



UFG

UNIVERSIDADE FEDERAL DE GOIÁS
INSTITUTO DE FÍSICA

Pós-graduação em Física

Dissertação de Mestrado

**Sobre os formalismos matemáticos da Mecânica
Quântica: Dirac, von Neumann e Álgebra C^***

Frederico Rodrigues Pfrimer

GOIÂNIA

30 de agosto de 2013

Sobre os formalismos matemáticos da Mecânica Quântica: Dirac, von Neumann e Álgebra C^*

Frederico Rodrigues Pfrimer

Dissertação de Mestrado apresentada ao programa de Pós-graduação em Física do Instituto de Física da Universidade Federal de Goiás como requisito parcial para obtenção do grau de Mestre em Física.

ORIENTADOR: *Prof. Dr. Ardiley Torres Avelar*

GOIÂNIA

2013

**Dados Internacionais de Catalogação na Publicação na (CIP)
(GPT/BC/UFG)**

P531s Pfrimer, Frederico Rodrigues.
Sobre os formalismos matemáticos da Mecânica Quântica: Dirac, von Neumann e Álgebra C* [manuscrito] / Frederico Rodrigues Pfrimer. - 2013.
x, 94 f.

Orientador: Prof. Dr. Ardiley Torres Avelar.
Dissertação (Mestrado) – Universidade Federal de Goiás, Instituto de Física, 2013.

Bibliografia.

1. Formalismo matemático – Mecânica quântica. 2. Álgebra C*. 3. Dirac, P. A. M., 1902-1984 – Formalismo. 4. Von Neumann, John – Formalismo.

CDU: 530.145

Aos meus pais Claudio R. da Silva e Simone A. Pfrimer e a todos os outros que me apoiaram ao longo dessa jornada de compreensão da natureza e de mim mesmo.

Agradecimentos

- Ao meu orientador prof. Dr. Ardiley Torres Avelar;
- Aos meus colegas do IF-UFG, pela amizade e pelo apoio;
- E principalmente a minha família, que estava comigo sempre, e nas horas mais difíceis.

Este trabalho foi financiado pelo CNPq.

Resumo

Os trabalhos originais de Dirac, “The Principles of Quantum Mechanics” e von Neumann, “The Mathematical Foundations of Quantum Mechanics”, estabeleceram as bases do formalismo matemático atual da mecânica quântica. Junto com a álgebra C^* , essas são três diferentes abordagens matemáticas e conceituais que podem ser utilizadas para a caracterização matemática da mecânica quântica. Fizemos uma releitura comparativa e mais matemática desses dois clássicos originais buscando clarificar suas semelhanças e diferenças, e em seguida expusemos de forma didática a álgebra C^* , apresentando todos os pré-requisitos de estruturas algébricas e outros conceitos necessários para a sua compreensão. O trabalho é ainda uma exposição mais didática dos formalismos de Dirac, von Neumann, e Álgebra C^* , permitindo uma compreensão mais rigorosa da matemática da mecânica Quântica.

Palavras-chave: formalismo matemático da mecânica quântica, formalismo de Dirac, formalismo de von Neumann, álgebra C^* , estruturas algébricas.

Sumário

| | | |
|----------|---|-----------|
| 1 | Introdução | 1 |
| 2 | O Formalismo de Dirac da Mecânica Quântica | 5 |
| 2.1 | Vetores Ket | 6 |
| 2.2 | Vetores Bra | 7 |
| 2.3 | Variáveis dinâmicas e observáveis | 11 |
| 2.4 | Relações Conjugadas | 16 |
| 2.5 | Autovalores e autovetores | 17 |
| 2.6 | Observáveis | 20 |
| 2.6.1 | Observáveis que Comutam | 21 |
| 2.7 | Funções de Observáveis | 22 |
| 2.8 | A “Função” Delta de Dirac | 24 |
| 2.9 | Representações | 26 |
| 2.9.1 | Conjunto Completo de Observáveis que comutam entre si | 28 |
| 2.10 | Propriedades dos Vetores Básicos | 29 |
| 2.11 | O Operador Identidade | 31 |
| 2.12 | A representação dos Operadores Lineares | 32 |
| 2.13 | Funções de Onda | 33 |
| 3 | O Formalismo de von Neumann da Mecânica Quântica | 35 |
| 3.1 | Limite e Continuidade | 39 |
| 3.2 | A Geometria do Espaço de Hilbert | 42 |
| 3.3 | Operadores Lineares | 46 |
| 3.4 | Continuidade de Operadores | 51 |
| 3.5 | Autovalores e Autovetores | 52 |
| 3.6 | Variedades Lineares e Projetores | 54 |
| 3.7 | O Traço | 57 |
| 4 | Estruturas Algébricas e Álgebra C^* | 60 |
| 4.1 | Noções Fundamentais | 61 |
| 4.2 | Estruturas Algébricas | 63 |
| 4.2.1 | Grupos, Semigrupos e Monoides | 64 |
| 4.2.2 | Anéis | 65 |

| | | |
|---------|---|----|
| 4.2.3 | Corpos | 66 |
| 4.2.4 | Espaços Vetoriais e Álgebras | 67 |
| 4.2.4.1 | Espaços Vetoriais | 67 |
| 4.2.4.2 | Álgebras | 68 |
| 4.2.4.3 | Álgebras-* | 69 |
| 4.3 | Espaços Métricos, Norma e Produto Interno | 70 |
| 4.3.1 | Métrica | 70 |
| 4.3.2 | Norma | 71 |
| 4.3.3 | Produto Interno | 72 |
| 4.3.4 | Convergência e Completeza em Espaços Métricos | 72 |
| 4.3.4.1 | Limite e Convergência | 73 |
| 4.3.4.2 | Sequência de Cauchy | 73 |
| 4.3.4.3 | Completeza | 74 |
| 4.3.5 | Espaços de Banach e de Hilbert | 74 |
| 4.3.5.1 | Espaços de Banach | 74 |
| 4.3.5.2 | Espaços de Hilbert | 75 |
| 4.4 | Álgebra C^* | 75 |
| 4.4.1 | Álgebra de Banach | 76 |
| 4.4.2 | Álgebra C^* | 77 |
| 4.4.3 | Espectro | 79 |
| 4.4.4 | Elementos Positivos | 81 |
| 4.4.5 | Estados | 83 |
| 5 | Conclusão | 87 |

Introdução

*To those who do not know mathematics it is difficult to get
across a real feeling as to the beauty,
the deepest beauty, of nature ...
If you want to learn about nature, to appreciate nature, it
is necessary to understand
the language that she speaks in.*

—RICHARD FEYNMAN

Desde o surgimento da mecânica quântica já passaram mais de 100 anos, e 80 anos da publicação do trabalho seminal de John von Neumann “*The Mathematical Foundations of Quantum Mechanics*” (1932) [1], o qual considera-se que estabeleceu com maior rigor as bases do formalismo matemático utilizado até os dias de hoje [2]. A teoria consiste atualmente na base de quase toda a física contemporânea, e seus resultados já foram verificados experimentalmente com uma precisão sem precedentes. Contudo, a teoria quântica é ainda fonte de muitas controvérsias e discussões no que tange a sua interpretação e seu significado. Existe um número de linhas de interpretações divergentes e pouco ou nenhum consenso nesse aspecto [3]. A tradicional interpretação de Copenhagen já não pode ser considerada a interpretação padrão olhando a nível global, e, recentemente, ainda mais visões diferentes estão surgindo derrubando qualquer ideia de um consenso.

Tudo isso ressalta a importância do formalismo matemático da teoria quântica, pois ele é o único aspecto em que existe um maior consenso. Mesmo havendo algumas divergências, o formalismo matemático é basicamente o que existe em comum entre as diversas interpretações. Apesar de existirem ‘interpretações’ que utilizam um formalismo um pouco diferente, como a mecânica boohmiana [4,5], o formalismo matemático é precisamente o que deve existir em comum entre duas interpretações diferentes da mesma teoria. Na verdade, a forma mais correta de nos expressarmos seria ‘interpretação de um formalismo matemático’, não de um teoria, sendo a teoria composta pelo formalismo mais a interpretação.

A chamada ‘velha mecânica quântica’ não era sequer uma teoria física propriamente dita, mas uma coleção de resultados e descrições heurísticas obtidos entre os anos de 1900 e 1925.

Nesse contexto, os resultados eram vistos como sendo correções à mecânica clássica, e tentava-se, sem muito sucesso, manter o formalismo matemático e a linguagem da mecânica clássica. A teoria era completamente fragmentada, tanto matematicamente, tanto nos princípios fundamentais. Nessa época, a teoria crescia sem muito rigor matemático, sendo a principal tarefa das teorias apenas descrever os fenômenos matematicamente. Muitos resultados importantes eram apenas leis matemáticas que conseguiam descrever resultados experimentais, como o espectro de emissão do átomo de hidrogênio, descrito pela equação:

$$E = -\frac{1}{2(k+l)^2}.$$

Um dos importantes princípios da velha teoria é a quantização da ação, segundo o qual, os sistemas obedecem as leis da mecânica clássica, mas apenas as trajetórias que obedecem

$$\oint p_i dq_i = n_i \hbar,$$

onde p_i é o momento e q_i a coordenada generalizada, são permitidas. Esse princípio no entanto é muito limitado pois só se aplica a sistemas em que o movimento é periódico, como é o caso da movimentação dos elétrons, não podendo, portanto, ser aplicado a muitos sistemas bem simples. Mais tarde, em 1924, de Broglie introduziu a ideia de ondas de matéria e dualidade onda-partícula, estendendo fenômenos já conhecidos dos fótons para toda as outras formas de matéria.

Até então não existia uma teoria unificada, apenas conjuntos de resultados, sem nenhum formalismo que conseguisse representar matematicamente os diferentes resultados. A mecânica matricial de Heisenberg-Born-Jordan [6–8] (1925) foi a primeira teoria a conseguir reproduzir alguns dos resultados conhecidos. Entretanto, na época havia pouco conhecimento sobre álgebra linear por parte dos físicos, não sendo a mesma popular como nos dias de hoje. O próprio Heisenberg não reconheceu de início que seu formalismo de índices tratava, na verdade, de matrizes. Isso fez com que a teoria perdesse em visibilidade, dando espaço à mecânica ondulatória de Schroedinger [9–11] que surgiria um pouco mais tarde ainda no mesmo ano.

A mecânica ondulatória deu origem ao conceito de função de onda e baseava-se na matemática de equações diferenciais. A evolução do sistema era descrita por uma equação diferencial parcial linear, a famosa equação de Schroedinger. Isso fez com que ela se tornasse muito mais popular, pois essa matemática já era conhecida pelos físicos, sendo portanto mais fácil de compreender e manipular. Pouco tempo depois o próprio Schroedinger mostrou [12] a equivalência dessas duas teorias.

Uma das primeiras questões quanto à interpretação surgiu na teoria de Schroedinger: qual era o significado do módulo ao quadrado da função de onda? Inicialmente pensava-se que era uma densidade de cargas, e depois foi surgindo uma interpretação probabilística dessa quantidade, o que daria origem à famosa interpretação de Copenhagen.

A verdadeira reconciliação entre os dois formalismos, mecânica matricial e mecânica ondulatória, é creditada a Dirac no seminal livro de 1930 “*Principles of Quantum Mechanics*” [13]. Nele fica claro que as duas formulações são apenas diferentes representações de uma teoria mais geral. Dirac utilizou espaços vetoriais e introduziu a notação de bras e kets, uma das principais e mais importantes ferramentas da área. Também considera-se [2] que o formalismo foi então consolidado no livro de von Neumann de 1932 “*Mathematical Foundations of Quantum Mechanics*” [1], o qual serve de base para as principais abordagens contemporâneas. Von Neumann trouxe uma formulação mais rigorosa, desenvolvendo de forma mais criteriosa o formalismo de operadores e aplicando o conceito de espaços de Hilbert.

Posteriormente, von Neumann publicou uma série de quatro artigos “On Rings of Operators” onde ele estuda especificamente a álgebra dos operadores, originando o que atualmente é conhecido como álgebra de von Neumann. Algumas dessas ideias foram generalizadas por Gelfand e Naimark em 1943 [14] originando a chamada álgebra C^* .

Os trabalhos de Dirac [13] e de von Neumann [1] foram os dois trabalhos originais que resultaram na formulação matemática da mecânica quântica atual. Eles servem de base para os diversos livros e publicações da mecânica quântica (MQ), e foram os trabalhos que deram origem a uma teoria unificada da MQ, pois até então não existia uma *única* teoria do ponto de vista matemático. Portanto, estudar e compreender diretamente esses trabalhos originais é de grande importância para uma compreensão mais profunda da MQ e de seu formalismo matemático. É beber diretamente a água da fonte, sem intermediários, e sentir o seu verdadeiro sabor. Esse estudo pode permitir uma percepção mais direta, e com menos influência, dos trabalhos originais desses dois grandes físicos e assim compreender melhor as raízes da MQ.

Pelo trabalho de Dirac podemos conhecer e compreender melhor sua notação, que é dominante na MQ moderna e em suas ramificações, como teoria da informação e da computação quântica; podemos até mesmo conhecer outros recursos interessantes mas que aparentemente caíram no esquecimento. Pelo trabalho de von Neumann podemos ter uma visão mais rigorosa do formalismo matemático da teoria quântica, compreender a sua noção de espaço de Hilbert e vantagens e desvantagens da sua formulação.

Em nosso trabalho fizemos uma releitura do formalismo matemático contido nesses dois livros seminais: “*Principles of Quantum Mechanics*” e “*Mathematical foundations of Quantum Mechanics*”, utilizando notações matemáticas mais modernas e formais, e buscando uma maior

clareza e facilidade de compreensão. O livro de von Neumann, em especial, devido à sua publicação antiga, sem os recursos tecnológicos atuais, é datilografado e com uma notação e aspecto visual muito pouco atrativos para o leitor contemporâneo, o que dificulta a leitura e compreensão do livro. Assim utilizamos a notação de kets no formalismo de von Neumann o que facilita a sua leitura e compreensão, e permite a comparação com as formulações da teoria que utilizam da notação de Dirac de bras e kets.

Apresentamos também as noções básicas e as principais estruturas algébricas de forma a mostrar o conceito moderno de espaços vetoriais e alguns outros conceitos utilizados no formalismo de Dirac e von Neumann, e com isso preparamos todo os pré-requisitos teóricos necessários para a compreensão do que é a álgebra C^* . Na sequência definimos a álgebra C^* e alguns dos seus principais resultados, buscando mostrar a sua correspondência com a álgebra dos operadores.

O Formalismo de Dirac da Mecânica Quântica

No seu famoso livro “The Principles of Quantum Mechanics” [13], Dirac propõe um formalismo matemático para a mecânica quântica voltado mais para a elegância e facilidade de uso, não frisando tanto a questão do rigor matemático, em oposição a von Neumann [1] que dá maior importância ao rigor matemático mas não alcança o mesmo nível de simplicidade e facilidade de uso. Dirac também é mais discursivo em seu tratamento matemático, no geral preferindo descrever mais com palavras do que com símbolos e equações matemáticas.

Em nossa abordagem buscamos retratar as visões de Dirac referentes ao formalismo matemático que estão essencialmente contidas nos três primeiros capítulos da 4ª edição de seu livro [13]. Essa edição tornou-se consagrada porque nela Dirac apresenta o formalismo usando a conhecida notação de bras e kets, notação que mnemonicamente unifica diversas noções permitindo um tratamento matemático mais elegante. Os três primeiros capítulos são intitulados:

1. O Princípio da Superposição
2. Variáveis dinâmicas e Observáveis
3. Representações

Fizemos uma releitura resumida desses capítulos com ênfase no formalismo matemático tornando-o um pouco mais simbólico e acrescentando alguns elementos de forma a clarificar as ideias contidas na formulação de Dirac e permitir uma comparação com outras abordagens. Justamente por isso buscamos ser o mais fiel possível às visões de Dirac e ressaltar desde já alguns pontos chaves para compreendermos as diferenças entre as diversas visões. Como a ênfase é na parte matemática, os aspectos referentes à interpretação do formalismo não são abordados.

2.1 Vetores Ket

Dirac começa a abordagem do formalismo matemático apresentando o princípio da superposição [13, p. 10]. Segundo ele, o princípio da superposição é uma espécie de processo aditivo implicando que os estados de um sistema físico podem de alguma forma ser somados resultando em um novo estado. Disso conclui-se que os estados devem ser representados por elementos matemáticos de um certo tipo que possam ser somados resultando em elementos matemáticos desse mesmo tipo. Por essa e outras motivações, a escolha natural é que esses elementos sejam vetores. Dirac afirma que vetores ordinários, existentes em um espaço de dimensão finita, não são suficientes para muitos dos sistemas da MQ, sendo necessária a generalização para espaços de dimensão infinita, o que gera algumas complicações no tratamento matemático, especialmente pelas questões de convergência. Ele, no entanto, evita um tratamento mais profundo dessas questões, e nesse ponto o trabalho de von Neumann se destaca por abordar precisamente essas questões.

Dirac chama os vetores que podem estar associados aos estados de um sistema físico de *kets* ou *vetores kets*, independente do espaço ser ou não de dimensão infinita. Nós associaremos a esses vetores o conjunto dos kets \mathcal{K} o que permitirá um tratamento mais matemático dos teoremas e definições. Um ket ou vetor ket também recebe uma notação especial $|\ \rangle$, sendo que um ket específico pode ser denotado inserindo-se um rótulo no seu interior, por exemplo, $|a\rangle$ ou $|\psi\rangle$.

Os kets podem ser somados ou ainda multiplicados por um número complexo resultando em um outro ket, ou seja:

Axioma 1.

$$\forall |a\rangle, |b\rangle \in \mathcal{K} : |a\rangle + |b\rangle \in \mathcal{K} \quad (2.1)$$

$$\forall |a\rangle \in \mathcal{K}, \alpha \in \mathbb{C} : \alpha |a\rangle \in \mathcal{K} \quad (2.2)$$

Podemos também somar uma sequência infinita de kets. Se tivermos, por exemplo, um conjunto de vetores ket $|x\rangle$ determinados pelo parâmetro real x que pode assumir todos os valores em um certo intervalo Ω , podemos integrá-lo na variável x em Ω obtendo um novo ket:

Axioma 2.

$$(\forall x \in \Omega : |x\rangle \in \mathcal{K}) \implies \int_{\Omega} |x\rangle dx \in \mathcal{K} \quad (2.3)$$

Como os kets são vetores, o conjunto dos kets \mathcal{K} satisfaz a estrutura algébrica de um espaço vetorial sobre os números complexos, ou seja, temos os seguintes axiomas:

Axioma 3. $\exists 0 \in \mathcal{K} : \forall |a\rangle, |b\rangle, |c\rangle \in \mathcal{K}, \alpha, \beta \in \mathbb{C} :$

$$|a\rangle + |b\rangle = |b\rangle + |a\rangle \quad (2.4)$$

$$|a\rangle + (|b\rangle + |c\rangle) = (|a\rangle + |b\rangle) + |c\rangle \quad (2.5)$$

$$|a\rangle + 0 = |a\rangle \quad (2.6)$$

$$\exists -|a\rangle \in \mathcal{K} : |a\rangle + (-|a\rangle) = 0 \quad (2.7)$$

$$\alpha(\beta|a\rangle) = (\alpha\beta)|a\rangle \quad (2.8)$$

$$\alpha(|a\rangle + |b\rangle) = \alpha|a\rangle + \alpha|b\rangle \quad (2.9)$$

$$(\alpha + \beta)|a\rangle = \alpha|a\rangle + \beta|a\rangle \quad (2.10)$$

$$1|a\rangle = |a\rangle \quad (2.11)$$

2.2 Vetores Bra

É possível associar um segundo conjunto de vetores, os vetores duais, aos vetores kets. Os vetores duais aos kets são o que Dirac denominou como *bras* ou *vetores bra*, para os quais utiliza-se a notação $\langle \ |$ de forma análoga ao que utilizamos para os kets. Denotaremos o conjunto dos bras por \mathcal{B} . Sempre que temos uma função linear que leva vetores ket a um número complexo, essa função pode ser vista como um produto escalar com algum vetor bra, ou seja, qualquer função f satisfazendo

$$\begin{aligned} \forall |a\rangle, |b\rangle \in \mathcal{K} : \quad f(|a\rangle + |b\rangle) &= f(|a\rangle) + f(|b\rangle) \\ f(\alpha|a\rangle) &= \alpha f(|a\rangle), \end{aligned}$$

é um funcional linear e pode ser escrita como o produto escalar de um bra com um ket.

De acordo com a notação de Dirac, podemos inferir que o produto escalar de um bra $\langle a|$ com um ket $|b\rangle$ é denotado pela justaposição dos símbolos do bra e do ket, ou seja, $\langle a||b\rangle$, e que podemos, nos casos em que não haja ambiguidade contrair as duas barras verticais em apenas uma, ou seja, $\langle a|b\rangle$, assim:

$$\langle a||b\rangle = \langle a|b\rangle, \quad (2.12)$$

ou seja, uma notação abreviada para o produto escalar. Dirac deixa isso um pouco que implícito, porém isso faz mais sentido sendo uma forma mais compreensiva de se entender a notação e facilita em casos especiais como por exemplo no produto escalar do bra nulo 0 com um ket $|a\rangle$, o qual é dado por $0|a\rangle$.

O produto escalar sempre leva um bra e um ket a um número complexo, o que pode ser

expresso simbolicamente pelo seguinte axioma:

Axioma 4.

$$\forall \langle a| \in \mathcal{B}, |b\rangle \in \mathcal{K} : \quad \langle a||b\rangle \in \mathbb{C} \quad (2.13)$$

A linearidade do produto escalar é expressa pelos axiomas:

Axioma 5.

$$\forall \langle a| \in \mathcal{B}, |b\rangle, |c\rangle \in \mathcal{K} : \quad \langle a|(|b\rangle + |c\rangle) = \langle a||b\rangle + \langle a||c\rangle \quad (2.14)$$

$$\forall \langle a| \in \mathcal{B}, |b\rangle \in \mathcal{K}, \alpha \in \mathbb{C} : \quad \langle a|(\alpha|b\rangle) = \alpha\langle a||b\rangle, \quad (2.15)$$

em que α é um número complexo.

Dirac postula que um bra está completamente especificado se o seu produto escalar com todos os kets for dado, o que pode ser expresso simbolicamente pelo seguinte axioma:

Axioma 6.

$$\forall \langle a|, \langle b| \in \mathcal{B} : \quad \langle a| = \langle b| \iff (\forall |c\rangle \in \mathcal{K} : \langle a|c\rangle = \langle b|c\rangle). \quad (2.16)$$

E ainda temos que um bra é nulo se, e somente se, o seu produto escalar com qualquer ket for nulo, simbolicamente a definição do bra nulo:

Axioma 7.

$$\forall \langle a| \in \mathcal{B} : \quad \langle a| = 0 \iff (\forall |b\rangle \in \mathcal{K} : \langle a|b\rangle = 0). \quad (2.17)$$

Dirac define a soma de dois bras $\langle a|$ e $\langle b|$ pela condição de que o seu produto escalar com qualquer ket $|c\rangle$ seja a soma dos produtos escalares de $\langle a|$ e $\langle b|$ com $|c\rangle$, ou seja:

Axioma 8.

$$\forall \langle a|, \langle b| \in \mathcal{B}, |c\rangle \in \mathcal{K} : \quad (\langle a| + \langle b|)|c\rangle = \langle a||c\rangle + \langle b||c\rangle; \quad (2.18)$$

e o produto de um vetor bra $\langle a|$ com um número α é definido pela condição de que o seu produto escalar com qualquer ket $|b\rangle$ é α vezes o produto escalar de $\langle a|$ com $|b\rangle$, ou seja,

Axioma 9.

$$\forall \langle a|, \langle b| \in \mathcal{B}, |c\rangle \in \mathcal{K}, \alpha \in \mathbb{C} : \quad (\alpha\langle a|)|b\rangle = \alpha(\langle a||b\rangle). \quad (2.19)$$

As Equações (2.14) e (2.18) mostram que os produtos de bras e kets satisfazem uma propriedade análoga à distributividade da multiplicação, e as Equações (2.19) e (2.15) mostram que a multiplicação por um número complexo satisfaz propriedades análogas à associatividade

e à comutatividade. Porém, note que, não foi definido uma operação de multiplicação, o que foi definido até o momento foi o produto escalar e a multiplicação de bras ou kets por escalares.

A beleza e simplicidade da notação de Dirac advêm do fato de que muitas de suas propriedades correspondem às propriedades usuais da adição e multiplicação. No entanto, é importante ressaltar que isso só ocorre devido à notação específica que foi utilizada, e que se modificarmos essa notação podemos perder essa correspondência com as propriedades da adição e multiplicação. Assim, por exemplo, se a notação para o produto escalar fosse $\langle a||b \rangle = \langle \langle a|, |b \rangle \rangle$, teríamos em vez da equação (2.14):

$$\langle \langle a|, |b \rangle + |b' \rangle \rangle = \langle \langle a|, |b \rangle \rangle + \langle \langle a|, |b' \rangle \rangle;$$

e não mais poderíamos dizer que essa propriedade corresponde à distributividade da multiplicação.

Com todos esses axiomas, demonstra-se que os bras formam um espaço vetorial sobre os números complexos, ou seja:

Teorema 1. $\exists 0 \in \mathcal{B}, \forall \langle a|, \langle b|, \langle c| \in \mathcal{B}, \alpha, \beta \in \mathbb{C} :$

$$\langle a| + \langle b| = \langle b| + \langle a| \quad (2.20)$$

$$\langle a| + (\langle b| + \langle c|) = (\langle a| + \langle b|) + \langle c| \quad (2.21)$$

$$\langle a| + 0 = \langle a| \quad (2.22)$$

$$\exists -\langle a| \in \mathcal{B} : \langle a| + (-\langle a|) = 0 \quad (2.23)$$

$$\alpha(\beta \langle a|) = (\alpha\beta) \langle a| \quad (2.24)$$

$$\alpha(\langle a| + \langle b|) = \alpha \langle a| + \alpha \langle b| \quad (2.25)$$

$$(\alpha + \beta) \langle a| = \alpha \langle a| + \beta \langle b| \quad (2.26)$$

$$1 \langle a| = \langle a|. \quad (2.27)$$

Prova 1. A demonstração de todas essas propriedades utiliza-se da mesma estrutura básica: aplicar em um ket genérico, aplicar os axiomas e depois fatorar de forma a obter um novo produto com um ket genérico, depois disso basta utilizar o axioma da igualdade de dois bras (Axioma 6) para obter o resultado desejado. Observe que sem esse axioma nenhuma propriedade poderia ser obtida.

$$(\langle a| + \langle b|) |\psi\rangle = \langle a||\psi\rangle + \langle b||\psi\rangle = \langle b||\psi\rangle + \langle a||\psi\rangle = (\langle b| + \langle a|) |\psi\rangle$$

$$\begin{aligned}
\langle (a + (b + c)) | \psi \rangle &= \langle a | \psi \rangle + \langle (b + c) | \psi \rangle \\
&= \langle a | \psi \rangle + (\langle b | \psi \rangle + \langle c | \psi \rangle) = (\langle a | \psi \rangle + \langle b | \psi \rangle) + \langle c | \psi \rangle \\
&= \langle (a + b) + c | \psi \rangle
\end{aligned}$$

$$\langle (a + 0) | \psi \rangle = \langle a | \psi \rangle + 0 | \psi \rangle = \langle a | \psi \rangle + 0 = \langle a | \psi \rangle$$

$$\langle (a + (-a)) | \psi \rangle = \langle a | \psi \rangle + (-a) | \psi \rangle = \langle a | \psi \rangle - \langle a | \psi \rangle = 0 = 0 | \psi \rangle$$

$$\langle (\alpha(\beta a)) | \psi \rangle = \alpha(\langle \beta a | \psi \rangle) = \alpha(\beta \langle a | \psi \rangle) = ((\alpha\beta) \langle a | \psi \rangle)$$

$$\langle \alpha(a + b) | \psi \rangle = \alpha(\langle (a + b) | \psi \rangle) = \alpha(\langle a | \psi \rangle + \langle b | \psi \rangle) = (\alpha \langle a | \psi \rangle) + (\alpha \langle b | \psi \rangle) = (\alpha \langle a + b | \psi \rangle)$$

$$\langle (\alpha + \beta) \langle a | \psi \rangle \rangle = (\alpha + \beta) \langle a | \psi \rangle = \alpha \langle a | \psi \rangle + \beta \langle a | \psi \rangle = (\alpha \langle a + \beta \langle a | \psi \rangle)$$

$$\langle (1 a) | \psi \rangle = 1 \langle a | \psi \rangle = \langle a | \psi \rangle$$

Até o momento não existe nenhuma conexão maior entre bras e kets fora a existência de um produto escalar. Dirac então postula que existe uma correspondência biunívoca entre bras e kets. Essa conexão é tal que é razoável chamar o bra correspondente ao ket de conjugado imaginário do ket. Dirac usa o termo ‘conjugado imaginário’ por que um ket ou um bra não pode ser separado em partes real e imaginária como os números complexos; como a adição de bras com kets não está definida, a forma usual de se obter as partes real e imaginária não pode ser aplicada.

Apesar de omitido por Dirac, não há problema em utilizarmos para o complexo imaginário a mesma notação utilizada para o complexo conjugado de um número, logo podemos enunciar os axiomas relativos à conexão entre bras e kets:

Axioma 10.

$$\forall |a\rangle, |b\rangle \in \mathcal{K} : \quad \overline{|a\rangle + |b\rangle} = \overline{|a\rangle} + \overline{|b\rangle}, \quad (2.28)$$

$$\forall |a\rangle \in \mathcal{K}, \alpha \in \mathbb{C} : \quad \overline{\alpha |a\rangle} = \bar{\alpha} \overline{|a\rangle}. \quad (2.29)$$

Em sua notação, Dirac também define a regra geral que o conjugado imaginário de um ket qualquer $|a\rangle$ é o bra $\langle a|$, e vice versa, ou seja,

Definição 1.

$$\forall |a\rangle \in \mathcal{K} : \quad \overline{\langle a|} = |a\rangle, \quad \overline{|a\rangle} = \langle a|. \quad (2.30)$$

Dados dois kets $\langle a|$ e $\langle b|$ é possível gerar o número $\langle a|b\rangle$ tomando o produto escalar do conjugado imaginário do segundo com o primeiro, o qual depende linearmente do primeiro e antilinearmente do segundo. Apesar de não mencionar explicitamente, Dirac está nesse momento definindo o produto interno de dois vetores kets:

Definição 2.

$$\forall |a\rangle, |b\rangle \in \mathcal{K} : \quad \langle |a\rangle, |b\rangle \rangle = \overline{|b\rangle} |a\rangle = \langle b|a\rangle. \quad (2.31)$$

Dirac então postula que

Axioma 11.

$$\forall |a\rangle \in \mathcal{K}, \langle b| \in \mathcal{B} : \quad \langle b||a\rangle = \overline{\langle a||b\rangle}, \quad (2.32)$$

onde ele está recuperando a propriedade do produto interno:

$$\forall |a\rangle, |b\rangle \in \mathcal{K} : \quad \langle |a\rangle, |b\rangle \rangle = \overline{\langle |b\rangle, |a\rangle \rangle}. \quad (2.33)$$

Também postula que:

Axioma 12.

$$\forall |a\rangle \in \mathcal{K} : \quad \langle a|a\rangle \geq 0, \quad \langle a|a\rangle = 0 \iff |a\rangle = 0, \quad (2.34)$$

o que recupera mais uma propriedade do produto interno:

$$\forall |a\rangle \in \mathcal{K} : \quad \langle |a\rangle, |a\rangle \rangle \in \mathbb{R} \geq 0, \quad \langle |a\rangle, |a\rangle \rangle = 0 \iff |a\rangle = 0 \quad (2.35)$$

Um bra e um ket são ditos ortogonais quando o seu produto escalar for nulo; dois bras ou dois kets são ortogonais quando o produto escalar de um com o conjugado imaginário do outro for nulo.

Dirac define ainda a norma ou comprimento de um bra $\langle a|$ ou de um ket $|a\rangle$ como sendo a raiz quadrada de $\langle a|a\rangle$, ou seja:

Definição 3.

$$\forall |a\rangle \in \mathcal{K} : \quad \| |a\rangle \| = \sqrt{\langle a|a\rangle}, \quad (2.36)$$

$$\forall \langle a| \in \mathcal{B} : \quad \| \langle a| \| = \sqrt{\langle a|a\rangle}. \quad (2.37)$$

2.3 Variáveis dinâmicas e observáveis

Dirac introduz o conceito de operador linear a partir da noção de uma função linear que leva um vetor ket em um outro vetor ket [13, p. 23]. Suponha que o ket $|F\rangle$ seja uma função linear de algum ket $|a\rangle$, então podemos dizer que a conexão entre esses dois kets é dada por

um operador linear, ou seja, que obtemos $|F\rangle$ aplicando um operador linear em $|a\rangle$. Seja X o símbolo do operador, então escrevemos:

$$|F\rangle = X|a\rangle.$$

Observe que, nessa notação, o operador X atuando no ket $|a\rangle$ é expresso pela simples justaposição dos símbolos com o operador sempre à esquerda do ket, o que soa como o simples produto do operador com o ket. Vamos denotar o conjunto \mathcal{L} como sendo o conjunto dos operadores lineares, assim, o fato de um operador linear ser uma função que leva um ket em outro ket se expressa pelo axioma:

Axioma 13.

$$\forall |a\rangle \in \mathcal{K}, \forall X \in \mathcal{L} : X|a\rangle \in \mathcal{K}, \quad (2.38)$$

e as propriedades de linearidade se expressam pelos seguintes axiomas:

Axioma 14.

$$\forall X \in \mathcal{L}, |a\rangle, |b\rangle \in \mathcal{K} : X(|a\rangle + |b\rangle) = X|a\rangle + X|b\rangle, \quad (2.39)$$

$$\forall X \in \mathcal{L}, |a\rangle \in \mathcal{K}, \alpha \in \mathbb{C} : X(\alpha|a\rangle) = \alpha(X|a\rangle). \quad (2.40)$$

Nessa notação, essas propriedades soam como a distributividade da multiplicação sobre a adição e a comutação entre operadores e números.

Considera-se que um operador linear está completamente definido quando o resultado da sua aplicação sobre cada ket for dada, ou seja, dois operadores lineares são iguais se, e somente se, os resultados de suas aplicações em todos os kets forem iguais, matematicamente:

Axioma 15.

$$\forall X, Y \in \mathcal{L} : X = Y \iff (\forall |a\rangle \in \mathcal{K} : X|a\rangle = Y|a\rangle). \quad (2.41)$$

E um operador é nulo se, e somente se, a sua aplicação em cada ket for nula, ou seja:

Axioma 16.

$$\forall X \in \mathcal{L} : X = 0 \iff (\forall |a\rangle \in \mathcal{K} : X|a\rangle = 0). \quad (2.42)$$

Assim como bras e kets, operadores lineares também podem ser somados resultando em novos operadores. A soma de dois operadores lineares é definida como sendo o operador linear que operando em algum ket produz a soma do que resultaria se atuasse em cada operador separadamente. Matematicamente isso nos leva ao seguinte axioma:

Axioma 17.

$$\forall X, Y \in \mathcal{L}, |a\rangle \in \mathcal{K} : (X + Y)|a\rangle = X|a\rangle + Y|a\rangle. \quad (2.43)$$

Agora temos a distributividade à esquerda do produto de operadores lineares com kets, e junto com a Equação (2.39) temos a propriedade de distributividade da multiplicação sobre a adição.

O produto de dois operadores lineares é definido como o operador linear cuja aplicação em um ket produz o mesmo resultado que a aplicação sucessiva dos dois operadores. Seja AB a notação para o produto dos operadores A e B , então temos o axioma:

Axioma 18.

$$\forall X, Y \in \mathcal{L}, |a\rangle \in \mathcal{K} : (XY)|a\rangle = X(Y|a\rangle). \quad (2.44)$$

Esse axioma soa como a propriedade associativa da multiplicação para as triplas $\mathcal{L} \times \mathcal{L} \times \mathcal{K}$, e assim podemos escrever esse produto omitindo os parênteses.

Aqui é importante ressaltar que a comutatividade da multiplicação não vale de forma geral para os operadores lineares. No caso especial em que dois operadores A e B satisfazem $AB = BA$, dizemos que A comuta com B , ou que A e B comutam.

Dirac então afirma que os operadores formam uma álgebra e que temos muitas das propriedades da álgebra dos números complexos, mas o axioma da comutatividade da multiplicação não vale. Mais precisamente ele quer dizer que os operadores lineares satisfazem a estrutura algébrica de um anel e de uma álgebra sobre os números complexos, ou seja, $\forall X, Y, Z \in \mathcal{L}, \alpha, \beta \in \mathbb{C} :$

$$X + Y = Y + Z \quad (2.45)$$

$$X + (Y + Z) = (X + Y) + Z \quad (2.46)$$

$$X + 0 = X \quad (2.47)$$

$$X + (-X) = 0 \quad (2.48)$$

$$\alpha(\beta)X = (\alpha\beta)X \quad (2.49)$$

$$\alpha(X + Y) = \alpha X + \alpha Y \quad (2.50)$$

$$(\alpha + \beta)X = \alpha X + \beta X \quad (2.51)$$

$$(\alpha X)(\beta Y) = (\alpha\beta)(XY) \quad (2.52)$$

$$1X = X \quad (2.53)$$

Nessa sequência, Dirac faz uma afirmação que pode gerar contradições se olhada com maiores critérios:

“Se pegarmos um número α e o multiplicarmos em um vetor ket, isso parece como um operador linear operando em vetores kets, as condições [de linearidade eq. (2.39) e (2.40)] sendo satisfeitas com X substituído por α . Um número é então um caso especial de operador linear. Ele tem a propriedade de que ele comuta com todos os operadores lineares e essa propriedade o distingue de um operador linear geral.” [13, p. 24]

Se considerarmos que Dirac está realmente dizendo que números são operadores lineares, então temos uma contradição com o axioma da igualdade de dois operadores lineares. A contradição é obtida quando nosso conjunto de operadores possui um operador identidade $\hat{1}$, nesse caso, para qualquer ket $|a\rangle \in \mathcal{K}$, temos que $\hat{1}|a\rangle = |a\rangle$ e, conseqüentemente, para qualquer número α , temos que $(\alpha\hat{1})|a\rangle = \alpha|a\rangle$ de onde infere-se, usando o referido axioma, que $\alpha\hat{1} = \alpha$, e ainda que $\hat{1} = 1$, ou seja, estamos afirmando que qualquer operador identidade é o próprio número 1, o que não pode ser necessariamente válido.

Aparentemente, é então necessário acrescentarmos um novo axioma para caracterizar o produto de um número por um operador de forma análoga ao que foi feito para o produto de dois operadores:

$$\forall X \in \mathcal{L}, |a\rangle \in \mathcal{K}, \alpha \in \mathbb{C} : (\alpha X)|a\rangle = \alpha(X|a\rangle).$$

Vamos agora definir a atuação de um operador linear sobre um bra. Considere o produto escalar de um bra $\langle a|$ com um ket $X|b\rangle$. Esse resultado, $\langle a|(X|b\rangle)$ depende linearmente de $|b\rangle$ e, conseqüentemente, da definição de bras, ele deve ser o produto escalar de algum bra com $|b\rangle$, ou seja, $\exists \langle c| : \langle c||b\rangle = \langle a|(X|b\rangle)$. Esse novo bra $\langle c|$ depende linearmente de $\langle a|$ de forma que podemos tratá-lo como o resultado de algum operador linear aplicado a $\langle a|$. Esse operador é determinado unicamente pelo operador linear original X , sendo razoável considerar que temos o mesmo operador atuando no bra.

Segundo Dirac, a notação adequada para expressar a atuação do operador no bra é precisamente aquela que torna válida a propriedade da associatividade da multiplicação para triplas de bras, operadores e kets, de onde resulta o seguinte axioma:

Axioma 19.

$$\forall \langle a| \in \mathcal{B}, X \in \mathcal{L}, |b\rangle \in \mathcal{K} : \langle a|(X|b\rangle) = (\langle a|X)|b\rangle. \quad (2.54)$$

Nessa notação, é necessário então que o operador esteja sempre à direita do bra. Note que, a escolha de uma notação diferente pode fazer com que percamos essa propriedade de associatividade, por exemplo: se a notação para o operador X atuando no bra $\langle a|$ fosse $X\langle a|$ ou $X(\langle a|)$ teríamos:

$$\langle a|(X|b\rangle) = X(\langle a||b\rangle) \quad \text{ou} \quad \langle a|(X|b\rangle) = (X\langle a||b\rangle),$$

e conseqüentemente perderíamos a propriedade de associatividade.

Por essa propriedade podemos simplificar a notação excluindo os parênteses:

$$\langle a|(X|b) = (\langle a|X)|b) = \langle a|X|b).$$

Dirac posteriormente busca dar sentido a um outro tipo de “produto”, o produto de um ket com um bra, como $|a\rangle\langle b|$. Note que Dirac não definiu a operação geral de multiplicação, apenas uma notação que se assemelha à multiplicação; ainda assim ele pressupõe que o axioma da associatividade da multiplicação vale para obter que $|a\rangle\langle b|$ é um operador linear. Logo Dirac está postulando que

Axioma 20.

$$\forall |a\rangle, |c\rangle \in \mathcal{K}, \langle b| \in \mathcal{B} : |a\rangle(\langle b|c\rangle) = (|a\rangle\langle b|)|c\rangle. \quad (2.55)$$

Desse postulado segue então que o produto $|a\rangle\langle b|\psi\rangle = \langle b|\psi\rangle|a\rangle$ é um ket que depende linearmente de $\langle\psi|$, e portanto $|a\rangle\langle b|$ é um operador linear:

$$\forall |a\rangle \in \mathcal{K}, \langle b| \in \mathcal{B} : |a\rangle\langle b| \in \mathcal{L}. \quad (2.56)$$

Nesse ponto temos uma importante afirmação de Dirac:

“Nós agora temos um completo esquema algébrico envolvendo três tipos de quantidades, vetores bra, vetores ket e operadores lineares. Eles podem ser multiplicados juntos nos diversos modos descritos até agora, e os axiomas da distributividade e associatividade da multiplicação sempre valem, mas o axioma da comutatividade não vale.” [13, p. 25]

Essa afirmação revela a beleza e também alguns pormenores da notação de Dirac. Em nenhum momento ele definiu uma operação geral de multiplicação, o que ele fez foi definir diversas operações e dar a mesma notação para todas elas; ele definiu a multiplicação por escalar, o produto escalar, a atuação de um operador em um ket, a atuação de um operador em um bra, e um ‘produto’ especial de um ket com um bra. No entanto, Dirac percebeu que essas operações satisfaziam propriedades semelhantes e que existia uma notação especial e que todas essas operações satisfazem propriedades correspondentes a uma multiplicação não comutativa, e de certa forma passa a considerar todas como uma forma de multiplicação.

2.4 Relações Conjugadas

Seja X um operador linear, então qual seria o bra correspondente ao ket $X|a\rangle$? Em outras palavras, qual o imaginário conjugado de $X|a\rangle$? Para responder a essa pergunta, Dirac introduz o conceito de adjunto de um operador linear ou complexo conjugado de um operador linear. O adjunto do operador X é denotado por \bar{X} , a mesma notação usada para o complexo conjugado de um número. Sendo assim, Dirac postula o seguinte:

Axioma 21.

$$\forall X \in \mathcal{L}, |a\rangle \in \mathcal{K} : \quad \overline{X|a\rangle} = \langle a| \bar{X}. \quad (2.57)$$

Usando que $\langle a|b\rangle = \overline{\langle b|a\rangle}$, obtemos a importante propriedade:

$$\forall \langle a| \in \mathcal{B}, X \in \mathcal{L}, |a\rangle \in \mathcal{K} : \quad \langle a| \bar{X} |b\rangle = \overline{\langle b|X|a\rangle}, \quad (2.58)$$

a qual será essencial para muitas demonstrações.

Substituindo X por \bar{X} na equação anterior obtemos que

$$\langle a| \overline{\bar{X}} |b\rangle = \overline{\langle b| \bar{X} |a\rangle} = \overline{\overline{\langle a|X|b\rangle}} = \langle a|X|b\rangle,$$

e como isso é válido para qualquer bra $\langle a|$ e qualquer ket $|b\rangle$, podemos afirmar que

$$\forall X \in \mathcal{L} : \quad \overline{\bar{X}} = X, \quad (2.59)$$

ou seja, que o adjunto do adjunto de um operador é o operador linear original, como no caso dos números complexos.

Quando um operador é igual ao seu adjunto dizemos que o mesmo é autoadjunto. Como no caso dos números complexos, um operador pode ser separado em partes real e imaginária, ou seja, qualquer operador X pode ser escrito como

$$X = A + iB, \quad (2.60)$$

em que A e B são dois operadores autoadjuntos e i é a unidade imaginária.

O adjunto ou complexo conjugado da soma de dois operadores pode ser obtido por

$$\langle a| \overline{(X + Y)} |b\rangle = \overline{\langle b|(X + Y)|a\rangle} = \overline{\langle b|X|a\rangle + \langle b|Y|a\rangle} = \langle a| \bar{X} |b\rangle + \langle a| \bar{Y} |b\rangle = \langle a| (\bar{X} + \bar{Y}) |b\rangle$$

Como a equação é válida para quaisquer $\langle a|$ e $|b\rangle$, obtemos que o adjunto da soma de dois

operadores é a soma dos adjuntos de cada operador, matematicamente:

$$\forall X, Y \in \mathcal{L} : \quad \overline{X + Y} = \overline{X} + \overline{Y}. \quad (2.61)$$

De forma semelhante obtêm-se o adjunto do produto de dois operadores lineares:

$$\langle a | \overline{(XY)} | b \rangle = \overline{\langle b | (XY) | a \rangle} = \overline{(\langle b | X)(Y | a \rangle)} = \overline{Y | a \rangle} \overline{\langle b | X} = \langle a | (\overline{Y} \overline{X}) | b \rangle,$$

como isso vale para qualquer bra $\langle a |$ e qualquer ket $| b \rangle$, podemos afirmar que

$$\forall X, Y \in \mathcal{L} : \quad \overline{XY} = \overline{Y} \overline{X}, \quad (2.62)$$

ou seja, o complexo conjugado do produto de dois operadores lineares é igual ao produto dos complexos conjugados dos fatores na ordem reversa.

Falta agora obter o complexo conjugado de operadores do tipo $| a \rangle \langle b |$. Fazendo

$$\langle a | \overline{(|\psi\rangle\langle\phi|)} | b \rangle = \overline{\langle b | (|\psi\rangle\langle\phi|) | a \rangle} = \overline{(\langle b | |\psi\rangle)(\langle\phi | | a \rangle)} = \overline{\langle a | |\phi\rangle} \overline{\langle\psi | | b \rangle} = \langle a | (|\phi\rangle\langle\psi|) | b \rangle,$$

obtemos de forma geral que

$$\forall | a \rangle \in \mathcal{K}, \langle a | \in \mathcal{B} : \quad \overline{| a \rangle \langle b |} = | b \rangle \langle a |. \quad (2.63)$$

Segundo Dirac é possível resumir todas as regras referentes ao complexo conjugado ou imaginário em uma única regra “o conjugado complexo ou conjugado imaginário de qualquer produto de vetores bras, vetores kets e operadores lineares é obtido tomando o complexo conjugado ou imaginário de cada fator e revertendo a ordem dos fatores.” Dirac afirma ainda que essa regra é bem geral, sendo válida mesmo para os casos não explicitados. De forma matemática, temos que a propriedade $\overline{xy} = \overline{y} \overline{x}$ é válida para qualquer par ordenado (x, y) pertencente a um dos conjuntos $\mathcal{B} \times \mathcal{K}, \mathcal{K} \times \mathcal{B}, \mathcal{L} \times \mathcal{L}, \mathcal{L} \times \mathcal{K}, \mathcal{B} \times \mathcal{L}$, simbolicamente:

$$\forall (x, y) \in \mathcal{B} \times \mathcal{K} \cup \mathcal{K} \times \mathcal{B} \cup \mathcal{L} \times \mathcal{L} \cup \mathcal{L} \times \mathcal{K} \cup \mathcal{B} \times \mathcal{L} : \quad \overline{xy} = \overline{y} \overline{x}. \quad (2.64)$$

2.5 Autovalores e autovetores

Equações do tipo

$$X | a \rangle = \alpha | a \rangle, \quad (2.65)$$

em que X é um operador linear, α é um número e $|a\rangle$ é um ket não nulo, ou ainda o conjugado dessa equação:

$$\langle b|X = \beta\langle b|, \quad (2.66)$$

em que β é um número e $\langle b|$ é um bra não nulo, são de grande importância para o desenvolvimento da teoria e são chamadas equações de autovalores. Quando essa primeira equação é satisfeita, dizemos que α é um autovalor do operador X e que $|a\rangle$ é um autoket ou um autovetor do operador X . Podemos ainda dizer que o autoket $|a\rangle$ pertence ao autovalor α . De forma análoga, β é um autovalor do operador X e $\langle b|$ é um autobra desse operador.

Agora restringiremos nosso estudo a operadores lineares reais, os quais são extensamente utilizados na teoria.

Seja um operador linear real X , um número x' e um ket $|a\rangle$ tais que

$$X|a\rangle = x'|a\rangle,$$

ou seja, x' é um autovetor e $|a\rangle$ é um autoket do operador X , então verifica-se que

1. Todos os autovalores são números reais:

$$\overline{x'} = x'.$$

2. Os autovalores associados com os autokets são os mesmos que os autovalores associados aos autobras:

$$X|a\rangle = x'|a\rangle \iff \langle a|X = x'\langle a|.$$

3. O conjugado imaginário de um autoket é um autobra associado ao mesmo autovalor:

$$\overline{X|a\rangle} = \overline{x'|a\rangle} = \langle a|X = x'\langle a|.$$

Dirac então define uma notação especial para os autovalores e autovetores de operadores associados a variáveis dinâmicas reais, pois os mesmos são amplamente utilizados na mecânica quântica. Seja X uma variável dinâmica (um operador) real, então x' , x'' , x''' , $x^{(n)}$ são autovalores diferentes associados ao operador X . E ainda, $|x'\rangle$ denota um autoket do operador X com autovalor x' , e analogamente para os outros autovalores. Simbolicamente temos:

$$X|x'\rangle = x'|x'\rangle, \quad X|x''\rangle = x''|x''\rangle, \quad X|x'''\rangle = x'''|x'''\rangle, \quad X|x^{(n)}\rangle = x^{(n)}|x^{(n)}\rangle. \quad (2.67)$$

Se estivermos lidando com diversos autokets pertencendo a um mesmo operador, pode-

mos distingui-los utilizando um sub-índice, por exemplo: $|x'_i\rangle$, assim temos:

$$X|x'_i\rangle = x^{(n)}|x'_i\rangle. \quad (2.68)$$

No caso, por exemplo, em que tivermos dois autokets para o autovalor x' , podemos diferencia-los por $|x'_1\rangle$ e $|x'_2\rangle$. Toda essa notação se estende de forma análoga para os bras, ou seja:

$$\langle x'|X = x'\langle x'|, \quad \langle x''|X = x''\langle x''|, \quad \langle x'''|X = x'''\langle x'''|, \quad \langle x^{(n)}|X = x^{(n)}\langle x^{(n)}|. \quad (2.69)$$

Dirac generaliza essa notação para autokets e autobras simultâneos de dois ou mais operadores. Um ket é um autoket simultâneo de dois operadores se for simultaneamente um autoket de cada um dos operadores. Um autoket simultâneo dos operadores X e Y com autovalores x' e y' respectivamente, é denotado por $|x'y'\rangle$, ou seja,

$$X|x'y'\rangle = x'|x'y'\rangle \quad (2.70)$$

$$Y|x'y'\rangle = y'|x'y'\rangle. \quad (2.71)$$

Essa notação se estende naturalmente para os casos mais gerais em que temos diferentes autovalores: $x'y', x''y', x'y'' \dots, x^{(n)}y^{(m)}$, e para os respectivos bras:

$$\langle x^{(n)}y^{(m)}|X = x^{(n)}\langle x^{(n)}y^{(m)}| \quad (2.72)$$

$$\langle x^{(n)}y^{(m)}|Y = y^{(m)}\langle x^{(n)}y^{(m)}|. \quad (2.73)$$

Dirac menciona dois importantes teoremas relativos a autovetores e autovalores:

Teorema 2. *Autovetores de um mesmo operador real pertencentes a autovalores diferentes são ortogonais:*

$$X|x'\rangle = x'|x'\rangle, \quad X|x''\rangle = x''|x''\rangle, \quad x' \neq x'' \implies \langle x'|x''\rangle = 0. \quad (2.74)$$

Prova 2.

$$\langle x'|X|x''\rangle = \langle x'|x'|x''\rangle = \langle x'|x''|x''\rangle \implies x'\langle x'|x''\rangle = x''\langle x'|x''\rangle \implies x' = x'' \text{ ou } \langle x'|x''\rangle = 0$$

Teorema 3. *Qualquer combinação linear de autovetores de um operador linear X pertencentes a um mesmo autovalor x' é um novo autovetor desse mesmo operador e pertencente ao*

mesmo autovalor:

$$X|x'\rangle = x'|x'\rangle \implies X\left(\sum_i \alpha|x'_i\rangle\right) = x'\left(\sum_i \alpha|x'_i\rangle\right). \quad (2.75)$$

2.6 Observáveis

Para definir matematicamente o conceito de observável, Dirac precisa acrescentar algumas interpretações físicas de forma a conectar a matemática com a natureza física [13, p. 34]. De acordo com Dirac, quando fazemos uma observação estamos medindo uma variável dinâmica, e o resultado deve necessariamente ser real. Disso conclui-se que as variáveis dinâmicas que podem ser medidas devem ser reais.

Para obter a definição de observável, Dirac precisa ir mais além e postular outras interpretações. Ele afirma que, se o sistema estiver em um autoestado de uma variável dinâmica X pertencente ao autovalor x' , então a medição de X vai com certeza resultar no número x' . E ainda, se o sistema está em um estado tal que o resultado da medição da variável dinâmica X é certo que irá resultar em determinado valor, ou seja, só há um resultado possível para essa medição, então o sistema está em um autoestado de X e o resultado da medição irá fornecer o autovalor de X ao qual o autoestado pertence.

Seguindo as suas considerações, Dirac afirma que quando medimos uma variável dinâmica X , a perturbação ocorrida no ato da medição faz com que o estado do sistema salte para um dos autoestados dessa variável, pois, devido à continuidade física, se fizermos uma segunda medição logo em seguida, devemos, com total certeza, obter o mesmo resultado. Assim, depois que a primeira medição foi realizada, não pode haver nenhuma incerteza quanto ao resultado da segunda.

Podemos então inferir que, não importa em qual estado o sistema físico esteja, qualquer resultado da medição de uma variável dinâmica real é um dos seus autovalores; e ainda, qualquer autovalor é um possível resultado de uma medição.

Dirac faz então uma afirmação ainda mais forte: se uma certa variável dinâmica real X for medida com o sistema em um estado particular, os estados aos quais o sistema pode saltar são tais que o estado original é linearmente dependente deles. Dirac então conclui que qualquer estado deve ser uma combinação linear dos autoestados da variável dinâmica.

Definição 4. *Define-se um conjunto completo de estados como sendo um conjunto tal que qualquer estado possível é uma combinação linear deles.*

Nem todas as variáveis dinâmicas reais possuem autoestados suficientes para formarem um conjunto completo, essas portanto não podem ser medidas. Assim sendo, Dirac define um Observável:

Definição 5. *Um observável é um operador linear real cujos autoestados formam um conjunto completo de estados.*

Os autovalores de uma variável dinâmica real podem consistir em um conjunto discreto de números, ou um contínuo, ou seja, todos os números em um certo intervalo, e de forma geral, os dois casos. A condição para que qualquer ket $|p\rangle$ seja uma combinação linear dos autovalores de um operador X , ou seja, para que os autokets do operador X formem um conjunto completo, se expressa como:

$$\forall |p\rangle \in \mathcal{K} : |p\rangle = \int |x'_c\rangle dx' + \sum_r |x'_d\rangle, \quad (2.76)$$

em que a integral é sobre toda a parte contínua do espectro e o somatório sobre a parte discreta. Quando essa condição é satisfeita, o operador X é um observável.

2.6.1 Observáveis que Comutam

Dirac prova que quando dois observáveis comutam, seus autoestados simultâneos são tantos que eles formam um conjunto completo. E ainda, quando os autoestados simultâneos de dois observáveis formam um conjunto completo, os dois observáveis comutam. Isso é formulado mais precisamente no seguinte teorema:

Teorema 4. *Se dois observáveis X e Y comutam, então os autoestados conjuntos dos dois observáveis formam um conjunto completo. Se os autoestados simultâneos de dois observáveis formam um conjunto completo, então os observáveis comutam.*

Matematicamente temos que, se

$$\begin{aligned} X|x'y'\rangle &= x'|x'y'\rangle & X|x^{(r)}y^{(s)}\rangle &= x^{(r)}|x^{(r)}y^{(s)}\rangle \\ Y|x'y'\rangle &= y'|x'y'\rangle & Y|x^{(r)}y^{(s)}\rangle &= y^{(s)}|x^{(r)}y^{(s)}\rangle, \end{aligned}$$

então:

$$XY - YX = 0 \iff \forall |p\rangle \in \mathcal{K} : |p\rangle = \iint |x'y'\rangle dx' dy' + \sum_{r,s} |x^{(r)}y^{(s)}\rangle. \quad (2.77)$$

Esse teorema se estende naturalmente para conjuntos maiores de observáveis em que

todos os observáveis comutam entre si.

2.7 Funções de Observáveis

Dirac considera que podemos tomar qualquer função f de um observável X e tratá-la como sendo um novo observável tal que é automaticamente medido quando X é medido. Para que isso possa ocorrer, é preciso que quando for certo que a medição de X resultará no valor x' , então deve ser certo que a medição da função desse observável, $f(X)$, resultará no valor $f(x')$. Assim, Dirac define de forma geral $f(X)$ como sendo o operador linear que satisfaz

$$f(X)|x'\rangle = f(x')|x'\rangle, \quad (2.78)$$

para qualquer autoket $|x'\rangle$ do operador X sendo que $f(x')$ deve ser um número para qualquer autovalor x' .

Quando $f(X)$ é expressível como uma série de potências verifica-se que essa condição é imediatamente satisfeita, ou seja, se

$$f(X) = c_0 + c_1X + c_2X^2 + \dots, \quad (2.79)$$

em que as constantes são todas números, então verifica-se que

$$f(X)|x'\rangle = f(x')|x'\rangle.$$

O conjugado complexo de $f(X)$ é definido pelo conjugado imaginário da equação que define esse operador, ou seja,

$$\langle x' | \overline{f(X)} = \overline{f(x')} \langle x' |, \quad (2.80)$$

sendo válido para qualquer autobra $\langle x'|$.

Demonstra-se então que o complexo conjugado do operador linear $f(X)$ é o complexo conjugado da função f aplicado ao operador X , ou seja,

$$\overline{f(X)} = \overline{f}(X). \quad (2.81)$$

Segue automaticamente que, se a função f for real, então $f(X)$ será também um observável, pois é agora um operador real e seus autoestados formam um conjunto completo.

Como existem autoestados simultâneos de dois ou mais observáveis que comutam entre

si, podemos definir operadores lineares que são funções simultâneas de dois ou mais observáveis. Essa generalização corresponde ao conceito de funções de múltiplas variáveis aplicadas a operadores. Sejam X, Y, Z, \dots observáveis que comutam entre si, então define-se a função desses operadores $f(X, Y, Z, \dots)$ como sendo o operador linear que satisfaz:

$$f(X, Y, Z, \dots)|x'y'z' \dots\rangle = f(x', y', z', \dots)|x'y'z' \dots\rangle, \quad (2.82)$$

em que $|x'y'z' \dots\rangle$ é qualquer autoket simultâneo dos observáveis X, Y, Z, \dots .

De forma análoga ao que ocorre para funções de uma única variável, demonstra-se que

$$\overline{f(X, Y, Z, \dots)} = \bar{f}(X, Y, Z, \dots), \quad (2.83)$$

e que, se f for uma função real, então $f(X, Y, Z, \dots)$ é um operador linear real e também um observável. Essa definição de função de observável tem duas aplicações práticas de especial importância: o recíproco ou inverso de um observável e a raiz quadrada de um observável.

O recíproco existe se o observável não possuir autovalores nulos. Seja X um observável, então o seu recíproco X^{-1} satisfaz:

$$X^{-1}|x'\rangle = x'^{-1}|x'\rangle, \quad (2.84)$$

para qualquer autoket $|x'\rangle$ do operador X com autovalor x' . Como

$$X^{-1}X|x'\rangle = XX^{-1}|x'\rangle = |x'\rangle, \quad (2.85)$$

vale para qualquer autoket $|x'\rangle$ e o operador X é um observável, devemos ter que

$$X^{-1}X = 1 \quad (2.86)$$

$$XX^{-1} = 1. \quad (2.87)$$

Note que Dirac nesse ponto não é muito rigoroso pois trata o que deveria ser um operador identidade como o próprio número 1. Dirac não menciona até então o operador identidade e seu tratamento simplificado pode resultar em inconsistências em alguns casos. Um dos problemas é que o número 1 não é propriamente um operador linear não podendo ser escrito em termos dos autoestados de um observável.

Dirac então generaliza essa definição de inverso para operadores lineares gerais, e como queremos ser mais criteriosos nesse momento, é necessário definirmos o que seria o operador identidade. O operador identidade para um dado conjunto de operadores lineares e kets, se

existir, é definido como sendo o operador linear $\hat{\mathbf{1}}$ tal que:

$$\forall |a\rangle \in \mathcal{K} : \quad \hat{\mathbf{1}}|a\rangle = |a\rangle, \quad (2.88)$$

o que leva a:

$$\forall X \in \mathcal{L} : \quad \hat{\mathbf{1}}X = X\hat{\mathbf{1}} = X. \quad (2.89)$$

Uma vez definido o operador identidade, podemos definir o inverso ou recíproco de um operador de forma mais rigorosa. O operador linear X^{-1} é o inverso do operador linear X se, e somente se,

$$X^{-1}X = \hat{\mathbf{1}}, \quad (2.90)$$

$$XX^{-1} = \hat{\mathbf{1}}. \quad (2.91)$$

Uma importante propriedade é que o inverso do produto de operadores é o produto dos inversos dos operadores com a ordem invertida, ou seja:

$$\forall X, Y \in \mathcal{L} : \quad (XY)^{-1} = Y^{-1}X^{-1}, \quad (2.92)$$

$$\forall X, Y, Z \in \mathcal{L} : \quad (XYZ)^{-1} = Z^{-1}Y^{-1}X^{-1}, \quad (2.93)$$

⋮

2.8 A “Função” Delta de Dirac

No formalismo proposto por Dirac, a atualmente bem conhecida função Delta de Dirac δ desempenha uma tarefa de extrema importância para o desenvolvimento da teoria no caso de autovalores contínuos. Von Neumann critica fortemente esse artifício matemático e abomina o seu uso buscando propor um formalismo alternativo em que o mesmo não apareça. Ele o consegue, mas no entanto seu formalismo é limitado e não abrange os casos em que temos espectros contínuos.

Para verificarmos a necessidade desse artifício matemático (dizemos artifício porque ela não é rigorosamente uma função nas definições clássicas, mas uma distribuição), suponhamos que X seja um observável de espectro contínuo, então os kets $|a\rangle$ e $|b\rangle$ podem ser escritos como:

$$|a\rangle = \int |x'_\alpha\rangle dx'$$

$$|b\rangle = \int |x''_\beta\rangle dx'',$$

em que α e β são sub-índices utilizados para distinguir os dois integrandos. Tomando o produto interno dos dois kets temos

$$\langle a|b\rangle = \iint \langle x'_\alpha|x''_\beta\rangle dx' dx''.$$

Pelo teorema da ortogonalidade temos que

$$x' \neq x'' \implies \langle x'_\alpha|x''_\beta\rangle = 0,$$

então, considerando apenas a integral interna

$$\int \langle x'_\alpha|x''_\beta\rangle dx',$$

verifica-se que o integrando é nulo em todo o domínio de integração, exceto no ponto em que $x' = x''$. Se o integrando for finito nesse ponto, então a integral inteira iria convergir para zero e $\langle a|b\rangle = 0$. Para que o resultado da integral seja finito, o integrando então deve ser infinito, e fazendo $|a\rangle = |b\rangle \neq 0$, verifica-se que de forma geral $\langle x'|x'\rangle$ é infinito.

Tendo isso em vista, Dirac define a chamada delta de Dirac como uma quantidade $\delta(x)$ dependendo do parâmetro real x satisfazendo os seguintes axiomas:

$$\int_{-\infty}^{\infty} \delta(x) dx = 1 \quad (2.94)$$

$$x \neq 0 \implies \delta(x) = 0 \quad (2.95)$$

Para tornar a Delta um pouco mais compreensível e aceitável, Dirac afirma [13, p. 58] que ela pode ser vista como o limite de uma sequência de funções que se anulam em todos os pontos exceto em uma pequena região de largura a e é tão grande nessa pequena região que a integral inteira tem valor 1. No limite de a tendendo a zero, essa função converge para a Delta de Dirac. A forma exata da função não importa, desde que ela seja bem comportada. A Delta não é uma função de acordo com as definições matemáticas usuais porque não possui um valor definido para cada ponto em seu domínio, ela diverge no ponto $x = 0$. Dirac afirma que o seu uso deve ser limitado a certas expressões matemáticas simples em que nenhuma inconsistência pode surgir, de forma geral ela só pode ser usada dentro de integrais.

A mais importante propriedade dessa função é a propriedade de filtragem expressa na seguinte equação:

$$\int_{-\infty}^{\infty} f(x)\delta(x)dx = f(0), \quad (2.96)$$

em que $f(x)$ é qualquer função contínua de x .

Modificando a origem da equação anterior obtemos a formulação mais prática dessa propriedade:

$$\int_{-\infty}^{\infty} f(x)\delta(x-a)dx = f(a), \quad (2.97)$$

em que a é um número real qualquer.

Enumeraremos agora algumas importantes equações relativas à Delta de Dirac, cujo significado preciso é que os dois lados das igualdades dão os mesmos resultados dentro de um integrando:

$$\delta(-x) = \delta(x) \quad (2.98)$$

$$x\delta(x) = 0 \quad (2.99)$$

$$\delta(ax) = a^{-1}\delta(x) \quad (a > 0) \quad (2.100)$$

$$\int \delta(a-x)\delta(x-b)dx = \delta(a-b) \quad (2.101)$$

$$f(x)\delta(x-a) = f(a)\delta(x-a) \quad (2.102)$$

2.9 Representações

Depois de todo esse tratamento mais abstrato, Dirac desenvolve as representações dos elementos do formalismo. Uma representação faz a correspondência entre um elemento abstrato e uma estrutura definida em termos de números, algo portanto mais concreto. Os diversos modos de fazer a correspondência entre o elemento abstrato e o ‘conjunto’ de números são chamados de representações, e a estrutura concreta associada ao elemento é chamada de representativo.

Para caracterizarmos uma representação, basta um conjunto completo de bras, os quais serão denominados bras básicos, sendo eles suficientes para fixar a representação. Esse conjunto de bras caracteriza uma base do espaço vetorial de bras.

Os produtos escalares do ket $|a\rangle$ com cada bra básico define o representativo desse ket. Conhecendo o resultado do produto escalar de um ket com todos os bras básicos podemos determinar univocamente o ket.

A prova é simples, suponha que tenhamos dois kets $|a\rangle$ e $|a'\rangle$ para os quais esse conjunto de números seja o mesmo, ou seja, para qualquer bra básico $|\lambda\rangle$,

$$\langle\lambda|a\rangle = \langle\lambda|a'\rangle$$

então, a diferença $|a\rangle - |a'\rangle$ terá produto escalar com qualquer outro bra básico nulo, e portanto, com qualquer bra:

$$\langle \lambda | (|a\rangle - |a'\rangle) = 0 \implies \forall \langle \psi | \in \mathcal{B} : \langle \psi | (|a\rangle - |a'\rangle) = 0,$$

sendo assim, os dois kets são iguais.

Dirac supõe que cada bra básico pode ser enumerado por um ou mais parâmetros, $\lambda_1, \lambda_2, \dots, \lambda_u$, cada um tendo um certo valor numérico. Um bra básico é então escrito em termos dos valores de cada um dos parâmetros λ : $\langle \lambda_1 \lambda_2 \dots \lambda_u |$, e o conjugado complexo de um bra básico, um ket básico, como $|\lambda_1 \lambda_2 \dots \lambda_u\rangle$:

$$\overline{\langle \lambda_1 \lambda_2 \dots \lambda_u |} = |\lambda_1 \lambda_2 \dots \lambda_u\rangle, \quad \overline{|\lambda_1 \lambda_2 \dots \lambda_u\rangle} = \langle \lambda_1 \lambda_2 \dots \lambda_u |. \quad (2.103)$$

O representativo de um ket qualquer $|a\rangle$ será o conjunto de números dados por $\langle \lambda_1 \lambda_2 \dots \lambda_u | a \rangle$, para cada parâmetro λ_i varrendo todo o seu domínio.

Podemos então denotar o representativo de um ket $|a\rangle$ como $\langle \lambda_1 \lambda_2 \dots \lambda_u | a \rangle$, podendo o mesmo ser visto como uma função complexa das variáveis $\lambda_1, \lambda_2, \dots, \lambda_u$, cada variável estando restrita a determinado domínio.

De forma geral, os bras básicos de uma representação não precisam ser todos linearmente independentes, no entanto, nas representações mais usuais eles são todos linearmente independentes, e ainda são ortogonais uns aos outros. Nesse caso a representação é uma representação ortogonal, ou podemos dizer que os bras básicos formam uma base ortogonal.

Consideremos agora uma representação ortogonal com bras básicos $\langle \lambda_1 \lambda_2 \dots \lambda_u |$, caracterizados pelos parâmetros $\lambda_1, \lambda_2, \dots, \lambda_u$ cujos domínios são todos reais. Dirac então define para cada parâmetro λ_i um operador linear L_i definido pela equação:

$$\langle \lambda_1 \lambda_2 \dots \lambda_u | L_i = \lambda_i \langle \lambda_1 \lambda_2 \dots \lambda_u |, \quad (2.104)$$

Tomando o conjugado dessa equação, temos também que:

$$\overline{L_i} |\lambda_1 \lambda_2 \dots \lambda_u\rangle = \lambda_i |\lambda_1 \lambda_2 \dots \lambda_u\rangle. \quad (2.105)$$

Como os diferentes bras básicos são todos ortonormais, demonstra-se que os operadores L_i são todos autoadjuntos. Para verificarmos essa afirmação, fazemos

$$\langle \lambda_1 \lambda_2 \dots \lambda_u | L_i | \lambda'_1 \lambda'_2 \dots \lambda'_u \rangle = \lambda_i \langle \lambda_1 \lambda_2 \dots \lambda_u | \lambda'_1 \lambda'_2 \dots \lambda'_u \rangle = \lambda_i \delta_{\lambda_1, \lambda'_1} \delta_{\lambda_2, \lambda'_2} \dots \delta_{\lambda_u, \lambda'_u} \quad (2.106)$$

e

$$\langle \lambda'_1 \lambda'_2 \dots \lambda'_u | L_i | \lambda_1 \lambda_2 \dots \lambda_u \rangle = \lambda_i \langle \lambda'_1 \lambda'_2 \dots \lambda'_u | \lambda_1 \lambda_2 \dots \lambda_u \rangle = \lambda'_i \delta_{\lambda'_1, \lambda_1} \delta_{\lambda'_2, \lambda_2} \dots \delta_{\lambda'_u, \lambda_u} \quad (2.107)$$

$$= \lambda_i \delta_{\lambda_1, \lambda'_1} \delta_{\lambda_2, \lambda'_2} \dots \delta_{\lambda_u, \lambda'_u}. \quad (2.108)$$

Obtemos então que,

$$\overline{\langle \lambda'_1 \lambda'_2 \dots \lambda'_u | L_i | \lambda_1 \lambda_2 \dots \lambda_u \rangle} = \langle \lambda_1 \lambda_2 \dots \lambda_u | \overline{L_i} | \lambda'_1 \lambda'_2 \dots \lambda'_u \rangle = \langle \lambda_1 \lambda_2 \dots \lambda_u | L_i | \lambda'_1 \lambda'_2 \dots \lambda'_u \rangle, \quad (2.109)$$

e, como isso é válido para qualquer bra e qualquer ket básico, conclui-se que $L_i = \overline{L_i}$ sendo L_i portanto autoadjunto. Como os autoestados dos operadores L_i são os bras e kets básicos, podemos agora concluir que cada operador L_i é um observável. Como os bras e kets básicos são autoestados simultâneos de todos os operadores L_i , decorre do teorema que todos os operadores L_i comutam entre si.

2.9.1 Conjunto Completo de Observáveis que comutam entre si

Dirac define um *Conjunto Completo de Observáveis que comutam entre si* como sendo um conjunto de observáveis que comutam uns com os outros e para os quais só existe um autoestado linearmente independente para cada conjunto de autovalores. Em outras palavras, dado os autovalores associados a cada observável, não existem dois ou mais autoestados linearmente associados a esse conjunto de autovalores, e o conjunto de autovalores é portanto não degenerado.

Assim, o conjunto dos observáveis L_1, L_2, \dots, L_u forma um conjunto completo de observáveis. Por outro lado, um conjunto completo de observáveis que comutam entre si X_1, X_2, \dots, X_u pode ser associado a uma representação ortogonal $\langle x'_1 x'_2 \dots x'_u |$ com

$$\langle x'_1 x'_2 \dots x'_u | X_i = x'_i \langle x'_1 x'_2 \dots x'_u |,$$

fazendo os observáveis X_i um papel correspondente aos operadores L_i definidos anteriormente.

Quando temos um conjunto de observáveis que comutam entre si X_1, X_2, \dots, X_u , mas o conjunto não é completo, apenas conhecer os autovalores de cada um desses observáveis pode não ser suficiente para especificar um único bra básico, pois podemos ter vários bras básicos pertencentes ao mesmo conjunto de autovalores. Assim, o conjunto de parâmetros x'_1, x'_2, \dots, x'_u não é suficiente para especificar univocamente um ket básico, mas podemos acrescentar um outro conjunto de parâmetros $\lambda_1, \lambda_2, \dots, \lambda_v$ de forma que o conjunto total dos

parâmetros sejam agora capazes de especificar univocamente um bra básico. Qualquer bra básico pode então ser denotado por $\langle x'_1 x'_2 \dots x'_u \lambda'_1 \lambda'_2 \dots \lambda'_v |$. Associado a esses parâmetros adicionais λ , temos os operadores L_1, L_2, \dots, L_u definidos de forma análoga ao que foi feito antes:

$$\langle x'_1 x'_2 \dots x'_u \lambda'_1 \lambda'_2 \dots \lambda'_v | L_i = \lambda'_i \langle x'_1 x'_2 \dots x'_u \lambda'_1 \lambda'_2 \dots \lambda'_v |. \quad (2.110)$$

Juntos, o conjunto de observáveis X_1, X_2, \dots, X_u e o conjunto de observáveis L_1, L_2, \dots, L_u formam um conjunto completo de observáveis que comutam entre si. Assim, de qualquer conjunto de observáveis que comutam entre si pode ser feito um conjunto completo de observáveis que comutam entre si acrescentando-se alguns outros observáveis.

2.10 Propriedades dos Vetores Básicos

Até o momento não abordamos a norma dos vetores básicos, e por isso os mesmos não estão completamente caracterizados. A abordagem mais simples e usual é fixar a norma de todos os vetores fazendo o processo chamado normalização, no entanto, só podemos realmente normalizar os vetores básicos quando eles são autoestados de um observável com autovalores discretos, caso contrário, vimos anteriormente no estudo da função delta que a norma diverge indo com um valor infinito. Justamente por isso, Dirac generaliza as noções convencionais de um espaço vetorial utilizando a Delta de Dirac e com isso ele nos permite lidarmos com essa divergência na norma sem maiores complicações matemáticas.

Vamos supor que temos um único observável X e que só existe um único autovetor para cada autovalor de X , ou seja, o observável X sozinho forma um conjunto completo de observáveis e nos permite construir uma representação ortonormal. Os vetores básicos serão então escritos como $\langle x' |$ e $| x' \rangle$.

No caso dos autovalores forem discretos, simplesmente normalizamos os vetores básicos:

$$\langle x' | x'' \rangle = \begin{cases} 1, & \text{se } x' = x'' \\ 0, & \text{caso contrário.} \end{cases} \quad (2.111)$$

Definindo a função conhecida como Delta de Kronecker:

$$\delta_{x,y} = \begin{cases} 1, & \text{se } x = y \\ 0, & \text{caso contrário;} \end{cases} \quad (2.112)$$

a equação da normalização pode ser escrita como:

$$\langle x' | x'' \rangle = \delta_{x', x''}. \quad (2.113)$$

Vamos agora considerar o caso em que o observável X possui autovalores contínuos. Nesse caso vimos que

$$\int_{\Omega} \langle x' | x'' \rangle dx'' = \alpha(x', x'') > 0$$

quando o domínio de integração inclui o ponto x' , e que

$$\langle x' | x'' \rangle = 0 \text{ se } x' \neq x''$$

devido à ortogonalização. Então podemos considerar que

$$\langle x' | x'' \rangle = \alpha(x', x'') \delta(x' - x''),$$

e aqui vemos a importância da Delta de Dirac, pois sem ela não seria possível expressar essa equação.

O que fazemos então é fixar o valor de $\alpha(x', x'')$ como sendo 1, então obtemos o correspondente à normalização dos vetores básicos no caso de autovalores contínuos:

$$\langle x' | x'' \rangle = \delta(x' - x''), \quad (2.114)$$

temos agora um grande paralelo entre os casos discreto e contínuo.

Dirac também generaliza essas definições para o caso em que temos tanto autovalores contínuos como discretos. Vamos utilizar $x^{(r)}$ e $x^{(s)}$ para denotar autovalores discretos, e x' e x'' para denotar autovalores contínuos, então temos a generalização:

$$\langle x^{(r)} | x^{(s)} \rangle = \delta_{x^{(r)}, x^{(s)}}, \quad (2.115)$$

$$\langle x^{(r)} | x' \rangle = 0, \quad (2.116)$$

$$\langle x' | x'' \rangle = \delta(x' - x''). \quad (2.117)$$

2.11 O Operador Identidade

Os bras e kets básicos possuem uma importante relação com o operador identidade $\hat{\mathbf{1}}$. No caso discreto ela se expressa como

$$\sum_{x'} |x'\rangle \langle x'| = \hat{\mathbf{1}}; \quad (2.118)$$

e no caso contínuo:

$$\int |x'\rangle \langle x'| dx' = \hat{\mathbf{1}}. \quad (2.119)$$

Essas relações decorrem do fato de que, no caso discreto

$$\sum_{x'} |x'\rangle \langle x'| |x''\rangle = \sum_{x'} |x'\rangle \delta_{x',x''} = |x''\rangle, \quad (2.120)$$

e no caso contínuo

$$\int |x'\rangle dx' \langle x'| |x''\rangle = \int |x'\rangle dx' \delta(x' - x'') = |x''\rangle. \quad (2.121)$$

Como isso é válido para qualquer ket básico $|x''\rangle$, e qualquer ket pode ser expandido em termos dos kets básicos, infere-se que é válido no geral para qualquer ket.

Essas equações nos mostram a forma de se expandir bras ou kets em termos dos bras ou kets básicos. No caso discreto, fazemos:

$$\forall |P\rangle \in \mathcal{K} : |P\rangle = \hat{\mathbf{1}} |P\rangle = \sum_{x'} |x'\rangle \langle x'| P\rangle, \quad (2.122)$$

e no caso contínuo:

$$\forall |P\rangle \in \mathcal{K} : |P\rangle = \hat{\mathbf{1}} |P\rangle = \int |x'\rangle dx' \langle x'| P\rangle. \quad (2.123)$$

O produto escalar também ganha uma representação importante:

$$\langle Q|P\rangle = \sum_{x'} \langle Q|x'\rangle \langle x'|P\rangle, \quad (2.124)$$

no caso discreto, e

$$\langle Q|P\rangle = \int \langle Q|x'\rangle dx' \langle x'|P\rangle, \quad (2.125)$$

no caso contínuo.

Quando temos autovalores discretos e contínuos, o operador identidade é generalizado como:

$$\sum_{x'} |x'\rangle\langle x'| + \int |x'\rangle\langle x'| dx' = \hat{\mathbf{1}}, \quad (2.126)$$

e as outras generalizações decorrem de imediato.

2.12 A representação dos Operadores Lineares

De forma análoga ao que foi feito com bras e kets, a representação de um operador linear é um conjunto de números que define completamente o operador, e é fixada uma vez que tenhamos um conjunto de bras básicos.

Consideremos que os vetores básicos são autovetores simultâneos do conjunto completo de observáveis que comutam entre si X_1, X_2, \dots, X_u . Seja A um operador linear qualquer, então tomamos um ket básico genérico $\langle x'_1 x'_2 \dots x'_u |$, e um ket básico genérico $|x''_1 x''_2 \dots x''_u\rangle$, então os números

$$\langle x'_1 x'_2 \dots x'_u | A | x''_1 x''_2 \dots x''_u \rangle \quad (2.127)$$

definem um representativo do operador A pois são suficientes para definir A completamente, uma vez que caracterizam completamente a atuação desse operador em qualquer ket.

No caso particular em que temos apenas um observável X com autovalores discretos, o representativo do operador A é um conjunto discreto de números $\langle x' | A | x'' \rangle$. A forma mais natural de se escrever esses números é em uma tabela bidimensional:

$$\begin{pmatrix} \langle x^1 | A | x^1 \rangle & \langle x^1 | A | x^2 \rangle & \cdots \\ \langle x^2 | A | x^1 \rangle & \langle x^2 | A | x^2 \rangle & \cdots \\ \vdots & \vdots & \vdots \end{pmatrix}, \quad (2.128)$$

tabelas assim são chamadas matrizes, e os números, elementos de matriz. Sendo assim, os operadores podem ser representados por matrizes.

Se o operador A for o operador identidade, os elementos de matriz serão

$$\langle x' | \hat{\mathbf{1}} | x'' \rangle = \delta_{x', x''}, \quad (2.129)$$

e a matriz resultante é a matriz identidade. Se operador for o próprio observável X , os elementos de matriz serão:

$$\langle x' | X | x'' \rangle = x' \delta_{x', x''}, \quad (2.130)$$

e todos os elementos fora da diagonal principal serão nulos, a matriz portanto será diagonal.

Agora vamos examinar o representativo do produto AB de dois operadores lineares A e B em termos dos representativos de cada fator:

$$\langle x' | AB | x'' \rangle = \langle x' | A \sum_{x'''} | x''' \rangle \langle x''' | B | x'' \rangle = \sum_{x'''} \langle x' | A | x''' \rangle \langle x''' | B | x'' \rangle, \quad (2.131)$$

o que corresponde à multiplicação usual das matrizes correspondentes à cada fator. O mesmo ocorre com a adição e as outras operações, sendo que a operação correspondente a tomar o adjunto de um operador é tomar o transposto conjugado da matriz.

Dirac também generaliza o conceito de matriz para o caso dos operadores com autovalores contínuos e para o caso misto, em que temos autovalores contínuos e discretos.

2.13 Funções de Onda

A representação de um ket $|P\rangle$ quando o conjunto completo de observáveis é composto apenas pelo observável X é $\langle x' | P \rangle$, sendo portanto uma função. Seja ψ essa função, ou seja:

$$\langle x' | P \rangle = \psi(x'), \quad (2.132)$$

então concluí-se que ψ determina completamente o ket $|P\rangle$. Sendo assim, é interessante utilizarmos a própria função ψ para denotar o ket $|P\rangle$. Para isso Dirac define a notação $|\psi(X)\rangle$ com a seguinte propriedade:

$$\langle x' | P \rangle = \psi(x') \implies |P\rangle = |\psi(X)\rangle \quad (2.133)$$

Essa notação complementa a representação dos kets, correspondendo a uma representação explícita de um ket:

$$\langle x' | \psi(X) \rangle = \psi(x'). \quad (2.134)$$

Essa função $\psi(X)$ utilizada para designar o ket é a chamada função de onda. Essa notação pode ser utilizada em qualquer caso, mas é especialmente útil quando os autovalores de X são contínuos, só nesse caso podemos ter uma função de onda contínua e propriedades interessantes.

Uma função $f(X)$ do observável X satisfaz

$$\langle x' | f(X) | \psi(X) \rangle = f(x') \psi(x'), \quad (2.135)$$

de onde concluímos que

$$f(X) |\psi(X)\rangle = |f(X)\psi(X)\rangle. \quad (2.136)$$

Esse resultado é geral e vale para quaisquer funções f e ψ . Por esse resultado, Dirac conclui que a barra vertical $|$ não é necessária nessa notação, e que podemos removê-la. Assim temos as novas regras para a notação:

$$\langle x'|P\rangle = \psi(x') \implies |P\rangle = \psi(X)\rangle. \quad (2.137)$$

Podemos ainda considerar o ket $\psi(X)\rangle$ como o produto do operador linear $\psi(X)$ com um ket especial denotado simplesmente por \rangle . Esse ket é denominado ket padrão, e é caracterizado pela seguinte propriedade:

$$\forall x' : \langle x'|\rangle = 1, \quad (2.138)$$

em que $\langle x'|$ é um bra básico.

De forma correspondente ao ket padrão \rangle , definimos também o bra padrão \langle

$$\forall x' : \langle|x'\rangle = 1, \quad (2.139)$$

e obtemos o correspondente à essa notação para o caso dos bras

$$|\overline{\psi(X)}\rangle = \langle\overline{\psi}(X)| = \langle\overline{\psi}(X), \quad (2.140)$$

como no caso dos kets, podemos considerar o bra $\langle\overline{\psi}(X)| = \langle\overline{\psi}(X)$ como sendo o produto do bra padrão \langle com o operador linear $\overline{\psi}(X)$.

Quando os autovalores de X são contínuos, temos a importante propriedade no cálculo do produto interno

$$\langle\overline{\psi}(X)|\phi(X)\rangle = \int \overline{\psi(x')} \phi(x') dx'. \quad (2.141)$$

Todas essas definições e propriedades se estendem naturalmente para o caso em que temos mais observáveis compondo o conjunto completo de observáveis.

O Formalismo de von Neumann da Mecânica Quântica

Von Neumann inicia o seu livro, “The Mathematical Foundations of Quantum Mechanic” [1] afirmando que o seu objetivo é “*apresentar a nova mecânica quântica em uma representação uniforme que é, até onde isso for possível e útil, matematicamente rigorosa*” [1, p. vii]. Desde o início de sua obra, von Neumann se coloca claramente em contraposição ao trabalho de Dirac [13]. Ele reconhece o valor e importância trabalho de Dirac:

Dirac, em diversos artigos, e também em seu recentemente publicado livro, deu uma representação da mecânica quântica que dificilmente será superada em concisão e elegância, e que é ao mesmo tempo de caráter invariante. [1, p. viii]

Porém coloca o seu livro como sendo uma resposta ao trabalho de Dirac. Ele perfaz duras críticas quanto ao rigor matemático:

O método de Dirac, mencionado acima, (e isso é negligenciado atualmente em grande parte da literatura da mecânica quântica, por causa da clareza e elegância da teoria) de forma alguma satisfaz os requerimentos de rigor matemático – nem mesmo se esses forem reduzidos de forma natural e adequada na medida comum em outras partes da física teórica. [1, p. viii - ix]

e critica especialmente a conhecida ‘função’ delta de Dirac:

[...] isso requer a introdução de uma função “imprópria” com propriedades auto-contraditórias. A inserção de tal “ficcção” matemática é freqüentemente necessária na abordagem de Dirac, mesmo se o problema em mãos for meramente o cálculo numérico de resultados de um bem definido resultado experimental. Não haveria nenhuma objeção aqui se esses conceitos, os quais não podem ser incorporados no formalismo atual da análise, fossem intrinsecamente necessários para a teoria física. [1, p. xi]

Von Neumann então coloca a sua formulação da mecânica quântica como sendo capaz de substituir a formulação de Dirac conservando-se livre dessas ‘inconsistências’ matemáticas que ele critica tão profundamente:

Deveria sim ser salientado que a “Teoria Quântica das Transformações” pode ser estabelecida em uma maneira que é tão clara e unificada, mas também sem objeções matemáticas. [1, p. xi]

e ainda ressalta que o seu trabalho difere essencialmente do anterior:

Deve ser enfatizado que a estrutura correta não precisa consistir em um refinamento matemático e explicação do método de Dirac, mas sim ela requer um procedimento que difere desde o começo, ou seja, a estruturação sobre a teoria de Hilbert dos operadores. [1, p. xi]

Observe que, não apenas escolhemos comparar essas duas formulações originais, como o próprio von Neumann escreveu seu livro num clima de comparação e contraposição ao trabalho de Dirac. Veremos no entanto que, apesar de suas alegações, o formalismo de von Neumann não é suficiente para descrever alguns importantes elementos descritos pelo outro formalismo: os sistemas com espectro contínuo, para os quais a função Delta tem se mostrado essencial.

No primeiro capítulo, von Neumann apresenta e discute as formulações originais da mecânica quântica, mais precisamente ele discute a mecânica matricial de Heisenberg-Born-Jordan [6–8] e a mecânica ondulatória de Schrodinger e como as duas foram unificadas na “Teoria das Transformações” por Dirac e Jordan. Esse capítulo é interessante para compreendermos o contexto e as motivações para esse livro clássico.

Grande parte do trabalho de von Neumann é especialmente dedicada à caracterização do espaço de Hilbert abstrato e dos principais elementos que decorrem desse aparato matemático, que é, segundo von Neumann, a base de todo o formalismo matemático da mecânica quântica.

Dentro do conceito de espaço de Hilbert, existem duas distinções principais às quais daremos uma notação distinta: um espaço de Hilbert de dimensão finita n será denotado por \mathcal{H}_n , e um espaço de Hilbert de dimensão infinita será denotado por \mathcal{H}_∞ . Quando não queremos fazer distinção entre cada caso, utilizaremos o símbolo \mathcal{H} , assim sendo, as propriedades válidas em \mathcal{H} são válidas tanto nos espaços de Hilbert de dimensão finita \mathcal{H}_n , como infinita \mathcal{H}_∞ .

Primeiramente, von Neumann postula para \mathcal{H} as propriedades vetoriais típicas: \mathcal{H} é um espaço vetorial com produto interno hermitiano [1, p. 34-35]. Utilizaremos a definição mais usual de um espaço vetorial, a mesma que utilizamos no formalismo de Dirac, diferindo em alguns axiomas da definição de von Neumann. As duas formulações no entanto são rigorosamente equivalentes como o próprio von Neumann afirma.

A. \mathcal{H} é um espaço vetorial sobre os números complexos [1, p. 36]. Ou seja, as operações de adição e multiplicação por escalar estão definidas em \mathcal{H} e resultam em elementos em \mathcal{H} :

Axioma 1.

$$\forall |a\rangle, |b\rangle \in \mathcal{H} : |a\rangle + |b\rangle \in \mathcal{H} \quad (3.1a)$$

$$\forall |a\rangle \in \mathcal{H}, \alpha \in \mathbb{C} : \alpha |a\rangle \in \mathcal{H} \quad (3.1b)$$

e essas operações satisfazem todas as propriedades de um espaço vetorial sobre os números complexos:

Axioma 2. $\exists -|a\rangle \in \mathcal{H}, \forall |a\rangle, |b\rangle, |c\rangle \in \mathcal{H}, \forall \alpha, \beta \in \mathbb{C} :$

$$|a\rangle + |b\rangle = |b\rangle + |a\rangle \quad (3.2a)$$

$$|a\rangle + (|b\rangle + |c\rangle) = (|a\rangle + |b\rangle) + |c\rangle \quad (3.2b)$$

$$\exists 0 \in \mathcal{H} : |a\rangle + 0 = |a\rangle \quad (3.2c)$$

$$|a\rangle + (-|a\rangle) = 0 \quad (3.2d)$$

$$\alpha(\beta |a\rangle) = (\alpha\beta) |a\rangle \quad (3.2e)$$

$$\alpha(|a\rangle + |b\rangle) = \alpha |a\rangle + \alpha |b\rangle \quad (3.2f)$$

$$(\alpha + \beta) |a\rangle = \alpha |a\rangle + \beta |a\rangle \quad (3.2g)$$

$$1 |a\rangle = |a\rangle \quad (3.2h)$$

Estamos utilizando a mesma notação e buscando aproximar ao máximo a forma de apresentação dos formalismos a fim de facilitar a percepção das diferenças e semelhanças entre ambos. Assim estamos fazendo o trabalho de escrever os dois formalismos “na mesma linguagem”.

Na sequência, von Neumann define alguns conceitos da álgebra linear:

Definição 1. *Os elementos $|a_1\rangle, |a_2\rangle, \dots, |a_k\rangle$ são linearmente independentes se $\alpha_1 |a_1\rangle + \alpha_2 |a_2\rangle + \dots + \alpha_k |a_k\rangle = 0$ implica que $\alpha_1 = \alpha_2 = \dots = \alpha_k = 0$, sendo os diferentes α_i números complexos.*

Definição 2. *Um subconjunto \mathcal{I} de um espaço de Hilbert \mathcal{H} é denominado uma variedade linear se ele contém todas as combinações lineares de seus próprios elementos, ou seja, se qualquer combinação $\alpha_1 |a_1\rangle + \alpha_2 |a_2\rangle + \dots + \alpha_k |a_k\rangle$ de quaisquer elementos $|a_i\rangle$ do conjunto \mathcal{H} e para quaisquer valores complexos dos elementos α_i , pertence ao próprio conjunto \mathcal{I} . Se \mathcal{U} é um subconjunto arbitrário de \mathcal{H} , então o conjunto de todas as combinações lineares dos elementos de \mathcal{U} é uma variedade linear, “a variedade linear gerada por \mathcal{U} ”.*

B. Na segunda etapa da caracterização de um espaço de Hilbert, von Neumann define o produto interno em \mathcal{H} [1, p. 38]:

Definição 3. O produto interno em \mathcal{H} é definido como sendo uma função $\langle \cdot, \cdot \rangle$ que leva dois vetores em um número complexo $\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{C}$ e que satisfaz as seguintes propriedades: $\forall |a\rangle, |b\rangle, \in \mathcal{H}, \forall \alpha \in \mathbb{C}$:

$$\langle |a\rangle + |a'\rangle, |b\rangle \rangle = \langle |a\rangle, |b\rangle \rangle + \langle |a'\rangle, |b\rangle \rangle \quad (3.3a)$$

$$\langle \alpha |a\rangle, |b\rangle \rangle = \alpha \langle |a\rangle, |b\rangle \rangle \quad (3.3b)$$

$$\langle |a\rangle, |b\rangle \rangle = \overline{\langle |b\rangle, |a\rangle \rangle} \quad (3.3c)$$

$$\langle |a\rangle, |a\rangle \rangle \geq 0, \quad \langle |a\rangle, |a\rangle \rangle = 0 \iff |a\rangle = 0 \quad (3.3d)$$

Dessa definição decorrem as seguintes propriedades para o segundo fator do produto interno:

$$\langle |a\rangle, |b\rangle + |b'\rangle \rangle = \langle |a\rangle, |b\rangle \rangle + \langle |a\rangle, |b'\rangle \rangle \quad (3.4)$$

$$\langle |a\rangle, \alpha |b\rangle \rangle = \bar{\alpha} \langle |a\rangle, |b\rangle \rangle, \quad (3.5)$$

ou seja, o produto interno é linear no primeiro fator e anti-linear no segundo, Note que $\langle |a\rangle, |a\rangle \rangle$ é necessariamente um número real uma vez que $\overline{\langle |a\rangle, |a\rangle \rangle} = \langle |a\rangle, |a\rangle \rangle$.

Podemos agora definir o comprimento, a magnitude ou a norma de um vetor:

Definição 4. A norma $\| |a\rangle \|$ de um vetor $|a\rangle$ é um número real positivo definido por:

$$\forall |a\rangle \in \mathcal{H} : \quad \| |a\rangle \| = \sqrt{\langle |a\rangle, |a\rangle \rangle} \quad (3.6)$$

A distância entre dois vetores é então definida em termos da norma:

Definição 5. A distância entre dois vetores $|a\rangle$ e $|b\rangle$ é denotada por $d(|a\rangle, |b\rangle)$ e é definida como

$$\forall |a\rangle, |b\rangle \in \mathcal{H} : \quad d(|a\rangle, |b\rangle) = \| |a\rangle - |b\rangle \|$$

Como \mathcal{H} é um espaço vetorial com produto interno, diversos teoremas da álgebra linear se aplicam diretamente, e suas provas serão omitidas. Um dos principais teoremas é a conhecida desigualdade de Cauchy-Schwarz:

Teorema 1.

$$\forall |a\rangle, |b\rangle \in \mathcal{H} : \quad \langle |a\rangle, |b\rangle \rangle \leq \| |a\rangle \| \cdot \| |b\rangle \| \quad (3.7)$$

Essas definições de norma e de distância satisfazem todas as propriedades usuais que caracterizam uma norma e uma distância. Temos a homogeneidade da norma:

$$\forall |a\rangle \in \mathcal{H}, \alpha \in \mathbb{C} : \quad \|\alpha |a\rangle\| = |\alpha| \| |a\rangle \|, \quad (3.8)$$

a positividade da norma:

$$\forall |a\rangle \in \mathcal{H} : \quad \| |a\rangle \| \geq 0, \quad \| |a\rangle \| = 0 \iff |a\rangle = 0, \quad (3.9)$$

e a desigualdade triangular:

$$\forall |a\rangle, |b\rangle \in \mathcal{H} : \quad \| |a\rangle + |b\rangle \| \leq \| |a\rangle \| + \| |b\rangle \|. \quad (3.10)$$

3.1 Limite e Continuidade

Uma vez definido o conceito de distância, podemos abordar os conceitos de limite e continuidade. Assim von Neumann define o conceito de funções contínuas no espaço de Hilbert:

Definição 6. *Uma função $F(|a\rangle)$ definida sobre \mathcal{H} e cujos valores são números complexos é contínua no ponto $|a_0\rangle \in \mathcal{H}$ se para cada número real $\varepsilon > 0$ existir um número real $\delta > 0$, tal que $\| |a\rangle - |a_0\rangle \| < \delta$ implica que $|F(|a\rangle) - F(|a_0\rangle)| < \varepsilon$. Matematicamente, $F : \mathcal{H} \rightarrow \mathbb{C}$ é contínua no ponto $|a_0\rangle$ se:*

$$\forall |a\rangle \in \mathcal{H}, \varepsilon \in \mathbb{R} > 0 : \quad \exists \delta \in \mathbb{R} > 0 : \quad \| |a\rangle - |a_0\rangle \| < \delta \implies |F(|a\rangle) - F(|a_0\rangle)| < \varepsilon \quad (3.11)$$

Definição 7. *Uma função $F(|a\rangle)$ definida sobre \mathcal{H} e cujos valores são vetores no espaço de Hilbert é contínua no ponto $|a_0\rangle \in \mathcal{H}$ se para cada $\varepsilon > 0$ existir um $\delta > 0$, tal que $\| |a\rangle - |a_0\rangle \| < \delta$ implica que $\| F(|a\rangle) - F(|a_0\rangle) \| < \varepsilon$. Matematicamente, $F : \mathcal{H} \rightarrow \mathbb{C}$ é contínua no ponto $|a_0\rangle$ se:*

$$\forall |a\rangle \in \mathcal{H}, \varepsilon \in \mathbb{R} > 0 : \quad \exists \delta \in \mathbb{R} > 0 : \quad \| |a\rangle - |a_0\rangle \| < \delta \implies \| F(|a\rangle) - F(|a_0\rangle) \| < \varepsilon \quad (3.12)$$

As definições são análogas, mas uma é válida para quando a função tem imagem nos números complexos \mathbb{C} , e a outra no espaço de Hilbert \mathcal{H} . Temos agora a definição de função limitada:

Definição 8. *Uma função F é dita limitada em \mathcal{H} ou em um dado subconjunto \mathcal{U} de \mathcal{H} se, existe um número real positivo C tal que, em qualquer ponto $|a\rangle$ desse conjunto, $\| F(|a\rangle) \| <$*

C ou $|F(|a\rangle)| < C$, dependendo se a função possui valores complexos ou vetoriais, simbolicamente, uma função F é limitada em \mathcal{U} se

$$\exists C \in \mathbb{R} : \quad \forall |a\rangle \in \mathcal{U} : \quad \|F(|a\rangle)\| < C \text{ ou } |F(|a\rangle)| < C \quad (3.13)$$

e a definição de limite de uma sequência de vetores:

Definição 9. Uma sequência de vetores $|a_1\rangle, |a_2\rangle, \dots$ converge para um vetor $|a\rangle$, ou tem limite $|a\rangle$, se a sequência de números $\| |a_1\rangle - |a\rangle \|, \| |a_2\rangle - |a\rangle \|, \dots$ converge para zero. Simbolicamente,

$$\lim_{i \rightarrow \infty} |a_i\rangle = |a\rangle \iff \lim_{i \rightarrow \infty} \| |a_i\rangle - |a\rangle \| = 0 \quad (3.14)$$

Uma outra definição importante é a de ponto limite de um subconjunto \mathcal{U} do espaço de Hilbert \mathcal{H} :

Definição 10. Um ponto $|a\rangle$ é um ponto limite de um conjunto $\mathcal{U} \subseteq \mathcal{H}$ se ele for o limite de uma sequência de \mathcal{U} , ou seja, se existir uma sequência $|a_1\rangle, |a_2\rangle, \dots \subseteq \mathcal{U}$ tal que

$$\lim_{i \rightarrow \infty} |a_i\rangle = |a\rangle, \quad (3.15)$$

então o vetor $|a\rangle$ é um ponto limite de \mathcal{U} .

Alternativamente, um vetor $|a\rangle$ é um ponto limite de \mathcal{U} se ele for arbitrariamente próximo de algum ponto de \mathcal{U} , ou seja, se para qualquer $\varepsilon > 0$ existir um ponto $|a'\rangle \in \mathcal{U}$ tal que $\| |a\rangle - |a'\rangle \| < \varepsilon$. Simbolicamente, um vetor $|a\rangle$ é um ponto limite de \mathcal{U} se

$$\forall \varepsilon \in \mathbb{R} > 0 : \quad \exists |a'\rangle \in \mathcal{U} : \quad \| |a\rangle - |a'\rangle \| < \varepsilon \quad (3.16)$$

Definição 11. Um conjunto $\mathcal{U} \subseteq \mathcal{H}$ é dito fechado se ele contém todos os seus pontos limites, ou seja, se

$$\forall |a\rangle \in \mathcal{H} : \quad (\forall \varepsilon \in \mathbb{R} > 0 : \quad \exists |a'\rangle \in \mathcal{U} : \quad \| |a\rangle - |a'\rangle \| < \varepsilon) \implies |a\rangle \in \mathcal{U} \quad (3.17)$$

Agora temos a importante definição de conjunto denso, que é essencial na caracterização rigorosa de um espaço de Hilbert de dimensão infinita:

Definição 12. Um conjunto $\mathcal{U} \subseteq \mathcal{H}$ é denso em toda parte, ou simplesmente denso, se seus pontos limites englobam todo o espaço de Hilbert \mathcal{H} , ou seja, se todos os pontos de \mathcal{H} forem pontos limites de \mathcal{U} . Simbolicamente, \mathcal{U} é denso em toda parte se

$$\forall |a\rangle \in \mathcal{H} : \quad \forall \varepsilon \in \mathbb{R} > 0 : \quad \exists |a'\rangle \in \mathcal{U} : \quad \| |a\rangle - |a'\rangle \| < \varepsilon. \quad (3.18)$$

Para facilitar a compreensão podemos ainda dizer que um conjunto \mathcal{U} é denso em toda a parte se qualquer ponto de \mathcal{H} for arbitrariamente próximo de algum ponto de \mathcal{U} . Um exemplo são os números racionais que são densos sobre os números reais: cada número real é um número racional ou é arbitrariamente próximo a algum número racional.

Uma sequência de Cauchy é uma sequência em que os elementos ficam arbitrariamente próximos uns dos outros a medida que a sequência progride, assim definimos:

Definição 13. *Uma sequência $|a_1\rangle, |a_2\rangle, \dots \subseteq \mathcal{H}$ é uma sequência de Cauchy se para qualquer $\varepsilon > 0$, existe um N tal que $\| |a_m\rangle - |a_n\rangle \| < \varepsilon$, para quaisquer $m, n > N$, simbolicamente, temos uma sequência de Cauchy se*

$$\forall \varepsilon > 0 : \exists n \in \mathbb{Z} : \forall m, n \in \mathbb{Z} \geq N : \| |a_m\rangle - |a_n\rangle \| < \varepsilon \quad (3.19)$$

Von Neumann afirma que as propriedades **A** e **B** nos permitem uma boa caracterização de \mathcal{H} , mas ainda não são suficientes para que possamos distinguir os \mathcal{H}_n de \mathcal{H}_∞ , ou seja, até o momento não é possível fazer qualquer consideração quanto ao número de dimensões do espaço de Hilbert que estamos tratando. Von Neumann então postula a terceira propriedade diferenciando espaços de dimensão finita e infinita:

Axioma 3. $\mathbf{C}^{(n)}$. *Qualquer subconjunto $\mathcal{I} \subset \mathcal{H}_n$ de vetores linearmente independentes possui no máximo n elementos [1, p. 45]. Ou seja, só é possível construir um conjunto de no máximo n vetores linearmente independentes.*

Axioma 4. $\mathbf{C}^{(\infty)}$ *Existem subconjuntos $\mathcal{I} \subset \mathcal{H}_\infty$ de vetores linearmente independentes com infinitos elementos. Ou seja, é possível construir um conjunto de vetores linearmente independentes com uma quantidade arbitrária de elementos.*

Se as propriedades A e B são válidas, então necessariamente um dos dois axiomas $\mathbf{C}^{(n)}$ ou $\mathbf{C}^{(\infty)}$ é válido; eles servem então apenas para diferenciarmos se estamos tratando de um espaço de dimensão finita ou infinita.

Até o momento, esses axiomas são suficientes para caracterizar completamente um espaço de Hilbert de dimensão finita \mathcal{H}_n , no entanto, alguns axiomas a mais são necessários para a caracterização completa de um espaço de Hilbert de dimensão infinita \mathcal{H}_∞ . Assim von Neumann postula:

Axioma 5. **D.** \mathcal{H} é **completo**, ou seja, todas as sequências de Cauchy convergem [1, p. 46].

Axioma 6. **E.** \mathcal{H} é **separável**, ou seja, existe uma sequência $|a_1\rangle, |a_2\rangle, \dots \subset \mathcal{H}$ que é densa em \mathcal{H} . Em outras palavras, existe um conjunto contável $\{|a_1\rangle, |a_2\rangle, \dots\}$ que é denso em toda parte.

Este último postulado de certa forma limita o ‘tamanho’ do espaço de Hilbert mesmo ele tendo dimensão possivelmente infinita, pois associa o mesmo a um conjunto contável. É importante percebermos que esse axioma já é automaticamente válido em um espaço de dimensão finita. Para isso vamos verificar que um espaço \mathbb{R}^n é separável: tomando como sequência os vetores (r_1, r_2, \dots, r_n) em que os parâmetros r_i são todos racionais, observa-se que qualquer vetor de \mathbb{R}^n é arbitrariamente próximo a algum elemento dessa sequência e portanto \mathbb{R}^n é separável. Pelo isomorfismo entre \mathbb{R}^n e \mathcal{H}_n pode-se então demonstrar que \mathcal{H}_n também é separável. Assim sendo, von Neumann percebeu que os espaços de dimensão finita tinham essa propriedade, e tentou generaliza-los conservando essa propriedade. Note ainda que essa propriedade não aparece no formalismo de Dirac, e faz com que os dois formalismos não sejam rigorosamente equivalentes.

3.2 A Geometria do Espaço de Hilbert

Seguindo a abordagem de von Neumann, definimos o conceito de ortogonalidade:

Definição 14. *Dois vetores $|a\rangle, |b\rangle$ são ortogonais se $\langle |a\rangle, |b\rangle \rangle = 0$.*

Definição 15. *Dois subconjuntos lineares \mathcal{I}, \mathcal{U} são ortogonais se cada elemento de \mathcal{I} for ortogonal a cada elemento de \mathcal{U} .*

Definição 16. *Um conjunto de vetores $\mathcal{U} \subset \mathcal{H}$ é dito ortonormal se*

$$\forall |a\rangle, |b\rangle \in \mathcal{U} : \quad \langle |a\rangle, |b\rangle \rangle = \begin{cases} 1, & \text{se } |a\rangle = |b\rangle, \\ 0, & \text{se } |a\rangle \neq |b\rangle. \end{cases} \quad (3.20)$$

Por essa definição, veja que um conjunto ortonormal é também um conjunto de vetores linearmente independentes

Von Neumann então define o que é um conjunto ortonormal completo, uma definição que se mostrará equivalente à definição de Dirac de ‘conjunto completo de estados’:

Definição 17. *Um conjunto ortonormal \mathcal{U} é completo se ele não é subconjunto de nenhum outro conjunto ortonormal com elementos adicionais, ou seja, se*

$$\nexists \mathcal{D} \text{ ortonormal} : \quad \mathcal{U} \subset \mathcal{D}.$$

Alternativamente, conjunto ortonormal \mathcal{U} é completo se não existe um outro vetor

normalizado ortogonal a todo o conjunto \mathcal{U} , isto é,

$$\nexists |a\rangle \in \mathcal{H}, \| |a\rangle \| = 1 : (\forall |b\rangle \in \mathcal{U} : \langle |a\rangle, |b\rangle \rangle = 0)$$

Note que este é um outro conceito de conjunto ‘completo’, que não tem nenhuma correlação com a definição dada a esse termo anteriormente.

Vamos acompanhar a abordagem de von Neumann e acrescentar o índice (n) ou (∞) nos teoremas para identificar se o mesmo é válido no caso de dimensão finita ou infinita.

Agora temos dois importantes teoremas que limitam o tamanho de conjuntos ortonormais. O primeiro decorre em especial do axioma $\mathbf{C}^{(n)}$, e o segundo decorre dos axiomas $\mathbf{C}^{(\infty)}$ e \mathbf{E} :

Teorema 2. (n) *Todo conjunto ortonormal possui no máximo n elementos, e é completo apenas se possuir n elementos.*

Teorema 3. (∞) *Todo conjunto ortonormal é finito ou infinito contável, e se é completo, então é necessariamente infinito.*

Esse último teorema é importante para compreendermos um pouco mais do que significa o espaço de Hilbert ser separável, pois essa condição implica que os conjuntos ortonormais são no máximo infinito contáveis. Com isso, não podemos, por exemplo, representar a base das posições $\{|x\rangle : x \in \mathbb{R}\}$, pois nesse caso o conjunto não seria contável, possuindo a cardinalidade dos números reais. O que podemos representar no espaço de Hilbert são apenas bases como $\{|n\rangle : n \in \mathbb{N}\}$, uma base infinita utilizada no oscilador harmônico. Sendo assim, o formalismo de von Neuman permite conjuntos ortogonais com cardinalidade no máximo \aleph_0 , que é a cardinalidade dos números naturais, enquanto que o formalismo de Dirac permite \aleph_1 , que é a cardinalidade dos números reais.

Para clarificar esse teorema vamos demonstrá-lo: Seja \mathcal{D} um conjunto ortonormal, e $|a\rangle$ e $|b\rangle$ dois elementos diferentes desse conjunto; assim,

$$\langle |a\rangle - |b\rangle, |a\rangle - |b\rangle \rangle = \langle |a\rangle, |a\rangle \rangle + \langle |b\rangle, |b\rangle \rangle - \langle |a\rangle, |b\rangle \rangle - \langle |b\rangle, |a\rangle \rangle = 2,$$

logo $\| |a\rangle - |b\rangle \| = \sqrt{2}$. Agora seja $\{s_1, s_2, \dots\}$ sequência que é densa em \mathcal{H} , a qual existe pelo postulado \mathbf{E} . Assim, para cada elemento

$$\forall |a\rangle \in \mathcal{D} : \exists |s_i\rangle \in \{s_1, s_2, \dots\} : \| |a\rangle - |s_i\rangle \| < \frac{1}{2}.$$

Então, consideremos os elementos da sequência correspondentes a $|a\rangle$ e $|b\rangle$, como sendo respectivamente $|s_i\rangle$ e $|s_j\rangle$; nesse caso, esses elementos necessariamente são diferentes, pois caso

contrário:

$$|s_i\rangle = |s_j\rangle \implies \| |a\rangle - |b\rangle \| = \| (|a\rangle - |s_i\rangle) - (|b\rangle - |s_i\rangle) \| \leq \| |a\rangle - |s_i\rangle \| + \| |b\rangle - |s_i\rangle \| < \frac{1}{2} + \frac{1}{2} = 1,$$

o que nos leva a uma contradição. Assim, a cada elemento de \mathcal{D} corresponde um único elemento da sequência, e portanto \mathcal{D} é finito ou é também uma sequência.

Von Neumann então aborda um aspecto fundamental de seu trabalho: as questões de convergência, que dificilmente são abordadas no trabalho de Dirac. As questões de convergência tem importância apenas para espaço de dimensão infinita, pois em dimensão finita elas são triviais.

Vamos utilizar de forma geral o símbolo

$$\sum_i$$

para denotar um somatório genérico sobre uma sequência, independente de ela ser ou não finita. Assim podemos utilizar a mesma notação para os casos de dimensão finita ou infinita.

Teorema 4. *Seja $\{|x_1\rangle, |x_2\rangle, \dots\}$ um conjunto ortonormal, então*

$$\forall |a\rangle, |b\rangle \in \mathcal{H} : \sum_i \langle |a\rangle, |x_i\rangle \rangle \overline{\langle |b\rangle, |x_i\rangle \rangle}$$

converge. Alternativamente,

$$\forall |a\rangle, |b\rangle \in \mathcal{H} : \sum_i \langle |a\rangle, |x_i\rangle \rangle \overline{\langle |b\rangle, |x_i\rangle \rangle} \in \mathbb{C}.$$

Teorema 5. *Seja $\{|x_1\rangle, |x_2\rangle, \dots\}$ um conjunto ortonormal, então a série*

$$\sum_i \alpha_i |x_i\rangle, \quad \alpha_i \in \mathbb{C},$$

converge se, e somente se, a série

$$\sum_i |\alpha_i|$$

converge. Alternativamente,

$$\sum_i \alpha_i |x_i\rangle \in \mathcal{H} \iff \sum_i |\alpha_i| \in \mathbb{R}.$$

Um resultado trivial é que, se $|a\rangle = \sum_i \alpha_i |x_i\rangle$, então $\langle |a\rangle, |x_i\rangle \rangle = \alpha_i$.

Vamos agora definir o que é uma variedade linear fechada, um conceito que será muito utilizado por possuir mais propriedades que uma variedade linear:

Definição 18. *Uma variedade linear que também é fechada é chamado uma variedade linear fechada.*

E definimos uma notação para gerar variedades lineares fechadas ou não:

Definição 19. *Seja um conjunto de vetores $\mathcal{U} \subseteq \mathcal{H}$, então denominamos $\{\mathcal{U}\}$ como sendo a variedade linear gerada por \mathcal{U} , ou seja, o conjunto de todas as combinações lineares de elementos de \mathcal{U} .*

Definição 20. *Seja um conjunto de vetores $\mathcal{U} \subseteq \mathcal{H}$, então denominamos $[\mathcal{U}]$ como sendo a variedade linear fechada gerada por \mathcal{U} , ou seja, o variedade linear gerada por \mathcal{U} ascrescida de todos os seus pontos limites.*

Teorema 6. *Seja $\{|x_1\rangle, |x_2\rangle, \dots\rangle$ um conjunto ortonormal, então, para qualquer $|a\rangle \in \mathcal{H}$, temos a convergência:*

$$|a'\rangle = \sum_i \langle |a\rangle, |x_i\rangle \rangle |x_i\rangle \in \mathcal{H}.$$

Temos também que $|a\rangle - |a'\rangle$ é ortogonal ao conjunto ortonormal $\{|x_1\rangle, |x_2\rangle, \dots\rangle$.

De acordo com von Neumann, esses teoremas nos levam um teorema para a caracterização de um conjunto ortonormal completo, reobtendo a definição de Dirac de um conjunto completo:

Teorema 7. *Seja um conjunto ortonormal $\{|x_1\rangle, |x_2\rangle, \dots\rangle$, para esse conjunto ser completo, cada uma dessas condições são necessárias e suficientes:*

1. *A variedade linear fechada gerada por $\{|x_1\rangle, |x_2\rangle, \dots\rangle$ é o próprio espaço de Hilbert \mathcal{H} , ou seja, $[\{|x_1\rangle, |x_2\rangle, \dots\rangle] = \mathcal{H}$*
2. $\forall |a\rangle \in \mathcal{H} : |a\rangle = \sum_i \langle |a\rangle, |x_i\rangle \rangle |x_i\rangle;$
3. $\forall |a\rangle, |b\rangle \in \mathcal{H} : \langle |a\rangle, |b\rangle \rangle = \sum_i \langle |a\rangle, |x_i\rangle \rangle \overline{\langle |b\rangle, |x_i\rangle \rangle}.$

Von Neumann então demonstra dois teoremas naturais porém muito importantes:

Teorema 8. *Para cada conjunto $\{|a_1\rangle, |a_2\rangle, \dots\rangle \subseteq \mathcal{H}$ existe um conjunto ortonormal $\{|\psi_1\rangle, |\psi_2\rangle, \dots\rangle \subseteq \mathcal{H}$ que gera a mesma variedade linear, ou seja, $[\{|a_1\rangle, |a_2\rangle, \dots\rangle] = [|\psi_1\rangle, |\psi_2\rangle, \dots\rangle]$*

Teorema 9. *Para cada variedade linear fechada \mathcal{I} existe um conjunto ortonormal $\{|\psi_1\rangle, |\psi_2\rangle, \dots\rangle$ que gera a mesma variedade linear fechada.*

3.3 Operadores Lineares

É de grande importância para o desenvolvimento da teoria, o conceito de operadores, particularmente, de operadores lineares. Von Neumann aborda [1, p. 87] esses conceitos de forma um pouco mais criteriosa, porém isso acarreta maiores complexidades no desenvolvimento rigoroso da teoria. Um operador é essencialmente um tipo de função que leva vetores em vetores, e vamos utilizar a justaposição para denotar a aplicação de um operador em um vetor, assim, a atuação do operador F no vetor $|a\rangle$ é denotada por $F|a\rangle$.

Escolhemos abordar primeiramente o conceito de operador linear para tornar o formalismo de von Neumann mais fácil de compreender e menos repetitivo. No seu livro, ele apresenta primeiramente o conceito de projetores, os quais já são operadores lineares. Buscamos também simplificar a sua exposição e torná-la mais compatível com a abordagem de Dirac

Como os operadores são funções, cada operador possui um domínio que é um subconjunto de \mathcal{H} . O domínio de um operador F qualquer é denotado por $\mathfrak{D}(F)$.

Definição 21. *Um operador F é uma função definida em um subconjunto $\mathfrak{D}(F)$ do espaço de Hilbert \mathcal{H} com valores em \mathcal{H} , assim,*

$$\forall |a\rangle \in \mathfrak{D}(F) : F|a\rangle \in \mathcal{H}.$$

Para o desenvolvimento da teoria, um tipo especial de operadores, os operadores lineares, são fundamentais:

Definição 22. *Um operador X é dito linear se o seu domínio $\mathfrak{D}(X)$ for uma variedade linear, isto é, se*

$$\forall |a\rangle, |b\rangle \in \mathfrak{D}(X), \alpha, \beta \in \mathbb{C} : \alpha|a\rangle + \beta|b\rangle \in \mathfrak{D}(X) \quad (3.21)$$

e se

$$\forall |a\rangle, |b\rangle \in \mathfrak{D}(X) : X(|a\rangle + |b\rangle) = X|a\rangle + X|b\rangle \quad (3.22)$$

$$\forall |a\rangle \in \mathfrak{D}(X) \alpha \in \mathbb{C} : X(\alpha|a\rangle) = \alpha X|a\rangle. \quad (3.23)$$

Como o domínio do operador deve ser uma variedade linear, então, se $X|a\rangle \in \mathcal{H}$ e $X|b\rangle \in \mathcal{H}$ temos também que $X(\alpha|a\rangle + \beta|b\rangle) \in \mathcal{H}$, sendo α e β números complexos. Observe que, sem essa condição, poderia acontecer de $\alpha|a\rangle + \beta|b\rangle$ não pertencer ao domínio do operador, e perderíamos essa propriedade.

Vamos agora limitar nosso estudo aos operadores lineares, mais ainda, vamos restringir

à operadores cujo domínio é denso em toda a parte. Com essas características definimos o conjunto \mathcal{L} como sendo o conjunto dos operadores lineares cujo domínio é denso em toda a parte. Esta última condição substitui e relaxa o requerimento de que os operadores devem ser definidos em todo o espaço de Hilbert, o qual deve ser abandonado na mecânica quântica. Como o domínio é denso em toda parte, cada ponto do espaço de Hilbert ou pertence ao domínio do operador ou está arbitrariamente próximo a algum ponto do domínio desse operador. Um domínio, por exemplo, que inclui todos os pontos do espaço de Hilbert exceto um único ponto, ou um conjunto finito de pontos, é denso em toda a parte. Assim temos os axiomas que caracterizam o domínio e o contra-domínio de um operador linear:

Axioma 7.

$$\forall X \in \mathcal{L} : \quad \forall |a\rangle \in \mathfrak{D}(X) : X|a\rangle \in \mathcal{H}, \quad (3.24)$$

$$\forall X \in \mathcal{L} : \quad \forall |a\rangle, |b\rangle \in \mathfrak{D}(X), \alpha, \beta \in \mathbb{C} : \quad \alpha|a\rangle + \beta|b\rangle \in \mathfrak{D}(X), \quad (3.25)$$

O domínio de qualquer operador linear $X \in \mathcal{L}$ é denso em \mathcal{H} , ou seja, é denso em toda a parte. Quando um operador tem como domínio o espaço de Hilbert inteiro, dizemos que o mesmo é definido em toda a parte

Duas funções são iguais se possuírem o mesmo domínio e dentro desse domínio, suas aplicações em quaisquer elementos resultarem no mesmo valor, ou seja, as funções f , e g são iguais se

$$\mathfrak{D}(f) = \mathfrak{D}(g), \text{ e } \forall x \in \mathfrak{D}(f) : f(x) = g(x).$$

Assim, acrescentamos um axioma explicitando essa condição de igualdade entre dois operadores lineares:

Definição 23.

$$\forall X, Y \in \mathcal{L} : \quad X = Y \iff (\mathfrak{D}(X) = \mathfrak{D}(Y), \text{ e } \forall |a\rangle \in \mathfrak{D}(X) : X|a\rangle = Y|a\rangle) \quad (3.26)$$

Em seu livro, von Neumann não menciona explicitamente essa condição para a igualdade de dois operadores, no entanto é importante mencioná-la para que possamos traçar um paralelo maior entre o formalismo de Dirac e de von Neumann.

Precisamos agora definir os dois operadores lineares especiais, o operador nulo 0 , e o operador identidade ou unidade $\hat{1}$, os quais são definidos em toda a parte. O operador nulo é definido por:

Definição 24.

$$\mathfrak{D}(0) = \mathcal{H}, \quad \forall |a\rangle \in \mathcal{H} : \quad 0|a\rangle = 0, \quad (3.27)$$

e o operador identidade é definido por

Definição 25.

$$\mathfrak{D}(\hat{\mathbf{1}}) = \mathcal{H}, \quad \forall |a\rangle \in \mathcal{H} : \hat{\mathbf{1}}|a\rangle = |a\rangle. \quad (3.28)$$

Esses dois operadores são únicos, isto é, não existe nenhum outro operador linear que satisfaça as mesmas propriedades.

Precisamos agora definir as operações básicas sobre os operadores. A soma de dois operadores é definida por

Axioma 8.

$$\forall X, Y \in \mathcal{L}, |a\rangle \in \mathfrak{D}(X) \cap \mathfrak{D}(Y) : (X + Y)|a\rangle = X|a\rangle + Y|a\rangle; \quad (3.29)$$

a multiplicação por escalar é definida por

Axioma 9.

$$\forall X \in \mathcal{L}, \alpha \in \mathbb{C}, |a\rangle \in \mathfrak{D}(X) : (\alpha X)|a\rangle = \alpha(X|a\rangle); \quad (3.30)$$

a multiplicação de dois operadores é definida por

Axioma 10.

$$\forall X, Y \in \mathcal{L}, |a\rangle \in \mathfrak{D}(Y) \text{ e } Y|a\rangle \in \mathfrak{D}(X) : (XY)|a\rangle = X(Y|a\rangle); \quad (3.31)$$

e o elemento inverso da adição é definido por

Axioma 11.

$$\forall X \in \mathcal{L}, |a\rangle \in \mathfrak{D}(X) : (-X)|a\rangle = -(X|a\rangle). \quad (3.32)$$

Comparando com a formulação de Dirac, observe que a questão dos domínios acrescenta um complexidade extra, e que o lado esquerdo dessas equações só está definido quando o lado direito também está definido.

Também define-se as potências inteiras:

$$\begin{aligned} \forall X \in \mathcal{L} : \quad X^0 &= \hat{\mathbf{1}} \\ X^1 &= X \\ X^2 &= XX \\ X^3 &= XXX \\ &\vdots \end{aligned}$$

Define-se o inverso da multiplicação X^{-1} por:

$$XX^{-1} = \hat{1}, \quad (3.33)$$

$$X^{-1}X = \hat{1}; \quad (3.34)$$

e analogamente as potências inteiras negativas:

$$X^{-2} = X^{-1}X^{-1}, \quad X^{-3} = X^{-1}X^{-1}X^{-1}, \dots \quad (3.35)$$

Como no caso dos números, temos que as potências se somam quando os fatores são multiplicados:

$$\forall X \in \mathcal{L}, \forall m, n \in \mathbb{Z} : \quad X^m X^n = X^{m+n}. \quad (3.36)$$

Um importante resultado para potências de operadores diz respeito a operadores que comutam: se dois operadores comutam, então suas potências também comutam, ou seja,

$$\forall X, Y \in \mathcal{L} : \quad XY = YX \implies \forall m, n \in \mathbb{Z} \quad X^m Y^n = Y^n X^m. \quad (3.37)$$

Vamos agora a outra importante definição, a definição de adjuntos:

Definição 26. *Dois operadores lineares X e X^* são ditos adjuntos se eles possuem o mesmo domínio, $\mathfrak{D}(X) = \mathfrak{D}(X^*)$, e se, nesse domínio*

$$\forall |a\rangle, |b\rangle \in \mathfrak{D}(X) : \quad \langle X|a\rangle, |b\rangle \rangle = \langle |a\rangle, X^*|b\rangle \rangle, \quad (3.38)$$

$$\langle X^*|a\rangle, |b\rangle \rangle = \langle |a\rangle, X|b\rangle \rangle. \quad (3.39)$$

Dessa definição decorre que essa relação é simétrica, ou seja, se X e X^* são adjuntos, então X^* e X também são adjuntos. Prova-se também que, se X é adjunto a X^*_1 e também a X^*_2 , então $X^*_1 = X^*_2$, ou seja, só existe um único adjunto associado à cada operador linear, e um elemento define univocamente o seu adjunto.

Assim podemos considerar a existência de uma função (*) que associa à cada operador linear X o seu adjunto X^* :

$$\forall X \in \mathcal{L} : \quad X^* \in \mathcal{L}, \quad \mathfrak{D}(X) = \mathfrak{D}(X^*), \quad (3.40)$$

$$\forall |a\rangle, |b\rangle \in \mathfrak{D}(X) : \quad \langle X|a\rangle, |b\rangle \rangle = \langle |a\rangle, X^*|b\rangle \rangle, \quad (3.41)$$

$$\langle X^*|a\rangle, |b\rangle \rangle = \langle |a\rangle, X|b\rangle \rangle. \quad (3.42)$$

Dessa definição decorre que essa operação é um involução, ou seja,

$$\forall X \in \mathcal{L} : X^{**} = X. \quad (3.43)$$

Demonstra-se também as seguintes propriedades, para a soma, produto, e multiplicação por escalar:

$$\forall X, Y \in \mathcal{L} : (X + Y)^* = X^* + Y^*, \quad (3.44)$$

$$\forall X, Y \in \mathcal{L} : (XY)^* = Y^* X^*, \quad (3.45)$$

$$\forall X \in \mathcal{L}, \alpha \in \mathbb{C} : (\alpha X)^* = \bar{\alpha} X^*. \quad (3.46)$$

Vamos agora a três definições importantes:

Definição 27. Um operador X é dito **hermitiano**¹ se $X^* = X$.

Definição 28. Um operador X é dito **definido**² se

$$\forall |a\rangle \in \mathfrak{D}(X) : \langle X|a\rangle, |a\rangle \rangle \geq 0.$$

Definição 29. Um operador U é dito **unitário** se

$$U^*U = UU^* = \hat{1}.$$

Dessa última definição, decorre que o produto interno de dois vetores nunca se altera pela atuação simultânea de um operador unitário nos dois vetores, ou seja, para um operador unitário U :

$$\forall |a\rangle, |b\rangle \in \mathcal{H} : \langle U|a\rangle, U|b\rangle \rangle = \langle U^*U|a\rangle, |b\rangle \rangle = \langle |a\rangle, |b\rangle \rangle. \quad (3.47)$$

Naturalmente, a norma de um vetor também não se altera pela atuação de um operador unitário U :

$$\forall |a\rangle \in \mathcal{H} : \|U|a\rangle\| = \||a\rangle\|. \quad (3.48)$$

¹Dirac refere-se à mesma definição como operador autoadjunto.

²Outros autores se referem a mesma definição como operador positivo

3.4 Continuidade de Operadores

A continuidade de operadores lineares já foi definida anteriormente, uma vez que operadores lineares são funções cujos valores são vetores do espaço de Hilbert. Assim, um operador X é contínuo no ponto $|a_0\rangle \in \mathfrak{D}(X)$ se, para cada $\varepsilon > 0$ existir um $\delta > 0$, tal que $\| |a\rangle - |a_0\rangle \| < \delta$ implica que $\| X|a\rangle - X|a_0\rangle \| < \varepsilon$. Matematicamente, X é contínuo no ponto $|a_0\rangle$ se:

$$\forall |a\rangle \in \mathfrak{D}(X), \varepsilon \in \mathbb{R} > 0 : \quad \exists \delta \in \mathbb{R} > 0 : \quad \| |a\rangle - |a_0\rangle \| < \delta \implies \| X|a\rangle - X|a_0\rangle \| < \varepsilon. \quad (3.49)$$

Porém verifica-se um importante teorema:

Teorema 10. *Um operador linear X é contínuo em toda a parte se ele for contínuo no ponto $|a\rangle = 0$.*

A demonstração é simples: continuidade no ponto $|a_0\rangle = 0$ significa

$$\forall |a\rangle \in \mathcal{H}, \varepsilon \in \mathbb{R} > 0 : \quad \exists \delta \in \mathbb{R} > 0 : \quad \| |a\rangle \| < \delta \implies \| X|a\rangle \| < \varepsilon.$$

Assim temos que:

$$\| |a\rangle - |a_0\rangle \| < \delta \implies \| X(|a\rangle - |a_0\rangle) \| = \| X|a\rangle - X|a_0\rangle \| < \varepsilon,$$

e portanto X é contínuo no ponto $|a_0\rangle$, e conseqüentemente, em toda a parte.

Von Neumann também demonstra que uma condição necessária e suficiente para que um operador X seja contínuo em toda a parte é a existência de uma constante real C tal que,

$$\forall |a\rangle \in \mathcal{H} : \quad \| X|a\rangle \| \leq C \| |a\rangle \|,$$

e, alternativamente,

$$\forall |a\rangle, |b\rangle \in \mathcal{H} : \quad |\langle X|a\rangle, |b\rangle| \leq C \| |a\rangle \| \cdot \| |b\rangle \|.$$

Essa primeira condição equivale à condição para um operador ser considerado limitado, então podemos dizer um operador é contínuo se, e somente se, ele for limitado.

3.5 Autovalores e Autovetores

Após discutir e caracterizar os operadores lineares, von Neumann discute [1, p. 102] o mais importante tipo de equação envolvendo essas entidades: as equações de autovalores, analisando uma de excepcional importância para a teoria: a equação de Schrödinger.

A equação de Schrödinger descreve os autoestados de um sistema físico e a sua evolução no tempo. Ela é dada por uma equação do tipo:

$$H|\psi\rangle = \lambda|\psi\rangle, \quad (3.50)$$

sendo que $|\psi\rangle \neq 0$, e H é o operador hermitiano correspondente à hamiltoniana do sistema físico. Cada vetor $|\psi\rangle$ que satisfaz essa equação é um autoestado desse sistema. Vamos utilizar as mesmas notações que Dirac para os autovalores e autovetores:

$$H|\psi^{(i)}\rangle = \lambda^{(i)}|\psi^{(i)}\rangle, \quad (3.51)$$

ou seja, $|\psi^{(i)}\rangle$ é o autovetor associado ao autovalor $\lambda^{(i)}$. Quando temos diferentes autovetores linearmente independente associados a um mesmo autovalor, acrescentamos um subíndice, ou seja,

$$\forall j : \quad H|\psi_m^{(i)}\rangle = \lambda^{(i)}|\psi_m^{(i)}\rangle. \quad (3.52)$$

Como H é hermitiano, temos que

$$\langle H|\psi\rangle, |\psi\rangle \rangle = \langle \lambda|\psi\rangle, |\psi\rangle \rangle = \lambda\langle |\psi\rangle, |\psi\rangle \rangle = \langle |\psi\rangle, H|\psi\rangle \rangle = \langle |\psi\rangle, \lambda|\psi\rangle \rangle = \langle |\psi\rangle, |\psi\rangle \rangle \bar{\lambda},$$

e portanto $\lambda = \bar{\lambda}$, ou seja, os autovalores são todos reais.

Precisamos ainda restringir a norma de nossas soluções a $\| |\psi\rangle \| = 1$, pois se $|\psi\rangle$ é uma solução com autovalor λ , então qualquer múltiplo $\alpha|\psi\rangle$, $\alpha \neq 0$ também é uma solução com o mesmo autovalor.

Outro resultado importante é que os autovetores associados a autovalores diferentes são ortogonais. Para verificar isso, basta fazer

$$\begin{aligned} \langle H|\psi^{(i)}\rangle, |\psi^{(j)}\rangle \rangle &= \lambda^{(i)}\langle |\psi^{(i)}\rangle, |\psi^{(j)}\rangle \rangle \\ \langle H|\psi^{(i)}\rangle, |\psi^{(j)}\rangle \rangle &= \langle |\psi^{(i)}\rangle, H|\psi^{(j)}\rangle \rangle = \lambda^{(j)}\langle |\psi^{(i)}\rangle, |\psi^{(j)}\rangle \rangle, \end{aligned}$$

igualando as duas equações temos e considerando que $\lambda^{(i)} \neq \lambda^{(j)}$:

$$\lambda^{(i)}\langle |\psi^{(i)}\rangle, |\psi^{(j)}\rangle \rangle = \lambda^{(j)}\langle |\psi^{(i)}\rangle, |\psi^{(j)}\rangle \rangle \implies \langle |\psi^{(i)}\rangle, |\psi^{(j)}\rangle \rangle = 0.$$

Associado a um mesmo autovalor $\lambda^{(i)}$ podemos ter diversos autovetores linearmente independentes. Esse número de autovetores linearmente independentes é denominado a multiplicidade do autovalor $\lambda^{(i)}$. Cada um desses conjuntos de autovetores linearmente independentes podem ser ortogonalizados resultando em um conjunto ortonormal, assim sendo, podemos considerar os diferentes $|\psi_m^{(i)}\rangle$ como sendo ortonormais, ou seja,

$$\langle \psi_m^{(i)} | \psi_n^{(j)} \rangle = \delta_{i,j} \delta_{m,n}.$$

Para que tenhamos consistência física, é necessário que o conjunto $\{\psi_1, \psi_2, \dots\}$ de todos os autoestados desse sistema seja capaz de gerar todo o espaço de Hilbert \mathcal{H} , ou ainda que, qualquer vetor $|\psi\rangle$ possa ser aproximado, com precisão arbitrária, por uma combinação linear de autoestados. A necessidade física disso advém do fato de qualquer condição inicial deve ser passível de ser representada em termos dos autoestados do sistema, para que assim possamos realizar a sua evolução temporal. Se os autoestados não forem capazes de gerar todo o espaço de Hilbert, então existiriam estados dentro do espaço de Hilbert que o sistema não admitiria como condição inicial.

Podemos assim afirmar que a variedade linear fechada gerada pelo conjunto $\{|\psi_n^{(j)}\rangle\}$ é o próprio espaço de Hilbert, e então esse é um conjunto ortonormal completo. Sendo assim, qualquer hamiltoniano deve possuir autoestados suficientes para a formação de um conjunto completo, e os operadores que não satisfazem essa condição não podem ser utilizados nesse formalismo. Essa condição de formar um conjunto completo corresponde à mesma consideração de Dirac segundo a qual os autovetores de um observável devem formar um conjunto completo, desta forma, na linguagem de Dirac, von Neumann está então dizendo que o hamiltoniano é um observável.

Von Neumann também afirma que, quando o conjunto ortonormal de autovetores do hamiltoniano não é completo, um “espectro contínuo de H existe”, e esses autovetores formam apenas a parte discreta do espectro. Essa afirmação é importante porque revela uma importante característica e limitação do formalismo de von Neumann: ele não é capaz de descrever sistemas de espectro contínuo. Isso revela uma limitação desse formalismo comparado ao formalismo de Dirac, o qual é capaz sim de descrever espectros contínuos; e prova que, apesar de von Neumann ser mais rigoroso, o seu formalismo não é tão abrangente quanto o formalismo de Dirac, e a delta de Dirac, tão criticada por von Neumann mostra-se como sendo o elemento fundamental para o tratamento de espectros contínuos, aspecto que ele deixa de abordar.

Na sequência, von Neumann analisa diversas peculiaridades do problema de autovalores, as quais não muito relevantes para a compreensão geral de sua obra, e portanto não serão aqui abordadas.

3.6 Variedades Lineares e Projetores

Von Neumann faz uma interessante abordagem da relação entre projetores e variedades lineares fechadas e um extenso tratamento de projetores, aos quais relativamente pouca importância é dada em livros modernos. A fim de definir o que seria uma projeção, definimos primeiramente:

Definição 30. *Seja $\mathcal{I} \subseteq \mathcal{H}$ uma variedade linear fechada, então definimos o conjunto $\mathcal{H} - \mathcal{I}$ como sendo o conjunto de todos os elementos de \mathcal{H} que são ortogonais a \mathcal{I} :*

$$\mathcal{H} - \mathcal{I} = \{|a\rangle \in \mathcal{H} : \forall |b\rangle \in \mathcal{I} : \langle |a\rangle, |b\rangle \rangle = 0\} \quad (3.53)$$

O conjunto $\mathcal{H} - \mathcal{I}$ é naturalmente um variedade linear fechada, e podemos agora apresentar o teorema chave para a definição de projeção:

Teorema 11. *Seja $\mathcal{I} \subseteq \mathcal{H}$ uma variedade linear fechada, então, qualquer vetor $|a\rangle$ pode ser separado de uma única maneira em duas componentes, uma pertencente a \mathcal{I} , e a outra a $\mathcal{H} - \mathcal{I}$, ou seja,*

$$\forall |a\rangle \in \mathcal{H} : \exists ! |p\rangle \in \mathcal{I}, |q\rangle \in \mathcal{H} - \mathcal{I} : |a\rangle = |p\rangle + |q\rangle. \quad (3.54)$$

O vetor $|p\rangle$ é denominado a projeção do vetor $|a\rangle$ em \mathcal{I} , e é único. Assim definimos a operação de projeção:

Definição 31. *A projeção $\mathcal{P}_{\mathcal{I}}|a\rangle$ do vetor $|a\rangle$ na variedade linear \mathcal{I} é definida pelas seguintes propriedades:*

$$\forall |a\rangle \in \mathcal{H} : \mathcal{P}_{\mathcal{I}}|a\rangle \in \mathcal{I}, \quad (3.55)$$

$$|a\rangle - \mathcal{P}_{\mathcal{I}}|a\rangle \in \mathcal{H} - \mathcal{I}. \quad (3.56)$$

Seja $\{\psi_1, \psi_2, \dots\}$ o conjunto ortonormal que gera a variedade linear fechada \mathcal{I} , então podemos obter a forma explícita de $\mathcal{P}_{\mathcal{I}}|a\rangle$:

$$\forall |a\rangle \in \mathcal{H} : \mathcal{P}_{\mathcal{I}}|a\rangle = \sum_i \langle |a\rangle, |\psi_i\rangle \rangle |\psi_i\rangle. \quad (3.57)$$

Assim, a operação $\mathcal{P}_{\mathcal{I}}|a\rangle$ leva cada vetor $|a\rangle$ do espaço de Hilbert \mathcal{H} à sua projeção em \mathcal{I} , portanto $\mathcal{P}_{\mathcal{I}}$ é um operador cujo domínio é todo o espaço de Hilbert \mathcal{H} , o chamado operador de projeção de \mathcal{I} . Para esses elementos definimos o conjunto dos projetores \mathbb{P} , assim, para qualquer variedade linear $\mathcal{I} \subseteq \mathcal{H}$

$$\mathcal{P}_{\mathcal{I}} \in \mathbb{P}$$

Os projetores são naturalmente operadores lineares, ou seja, o operador $\mathcal{P}_{\mathcal{I}}$ satisfaz a propriedade de linearidade:

$$\forall \alpha_1, \alpha_2 \in \mathbb{C}, \forall |a_1\rangle, |a_2\rangle \in \mathcal{H} : \quad \mathcal{P}_{\mathcal{I}}(\alpha_1 |a_1\rangle + \alpha_2 |a_2\rangle) = \mathcal{P}_{\mathcal{I}}(\alpha_1 |a_1\rangle) + \mathcal{P}_{\mathcal{I}}(\alpha_2 |a_2\rangle), \quad (3.58)$$

e tem como domínio todo o espaço de Hilbert, isto é, $\mathfrak{D}(\mathcal{P}_{\mathcal{I}}) = \mathcal{H}$, e portanto $\mathcal{P}_{\mathcal{I}} \in \mathcal{L}$. Os projetores também satisfazem:

$$\forall |a\rangle, |b\rangle \in \mathcal{H} : \quad \langle \mathcal{P}_{\mathcal{I}}|a\rangle, |b\rangle \rangle = \langle |a\rangle, \mathcal{P}_{\mathcal{I}}|b\rangle \rangle \quad (3.59)$$

$$\forall |a\rangle \in \mathcal{H} : \quad \mathcal{P}_{\mathcal{I}}\mathcal{P}_{\mathcal{I}}|a\rangle = \mathcal{P}_{\mathcal{I}}|a\rangle, \quad (3.60)$$

sendo portanto hermitianos e idempotentes.

Von Neumann demonstra ainda que uma variedade linear \mathcal{I} é completamente caracterizado pelo seu projetor $\mathcal{P}_{\mathcal{I}}$:

$$\mathcal{I} = \{ \mathcal{P}_{\mathcal{I}}|a\rangle : |a\rangle \in \mathcal{H} \}, \quad (3.61)$$

$$\mathcal{I} = \{ |a\rangle \in \mathcal{H} : \mathcal{P}_{\mathcal{I}}|a\rangle = |a\rangle \}, \quad (3.62)$$

$$(3.63)$$

o que revela a grande conexão entre variedades e projetores.

Ele então busca caracterizar um projetor independentemente da variedade linear ao qual ele se refere, considerando que um operador P definido em toda a parte é um projetor se existe uma variedade linear \mathcal{I} tal que $P = \mathcal{P}_{\mathcal{I}}$, temos:

Teorema 12. *Um operador P definido em toda a parte é um projetor se, e somente se, ele satisfaz as seguintes propriedades:*

$$\forall |a\rangle, |b\rangle \in \mathcal{H} : \quad \langle P|a\rangle, |b\rangle \rangle = \langle |a\rangle, P|b\rangle \rangle \quad (3.64)$$

$$P^2 = P \quad (3.65)$$

Esse projetor P gera a variedade linear \mathcal{I} :

$$\mathcal{I} = \{ P|a\rangle : |a\rangle \in \mathcal{H} \},$$

ou seja, $P = \mathcal{P}_{\mathcal{I}}$ e para qualquer elemento $|a\rangle \in \mathcal{I}$, $P|a\rangle = |a\rangle$.

Uma propriedade interessante é que a projeção nunca aumenta a norma de um vetor, ou seja,

$$\forall |a\rangle \in \mathcal{K}, \forall P \in \mathcal{P} : \quad \|P|a\rangle\| \leq \| |a\rangle \|. \quad (3.66)$$

Verifica-se que o operador identidade é um projetor especial, o projetor do próprio espaço de Hilbert \mathcal{H} :

$$\hat{\mathbf{1}} = \mathcal{P}_{\mathcal{H}} \quad (3.67)$$

Von Neumann então discute as condições para que a soma e o produto de projetores também seja um projetor. O produto de dois projetores é um projetor se, e somente se, os dois comutarem:

$$\forall P, Q \in \mathcal{P} : PQ \in \mathcal{P} \iff PQ = QP. \quad (3.68)$$

E a soma de dois projetores é um projetor se, e somente se, o produto de ambos for nulo:

$$\forall P, Q \in \mathcal{P} : P + Q \in \mathcal{P} \iff PQ = 0. \quad (3.69)$$

A subtração de dois projetores P e Q é um projetor apenas se $PQ = Q$:

$$\forall P, Q \in \mathcal{P} : P - Q \in \mathcal{P} \iff PQ = Q. \quad (3.70)$$

No seu extenso tratamento de projetores, von Neumann define duas importantes relações entre projetores (\leq e \geq) que são pouco divulgadas nos livros atuais:

$$\forall P, Q \in \mathcal{P} : P \leq Q \iff PQ = P \quad (3.71)$$

$$P \geq Q \iff PQ = Q. \quad (3.72)$$

Essas relações são relações de ordem parcial, isto é, satisfazem as seguintes propriedades:

$$\forall P, Q, R \in \mathcal{P} : P \leq Q \text{ e } Q \leq R \implies P \leq R \quad (3.73)$$

$$P \leq Q \text{ e } Q \leq P \implies P = Q \quad (3.74)$$

$$P \leq P, \quad (3.75)$$

e são intrinsecamente conectadas por:

$$\forall P, Q \in \mathcal{P} : P \leq Q \iff Q \geq P. \quad (3.76)$$

Temos ainda que:

$$\forall P, Q \in \mathcal{P} : P \leq Q \implies \hat{\mathbf{1}} - P \geq \hat{\mathbf{1}} - Q \quad (3.77)$$

$$\forall P \in \mathcal{P} : 0 \leq P \leq \hat{\mathbf{1}} \quad (3.78)$$

Essa relação de ordem é equivalente à seguinte condição:

Teorema 13.

$$\forall P, Q \in \mathcal{P} : P \leq Q \iff (\forall |a\rangle \in \mathcal{H} : \|P|a\rangle\| \leq \|Q|a\rangle\|) \quad (3.79)$$

3.7 O Traço

O traço de um operador ou uma matriz é um invariante, isto é, uma grandeza que não se altera com uma mudança de coordenadas ou uma transformação unitária. Para o caso de uma matriz $A_{i,j}$, o traço é dado por $\sum_{i=1}^n A_{i,i}$. Assim sendo, considerando a representação matricial de um operador utilizando conjunto completo ortonormal $\{|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_n\rangle\}$, temos que o traço de A será dado, em \mathcal{H}_n , por

$$Tr(A) = \sum_{i=1}^n \langle A|\psi_i\rangle, |\psi_i\rangle \rangle,$$

cujos valores são unitariamente invariantes, isto é, o valor é o mesmo para qualquer conjunto ortonormal completo.

Esse resultado pode ser naturalmente estendido para um espaço de Hilbert de dimensão infinita \mathcal{H}_∞ . Para isso tomamos qualquer conjunto ortonormal completo $\{|\psi_1\rangle, |\psi_2\rangle, \dots\}$ para o qual todos os $A|\psi_i\rangle$ estejam definidos, ou seja, $\forall i : |\psi_i\rangle \in \mathfrak{D}(A)$. Se o domínio de A for denso em toda a parte, basta ortogonalizarmos uma sequência que seja densa nesse domínio. Assim sendo, o traço de A fica definido como:

$$Tr(A) = \sum_{i=1}^{\infty} \langle A|\psi_i\rangle, |\psi_i\rangle \rangle, \quad (3.80)$$

e independe da escolha do conjunto ortonormal completo.

Para verificarmos essa última afirmação, tomamos um outro conjunto ortonormal completo $\{|\varphi_1\rangle, |\varphi_2\rangle, \dots\}$, e expandimos os vetores $|\psi_i\rangle$ do primeiro conjunto em termos dos elementos do novo conjunto:

$$|\psi_i\rangle = \sum_{j=1}^{\infty} \langle |\psi_i\rangle, |\varphi_j\rangle \rangle |\varphi_j\rangle,$$

substituindo na equação do traço, obtemos a igualdade desejada:

$$\begin{aligned}
Tr(A) &= \sum_{i=1}^{\infty} \langle A|\psi_i\rangle, \sum_{j=1}^{\infty} \langle |\psi_i\rangle, |\varphi_j\rangle \rangle |\varphi_j\rangle \rangle = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \overline{\langle |\psi_i\rangle, |\varphi_j\rangle \rangle} \langle A|\psi_i\rangle, |\varphi_j\rangle \rangle \\
&= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \langle A|\psi_i\rangle, |\varphi_j\rangle \rangle \langle |\varphi_j\rangle, |\psi_i\rangle \rangle = \sum_{j=1}^{\infty} \langle A \sum_{i=1}^{\infty} \langle |\varphi_j\rangle, |\psi_i\rangle \rangle |\psi_i\rangle, |\varphi_j\rangle \rangle \\
&= \sum_{j=1}^{\infty} \langle A|\varphi_j\rangle, |\varphi_j\rangle \rangle.
\end{aligned}$$

O traço é ainda uma operação linear:

$$\forall A \in \mathcal{L}, \alpha \in \mathbb{C} : \quad Tr(\alpha A) = \alpha Tr(A) \quad (3.81)$$

$$\forall A, B \in \mathcal{L} : \quad Tr(A + B) = Tr(A) + Tr(B), \quad (3.82)$$

e é invariante por permutação cíclica dos fatores, ou seja:

$$\forall A, B \in \mathcal{L} : \quad Tr(AB) = Tr(BA). \quad (3.83)$$

Essa é uma demonstração simples que convém ser apresentada pois nos permite verificar algumas das facilidades da notação de Dirac. Utilizando algumas passagens da demonstração anterior:

$$\begin{aligned}
Tr(AB) &= \sum_{i=1}^{\infty} \langle AB|\psi_i\rangle, |\psi_i\rangle \rangle = \sum_{i=1}^{\infty} \langle B|\psi_i\rangle, A^*|\psi_i\rangle \rangle \\
&= \sum_{i=1}^{\infty} \langle B|\psi_i\rangle, \sum_{j=1}^{\infty} \langle A^*|\psi_i\rangle, |\varphi_j\rangle \rangle |\varphi_j\rangle \rangle = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \overline{\langle A^*|\psi_i\rangle, |\varphi_j\rangle \rangle} \langle B|\psi_i\rangle, |\varphi_j\rangle \rangle \\
&= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \langle B|\psi_i\rangle, |\varphi_j\rangle \rangle \langle |\varphi_j\rangle, A^*|\psi_i\rangle \rangle = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \langle B|\psi_i\rangle, |\varphi_j\rangle \rangle \langle A|\varphi_j\rangle, |\psi_i\rangle \rangle \\
&= \sum_{j=1}^{\infty} \langle B \sum_{i=1}^{\infty} \langle A|\varphi_j\rangle, |\psi_i\rangle \rangle |\psi_i\rangle, |\varphi_j\rangle \rangle = \langle BA|\varphi_j\rangle, |\varphi_j\rangle \rangle \\
&= Tr(BA).
\end{aligned}$$

A fim de podermos comparar e verificar a elegância do formalismo de Dirac, vamos

apresentar a demonstração do mesmo resultado na notação de Dirac:

$$\begin{aligned}
 \text{Tr}(AB) &= \sum_{i=1}^{\infty} \langle \psi_i | AB | \psi_i \rangle = \sum_{i=1}^{\infty} \langle \psi_i | A \left(\sum_{j=1}^{\infty} |\varphi_j\rangle\langle\varphi_j| \right) B | \psi_i \rangle \\
 &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \langle \varphi_j | B | \psi_i \rangle \langle \psi_i | A | \varphi_j \rangle = \sum_{j=1}^{\infty} \langle \varphi_j | B \left(\sum_{i=1}^{\infty} |\psi_i\rangle\langle\psi_i| \right) A | \varphi_j \rangle \\
 &= \sum_{j=1}^{\infty} \langle \varphi_j | BA | \varphi_j \rangle.
 \end{aligned}$$

Vamos agora calcular o traço de um projetor. Seja $\mathcal{I} \subseteq \mathcal{H}$ uma variedade linear fechada, e $\mathcal{P}_{\mathcal{I}}$ o seu projetor; vamos agora considerar a sequência $\{|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_k\rangle\}$ como sendo a sequência ortonormal que gera \mathcal{I} , e $\{|x_1\rangle, |x_2\rangle, \dots, |x_l\rangle\}$ como sendo a sequência ortonormal geradora da variedade ortogonal $\mathcal{H}_{\infty} - \mathcal{I}$. Juntas, as duas sequências naturalmente geram o espaço de Hilbert \mathcal{H}_{∞} , sendo portanto um conjunto ortonormal completo. Podemos então utilizar essas duas sequências para calcular o traço do projetor:

$$\begin{aligned}
 \text{Tr}(\mathcal{P}_{\mathcal{I}}) &= \sum_{i=1}^k \langle \mathcal{P}_{\mathcal{I}} |\psi_i\rangle, |\psi_i\rangle \rangle + \sum_{i=1}^l \langle \mathcal{P}_{\mathcal{I}} |x_i\rangle, |x_i\rangle \rangle = \sum_{i=1}^k \langle |\psi_i\rangle, |\psi_i\rangle \rangle + \sum_{i=1}^l \langle 0, |x_i\rangle \rangle \\
 &= \sum_{i=1}^k 1 = k,
 \end{aligned} \tag{3.84}$$

ou seja, o traço de um projetor é o número de dimensões do espaço que ele gera. Por esse exemplo, é imediato que o traço de qualquer projetor de um espaço de dimensão infinita diverge, e que portanto, o traço de operadores muito bem definidos pode divergir.

Von Neumann então discute a convergência ou divergência de alguns tipos de operadores, mas é importante termos em mente que não existe uma prova geral de convergência do traço simplesmente porque ele não converge de forma geral, existem casos importantes em que o mesmo diverge.

O caso dos operadores definidos, ou seja, operadores X tal que $\forall |a\rangle \in \mathcal{H} : \langle X|a\rangle, |a\rangle \rangle \geq 0$, é interessante porque esses operadores sempre possuem um traço não negativo, ou seja $\text{Tr}(X) \geq 0$.

Estruturas Algébricas e Álgebra C^*

Na década de 30, uma série de paper de von Neumann e de Murray com von Neumann originou o estudo das álgebras de operadores em um Espaço de Hilbert. Eles analisaram em detalhes uma família de álgebras que é um caso particular da álgebra C^* : as álgebras de von Neumann ou álgebras W^* . Esses estudos culminaram em 1943 [14], por Gelfand e Naimark, numa caracterização abstrata de uma álgebra de operadores sem fazer referência a um espaço de Hilbert, a chamada álgebra C^* .

Essa pequena visão histórica nos permite perceber a essência do que é a álgebra C^* e um pouco de sua importância: a álgebra C^* é uma caracterização abstrata da álgebra dos operadores em um espaço de Hilbert que não faz referência ao espaço de Hilbert em si. A conexão com o espaço de Hilbert se dá pelo fato de que uma álgebra C^* admite uma representação em um espaço de Hilbert, mas a álgebra não carece necessariamente de um espaço de Hilbert. Esse é o conteúdo essencial do teorema de Gelfand-Naimark [14]: toda álgebra C^* é isomórfica a uma álgebra C^* dos operadores limitados de um espaço de Hilbert.

O formalismo da mecânica quântica via álgebra C^* difere-se essencialmente dos formalismos de Dirac e von Neumann pela ênfase dada: os formalismos de Dirac e von Neumann são totalmente estruturados sobre um espaço vetorial, e os operadores e outros conceitos são definidos em cima do conceito de vetor. Por outro lado, na álgebra C^* , a estrutura fundamental são os próprios operadores, os quais são tratados independente de termos ou não um espaço de Hilbert.

Atualmente, as principais aplicações da álgebra C^* na física se dão no tratamento de sistemas com um infinito número de partículas, especialmente nas teorias de campos. Nessa área de estudos mais avançados, a álgebra C^* é fundamental para a compreensão e obtenção de diversos resultados importantes, como por exemplo a existência de representações regulares não equivalentes [15].

Outra importância da álgebra C^* é a possibilidade de aplicação de seus resultados nos formalismos usuais da mecânica quântica, pois os operadores limitados de um espaço de Hilbert são uma álgebra C^* . Assim podemos aproveitar o vasto arsenal matemático existente a cerca das álgebras C^* e aplicá-lo na nos problemas da mecânica quântica usual.

As estruturas matemáticas, das quais as estruturas algébricas e as estruturas relacionais são casos particulares, são os objetos de estudo essenciais da álgebra abstrata, álgebra moderna, ou ainda simplesmente álgebra. A evolução da matemática é marcada por crescentes abstrações e melhorias na notação e simbolismo matemático, e nesse processo, por volta do início do século XX, teve início a inicialmente chamada álgebra moderna. Devido à tamanha importância dessa nova área, ela passou a ser chamada simplesmente álgebra, conferindo um novo significado ao termo.

A álgebra surgiu trazendo um novo patamar de abstração e rigor matemático, e é atualmente um dos ramos mais abstratos da matemática. Por meio da álgebra, a matemática passou a tratar objetos cada vez mais abstratos, objetos que poderiam até mesmo serem simplesmente inventados, não mais se limitando a objetos concretos. Os matemáticos passaram a estudar elementos e teorias mais gerais, estruturados na forma de sistemas axiomáticos. Isso permitiu tanto mais rigor como mais abstração, pois agora os objetos são definidos formalmente, e não precisam ser “construídos” em termos de objetos mais ‘concretos’. Nossa abordagem desse tema segue como referência [16].

4.1 Noções Fundamentais

Dois conceitos são fundamentais para o estudo da álgebra moderna, particularmente das estruturas matemáticas: operações e relações. Os conceitos de função e conjunto serão tomados como ponto de partida, e portanto são também fundamentais.

Vamos definir a seguinte notação para o que seria uma potência cartesiana:

Definição 1. *Seja um conjunto qualquer \mathbb{A} , e n um número natural, então $\mathbb{A}^n = \mathbb{A} \times \mathbb{A} \dots n$ vezes, sendo que $\mathbb{A} \times \mathbb{A}$ é o produto cartesiano do conjunto \mathbb{A} com o próprio conjunto \mathbb{A} .*

Definimos as operações n -árias, e os casos particulares, operações unárias e binárias:

Definição 2. *Uma **operação de aridade n** , ou ainda, uma **operação n -ária**, sobre um conjunto qualquer \mathbb{A} é qualquer função que leva n -uplas de elementos de \mathbb{A} em um elemento de \mathbb{A} ; ou seja qualquer função f da forma*

$$f : \mathbb{A}^n \longrightarrow \mathbb{A}$$

é uma operação n -ária sobre \mathbb{A} .

- *Uma operação de aridade 1 é chamada **operação unária**.*

- Uma operação de aridade 2 é chamada **operação binária**.

De acordo com a definição, a adição “+” e a multiplicação “.” são operações binárias sobre os números complexos, e também sobre os números reais.

Na sequência definimos as relações n -árias, e os casos particulares, relações unárias e binárias:

Definição 3. Uma **relação de aridade n** , ou ainda, uma **relação n -ária**, sobre um conjunto qualquer \mathbb{A} é qualquer subconjunto de \mathbb{A}^n ; ou seja, qualquer conjunto $R \subseteq \mathbb{A}^n$ é uma relação n -ária em \mathbb{A} .

- Uma relação de aridade 1 é chamada **relação unária**.
- Uma relação de aridade 2 é chamada **relação binária**.

Quando R é uma operação binária sobre \mathbb{A} , usamos a notação xRy , que lê-se “ x está R -relacionado a y ”, para indicar que $(x, y) \in R$, ou seja,

$$\forall x, y \in \mathbb{A} : \quad xRy \iff (x, y) \in R. \quad (4.1)$$

Vamos agora definir de forma geral o que é uma estrutura:

Definição 4. Seja \mathbb{A} um conjunto, \mathfrak{F} uma coleção de operações em \mathbb{A} , e \mathfrak{R} uma coleção de relações em \mathbb{A} , então a tripla $\langle \mathbb{A}, \mathfrak{F}, \mathfrak{R} \rangle$ é uma **estrutura sobre \mathbb{A}** .

Observe que, tanto \mathfrak{F} como \mathfrak{R} podem ser vazios. Como exemplo, $\langle \mathbb{R}, \{+, \cdot\}, \{\geq, \leq\} \rangle$ é uma estrutura sobre os números reais.

As estruturas algébricas aparecem no caso particular em que não temos relações:

Definição 5. Seja \mathbb{A} um conjunto e \mathfrak{F} uma coleção de operações em \mathbb{A} , então o par $\langle \mathbb{A}, \mathfrak{F} \rangle$ é uma **estrutura algébrica sobre \mathbb{A}** . Para uma coleção de operações $\mathfrak{F} = \{o_1, o_2, \dots, o_n\}$ utilizaremos alternativamente a notação $\langle \mathbb{A}; o_1, o_2, \dots, o_n \rangle$ para representar a estrutura algébrica.

Como exemplo, temos que $\langle \mathbb{C}; +, \cdot \rangle$ é uma estrutura sobre os números complexos.

Quando, pelo contrário, não temos operações, apenas relações, a estrutura é dita uma estrutura relacional:

Definição 6. Seja \mathbb{A} um conjunto e \mathfrak{R} uma coleção de relações em \mathbb{A} , então o par $\langle \mathbb{A}, \mathfrak{R} \rangle$ é uma **estrutura relacional sobre \mathbb{A}** .

Frequentemente utilizamos uma notação especial no caso das operações binárias, a chamada notação mesofixa. Seja ψ uma operação binária em \mathbb{A} , então podemos denotar a aplicação de ψ no par (a, b) por $a\psi b$, alternativamente à notação usual $\psi(a, b)$. Sendo assim, denotamos $+(a, b)$ por $a + b$, e denotamos $\cdot(a, b)$ por $a \cdot b$.

Passamos agora à definição de algumas das principais propriedades que uma operação binária pode satisfazer: comutatividade e associatividade, e de uma propriedade que envolve duas operações binárias: distributividade.

Definição 7. Uma operação binária χ sobre um conjunto \mathbb{A} é dita **comutativa**, ou satisfaz a comutatividade se:

$$\forall a, b \in \mathbb{A} : a\chi b = b\chi a.$$

Definição 8. Uma operação binária χ sobre um conjunto \mathbb{A} é dita **associativa**, ou satisfaz a associatividade se:

$$\forall a, b, c \in \mathbb{A} : a\chi(b\chi c) = (a\chi b)\chi c.$$

Definição 9. Uma operação binária χ_1 sobre um conjunto \mathbb{A} é dita **distributiva** sobre uma outra operação binária χ_2 sobre o mesmo conjunto se:

$$\begin{aligned} \forall a, b, c \in \mathbb{A} : \quad a\chi_1(b\chi_2 c) &= (a\chi_1 b)\chi_2(a\chi_1 c) \\ (a\chi_2 b)\chi_1 c &= (a\chi_1 c)\chi_2(b\chi_1 c). \end{aligned}$$

Outras duas propriedades importantes de operações binárias são a de existência elemento neutro ou identidade, e de elemento inverso:

Definição 10. Dizemos que uma operação binária χ sobre um conjunto \mathbb{A} possui um **elemento neutro** (ou **identidade**) $\psi \in \mathbb{A}$ se:

$$\forall a \in \mathbb{A} : a\chi\psi = a = \psi\chi a.$$

Definição 11. Dizemos que uma operação binária χ sobre um conjunto \mathbb{A} , que possui um elemento neutro (ou identidade) $\psi \in \mathbb{A}$, possui os **elementos inversos** se:

$$\forall a \in \mathbb{A} : \exists a' \in \mathbb{A} : a\chi a' = \psi = a'\chi a.$$

4.2 Estruturas Algébricas

Vamos agora enunciar as principais estruturas algébricas em conexão com este trabalho. Por razões pedagógicas, buscamos enunciar cada definição expondo cada propriedade ou

axioma em vez de utilizar as definições anteriores.

4.2.1 Grupos, Semigrupos e Monoides

Grupos, semigrupos e monoides são algumas das estruturas algébricas mais simples e forma uma cadeia: semigrupos \subset monoides \subset grupos. Começando da mais simples, temos a definição de semi-grupo:

Definição 12. Um *semi-grupo* é uma estrutura algébrica $\langle \mathbb{G}; \cdot : \mathbb{G}^2 \longrightarrow \mathbb{G} \rangle$ (um conjunto \mathbb{G} munido de uma operação binária “ \cdot ”) que satisfaz os seguinte axioma:

$$\forall a, b, c \in \mathbb{G} : \quad a \cdot (b \cdot c) = (a \cdot b) \cdot c \quad \text{Associatividade} \quad (4.2)$$

Ou seja, um semi-grupo é simplesmente uma estrutura algébrica com uma operação binária que é associativa. Acrescentando a existência de um elemento neutro, temos o chamado Monoide:

Definição 13. Um *monóide* é uma estrutura algébrica $\langle \mathbb{G}; \cdot : \mathbb{G}^2 \longrightarrow \mathbb{G} \rangle$ (um conjunto \mathbb{G} munido de uma operação binária “ \cdot ”) que satisfaz os seguintes axiomas:

$$\exists e \in \mathbb{G} : \forall a, b, c \in \mathbb{G} :$$

$$a \cdot (b \cdot c) = (a \cdot b) \cdot c \quad \text{Associatividade} \quad (4.3)$$

$$e \cdot a = a = a \cdot e \quad \text{Elemento neutro ou identidade} \quad (4.4)$$

Proposição 1. Em todo monóide o elemento neutro é único.

Acrescentando a propriedade de existência de um elemento inverso, obtemos a conhecida estrutura algébrica dos grupos:

Definição 14. Um *grupo* é uma estrutura algébrica $\langle \mathbb{G}; \cdot : \mathbb{G}^2 \longrightarrow \mathbb{G} \rangle$ (um conjunto \mathbb{G} munido de uma operação binária “ \cdot ”) que satisfaz os seguintes axiomas:

$$\exists e \in \mathbb{G}, \forall a, b, c \in \mathbb{G} :$$

$$a \cdot (b \cdot c) = (a \cdot b) \cdot c \quad \text{Associatividade} \quad (4.5)$$

$$e \cdot a = a = a \cdot e \quad \text{Elemento neutro ou identidade} \quad (4.6)$$

$$\exists a^{-1} \in \mathbb{G} : \quad aa^{-1} = e = a^{-1}a \quad \text{Elemento inverso} \quad (4.7)$$

Proposição 2. Em todo grupo, o elemento inverso para um dado elemento do grupo é único.

Acrescentando a propriedade de comutatividade, temos o chamado grupo abeliano ou

grupo comutativo:

Definição 15. Um **grupo abeliano ou comutativo** é uma estrutura algébrica $\langle \mathbb{G}; \cdot : \mathbb{G}^2 \longrightarrow \mathbb{G} \rangle$ (um conjunto \mathbb{G} munido de uma operação binária “ \cdot ”) que satisfaz os seguintes axiomas:

$$\exists e \in \mathbb{G}, \forall a, b, c \in \mathbb{G} \quad (a \cdot b) \cdot c = a \cdot (b \cdot c) \quad \text{Associatividade} \quad (4.8)$$

$$a \cdot b = b \cdot a \quad \text{Comutatividade} \quad (4.9)$$

$$e \cdot a = a = a \cdot e \quad \text{Elemento neutro ou identidade} \quad (4.10)$$

$$\exists a^{-1} \in \mathbb{G} : \quad a a^{-1} = e = a^{-1} a \quad \text{Elemento inverso} \quad (4.11)$$

Como exemplo temos que os diversos conjuntos de números (com exceção dos naturais) são grupos abelianos com relação à adição: $\langle \mathbb{Z}; + \rangle$, $\langle \mathbb{Q}; + \rangle$, $\langle \mathbb{R}; + \rangle$ e $\langle \mathbb{C}; + \rangle$ são grupos abelianos. Considerando que \mathbb{Q}_* , \mathbb{R}_* e \mathbb{C}_* são respectivamente os conjuntos dos números racionais, reais e complexos excluindo o zero, temos que eles são grupos abelianos com relação à multiplicação: $\langle \mathbb{Q}_*; \cdot \rangle$, $\langle \mathbb{R}_*; \cdot \rangle$ e $\langle \mathbb{C}_*; \cdot \rangle$ são grupos abelianos.

Observe que todo grupo abeliano é um grupo, todo grupo é um monoide, e todo monoide é um semi-grupo.

4.2.2 Anéis

Quando passamos a ter duas operações binárias, uma das estruturas mais importantes é o chamado anel:

Definição 16. Um **anel** é uma estrutura algébrica $\langle \mathbb{G}; \cdot : \mathbb{G}^2 \longrightarrow \mathbb{G}, + : \mathbb{G}^2 \longrightarrow \mathbb{G} \rangle$ (um conjunto \mathbb{G} munido de uma operação binária “ \cdot ” denominada multiplicação e outra operação binária “ $+$ ” denominada adição) que satisfaz os seguintes axiomas:

$$\exists 0 \in \mathbb{G}, \forall a, b, c \in \mathbb{G} \quad (a + b) + c = a + (b + c) \quad \text{Associatividade da adição} \quad (4.12)$$

$$a + b = b + a \quad \text{Comutatividade da adição} \quad (4.13)$$

$$0 + a = a = a + 0 \quad \text{Elemento neutro ou identidade da adição} \quad (4.14)$$

$$\exists -a \in \mathbb{G} : \quad a + (-a) = 0 = (-a) + a \quad \text{Elemento inverso da adição} \quad (4.15)$$

$$a \cdot (b \cdot c) = (a \cdot b) \cdot c \quad \text{Associatividade da multiplicação} \quad (4.16)$$

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad \text{Distributividade da mult. sobre a adição} \quad (4.17)$$

$$(a + b) \cdot c = a \cdot c + b \cdot c \quad \text{Distributividade da mult. sobre a adição} \quad (4.18)$$

Adicionando a propriedade de existência de um elemento identidade da multiplicação, obtemos o chamado anel com identidade:

Definição 17. *Um anel com identidade é uma estrutura algébrica $\langle \mathbb{G}; \cdot : \mathbb{G}^2 \rightarrow \mathbb{G}, + : \mathbb{G}^2 \rightarrow \mathbb{G} \rangle$ (um conjunto \mathbb{G} munido de uma operação binária “ \cdot ” denominada multiplicação e outra operação binária “ $+$ ” denominada adição) que satisfaz os seguintes axiomas:*

$$\exists 1, 0 \in \mathbb{G}, \forall a, b, c \in \mathbb{G} :$$

$$a + (b + c) = (a + b) + c \quad \text{Associatividade da adição} \quad (4.19)$$

$$a + b = b + a \quad \text{Comutatividade da adição} \quad (4.20)$$

$$0 + a = a = a + 0 \quad \text{Elemento neutro ou identidade da adição} \quad (4.21)$$

$$\exists -a \in \mathbb{G} : a + (-a) = 0 = (-a) + a \quad \text{Elemento inverso da adição} \quad (4.22)$$

$$a \cdot (b \cdot c) = (a \cdot b) \cdot c \quad \text{Associatividade da multiplicação} \quad (4.23)$$

$$a \cdot 1 = a = 1 \cdot a \quad \text{Elemento identidade da multiplicação} \quad (4.24)$$

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad \text{Distributividade da mult. sobre a adição} \quad (4.25)$$

$$(a + b) \cdot c = a \cdot c + b \cdot c \quad \text{Distributividade da mult. sobre a adição} \quad (4.26)$$

Como exemplo, temos que os diversos conjuntos de números (com exceção dos naturais) são anéis com identidade, ou seja, $\langle \mathbb{Z}; \cdot, + \rangle$, $\langle \mathbb{Q}; \cdot, + \rangle$, $\langle \mathbb{R}; \cdot, + \rangle$ e $\langle \mathbb{C}; \cdot, + \rangle$ são anéis com identidade. E ainda, as matrizes quadradas $m \times m$ são anéis com identidade.

Observe que, se $\langle \mathbb{G}; \cdot, + \rangle$ é um anel, então

- $\langle \mathbb{G}; \cdot \rangle$ é um semigrupo;
- $\langle \mathbb{G}; + \rangle$ é um grupo abeliano;

e ainda que se $\langle \mathbb{G}; \cdot, + \rangle$ é um anel com identidade, então

- $\langle \mathbb{G}; \cdot \rangle$ é um monoide;
- $\langle \mathbb{G}; + \rangle$ é um grupo abeliano.

4.2.3 Corpos

Avançando um pouco mais, chegamos à estrutura algébrica chamada corpo ou campo, que é de extrema importância pois caracteriza as principais propriedades dos números racionais, reais e complexos:

Definição 18. Um *corpo* é uma estrutura algébrica $\langle \mathbb{K}; \cdot : \mathbb{K}^2 \rightarrow \mathbb{K}, + : \mathbb{K}^2 \rightarrow \mathbb{K} \rangle$ (um conjunto \mathbb{A} munido de uma operação binária “ \cdot ” denominada multiplicação e outra operação binária “ $+$ ” denominada adição) que satisfaz os seguintes axiomas:

$$\exists 1, 0 \in \mathbb{K}, \forall a, b, c \in \mathbb{A} :$$

$$a + (b + c) = (a + b) + c \quad \text{Associatividade da adição} \quad (4.27)$$

$$a + b = b + a \quad \text{Comutatividade da adição} \quad (4.28)$$

$$0 + a = a \quad \text{Elemento neutro ou identidade da adição} \quad (4.29)$$

$$\exists -a \in \mathbb{K} : a + (-a) = 0 \quad \text{Elemento inverso da adição} \quad (4.30)$$

$$a \cdot (b \cdot c) = (a \cdot b) \cdot c \quad \text{Associatividade da multiplicação} \quad (4.31)$$

$$a \cdot b = b \cdot a \quad \text{Comutatividade da multiplicação} \quad (4.32)$$

$$a \cdot 1 = a \quad \text{Elemento identidade da multiplicação} \quad (4.33)$$

$$a \neq 0 \implies \exists a^{-1} \in \mathbb{G} : aa^{-1} = 1 \quad \text{Elemento inverso da multiplicação} \quad (4.34)$$

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad \text{Distributividade da mult. sobre a adição} \quad (4.35)$$

Os mais clássicos exemplos de corpos são os números racionais, reais, e complexos, ou seja, $\langle \mathbb{Q}; \cdot, + \rangle, \langle \mathbb{R}; \cdot, + \rangle$ e $\langle \mathbb{C}; \cdot, + \rangle$ são corpos. Observe que $\langle \mathbb{N}; \cdot, + \rangle$ e $\langle \mathbb{Z}; \cdot, + \rangle$ não são corpos.

Nessa cadeia de estruturas algébricas, temos que todo corpo $\langle \mathbb{K}; \cdot, + \rangle$ é também um anel com identidade.

4.2.4 Espaços Vetoriais e Álgebras

Espaços vetoriais e álgebras são conceitos um pouco mais complexos que envolvem dois conjuntos, e não somente um como consta na definição de estruturas algébricas, e envolvem funções binárias, e não somente operações.

4.2.4.1 Espaços Vetoriais

Um espaço vetorial ou uma álgebra é definido em termos de um outro conjunto, assim sendo, define-se um “espaço vetorial sobre o conjunto \mathbb{K} ” e não simplesmente um “espaço vetorial”.

Definição 19. Um *espaço vetorial sobre o conjunto* \mathbb{K} é uma tripla $\langle \mathbb{V}; \cdot : \mathbb{K} \times \mathbb{V} \rightarrow \mathbb{V}, + : \mathbb{V}^2 \rightarrow \mathbb{V} \rangle$ em que $\mathbb{K} \times \mathbb{V} \rightarrow \mathbb{V}$ é denominado produto por escalares, e a operação $+ : \mathbb{V}^2 \rightarrow \mathbb{V}$ é denominada soma vetorial, sendo que $\langle \mathbb{K}; \cdot, + \rangle$ é um corpo, e os seguintes

$$u + (v + x) = (u + v) + x \quad \text{Associatividade da adição (4.36)}$$

$$u + v = v + u \quad \text{Comutatividade da adição (4.37)}$$

$$\exists 0 \in \mathbb{V}, \forall \alpha, \beta \in \mathbb{K}, \forall u, v, x \in \mathbb{V} : \quad \text{Elemento neutro ou identidade da adição (4.38)}$$

$$\exists -u \in \mathbb{V} : \quad u + (-u) = 0 \quad \text{Elemento inverso da adição (4.39)}$$

$$\alpha \cdot (\beta \cdot u) = (\alpha \cdot \beta) \cdot u \quad \text{Associatividade da mult. por escalar (4.40)}$$

$$1 \cdot u = u \quad \text{Elemento identidade da mult. por escalar (4.41)}$$

$$\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v \quad \text{Distributividade da mult. sobre a adição (4.42)}$$

$$(\alpha + \beta) \cdot u = \alpha \cdot u + \beta \cdot u \quad \text{Distributividade da mult. sobre a adição (4.43)}$$

em que 1 é a unidade do corpo \mathbb{K} .

Se $\langle \mathbb{V}; \cdot, + \rangle$ é um espaço vetorial sobre algum corpo \mathbb{K} , então $\langle \mathbb{V}; + \rangle$ é um grupo abeliano.

4.2.4.2 Álgebras

Acrescentando a multiplicação ou produto e as propriedades convenientes a um espaço vetorial obtemos uma álgebra. Como nenhuma ambiguidade maior resulta de utilizarmos a mesma notação para o produto de dois elementos da álgebra e o produto por um escalar, e é comum na prática matemática utilizarmos a mesma notação para ambos (a simples justaposição dos símbolos), assim vamos, para simplificar, utilizar a mesma notação “ \cdot ” para ambas as funções.

Definição 20. Uma *álgebra sobre o conjunto* \mathbb{K} é um conjunto \mathbb{V} , munido das funções $\cdot : \mathbb{K} \times \mathbb{V} \rightarrow \mathbb{V}$ denominada produto por escalares, $\cdot : \mathbb{V}^2 \rightarrow \mathbb{V}$ denominada produto da álgebra, e $+$: $\mathbb{V}^2 \rightarrow \mathbb{V}$ denominada adição vetorial, sendo que $\langle \mathbb{K}; \cdot, + \rangle$ é um corpo e os seguintes axiomas são satisfeitos:

$$\exists 0 \in \mathbb{V}, \forall \alpha, \beta \in \mathbb{K}, \forall u, v, x \in \mathbb{V} :$$

$$u + (v + x) = (u + v) + x \quad \text{Associatividade da adição (4.44)}$$

$$u + v = v + u \quad \text{Comutatividade da adição (4.45)}$$

$$0 + u = u \quad \text{Elemento neutro ou identidade da adição (4.46)}$$

$$\exists -u \in \mathbb{V} : u + (-u) = 0 \quad \text{Elemento inverso da adição (4.47)}$$

$$\alpha \cdot (\beta \cdot u) = (\alpha \cdot \beta) \cdot u \quad \text{Associatividade da mult. por escalar (4.48)}$$

$$1 \cdot u = u \quad \text{Identidade da mult. por escalar (4.49)}$$

$$\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v \quad \text{Distributividade da mult. sobre a adição (4.50)}$$

$$(\alpha + \beta) \cdot u = \alpha \cdot u + \beta \cdot u \quad \text{Distributividade da mult. sobre a adição (4.51)}$$

$$u \cdot (v + x) = u \cdot v + u \cdot x \quad \text{Distributividade da mult. sobre a adição (4.52)}$$

$$(u + v) \cdot x = u \cdot x + v \cdot x \quad \text{Distributividade da mult. sobre a adição (4.53)}$$

$$(\alpha \cdot \beta) \cdot (u \cdot v) = (\alpha \cdot u) \cdot (\beta \cdot v) \quad (4.54)$$

em que 1 é a unidade do corpo \mathbb{K} .

Estendendo a definição de álgebra, temos as álgebras associativas, abelianas, e unitais que serão definidas de forma simplificada:

Definição 21. Uma álgebra \mathbb{V} é dita uma **álgebra associativa** se

$$\forall u, v, x \in \mathbb{V} : u \cdot (v \cdot x) = (u \cdot v) \cdot x.$$

Definição 22. Uma álgebra \mathbb{V} é dita uma **álgebra abeliana ou comutativa** se a multiplicação da álgebra é comutativa, ou seja,

$$\forall u, v \in \mathbb{V} : u \cdot v = v \cdot u.$$

Definição 23. Uma álgebra \mathbb{V} é dita uma **álgebra unital** se existir um elemento identidade da multiplicação:

$$\exists 1_{\mathbb{V}} \in \mathbb{V} : \forall u \in \mathbb{V} : u \cdot 1_{\mathbb{V}} = u = 1_{\mathbb{V}} \cdot u.$$

4.2.4.3 Álgebras-*

A fim de estudarmos a álgebra C^* , uma outra estrutura algébrica de fundamental importância é a álgebra-*, ou álgebra com involução. Vamos nos limitar à definição de álgebras-* sobre os números complexos, e considerar a operação $\alpha \rightarrow \bar{\alpha}$ como sendo a conjugação do número complexo α , assim sendo definimos:

Definição 24. Uma álgebra associativa \mathbb{V} sobre o corpo dos complexos \mathbb{C} é dita ter uma

involução ou ser uma **álgebra-*** se existir uma operação unária $*$: $\mathbb{V} \rightarrow \mathbb{V}$ denominada involução que satisfaz as seguintes propriedades:

$$\forall a, b \in \mathbb{V}, \forall \alpha, \beta \in \mathbb{V} : \quad (a^*)^* = a \quad (4.55)$$

$$(ab)^* = b^* a^* \quad (4.56)$$

$$(\alpha \cdot a + \beta \cdot b)^* = \bar{\alpha} \cdot a^* + \bar{\beta} \cdot b^* \quad (4.57)$$

Definição 25. Quando um elemento x de álgebra- $*$ \mathbb{V} satisfaz $x^* = x$ dizemos que x é **auto-adjunto**.

Um resultado fundamental é que quando uma álgebra- $*$ \mathbb{V} possui uma identidade $1_{\mathbb{V}}$, necessariamente essa identidade é autoadjunta, ou seja,

$$1_{\mathbb{V}}^* = 1_{\mathbb{V}}.$$

4.3 Espaços Métricos, Norma e Produto Interno

Vamos agora nos restringir a espaços vetoriais e álgebras sobre o corpo dos complexos e definir algumas noções importantes da álgebra linear. Nesse ponto partimos para conceitos um pouco mais conhecidos e frequentemente omitiremos o símbolo da multiplicação “ \cdot ”.

4.3.1 Métrica

Primeiro definiremos uma métrica que é uma função que define a ‘distância’ entre dois pontos de algum conjunto, de forma mais precisa, ela capta as propriedades essenciais da nossa noção intuitiva de distância entre dois pontos:

Definição 26. Uma **métrica** em um conjunto \mathbb{V} é uma função $d : \mathbb{V} \times \mathbb{V} \rightarrow \mathbb{R}$ que satisfaz os seguintes axiomas:

$$\forall u, v, w \in \mathbb{V}$$

$$d(u, v) \geq 0 \quad \text{Não-negatividade} \quad (4.58)$$

$$d(u, v) = 0 \iff u = v \quad (4.59)$$

$$d(u, v) = d(v, u) \quad \text{Simetria} \quad (4.60)$$

$$d(u, v) \leq d(u, w) + d(w, v) \quad \text{Desigualdade triangular} \quad (4.61)$$

Vamos agora uma outra importante definição, espaços métricos:

Definição 27. Se X é um conjunto não vazio e d é uma métrica em X , o par $\langle X, d \rangle$ é denominado um **espaço métrico**.

4.3.2 Norma

A definição de métrica é muito geral e não se limita a espaços vetoriais. Conectada esse conceito temos a definição de norma sobre um espaço vetorial:

Definição 28. Uma **norma** sobre um espaço vetorial \mathbb{V} definido sobre o corpo dos complexos \mathbb{C} é uma função $\| \cdot \| : \mathbb{V} \rightarrow \mathbb{R}$ que satisfaz as seguintes propriedades:

$$\forall u, v \in \mathbb{V}, \forall \alpha \in \mathbb{C} : \quad \|u\| \geq 0 \quad \text{Não-negatividade} \quad (4.62)$$

$$\|u\| = 0 \iff u = 0 \quad (4.63)$$

$$\|\alpha u\| = |\alpha| \|u\| \quad (4.64)$$

$$\|u + v\| \leq \|u\| + \|v\| \quad \text{Desigualdade triangular} \quad (4.65)$$

Definição 29. Um espaço vetorial que possui uma norma é dito um **espaço vetorial normado**.

Todo espaço vetorial normado possui naturalmente uma métrica especial, a métrica induzida pela norma:

Definição 30. A **métrica induzida pela norma** $\| \cdot \|$ em um espaço vetorial \mathbb{V} é dada por:

$$\forall u, v \in \mathbb{V}; \quad d(u, v) = \|u - v\|$$

A métrica d induzida pela norma $\| \cdot \|$ de um espaço vetorial normado \mathbb{V} satisfaz:

$$\forall u, v \in \mathbb{V}, \alpha \in \mathbb{C} :$$

$$d(u+x, v+x) = d(u, v) \quad \text{Invariância Translacional} \quad (4.66)$$

$$d(\alpha u, \alpha v) = |\alpha|d(u, v) \quad \text{Transformação de Escala,} \quad (4.67)$$

E essas duas propriedades, quando satisfeitas, garantem que a métrica d é induzida por uma norma, e essa norma satisfaz $\|u\| = d(u, 0)$.

4.3.3 Produto Interno

Outra definição importante é a de produto escalar ou produto interno:

Definição 31. *Um produto escalar ou produto interno em um espaço vetorial \mathbb{V} definido sobre o corpo dos complexos \mathbb{C} é uma função $\langle \cdot, \cdot \rangle : \mathbb{V} \times \mathbb{V} \rightarrow \mathbb{C}$ que satisfaz as seguintes propriedades:*

$$\forall u, v, w \in \mathbb{V}, \forall \alpha, \beta \in \mathbb{C} : \\ \langle u, \alpha v + \beta w \rangle = \alpha \langle u, v \rangle + \beta \langle u, w \rangle \quad \text{Linearidade na segunda variável} \quad (4.68)$$

$$\langle u, v \rangle = \overline{\langle v, u \rangle} \quad \text{Conjugação complexa} \quad (4.69)$$

$$u \neq 0 \implies \langle u, u \rangle > 0 \quad \text{Positividade} \quad (4.70)$$

Um produto interno também define naturalmente uma norma, a norma induzida pelo produto interno:

Definição 32. *A norma $\|\cdot\|$ induzida pelo produto interno $\langle \cdot, \cdot \rangle$ em um espaço vetorial \mathbb{V} é dada por:*

$$\forall u \in \mathbb{V} : \quad \|u\| = \sqrt{\langle u, u \rangle}.$$

Assim sendo, todo espaço vetorial normado possui uma distância e forma um espaço métrico; e todo espaço vetorial com produto interno, é também um espaço vetorial normado e forma um espaço métrico.

4.3.4 Convergência e Completeza em Espaços Métricos

Espaços métricos são importantes porque dentro deles podemos tratar de distâncias e convergência, para isso definiremos sequências e limite de sequências:

Definição 33. *Se \mathbb{X} é um conjunto não-vazio, uma função $a : \mathbb{N} \rightarrow \mathbb{X}$ é dita ser uma sequência em \mathbb{X} . Utiliza-se usualmente a notação a_n para denotar o valor de a no ponto n , e*

a notação $\{a_n : n \in \mathbb{N}\}$ ou ainda, numa abuso de linguagem, simplesmente a_n para denotar a sequência a .

4.3.4.1 Limite e Convergência

Definição 34. *Seja $\langle \mathbb{X}, d \rangle$ um espaço métrico; dizemos que uma sequência $\{a_n : n \in \mathbb{N}\}$ em \mathbb{X} converge para um elemento (ou tem limite) $x \in \mathbb{X}$ em relação à métrica d se para qualquer $\varepsilon > 0$ existir um número natural N tal que $d(x, a_n) < \varepsilon$ para todo $n > N$; simbolicamente,*

$$\lim_{n \rightarrow \infty} a_n = x \iff \forall \varepsilon \in \mathbb{R} > 0 : \exists N \in \mathbb{N} : \forall n \in \mathbb{N} : n > N \implies d(x, a_n) < \varepsilon.$$

Utilizando as propriedades da métrica demonstra-se que, se o limite existe, ele é único, e essa demonstração requer o uso de todas as propriedades da métrica, o que mostra a importância dessa definição e de cada uma de suas propriedades. Observe também que sem uma métrica não é possível falar de convergência e a noção de limite não tem sentido.

4.3.4.2 Sequência de Cauchy

Definição 35. *Seja um espaço métrico $\langle \mathbb{X}, d \rangle$; uma sequência $\{a_n : n \in \mathbb{N}\}$ em \mathbb{X} é dita ser uma sequência de Cauchy em relação à métrica d se para todo $\varepsilon > 0$ existir um número natural N tal que $d(a_i, a_j) < \varepsilon$ para todos $i > N$ e $j > n$, simbolicamente,*

$$\forall \varepsilon \in \mathbb{R} > 0 : \exists N \in \mathbb{N} : \forall i, j \in \mathbb{N} : i, j > N \implies d(a_i, a_j) < \varepsilon.$$

As sequências de Cauchy são de grande importância no estudo dos espaços métricos, sendo essa uma de suas mais importantes propriedades:

Proposição 3. *Seja um espaço métrico $\langle \mathbb{X}, d \rangle$ e seja $\{a_n : n \in \mathbb{N}\}$ um sequência convergente em relação à métrica d a algum elemento de \mathbb{X} , então $\{a_n : n \in \mathbb{N}\}$ é uma sequência de Cauchy.*

Como toda sequência convergente é de Cauchy, fica a pergunta: será que toda sequência de Cauchy é convergente? Essa pergunta é importante porque para saber se uma sequência é de Cauchy ou não só precisamos da própria sequência, enquanto que para saber se ela converge é preciso encontrar um elemento para o qual ela satisfaça o critério de convergência. Mas a resposta é negativa, existem sequências de Cauchy que não são convergentes. Um exemplo

clássico é a sequência que converge para o número de Euler:

$$a_n = \sum_{i=1}^n \frac{1}{i!}$$

e considerando que o nosso espaço métrico é composto pelos racionais \mathbb{Q} com a métrica $d(x, y) = |x - y|$, verifica-se que essa sequência é de Cauchy, mas ela converge para o número de Euler, o qual não é um número racional, e portanto ela não tem limite nos racionais.

4.3.4.3 Completeza

Mas em alguns espaços especiais como nos reais e complexos, essa propriedade é válida: todas as sequências de Cauchy convergem. Esses espaços recebem um nome especial, eles são ditos completos:

Definição 36. Dizemos que um espaço métrico (\mathbb{X}, d) é **completo** se toda sequência de Cauchy em \mathbb{X} convergir para um valor em \mathbb{X} .

4.3.5 Espaços de Banach e de Hilbert

Uma vez que tenhamos definido os conceitos de norma, produto interno, convergência e completeza, podemos tratar dos espaços de Banach e de Hilbert. Ambos são espaços vetoriais com propriedades métricas e de convergência: são espaços métricos completos.

4.3.5.1 Espaços de Banach

Quando um espaço vetorial \mathcal{B} é um espaço métrico completo, ou seja, todas as sequências de Cauchy convergem, e cuja métrica é induzida pela norma,

$$\forall u, v \in \mathcal{B} : d(x, u) = \|u - v\|,$$

o espaço \mathcal{B} é o que chamamos de espaço de Banach:

Definição 37. Um **espaço de Banach** \mathcal{B} é um espaço vetorial normado que é um espaço métrico completo com relação à métrica induzida pela norma.

4.3.5.2 Espaços de Hilbert

Restringindo um pouco mais essa definição, obtemos os chamados espaços de Hilbert. Quando um espaço vetorial \mathcal{H} é um espaço métrico completo cuja métrica é induzida pelo produto interno,

$$\forall u, v \in \mathcal{B} : \quad d(x, u) = \|u - v\| = \sqrt{\langle u - v, u - v \rangle}$$

dizemos que \mathcal{H} é um espaço de Hilbert:

Definição 38. *Um espaço de Hilbert \mathcal{H} é um espaço vetorial com produto interno que é um espaço métrico completo com relação à métrica induzida pelo produto interno.*

Todo espaço de Hilbert é naturalmente um espaço de Banach, mas a recíproca não é verdadeira, pois nem todo espaço de Banach possui um produto interno. Note ainda que essa é a definição moderna de espaços de Hilbert, e que ela difere da definição original dada por von Neumann, que foi quem cunhou o termo ‘espaço de Hilbert’. Segundo von Neumann, um espaço de Hilbert, além das propriedades aqui enunciadas, deve ser separável; ocorreu então uma mudança de significado desse termo que passou a ser mais geral e o espaço de Hilbert de von Neumann passou a ser chamado de *espaço de Hilbert separável*.

4.4 Álgebra C^*

Após todo esse instrumental matemático, podemos finalmente definir a álgebra C^* e abordar algumas de suas propriedades fundamentais. Primeiramente definimos o que é uma álgebra normada:

Definição 39. *Uma álgebra \mathbb{V} sobre os números complexos \mathbb{C} é dita ser uma **álgebra normada** se possuir uma função $\|\cdot\| : \mathbb{V} \rightarrow \mathbb{R}$ denominada norma que satisfaz:*

$$\forall u, v \in \mathbb{V}, \forall \alpha \in \mathbb{C} :$$

$$\|u\| \geq 0 \quad \text{Não-negatividade} \quad (4.71)$$

$$\|u\| = 0 \iff u = 0 \quad (4.72)$$

$$\|\alpha u\| = |\alpha| \|u\| \quad (4.73)$$

$$\|u + v\| \leq \|u\| + \|v\| \quad (4.74)$$

$$\|uv\| \leq \|u\| \|v\|. \quad (4.75)$$

Alternativamente, uma álgebra \mathbb{V} sobre os números complexos \mathbb{C} é dita ser uma álgebra normada se possuir uma norma $\|\cdot\| : \mathbb{V} \rightarrow \mathbb{R}$ que satisfaça:

$$\forall a, b \in \mathbb{V} : \quad \|ab\| \leq \|a\| \|b\| \quad \text{Desigualdade do produto.}$$

Definimos agora que seria uma involução sobre uma álgebra associativa:

Definição 40. Uma operação unária $*$: $\mathbb{V} \rightarrow \mathbb{V}$ é dita ser uma **involução**, ou **operação de adjunto** sobre uma álgebra associativa \mathbb{V} definida sobre o corpo dos complexos se satisfaz as seguintes propriedades:

$$\forall \alpha, \beta \in \mathbb{C}, \forall A, B \in \mathbb{V} :$$

$$A^* = A \quad (4.76)$$

$$(AB)^* = B^* A^* \quad (4.77)$$

$$(\alpha A + \beta B)^* = \bar{\alpha} A^* + \bar{\beta} B^*, \quad (4.78)$$

em que $\bar{\alpha}$ é o complexo conjugado de α .

As álgebras associativas que possuem uma involução são chamadas *álgebras com involução* ou *álgebras-**. A operação de conjugação dos números complexos é uma involução e é a que dá origem a essa definição geral de involução

Um importante resultado é que quando álgebra- $*$ possui identidade, a sua identidade é autoadjunta.

4.4.1 Álgebra de Banach

Com todos esses recursos em mãos podemos facilmente definir as álgebras de Banach, Banach- $*$ e C^* .

Definição 41. Uma **álgebra de Banach** é uma álgebra associativa normada \mathcal{B} que é completa em relação à métrica induzida pela norma.

Em uma álgebra de Banach sempre é possível redefinir a norma de forma a termos que $\|\hat{\mathbf{1}}\| = 1$, basta definirmos a nova norma $\|\cdot\|'$:

$$\|A\|' = \frac{\|a\|}{\|\hat{\mathbf{1}}\|},$$

e obtemos $\|\hat{\mathbf{1}}\|' = \frac{\|\hat{\mathbf{1}}\|}{\|\hat{\mathbf{1}}\|} = 1$. Outra propriedade importante das álgebras com identidade é que

$$\|\hat{\mathbf{1}}\| \geq 1,$$

supondo que não estejamos no caso trivial $\|\hat{\mathbf{1}}\| \neq 0$, pois $\|\hat{\mathbf{1}}\| \leq \|\hat{\mathbf{1}}\|^2$.

Acrescentando a operação de involução, e uma propriedade conveniente, obtemos a álgebra de Banach- $*$:

Definição 42. *Uma álgebra de Banach- $*$ é uma álgebra de Banach \mathcal{B} com involução que satisfaz a propriedade adicional:*

$$\forall A \in \mathcal{V} : \quad \|A\| = \|A^*\|$$

Todas as álgebra de Banach ou Banach- $*$ são também espaços de Banach, o que justifica a origem do nome.

4.4.2 Álgebra C^*

Definição 43. *Uma álgebra C^* é uma álgebra de Banach- $*$ \mathcal{C} que satisfaz a propriedade adicional:*

$$\forall A \in \mathcal{V} : \quad \|A^*A\| = \|A\|^2,$$

denominada propriedade C^* .

Alternativamente, uma álgebra C^ é uma álgebra associativa normada com involução \mathcal{C} que é completa em relação à métrica induzida pela norma e satisfaz a propriedade adicional:*

$$\forall A \in \mathcal{V} : \quad \|A^*A\| = \|A\|^2.$$

Note que, a propriedade C^* $\|A^*A\| = \|A\|^2$ automaticamente garante a propriedade da álgebra de Banach- $*$ $\|A\| = \|A^*\|$, pois combinando-a com a inequação do produto $\|AB\| \leq \|A\|\|B\|$, obtemos:

$$\|A\|^2 = \|A^*A\| \leq \|A^*\| \|A\| \implies \|A\| \leq \|A^*\|,$$

por outro lado,

$$\|A^*\|^2 = \|AA^*\| \leq \|A\| \|A^*\| \implies \|A^*\| \leq \|A\|,$$

de onde conclui-se que

$$\|A\| = \|A^*\|.$$

O principal exemplo de álgebra C^* de nosso interesse é formada pelos operadores limitados $\mathcal{L}(\mathcal{H})$ de um espaço de Hilbert \mathcal{H} . Por razões didáticas é importante definirmos antes o que é o supremo de um conjunto. Seja $\mathbb{A} \subseteq \mathbb{R}$ um subconjunto dos reais, então

1. Um elemento $M \in \mathbb{R}$ é dito **majorante** ou limite superior de \mathbb{A} se

$$\forall x \in \mathbb{A} : x \leq M.$$

2. Um elemento $M \in \mathbb{R}$ é dito **supremo** de \mathbb{A} se ele for o menor dos majorantes:

$$\sup \mathbb{A} = M \iff (\forall x \in \mathbb{A} : x \leq M') \implies M \leq M'$$

Assim, definindo a norma por:

$$\forall A \in \mathcal{L}(\mathcal{H}) : \|A\| = \sup \{ \|A|\psi\rangle\| : |\psi\rangle \in \mathcal{H}, \|\psi\rangle\| = 1 \}, \quad (4.79)$$

e utilizando as operações binárias de soma e produto de operadores, e como involução a operação de adjunto, verifica-se que $\mathcal{L}(\mathcal{H})$ é uma álgebra C^* .

A propriedade da norma C^* pode ser verificada fazendo:

$$\|A\|^2 = \left(\sup \left\{ \sqrt{\langle A|\psi\rangle, A|\psi\rangle} : |\psi\rangle \in \mathcal{H}, \|\psi\rangle\| = 1 \right\} \right)^2 \quad (4.80)$$

$$= \sup \{ \langle A|\psi\rangle, A|\psi\rangle : |\psi\rangle \in \mathcal{H}, \|\psi\rangle\| = 1 \} \quad (4.81)$$

$$= \sup \{ \langle |\psi\rangle, A^*A|\psi\rangle : |\psi\rangle \in \mathcal{H}, \|\psi\rangle\| = 1 \} \quad (4.82)$$

$$\leq \sup \{ \|\psi\rangle\| \|A^*A|\psi\rangle\| : |\psi\rangle \in \mathcal{H}, \|\psi\rangle\| = 1 \} \quad (4.83)$$

$$= \|A^*A\|. \quad (4.84)$$

É importante aqui definirmos o que é um operador limitado: um operador linear A atuando em um espaço de Hilbert \mathcal{H} é limitado se existir um número real M tal que

$$\forall |\psi\rangle \in \mathcal{H} : \|A|\psi\rangle\| \leq M\|\psi\rangle\|.$$

Para verificarmos a relação da álgebra C^* com os operadores limitados, fazemos:

$$\frac{\|A|\psi\rangle\|}{\|\psi\rangle\|} = \|A \frac{|\psi\rangle}{\|\psi\rangle}\| \leq \sup \{ \|A|\psi\rangle\| : \|\psi\rangle\| = 1, |\psi\rangle \in \mathcal{H} \} = \|A\|$$

de onde obtemos que

$$\forall |\psi\rangle \in \mathfrak{D}(A) : \quad \|A|\psi\rangle\| \leq \|A\| \|\psi\|, \quad (4.85)$$

e inferimos que A é limitado. Isso nos mostra claramente que a álgebra C^* só admite operadores limitados, pois $\|A\|$ é um número real.

A partir de agora, vamos tratar somente das álgebras C^* com identidade, e considerar $\hat{1}$ como sendo a identidade da álgebra. Sabemos que a identidade, quando existe, é única, e é definida por

$$\forall A \in \mathcal{C} : \quad \hat{1}A = A = A\hat{1}.$$

Tomando o adjunto da equação acima,

$$\forall A^* \in \mathcal{C} : \quad \hat{1}^*A^* = A^* = A^*\hat{1}^* \implies \forall A \in \mathcal{C} : \quad \hat{1}^*A = A = A\hat{1},$$

verifica-se que $\hat{1}^*$ também é uma identidade, e portanto a identidade é autoadjunta, $\hat{1}^* = \hat{1}$. Quanto à norma da identidade, verifica-se, a partir das relações

$$\|\hat{1}\| = \|\hat{1}^*\hat{1}\| = \|\hat{1}\|^2 \quad (4.86)$$

$$\|A\| = \|\hat{1}A\| \leq \|\hat{1}\| \|A\|, \quad (4.87)$$

que $\|\hat{1}\| = 0$ ou 1 , no entanto, o caso $\|\hat{1}\| = 0$ nos leva ao caso trivial em que a identidade é nula. Sendo assim, definimos que

$$\|\hat{1}\| = 1. \quad (4.88)$$

4.4.3 Espectro

Vamos tratar agora do espectro de operadores de uma álgebra C^* , para isso é fundamental abordarmos primeiro a questão do inverso de operadores. Dizemos que um elemento $A \in \mathcal{C}$ é inversível se existir um elemento $A^{-1} \in \mathcal{C}$ que satisfaça:

$$A^{-1}A = \hat{1} = AA^{-1}, \quad (4.89)$$

nesse caso, o elemento A^{-1} é o inverso de A , e esse elemento é único. O conjunto de todos elementos invertíveis de uma álgebra \mathcal{C} é denotado por $\text{Inv}(\mathcal{C})$. Assim temos as seguintes propriedades:

$$\forall A, B \in \text{Inv}(\mathcal{C}) :$$

$$A^* \in \text{Inv}(\mathcal{C}) \quad (4.90)$$

$$(A^*)^{-1} = (A^{-1})^* \quad (4.91)$$

$$A^{-1} \in \text{Inv}(\mathcal{C}) \quad (4.92)$$

$$(A^{-1})^{-1} = A \quad (4.93)$$

$$AB \in \text{Inv}(\mathcal{C}) \quad (4.94)$$

$$(AB)^{-1} = B^{-1}A^{-1} \quad (4.95)$$

Podemos agora definir o resolvente e o espectro de um operador:

Definição 44. *Seja \mathcal{C} uma álgebra C^* com identidade e um elemento $A \in \mathcal{C}$; o conjunto resolvente de A é denotado por $\rho(A)$ e definido por:*

$$\rho(A) = \{\lambda \in \mathcal{C} : \lambda \hat{\mathbf{1}} - A \in \text{Inv}(\mathcal{C})\} \quad (4.96)$$

Definição 45. *Seja \mathcal{C} uma álgebra C^* com identidade e um elemento $A \in \mathcal{C}$; o espectro A é denotado por $\sigma(A)$ e definido por:*

$$\sigma(A) = \{\lambda \in \mathcal{C} : \lambda \hat{\mathbf{1}} - A \notin \text{Inv}(\mathcal{C})\} \quad (4.97)$$

Para facilitar a compreensão dessa noção de espectro, vamos correlacionar essa definição com a noção de autovalores e autovetores existente nos formalismos de Dirac e von Neumann. Sabemos que nesses formalismos, o espectro de um operador é o conjunto dos autolavores do operador, assim vamos mostrar a correspondência dessa noção com a definição na álgebra C^* . Consideremos que um operador A tem autovalor λ e autoestado $|\psi\rangle$, então:

$$A|\psi\rangle = \lambda|\psi\rangle,$$

o que nos leva a

$$(\lambda \hat{\mathbf{1}} - A)|\psi\rangle = 0.$$

Porém, se $(\lambda \hat{\mathbf{1}} - A)$ for inversível, obtemos uma contradição:

$$(\lambda \hat{\mathbf{1}} - A)^{-1}(\lambda \hat{\mathbf{1}} - A)|\psi\rangle = \hat{\mathbf{1}}|\psi\rangle = |\psi\rangle = 0,$$

como $|\psi\rangle$ é não nulo, para que λ seja um autovalor, $(\lambda \hat{\mathbf{1}} - A)$ não pode ser inversível. Isso justifica a definição de espectro em uma álgebra C^* e mostra que ela corresponde com a noção de espectro nos formalismos de Dirac e von Neumann.

Definimos algumas classes importantes de operadores:

Definição 46. Um elemento $X \in \mathcal{C}$ é **normal** se ele comuta com o seu adjunto, ou seja, se

$$XX^* = X^*X.$$

Definição 47. Um elemento $X \in \mathcal{C}$ é uma **isometria** se satisfaz

$$X^*X = \hat{1},$$

e X é **unitário** se satisfaz

$$X^*X = \hat{1} = XX^*.$$

O espectro de um elemento de uma álgebra C^* é limitado pela norma do elemento:

Proposição 4. Se $A \in \mathcal{C}$ é autoadjunto,

$$\sigma(A) \subseteq [-\|A\|, \|A\|], \quad \sigma(A^2) \subseteq [0, \|A\|^2]$$

o que mostra que em uma álgebra C^* não podemos ter elementos cujo espectro contém números arbitrariamente grandes, como por exemplo os operadores de posição e momentum da mecânica quântica.

4.4.4 Elementos Positivos

Uma das classes mais importantes de elementos de uma álgebra C^* é a classe dos elementos positivos, noção que introduz uma relação de ordem parcial com propriedades muito semelhantes à relação de ordem dos números reais.

Definição 48. Um elemento A de uma álgebra C^* \mathcal{C} é dito **positivo** se existir um elemento $B \in \mathcal{C}$ tal que $A = B^*B$.

Definição 49. O conjunto de todos os elementos positivos de \mathcal{C} é denotado por \mathcal{C}_+ :

$$\mathcal{C}_+ = \{B^*B : B \in \mathcal{C}\}. \quad (4.98)$$

Proposição 5. Todos os elementos positivos são autoadjuntos,

$$\forall A \in \mathcal{C}_+ : A^* = A. \quad (4.99)$$

Temos ainda as seguintes caracterizações alternativas de um elemento positivo:

Proposição 6. *Um elemento A de uma álgebra C^*C é positivo se, e somente se, A for autoadjunto e seu espectro for formado somente por números não negativos.*

Proposição 7. *Um elemento A de uma álgebra C^*C é positivo se, e somente se, existir um elemento autoadjunto $B \in C$ tal que $A = B^2$.*

Esta última proposição nos permite definir a raiz quadrada de um operador positivo A , a qual é denotada por \sqrt{A} e é definida como sendo o (único) elemento positivo B tal que $A = B^2$, ou seja

$$A = B^2 \text{ e } B \in C_+ \implies \sqrt{A} = B. \quad (4.100)$$

Todos os elementos positivos possuem uma raiz quadrada, e essa raiz é única. Podemos ainda definir o módulo de um operador autoadjunto A por

$$|A| = \sqrt{A^2}. \quad (4.101)$$

Uma outra propriedade importante é que \sqrt{A} comuta com todo elemento que comuta com A .

Vamos tratar agora da ordem parcial de elementos autoadjuntos induzida pela noção de positividade. Primeiro, deve-se observar que a intersecção entre C_+ e $-C_+$ contém somente o elemento nulo,

$$C_+ \cap (-C_+) = \{0\}, \quad (4.102)$$

e que podemos associar a cada elemento autoadjunto A , os elementos A_+ e A_- definidos por

$$A_+ = \frac{|A| + A}{2}, \quad A_- = \frac{|A| - A}{2} \quad (4.103)$$

que satisfazem as seguintes propriedades:

$$A_+, A_- \in C_+ \quad (4.104)$$

$$A = A_+ - A_- \quad (4.105)$$

$$A_+A_- = A_-A_+ = 0. \quad (4.106)$$

Verifica-se ainda que esses dois elementos são os únicos que satisfazem essas duas propriedades. Isso nos permite decompor qualquer elemento autoadjunto de uma álgebra C^* em partes positiva e negativa, e essas partes são ortogonais, isto é, o produto delas é nulo.

Podemos agora introduzir as relações de ordem parcial \geq e \leq entre dois elementos autoadjuntos:

Definição 50. *Sejam dois elementos autoadjuntos A e B da álgebra \mathcal{C} , então*

$$A \geq B \iff A - B \in \mathcal{C}_+, \quad (4.107)$$

$$A \leq B \iff B - A \in \mathcal{C}_+. \quad (4.108)$$

As duas relações estão correlacionadas por $A \geq B \iff B \leq A$, e a relação \geq satisfaz, para quaisquer elementos autoadjuntos $A, B, C \in \mathcal{C}$:

1. Reflexividade: $A \geq A$
2. Antissimetria: $A \geq B$ e $B \geq A \implies A = B$
3. Transitividade: $A \geq B$ e $B \geq C \implies A \geq C$

Essa relação é compatível com a operação de soma, $A, B, C \in \mathcal{C}$:

$$A \geq B \implies A + C \geq B + C \quad (4.109)$$

e quando os elementos A e B comutam, ela é compatível com a multiplicação:

$$A \geq 0 \text{ e } B \geq 0 \implies AB \geq 0. \quad (4.110)$$

Essas propriedades tornam essa relação uma generalização da relação de ordem dos números reais. Outras propriedades importantes são:

$\forall A, B, C \in \mathcal{C}$:

$$A \geq B \geq 0 \implies \|A\| \geq \|B\| \quad (4.111)$$

$$A \geq 0 \implies A\|A\| \geq A^2 \quad (4.112)$$

$$\|A\|\hat{\mathbf{1}} \geq A \quad (4.113)$$

$$A \geq B \geq 0 \implies C^*AC \geq C^*BC \geq 0. \quad (4.114)$$

4.4.5 Estados

Podemos agora finalmente chegar à definição de estados, que é o que permite uma conexão maior da álgebra C^* com os outros formalismos da mecânica quântica. Os estados de álgebra C^* correspondem aos operadores densidade (operador positivo com traço 1) no formalismo de von Neumann, e, em um caso particular, aos vetores de estado no formalismo de Dirac.

Para definir o que é um estado, precisamos primeiramente definir o que é um funcional linear:

Definição 51. *Um funcional linear f sobre uma álgebra C^*C é uma função linear que mapeia elementos da álgebra em números complexos, ou seja,*

$$f : C \longrightarrow \mathbb{C} \quad (4.115)$$

$$\forall A, B \in C, \alpha, \beta \in \mathbb{C} : f(\alpha A + \beta B) = \alpha f(A) + \beta f(B). \quad (4.116)$$

Definimos agora os funcionais lineares positivos e a norma de um funcional linear:

Definição 52. *Um funcional linear ω sobre C é dito positivo se satisfaz:*

$$\forall A \in C : \omega(A^*A) \in \mathbb{R} \geq 0.$$

Note que, um elemento positivo é sempre da forma A^*A , logo um funcional linear positivo sempre mapeia elementos positivos em números reais não-negativos, então a condição de positividade de ω é equivalente a ω ser não-negativo em elementos positivos.

Uma propriedade muito importante dos funcionais lineares positivos é a conexão deles com a relação de ordem: seja um funcional linear positivo ω :

$$\forall A, B \in C : A \geq B \implies \omega(A) \geq \omega(B). \quad (4.117)$$

Essa propriedade decorre do fato de que, se $A \geq B$, então, $A - B$ é positivo e $\omega(A - B) = \omega(A) - \omega(B) \geq 0$. Outra propriedade que decorre da positividade é que os funcionais lineares são limitados pela norma:

$$\forall A \in C : \omega(A) \leq \|A\| \omega(\hat{\mathbf{1}}). \quad (4.118)$$

Definição 53. *Seja f um funcional linear sobre C , então define-se a norma de f como:*

$$\|f\| = \sup \{ |f(A)| : \|A\| = 1, A \in C \}. \quad (4.119)$$

A norma de um funcional linear positivo ω é dada pelo valor desse funcional na identidade:

$$\|\omega\| = \sup \{ \omega(A) : \|A\| = 1, A \in C \} = \sup \{ \|A\| \omega(\hat{\mathbf{1}}) : \|A\| = 1, A \in C \} = \omega(\hat{\mathbf{1}})$$

Podemos agora definir o que é um estado:

Definição 54. *Um estado de C é um funcional linear positivo ω sobre C com $\|\omega\| = 1$, ou*

seja, um funcional linear ω é um estado se

$$\begin{aligned} \forall A \in \mathcal{C} : \quad \omega(A^*A) &\geq 0 \\ \|\omega\| &= 1, \end{aligned}$$

alternativamente,

$$\begin{aligned} \forall A \in \mathcal{C} : \quad \omega(A^*A) &\geq 0 \\ \omega(\hat{\mathbf{1}}) &= 1. \end{aligned}$$

É interessante que os funcionais lineares positivos satisfazem propriedades análogas ao produto interno, inclusive a desigualdade de Cauchy-Schwarz:

Proposição 8. *Seja ω um funcional linear positivo sobre \mathcal{C} , então, para $\forall A, B \in \mathcal{C}, \alpha, \beta \in \mathbb{C} ::$*

$$\omega(A(\alpha B + \beta B')) = \alpha\omega(AB) + \beta\omega(AB') \quad (4.120)$$

$$\omega(A^*B) = \overline{\omega(B^*A)} \quad (4.121)$$

$$|\omega(A^*B)|^2 \leq \omega(A^*A)\omega(B^*B) \quad (4.122)$$

No formalismo de von Neumann, os funcionais lineares estão associados a operadores densidade, isto é, operadores positivos com traço 1, e podemos fazer a correspondência:

$$\omega(A) = Tr(\rho A),$$

em que ρ é um operador densidade. É imediato que ω definido dessa maneira é um funcional linear, mas precisamos verificar que ele é positivo e $\|\omega\| = 1$. Sabe-se que o traço de um operador positivo, que von Neuman chama de definido, é sempre não negativo, assim, verifica-se a positividade fazendo:

$$\omega(A^*A) = Tr(\rho A^*A) = Tr(A\rho A^*) \geq 0,$$

pois $A\rho A^*$ é positivo. Como a norma de um funcional linear positivo é dada pelo seu valor na identidade,

$$\|\omega\| = \omega(\hat{\mathbf{1}}) = Tr(\rho\hat{\mathbf{1}}) = Tr(\rho) = 1.$$

Com isso demonstramos que qualquer funcional da forma $\omega(A) = Tr(\rho A)$ em que ρ é um operador positivo com traço igual a unidade é um estado de acordo com a definição da álgebra C^*

No formalismo original de Dirac, ele considera que os estados são vetores kets. Esse é

um caso particular que atualmente chamamos de estados puros, e a correspondência com a definição de estados de uma álgebra C^* é dada da seguinte forma: qualquer ket normalizado $|\psi\rangle$ define um estado $\omega_{|\psi\rangle}$ dado por

$$\omega_{|\psi\rangle}(A) = \langle\psi|A|\psi\rangle,$$

que corresponde ao estado definido pelo operador densidade $\rho = |\psi\rangle\langle\psi|$:

$$\omega_{|\psi\rangle}(A) = \text{Tr}(|\psi\rangle\langle\psi|A) = \langle\psi|A|\psi\rangle.$$

Isso faz a correspondência do formalismo da álgebra C^* com os formalismos de Dirac e von Neumann. Todos esses formalismos possuem operadores lineares, e em todos os formalismos, os operadores autoadjuntos representam observáveis do sistema. Os estados, por sua vez, são representados de forma diferente em cada formalismo:

- no formalismo de Dirac: vetores kets normalizados;
- no formalismo de von Neumann: operadores positivos com traço 1;
- no formalismo da álgebra C^* : funcionais lineares positivos com norma 1.

Com isso concluímos a apresentação dos fundamentos básicos da álgebra C^* e de como ela se relaciona com os outros formalismos da Mecânica Quântica.

CAPÍTULO 5

Conclusão

Expusemos as duas formulações originais do formalismo matemático atual da MQ: a formulação de Dirac e a de von Neumann. Em seguida expusemos a forma matemática atual de alguns dos conceitos previamente definidos e da álgebra C^* , apresentando as estruturas algébricas fundamentais e os conceitos básicos necessários para a compreensão do tema. Definimos passo a passo todos os conceitos necessários para a compreensão da álgebra C^* , de forma que nossa abordagem foi autocontida, sendo necessário ao leitor pouco ou nenhum conhecimento prévio do tema, exceto uma compreensão básica do linguajar matemático.

Essas duas formulações originais foram um marco para a matemática da MQ: fundamentaram os aspectos matemáticos da teoria quântica, unificando a mecânica matricial de Heisenberg-Born-Jordan e a mecânica ondulatória de Schrodinger originando o que é a base matemática comum à maioria das formulações e interpretações da MQ. Até então existiam duas teorias distintas cuja equivalência havia sido provada. Após esses trabalhos, mecânica matricial e mecânica ondulatória passaram a ser apenas duas representações diferentes de uma mesma teoria.

Usualmente os dois formalismos são meio que considerados equivalentes, o formalismo de von Neumann apenas sendo mais rigoroso, ou seja, considera-se que os dois são capazes de descrever os mesmos fenômenos físicos, porém um é matematicamente mais rigoroso. O próprio von Neumann sugere essa ideia no prefácio de seu livro: ele afirma que a mesma teoria pode ser estabelecida de forma também clara e estabelecida mas sem objeções matemáticas, o que nos leva a entender que em essência se trata da mesma teoria. No entanto, elas foram criadas num clima de grande rivalidade e muitas críticas por parte de von Neumann, ele até mesmo ressaltou que a sua formulação difere *desde o princípio* da formulação de Dirac.

O formalismo de Dirac [13] não utiliza originalmente a estrutura de espaços de Hilbert; os kets não são, *originalmente*, elementos de um espaço de Hilbert. É muito comum utilizarmos o formalismo matemático de Dirac, (o de von Neumann por sua vez é pouco utilizado devido às suas complexidades), e ainda assim dizermos que a MQ é feita no espaço de Hilbert; o que nos leva a acreditar que o formalismo de Dirac utiliza o conceito de espaços de Hilbert. O Próprio Dirac afirma [13, p. 40] que o espaço que ele utiliza não é um ‘espaço de Hilbert’, e não considera em suas definições as propriedades matemáticas que caracterizam realmente

essa estrutura. No caso, Dirac não impõe a condição de completeza do espaço, que é a propriedade fundamental na definição moderna que caracteriza um espaço de Hilbert; na definição de von Neumann é necessário ainda que o espaço seja separável.

O trabalho de von Neumann “On the foundations of Quantum Mechanics” [1] não corrige ou substitui completamente o formalismo de Dirac, e portanto não é realmente correto dizer que seu trabalho consolidou a formulação matemática da MQ atual. Von Neumann criticou a ‘função’ Delta e o rigor do formalismo de Dirac no geral, mas não conseguiu realmente resolver os seus problemas. Todo o formalismo de von Neumann abrange apenas a parte discreta do formalismo de Dirac. Mesmo com tantas críticas à ‘função’ Delta, e afirmando que irá fazer um formalismo em que ela não é mais necessária, von Neumann realmente desenvolve um formalismo sem as mesmas objeções matemáticas e sem a Delta, porém seu formalismo contempla apenas uma parte do formalismo de Dirac: *a parte em que a Delta não é necessária!* Por incrível que pareça, a parte de formalismo de Dirac que podemos dizer que von Neumann tornou mais rigorosa é a parte de espectros discretos sendo que a Delta é utilizada apenas em espectros contínuos. Note ainda que esse elemento tão criticado foi no entanto rigorosamente estabelecida pela teoria das distribuições de Schwartz.

A MQ atual não é, e não pode ser, formulada inteiramente no formalismo rigoroso de von Neumann de [1]. O conceito de espaços de Hilbert usado e cunhado por von Neumann é diferente do conceito atual, e não admite sistemas com espectros contínuos. Von Neumann utiliza o que é atualmente denominado espaço de Hilbert *separável*, condição que limita as bases ortogonais possíveis a conjuntos contáveis, como $\{|x\rangle : x \in \mathbb{N}\}$, e por isso, por exemplo, a base das posições $\{|x\rangle : x \in \mathbb{R}\}$ não pode ser tratada nesse formalismo.

A Álgebra C^* pode ser utilizada para descrever a MQ de operadores limitados. No entanto, a MQ como um todo não pode ser feita no espaço dos operadores limitados e isso pode gerar complicações no uso da álgebra C^* pois os operadores de posição e momento não são limitados. Seguindo a demonstração de [17] e considerando dois operadores X e P que satisfaçam a relação de comutação canônica

$$[X, P] = -i\hat{1},$$

em que consideramos $\hbar = 1$, fazendo:

$$\begin{aligned} \forall n \in \mathbb{N} : \quad [X, P^n] &= XP^n - P^n X = inP^{n-1} \\ \|XP^n\| + \|P^n X\| &\geq n\|P^{n-1}\| \\ \|X\|\|P\|^n + \|P\|^n\|X\| &\geq n\|P\|^{n-1} \\ \forall n \in \mathbb{N} : \quad \|X\|\|P\| &\geq \frac{n}{2} \end{aligned}$$

obtemos que X ou P devem ter uma norma infinita não sendo portanto limitados. Por outro lado, como X e P podem ter autovalores arbitrariamente grandes, eles não possuem uma norma definida, e não são limitados.

Esso mostra em síntese que os pormenores do formalismo matemático da MQ não estão tão bem estabelecidos como acredita-se ou não foram estabelecidos tão cedo quanto imaginamos. O formalismo de Dirac parece [18] ter sido rigorosamente estabelecido com a teoria do ‘espaço de Hilbert equipado’ (*rigged Hilbert space*) que surgiu na década de 60 [19, 20]. Como nenhum desses formalismos, Dirac, von Neumann e álgebra C^* oferecem uma solução rigorosa e definitiva para toda a MQ, indagamos: existe algum formalismo matemático estabelecido rigoroso, algum livro ou artigo que estabelece uma fundamentação matemática unificada válida para toda a MQ, cobrindo todas as lacunas deixadas pelos formalismos de Dirac, von Neumann e álgebra C^* ? Existe um formalismo rigoroso que cubra espectros discretos e contínuos de operadores limitados ou não? Seria a teoria do *rigged Hilbert space* quem cumpre completamente esse papel?

Observa-se que von Neumann enfatiza o rigor matemático e Dirac enfatiza a pragmática e facilidade de uso [2], desenvolvendo uma notação com diversas propriedades interessantes e muito simples; mas isso nos leva a uma pergunta interessante: seria possível, e como podemos, unificar o rigor de von Neumann com a pragmática e simplicidade de Dirac em um único formalismo matemático? Como podemos incorporar *de forma rigorosa* a notação de Dirac de bras e kets ao formalismo de von Neumann? Uma resposta afirmativa a essas perguntas mostraria que é possível reunir os ideais de rigor, pragmatismo e simplicidade numa única teoria resultando em um formalismo que reúne o melhor de Dirac e o melhor de von Neumann. Esse formalismo poderia ser restrito ao domínio dos espaços de Hilbert separáveis, mas nesse domínio ele seria simultaneamente rigoroso e elegante.

Vimos que o formalismo/notação de bras e kets é apenas uma notação, e não forma até então uma verdadeira álgebra; para isso seria necessário definir as operações binárias de adição (+) e de multiplicação (\cdot) de forma que tivéssemos:

$$\begin{array}{ll} \langle a|b \rangle = \langle a| \cdot |b \rangle & |a\rangle\langle b| = |a\rangle \cdot \langle a| \\ X|a\rangle = X \cdot |a\rangle & \langle a|X = \langle a| \cdot X \\ \alpha X = \alpha \cdot X & X\alpha = X \cdot \alpha \\ \alpha|a\rangle = \alpha \cdot |a\rangle & \langle a|\alpha = \langle a| \cdot \alpha \\ |a\rangle\alpha = |a\rangle \cdot \alpha & \alpha\langle a| = \alpha \cdot \langle a| \end{array}$$

Dirac construiu a notação intencionalmente de forma a tornar essa correspondência possível, pois graças a ela as propriedades da notação de Dirac se assemelham às propriedades das operações algébricas usuais. Mas como Dirac não definiu e caracterizou formalmente tais ope-

rações, temos apenas uma notação que simplesmente aparenta ser uma álgebra em que bras, kets e operadores podem ser somados e multiplicados. Porém isso nos leva à seguinte indagação: não seria possível definir as operações de adição e multiplicação e criar um formalismo em que bras, kets, operadores e números estão formalmente sendo multiplicados e somados? Com isso reduziríamos o número de axiomas ou propriedades necessárias para a definição do formalismo de Dirac tornando o mais elegante e formal.

Pelo que observamos, ainda há espaço para melhorias no formalismo matemático da MQ. Além disso, verificamos a importância de se estudar os trabalhos originais e de se integrar os conhecimentos das diferentes abordagens. O estudo dos projetores por von Neumann e dos operadores positivos pela álgebra C^* , por exemplo, raramente são discutidos nos livros texto de MQ [21–23] mas fornecem ferramentas que podem ser de grande utilidade, uma vez que projetores e operadores positivos são os elementos fundamentais para as medidas projetivas ou POVMs.

Referências Bibliográficas

- [1] J. von Neumann, “Mathematical foundations of quantum mechanics” (Princeton University Press, 1955)
- [2] Kronz, Fred and Luper, Tracy, “Quantum Theory: von Neumann vs. Dirac”, The Stanford Encyclopedia of Philosophy (Summer 2012 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/sum2012/entries/qt-nvd/>>.
- [3] Fuchs, Christopher A., “Quantum Mechanics as Quantum Information (and only a little more)”, *quant-ph/0205039v1*, (2002).
- [4] D. Bohm, “A Suggested Interpretation of the Quantum Theory in Terms of ‘Hidden Variables’ I”. *Physical Review* **85**: 166-179,(1952)
- [5] D. Bohm, ‘A Suggested Interpretation of the Quantum Theory in Terms of ‘Hidden Variables’ II”, *Physical Review* **85**: 180-193,(1952)
- [6] W. Heisenberg, “Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen”, *Zeitschrift für Physik*, 33, 879-893, 1925 (received July 29, 1925). [English translation in: B. L. van der Waerden, editor, *Sources of Quantum Mechanics* (Dover Publications, 1968) ISBN 0-486-61881-1 (English title: *Quantum-Theoretical Reinterpretation of Kinematic and Mechanical Relations*).]
- [7] M. Born and P. Jordan, “Zur Quantenmechanik”, *Zeitschrift für Physik*, 34, 858-888, 1925 (received September 27, 1925). [English translation in: B. L. van der Waerden, editor, *Sources of Quantum Mechanics* (Dover Publications, 1968) (English title: *On Quantum Mechanics*).]
- [8] M. Born, W. Heisenberg, and P. Jordan, “Zur Quantenmechanik II”, *Zeitschrift für Physik*, 35, 557-615, 1926 (received November 16, 1925). [English translation in: B. L. van der Waerden, editor, *Sources of Quantum Mechanics* (Dover Publications, 1968) (English title: *On Quantum Mechanics II*).]
- [9] E. Schrodinger, “Quantization as a Problem of Proper Values (Part I)” *Annalen der Physik* (4), vol. 79, (1926)

- [10] E. Schrodinger, “Quantization as a Problem of Proper Values (Part II)” *Annalen der Physik* (4), vol. 79, (1926)
- [11] E. Schrodinger, “Quantization as a Problem of Proper Values (Part III)” *Annalen der Physik* (4), vol. 80, (1926)
- [12] E. Schrodinger, “On the relation between the Quantum Mechanics of Heisenberg, Born and Jordan, and that of Schrodinger” *Annalen der Physik* (4), vol. 79, (1926)
- [13] P. A. M. Dirac, “The Principles of Quantum Mechanics”, (Oxford Unifersity Press, 4ed., 1958)
- [14] I. M. Gelfand, M. A. Naimark, “On the imbedding of normed rings into the ring of operators on a Hilbert space”, *Math. Sbornik* 12 (2) 197-217.(1943)
- [15] O. Bratteli, D. W. Robinson, “Operator Algebras and Quantum Statistical Mechanics 1”, (Springer, 1987)
- [16] J. C. A. Barata, “Curso de Física Matemática”, http://denebola.if.usp.br/~jbarata/Notas_de_aula (2013)
- [17] M. Reed, and B. Simon, “Methods of Modern Mathematical Physics Vol. I” (Academic Press, 1972)
- [18] R. de la Madrid, “The role of the rigged Hilbert space in Quantum Mechanics”, *Eur. J. Phys.* 26, 287 (2005); arXiv:quant-ph/0502053
- [19] I. M. Gelfand, N. Y. Vilenkin, “Generalized Functions, Vol. IV”, (Academic Press, 1964).
- [20] K. Maurin, “Generalized Eigenfunction Expansions and Unitary Representations of Topological Groups”, (Polish Scientific Publishers, 1968).
- [21] L. I. Schiff, “Quantum Mechanics”, (McGraw-Hill, 3ed., 1968)
- [22] J. J. Sakurai, “Modern Quantum Mechanics”, (Addison-Wesley, 1994)
- [23] C. Tannoudji, “Quantum Mechanics Vol. 1”, (John Willey, 1985).
- [24] P. A. M. Dirac, “A new notation for quantum mechanics”, *Mathematical Proceedings of the Cambridge Philosophical Society* 35 (3): pp. 416-418, (1939)