CAPÍTULO 6

CONCLUSÃO

Neste trabalho, apresentou-se as bases matemáticas para a determinação de uma teoria probabilística não-comutativa, quântica, ou livre. Resultados em Análise Matemática, Álgebra de Operadores, Análise Funcional entre outros foram apresentados.

De modo especial, como um bom exemplo buscou-se descrever a probabilidade não comutativa a fim de se entender matematicamente as propriedades quânticas do Algoritmo de Grover. Tal descrição tinha como principal propósito calcular as $*$-distribuições dos operadores que atuam nesse algoritmo e indicar como pode-se estabelecer resultados a partir desta descrição.

Também, ficou indicada a utilização de propriedades de variáveis aleatórias não comutativas como passos para modelagem de experimentos clássicos. Esta utilização não foi inteiramente explicitada, mas concorda com os trabalhos apresentados na literatura.

Como proposta de trabalhos futuros fica a de se complementar as análises sobre os casos concretos, estabelecendo-se os passos da avaliação de propriedades, determinação de correlações e estudos inferenciais (à luz da teoria apresentada).

REFERÊNCIAS BIBLIOGRÁFICAS

[Accardi, 1991] Accardi, L. (1991). *Quantum probability and related topics*, volume 6. World Scientific.

[Accardi and Bożejko, 1998] Accardi, L. and Bożejko, M. (1998). Interacting fock spaces and gaussianization of probability measures. *Infinite dimensional analysis, quantum probability and related topics*, 1(04):663–670.

[Ávila, 1999] Ávila, G. (1999). *Introdução à Análise Matemática*. Editora Edgard Blücher Ltda., Rio de Janeiro, 2ª ed. edition.

[Barroso, 2014] Barroso, C. S. (2014). Análise funcional: Uma introdução. Escola de Matemática da América Latina e do Caribe. Universidade Federal do Ceará.

[Bierens, 2014] Bierens, H. J. (2014). Introduction to hilbert spaces. `http://grizzly.la.psu.edu/~hbierens/HILBERT.PDF`.

[Blyth and Robertson, 2006] Blyth, T. S. and Robertson, E. F. (2006). *Further Linear Algebra*. Undergraduate Mathematics Series, Springer, New York.

[Casella and Berger, 2002] Casella, G. and Berger, R. L. (2002). *Statistical Inference*. Duxbury Pacific Grove, CA, 2ª ed. edition.

[Cormen et al., 1990] Cormen, T. H., Leiserson, C. E., and Rivest, R. L. (1990). *Introduction to Algorithms*. Massachusetts Institute of Technology.

[DeGroot, 1989] DeGroot, M. H. (1989). *Probability and Statistics*. Addison-Wesley Publishing Company, New York, 2ª ed. edition.

[García et al., 2007] García, J. C., Quezada, R., and Sonts, S. B., editors (2007). *Proceedings of the 28th Conference on Quantum Probability and Related Topics*, Guanajuato, Mexico. CIMAT, Word Scientific.

[Griffiths, 2005] Griffiths, D. J. (2005). *Introduction to Quantum Mechanics*. Pearson Education International, 2ª edition.

[Grover, 1997] Grover, L. K. (1997). Quantum mechanics helps in searching for a needle in a haystack. *Physical Review Letters*, 79(2):4.

[Guerra, 2014] Guerra, M. (2014). Teoria da medida e integral de lebesgue; análise matemática iii. `http://pascal.iseg.utl.pt/~mguerra/am3/docs/lebesgue.pdf`.

[Hora and Obata, 2007] Hora, A. and Obata, N. (2007). *Quantum probability and spectral analysis of graphs*. Springer Science & Business Media.

[Kolmogorov and Fomin, 1960] Kolmogorov, A. N. and Fomin, S. V. (1960). *Measure, Lebesgue Integrals and Hilbert Space*. Academic Press - London, New York.

[Kuperberg, 2005] Kuperberg, G. (2005). A concise introduction to quantum probability, quantum mechanics, and quantum computation. UC Davis, visiting Cornell University. Disponível em: `https://www.math.ucdavis.edu/~greg/intro.pdf`.

[Leadbetter et al., 2014] Leadbetter, R., Cambanis, S., and Pipiras, V. (2014). *A Basic Course in Measure and Probability. Theory for Applications.* Cambridge University Press. New York.

[Lima, 2006] Lima, E. L. (2006). *Análise Real - Funções de uma variável*, volume 1. IMPA, Rio de Janeiro, 8ª ed. edition.

[Magalhães, 2006] Magalhães, M. N. (2006). *Probabilidade e Variáveis Aleatórias*. EDUSP, 2ª edition.

[Massen, 2004] Massen, H. (2004). Quantum probability, quantum information theory, quantum computing. `http://www.math.ru.nl/~massen/lectures/qpqiqc.pdf`.

[Mishchenko et al., 2008] Mishchenko, A., Brayton, R., and Chatterjee, S. (2008). Boolean factoring and decomposition of logic networks. In *Proceedings of the 2008 IEEE/ACM International Conference on Computer-Aided Design*, pages 38–44. IEEE Press.

[Nica and Speicher, 2006] Nica, A. and Speicher, R. (2006). *Lectures on the combinatorics of free probability*, volume 13. Cambridge University Press.

[Nielsen and Chuang, 2010] Nielsen, M. A. and Chuang, I. L. (2010). *Quantum Computation and Quantum Information*. Cambridge University Press.

[Nowosad, 1967] Nowosad, P. (1967). *Introdução à Análise Funcional*. 6º Colóquio Brasileiro de Matemática. SBM, Poços de Caldas.

[Rudin, 1976] Rudin, W. (1976). *Principles of Mathematical Analysis*. Third Edition. International series in pure and applied mathematics. McGraw-Hill.

[Shohat and Tamarkin, 1943] Shohat, J. A. and Tamarkin, J. D. (1943). *The problem of moments*. Number 1. American Mathematical Soc.

[Varizani, 2013] Varizani, U. (2013). *Quantum Mechanics and Quantum Computation*. Notas de aula do curso oferecido pela Berkeley University of California no site do edx: `http:https://www.edx.org`.

[Woronowicz, 1970] Woronowicz, S. (1970). The quantum problem of moments i. *Reports on Mathematical Physics*, 1(2):135–145.