



NEL

Álgebra Linear

Com um pouco de Mecânica Quântica

Décio Krause

Coleção Rumos da Epistemologia 15

Álgebra Linear
Com um pouco de
Mecânica Quântica

Universidade Federal de Santa Catarina

Reitor: Luis Carlos Cancellier de Olivo

Departamento de Filosofia

Chefe: Nazareno Eduardo de Almeida

Programa de Pós-Graduação em Filosofia

Coordenador: Roberto Wu

NEL – Núcleo de Epistemologia e Lógica

Coordenador: Jonas Rafael Becker Arenhart

GLFC – Grupo de Lógica e Fundamentos da Ciência – UFSC/CNPq

Coordenador: Décio Krause

COLEÇÃO RUMOS DA EPISTEMOLOGIA, VOL. 15

DÉCIO KRAUSE

Álgebra Linear
Com um pouco de Mecânica
Quântica

NEL – Núcleo de Epistemologia e Lógica
GLFC – Grupo de Lógica e Fundamentos da Ciência – UFSC/CNPq
Universidade Federal de Santa Catarina
Florianópolis, 2016

© 2016, NEL – Núcleo de Epistemologia e Lógica, UFSC

ISBN: 978-85-87253-29-3 (papel)

978-85-87253-28-6 (e-book)

UFSC, Centro de Filosofia e Ciências Humanas, NEL

Caixa Postal 476

Bloco D, 2 andar, sala 209

Florianópolis, SC, 88010-970

(48) 3721-8612

nel@cfh.ufsc.br

<http://nel.ufsc.br>

FICHA CATALOGRÁFICA

Catálogo na fonte pela Biblioteca Universitária

da

Universidade Federal de Santa Catarina

K91a Krause, Décio

Álgebra Linear com um pouco de Mecânica Quântica
/ Florianópolis : NEL/UFSC, 2016.

XIV, 131 p. : gráfs., tabs.

(Coleção Rumos da Epistemologia ; v. 15)

Inclui bibliografia.

ISBN 978-85-87253-29-3 (papel)

ISBN 978-85-87253-28-6 (e-book)

1. Álgebra Linear 2. Mecânica Quântica. I. Título

CDU: 512.64

Reservados todos os direitos de reprodução total ou parcial por

NEL – Núcleo de Epistemologia e Lógica, UFSC.

Impresso no Brasil

Para Mercedes, a mesma pessoa de sempre.
Obrigado por uma vida.

Para o Professor Leo Barsotti, com meu agradecimento pelos ensinamentos de Álgebra Linear.

coleção
rumos da epistemologia

Editor: Jaimir Conte

Conselho Editorial: Alberto O. Cupani
Alexandre Meyer Luz
Cezar A. Mortari
Décio Krause
Gustavo A. Caponi
José A. Angotti
Luiz Henrique A. Dutra
Marco A. Franciotti
Sara Albieri



nel@cfh.ufsc.br
(48) 3721-8612

Núcleo de Epistemologia e Lógica
Universidade Federal de Santa Catarina

<http://nel.ufsc.br>
fax: (48) 3721-9751

Criado pela portaria 480/PRPG/96, de 2 de outubro de 1996, o NEL tem por objetivo integrar grupos de pesquisa nos campos da lógica, teoria do conhecimento, filosofia da ciência, história da ciência e outras áreas afins, na própria UFSC ou em outras universidades. Um primeiro resultado expressivo de sua atuação é a revista *Principia*, que iniciou em julho de 1997 e já tem dezessete volumes publicados, possuindo corpo editorial internacional. *Principia* aceita artigos inéditos, além de resenhas e notas, sobre temas de epistemologia e filosofia da ciência, em português, espanhol, francês e inglês. A Coleção Rumos da Epistemologia é publicada desde 1999, e a série Nel-lógica inicia sua publicação em 2014. Ambas aceitam textos inéditos, coletâneas e monografias, nas mesmas línguas acima mencionadas.

“A matemática [e poderíamos dizer que também pelo menos parte da física] origina-se de intuições, mas não pode ser nelas fundamentada.”

Ernst Zermelo

“Nos últimos anos, tornou-se-me cada vez mais difícil acompanhar e compreender os desenvolvimentos em Física [ele se referia em particular aos resultados da mecânica quântica]. Depois de tentar, cada vez mais exasperado e indeciso, finalmente me rendi em desespero. Isso deixou-me completamente exausto da vida. Sentia-me de fato condenado a continuar vivendo principalmente para prover às crianças os meios de subsistência. Tentei outras coisas, mas isso ajudava apenas momentaneamente. Portanto, concentro-me cada vez mais nos detalhes precisos do suicídio. Não tenho nenhuma outra opção viável senão o suicídio. [...] Perdoem-me ... Possam vocês e os seus ficarem bem.”

Paul Ehrenfest, físico austríaco, que se suicidou em 1933.

Sumário

Prefácio	x
Sobre o autor	xii
1 Espaços Vetoriais	1
1 Espaços vetoriais	1
2 Combinações lineares, superposições	8
3 Sub-espços vetoriais	10
4 Espaço gerado, base	13
5 Sobre espaços de dimensão infinita	16
6 Coordenadas de um vetor	17
7 Matriz de mudança de coordenadas	18
8 Existência de base	21
9 Espaços vetoriais isomorfos	23
10 Mais sobre dimensão infinita	26
10.1 Adendo: o axioma da escolha	27
2 Produtos Internos	29
1 Produtos internos	29
2 Espaços de Hilbert	35
2.1 Espaços de Hilbert e mecânica quântica	37
2.2 Lógica quântica	39
3 Ortogonalidade	41
3.1 Processo de Gram-Schmidt	42
3.2 Coeficientes de Fourier	43
4 A condição de normalização na teoria quântica	45
5 A notação de Dirac	47

3	Operadores Lineares	48
1	Representação matricial	50
1.1	O espaço dos operadores	53
1.2	O comutador	57
2	Funcionais Lineares	59
3	O espaço dual	61
4	Digressão ‘quântica’: spin	63
5	Auto-vetores e auto-valores	67
6	O exemplo das matrizes de Pauli	69
6.1	Matrizes de Pauli	70
7	Diagonalização	72
8	Matrizes e operadores ortogonais e unitários	75
4	Somas, somas diretas e projeções	80
1	Soma de sub-espços	80
2	Projeções	82
2.1	Notação para projeções	84
3	Resolução da identidade	86
4	A função traço	86
5	Complemento ortogonal	87
5	Produto tensorial	90
1	Produto tensorial	90
2	Emaranhamento	94
2.1	Posições definidas ?	96
2.2	O gato de Schrödinger	98
2.3	O fascínio pelo emaranhamento	100
6	Álgebra linear e mecânica quântica	106
1	Ignorância do estado: operador de densidade	108
2	O Princípio da Indeterminação	110
3	Uma formulação	112
3.1	Os postulados	115
	Referências Bibliográficas	123
	Índice Remissivo	126

Prefácio

UMA ESTRUTURA FUNDAMENTAL em matemática é a de espaço vetorial. Por mais abstrata que possa parecer, este tipo de estrutura é utilizada em muitas áreas da matemática aplicada e da física, podendo-se dizer que é extremamente relevante para a ciência contemporânea. Esta estrutura fundamenta a mais conhecida e utilizada formulação da mecânica quântica não relativista (doravante simplesmente ‘mecânica quântica’), disciplina que tem elevada importância também na filosofia da ciência atual. Por este motivo, e para adentrar aos problemas filosóficos trazidos por esta importante área do conhecimento, é fundamental que o estudante de filosofia se acerque de uma matemática mínima para entender as bases desta disciplina física.

Nestas notas, que são basicamente sobre álgebra linear, introduzimos com precisão o conceito de espaço vetorial e muitos outros relacionados a esta estrutura, mas não adentramos aos problemas filosóficos da mecânica quântica em profundidade, chegando apenas a apontar alguns deles, como o emaranhamento, a questão do realismo e outros poucos. Porém, chegamos a caracterizar precisamente o conceito de *espaço de Hilbert*, de enorme importância no formalismo usual da mecânica quântica,¹ com o propósito do estudante ir se acostumando com a terminologia e com as relações dos conceitos matemáticos com o desenvolvimento dessa importante teoria da física. Esperamos que o que aqui se apresenta seja útil para motivar o leitor a procurar textos mais completos, alguns dos quais indicados na bibliografia ao final.

Como este é um texto introdutório, demos destaque a espaços de dimensão

¹ A formulação da mecânica quântica utilizando-se espaços de Hilbert é a mais utilizada seja em física, seja em discussões filosóficas, mas há alternativas, algumas das quais muitos preferem, como a célebre formulação feita por R. Feynman via *integrais de caminho*. Em [Sty.02] são apresentadas *nove* maneiras de se obter a mecânica quântica padrão.

finita, ainda que os de dimensão infinita sejam mencionados aqui e acolá, principalmente nos exemplos. Este livro é dedicado prioritariamente a estudantes de filosofia, motivo pelo qual damos explicações informais (quando cabem) e introduzimos alguns conceitos filosóficos e lógicos. No entanto, estas notas podem igualmente ser de utilidade para estudantes de matemática, física ou de qualquer área que faça uso desses conceitos. Para seu uso em um curso de Álgebra Linear propriamente dita, na maioria das vezes, o texto poderá e talvez deva ser complementado com temas como matrizes e operações com matrizes, sistemas de equações lineares e outros assuntos não aqui cobertos porque demos atenção mais à parte conceitual.

Agradeço a meus alunos por terem suportado exposições longas sobre esses assuntos, e em especial a Lauro de Matos Nunes Filho, Joanne Simon Flausino e Raoni Wohnrath Arroyo, que tiveram a boa vontade de oferecer correções ao texto, e a Guilherme Mäder pelo auxílio na tradução da frase de Ehrenfest. As falhas que ainda permanecem são de minha inteira responsabilidade. Finalmente, agradeço aos responsáveis pela Coleção Rumos da Epistemologia e à direção do NEL pelo acolhimento deste trabalho. Agradeço ademais a Cezar A. Mortari pela gentil ajuda na formatação deste texto; os defeitos que ainda permanecem devem-se exclusivamente a mim mesmo.

Florianópolis, Julho de 2016
Décio Krause

Sobre o autor

DÉCIO KRAUSE é graduado em matemática (PUC/PR), fez o Mestrado em Educação na UFPR e aposentou-se como Professor Titular do Departamento de Matemática da UFPR. Fez o doutorado no Departamento de Filosofia da USP com uma tese em lógica, supervisionada por Newton C. A. da Costa, tendo se especializado em lógica e nos fundamentos lógicos e metafísicos da física quântica. Realizou estudos de pós-doutoramento nas universidades de Florença, Leeds e Oxford. É autor, junto com Steven French, de *Identity in Physics: A Historical, Philosophical, and Formal Analysis*, publicado pela Oxford U. Press (2006). Também é autor do livro *Introdução aos Fundamentos Axiomáticos da Ciência* (S. Paulo, EPU 2002). Tem no prelo o livro *Tópicos em Ontologia Analítica*, a sair pela editora da UNESP e, em conjunto com Jonas R. B. Arenhart, *The Logical Foundations of Scientific Theories*, a aparecer pela Routledge. É autor de mais de uma centena de artigos, capítulos de livros, resenhas, artigos de divulgação, organização de livros e de volumes, bem como de outros escritos, centrando-se na filosofia da lógica e dos fundamentos lógicos e metafísicos da mecânica quântica. Orientou 13 dissertações de mestrado e 6 de doutorado. Atualmente é Professor Titular do Departamento de Filosofia da UFSC e pesquisador do CNPq desde 1992 (atualmente, nível 1B).

ESPAÇOS VETORIAIS

1. Espaços vetoriais

O CONCEITO DE **espaço vetorial**, ou **espaço linear**, remonta a Hermann Grassmann (1809-1877), apesar de ter sido introduzido da forma como conhecemos hoje somente na década de 1920. Trata-se de um conceito fundamental em muitas áreas da matemática, pura e aplicada, e tem particular interesse na formulação usual da mecânica quântica, através de uma variante da teoria de tais espaços, conhecida como *teoria dos espaços de Hilbert*. Espaços de Hilbert, como teremos oportunidade de ver no que se segue, constituem um tipo particular de espaço vetorial.

Neste capítulo, algumas das principais noções relacionadas ao conceito de espaço vetorial serão introduzidas, sempre tendo-se em vista o caráter introdutório deste texto e sua destinação prioritária a estudantes de filosofia.

Um alerta inicial: grosso modo, uma **estrutura**, em termos matemáticos, nada mais é do que uma coleção adequada de domínios (que matematicamente são tomados como conjuntos) e de operações e elementos distinguidos desses domínios, sujeitos a axiomas (ou postulados) convenientes. Esta é basicamente a abordagem de N. Bourbaki (veja-se [Cor.92] para uma abordagem introdutória) à matemática, e que foi levada à física por vários cientistas, dentre os quais destacamos Patrick Suppes (1922-2014) [Sup.57, cap.12]. Basicamente, suas ideias permeiam nossa abordagem.

Iniciamos introduzindo o conceito básico:

Definição 1.1 (Espaço vetorial). Um **espaço vetorial** (ou **espaço linear**) é uma estrutura

$$\mathcal{E} = \langle \mathcal{V}, \mathcal{K}, +, \cdot \rangle,$$

onde:

1. \mathcal{V} é um conjunto não vazio cujos elementos são chamados de **vetores**.¹ Tais elementos serão designados por letras gregas minúsculas $\alpha, \beta, \psi, \dots$, mas mais tarde usaremos a notação de Dirac, escrevendo $|\alpha\rangle, |\beta\rangle, |\psi\rangle, \dots$, e os denominaremos de **kets** (por motivos que serão apresentados oportunamente). Excepcionalmente em exemplos envolvendo a mecânica quântica, esta notação será relaxada, mas o contexto deixará claro que se tratam de vetores, e daremos as explicações devidas.
2. \mathcal{K} é um corpo, ou seja, uma estrutura $\mathcal{K} = \langle K, +, \cdot, 0, 1 \rangle$ satisfazendo axiomas conhecidos.² Os elementos de K são denominados de **escalares**, e serão denotados por letras latinas minúsculas com ou sem índices. Os corpos dos quais faremos uso no que segue serão o corpo dos reais e o dos complexos. Outros corpos serão mencionados explicitamente.
3. $+$ é uma operação binária³ sobre \mathcal{V} , dita **adição de vetores**, de sorte que $\langle \mathcal{V}, + \rangle$ é um grupo comutativo. O elemento neutro deste grupo é

¹Como praticamente tudo em matemática, há uma história (ou pelo menos uma estória) de cada conceito, e portanto cabe mencionar que ‘vetor’, aqui, nada tem a ver com o que aprendemos na escola elementar, como algo que tem um sentido, uma direção e um comprimento, mas trata-se meramente do nome dos elementos do conjunto \mathcal{V} . Essas coisas, quando adequadamente formuladas, servem apenas como um *modelo* do que aqui se postula, a despeito de que certamente serviram de motivação para a definição.

²Os postulados são os seguintes: munido da adição escalares, K é um grupo comutativo com elemento neutro 0; $K - \{0\}$, munido da multiplicação, também é um grupo comutativo, sendo 1 o seu elemento neutro. Finalmente, a multiplicação é distributiva em relação à adição.

³Uma operação binária sobre um conjunto A é uma aplicação (função) de $A \times A$ em A . É um modo de se dizer que se está ‘pegando’ dois elementos de A em uma certa ordem e *operando* com eles, o resultado sendo ainda um elemento de A , o *composto* dos dois elementos. Se a operação for denotada aditivamente ($+$), o composto é chamado de *soma* dos elementos, e se for denotada multiplicativamente (\cdot), é chamado de *produto* dos elementos.

chamado de vetor nulo, e designado por $\mathbf{0}$ (não confundir $\mathbf{0}$ com o escalar 0).

4. \cdot é uma lei de composição externa⁴ sobre \mathcal{V} , ou seja, uma aplicação de $\mathcal{K} \times \mathcal{V}$ em \mathcal{V} , dita **multiplicação de vetor por escalar**. Esta operação satisfaz os seguintes postulados, para todos α e β em \mathcal{V} e todos $a, b \in K$:

$$(a) \quad a \cdot (\alpha + \beta) = a \cdot \alpha + a \cdot \beta$$

$$(b) \quad (a + b) \cdot \alpha = a \cdot \alpha + b \cdot \alpha$$

$$(c) \quad (a \cdot b) \cdot \alpha = a \cdot (b \cdot \alpha)$$

$$(d) \quad 1 \cdot \alpha = \alpha$$

Observação terminológica Doravante, escreveremos simplesmente $a\alpha$ para denotar $a \cdot \alpha$, bem como ab para $a \cdot b$. Observe que apesar de usarmos a mesma notação " \cdot " tanto para a multiplicação de vetor por escalar quanto para a multiplicação de escalares, elas não são a mesma operação. Usar símbolos distintos tornaria o texto muito carregado, de forma que prosseguiremos com a prática matemática usual de usar o mesmo símbolo para coisas diferentes. O contexto, no entanto, deixará claro quando se trata de uma ou de outra operação. O mesmo se aplica para a adição de vetores e para a adição de escalares, ambas denotadas por "+". Salientamos que no momento não há qualquer operação de multiplicação entre vetores. Isso será feito abaixo com a introdução das noções de produto interno, do produto de operadores e matrizes e de produto tensorial.

Quando temos um espaço vetorial $\mathcal{E} = \langle \mathcal{V}, \mathcal{K}, +, \cdot \rangle$, dizemos, mais uma vez por abuso de linguagem, que \mathcal{V} é **um espaço vetorial sobre K** , referindo-nos unicamente ao conjunto dos vetores e ao domínio do corpo de escalares (não confundir K com \mathcal{K}), ou que é um **K -espaço vetorial**. Nos casos particulares

⁴Uma lei de composição externa sobre um conjunto A com 'operadores' em um conjunto B , em geral denotada multiplicativamente, é via de regra denominada de *produto dos elementos de A pelos de B* , e pode ser à esquerda ou à direita, dependendo da posição dos operadores. Mais precisamente, uma *lei de composição externa* à esquerda sobre A é uma aplicação de $B \times A$ em A , e uma lei à direita é uma aplicação de $A \times B$ em A . Em física, como em geral são utilizados números como operadores, os físicos confundem as duas, deixando de fazer a distinção. Assim, se os elementos de A são denotados por letras gregas minúsculas e os operadores por letras latinas minúsculas, para eles $a \cdot \alpha$ é o mesmo que $\alpha \cdot a$.

de $K = \mathbb{R}$ ou de $K = \mathbb{C}$, falamos de **espaços vetoriais reais** ou de **espaços vetoriais complexos** respectivamente. Repetidas vezes usaremos esta terminologia.

O último postulado dado acima pode parecer o mais estranho e menos evidente de todos. Mas ele é fundamental. Se escrevemos $-\alpha$ para denotar o oposto de α , então parece sensato pedir que $-\alpha = (-1)\alpha$. No entanto, para provar este fato necessitamos do referido axioma. Antes, provamos o seguinte:

Teorema 1.1. *Para todo vetor α , tem-se que $0\alpha = \mathbf{O}$.*

Demonstração. De fato, $0\alpha = (0+0)\alpha = 0\alpha + 0\alpha$, donde $0\alpha = \mathbf{O}$. □

Nota 1.1. O símbolo \square chama-se *barra de Halmos* e foi inventada pelo matemático americano Paul R. Halmos (1916-2006) para indicar o fim de uma demonstração. É equivalente ao célebre ‘qed’ (*quod erat demonstrandum*, ou ‘como queríamos demonstrar’, *qcd* em nossa língua).

Exercício 1.1. Procure dar explicações detalhadas para a demonstração precedente.

Solução: A primeira igualdade vale-se de uma propriedade do corpo \mathcal{K} , a saber, que 0 é o elemento neutro da adição de escalares. A segunda igualdade faz uso do axioma (b) do item 4 da definição. A conclusão segue-se em virtude da unicidade do vetor nulo (que resulta do fato de que o conjunto dos vetores é um *grupo* quando munido da adição de vetores) e do fato de ele ser o único vetor que, somado a outro qualquer (no caso, ao vetor 0α), dá como resultado esse outro vetor. Ainda assim, convença-se de todos os detalhes; é um excelente exercício.

Teorema 1.2 (Unicidade do oposto). *O oposto $-\alpha$ de um vetor α é único.*

Demonstração. Usando **redução ao absurdo**,⁵ e sendo $-\alpha$ o oposto de α , suponha que α tem mais de um oposto; seja $-\alpha'$ um ‘outro’ oposto de α . Mostraremos que eles são iguais, ou seja, há um só oposto. Com efeito, $-\alpha = -\alpha + \mathbf{O} = -\alpha + (\alpha + -\alpha') = (-\alpha + \alpha) + -\alpha' = \mathbf{O} + -\alpha' = -\alpha'$. □

⁵Repare aqui a importância da lógica subjacente, nem sempre explicitada em textos matemáticos ou de física. Como há lógicas (por exemplo, a intuicionista) em que a redução ao absurdo não vale em geral, uma demonstração como esta não poderia ser feita caso esta fosse a lógica utilizada. Guarde isso: lógica é importante.

Exercício 1.2. Explique a demonstração precedente explicitando todos os seus passos.⁶

Exercício 1.3. Preencha os detalhes de uma afirmação feita acima, a saber, que o conjunto \mathcal{V} de vetores é um grupo (na verdade, um grupo comutativo) quando munido da operação de adição de vetores. Prove usando redução ao absurdo que o vetor nulo, que é o elemento neutro deste grupo, é único.

Agora, podemos estabelecer o pretendido:

Teorema 1.3. *Para todo vetor α , seu oposto $-\alpha$ é obtido multiplicando-se o vetor por -1 , ou seja, $-\alpha = (-1)\alpha$.*

Demonstração. Temos que $\alpha + (-1)\alpha = 1\alpha + (-1)\alpha = (1 - 1)\alpha = 0\alpha = \mathbf{O}$. Portanto, dada a unicidade do oposto de α , resulta que $-\alpha = (-1)\alpha$. \square

Exemplos importantes de espaços vetoriais são os seguintes. O estudante ganharia muito se preenchesse todos os detalhes que achar necessários para esclarecimento em cada um dos exemplos dados.

Exemplo 1.1. Seja $\mathbb{R}^n = \{(x_1, \dots, x_n) : x_i \in \mathbb{R}\}$ o conjunto das n -uplas de números reais. Municemos este conjunto com as operações seguintes, onde $k \in \mathbb{R}$, para obter um espaço vetorial real:⁷

⁶ O 'método 'clássico' de redução ao absurdo — porque há, por exemplo, a **redução ao absurdo intuicionista** que não é a mesma coisa — apregoa que se queremos provar A , iniciamos supondo que A seja falso, ou seja (de acordo com a lógica clássica), que a negação de A seja verdadeira. Então, da negação de A , derivamos uma contradição. Como nada, ainda na lógica clássica, que seja verdadeiro pode implicar uma contradição, então a negação de A deve ser falsa, o que implica que A é verdadeira. Bem, você é desafiado a entrar nos detalhes desta explicação superficial e identificar todas as hipóteses que estão sendo pressupostas, que em geral são tomadas como assentadas, mas que dependem da lógica utilizada, inclusive na palavra *verdadeira*, usada livremente (e indevidamente! — com efeito, em lógica estabelece-se uma diferença fundamental entre 'verdade' e 'demonstrabilidade'. Para uma visão geral, ver [Hen.79]). Dito de modo breve, a redução ao absurdo intuicionista diz que se de uma hipótese A derivamos duas sentenças contraditórias, então podemos inferir a negação de A . Note que não supomos, como no caso clássico, que A é falsa a princípio, e não inferimos A , mas sua negação. Ver [Kle.52, p.99].

⁷Mais uma notação comum em matemática. O símbolo $:=$ significa **igual por definição**, e trata-se de uma equivalência metalinguística, indicando que a operação à sua

1. $(x_1, \dots, x_n) + (y_1, \dots, y_n) := (x_1 + y_1, \dots, x_n + y_n)$ (o leitor deve perceber que, à esquerda da igualdade, "+" denota a adição de vetores, mas à direita denota a adição de escalares).
2. $k(x_1, \dots, x_n) := (kx_1, \dots, kx_n)$ (idem observação acima com respeito à multiplicação de vetor por escalar, à esquerda da igualdade e a multiplicação de escalares, à direita).

O espaço vetorial do exemplo precedente será denominado de \mathbb{R}^n . De maneira semelhante, definimos o espaço complexo \mathbb{C}^n tomando operações análogas às acima, somente que consideradas agora sobre \mathbb{C} .

Exemplo 1.2. Considere o conjunto \mathcal{F} das funções reais de variável real com mesmo domínio, digamos o intervalo $[a, b] \subseteq \mathbb{R}$. Para $f, g \in \mathcal{F}$, definimos as operações seguintes, aqui colocadas porque muitas vezes serão requisitadas mais tarde:

1. $(f + g)(x) := f(x) + g(x)$
2. $(kf)(x) := kf(x)$

É fácil ver que resulta um espaço vetorial real, cujo vetor nulo é a função nula $n(x) = 0$.

Exemplo 1.3. Seja $\mathbb{R}^{n \times m}$ o conjunto das matrizes reais de ordem $n \times m$, munido das operações usuais de adição de matrizes e de multiplicação de matriz por escalar real. Neste caso, tem-se um espaço vetorial real, cujo vetor nulo é a matriz nula $n \times m$.⁸

Exercício 1.4. Mostre que temos um espaço vetorial real se tomarmos o conjunto \mathbb{R}^+ dos reais não negativos munido das operações $x + y := xy$ e $kx := x^k$, esquerda está sendo introduzida (na metalinguagem) pela expressão à sua direita. Equivale a outros que aparecem também frequentemente, como $=_{\text{def}}$ ou $\stackrel{\text{def}}{=}$. Um alerta aos interessados em fundamentos: podemos com efeito introduzir novos símbolos na linguagem da teoria que estamos considerando, mas é preciso certo cuidado, que é explicado em [Sup.57, Cap.8].

⁸Bom, não dá para explicar tudo. Aquilo que achamos que o leitor pode procurar por si mesmo será deixado ao seu encargo, como conhecer fatos básicos sobre matrizes, funções, etc.. (mas mais abaixo explicamos o que é o produto de matrizes ... – página 1).

para $x, y \in \mathbb{R}^+$ e $k \in \mathbb{R}$. Repare que, nos primeiros membros, estão as operações de espaços vetoriais, enquanto que, nos segundos membros, estão as propriedades de adição e de exponenciação nos reais. Veja o paralelo que há entre as operações definidas e as propriedades do logaritmo.

Exercício 1.5 (Importante). Todo corpo $\mathcal{K} = \langle K, +, \cdot, 0, 1 \rangle$ pode ser visto como um espaço vetorial sobre K . Dizemos que **todo corpo é um espaço vetorial sobre si mesmo**. Procure entender este fato e só depois, se necessário, leia a explicação a seguir.

Explicação O espaço vetorial em questão, em termos da definição dada, torna-se a estrutura $\mathfrak{R} = \langle K, \mathcal{K}, +, \cdot \rangle$, isto é, os *vetores* confundem-se com os escalares, assim como as operações entre vetores e entre vetores e escalares confundem-se com as correspondentes operações entre escalares. É um exercício verificar que dá tudo certo (ou seja, que a estrutura é de fato um K -espaço vetorial).

Exercício 1.6. Mostre que se tomarmos \mathbb{C} como conjunto de vetores e \mathbb{R} como conjunto de escalares, e considerando a adição de números complexos como adição de vetores e a multiplicação de número complexo por número real como a multiplicação de vetor por escalar, resulta um espaço vetorial real (lembre que o espaço é *real* se o corpo de escalares for \mathbb{R}).

Exercício 1.7. Mostre que se tomarmos \mathbb{C} como conjunto de vetores e o próprio \mathbb{C} como conjunto de escalares, e considerando a adição de números complexos como adição de vetores e a multiplicação de números complexos como a multiplicação de vetor por escalar, resulta um espaço vetorial complexo.

Exercício 1.8. Mostre que se tomarmos \mathbb{R} como conjunto de vetores e \mathbb{R} como conjunto de escalares, e considerando a adição de números reais como adição de vetores e a multiplicação de números reais como a multiplicação de vetor por escalar, resulta um espaço vetorial real.

Exercício 1.9. Justifique porque não resulta espaço vetorial se tomarmos \mathbb{R} como conjunto de vetores e \mathbb{C} como conjunto de escalares, e considerando a adição de números reais como adição de vetores e a multiplicação de número real por número complexo como a multiplicação de vetor por escalar.

Exercício 1.10. Justifique porque o conjunto dos polinômios de grau 3 com coeficientes reais, ou seja, entidades da forma $p(x) = a_0 + a_1x + a_2x^2 + a_3x^3$, com

$a_3 \neq 0$, munido das operações de adição de polinômios e de multiplicação de polinômio por número real **não é** um espaço vetorial real.⁹

Exercício 1.11. Mostre agora que o conjunto dos polinômios reais de grau **menor ou igual** a 3 (ou seja, sem se exigir que a_3 não seja nulo), munido das mesmas operações, agora é um espaço vetorial real.

2. Combinações lineares, superposições

Uma **combinação linear** de vetores é a soma desses vetores, eventualmente multiplicados por escalares, como $\beta = x_1\alpha_1 + x_2\alpha_2 + \dots + x_n\alpha_n$. Uma tal expressão é por vezes denominada de **superposição** desses vetores. Isso terá importância em física. Em especial, essa disciplina interessa-se por superposições nas quais se tenha

$$\sum_{i=1}^n |x_i|^2 = 1,$$

que chamaremos de *condição de normalização*. O motivo é que os escalares x_i representarão probabilidades, e sua soma deverá ser igual à unidade. Este conceito, no entanto, poderá ser introduzido somente mais tarde.

Definição 2.1 (Dependência e Independência linear). Um conjunto de vetores $A = \{\alpha_1, \dots, \alpha_n\}$ é **linearmente independente** (ou os vetores de A são linearmente independentes) se uma combinação linear da forma

$$x_1\alpha_1 + \dots + x_n\alpha_n = \mathbf{O}$$

acarreta que $x_1 = \dots = x_n = 0$ (Não confunda o escalar 0 com o vetor nulo \mathbf{O}). Caso contrário, o conjunto A (ou os seus vetores) são **linearmente dependentes**.

⁹Dica: não é fechado para as operações. Um conjunto A sobre o qual está definida uma operação binária \star é **fechado** para esta operação se a composição de quaisquer dois elementos a, b de A fornece um elemento $a \star b$ ainda em A . No caso da multiplicação de vetor por escalar do exemplo, a soma de quaisquer dois polinômios de grau 3 deve dar ainda um polinômio de grau 3 e a multiplicação de um desses polinômios por um escalar deve ainda dar um de tais polinômios.

Equivalentemente, A (ou os seus vetores) é linearmente dependente se podemos encontrar uma combinação linear nula (como a acima) com pelo menos um dos escalares x_j diferente de 0. Por exemplo, o conjunto (os vetores) $A = \{(1, 2), (-1, 1)\}$ do \mathbb{R}^2 é linearmente independente, já que

$$x_1(1, 2) + x_2(-1, 1) = (0, 0)$$

acarreta $x_1 = x_2 = 0$. Por outro lado os vetores $(1, 3)$ e $(-2, -6)$ são linearmente dependentes, como é fácil verificar, pois podemos escrever a combinação linear nula $-2 \cdot (1, 3) + 1 \cdot (-2, -6) = (0, 0)$ sem que os coeficientes sejam todos necessariamente nulos.

A recíproca, porém, não vale em geral. Ou seja, podemos ter uma combinação linear nula com vetores linearmente dependentes, bastando para tanto tomar todos os coeficientes iguais a zero.

Exercício 2.1. Comente os detalhes da afirmação feita na última frase acima.

Exercício 2.2. Verifique se cada conjunto de vetores a seguir, de algum espaço \mathbb{R}^n , é linearmente dependente ou independente.

1. Em \mathbb{R}^2 , o conjunto $A = \{(1, -2), (2, 1)\}$
2. Em \mathbb{R}^2 , o conjunto $A = \{(1, -2), (2, -4)\}$
3. Em \mathbb{R}^3 , o conjunto $A = \{(1, -2, 0), (2, -1, -4), (0, 0, 1)\}$
4. Em \mathbb{R}^3 , o conjunto $A = \{(0, -1, 0), (1, 0, -1), (0, 0, 1)\}$

Convenção Neste contexto, convencionou-se que **o conjunto vazio forma um conjunto linearmente independente de vetores.**

Exercício 2.3. Mostre que qualquer conjunto de vetores que contenha o vetor nulo é linearmente dependente.

Exercício 2.4. Idem para qualquer conjunto de vetores que tenha um dos seus elementos como combinação de outros vetores do conjunto.

Exercício 2.5. Mostre que o conjunto de matrizes reais abaixo é linearmente independente

$$A = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \right\}$$

3. Sub-espaços vetoriais

Informalmente, definiremos o que significa *restringir* uma operação a um conjunto. Primeiramente, vejamos o caso de operações binárias. Seja A um conjunto sobre o qual está definida a operação binária $*$, e seja $B \subseteq A$. Alternativamente, $*$ pode ser identificada com o conjunto das triplas ordenadas da forma $(a, b, a * b)$, com $a, b \in A$. Restrinjamos agora o conjunto dessas triplas, considerando unicamente aquelas tais que $a, b \in B$. É imediato que tal coleção é uma função de $B \times B$ em B , dita **restrição** da operação $*$ (definida sobre A), ao subconjunto B . Da mesma forma, se \diamond é uma lei de composição externa sobre A , se restringirmos os elementos considerados a apenas aqueles que pertencem a B , obteremos uma lei de composição externa sobre B , também dita *restrição* (a B) da lei \diamond .

Por exemplo, considere a adição de números reais e agora considere esta operação aplicada unicamente ao subconjunto dos reais que é isomorfo ao conjunto dos números inteiros.¹⁰ Temos então (novamente por abuso de linguagem) uma restrição da operação de adição aos inteiros.

Definição 3.1 (Sub-espaço vetorial). Sejam $\mathcal{E} = \langle \mathcal{V}, \mathcal{K}, +, \cdot \rangle$ e $\mathcal{W} \subseteq \mathcal{V}$, $\mathcal{W} \neq \emptyset$. Então \mathcal{W} é **sub-espaço vetorial** de \mathcal{V} se $\langle \mathcal{W}, \mathcal{K}, +_{\mathcal{W}}, \cdot_{\mathcal{W}} \rangle$ é um espaço vetorial sobre \mathcal{K} , sendo $+_{\mathcal{W}}$ e $\cdot_{\mathcal{W}}$ restrições das operações de $+$ e \cdot a \mathcal{W} .

Teorema 3.1 (Importante). *Uma condição necessária e suficiente para que \mathcal{W} seja sub-espaço vetorial de \mathcal{V} é que, para todos $\alpha, \beta \in \mathcal{W}$ e para todo $k \in \mathcal{K}$, se tenha:*

1. Se $\alpha, \beta \in \mathcal{W}$, então $\alpha + \beta \in \mathcal{W}$
2. Se $\alpha \in \mathcal{W}$, então $k\alpha \in \mathcal{W}$

Demonstração. Quanto à necessidade dessas condições, ela segue do fato de que um espaço vetorial deve ser fechado relativamente às operações de adição de vetores e de multiplicação de vetor por escalar.¹¹ Quanto à suficiência dessas condições, basta verificar que delas resultam as condições da definição de espaço vetorial para $\langle \mathcal{W}, \mathcal{K}, +_{\mathcal{W}}, \cdot_{\mathcal{W}} \rangle$. Por exemplo, como $\mathcal{W} \neq \emptyset$, existe $\alpha \in \mathcal{W}$, logo $-\alpha = (-1)\alpha \in \mathcal{W}$ (pela segunda condição). Analogamente, pela primeira

¹⁰Ou seja, considere somente aqueles reais que são inteiros.

¹¹Ou seja, o resultado da adição de vetores do espaço tem que ser ainda um vetor do espaço e a multiplicação de um vetor do espaço por um escalar tem que dar ainda um vetor no espaço.

condiçaõ, $O = \alpha - \alpha \in \mathcal{W}$. Quanto às demais propriedades, em virtude de elas valerem para todos os vetores de \mathcal{V} , valerão em particular para os vetores de \mathcal{W} , ou seja, são "herdadas" por \mathcal{W} . \square

Da mesma forma que antes, o estudante deve completar os detalhes nos exemplos a seguir de forma que os compreenda adequadamente.

Exemplo 3.1. Sendo $\mathcal{W} = \{O\}$, sendo O o vetor nulo de \mathcal{V} , então \mathcal{W} é sub-espço vetorial de \mathcal{V} . Este sub-espço é chamado de **sub-espço trivial** de \mathcal{V} . Você é desafiado a dar uma demonstração deste fato.

Exemplo 3.2. O conjunto das matrizes simétricas de ordem n sobre o corpo dos números complexos é um sub-espço do espaço das matrizes reais de ordem n . Especifique o exemplo indicando o que são os vetores, os escalares e as operações relevantes.

Exemplo 3.3. O conjunto das funções reais contínuas no intervalo $[a, b]$ é um sub-espço do espaço vetorial dado no exemplo 1.2. Idem observação do exemplo precedente.

Exemplo 3.4. Consideremos o espaço real \mathbb{R}^3 das triplas ordenadas de números reais (um caso particular do espaço \mathbb{R}^n do exemplo 1.1). Os seguintes sub-conjuntos são sub-espços do \mathbb{R}^3 , e serão importantes abaixo para exemplos. Os nomes dados a esses espaços têm em mente (intuitivamente) um sistema de coordenadas cartesianas ortogonais para o espaço tridimensional.

1. $X = \{(x, 0, 0) : x \in \mathbb{R}\}$ ("eixo X").
2. $Y = \{(0, y, 0) : y \in \mathbb{R}\}$ ("eixo Y").
3. $Z = \{(0, 0, z) : z \in \mathbb{R}\}$ ("eixo Z").
4. $XY = \{(x, y, 0) : x, y \in \mathbb{R}\}$ ("plano XY").
5. $XX = \{(x, 0, z) : x, z \in \mathbb{R}\}$ ("plano XZ").
6. $YZ = \{(0, y, z) : y, z \in \mathbb{R}\}$ ("plano YZ").
7. $P = \{(x, y, z) : ax + by + cz = 0, a, b, c \neq 0\}$ (plano passando pela origem).
8. $R = \{(x, y, z) : \frac{x}{a} = \frac{y}{b} = \frac{z}{c}, a, b, c \neq 0\}$ (reta passando pela origem).

Exercício 3.1. Prove que cada um dos casos do exemplo anterior de fato define um sub-espaço do \mathbb{R}^3 .

Exercício 3.2. Mostre que o conjunto das quádruplas de números reais com as duas últimas componentes nulas, ou seja, da forma $(x_1, x_2, 0, 0)$, com x_1 e x_2 em \mathbb{R} , munido das operações usuais de adição e de multiplicação é um sub-espaço vetorial do \mathbb{R}^4 .

Exercício 3.3. Mostre que o conjunto das matrizes reais simétricas de ordem 2 é um sub-espaço de $\mathbb{R}^{2 \times 2}$.

Para os fundamentos da física quântica, é importante observarmos o seguinte.

Teorema 3.2. *A interseção de sub-espaços de um espaço vetorial é ainda um sub-espaço desse espaço.*

Demonstração. Sejam \mathcal{W}_1 e \mathcal{W}_2 sub-espaços de um K -espaço vetorial \mathcal{V} , e seja $\mathcal{W} = \mathcal{W}_1 \cap \mathcal{W}_2$. Então, se α e β pertencem a \mathcal{W} , pertencem a \mathcal{W}_1 e a \mathcal{W}_2 . Como por hipótese ambos são sub-espaços de \mathcal{V} , $\alpha + \beta$ pertence a ambos (um sub-espaço é fechado para a adição de vetores). Logo, ambos pertencem a \mathcal{W} . Agora, suponha que $\alpha \in \mathcal{W}$ e que $k \in K$. Logo $\alpha \in \mathcal{W}_1$ e $\alpha \in \mathcal{W}_2$. Como são ambos sub-espaços, segue-se que $k\alpha$ pertence a ambos os subespaços (pelo fechamento relativamente à multiplicação de vetor por escalar). Logo, $k\alpha \in \mathcal{W}$. \square

A união de sub-espaços, no entanto, não é em geral um sub-espaço (podendo ser eventualmente, dependendo dos sub-espaços). Por exemplo, seja $\mathcal{V} = \mathbb{R}^3$, e $\mathcal{W}_1 = \{(x, 0, 0) : x \in \mathbb{R}\}$ ("eixo X"), enquanto que $\mathcal{W}_2 = \{(0, y, 0) : y \in \mathbb{R}\}$ ("eixo Y"), como no exemplo 3.4 acima. Ora, $\mathcal{W}_1 \cup \mathcal{W}_2 = \{\alpha \in \mathbb{R}^3 : \alpha \in \mathcal{W}_1 \vee \alpha \in \mathcal{W}_2\}$, o que significa que esses vetores estão no eixo X ou no eixo Y (somente o vetor nulo está em ambos). Porém, a soma de dois vetores não nulos quaisquer $\alpha_1 \in \mathcal{W}_1$ e $\alpha_2 \in \mathcal{W}_2$ não pertence a nenhum dos sub-espaços, logo a união não é fechada para a adição de vetores.

Há porém um 'menor' sub-espaço de \mathcal{V} que contém a união de sub-espaços, a saber, o **espaço gerado pela união**. Estes fatos são importantes para o formalismo da mecânica quântica, pois nos interessa a estrutura algébrica do conjunto dos sub-espaços de um espaço vetorial munido de operações adequadas, inspiradas nos fatos acima: interseção de sub-espaços e o espaço gerado pela união de sub-espaços.

4. Espaço gerado, base

Seja $\mathcal{E} = \langle \mathcal{V}, \mathcal{K}, +, \cdot \rangle$ um espaço vetorial e $A = \{\alpha_1, \dots, \alpha_n\}$ um conjunto de vetores de \mathcal{V} . Temos então:

Definição 4.1. Chama-se **espaço gerado** por A ao conjunto

$$[A] = \{\beta \in \mathcal{V} : \beta = \sum_{i=1}^n x_i \alpha_i, x_i \in \mathcal{K}\}.$$

Ou seja, o conjunto $[A]$ é conjunto de todos os vetores que são combinações lineares dos vetores de A . Mostrar-se-á agora que tal conjunto, munido das operações do espaço vetorial do qual é um sub-conjunto, é por si um sub-espaço vetorial do espaço dado.

Teorema 4.1. *O conjunto $[A]$ é um sub-espaço vetorial de \mathcal{E} .*

Demonstração. Basta notar que a soma de vetores de $[A]$ é ainda um vetor de $[A]$, bem como a multiplicação de qualquer de seus vetores por um escalar (assim cumprindo as condições do teorema (3.1). \square

Convenção Convencionou-se que $[\emptyset] = \{O\}$.

O espaço gerado por um conjunto de vetores é, portanto, o conjunto de todas as combinações lineares desses vetores. Perceba que se o conjunto for linearmente dependente, alguns de seus vetores podem ser escritos como combinações lineares dos demais, de forma que, para obter o espaço gerado, esses vetores podem ser suprimidos, resultando o seguinte

Teorema 4.2. *Dado um conjunto de vetores A , existe sempre um subconjunto de A linearmente independente que gera o mesmo espaço que A .*

Demonstração. Seja $A = \{\alpha_1, \dots, \alpha_n\}$ o conjunto em questão, que supomos ser linearmente dependente. Portanto, há um vetor α_j que pode ser escrito simplificadaamente como

$$\alpha_j = \sum_{i \neq j} k_i \alpha_i,$$

ou seja, ele é combinação linear dos demais vetores de A . Se $\beta \in [A]$, então

$$\beta = x_1 \alpha_1 + \dots + x_j \alpha_j + \dots + x_n \alpha_n = \sum_{i \neq j} x_i \alpha_i + x_j \alpha_j,$$

ou seja,

$$\beta = \sum_{i \neq j} x_i \alpha_i + x_j \sum_{i \neq j} k_i \alpha_i = \sum_{i \neq j} (x_i + k_i) \alpha_i,$$

o que mostra ser β combinação linear dos vetores de A , exceto α_j . Se $A - \{\alpha_j\}$ for linearmente independente, é o conjunto procurado (é fácil comprovar que A e $A - \{\alpha_j\}$ geram o mesmo espaço). Se ainda for linearmente dependente, há um vetor nesse conjunto que é combinação linear dos demais, e o processo pode ser repetido até que restem unicamente vetores linearmente independentes, que continuarão gerando o mesmo espaço. \square

Definição 4.2 (Base de um espaço vetorial). Uma **base** para um espaço vetorial \mathcal{E} é um conjunto \mathcal{A} de vetores de \mathcal{V} que satisfaz as condições seguintes:

1. \mathcal{A} é linearmente independente
2. \mathcal{A} gera \mathcal{E} , ou seja, todo vetor de \mathcal{V} é combinação linear dos vetores de \mathcal{A} .

O conjunto \mathcal{A} tem um cardinal, que no caso finito pode ser entendido intuitivamente como designando a quantidade de elementos de \mathcal{A} . Pode-se demonstrar¹² que todas as bases de um espaço vetorial têm a mesma cardinalidade (porém, veja a discussão abaixo sobre a existência de bases). Este cardinal chama-se *dimensão* do espaço vetorial.

Definição 4.3 (Dimensão). Chama-se **dimensão** de um espaço vetorial ao cardinal de uma base desse espaço.

Por exemplo, o espaço \mathbb{R}^n tem dimensão n , pois tem o conjunto (com n elementos)

$$\Xi = \{\epsilon_1, \dots, \epsilon_n\}, \tag{1.1}$$

onde $\epsilon_i = (0, \dots, 1, \dots, n)$ (com o 1 na i -ésima posição) como uma base. Esta base é dita **base canônica** do \mathbb{R}^n .

Da mesma forma, se olharmos agora os vetores ϵ_i como formados por números complexos, então Ξ também representa uma base (canônica) para o espaço complexo \mathbb{C}^n . Essas bases desempenharão papel relevante à frente.

A dimensão do sub-espaço trivial é zero.

¹²Como veremos depois, a demonstração desse fato depende do Axioma da Escolha.

Exercício 4.1. Mostre que o conjunto

$$\mathcal{X} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

é uma base para o espaço das matrizes complexas (reais) de ordem 2, também dita base canônica para esses espaços.

Exercício 4.2. Justifique informalmente porque a dimensão do espaço das matrizes de ordem $m \times n$ sobre K é $m.n$.

Há dois teoremas relacionados aos conceitos de dependência e independência linear e de base e dimensão que merecem destaque, e que serão utilizados mais tarde na demonstração de outros resultados. São os seguintes, aqui somente enunciados (para uma demonstração do primeiro, ver [Bar.76, p.81]). O primeiro afirma que **qualquer conjunto linearmente independente de vetores pode ser estendido a uma base**.

Teorema 4.3 (Teorema de Steinitz). *Suponha que a dimensão de \mathcal{V} seja n , e que $\mathcal{B} = \{\alpha_1, \dots, \alpha_k\}$ seja um conjunto linearmente independente de vetores. Então existem vetores $\alpha_{k+1}, \dots, \alpha_n$ tais que $\mathcal{A} = \{\alpha_1, \dots, \alpha_k, \alpha_{k+1}, \dots, \alpha_n\}$ é uma base para \mathcal{V} .*

Teorema 4.4. *Assuma que a dimensão de \mathcal{V} é n . Então:*

1. *Qualquer conjunto com mais de n vetores é linearmente dependente*
2. *Nenhum conjunto com menos de n vetores pode gerar \mathcal{V}*

Ou seja, uma base é um conjunto linearmente independente ‘maximal’.

Exercício 4.3. Aqui vamos resolver um exercício proposto acima, mostrando que qualquer conjunto de vetores que contenha o vetor nulo como um de seus elementos é linearmente dependente.

A solução é simples. Suponha o conjunto $\{\alpha_1, \dots, \alpha_k, \mathbf{O}\}$. Formemos a combinação linear nula

$$x_1\alpha_1 + \dots + x_k\alpha_k + k\mathbf{O} = \mathbf{O}.$$

É claro que isso não implica que todos os escalares devam ser nulos, pois por exemplo k pode ser diferente de zero.

Exercício 4.4. Encontre uma base para o espaço vetorial das matrizes reais simétricas de ordem 3 e indique sua dimensão.¹³

Exercício 4.5. Mostre que o espaço vetorial $\mathfrak{R} = \langle K, \mathcal{K}, +, \cdot \rangle$ dado acima no exercício 1.5 tem dimensão 1 e que qualquer conjunto contendo somente um escalar não nulo é um conjunto linearmente independente de vetores que pode ser tomado como uma base de tal espaço.

Exercício 4.6. Mostre que o conjunto $\{1, i\}$ é uma base para o espaço vetorial do exercício 1.6, que portanto tem dimensão 2.

Exercício 4.7. Mostre que o conjunto $\{1\}$ é uma base para o espaço vetorial do exercício 1.7, que portanto tem dimensão 1.

Exercício 4.8. Mostre que o conjunto $\{1\}$ é uma base para o espaço vetorial do exercício 1.8, que portanto também tem dimensão 1.

Note a diferença entre os espaços dos dois últimos exercícios. No penúltimo, 1 é visto como um número complexo, e os escalares são também números complexos, assim que qualquer complexo $a + bi$ pode ser obtido como uma combinação linear do vetor do conjunto proposto, a saber, $a + bi = 1 \cdot (a + bi)$. Já no último, 1 deve ser visto como um número real e os escalares são números reais. É claro que qualquer real a pode ser obtido como $a = 1 \cdot a$.

5. Sobre espaços de dimensão infinita

Há espaços vetoriais que têm dimensão infinita, como o acima mencionado espaço das funções reais com domínio no intervalo $[a, b]$. É preciso cuidado nesses casos, pois as operações de adição de vetor e de multiplicação por escalar são definidas para trabalharmos com um número finito de vetores. Mas, o que seria uma combinação linear de infinitos vetores? A convenção que se adota é a de que quando dizemos que um espaço é gerado por um conjunto infinito de vetores, queremos dizer que cada vetor do espaço é uma combinação linear *finita* de vetores desse conjunto. Da mesma forma, quando dizemos que um conjunto infinito de vetores é linearmente independente, isso significa que toda combinação linear *finita* desses vetores, quando igualada ao vetor nulo, implica

¹³Uma matriz quadrada é simétrica se for igual à sua transposta – a matriz obtida trocando-se ordenadamente as linhas pelas colunas.

que todos os coeficientes resultam nulos. Assim sendo, um conjunto infinito de vetores é uma base para um espaço se todo vetor pode ser escrito *de modo único* como combinação linear *finita* de vetores da base.

Um bom e simples exemplo é considerarmos o espaço vetorial dos polinômios com coeficientes reais. Um tal polinômio de grau n pode ser assim representado:

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n, \quad (1.2)$$

com todos os $a_j \in \mathbb{R}$. As operações são as seguintes (é um exercício mostrar que de fato temos um espaço vetorial): sendo $q(x) = b_0 + b_1x + b_2x^2 + \cdots + b_nx^n$, definimos

$$(p+q)(x) := (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_n + b_n)x^n$$

$$(kp)(x) := (ka_0) + (ka_1)x + \cdots + (ka_n)x^n, \text{ para todo } k \text{ real.}$$

Como podemos variar n como quisermos, obtendo polinômios da ordem que quisermos, e fica patente que este espaço não pode ser gerado por um conjunto finito de vetores, ainda que qualquer polinômio $r(x)$ possa ser obtido como combinação linear de polinômios da forma $p_0(x) = k$ (para k real), $p_1(x) = x$, $p_2(x) = x^2$, etc. É fácil ver que qualquer combinação linear $\lambda_0k + \lambda_1x + \lambda_2x^2 + \cdots + \lambda_nx^n = 0$ implica que $\lambda_0 = \lambda_1 = \cdots = \lambda_n = 0$, o que mostra que eles são linearmente independentes. Assim, o conjunto $\{p_0, p_1, \dots\}$ é uma base para o espaço dos polinômios com coeficientes reais, que tem dimensão infinita.

Exercício 5.1. Relativamente ao exemplo do espaço dos polinômios acima referido, qual seria o vetor nulo? O que seria o oposto de um polinômio $p(x)$?

6. Coordenadas de um vetor

Um conceito importante é o de **matriz das coordenadas** de um vetor em uma base ordenada. Seja $\mathcal{A} = \{\alpha_1, \dots, \alpha_n\}$ uma base ordenada para o espaço vetorial $\mathcal{E} = \langle \mathcal{V}, \mathcal{K}, +, \cdot \rangle$. Se $\beta \in \mathcal{V}$, podemos escrever $\beta = x_1\alpha_1 + \cdots + x_n\alpha_n$. Denomina-se de matriz das coordenadas do vetor β na base ordenada \mathcal{A} à matriz linha (com uma linha e n colunas)¹⁴

$$[\beta]_{\mathcal{A}} = [x_1, x_2, \dots, x_n]. \quad (1.3)$$

¹⁴Contrariando a notação usual de matrizes, usaremos vírgulas para separar seus elementos.

Por exemplo, $\mathcal{A} = \{(1, 1), (-1, 2)\}$ é uma base ordenada para o \mathbb{R}^2 , como é fácil provar (é linearmente independente e todo vetor $(x, y) \in \mathbb{R}^2$ pode ser escrito como combinação linear desses vetores). Seja $\beta = (2, 3)$. Então, $(2, 3) = x_1(1, 1) + x_2(-1, 2)$, ou $(2, 3) = (x_1 - x_2, x_1 + 2x_2)$, o que fornece $x_1 = 7/3$ e $x_2 = 1/3$. Assim,

$$[(2, 3)]_{\mathcal{A}} = [7/3, 1/3].$$

Importante observar que nem sempre podemos encontrar explicações intuitivas e ‘visuais’ como esta, em virtude da sofisticação dos casos e dos espaços com os quais trabalhamos. Mas é para isso também que serve a matemática: para servir de ‘piloto automático’ para nos guiar em campos onde não há nada ‘concreto’ para nos basearmos.

O que significa isso intuitivamente? No \mathbb{R}^2 , é fácil explicar. Podemos olhar $(2, 3)$ como constituindo a matriz das coordenadas de um vetor β na base canônica, ou seja, $\beta = 2(1, 0) + 3(0, 1)$. Então \mathcal{A} pode ser visto como ‘um outro sistema de coordenadas’ para o \mathbb{R}^2 (com efeito, é uma outra base). A matriz $[\beta]_{\mathcal{A}}$ indica como podemos escrever o vetor β nesta nova base, ou seja, $\beta = \frac{7}{3}(1, 1) + \frac{1}{3}(-1, 2)$.

Importante é o seguinte resultado.

Teorema 6.1 (Unicidade das coordenadas). *A matriz das coordenadas de um vetor em uma base ordenada é única.*

Demonstração. Suponha por absurdo que haja duas matrizes das coordenadas de β na base dada, $[\beta]_{\mathcal{A}} = [x_1, x_2, \dots, x_n]$ e $[\beta]_{\mathcal{A}} = [y_1, y_2, \dots, y_n]$. Então, pela igualdade de matrizes, segue que $x_i = y_i$, para todo i . \square

7. Matriz de mudança de coordenadas

Chama-se **matriz de mudança de coordenadas**, ou **matriz de mudança de base** a uma matriz que permite, dadas as coordenadas de um vetor em uma base, encontrar suas coordenadas em outra base (as bases são supostas sempre ordenadas). Se β é um vetor arbitrário do espaço em questão, então se $\mathcal{A} = \{\alpha_1, \dots, \alpha_n\}$ e $\mathcal{B} = \{\alpha'_1, \dots, \alpha'_n\}$ forem as bases consideradas, e $[\beta]_{\mathcal{A}} = [x_1, x_2, \dots, x_n]$ e $[\beta]_{\mathcal{B}} = [y_1, y_2, \dots, y_n]$ as matrizes das respectivas coordenadas de um vetor β , a matriz será uma matriz $M = [m_{ij}]$, tal que

$$[\beta]_{\mathcal{A}}^T = M[\beta]_{\mathcal{B}}^T. \quad (1.4)$$

Para constatar isso, utilizamos a notação precedente, obtendo

$$\beta = \sum_j y_j \alpha'_j = \sum_j y_j \left(\sum_i m_{ij} \alpha_i \right) = \sum_i \underbrace{\left(\sum_j m_{ij} y_j \right)}_{x_i} \alpha_i. \quad (1.5)$$

Perante a unicidade das coordenadas, temos

$$x_i = \sum_j m_{ij} y_j. \quad (1.6)$$

A expressão (1.4) pode ser escrita na forma matricial como

$$[\beta]_{\mathcal{A}}^T = M[\beta]_{\mathcal{B}}^T.$$

Neste caso, dizemos que ela ‘muda’ as coordenadas de β da base \mathcal{B} para a base \mathcal{A} . Uma tal matriz é inversível,¹⁵ como é possível provar (veja exemplo abaixo), resultando

$$[\beta]_{\mathcal{B}}^T = M^{-1}[\beta]_{\mathcal{A}}^T. \quad (1.7)$$

Algoritmo Para achar a matriz de mudança de coordenadas da base \mathcal{B} para a base \mathcal{A} , basta seguir o procedimento acima, que pode ser condensado nas seguintes regras:

1. Escreva os vetores de \mathcal{B} como combinações lineares dos vetores de \mathcal{A} (expressão 1.5).
2. Encontre os coeficientes dessas combinações lineares (resolvendo sistemas de equações lineares)
3. Forme a matriz M com esses coeficientes transpostos.

Exemplo 7.1. Encontre a matriz de mudança de coordenadas da base ordenada $\mathcal{B} = \{(1, -1), (1, 2)\}$ para a base $\mathcal{A} = \{(1, 1), (0, 1)\}$, ambas para o \mathbb{R}^2 .

¹⁵Uma matriz quadrada A é inversível se e somente se existe uma matriz de mesma ordem A^{-1} tal que $AA^{-1} = A^{-1}A = I$, onde I é a matriz identidade de mesma ordem. Uma condição necessária e suficiente para que A seja inversível é que seu determinante seja diferente de zero.

Solução: Seguiremos o algoritmo. Primeiro,

$$(1, -1) = m_{11}(1, 1) + m_{12}(0, 1)$$

ou

$$(1, -1) = (m_{11}, m_{11} + m_{12}),$$

o que fornece $m_{11} = 1$ e $m_{12} = -2$.

Analogamente,

$$(1, 2) = m_{21}(1, 1) + m_{22}(0, 1),$$

o que fornece $m_{21} = 1$ e $m_{22} = 1$.

Portanto,

$$M = \begin{pmatrix} m_{11} & m_{21} \\ m_{12} & m_{22} \end{pmatrix},$$

ou seja,

$$M = \begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix}.$$

Para observar o ‘efeito’ da matriz de mudança, seja $\beta = (3, 4)$. É fácil ver que suas coordenadas na base \mathcal{B} são $[\beta]_{\mathcal{B}} = [2/3, 7/3]$. Agora aplique M , obtendo

$$M[\beta]_{\mathcal{B}}^T = \begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2/3 \\ 7/3 \end{pmatrix}$$

Logo

$$[\beta]_{\mathcal{A}} = [3, 1].$$

Exercício 7.1 (A matriz de mudança é inversível). Acompanhe os detalhes atentamente. Considere novamente a expressão que caracteriza M , a saber, $[\beta]_{\mathcal{A}}^T = M[\beta]_{\mathcal{B}}^T$. Repare que $[\beta]_{\mathcal{A}}$ é igual à matriz nula (denotada $\mathbf{0}$) se e somente se $[\beta]_{\mathcal{B}}$ for também idêntica à matriz nula, pois sendo \mathcal{A} e \mathcal{B} bases, seus vetores são linearmente independentes e o sistema $M[\beta]_{\mathcal{B}}^T = \mathbf{0}$ tem solução única e portanto trivial (o determinante de M é diferente de zero). Resulta pois que M^{-1} existe e a expressão (1.7) faz sentido.

Teorema 7.1 (Unicidade da matriz de mudança). *A matriz de mudança de coordenadas de uma base para outra é única.*

Demonstração. Suponha, mais uma vez por absurdo, que existam duas matrizes M e M' tais que $[\beta]_{\mathcal{A}}^T = M[\beta]_{\mathcal{B}}^T$ e $[\beta]_{\mathcal{A}}^T = M'[\beta]_{\mathcal{B}}^T$, para todo vetor β . Logo, pela transitividade da relação de igualdade, temos $M[\beta]_{\mathcal{B}}^T = M'[\beta]_{\mathcal{B}}^T$, ou seja, $(M - M')[\beta]_{\mathcal{B}}^T = \mathbf{O}$. Ora, isso vale para toda matriz $[\beta]_{\mathcal{B}}$ se e somente se $M - M' = \mathbf{O}$, ou seja, se e somente se $M = M'$. \square

Exercício 7.2. Justifique a mudança da ordem dos somatórios em (1.5).

Exercício 7.3. Justifique (demonstre) o seguinte fato sobre matrizes, utilizado na demonstração precedente: se $A \cdot B = \mathbf{O}$ (matriz nula) para toda matriz coluna B , então $A = \mathbf{O}$.

8. Existência de base

Para alguns espaços de dimensão infinita, como o espaço das funções reais contínuas no intervalo $[a, b]$, não se pode *exibir* uma base, e nem mesmo indicá-la como fizemos acima no caso do espaço dos polinômios com coeficientes reais e para o próprio espaço dos reais visto como espaço vetorial sobre si mesmo. Mas assume-se que elas existem, e isso na verdade pode ser demonstrado para um espaço vetorial arbitrário, a saber,

Teorema 8.1. *Todo espaço vetorial possui uma base.*

A palavra ‘uma’ deve ser entendida literalmente, como artigo indefinido, e não como indicando ‘uma única’. Um espaço vetorial, com exceção do subespaço trivial constituído unicamente pelo vetor nulo (cuja base, como vimos, é o conjunto vazio), admite uma infinidade de bases, todas elas de mesma cardinalidade. A demonstração do teorema anterior faz uso do Axioma da Escolha (mais precisamente, do **Lema de Zorn**, que lhe é equivalente, que pode ser assim enunciado: *Seja $\langle A, \leq \rangle$ um conjunto não vazio parcialmente ordenado¹⁶ tal que toda cadeia em A admita um limitante superior. Então A tem elemento*

¹⁶Uma **ordem parcial** sobre um conjunto A é uma relação binária sobre A (um subconjunto de $A \times A$) que é (i) reflexiva, (ii) anti-simétrica e (iii) transitiva. A expressão ‘parcial’ vem do fato de que podem haver elementos de A que não estejam na relação.

maximal. Uma **cadeia** em A é um subconjunto C de A que é **linearmente ordenado** por \leq ;¹⁷ um **limitante superior** (cota superior, limite superior) de C é um elemento m tal que $c \leq m$ para todo $c \in C$; este elemento pode não pertencer a C , e pode não ser único. Diz-se que m é **elemento maximal** de A se não existe $x \in A$ tal que $m < x$ (onde $a < b := a \leq b \wedge a \neq b$). O elemento maximal de um conjunto, quando existe, pertence ao conjunto, e pode não ser único. A demonstração procede mostrando que se \mathcal{I} é uma família de conjuntos linearmente independentes de vetores, então se os elementos de \mathcal{I} são tais que se $B \in \mathcal{I}$ implica que todo subconjunto finito de B pertence a \mathcal{I} (neste caso, diz-se que \mathcal{I} tem característica finita), resulta que \mathcal{I} tem elemento maximal com respeito à inclusão \subseteq (isso é conhecido como Lema de Tukey, ou Princípio Maximal). Ora, uma família de conjuntos de vetores linearmente independente tem característica finita, e assim existe um conjunto maximal B , e pode-se então mostrar que B é uma base.

Um exemplo é facilmente visto considerando-se novamente o espaço vetorial dos polinômios com coeficientes reais visto antes. Tome qualquer polinômio não identicamente nulo p_0 , por exemplo $p_0(x) = 1$. O conjunto $S_0 = \{p_0\}$ é linearmente independente (prove isso como exercício). Se este conjunto gerasse o espaço, seria uma base, mas este não é o caso. Defina agora $S_1 = \{p_0, p_1\}$ com $p_1(x) = 1 + x$ (perceba que este polinômio não pertence ao espaço gerado por S_0). Ora, S_1 é linearmente independente (exercício), mas também não gera o espaço. Continuemos assim, obtendo $S_2 = \{p_0, p_1, p_2\}$ (onde $p_2(x) = 1 + x + x^2$), $S_3 = \{p_0, p_1, p_2, p_3\}$, e assim por diante. Evidentemente temos

$$S_0 \subseteq S_1 \subseteq S_2 \subseteq \dots,$$

e este conjunto é parcialmente (na verdade, é linearmente) ordenado por inclusão. Constata-se que todo subconjunto finito de elementos (que são cadeias) tem um limitante superior. Assim pelo Lema de Zorn, o conjunto tem um elemento maximal, que é a base procurada, um conjunto *maximal* linearmente independente que gera o espaço.

Para as finalidades da física, estaremos interessados em *bases ortonormais*, mas para tanto necessitamos introduzir noções métricas nos espaços vetoriais, o que fazemos por meio de um *produto interno*, como veremos no próximo capítulo. Resta no entanto uma observação. O esboço da demonstração do

¹⁷Uma ordem linear, ou total, é uma ordem parcial tal que quaisquer dois elementos do conjunto estão relacionados.

teorema (8.1) feita acima mostra a importância do Axioma da Escolha para as finalidades da física. Precisamos falar de bases, e precisamos que todas as bases tenham o mesmo número de elementos. Acontece que há ‘matemáticas’ nas quais o Axioma da Escolha não vale em geral,¹⁸ e em algumas delas, pode-se encontrar espaços vetoriais que não têm base, ou então espaços vetoriais que têm bases de cardinalidades distintas (veja [Jec.77, p.366]). Como se elaboraria a mecânica quântica tendo por alicerce uma tal matemática? Seria isso possível? Claro que até o momento ninguém sabe a resposta.

9. Espaços vetoriais isomorfos

Escreveremos $\mathcal{E} \text{ iso } \mathcal{F}$ para indicar que há um **isomorfismo** entre os espaços vetoriais $\mathcal{E} = \langle \mathcal{V}, \mathcal{K}, +, \cdot \rangle$ e $\mathcal{F} = \langle \mathcal{W}, \mathcal{K}, +, \cdot \rangle$, ambos sobre o mesmo corpo $\mathcal{K} = \langle K, +, \cdot, 0, 1 \rangle$, ou seja, existe uma aplicação (função) bijetiva $f: \mathcal{V} \mapsto \mathcal{W}$ tal que

$$f(\alpha + \beta) = f(\alpha) + f(\beta) \text{ e } f(k \cdot \alpha) = k \cdot f(\alpha) \quad (1.8)$$

para todos $\alpha, \beta \in \mathcal{V}$ e $k \in K$. Pode-se demonstrar (exercício) que a relação **iso** é uma relação de equivalência,¹⁹ o que implica em particular ser transitiva. Assim, se \mathcal{E} é isomorfo a \mathcal{F} e se \mathcal{F} é isomorfo a \mathcal{G} , então \mathcal{E} é isomorfo a \mathcal{G} . Disso tudo resulta trivial provar o seguinte resultado (a argumentação precisa, no entanto, dá um bom exercício), extremamente útil nas demonstrações:

$$K^n \text{ iso } K^{1 \times n} \text{ iso } K^{n \times 1}.$$

Para $n = 2$ e $K = \mathbb{R}$, podemos então representar a correspondência escrevendo

$$(a, b) \Leftrightarrow [ab] \Leftrightarrow \begin{pmatrix} a \\ b \end{pmatrix}.$$

Note que os conjuntos acima, munidos das respectivas operações de espaço vetorial, constituem espaços vetoriais sobre K . Assim, face o isomorfismo, é

¹⁸Há versões mais fracas do axioma, como aquela que fala unicamente de conjuntos enumeráveis. Ver [Jec.77].

¹⁹Uma **relação de equivalência** sobre um conjunto A é uma relação binária R sobre A que é reflexiva (ou seja, para todo elemento x de A tem-se que xRx), simétrica (se xRy então yRx para todos x e y) e transitiva (se xRy e yRz , então xRz para todos x, y, z).

indiferente (matematicamente falando) se operamos com n -uplas de elementos de K ou com matrizes linha formadas por elementos de K ou com matrizes coluna com tais elementos. Isso traz uma enorme vantagem, pois simplifica em muito as demonstrações, já que podemos ir de um espaço a outro sem maiores detalhes, ora operando com n -uplas, ora transformando-as em matrizes linha, como se fossem a mesma coisa.

Vamos dar um exemplo dessa utilidade, indicando os detalhes para que o leitor perceba como age o matemático, muitas vezes caminhando de uma estrutura para outra e voltando, desde que essas estruturas preservem as propriedades relevantes (o que é feito pela existência dos isomorfismos). Considere o espaço vetorial \mathbb{R}^3 . Vetores deste espaço são triplas ordenadas de números reais da forma (x, y, z) . Considere agora o seguinte problema: verificar se os vetores $(1, 2, 3)$, $(-1, 0, 1)$ e $(0, 1, 2)$ são linearmente independentes. Podemos fazer isso simplesmente considerando esses vetores como matrizes linhas: $[1\ 2\ 3]$, $[-1\ 0\ 1]$ e $[0\ 1\ 2]$. Olhe agora para a matriz abaixo como sendo uma matriz 3×1 cujas linhas são as três matrizes acima.

$$\begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix}$$

Olhe agora para esta nova matriz 3×3 e extraia fatos sobre ela. Se o determinante desta matriz for não nulo, nenhuma das linhas pode ser combinação linear das demais, segundo a (suposta) conhecida teoria dos determinantes. Ora, esta é exatamente a condição que esperamos para a independência linear dos vetores. Ou seja, "mudamos" de espaço vetorial para conseguir certos resultados: neste caso, se o determinante da matriz acima é não nulo, ela é inversível e portanto suas linhas ("logo", os vetores de \mathbb{R}^3) são linearmente independentes.

Observação O procedimento acima foi na verdade um truque, útil, mas um truque. Mudamos de espaços de n -uplas para matrizes, para outras matrizes e depois voltamos, e tudo funciona. No entanto, carece de um rigor matemático mais refinado, principalmente aquela parte em que consideramos o determinante da matriz 3×3 . Claro que podemos, neste caso, prover o que está faltando, mas não importa aqui. Isso é comum entre físicos principalmente. Houve uma espécie de debate entre o matemático von Neumann, rigoroso

para com a mecânica quântica, e o físico Dirac, mais ‘pragmático’ sacrificando algumas vezes o rigor em prol de resultados físicos mais expressivos. O caso em tela é a célebre **função delta**, que Dirac introduziu e que aparentemente levava a inconsistências (mas que depois encontrou uma fundamentação matemática precisa nas mãos de L. Schwartz). Para detalhes sobre esta polêmica, ver [Kro.12].

Prosseguiremos dentro do rigor matemático esperado.

Teorema 9.1. *Todo espaço vetorial de dimensão n sobre o corpo $\mathcal{K} = \langle K, +, \cdot, 0, 1 \rangle$ é isomorfo ao K^n .*

Demonstração. Como K^n , o conjunto das n -uplas de elementos de K munido das operações usuais, como as definidas para o \mathbb{R}^n (veja o exemplo (1.1), é isomorfo a $K^{1 \times n}$, o espaço das matrizes $1 \times n$ com elementos em K (munido das operações correspondentes entre matrizes), basta provar que o isomorfismo se dá com esse último espaço. Seja $\mathcal{A} = \{\alpha_1, \dots, \alpha_n\}$ uma base ordenada para o espaço vetorial $\mathcal{E} = \langle \mathcal{V}, \mathcal{K}, +, \cdot \rangle$ de dimensão n . Então, para $\beta \in \mathcal{V}$, podemos escrever

$$\beta = x_1 \alpha_1 + \dots + x_n \alpha_n.$$

Definimos a aplicação $f : \mathcal{V} \mapsto K^{1 \times n}$ por $f(\beta) = [\beta]_{\mathcal{A}} = [x_1 \dots x_n]$. Devemos agora provar que f é um isomorfismo, ou seja, que é bijetiva e que ‘preserva’ as operações. Quanto à primeira parte, inicialmente mostramos que f é injetiva. Com efeito, pela unicidade das coordenadas (teorema (6.1)), se $\gamma \neq \beta$, sua matriz das coordenadas em \mathcal{A} é distinta da de β .²⁰ Quanto a f ser sobrejetiva, dada uma matriz $[x_1 \ x_2 \ \dots \ x_n]$ de escalares de K , haverá um único vetor β (de novo, pela unicidade das coordenadas) tal que $\beta = x_1 \alpha_1 + \dots + x_n \alpha_n$.²¹ Assim, f é bijetiva. Agora, com relação a ela ‘manter’ as relações das estruturas, a saber, as operações de espaço vetorial. Sejam α e β vetores de \mathcal{V} . Assim, existem únicas $[\alpha]_{\mathcal{A}} = [x_1 \ x_2 \ \dots \ x_n]$ e $[\beta]_{\mathcal{A}} = [y_1 \ y_2 \ \dots \ y_n]$. Ora, é claro que $f(\alpha + \beta) = [\alpha + \beta]_{\mathcal{A}} = [x_1 + y_1 \ \dots \ x_n + y_n] = [x_1 \ \dots \ x_n] + [y_1 \ \dots \ y_n] = [\alpha]_{\mathcal{A}} + [\beta]_{\mathcal{A}} = f(\alpha) + f(\beta)$, e que $f(k\alpha) = [kx_1 \ \dots \ kx_n] = k[x_1 \ \dots \ x_n] = kf(\alpha)$. Assim, f é um isomorfismo. \square

Exercício 9.1. Exiba um isomorfismo entre o espaço \mathbb{R}^4 e o espaço $\mathfrak{P}(\mathbb{R}^3)$ dos polinômios reais de grau menor ou igual a 3. (Dica: você terá não só que exibir a função, mas mostrar que ela é de fato um isomorfismo.)

²⁰Recorde que uma função $f : A \mapsto B$ é injetiva se $x \neq y$ implica $f(x) \neq f(y)$.

²¹Uma função $f : A \mapsto B$ é sobrejetiva se para todo $y \in B$, existe $x \in A$ tal que $y = f(x)$.

Exercício 9.2. Mostre que os espaços vetoriais \mathbb{C} dos números complexos da forma $z = a + bi$, munido das operações usuais e o \mathbb{R}^2 , também munido das operações usuais, são isomorfos. (Isto é o que nos possibilita representar números complexos como pares ordenados de números reais).

Exercício 9.3. Encontre uma base para o sub-espaço do \mathbb{R}^3 gerado pelos vetores $(1, 0, 1)$, $(-1, 1, 0)$, $(1, 2, 1)$, e $(0, 0, -1)$. Verifique se o vetor $(2, -1, -2)$ pertence a tal sub-espaço.

Exercício 9.4. Mostre que o espaço $\mathbb{R}^{2 \times 2}$ das matrizes reais 2×2 é isomorfo ao espaço \mathbb{R}^4 .

10. Mais sobre dimensão infinita

Espaços vetoriais de dimensão infinita são extremamente relevantes em matemática e em física. Por exemplo, em mecânica quântica os espaços associados aos observáveis *posição* e *momento* de uma partícula, são infinitos. Acima, mencionamos somente um caso, o das funções reais de variável real com domínio no intervalo $[a, b]$ da reta real. Vimos que, por força do teorema (8.1), ele possui uma base. O problema é que não temos como exibir essa base. Uma justificativa informal pode ser a seguinte: ela teria que ser formada por uma infinidade de "vetores" (funções deste espaço) tais que *qualquer* dessas funções pudesse ser obtida como combinação linear das funções da base, com coeficientes reais. Intuitivamente, isso é claramente impossível de ser realizado tendo em vista a variedade de funções com domínio $[a, b]$. Resta acreditar no teorema e em suas consequências. O fato interessante é que a impossibilidade de exibir uma base não se deve a qualquer incapacidade nossa, mas é inerente à matemática envolvida.

Um outro exemplo é o seguinte. Uma vez que saibamos que o conjunto \mathbb{Q} dos números racionais munido das operações de adição de racionais e de multiplicação de racionais forma um corpo $\mathcal{Q} = \langle \mathbb{Q}, +, \cdot, 0, 1 \rangle$ (se preciso, vá atrás dos detalhes em um livro de Álgebra), podemos considerar a estrutura $\mathcal{R}^{\mathcal{Q}} = \langle \mathbb{R}, \mathcal{Q}, +, \cdot \rangle$ (as operações nesta estrutura sendo a adição de reais e a multiplicação de real por racional) e mostrar (Exercício) que se trata de um espaço vetorial *racional* (sobre o corpo dos racionais). Aqui, os vetores são os números reais, e os escalares, os números racionais. Uma base para tal espaço, que existe por força do teorema mencionado, terá que ser um conjunto de números reais de forma que qualquer real possa ser obtido como uma combinação linear de

tais números, desde que os coeficientes sejam racionais. Como no caso precedente, tal base não pode ser exibida, existindo única e exclusivamente por força do teorema. Chama-se *base de Hamel*, e pode-se mostrar que a dimensão desse espaço é a cardinalidade do contínuo.²²

10.1. Adendo: o axioma da escolha

Aqui, mais uma digressão. Muitas vezes, praticando uma ciência como a física, não nos damos conta de princípios que são assumidos como se fossem ‘naturais’ na base matemática e lógica das teorias consideradas, mas que uma reflexão mostra que devem ser tomados com cautela e, para o filósofo e para aquele interessado nos fundamentos da ciência que pratica, conhecidos ao menos em suas linhas gerais. Um típico exemplo é o do Axioma da Escolha, que fundamenta a matemática que subjaz toda a física presente e em particular a mecânica quântica. Trata-se de um postulado da teoria de conjuntos que pode-se dizer, sem perda de generalidade, serve de lógica subjacente a esta teoria. Em síntese (há dezenas de formas equivalentes de enunciá-lo e há formas mais fortes e mais fracas dele), assegura que dado um conjunto com determinadas condições, como tendo conjuntos não vazios e dois a dois disjuntos como elementos, existe um conjunto (o *conjunto escolha*) contendo um e só um elemento de cada um dos conjuntos que formam o conjunto dado. Se o número desses conjuntos for finito, este resultado pode ser demonstrado a partir dos demais postulados da teoria (que aqui assumiremos ser a teoria Zermelo-Fraenkel – ver [Kra.02]). O

²²Uma das grandes coisas que Georg Cantor, o criador da teoria dos conjuntos fez, foi mostrar que há infinitos de diferentes ‘tamanhos’, o que aparentemente já era conhecido de Leibniz (ver [SapAlc.15]). O ‘menor’ deles é o dos conjuntos *enumeráveis*, que podem ser colocados em correspondência 1×1 (por meio de uma *função bijetora*) com o conjunto dos números naturais, e é designado por \aleph_0 . O conjunto dos números reais, na matemática padrão, tem um cardinal estritamente maior (considerada a ordem dos cardinais), sendo igual a 2^{\aleph_0} , dita *cardinalidade do contínuo*, mas há ainda cardinais maiores, e maiores ... Uma observação complementar que pode ser útil para certos leitores: foi dito ‘na matemática padrão’ porque podemos supor que a teoria utilizada seja uma teoria de primeira ordem (tendo a lógica clássica de predicados de primeira ordem como lógica subjacente). Neste caso, ela está sujeita ao chamado Teorema de Löwenheim-Skolem Descendente que afirma que se a teoria tiver modelo (for consistente), terá modelo enumerável. Nesse modelo, o conjunto que representa os números reais será enumerável. Este fato, conhecido como ‘paradoxo’ de Skolem, nada tem de paradoxical, tratando-se apenas de um resultado contra intuitivo.

problema é que em muitos casos nada na teoria e muito menos no axioma nos dá a informação de *como* obter o conjunto escolhido. Ele simplesmente *existe* desde que assumamos o postulado, e somente em alguns casos podemos apresentar uma ‘função escolhida’ que faria a seleção.

Como vimos, este axioma é fundamental em mecânica quântica (pelo menos em uma de suas formas equivalentes, o Lema de Zorn), pois garante a existência de bases para os espaços de Hilbert relevantes. Vê-se mais uma vez a importância da lógica e da matemática subjacentes (que para muitos, como para nós aqui, são tomadas como uma só coisa); utilizamos a redução ao absurdo, um dos princípios básicos da lógica clássica, o axioma da escolha e vários outros postulados dos quais muitas vezes não nos damos conta. Uma questão interessante para os fundamentos da mecânica quântica é explicitar essa ‘base lógica’. Se formos olhar os livros desta disciplina e fizermos uma análise detalhada, encontraremos nada mais do que aquilo que usualmente chamamos de lógica clássica. No entanto, há questionamentos ‘quânticos’ sobre alguns dos princípios desta lógica, como a célebre lei distributiva, como veremos abaixo. Como ficamos? Ora, não ficamos. A área que se denomina hoje de *lógica quântica* está mais voltada para a *computação quântica* e para a teoria da *informação quântica* (ver [DalGiuGre.04]) do que para a procura de uma ‘lógica da mecânica quântica’. A questão permanece em aberto.

Talvez tenha passado o tempo de se buscar em ciência, e particularmente em física, o desenvolvimento das disciplinas como *ciências de princípios*, no sentido aristotélico do termo, como exposto em seus *Segundos Analíticos* [Ari.53]. Hoje, principalmente devido ao fato de que não se conhecem quais seriam os princípios (postulados) – os físicos falam em ‘equações’ – que fundamentariam as principais disciplinas a partir das teorias quânticas de campos,²³ e das teorias de cordas e outras, fica patente que a física presente (pelo menos) funciona como um conjunto de regras mais ou menos heurísticas que são juntadas para se atacar um problema particular, caso a caso. Não há até o momento, e talvez nunca venha haver, uma teoria unificadora. Os vários ramos muitas vezes parecem ser até mesmo inconsistentes uns com os outros. Mas desenvolver isso demandaria um outro texto. Fica a indicação para aguçar a curiosidade do leitor.

Falaremos um pouco mais sobre a lógica quântica mais à frente.

²³De fato há formulações ‘axiomáticas’ de algumas teorias quânticas de campos, mas esta é outra história.

PRODUTOS INTERNOS

COM A ESTRUTURA de espaço vetorial, o máximo que podemos expressar são combinações lineares (superposições) de vetores. Não há como considerar questões métricas, como ângulo entre vetores, ‘comprimento’ de um vetor e outro de mesma natureza. Para tanto, vamos estender a estrutura \mathcal{E} adicionando um *produto interno*. O espaço assim obtido é denominado de *espaço vetorial com produto interno*, ou *pré-espaço de Hilbert*. Mas, o que é um produto interno?

1. Produtos internos

Definição 1.1 (Produto interno). Um **produto interno** sobre um espaço vetorial $\mathcal{E} = \langle \mathcal{V}, \mathcal{K}, +, \cdot \rangle$ é uma aplicação de $\mathcal{V} \times \mathcal{V}$ em K (o domínio do corpo \mathcal{K}). A imagem do par de vetores α e β (nesta ordem), é denotada por $\langle \alpha | \beta \rangle$. A definição exige que sejam cumpridas as seguintes condições, para todos $\alpha, \beta, \gamma \in \mathcal{V}$ e $a \in K$, se tenha:¹

1. $\langle \alpha | \beta + \gamma \rangle = \langle \alpha | \beta \rangle + \langle \alpha | \gamma \rangle$
2. $\langle \alpha | a\beta \rangle = a\langle \alpha | \beta \rangle$

Estas duas condições dizem que o produto interno é linear na segunda variável (a que vem depois do traço vertical em $\langle \cdot | \cdot \rangle$).

¹Há autores que preferem postular as condições 1 e 2 de forma alternativa, a saber, (1') $\langle \alpha + \beta | \gamma \rangle = \langle \alpha | \gamma \rangle + \langle \beta | \gamma \rangle$ e (2') $\langle a\alpha | \beta \rangle = a\langle \alpha | \beta \rangle$. Isso é meramente uma questão de conveniência. Aqui, seguimos o procedimento usual dos textos de física.

3. $\langle \alpha | \beta \rangle = \overline{\langle \beta | \alpha \rangle}$ (Em física, é comum denotar o conjugado de um número complexo $z = a + bi$ não por $\bar{z} = a - bi$, mas por z^* . Mais à frente, usaremos essa notação).
4. $\langle \alpha | \alpha \rangle \geq 0$ e $\langle \alpha | \alpha \rangle = 0$ se e somente se $\alpha = 0$.

Teorema 1.1. *Dada a definição, temos:*

1. $\langle \alpha | \beta \rangle = \bar{a} \langle \alpha | \beta \rangle$
2. $\langle \alpha + \beta | \gamma \rangle = \langle \alpha | \gamma \rangle + \langle \beta | \gamma \rangle$

Demonstração. Com efeito, ² $\langle \alpha \alpha | \beta \rangle = \overline{\langle \beta | \alpha \alpha \rangle} = \overline{a \langle \beta | \alpha \rangle} = \bar{a} \langle \alpha | \beta \rangle$. Quanto ao segundo item, sugerimos que o leitor o faça como um exercício. \square

As condições do teorema dizem que o produto interno é **sesquilinear** na primeira variável, e **linear** na segunda.

Daremos agora alguns exemplos de produtos internos que interessarão ao nosso estudo.

Exemplo 1.1. Sobre o espaço real \mathbb{R}^n , sendo $\alpha = (x_1, \dots, x_n)$ e $\beta = (y_1, \dots, y_n)$, a aplicação seguinte é um produto interno:

$$\langle \alpha | \beta \rangle := \sum_{i=1}^n x_i y_i \quad (2.1)$$

Este produto interno é denominado de **produto interno canônico** sobre o \mathbb{R}^n .

Exemplo 1.2. Sobre o espaço real \mathbb{C}^n , sendo $\alpha = (x_1, \dots, x_n)$ e $\beta = (y_1, \dots, y_n)$ n -uplas de números complexos, a aplicação seguinte é um produto interno:

$$\langle \alpha | \beta \rangle := \sum_{i=1}^n x_i^* y_i \quad (2.2)$$

onde x_i^* é o conjugado de x_i , ou seja, se $x_i = a + bi$, então $x_i^* = a - bi$. Este produto interno é denominado de **produto interno canônico** sobre o \mathbb{C}^n .

Exemplo 1.3. Sobre o espaço das funções reais (de variável real) contínuas no intervalo $[a, b]$ (e isto vale para $a = -\infty$, $b = +\infty$), a aplicação seguinte é um produto interno:

$$\langle f | g \rangle := \int_a^b f(x)g(x)dx \quad (2.3)$$

²Observamos que, para números complexos z_1 e z_2 , tem-se que $\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$.

Uma explicação a respeito deste último exemplo. Em se tratando de funções de variável complexa, definimos a conjugada de f , denotada por f^* , da seguinte maneira, para cada número complexo z de seu domínio:

$$f^*(z) := (f(z^*))^*.$$

Exercício 1.1. Considere a seguinte função complexa sobre o espaço do exemplo precedente: $f(z) = a^2 + i(3a - b)$, para cada $z = a + ib$. Encontre $f(1 - 2i)$ e ache f^* .

Exemplo 1.4. Seja $\mathbb{C}^{n \times n}$ o espaço vetorial das matrizes complexas de ordem n . Se A^* denota a transposta conjugada de A , então a aplicação

$$\langle A|B \rangle := \text{Tr}(A^*B) \tag{2.4}$$

é um produto interno. No caso real (ou seja, se \mathbb{R}^n), então $\langle A|B \rangle := \text{Tr}(A^T B)$.

Neste exemplo, estamos fazendo uso da **função traço**, que aparecerá novamente à frente. Trata-se de uma função que associa a cada matriz quadrada um escalar que é a soma dos elementos de sua diagonal principal. Por outro lado, A^T designa a **transposta** de A , obtida simplesmente trocando-se as linhas da matriz pelas suas respectivas colunas.

Produto de Matrizes Uma pequena digressão aqui, principalmente para estudantes de filosofia, sobre o produto de matrizes. Sejam $A = [a_{ij}]$ de ordem $m \times n$ e $B = [b_{jk}]$ de ordem $n \times k$. Nessas condições, o **produto das matrizes** A e B (nesta ordem) é possível e é definido como segue. Antes, repare nas ordens das matrizes: *o número de colunas da primeira deve ser igual ao número de linhas da segunda*. A matriz resultante será denotada $C = [c_{ik}]$, e terá ordem $m \times k$. Seus elementos são obtidos da seguinte maneira:

$$c_{ik} = \sum_{j=1}^n a_{ij} b_{jk}.$$

Vejamos um exemplo:

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & -1 \\ -2 & 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 2 & -1 & 0 \\ -1 & 1 & 3 \\ 0 & 1 & -2 \end{pmatrix} = \begin{pmatrix} 3 & -2 & -3 \\ -2 & 1 & 4 \\ -5 & 2 & 1 \end{pmatrix}.$$

Informalmente, isso corresponde a multiplicarmos as linhas de A e as colunas de B como se fossem vetores, por meio do produto interno canônico. Mas isso é somente uma comparação grosseira, mas não é ‘absolutamente’ certo, porque se os elementos das matrizes forem números complexos, de acordo com a definição do produto, devemos multiplicar as linhas de A pelas colunas de B e não os conjugados complexos das linhas de A pelas colunas de B , como exigiria o produto canônico que utiliza números complexos. No exemplo 1.4, para obter $\langle A|B \rangle$, primeiramente obtemos a transposta conjugada de A e fazemos o produto A^*B para então tomar o traço da matriz resultante. Exercite com as matrizes seguintes.

Exercício 1.2. Obtenha o produto interno $\langle A|B \rangle$ e depois o produto AB , sendo

$$A = \begin{pmatrix} 2+2i & -1+i \\ -1 & 1+i \end{pmatrix}, B = \begin{pmatrix} 2+i & 1+i \\ -i & 1-i \end{pmatrix}.$$

Exercício 1.3. Mostre que o produto de matrizes é associativo, e encontre um exemplo para mostrar que o produto não é comutativo (basta exibir duas matrizes que não comutam, mas cujas ordens permitiriam o produto, pois ele precisa estar definido). Mostre ainda que se A é uma matriz qualquer de ordem $m \times n$ e se I_n é a matriz identidade de ordem n (isto é, $n \times n$), então $AI_n = A$; se I_m é a matriz identidade de ordem m , então $I_mA = A$. Ademais, se A for inversível (uma condição necessária e suficiente para que seja inversível é que seja quadrada e que seu determinante seja diferente de zero), sendo A^{-1} a sua inversa (que se pode provar ser única), tem-se que $AA^{-1} = A^{-1}A = I_n$. Deste modo, o conjunto das matrizes inversíveis de ordem n sobre um corpo \mathcal{K} , munido da operação de produto de matrizes, é um grupo não comutativo.

Definição 1.2 (Norma). Chama-se **norma** (ou ‘comprimento’) em um espaço vetorial \mathcal{V} a uma aplicação que, a cada vetor α associa um escalar, denotado $\|\alpha\|$, tal que:

1. $\|\alpha\| \geq 0$ e $\|\alpha\| = 0$ se e somente se $\alpha = \mathbf{0}$.
2. $\|k\alpha\| = |k| \cdot \|\alpha\|$, para k escalar.
3. $\|\alpha + \beta\| \leq \|\alpha\| + \|\beta\|$ (desigualdade triangular)

Um vetor α é **unitário** se $\|\alpha\| = 1$.

Importa aqui, dentre todas as possíveis normas, aquela que é definida a partir do produto interno, dita **norma advinda do produto interno**, a saber, a aplicação tal que

$$\|\alpha\| := \sqrt{\langle \alpha | \alpha \rangle} \quad (2.5)$$

Exemplo 1.5. Mostre que a aplicação recém definida é de fato uma norma.

A importância da observação acima, de que a norma definida por (2.5) é *advinda* do produto interno é que existem normas (funções que cumprem as condições da definição) sem que tenham sido originadas a partir do produto interno. Alguns exemplos sobre \mathbb{R}^2 são os seguintes (para distingui-las da norma acima, vamos usar sub-índices): para $\alpha = (x_1, x_2)$, temos

$$\|\alpha\|_1 := |x_1| + |x_2|$$

$$\|\alpha\|_2 := \max\{|x_1|, |x_2|\}.$$

Assim, $\|(2, 3)\|_1 = 5$, enquanto que $\|(2, 3)\|_2 = 3$. O que nos interessará, no entanto, será a norma induzida pelo produto interno, pois será com ela que a noção de espaço de Hilbert é introduzida. Os teoremas que se reportam a normas, no entanto, valem para qualquer que seja ela, por exemplo, o seguinte.

Teorema 1.2 (Desigualdade de Cauchy-Schwarz). *Para todos α e β , tem-se:*

$$|\langle \alpha | \beta \rangle| \leq \|\alpha\| \cdot \|\beta\| \quad (2.6)$$

Demonstração. Se $\alpha = 0$, é imediato. Caso contrário, seja

$$\gamma = \beta - \frac{\langle \beta | \alpha \rangle}{\|\alpha\|^2} \cdot \alpha.$$

Daí resulta

$$0 \leq \|\gamma\|^2 = \left\langle \beta - \frac{\langle \beta | \alpha \rangle}{\|\alpha\|^2} \alpha \left| \beta - \frac{\langle \beta | \alpha \rangle}{\|\alpha\|^2} \alpha \right. \right\rangle$$

Consequentemente,

$$0 \leq \|\beta\|^2 - \frac{|\langle \beta | \alpha \rangle|^2}{\|\alpha\|^2},$$

de onde se segue o resultado. □

Um fato relevante para a mecânica quântica é o seguinte resultado relativo à norma da soma de dois vetores:

$$\|\alpha + \beta\|^2 = \langle \alpha + \beta | \alpha + \beta \rangle = \|\alpha\|^2 + \|\beta\|^2 + 2\operatorname{Re}(\langle \alpha | \beta \rangle),$$

onde $\operatorname{Re}(\langle \alpha | \beta \rangle)$ indica a parte real do número complexo $\langle \alpha | \beta \rangle$. A relação com a mecânica quântica pode ser antecipada assim, somente para você se dar conta da importância dos conceitos sendo introduzidos. Os *estados* dos sistemas físicos são descritos por vetores unitários. Suponha que tenhamos um estado descrito por uma superposição (aqui é melhor falar assim do que em combinação linear) da forma $\alpha + \beta$, que é por hipótese unitário. O quadrado da norma deste vetor (ou função de onda) é importante pois nos dará uma probabilidade de encontrarmos o sistema (como uma partícula) em uma certa situação. A parte $\operatorname{Re}(\langle \alpha | \beta \rangle)$, que nem sempre pode ser eliminada, e que é chamada de **termo de interferência**, implica que há *superposição* das funções de onda α e β , e de acordo com a interpretação padrão, é o que explica a existência de franjas de interferência em experimentos como o das duas fendas.³

Definição 1.3 (Distância). Chama-se **distância** em um espaço vetorial \mathcal{V} com produto interno a toda aplicação $d : \mathcal{V} \times \mathcal{V} \mapsto K$ tal que:

1. $d(\alpha, \beta) = d(\beta, \alpha)$
2. $d(\alpha, \beta) \geq 0$ e $d(\alpha, \beta) = 0$ se e somente se $\alpha = \beta$.
3. $d(\alpha, \beta) \leq d(\alpha, \gamma) + d(\gamma, \beta)$ (desigualdade triangular)

Teorema 1.3. *Considere a norma advinda do produto interno (a única que consideraremos doravante). A aplicação $d(\alpha, \beta) = \|\alpha - \beta\|$, é uma distância.*

Demonstração. Imediata. □

No caso do espaço vetorial ser \mathbb{R} ou \mathbb{C} , costuma-se escrever $|\alpha|$ em vez de $\|\alpha\|$, deste modo confundindo-se a norma de um vetor com o módulo de um escalar. Do mesmo modo, escrevemos nesses casos $|\alpha - \beta|$ em vez de $\|\alpha - \beta\|$.

³Para tudo o que se refere à história da mecânica quântica não relativista, indicamos o livro de M. Kumar [Kum.09]. Para coisas mais técnicas, consultar os demais textos indicados na Bibliografia.

Importante notar que, sendo $z = a + bi$ um número complexo (ou ‘vetor’ no espaço vetorial dos complexos sobre si mesmo), resulta que

$$|z|^2 = \langle z|z \rangle = z^* z = (a - bi)(a + bi) = a^2 + b^2 \in \mathbb{R}.$$

Da mesma forma, levando em conta a **fórmula de Euler** $e^{i\theta} = \cos \theta + i \sin \theta$, um número complexo $z = a + bi$ pode ser escrito em **coordenadas polares** como $z = r.e^{i\theta}$, e temos que $z^* = r.e^{-i\theta}$, logo $z^* z = r^2$.

2. Espaços de Hilbert

Dizemos que uma seqüência de vetores $\alpha_1, \alpha_2, \dots$ de um espaço vetorial \mathcal{V} *converge* para um vetor β se os vetores da seqüência vão ficando cada vez mais próximos de β à medida em que avançamos na seqüência. Mais precisamente,

Definição 2.1 (Seqüência convergente). A seqüência de vetores $\alpha_1, \alpha_2, \dots$ de um espaço vetorial \mathcal{V} **converge** para um vetor β se, para todo $\epsilon > 0$ real, existe um número natural n tal que, se $i > n$, resulta que $\|\beta - \alpha_i\| < \epsilon$.

Definição 2.2 (Seqüência de Cauchy). Uma seqüência de vetores $\alpha_1, \alpha_2, \dots$ de um espaço vetorial \mathcal{V} é uma **seqüência de Cauchy** se, para todo $\epsilon > 0$ real, existe um número natural n tal que, para $i, j > n$, tem-se que $\|\alpha_i - \alpha_j\| < \epsilon$.

Intuitivamente, em uma seqüência de Cauchy, os elementos da seqüência vão ficando cada vez mais próximos uns dos outros à medida em que avançamos na seqüência.

Toda seqüência de Cauchy é convergente, como se pode mostrar. O problema é que uma seqüência pode convergir para um vetor que não pertença ao espaço considerado. Quando toda seqüência de Cauchy converge para um vetor ainda no espaço, dizemos que o espaço é (topologicamente) **completo**. Note que a noção de convergência depende da norma. Caso particularmente importante é quando a norma é a advinda do produto interno, resultando na seguinte

Definição 2.3 (Espaço de Hilbert). Um espaço vetorial com produto interno \mathcal{V} é um **espaço de Hilbert** se for completo em relação à norma induzida pelo produto interno.

Ou seja, a norma a ser considerada é aquela que se define por meio do produto interno, isto é, $\|\alpha\| := \sqrt{\langle \alpha | \alpha \rangle}$.

Definição 2.4. Um espaço de Hilbert é **separável** se contém uma base contável (finita ou enumerável).

Resulta que todo espaço de dimensão finita é obviamente separável.

Definição 2.5. Um conjunto A de vetores de um espaço \mathcal{E} é **fechado** relativamente a um sub-espaço \mathcal{W} se para todo $\epsilon > 0$ real e para qualquer $\beta \in \mathcal{W}$, existe $\alpha \in A$ tal que $d(\alpha, \beta) < \epsilon$. Pode-se em especial falar de um sub-espaço \mathcal{W} ser fechado (nele mesmo).

Uma outra forma de se conceituar sub-espaços fechados é dizer simplesmente que são sub-espaços do espaço de Hilbert que por sua vez também são espaços de Hilbert.

Exemplo 2.1. Considere o conjunto $\ell^2(\infty)$ de todas as matrizes coluna complexas (de ordem $n \times 1$ e elementos complexos) com uma quantidade enumerável de elementos, ou seja, da forma

$$A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix}$$

com $c_j \in \mathbb{C}$, satisfazendo

$$\sum_{k=1}^{\infty} |a_k|^2 < \infty.$$

Podemos obter um espaço de Hilbert separável com as operações usuais de soma de matrizes e de multiplicação de matriz por um número complexo, a saber,

$$A + B = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \end{pmatrix}$$

e

$$k.A = k \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} k.a_1 \\ k.a_2 \\ \vdots \end{pmatrix},$$

com o produto interno

$$\langle A|B \rangle = \sum_{k=1}^{\infty} a_k^* b_k.$$

Os dois exemplos a seguir são mais técnicos e podem ser descartados por estudantes de filosofia em um primeiro curso (eles generalizam o anterior).

Exemplo 2.2 (O espaço $\ell^2(\mathbb{N})$). Vamos denotar por \mathbb{N} o conjunto dos números naturais, e por $\ell^2(\mathbb{N})$ o conjunto de todas as funções $f: \mathbb{N} \rightarrow \mathbb{C}$ tais que a soma $\sum_{i=0}^{\infty} |f(i)|^2 < \infty$ (é finita). Definimos, para duas de tais funções f e g quaisquer,

$$\langle f|g \rangle := \sum_{i=0}^{\infty} f^*(i)g(i),$$

e verificamos que se trata de um produto interno. Temos então um espaço de Hilbert separável infinito dimensional que é isomorfo a todos os espaços de Hilbert separáveis e de dimensão infinita. Para comprovar que é separável, basta mostrar (vide [HeiZim.12, p.4]) que ele possui uma base enumerável (de funções chamadas de **funções de Kronecker**, que aparecerão novamente abaixo).

Exemplo 2.3. Outro exemplo importante é o espaço \mathcal{L}^2 das **funções quadrado integráveis**, ou seja, de todas as funções complexas f assumindo valores reais que satisfazem

$$\int_{-\infty}^{\infty} |f(x)|^2 dx < \infty,$$

com o produto interno sendo definido por

$$\langle f|g \rangle = \int_{-\infty}^{\infty} f^*(x)g(x)dx.$$

As operações de adição de vetores (funções) e de multiplicação de vetor (função) por escalar são as usuais. Este exemplo aparecerá novamente abaixo (página 118).

2.1. Espaços de Hilbert e mecânica quântica

O ‘formalismo’ usual da mecânica quântica (MQ), introduzido por von Neumann em 1932, faz uso da teoria de espaços de Hilbert complexos separáveis,

como veremos no último capítulo. A palavra ‘formalismo’ é usada pelos físicos para designar a formulação matemática da mecânica quântica (não relativista, a única que consideraremos) e nada tem a ver, em princípio, com sistemas formais que são tratados em lógica e em fundamentos da matemática.⁴

No formalismo da MQ, é importante descrever os **estados** dos sistemas físicos e considerar os **observáveis** físicos (como massa, carga elétrica, posição, momento, spin) que podem ser *medidos* quando o sistema encontra-se em um certo estado. Deixaremos por enquanto os observáveis de fora até que tenhamos introduzido o conceito de *operador linear*. Os estados de um sistema físico σ são representados por vetores de um espaço de Hilbert \mathcal{H} . Ou seja, a cada sistema físico σ associamos um \mathcal{H} cujos vetores representam seus estados. Convenciona-se que se α denota um estado, então $k.\alpha$ denota o mesmo estado, para todo escalar k , o que implica que podemos considerar apenas vetores unitários. Assim, qualquer múltiplo escalar de α representa o mesmo estado que α .

Ora, sabemos que dado um vetor não nulo, a coleção dos múltiplos escalares desse vetor, munido das restrições das operações do espaço, forma um sub-espaço vetorial de dimensão 1. Os físicos chamam esses sub-espaços de *rays*. Assim, dizer que os vetores *unitários* representam os estados é um abuso de linguagem, mas eles são suficientes, pois constituem uma base para cada *ray*. Em geral, este escalar é tomado como sendo $k = e^{i\theta}$, que já encontramos antes (página 35 – trata-se portanto de um número complexo $z = re^{i\theta}$ onde $r = 1$), e é chamado de **fator de fase** de sorte que α e $\beta = e^{i\theta}\alpha$ são considerados como representando o mesmo estado. O que se afirma quando dizemos que vetores múltiplos representam o mesmo estado equivale a ‘ignorar’ o fator de fase [SusFri.14, p.108].

Rays são importantes porque mais tarde mostraremos que a cada um deles está associado um **operador de projeção**, que ‘projeta’ um dado vetor do espaço no sub-espaço correspondente. Em mecânica quântica, interessa também a álgebra dos sub-espaços fechados do espaço de Hilbert, base para a área conhecida como **lógica quântica**.

⁴A mecânica quântica relativista é denominada de **teoria quântica de campos**. Há várias de tais teorias, e em suma trata-se da junção da mecânica quântica tradicional com a relatividade restrita. Não se sabe até o momento como compatibilizar a mecânica quântica com a relatividade geral.

2.2. Lógica quântica

Dito sem muitos detalhes, sabe-se que a álgebra dos sub-conjuntos de um conjunto dado, considerando-se as operações de união, interseção e de complementação, forma uma **álgebra de Boole**. Não sem surpresa, a álgebra subjacente ao cálculo proposicional clássico, alicerce da chamada **lógica clássica**, é também uma álgebra de Boole, assim como a **álgebra dos observáveis da mecânica clássica** (para referências, indico o meu texto ainda no prelo, *Introdução à Lógica Quântica*; referências adequadas podem ser buscadas em [Red.87]).⁵ Parece que tudo isso está em completo acordo. Mas, na mecânica quântica, considerando-se a abordagem via espaços de Hilbert (que são conjuntos), se quiséssemos proceder do mesmo modo, considerando os sub-espaços fechados e as operações de união, interseção e complemento entre conjuntos, não teríamos sucesso. Bom, já vimos, a união de sub-espaços nem sempre é um sub-espaço, assim que deveríamos tomar como operação não a união, mas o processo de tomar o *espaço gerado pela união* e, por complemento, o complemento ortogonal (como veremos a seguir). Consideremos alguns sub-espaços do \mathbb{R}^2 para constatar que a lei distributiva da interseção relativamente ‘a ‘união’ não vale. Ou seja, consideremos os seguintes sub-espaços (veja a Figura 2.1):

1. $X = \{(x, 0) : x \in \mathbb{R}\}$ ("eixo X").
2. $Y = \{(0, y) : y \in \mathbb{R}\}$ ("eixo Y").
3. $R = \{(x, x) : x \in \mathbb{R}\}$ ("reta $x = y$ ").

Tomemos \mathcal{H} como o espaço euclidiano \mathbb{R}^2 munido do produto interno canônico, U , V e W subespaços definidos respectivamente (e adequadamente) como correspondendo intuitivamente aos eixos OX , OY e à reta $x = y$. Nota-se então que $X \sqcup (Y \cap Z) = X$, ao passo que $(X \sqcup Y) \cap (X \sqcup Z) = \mathbb{R}^2$. Em outras palavras, o reticulado dos sub-espaços de \mathbb{R}^2 não é distributivo.

Esse resultado vale em geral: tomando \mathcal{F} como a coleção de todos os sub-espaços fechados de \mathcal{H} munido das operações indicadas, junto com o sub-espaço trivial trivial e do próprio \mathcal{H} , verifica-se que esta estrutura não é uma

⁵Uma observação para os filósofos. É preciso um cuidado aqui, pois pode-se ser levado a pensar que o único modo de se encontrar um modelo algébrico para a lógica proposicional clássica é apresentar uma álgebra de Boole, pois há outras estruturas, ou melhor, certos reticulados, que modelam essa lógica igualmente sem serem álgebras de Boole. Ver [PavMeg.99].

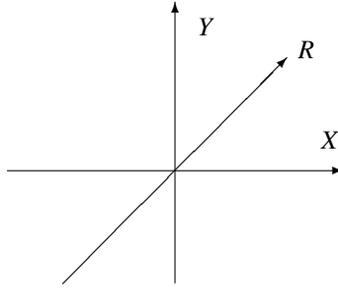


Figura 2.1: Sub-espacos (rays) do \mathbb{R}^2 .

álgebra de Boole, mas um **reticulado ortomodular** (veja o exercício 5.2). Para mais detalhes, recomendamos [DalGiuGre.04], [Red.87], [RonDomFre.16], e há vários sites na internet sobre o assunto.

Mas, por que razão costuma-se dizer que a lógica da mecânica quântica não é clássica? O problema é que nesta disciplina nem todas as propriedades de uma álgebra de Boole vigoram. Vejamos isso com um exemplo.

Consideremos o **spin** de uma determinada partícula, digamos um elétron, que pode ser avaliada segundo uma direção determinada. É um *fato* da física que o spin de um elétron pode assumir apenas um dentre dois valores, $1/2$ ou $-1/2$ (que vamos denotar simplesmente por $+$ e $-$). Consequentemente, chamando de s_e^x o spin do elétron e na direção x , em virtude do que se disse acima, $(s_e^x = +) \vee (s_e^x = -)$ é verdadeira. Outro fato aceito pela mecânica quântica é o Princípio de Indeterminação de Heisenberg (para spin), que assera que o spin de uma partícula não pode ser medido simultaneamente em duas direções distintas. Suponha agora que x e y sejam duas direções distintas e que obtivemos $s_e^y = +$ por medição. Então, podemos dizer que

$$(s_e^y = +) \wedge ((s_e^x = +) \vee (s_e^x = -)).$$

Mas então, usando a lógica proposicional (em especial, a Lei Distributiva $\alpha \wedge (\beta \vee \gamma) \leftrightarrow (\alpha \wedge \beta) \vee (\alpha \wedge \gamma)$), obtemos:

$$(s_e^y = + \wedge s_e^x = +) \vee (s_e^y = + \wedge s_e^x = -).$$

No entanto, qualquer uma dos componentes dessa disjunção ou é falsa ou sem sentido, em virtude do que se disse acima. Em outras palavras, a lei distributiva, uma das mais fundamentais da lógica tradicional, não valeria no contexto da mecânica quântica, como sustentaram von Neumann e Birkhoff e muitos depois deles. Em resumo, o que esses autores verificaram foi que o reticulado subjacente às proposições da mecânica quântica não é uma álgebra de Boole, mas um *reticulado ortomodular* (análise extensa em [DalGiuGre.04]).

3. Ortogonalidade

Daqui para frente, assumiremos sempre que \mathcal{H} é um espaço de Hilbert.

Definição 3.1 (Vetores Ortogonais e Ortonormais). Dois vetores α e β de \mathcal{H} são **ortogonais** se $\langle \alpha | \beta \rangle = 0$. Eles são **ortonormais** se, além de ortogonais, são unitários. Por vezes escreveremos $\alpha \perp \beta$ para indicar a ortogonalidade, ou então α^\perp para indicar o vetor ortogonal a α .

Exemplo 3.1. O espaço \mathbb{R}^n munido do produto interno canônico é um espaço de Hilbert (exercício). Os vetores da base canônica Ξ (veja à página 14) são ortonormais relativamente a esse produto interno. Situação análoga ocorre com o \mathbb{C}^n .

Uma base para \mathcal{H} formada por vetores ortonormais é uma **base ortonormal** de \mathcal{H} .

Exemplo 3.2 (Séries de Fourier). Considere o espaço de Hilbert das funções seccionalmente contínuas no intervalo $[-\pi, \pi]$ da reta real munido do produto interno

$$\langle f | g \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)g(x)dx.$$

Uma função seccionalmente contínua nesse intervalo é uma função que é contínua nesse intervalo ou que tem no máximo um número finito de descontinuidades de primeira espécie (se ela é descontínua em um ponto do intervalo, isso se deve ao fato de que seus limites laterais são distintos no ponto, mas existem, contrariamente às funções que são descontínuas de segunda espécie, como $f(x) = \tan(x)$, que não tem limites laterais em pontos como $a = \frac{\pi}{2}$).

Verifica-se que as funções $1, \sin x, \cos 2x, \sin 2x, \cos 2x, \dots$ são vetores ortonormais relativamente ao produto interno acima. Se f é uma função seccionalmente contínua no intervalo dado, ela é integrável no intervalo e é limite da série seguinte, que pode ser vista como sua expressão como combinação linear das funções ortonormais dadas, a saber,

$$f(x) = a_0 1 + a_1 \sin x + b_1 \cos x + a_2 \sin 2x + b_2 \cos 2x + \dots \quad (2.7)$$

O que necessitamos é aprender a calcular os coeficientes, os ‘coeficientes de Fourier’, o que faremos abaixo no caso geral de um \mathcal{H} qualquer. A expressão (2.7) é dita ser o desenvolvimento de $f(x)$ em série de Fourier.

3.1. Processo de Gram-Schmidt

Denomina-se de *Processo de Ortogonalização de Gram-Schmidt* o seguinte procedimento para, dada uma base ordenada $\mathcal{A} = \{\alpha_1, \dots, \alpha_n\}$ para um espaço \mathcal{H} , encontrarmos uma base ordenada ortogonal $\mathcal{B} = \{\beta_1, \dots, \beta_n\}$ para \mathcal{H} , do seguinte modo:

$$(1) \beta_1 = \alpha_1$$

$$(2) \beta_{m+1} = \alpha_{m+1} - \sum_{i=1}^m \frac{\langle \alpha_{m+1} | \beta_i \rangle}{\|\alpha_i\|^2} \alpha_i.$$

Constata-se sem dificuldade que os β_j são ortogonais entre si. Para exemplificar, mostraremos que β_2 é ortogonal a β_1 . Com efeito, lembrando que $\beta_1 = \alpha_1$, temos que

$$\langle \beta_1 | \beta_2 \rangle = \langle \alpha_1 | \alpha_2 - \frac{\langle \alpha_1 | \alpha_2 \rangle}{\|\alpha_1\|^2} \alpha_1 \rangle = \langle \alpha_1 | \alpha_2 \rangle - \frac{\langle \alpha_1 | \alpha_2 \rangle}{\|\alpha_1\|^2} \langle \alpha_1 | \alpha_1 \rangle = 0.$$

Uma vez obtida a base ortogonal \mathcal{B} , uma base ortonormal é obtida simplesmente dividindo-se cada vetor de \mathcal{B} pela sua norma, ou seja, obtendo

$$\mathcal{N} = \left\{ \frac{\beta_1}{\|\beta_1\|}, \dots, \frac{\beta_n}{\|\beta_n\|} \right\}.$$

O que o Processo de Gram-Schmidt nos garante é que, dada qualquer base, sempre podemos encontrar uma base ortonormal. Isso é relevante, pois podemos nos ater a essas bases unicamente, como fazemos em mecânica quântica, onde espaços de Hilbert são fundamentais. Vetores unitários de um espaço de Hilbert \mathcal{H} representarão os *estados* dos sistemas físicos, e certos operadores (que veremos à frente) representarão os *observáveis* físicos.

3.2. Coeficientes de Fourier

O ‘caso geral’ ao qual nos referimos acima no caso do exemplo das séries de Fourier, trata de encontrarmos as coordenadas de um vetor em uma base ortonormal. Tais coeficientes são denominados de **coeficientes de Fourier**.

Seja $\mathcal{A} = \{\alpha_1, \dots, \alpha_n\}$ uma base ortonormal ordenada para um espaço de Hilbert \mathcal{H} . Se β é um vetor qualquer desse espaço, existem escalares x_i ($i = 1, \dots, n$), tais que

$$\beta = x_1 \alpha_1 + \dots + x_n \alpha_n. \quad (2.8)$$

Ora, sabemos que

$$\langle \alpha_i | \alpha_j \rangle = \delta_{ij}, \quad (2.9)$$

sendo δ_{ij} o símbolo de Kronecker, logo

$$\langle \alpha_i | \beta \rangle = \langle \alpha_i | x_1 \alpha_1 + \dots + x_n \alpha_n \rangle = x_1 \langle \alpha_i | \alpha_1 \rangle + \dots + x_i \langle \alpha_i | \alpha_i \rangle + \dots + x_n \langle \alpha_i | \alpha_n \rangle.$$

Tendo em vista (2.9), resulta que

$$\langle \alpha_i | \beta \rangle = x_i. \quad (2.10)$$

Portanto, em (2.8), temos

$$\beta = \langle \alpha_1 | \beta \rangle \alpha_1 + \dots + \langle \alpha_n | \beta \rangle \alpha_n = \sum_{i=1}^n \langle \alpha_i | \beta \rangle \alpha_i. \quad (2.11)$$

Se os vetores da base \mathcal{A} não fossem ortonormais, mas ortogonais simplesmente, deveríamos dividir cada vetor da base por sua norma, de modo a torná-los unitários; assim,

$$\beta = \sum_{i=1}^n \frac{\langle \alpha_i | \beta \rangle}{\|\alpha_i\|} \cdot \alpha_i. \quad (2.12)$$

De maneira geral, os coeficientes de Fourier são portanto

$$x_i = \frac{\langle \alpha_i | \beta \rangle}{\|\alpha_i\|}, \quad (2.13)$$

e são exatamente esses que devem ser buscados na expressão (2.7), levando em conta o produto interno e as funções envolvidas.

Exercício 3.1. Calcule os coeficientes da expressão (2.7).

Exemplo 3.3. Considere os vetores $\alpha_1 = (1, 0, 1)$, $\alpha_2 = (0, 1, 2)$ e $\alpha_3 = (-1, 1, 0)$ do \mathbb{R}^3 , visto como espaço vetorial real com o produto interno canônico. Notamos que eles são linearmente independentes (exercício), logo formam uma base para o espaço. Usando o Processo de Gram-Schmidt, vamos encontrar uma base ortogonal para o mesmo espaço, ‘ortogonalizando’ a base acima. Usando o procedimento indicado acima, fazemos:

$$\beta_1 = \alpha_1 = (1, 0, 1)$$

$$\beta_2 = \alpha_2 - \frac{\langle \alpha_2 | \beta_1 \rangle}{\|\beta_1\|^2} \beta_1.$$

E por fim

$$\beta_3 = \alpha_3 - \frac{\langle \alpha_3 | \beta_1 \rangle}{\|\beta_1\|^2} \beta_1 - \frac{\langle \alpha_3 | \beta_2 \rangle}{\|\beta_2\|^2} \beta_2.$$

Agora, faça cuidadosamente as contas. Mas repare que os vetores $\beta_1, \beta_2, \beta_3$ não são ainda ortonormais. Para isso, deverá dividir cada um deles pela sua respectiva norma.

Se você prestar atenção no que está ocorrendo, notará o seguinte. Partimos de um dos vetores, digamos α_1 (pois fizemos β_1 ser igual a ele). Em seguida, obtivemos um vetor ortogonal a ele da seguinte forma: projetamos ortogonalmente (isso ficará mais claro à frente, mas sua intuição basta) α_2 sobre ele e tomamos o vetor diferença α_2 menos essa projeção, que é ortogonal a α_1 (isso precisa ser demonstrado, mas é fácil). Depois, tomamos α_3 e projetamos *sobre o sub-espaço gerado* pelos dois vetores obtidos, e de novo tomamos a diferença α_3 menos essa projeção. Projeções são muito importantes e serão consideradas mais tarde (seção 2).

Exercício 3.2. Considere a base ortonormal obtida no exemplo precedente. Expresse o vetor $\gamma = (-2, 1, 2)$ como combinação linear dos vetores dessa base e encontre os coeficientes de Fourier correspondentes.

Exercício 3.3. Considere no espaço $\ell^2(\mathbb{N})$ as chamadas **funções de Kronecker**, definidas assim:

$$\delta_k(j) = \begin{cases} 1 & \text{se } k = j \\ 0 & \text{se } k \neq j \end{cases}$$

Mostre que (i) $\langle \delta_k | \delta_\ell \rangle = 0$ sempre que $k \neq \ell$, logo, o conjunto das funções de Kronecker é um conjunto ortonormal de vetores; (ii) o argumento acima sobre o Lema de Zorn para mostrar que este conjunto é uma base para o espaço $\ell^2(\mathbb{N})$.

4. A condição de normalização na teoria quântica

Já vimos que o espaço vetorial \mathcal{L}^2 de todas as funções complexas $f(x)$ tomando valores em \mathbb{R} , tais que

$$\int_{-\infty}^{\infty} |f(x)|^2 dx < \infty,$$

ditas funções ‘quadrado-integráveis’, munido das operações usuais de adição de funções e de multiplicação de função por escalar real, e com o produto interno

$$\langle f | g \rangle = \int_{-\infty}^{\infty} f^*(x)g(x)dx$$

é relevante em mecânica quântica.

Na mecânica quântica de ondas, um sistema de partículas em uma dimensão tem seus estados descritos por uma **função de onda** $\psi(x, t)$, que satisfaz

$$\int_{-\infty}^{\infty} |\psi(x, t)|^2 dx = 1. \quad (2.14)$$

Podemos entender $\psi(x, t)$ como pertencente a \mathcal{L}^2 , para t (a coordenada temporal) fixado, e (2.14) é então dita **condição de normalização**, ou seja, $|\psi(x, t)|^2 = \psi^*(x, t)\psi(x, t) = 1$. Esta condição está associada ao papel desempenhado pelo conceito de probabilidade na teoria quântica.⁶ Com efeito, segundo a **interpretação probabilista** da função de onda devida a Max Born, dado um intervalo $[a, b]$ da reta real, a probabilidade de encontrarmos o valor da medida de um observável físico \hat{A} medido sobre um sistema no estado descrito por $\psi(x, t)$ neste intervalo é precisamente

$$\text{Prob}_{[a,b]}^{\psi(x,t)}(A) = \int_{[a,b]} |\psi(x, t)|^2 dx.$$

⁶Foge aos nossos objetivos discutir aqui o caso da probabilidade em mecânica quântica, que é bastante controverso. Mesmo assim, algumas palavras sobre isso serão ditas mais à frente.

Mais à frente, veremos o caso de dimensão finita, ao qual nos restringiremos. Os espaços ℓ^2 e \mathcal{L}^2 são isomorfos.

Digamos que o observável a ser medido para um sistema composto por uma partícula em um espaço de dimensão igual à unidade. Assim, a expressão simplificada

$$p = \int_a^b |\psi(x, t)|^2 dx \quad (2.15)$$

designa a probabilidade de encontrar a partícula, representada pela função de onda $\psi(x, t)$ no intervalo $[a, b]$ da reta real.

O valor $|\psi(x, t)|^2$ é denotado

$$\rho(x, t) \quad (2.16)$$

e denominado na forma de uma função é denotada de função de **densidade de probabilidade**, ou função de **probabilidades**. Informalmente, ela esta nos dando, para cada x e t , uma medida de como a probabilidade, fornecida pelo quadrado da função ψ se distribui em torno do ponto x . Na mecânica quântica, jamais temos uma partícula localizada em um ponto, como assumimos ser possível na física clássica, mas temos somente uma probabilidade de ela estar nas cercanias do ponto, o que é dado pela função mencionada. É assim que devemos ‘ler’ o resultado do exemplo a seguir.

Exemplo 4.1. Suponha (para t fixo), que $\rho(x) = 1/x^2$. A probabilidade de encontrarmos a partícula em $[1, 2]$, é dada por

$$p = \int_1^2 1/x^2 dx = \left. \frac{-1}{x} \right|_1^2 = 0,5.$$

Obviamente, se aumentarmos o tamanho do intervalo, por exemplo tomando $[1, 3]$, a probabilidade será maior (dando 0,67), pois será ainda mais provável encontrar a partícula dentro de um intervalo maior. Se integramos sobre a reta toda (de $-\infty$ a $+\infty$), a probabilidade dará 1.⁷

Voltaremos a essa e outras questões ‘quânticas’ mais tarde.

⁷Para alunos de filosofia, há sites que fornecem uma calculadora para a integral definida.

5. A notação de Dirac

Paul Dirac (1902-1984), notável físico britânico, Prêmio Nobel de física em 1933, introduziu uma notação extremamente sugestiva e útil para o desenvolvimento do formalismo da MQ. Primeiramente, devemos saber que em inglês, os limitadores $\langle \rangle$ são chamados de *brackets*, e note que são eles que utilizamos para representar o produto interno de dois vetores, escrevendo $\langle \alpha | \beta \rangle$. Dirac separou as duas partes, obtendo $\langle \alpha |$ e $| \beta \rangle$. Os segundos representam vetores, e são chamados de **kets**; os primeiros representam operadores (certas funções que associam vetores a vetores que veremos no próximo capítulo), os **bras**. Assim, uma expressão como $\langle \alpha | \beta \rangle$ pode ser vista tanto como o produto interno entre os vetores α e β como a ação (imagem) do vetor $| \beta \rangle$ pelo operador (de um tipo especial, chamado de **funcional linear**, dos quais falaremos no próximo capítulo) $\langle \alpha |$, ou seja,⁸

$$\langle \alpha | \beta \rangle = \langle \alpha | (| \beta \rangle).$$

No capítulo que segue, faremos uso desta notação, e outras explicações serão dadas.

⁸Repare que no segundo membro estamos utilizando uma notação comum para funções, como quando escrevemos $f(x)$, só que aqui f é $\langle \alpha |$ e x é $| \beta \rangle$.

OPERADORES LINEARES

OPERADORES LINEARES são particularmente importantes no formalismo da mecânica quântica. Alguns operadores (os auto-adjuntos, ou hermitianos) representarão os *observáveis* físicos. Iniciaremos com um caso mais geral, o das transformações lineares, ou homomorfismos entre espaços vetoriais. Operadores lineares são casos especiais de transformações lineares: são transformações lineares de um espaço nele mesmo.

Definição 0.1 (Transformação linear). Uma **transformação linear**, ou **homomorfismo** de um espaço vetorial \mathcal{V} em um espaço vetorial \mathcal{W} , ambos sobre um mesmo corpo \mathcal{K} , é uma aplicação $T : \mathcal{V} \mapsto \mathcal{W}$ tal que, para todos $\alpha, \beta \in \mathcal{V}$ e $a \in K$, se tenha que

1. $T(\alpha + \beta) = T(\alpha) + T(\beta)$
2. $T(a\alpha) = aT(\alpha)$.

As condições 1 e 2 são chamadas de **condições de linearidade**. Podemos substituir 1 e 2 por uma só condição, a seguinte, para b também em K :

$$T(a\alpha + b\beta) = aT(\alpha) + bT(\beta).$$

Notação Muitas vezes escreveremos simplesmente $T\alpha$ em vez de $T(\alpha)$, principalmente quando estivermos usando os kets de Dirac.

Definição 0.2 (Operador linear). Um **operador linear** sobre um espaço vetorial \mathcal{V} é uma transformação linear de \mathcal{V} em \mathcal{V} .

Exemplo 0.1. A aplicação $T : \mathbb{R}^2 \mapsto \mathbb{R}^2$ definida por $T(x, y) = (x + y, -x + 2y)$ é um operador linear sobre o \mathbb{R}^2 , como facilmente se verifica.

Exemplo 0.2. Considere o espaço vetorial das funções diferenciáveis em um intervalo $[a, b] \subseteq \mathbb{R}$. A aplicação D definida por

$$D(f(x)) = \frac{df(x)}{dx}$$

é um operador linear sobre o espaço referido.

Exemplo 0.3. Considere o espaço vetorial das matrizes complexas de ordem 2×1 munido das operações usuais de adição de matrizes e de multiplicação de matriz por número complexo. A função T definida abaixo é um operador linear sobre este espaço:

$$T \begin{pmatrix} x \\ z \end{pmatrix} = \begin{pmatrix} 2+i & -i \\ i & 1+i \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

O leitor deve recordar que o produto de matrizes é o produto ‘linha por coluna’. Assim, no produto do segundo membro, multiplicamos a primeira linha da matriz 2×2 como se ela fosse um vetor pela coluna que lhe segue de acordo com o produto interno canônico, dando como resultado o primeiro elemento de uma matriz 2×1 . O segundo elemento é obtido multiplicando-se do mesmo modo a segunda linha pelo vetor. O resultado é

$$T \begin{pmatrix} x \\ z \end{pmatrix} = \begin{pmatrix} (2+i)x - iy \\ ix + (1+i)y \end{pmatrix}$$

que podemos escrever, face ao isomorfismo entre o espaço das matrizes complexas 2×1 e 1×2 como¹

$$T(x, y) = (2x + (x - y)i, y + (x + y)i).$$

Notação Em física quântica, os físicos distinguem (acertadamente) entre os observáveis físicos, como massa, spin, carga elétrica, momento, etc. e os operadores que os representam, mudando um pouco a notação: se O é um observável físico, como o spin de uma partícula, então \hat{O} é o operador que o representa. No momento, não faremos essa distinção.

¹Na verdade, rigorosamente falando, os operadores não seriam os mesmos, já que atuam sobre espaços vetoriais distintos, mas isso é um detalhe que não importa matematicamente: o isomorfismo nos dá muitas vantagens.

1. Representação matricial

Nesta seção, utilizaremos a notação de Dirac, para que o leitor vá se acostumando a ela assim como com a notação que irá sendo introduzida. Seja \mathcal{V} espaço de dimensão finita n sobre \mathcal{K} (real ou complexo), e seja $\mathcal{A} = \{|\alpha_i\rangle\}$ uma base ordenada para \mathcal{V} .² Se T é um operador linear sobre \mathcal{V} e $|\beta\rangle \in \mathcal{V}$, podemos escrever $|\beta\rangle$ como combinação linear dos vetores da base, ou seja, há escalares x_1, \dots, x_n tais que

$$|\beta\rangle = x_1|\alpha_1\rangle + \dots + x_n|\alpha_n\rangle. \quad (3.1)$$

A transformada de $|\beta\rangle$ pelo operador T , sendo vetor de \mathcal{V} , pode também ser escrito como combinação linear dos vetores da base, ou seja, escrevendo, conforme nossa notação feita acima, $T|\gamma\rangle$ para $T(|\gamma\rangle)$,

$$T|\beta\rangle = y_1|\alpha_1\rangle + \dots + y_n|\alpha_n\rangle = \sum_{j=1}^n y_j|\alpha_j\rangle. \quad (3.2)$$

Conheceremos T se soubermos como encontrar, dados os x_i (as coordenadas de $|\beta\rangle$ na base), os escalares y_i (as coordenadas de $T|\beta\rangle$).

A partir a combinação linear (3.2), obtemos

$$T|\beta\rangle = T(x_1|\alpha_1\rangle + \dots + x_n|\alpha_n\rangle),$$

ou seja, pela linearidade de T ,

$$T|\beta\rangle = x_1T|\alpha_1\rangle + \dots + x_nT|\alpha_n\rangle.$$

em notação mais cômoda,

$$T|\beta\rangle = \sum_{i=1}^n x_iT|\alpha_i\rangle. \quad (3.3)$$

Por sua vez, as transformadas $T(|\alpha_i\rangle)$ dos vetores da base podem ser escritos como combinações lineares dos vetores da própria base, ou seja,

²Note que estamos simplificando a escrita, escrevendo $\{|\alpha_i\rangle\}$ em vez de $\{|\alpha_1\rangle, |\alpha_2\rangle, \dots, |\alpha_n\rangle\}$. Obviamente, $i = 1, \dots, n$, o que ficará implícito.

$$\begin{aligned}
 T|\alpha_1\rangle &= a_{11}|\alpha_1\rangle + a_{12}|\alpha_2\rangle + \cdots + a_{1n}|\alpha_n\rangle = \sum_{j=1}^n a_{1j}|\alpha_j\rangle \\
 T|\alpha_2\rangle &= a_{21}|\alpha_1\rangle + a_{22}|\alpha_2\rangle + \cdots + a_{2n}|\alpha_n\rangle = \sum_{j=1}^n a_{2j}|\alpha_j\rangle \\
 &\vdots \\
 T|\alpha_n\rangle &= a_{n1}|\alpha_1\rangle + a_{n2}|\alpha_2\rangle + \cdots + a_{nn}|\alpha_n\rangle = \sum_{j=1}^n a_{nj}|\alpha_j\rangle
 \end{aligned}$$

Ou simplesmente,

$$T|\alpha_i\rangle = \sum_{j=1}^n a_{ij}|\alpha_j\rangle, \quad i = 1, \dots, n \quad (3.4)$$

Substituindo (3.4) em (3.3), obtemos

$$T|\beta\rangle = \sum_{i=1}^n x_i \sum_{j=1}^n a_{ij}|\alpha_j\rangle. \quad (3.5)$$

ou seja, rearranjando os somatórios,

$$T|\beta\rangle = \sum_{j=1}^n \sum_{i=1}^n a_{ij}x_i|\alpha_j\rangle. \quad (3.6)$$

Igualando os coeficientes de (3.2) e (3.6), obtemos

$$y_j = \sum_{i=1}^n a_{ij}x_i, \quad j = 1, \dots, n \quad (3.7)$$

Em notação matricial, temos

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & & & \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

A matrix $A = [a_{ij}]$ dos coeficientes das combinações lineares acima age como se fosse o operador, permitindo que, dadas as coordenadas x_1, \dots, x_n de um vetor numa base ordenada, encontremos as coordenadas y_1, \dots, y_n de sua transformada na mesma base.

A matrix $A = [a_{ij}]$ chama-se **matriz representativa** do operador T na base \mathcal{A} , denotada $[T]_{\mathcal{A}}$. Temos então que

$$[T(\alpha)]_{\mathcal{A}}^T = [T]_{\mathcal{A}}[\alpha]_{\mathcal{A}}^T. \quad (3.8)$$

Ou seja, a matriz $[T]_{\mathcal{A}}$, agindo sobre a matriz das coordenadas do vetor α na base \mathcal{A} , a saber, $[\alpha]_{\mathcal{A}}$ (tomada transposta para que o produto seja possível), fornece a matriz das coordenadas na mesma base da transformada do vetor pelo operador, $[T(\alpha)]_{\mathcal{A}}$.

Regra prática Para achar a matriz representativa de um operador T em uma base finita $\{|\alpha_i\rangle\}$, proceda como segue:

1. Transforme os vetores da base usando T
2. Escreva as transformadas como combinações lineares dos vetores da mesma base.
3. Ache os coeficientes (o que pode fazer resolvendo sistemas de equações lineares)
4. A matriz é formada por esses coeficientes colocados como colunas.

Exemplo 1.1. Consideremos o operador $T(x, y) = (x + y, -x + 2y)$ sobre o \mathbb{R}^2 e a base ordenada $\mathcal{A} = \{(1, 2), (-1, 1)\}$. Seguiremos os passos indicados acima.

$$T(1, 2) = (3, 3) = a_{11}(1, 2) + a_{12}(-1, 1)$$

$$T(-1, 1) = (0, 3) = a_{21}(1, 2) + a_{22}(-1, 1)$$

Os sistemas lineares indicam que $a_{11} = 2$, $a_{12} = -1$, $a_{21} = 1$ e $a_{22} = 1$. Assim,

$$[T]_{\mathcal{A}} = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}$$

A seguinte notação é útil, quando $\{|\alpha_i\rangle\}$ for uma base ortonormal de \mathcal{V} . Neste caso, tomemos novamente as expressões (3.4). Observamos que

$$T|\alpha_i\rangle = a_{i1}|\alpha_1\rangle + a_{i2}|\alpha_2\rangle + \cdots + a_{in}|\alpha_n\rangle$$

Façamos agora o produto interno $\langle\alpha_i|T|\alpha_i\rangle$, que devido ao fato de que $\langle\alpha_i|\alpha_j\rangle = \delta_{ij}$, obtemos

$$\langle\alpha_i|T|\alpha_i\rangle = a_{ii}.$$

Isso mostra que a matriz representativa de T na base ortonormal $\{|\alpha_i\rangle\}$ pode ser escrita (esta matriz aparecerá novamente na seção 1)

$$[T]_{\mathcal{A}} = \begin{pmatrix} \langle \alpha_1 | T | \alpha_1 \rangle & \langle \alpha_1 | T | \alpha_2 \rangle & \cdots & \langle \alpha_1 | T | \alpha_n \rangle \\ \langle \alpha_2 | T | \alpha_1 \rangle & \langle \alpha_2 | T | \alpha_2 \rangle & \cdots & \langle \alpha_2 | T | \alpha_n \rangle \\ \vdots & & & \\ \langle \alpha_n | T | \alpha_1 \rangle & \langle \alpha_n | T | \alpha_2 \rangle & \cdots & \langle \alpha_n | T | \alpha_n \rangle \end{pmatrix} \quad (3.9)$$

A partir dessa matriz, definimos o **traço** do operador T como a soma dos elementos da diagonal principal da matrix acima, ou seja,

$$\text{Tr}(T) = \sum_{i=1}^n \langle \alpha_i | T | \alpha_i \rangle. \quad (3.10)$$

Em mecânica quântica, a expressão $\langle \alpha | T | \alpha \rangle$ é de grande importância. Como veremos na seção (2), a notação $\langle \alpha |$ denotará uma certa transformação linear (dita *funcional linear*) e então $\langle \alpha_i | T | \alpha_i \rangle$ pode ser visto como um abreviação da imagem do vetor $T | \alpha_i \rangle$ pelo funcional linear $\langle \alpha_i |$. A expressão ainda indica o **valor esperado** (um conceito estatístico) da medida do observável representado por T para o sistema físico no estado $|\alpha\rangle$, denotado $\langle T \rangle_{|\alpha\rangle}$, como veremos na seção 1.

1.1. O espaço dos operadores

Seja $\mathcal{E} = \langle \mathcal{V}, \mathcal{K}, +, \cdot \rangle$ um espaço vetorial e consideremos a coleção de todos os operadores lineares sobre \mathcal{E} . Dados dois de tais operadores, U e T , definamos as operações seguintes:

$$(U + T)(\alpha) = U(\alpha) + T(\alpha),$$

$$(kT)(\alpha) = kT(\alpha),$$

para k escalar.

É claro que, munido dessas operações, temos um espaço vetorial sobre \mathcal{K} , que denotaremos por $\mathcal{L}(\mathcal{E})$ (ou por $\mathcal{L}(\mathcal{V})$).³ Deste modo, às vezes escreveremos

³O leitor deve lembrar da convenção que fizemos anteriormente, de chamar indistintamente o espaço vetorial de \mathcal{E} , quando então nos referimos à estrutura, ou de \mathcal{V} , fazendo referência ao conjunto de vetores.

$T \in \mathcal{L}(\mathcal{E})$ (ou $T \in \mathcal{L}(\mathcal{V})$) para indicar que T é um operador linear sobre \mathcal{E} (sobre \mathcal{V}).

Pode-se desenvolver o estudo deste espaço e mostrar, por exemplo, que se \mathcal{V} tem dimensão n , então $\mathcal{L}(\mathcal{V})$ tem dimensão n^2 .⁴

Este espaço é interessante porque podemos considerar uma operação de produto de ‘vetores’ (operadores), definida do seguinte modo:

$$(T \cdot U)(\alpha) := T(U(\alpha)).$$

Abreviaremos $T \cdot U$ por TU . Na verdade, não há nenhum mistério nisso: trata-se da conhecida operação de composição de funções.

Definição 1.1. Um operador linear T sobre \mathcal{V} é **inversível** se existe um operador T^{-1} sobre \mathcal{V} tal que $T^{-1}T = TT^{-1} = I$, sendo I o operador identidade.

Neste caso, T^{-1} é chamado de **inverso** de T . Como T é em particular uma função, uma condição necessária e suficiente para T ser inversível é que seja bijetivo. Depois veremos outra condição.⁵

É agora fácil constatar que o conjunto de todos os operadores lineares inversíveis sobre \mathcal{V} munido do produto de operadores é um grupo não comutativo (exercício).

Definição 1.2 (Núcleo). Seja $T \in \mathcal{L}(\mathcal{V})$. Chama-se **núcleo** (em inglês, **kernel**) de T ao conjunto de vetores de \mathcal{V} cujas transformadas são o vetor nulo, ou seja,

$$\text{Ker}(T) := \{\alpha \in \mathcal{V} : T(\alpha) = \mathbf{0}\}.$$

Repare que $\text{Ker}(T)$ nunca é vazio, porque pelo menos $\mathbf{0} \in \text{Ker}(T)$. Você é capaz de provar isso?⁶ A questão é saber o que acontece quando o núcleo tem *somente* o vetor nulo. Os teoremas abaixo esclarecem a pergunta.

Teorema 1.1. O núcleo de T é um sub-espaço de \mathcal{V} .

Demonstração. Pelo teorema 3.1, basta mostrar que se dois vetores pertencem ao núcleo, sua soma e o seu produto por escalar também pertencem. Faça isso como exercício. □

⁴Um modo de entender isso, ainda que não seja uma demonstração, é o fato de que qualquer matriz de um operador T sobre \mathcal{V} será $n \times n$, ou seja, terá n^2 elementos.

⁵A saber, o núcleo de T , $\text{Ker}(T)$, contém somente o vetor nulo.

⁶Dica: lembre que $\mathbf{0} = 0T(\alpha)$ para todo α . Assim, pela linearidade de T , segue-se que $\mathbf{0} = 0T(\alpha) = T(0\alpha) = T(\mathbf{0})$. Logo $\mathbf{0} \in \text{Ker}(T)$.

Lembrando que um operador linear T sobre um espaço \mathcal{V} é uma função, podemos falar na **imagem** de T , a saber, o conjunto

$$\text{Img}(T) := \{\beta \in \mathcal{V} : \exists(\alpha \in \mathcal{V}) \wedge T(\alpha) = \beta\}.$$

Exercício 1.1. Mostre que $\text{Img}(T)$ é um sub-espaço de \mathcal{V} .

O seguinte resultado será útil nas demonstrações a seguir.

Teorema 1.2. *Um operador linear $T \in \mathcal{L}(\mathcal{V})$ é injetivo se e somente se $\text{Ker}(T) = \{0\}$.*

Demonstração. Se T é injetivo e $\alpha \neq 0$, então $T(\alpha) \neq T(0) = 0$, logo $\text{Ker}(T) = \{0\}$. Reciprocamente, suponha que $\text{Ker}(T) = \{0\}$. Ora, para quaisquer α e β , $T(\alpha) = T(\beta)$ equivale a $T(\alpha) - T(\beta) = T(\alpha - \beta) = 0$. Assim, dada a hipótese, $T(\alpha) = T(\beta)$ se e somente se $\alpha = \beta$, logo T é injetivo. \square

Daqui para a frente, usaremos o fato estabelecido por este teorema livremente.

Teorema 1.3 (Teorema do Núcleo e da Imagem). *Se $T \in \mathcal{L}(\mathcal{V})$, então*

$$\dim(\text{Ker}(T)) + \dim(\text{Img}(T)) = \dim(\mathcal{V}).$$

Demonstração. Seja $\mathcal{A} = \{\alpha_1, \dots, \alpha_k\}$ uma base para o núcleo de T . Se $k = n$, o resultado é imediato, porque o núcleo consistiria de todo o espaço (ou seja, todos os vetores seriam levados por T no vetor nulo), e a dimensão da imagem, que conteria unicamente o vetor nulo, seria zero (lembre que uma base para o espaço trivial é o conjunto vazio, que não tem elementos). Mas se $k < n$, então existem (pelo teorema 4.3) vetores $\alpha_{k+1}, \dots, \alpha_n$ de forma que $\mathcal{B} = \{\alpha_1, \dots, \alpha_k, \alpha_{k+1}, \dots, \alpha_n\}$ é uma base para \mathcal{V} . Bastará portanto provar que $\mathcal{C} = \{T(\alpha_{k+1}), \dots, T(\alpha_n)\}$ é uma base para a imagem de T . Ora, \mathcal{C} gera a imagem de T , pois dado um vetor qualquer β na imagem, existe um vetor α em \mathcal{V} tal que $\beta = T(\alpha)$. Como é $\alpha = x_1\alpha_1 + \dots + x_n\alpha_n$ para adequados x_j , vem que $\beta = x_1T(\alpha_1) + \dots + x_kT(\alpha_k) + x_{k+1}T(\alpha_{k+1}) + \dots + x_nT(\alpha_n)$, e como para $i = 1, \dots, k$ tem-se que $T(\alpha_i) = 0$, resulta o que se quer. Resta mostrar que os vetores de \mathcal{C} são linearmente independentes. Para tanto, suponha que

$$\sum_{i=k+1}^n c_i T(\alpha_i) = 0 \quad \text{ou} \quad T\left(\sum_{i=k+1}^n c_i \alpha_i\right) = 0,$$

o que mostra que $\sum_{i=k+1}^n c_i \alpha_i$ pertence ao núcleo. Logo, existem escalares que permitem exprimi-lo como combinação linear dos vetores da base \mathcal{A} do núcleo, os quais designaremos por y_1, \dots, y_k . Portanto, temos que

$$\sum_{i=k+1}^n c_i \alpha_i = \sum_{i=1}^k y_i \alpha_i,$$

ou seja,

$$\sum_{i=1}^k y_i \alpha_i - \sum_{i=k+1}^n c_i \alpha_i = \mathbf{O},$$

que é uma combinação linear nula de vetores da base \mathcal{V} que, por fazerem parte de uma base, são linearmente independentes. Logo os coeficientes dessa combinação linear devem ser todos nulos, e em particular são nulos os y_j , o que fornece o que se quer. Assim, a dimensão da imagem de T é $n - k$ e tem-se o teorema. \square

Teorema 1.4. $T \in \mathcal{L}(\mathcal{V})$ é inversível se e somente se $\text{Ker}(T) = \{\mathbf{O}\}$.

A demonstração deste teorema é consequência imediata (os matemáticos chamam a uma consequência imediata de um teorema de Corolário do teorema) do seguinte resultado, cuja demonstração usa o teorema do núcleo e da imagem.

Teorema 1.5. *Seja $T \in \mathcal{L}(\mathcal{V})$. Então as seguintes afirmativas são equivalentes:*

1. T é inversível
2. T é injetiva
3. T é sobrejetiva

Demonstração. Repare que as equivalências não valem para funções em geral, pois há funções injetivas que não são sobrejetivas e vice-versa. Mas, no caso de operadores, elas se equivalem. A demonstração consiste em um ciclo de implicações ([Bar.76, p.112]). Se T é inversível, é injetiva e sobrejetiva. Logo 1 implica 2 e 3. Se T é injetiva, $\text{Ker}(T) = \{\mathbf{O}\}$ (Teorema acima)⁷ e portanto a

⁷Sabemos que o vetor nulo se transforma no vetor nulo. Se mais algum além dele também se transformasse no vetor nulo, T não poderia ser injetiva, que é a hipótese.

dimensão da imagem de T é igual à dimensão de \mathcal{V} (pelo teorema do núcleo e da imagem). Mas isso mostra que T é sobrejetiva, pois a imagem coincidirá com todo o espaço. Agora mostraremos que 3 implica 2. Com efeito, se T é sobrejetiva, então sua imagem é todo o espaço e, de novo pelo teorema do núcleo e da imagem, vem que a dimensão de seu núcleo tem que ser zero, ou seja, o núcleo contém somente o vetor nulo. Assim, para quaisquer α e β , $T(\alpha) = T(\beta) = \mathbf{O}$, o que equivale a $T(\alpha - \beta) = \mathbf{O} = T(\mathbf{O})$, o que acarreta $\alpha = \beta$.⁸ Finalmente, provamos que 2 implica 1. De fato, como 2 equivale a 3, sendo T injetiva, é sobrejetiva, e portanto bijetiva, logo inversível. \square

1.2. O comutador

Definição 1.3 (Operadores que comutam). Dois operadores T e U sobre \mathcal{V} **comutam** se $TU = UT$.

Escrevemos $[T, U]$ (dito **comutador** de T e U) para denotar a diferença $TU - UT$, ou seja,

$$[T, U] := TU - UT \quad (3.11)$$

Claro que quando T e U comutam, seu comutador é nulo. Propriedades importantes são as seguintes, aqui só enunciadas:

Teorema 1.6. *O comutador obedece as seguintes propriedades, para quaisquer operadores A, B, C :*

1. $[A, B] = -[B, A]$, ou $[A, B] + [B, A] = \mathbf{O}$

2. $[A, A] = \mathbf{O}$

3. $[A, B + C] = [A, B] + [A, C]$

4. $[A + B, C] = [A, C] + [B, C]$

5. $[AB, C] = [A, C]B + A[B, C]$

6. $[A, BC] = [A, B] \cdot C + B \cdot [A, C]$

7. $[A, [B, C]] + [C, [A, B]] + [B, [C, A]] = \mathbf{O}$

⁸Lembre que podemos expressar que uma função f é injetiva de dois modos equivalentes (um é a contrapositiva do outro): $f(x) = f(y) \rightarrow x = y$ ou $x \neq y \rightarrow f(x) \neq f(y)$.

Demonstração. Exercício. □

Fato relevante em física é o seguinte. Como já dito, os observáveis físicos são representados no formalismo quântico por certos operadores, a saber, aqueles que são denominados de **auto-adjuntos**, ou **hermitianos**.⁹ Intuitivamente, um observável físico é algo que *pode ser medido*. Na física clássica, assume-se que quaisquer dois observáveis podem ‘ser medidos ao mesmo tempo’, ou seja, podem ter seus valores avaliados numa mesma situação física no mesmo instante. Já na física quântica, isso não ocorre, o que representa um fato distintivo desta disciplina. Porém, quando dois operadores que representam observáveis comutam, isso indica que os observáveis correspondentes *podem ser medidos simultaneamente*. O resultado abaixo, que pode ser estabelecido no contexto da física quântica, mostra que isso não acontece com dois observáveis básicos (dentre outros), a posição e o momento. Ou seja,

Se X é um operador que representa a posição de um sistema físico (como uma partícula elementar) e P representa o seu momento, então temos:

1. $[X, X] = 0$
2. $[P, P] = 0$
3. $[X, P] = i \cdot \hbar \cdot I$,

onde, nesta última relação, i é a unidade complexa e $\hbar = \frac{h}{2\pi}$, sendo h a constante de Planck, e I o operador identidade. Os textos de física em geral omitem o operador identidade, escrevendo simplesmente $[X, P] = i \cdot \hbar$, mas o leitor deve atentar para que isto é apenas um abuso de linguagem. Faremos o mesmo mais à frente (seção 2). Isso vai desempenhar um papel importante na física quântica. Esta relação foi introduzida por W. Heisenberg. Ela está dizendo que os operadores de **posição** e de **momento** não comutam, ou seja, não podem ser medidos simultaneamente. Este é um dos mais típicos fatos que caracterizam a mecânica quântica. Na física clássica, assume-se que todos os observáveis podem ter valores simultaneamente; se não conhecemos seus valores, isso deve a uma limitação epistemológica.

⁹Não faremos aqui a distinção entre esses operadores, que precisa ser feita em certos contextos. Assim, usaremos esses dois termos como sinônimos, exceto se for explicitamente mencionado o contrário.

Mais à frente (seção 2), faremos uma aplicação do conceito de comutador, mostrando como ele se relaciona com uma formulação do Princípio de Indeterminação de Heisenberg.

2. Funcionais Lineares

Um caso especialmente importante de transformações lineares é o seguinte. Seja \mathcal{V} um espaço vetorial com produto interno sobre $\mathcal{K} = \langle K, +, \cdot, 0, 1 \rangle$. Como já vimos, podemos considerar \mathcal{K} como um espaço vetorial sobre si mesmo (ou seja, considerar o espaço vetorial que identifica os vetores com os escalares, a adição de vetores com a adição de vetores e a multiplicação de vetor por escalar com o produto de escalares). Temos então a

Definição 2.1 (Funcional linear). Um **funcional linear** é uma transformação linear f de \mathcal{V} em \mathcal{K} (visto como espaço vetorial sobre si mesmo).

Ou seja, trata-se de uma função linear que associa um escalar a cada vetor do espaço. Usaremos letras latinas minúsculas como f, g, h para indicar os funcionais.

Exemplo 2.1. Consideremos o espaço \mathbb{R}^n . Para cada $\alpha = (x_1, \dots, x_i, \dots, x_n)$, definimos a função **projeção sobre a i -ésima coordenada**,

$$f(x_1, \dots, x_i, \dots, x_n) = x_i$$

é um funcional linear sobre \mathbb{R}^n .

Exemplo 2.2. Seja A matriz $n \times n$ sobre o corpo \mathcal{K} . Então a **função traço**, (a qual re-encontraremos na seção 4) que associa a cada matriz quadrada a soma dos elementos de sua diagonal principal, é um funcional linear sobre o espaço das matrizes $n \times n$ sobre \mathcal{K} .

Definição 2.2. Um operador linear (em particular, um funcional linear f) $T \in \mathcal{L}(\mathcal{V})$ é **contínuo em um ponto**¹⁰ $\alpha \in \mathcal{V}$ se para cada $\varepsilon > 0$ existe um $\delta > 0$ (que depende do ε) tal que $|T(\beta) - T(\alpha)| < \varepsilon$ sempre que $|\beta - \alpha| < \delta$. O operador é **contínuo** se for contínuo em todos os pontos (vetores) de \mathcal{V} . O operador é **limitado** se para cada natural $n > 0$ há uma constante C tal que $\|T(\alpha)\| \leq C\|\alpha\|$ para todos $\alpha \in \mathcal{V}$.

¹⁰'Ponto' aqui é sinônimo de 'vetor'. Trata-se de uma terminologia que vem de outras partes da matemática, que manteremos.

Teorema 2.1. (i) Se um operador (funcional) linear é contínuo em um ponto, é contínuo em todos. (ii) Um operador (funcional) linear é contínuo se e somente se é limitado.¹¹

Um caso de particular interesse é o seguinte funcional linear, suposto contínuo sobre um espaço de dimensão finita. Para cada vetor α de \mathcal{V} , associamos o escalar $\langle \alpha | \beta \rangle$, para β um vetor fixo de \mathcal{V} . Ou seja, temos

$$f_\beta(\alpha) = \langle \beta | \alpha \rangle.$$

Exemplo 2.3. Mostre que a aplicação recém definida é de fato linear.

O teorema a seguir consiste na recíproca do fato anterior, que vale se o espaço for de dimensão finita.

Teorema 2.2 (Teorema de representação de Riesz). Para todo funcional linear f sobre um espaço vetorial \mathcal{V} de dimensão finita, existe um único vetor β tal que, para todo α , tem-se

$$f(\alpha) = \langle \beta | \alpha \rangle.$$

Demonstração. Suponha que $\mathcal{A} = \{\alpha_1, \dots, \alpha_n\}$ é uma base ortonormal para o espaço em questão. Definimos $\beta = \sum_i c_i \alpha_i$ como candidato a ser o tal vetor do qual fala o teorema. Então devemos ter

$$f(\alpha_j) = \langle \beta | \alpha_j \rangle = \left\langle \sum_i c_i \alpha_i | \alpha_j \right\rangle = \sum_i c_i^* \langle \alpha_i | \alpha_j \rangle = c_j^*$$

porque a base é ortonormal. Portanto,

$$\beta = \sum_i f^*(\alpha_i) \cdot \alpha_i \tag{3.12}$$

Provaremos agora que o tal vetor β realmente faz o que dele se espera. Defina-se o funcional f_β do seguinte modo: $f_\beta(\alpha) = \langle \beta | \alpha \rangle$. Então temos (levando em conta a equação 3.12)

$$f_\beta(\alpha_j) = \langle \beta | \alpha_j \rangle = \left\langle \sum_i f^*(\alpha_i) | \alpha_j \right\rangle = f(\alpha_j).$$

¹¹A demonstração deste fato, para o caso geral, faz uso do Axioma da Escolha.

Quanto à unicidade, suponha que γ é tal que $\langle \beta | \alpha \rangle = \langle \gamma | \alpha \rangle$ (chame esta igualdade de (1)), para todo α . Então, pondo $\alpha = \gamma - \beta$, temos

$$\langle \gamma - \beta | \gamma - \beta \rangle = (\langle \gamma | \gamma \rangle - \langle \beta | \gamma \rangle) + (\langle \beta | \beta \rangle - \langle \gamma | \beta \rangle) = 0.$$

Com efeito, como a igualdade (1) deve valer para qualquer α , valerá em particular para β e γ , portanto as diferenças entre os parênteses dão zero. \square

O teorema de Riesz é uma das razões pelas quais a notação de Dirac é relevante. Note bem o significado deste teorema. O vetor β de certo modo caracteriza o funcional linear f (o espaço tem que ser de dimensão finita); para cada α , a imagem de α por f é determinada a partir de β efetuando-se o produto interno $\langle \beta | \alpha \rangle$ (recorde que a imagem de um vetor por um funcional é um escalar). Pelo Teorema de Riesz, que isso sempre ocorre para cada funcional T , e que tal β é único para cada f . Por isso Dirac usou a notação usando bras para expressar esse funcional, $\langle \beta |$. De maneira geral, $\langle \beta |$, $\langle \gamma |$ etc. denotam funcionais lineares T_β, T_γ , etc.. Assim, insistimos mais uma vez: a notação $\langle \beta | \alpha \rangle$ pode significar duas coisas: o produto interno dos vetores $|\beta\rangle$ e $|\alpha\rangle$ ou a imagem do vetor $|\alpha\rangle$ pelo funcional $\langle \beta |$. Genial, não?

A notação de Dirac é bem conveniente. Vejamos mais um exemplo. Suponha que desejamos conhecer a imagem do vetor $|2 - 1, i, 3\rangle$ pelo funcional $\langle 1 - i, 2 + 2i, -1 |$, considerando espaços complexos. Ora, pelo que se viu, basta acharmos

$$\langle 1 - i, 2 + 2i, -1 | 2 - i, i, 3 \rangle,$$

ou seja,

$$(1 - i)^* \cdot (2 - i) + (2 + 2i)^* \cdot (i) + (-1)^* \cdot 3,$$

e você pode facilmente prosseguir daí. Mais sobre isso abaixo.

3. O espaço dual

O conjunto de todos os funcionais lineares sobre \mathcal{V} , munido das operações de adição de funções e de multiplicação de função por escalar é um espaço vetorial sobre o mesmo corpo, denominado de **espaço dual** de \mathcal{V} , denotado \mathcal{V}^* , que tem a mesma dimensão que este espaço. Pode-se mostrar que o dual do dual de \mathcal{V} , isto é, \mathcal{V}^{**} , é igual a \mathcal{V} .

Exercício 3.1. Mostre que $\dim(\mathcal{V}^*) = \dim(\mathcal{V})$.

A solução do exercício segue imediatamente do fato de que se \mathcal{V} tem dimensão finita, então $\dim(\mathcal{L}(\mathcal{V})) = \dim(\mathcal{V})$. Você saberia demonstrar isso primeiro?

A notação matricial é útil também aqui. Tomemos um exemplo simples, sobre \mathbb{C}^2 . Sabemos pelo teorema de Riesz que cada funcional f pode ser determinado por um vetor, digamos $\beta = (4 - i, 2 - 3i)$. Assim, para $\alpha = (1 + i, 2 - i)$, $f(\alpha) = \langle \beta | \alpha \rangle = (4 - i)^*(1 + i) + (2 - 3i)^*(2 - i) = (4 + i)(1 + i) + (2 + 3i)(2 - i)$. Ora, podemos representar este produto interno da seguinte forma, usando o produto matricial (linha por coluna):

$$\langle \beta | \alpha \rangle = (4 + i, 2 + 3i) \cdot \begin{pmatrix} 1 + i \\ 2 - i \end{pmatrix}.$$

Deste modo, os bras serão algumas vezes como matrizes linha, e os vetores do espaço como vetores coluna.

Podemos então fazer a seguinte associação, bem útil em mecânica quântica. A cada ket (vetor) $|\alpha\rangle$ de \mathcal{V} , associamos um bra (funcional linear) $\langle \alpha|$ em \mathcal{V}^* , de forma que podemos pensar e, operações como na correspondente adição de bras $\langle \alpha| + \langle \beta|$ e na multiplicação $z \cdot \langle \alpha|$ como sugerindo uma análoga no espaço dos funcionais. Mas aqui é preciso cuidado. O bra correspondente *não será* $z \cdot \langle \alpha|$ devido às propriedades do complexo conjugado, mas $z^* \langle \alpha|$.

Por exemplo, se

$$|\alpha\rangle = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix},$$

então o bra correspondente seria

$$\langle \alpha| = (a_1^* a_2^* \dots a_n^*).$$

Assim, sendo $|\beta\rangle = (x_1, x_2, \dots, x_n)$ e $|\alpha\rangle = (y_1, y_2, \dots, y_n)$, podemos escrever

$$\langle \beta | \alpha \rangle = \langle x_1^*, x_2^*, \dots, x_n^* | y_1, y_2, \dots, y_n \rangle = \sum_{i=1}^n x_i^* y_i.$$

Ou seja, se os vetores são representados como colunas, os funcionais são representados como linhas, resultando

$$\langle \beta | \alpha \rangle = (x_1^* x_2^* \dots x_n^*) \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i=1}^n x_i^* y_i,$$

que é o produto interno canônico.

4. Digressão ‘quântica’: spin

Com a finalidade de motivar o estudo que vem sendo realizado, vamos dizer alguma coisa sobre a mecânica quântica de forma a utilizar alguns dos conceitos que estão sendo vistos.

No início do desenvolvimento dessa teoria, na década de 1920, havia um problema de se justificar porque o número de elétrons em um orbital completo deveria ser um certo valor e não outro que parecia mais evidente.¹² Conheciam-se três ‘números quânticos’, mas para dar uma justificativa, Wolfgang Pauli (1900-1958) postulou por volta de 1925 a existência de um quarto número a ser atribuído aos elétrons em um átomo a fim de determinar a posição (níveis de energia) dos elétrons no átomo. Este número, segundo Pauli, deveria assumir unicamente dois valores possíveis, mas não se sabia ao certo porque quatro números eram necessários. Ainda em 1925, dois jovens (Pauli também era jovem: nasceu em 1900) estudantes holandeses de pós-graduação, Samuel Goudsmit (1902-1972) e George Uhlenbeck (1900-1988) constataram que este ‘quarto número’ não deveria ser considerado como mais um número quântico, mas como uma característica (propriedade) intrínseca dos elétrons, a qual denominaram de *spin*. Como comenta Kumar [Kum.09, p.169], trata-se de um nome infeliz, porque pode ser associado ao movimento de rotação (*spinning*) dos objetos, enquanto que o conceito de spin era algo puramente teórico que resolvia alguns problemas da física de então. Os dois valores do spin que podem ser assumidos são sempre relativos a uma dada direção e, na direção dada, digamos a do eixo z de um sistema de coordenadas cartesianas ortogonais no \mathbb{R}^3 , denominam-se UP e DOWN. Como comentam Susskind e Friedman, “[i]ngenuamente, o spin pode ser representado por uma pequena flecha

¹²Para a história deste caso, ver [Kum.09, pp.167ss].

que aponta para alguma direção,¹³ mas esta analogia é demasiadamente clássica para representar acuradamente a situação real. O spin de um elétron diz respeito a quanto um sistema pode ser mecânico-quântico, e qualquer tentativa de visualizá-lo classicamente, erroneamente considerará este fato" [SusFri. 14, p. 3]. Com efeito, Pauli já havia antecipado que o seu 'novo número quântico' "não pode ser descrito de um ponto de vista clássico" (cf. [Kum.09, p. 171]).

Utilizaremos o nosso 'formalismo' para representar o spin, e vamos fazer isso seguindo [SusFri. 14, cap.2], livro que recomendamos para uma iniciação à física quântica. Usaremos vetores em um espaço de Hilbert adequado, e usaremos livremente a notação bra-ket vista acima. Vamos rotular os estados do spin (UP ou DOWN) ao longo dos três eixos cartesianos ortogonais x, y e z . Denominaremos de \mathcal{A} o aparato que mede o spin de um sistema físico, e representaremos os estados UP e DOWN que podem ser assumidos pelos kets $|u\rangle$ e $|d\rangle$ respectivamente. Ademais, tais valores de spin são escritos $\sigma_z = +1$ ou $\sigma_z = -1$, sendo σ algo que representa a quantidade física sendo avaliada. Quando o aparelho é orientado na direção z e a sua medida resulta $+1$, dizemos que o estado $|u\rangle$ foi preparado para ser observado; isso tem a ver com o modo pelo qual os físicos se expressam. Não importa muito aqui.

Para facilitar o entendimento, vamos chamar os up e down de uma outra direção x de RIGHT e LEFT respectivamente; assim, se o aparato for orientado na direção x e a medida registra o valor $\sigma_x = -1$, dizemos que o estado $|l\rangle$ é que foi preparado. Analogamente, ao longo do eixo y , vamos assumir que os estados que podem ser preparados são $|in\rangle$ e $|o\rangle$, para IN e OUT respectivamente. Se fossemos desenhar três eixos ortogonais, o eixo y é que estaria 'saindo ou entrando' na folha de papel. Como dizem Susskind e Friedman, "[y]ou get the idea" [SusFri. 14, p. 37].

Supõe-se ademais que não há mais nada para ser conhecido a respeito do spin nessas direções, ou seja, que não há outras 'variáveis ocultas' (*hidden variables*), assim que o espaço dos estados será bi-dimensional, e o que importa é que as bases escolhidas devem ser ortonormais.

Seja $|\alpha\rangle$ um estado (vetor) arbitrário, que pode representar qualquer estado de spin, preparado da maneira como quisermos. Escolhendo $\mathcal{Z} = \{|u\rangle, |d\rangle\}$ como base, e constando que ela é ortonormal, podemos escrever, lembrando

¹³[De fato, nos livros de física é comum utilizar-se da notação $|\uparrow\rangle$ e $|\downarrow\rangle$ para representá-los.]

que o espaço é o espaço complexo \mathbb{C}^2

$$|\alpha\rangle = a_u|u\rangle + a_d|d\rangle, \quad (3.13)$$

e sabemos que (coeficientes de Fourier) $a_u = \langle u|\alpha\rangle$ e $a_d = \langle d|\alpha\rangle$. Importante observar que as quantidades $a_u^* \cdot a_u = |a_u|^2$ e $a_d^* \cdot a_d = |a_d|^2$ têm a seguinte interpretação no formalismo quântico (para constatar este fato, há que se conhecer os axiomas da mecânica quântica não-relativista), supondo-se que o sistema foi preparado no estado $|\alpha\rangle$ e o aparato foi orientado ao longo do eixo z , a quantidade (veja que é um número real) $a_u^* \cdot a_u = |a_u|^2$ é a **probabilidade** de que, quando medirmos o spin, encontremos o valor $\sigma_z = +1$, isto é, seja UP ao longo da direção z . O mesmo se dá com $a_d^* \cdot a_d$, que é a probabilidade de que obtenha-se DOWN ao longo de z .

Perceba algo fundamental: *nada é dito sobre o valor do spin na direção z antes da medição*. Isso é característico da interpretação dominante da mecânica quântica (dita Interpretação de Copenhague) e de suas variantes. É somente *após* a medição que podemos saber o valor do spin; antes, o vetor $|\alpha\rangle$ pode nos fornecer unicamente *potencialidades* (termo que Heisenberg gostava de empregar) desses valores, mas nunca valores atuais.

O fato da base \mathcal{Z} ser ortogonal (todo conjunto ortonormal de vetores é ortogonal) nos diz que *estados (representados por vetores) ortogonais são fisicamente distintos e mutuamente exclusivos*. Se o spin (ou qualquer outra quantidade) está em um desses estados (digamos, UP), *não pode* estar no outro estado: a probabilidade disso acontecer é zero! [SusFri.14, p.40]. Uma observação importante feita por esses dois autores é a seguinte (id.ibid.): as direções UP e DOWN *não são* direções ortogonais no espaço.

Todas essas hipóteses nos conduzem ao fato de que

$$a_u^* \cdot a_u + a_d^* \cdot a_d = |a_u|^2 + |a_d|^2 = 1,$$

o que é equivalente a dizer que $|\alpha\rangle$ (da expressão 3.13) tem norma igual a 1 (é unitário), vindo então em dos postulados da mecânica quântica não-relativista, como frequentemente formulado:

O estado de um sistema quântico é representado por um vetor unitário em um adequado espaço de Hilbert, que é o espaço de seus estados – lembre que múltiplos escalares desses vetores representam o mesmo

estado, assim que os estados são na verdade dados por rays – e os quadrados dos coeficientes de qualquer combinação linear de um desses vetores em uma base ortonormal fornecem as probabilidades de obtenção dos resultados possíveis da medida dos observáveis.

Mais tarde veremos quais podem ser esses valores possíveis.¹⁴

O que acontece com o spin ao longo dos eixos x e y ? Vamos nos fixar primeiramente no eixo x . Convencionamos antes que os vetores que representam os estados de spin ao longo do eixo x são $|l\rangle$ e $|r\rangle$. Tomando um deles, e considerando a base \mathcal{Z} acima, podemos escrever (sendo agora $|u\rangle$ o nosso vetor $|\alpha\rangle$ de antes)

$$|r\rangle = a_u|u\rangle + a_d|d\rangle,$$

já que podemos escrever qualquer vetor do espaço como combinação linear dos vetores da base. Por outro lado, sabemos que este vetor deve estar normalizado, ou seja, ser unitário. Uma combinação linear que cumpre esta condição preservando iguais probabilidades é

$$|r\rangle = \frac{1}{\sqrt{2}}|u\rangle + \frac{1}{\sqrt{2}}|d\rangle. \quad (3.14)$$

A combinação linear que expressa $|l\rangle$ deve ser ortogonal a este vetor. Trabalhando um pouco com a matemática envolvida, chegamos a

$$|l\rangle = \frac{1}{\sqrt{2}}|u\rangle - \frac{1}{\sqrt{2}}|d\rangle. \quad (3.15)$$

Exercício 4.1. Prove que esses dois vetores, $|r\rangle$ e $|l\rangle$, são ortonormais.

Ao longo do eixo y , procedemos da mesma forma, obtendo os vetores¹⁵

$$|in\rangle = \frac{1}{\sqrt{2}}|u\rangle + \frac{i}{\sqrt{2}}|d\rangle. \quad (3.16)$$

e

$$|o\rangle = \frac{1}{\sqrt{2}}|u\rangle - \frac{i}{\sqrt{2}}|d\rangle. \quad (3.17)$$

¹⁴Serão os *auto-valores* do operador hermitiano que representa o observável em questão.

¹⁵Aqui o leitor saberá porque escolhemos $|in\rangle$ (com duas letras em vez de uma só, como nos outros) para representar IN; para não confundir com o número complexo i .

Exercício 4.2. Prove que esses dois vetores, $|in\rangle$ e $|o\rangle$, são ortonormais.

Poderíamos ter utilizado a notação matricial, representando os vetores como matrizes coluna complexas 2×1 ; assim, teríamos

$$|u\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, |d\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, |l\rangle = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}, |r\rangle = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix},$$

$$|in\rangle = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} \end{pmatrix}, |o\rangle = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} \end{pmatrix},$$

Esta pequena digressão ainda mostra a necessidade de se utilizar números complexos para representar o spin. Não há, dentro deste formalismo, modo alternativo. Mais à frente, falaremos dos operadores que representam os componentes de spin, mas para isso necessitamos dos conceitos introduzidos a seguir.

5. Auto-vetores e auto-valores

Deixemos novamente de lado a notação de Dirac por mais um momento.

Definição 5.1 (auto-vetores e auto-valores). Seja T um operador linear sobre o espaço vetorial \mathcal{V} (sobre o corpo \mathcal{K} , que como sabemos, no nosso caso é ou o corpo dos reais ou o dos complexos). Um **autovetor** de T é um vetor não nulo $\xi \in \mathcal{V}$ tal que $T(\xi) = \lambda\xi$, para $\lambda \in K$. O escalar λ é dito **autovalor** associado ao autovetor ξ .

Definição 5.2 (auto-vetores e auto-valores). Seja A matrix de ordem n sobre K . Um **autovetor** de A é uma matriz $1 \times n$ X tal que $AX^T = \lambda X^T$, para $\lambda \in K$. O escalar λ é dito **autovalor** associado ao autovetor X .

Definição 5.3 (Espectro). Denomina-se de **espectro** de um operador ou de uma matriz ao conjunto de seus auto-valores.

Exemplo 5.1. Seja D operador linear sobre o espaço vetorial das funções reais de variável real deriváveis no intervalo $[a, b]$ da reta, definido por

$$D(f(x)) = \frac{df(x)}{dx}.$$

Então $f(x) = e^{3x}$ é autovetor de D , cujo autovalor associado é $\lambda = 3$.

Exemplo 5.2. Seja T operador linear sobre o \mathbb{R}^2 definido por $T(x, y) = (2x - y, y)$. Impondo $T(x, y) = \lambda(x, y)$, obtemos o sistema de equações

$$\begin{cases} (2 - \lambda)x - y = 0 \\ (1 - \lambda)y = 0 \end{cases}$$

A matriz dos coeficientes do sistema é

$$A = \begin{pmatrix} 2 - \lambda & -1 \\ 0 & 1 - \lambda \end{pmatrix}$$

que deve ter determinante não nulo para que haja solução não trivial. Portanto, os auto-valores são as raízes do **polinômio característico** $\lambda^2 - 3\lambda + 2 = 0$, que são $\lambda_1 = 1$ e $\lambda_2 = 2$. Os auto-vetores associados são obtidos resolvendo-se o sistema acima para cada valor de λ , o que fornece dois conjuntos de vetores, a saber: $E_{\lambda_1} = \{\gamma \in \mathcal{V} : T(\gamma) = \lambda_1 \gamma\} = \{(x, y) \in \mathbb{R}^2 : x = y\}$ e $E_{\lambda_2} = \{\gamma \in \mathcal{V} : T(\gamma) = \lambda_2 \gamma\} = \{(x, y) \in \mathbb{R}^2 : y = 0\}$.

É fácil ver que o conjunto $E_\lambda \cup \{\mathbf{0}\}$ é um sub-espaço vetorial de \mathcal{V} , dito *espaço característico* associado ao autovalor λ , onde $\mathbf{0}$ é o vetor nulo de \mathcal{V} .

Seja A matriz quadrada de ordem n sobre o corpo \mathcal{K} . Impondo $AX^T = \lambda X^T$, para $X = [x, y]$, obtemos $AX^T - \lambda X^T = 0$, ou $(A - \lambda I)X^T = 0$, sendo I a matriz identidade de ordem n e 0 a matriz nula $n \times 1$. A matriz $A - \lambda I$ deve ser **singular** (não inversível) para que o sistema admita solução não trivial, portanto, $\det(A - \lambda I) = 0$, que é o **polinômio característico** de A . Suas raízes são os auto-valores de A , e os auto-vetores associados são obtidos resolvendo-se os sistemas $(A - \lambda I)X^T = 0$ para cada λ encontrado.

Exemplo 5.3. Seja

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

matrix real. Então $\det(A - \lambda I) = 0$ fornece

$$\det(A - \lambda I) = \begin{vmatrix} 1 - \lambda & 1 \\ 0 & 1 - \lambda \end{vmatrix} = (1 - \lambda)^2 = 0,$$

o que fornece $\lambda_1 = \lambda_2 = 1$. Resolvendo-se o sistema $AX^T = 1 \cdot X^T$, encontramos $y = 0$, portanto $E_1 = \{(x, y) \in \mathbb{R}^2 : y = 0\}$.

Pode-se demonstrar que se T é operador linear sobre \mathcal{V} e $A = [T]_{\mathcal{A}}$ é a matriz representativa de T na base ordenada \mathcal{A} , então ξ é autovetor de T se e somente se $X = [\xi]_{\mathcal{A}}$, a matriz das coordenadas de ξ na base \mathcal{A} , é autovetor de A . Os auto-valores são os mesmos.

Definição 5.4. Duas matrizes A e B são **semelhantes** e existe uma matriz inversível M tal que $B = M^{-1}AM$.

Teorema 5.1. *Matrizes semelhantes têm o mesmo polinômio característico, logo os mesmos auto-valores.*

Demonstração. $\det(B - \lambda I) = \det(M^{-1}AM - \lambda I) = \det(M^{-1}(A - \lambda I)M) = \det(A - \lambda I)$. □

Exercício 5.1. Mostre que o conjunto $E_{\lambda} \cup \{0\}$ é um subspaço vetorial de \mathcal{V} .

Definição 5.5. Um operador linear T sobre $\mathcal{E} = \langle \mathcal{V}, \mathcal{K}, +, \cdot \rangle$ de dimensão n (com ou sem produto interno) é **maximal**, ou **não-degenerado**, se possui n auto-vetores distintos. Caso contrário, é **degenerado**. Diz-se o mesmo de uma matriz de ordem $n \times n$ com elementos em \mathcal{K} .

No formalismo da mecânica quântica, teremos que levar em conta se os operadores têm ou não auto-valores repetidos. Guarde isso.

6. O exemplo das matrizes de Pauli

No formalismo da mecânica quântica, desempenham papel importante as equações da forma

$$T|\alpha\rangle = \lambda|\alpha\rangle,$$

que têm a seguinte interpretação. Se T representa um observável físico, medido em relação a um sistema que esteja em um estado representado por $|\alpha\rangle$, o valor λ é interpretado como um **valor possível** para a medida do observável (para o sistema no dado estado).

Isso se conforma com o que Michael Redhead chama de **algoritmo da quantização**, a saber, a questão de responder à pergunta: “Quais são os valores possíveis da medida de um observável relativamente a um sistema físico?”, que ele responde da seguinte forma: “São os auto-valores do operador associado ao observável a ser medido.” [Red.87, pp.5-7].

6.1. Matrizes de Pauli

Para exemplificar, vamos falar dos operadores que representam o spin, mais uma vez para vermos uma aplicação dos conceitos deste capítulo e dos precedentes. O objetivo é construir operadores que representem os componentes de spin σ_z , σ_x e σ_y , e mais uma vez seguiremos [SusFri.14, pp.75ss].

Consideremos a direção z . Sabemos que $\sigma_z = \pm 1$ (UP ou DOWN) e que $\mathcal{Z} = \{|u\rangle, |d\rangle\}$ é uma base ortonormal para os estados de spin medidos ao longo da direção considerada. Os vetores desta base devem ser auto-vetores normalizados do operador procurado, os quais têm $+1$ e -1 como auto-valores. ou seja, devemos ter

$$\sigma_z|u\rangle = +1\cdot|u\rangle, \sigma_z|d\rangle = -1\cdot|d\rangle.$$

Levando em conta o que sabemos sobre a representação matricial de operadores, o que queremos é um operador σ_z cuja matriz representativa na base

$$\mathcal{Z} = \{|u\rangle, |d\rangle\} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\},$$

ou seja, tal que

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = +1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -1 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

O que indica que os auto-vetores $\alpha_1 = (1, 0)$ e $\alpha_2 = (0, 1)$ têm auto-valores respectivamente iguais a $+1$ e -1 .

Como exercício, o leitor pode calcular esses valores, chegando à seguinte matriz:

$$\sigma_z = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Vamos agora determinar σ_x . Para tanto, lembremos que

$$|r\rangle = \frac{1}{\sqrt{2}}|u\rangle + \frac{1}{\sqrt{2}}|d\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Da mesma forma,

$$|l\rangle = \frac{1}{\sqrt{2}}|u\rangle - \frac{1}{\sqrt{2}}|d\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

Como queremos $\sigma_x|r\rangle = +1\cdot|r\rangle$ e $\sigma_x|l\rangle = -1\cdot|l\rangle$, matricialmente temos

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix},$$

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} = -\begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix},$$

Novas contas fornecem

$$\sigma_x = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Procedimento análogo é feito com relação a σ_y , obtendo-se sucessivamente

$$|in\rangle = \frac{1}{\sqrt{2}}|u\rangle + \frac{i}{\sqrt{2}}|d\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{i}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} \end{pmatrix}.$$

Da mesma forma,

$$|o\rangle = \frac{1}{\sqrt{2}}|u\rangle - \frac{i}{\sqrt{2}}|d\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \frac{i}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} \end{pmatrix}.$$

Como queremos $\sigma_y|in\rangle = +1\cdot|in\rangle$ e $\sigma_y|o\rangle = -1\cdot|o\rangle$, matricialmente temos

$$\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} \end{pmatrix},$$

$$\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} \end{pmatrix} = -\begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} \end{pmatrix},$$

Novas contas fornecem

$$\sigma_y = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

Essas três matrizes são conhecidas como *matrizes de Pauli*, ou seja,

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

O site *Pauli Matrices* na *Wikipedia* está muito bem escrito e contém uma boa quantidade de informações sobre elas. Na verdade, estamos simplificando as coisas já que nosso objetivo é mais a Álgebra Linear do que a mecânica quântica. As matrizes de Pauli são obtidas das acima multiplicando-as pelo fator $\hbar/2$. Assim, temos as ‘verdadeiras’ matrizes de Pauli:

$$S_x = \frac{\hbar}{2}\sigma_x, S_y = \frac{\hbar}{2}\sigma_y, S_z = \frac{\hbar}{2}\sigma_z.$$

Exercício 6.1. Explícite auto-vetores e auto-valores de cada uma das matrizes de Pauli.

Exercício 6.2. Comprove cada uma das igualdades seguintes: (1) $\det\sigma_i = -1$ ($i = 1, 2, 3$); (2) $\text{Tr}(\sigma_i) = 0$; (3) $\sigma_x = \sigma_y^2 = \sigma_z^2 = I_2$ (matriz identidade de ordem 2); (4) $[\sigma_x, \sigma_y] = 2i\sigma_z$; (5) $[\sigma_y, \sigma_z] = 2i\sigma_x$; (6) $[\sigma_z, \sigma_x] = 2i\sigma_y$.

7. Diagonalização

Definição 7.1 (Matrizes semelhantes). A matriz quadrada A de ordem n é **semelhante** a uma matriz quadrada B de ordem n , ambas com elementos no mesmo corpo, se existe uma matriz inversível M também com elementos no corpo, tal que

$$A = M^{-1}BM.$$

O seguinte teorema explica um importante resultado sobre essas matrizes

Teorema 7.1. *Duas matrizes A e B de ordem n sobre o corpo \mathcal{K} são semelhantes se e somente se existirem duas bases ordenadas para o \mathcal{K}^n em relação às quais A e B representam o mesmo operador linear sobre \mathcal{K}^n .*

Dito de outra forma, matrizes semelhantes podem ser vistas como matrizes *do mesmo* operador linear em bases distintas. Isto tem importância, porque, dado um operador linear T , estaremos interessados em encontrar uma base relativamente à qual a matriz de T seja diagonal. No caso mais específico que interessa à mecânica quântica, esta base será formada por auto-vetores ortogonais de T .

Importante é o seguinte teorema, aqui somente enunciado:

Teorema 7.2 (Teorema dos Eixos Principais). *Toda matriz hermitiana (ou somente simétrica no caso real) A é unitariamente (no caso real, ortogonalmente) semelhante a uma matriz diagonal real D , cujos elementos diagonais são os autovalores de A . Ou seja, existe M unitária (ortogonal) tal que $D = M^{-1}AM = M^*AM$ (respect., $D = M^TAM$).*

Vem então a seguinte definição:

Definição 7.2 (Operador diagonalizável). Um operador linear T sobre \mathcal{V} é **diagonalizável** se existe uma base de \mathcal{V} na qual a matriz representativa de T seja uma matriz diagonal.

Definição 7.3 (Matriz diagonalizável). Uma matriz A é **diagonalizável** se é semelhante a uma matriz diagonal.

É fácil ver que se $\mathcal{A} = \{\alpha_1, \dots, \alpha_n\}$ é uma base ordenada de \mathcal{V} formada por auto-vetores de T , então teremos $T(\alpha_i) = \lambda_i \alpha_i$, e portanto a matriz de T nessa base será a matriz diagonal

$$A = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

Teorema 7.3. *Uma matriz A de ordem n sobre o corpo K é diagonalizável se e somente se seus auto-vetores gerarem K^n .*

Demonstração. Primeiramente, recorde que $K^n = \{(x_1, \dots, x_n) : x_j \in K\}$ pode ser visto como um espaço vetorial sobre K . Quanto à demonstração, suponha inicialmente que A é diagonalizável, logo semelhante a muma matriz diagonal D que podemos supor tem em sua diagonal principal os escalares d_1, \dots, d_n . Então os vetores $\epsilon_i = (0, \dots, 1, \dots, 0)$ são auto-vetores de D , pois $D\epsilon_i^T = d_i\epsilon_i^T$. Isso ainda mostra que os d_i são auto-valores de D , logo de A . Ademais, se houvesse outros auto-valores λ de D então haveria vetores $X = (x_1, \dots, x_n) \neq 0$ tais que $DX^T = \lambda X^T$, o que daria $(d_1 x_1, \dots, d_n x_n) = (\lambda x_1, \dots, \lambda x_n)$, ou seja, $d_i x_i = \lambda_i x_i$. Mas como os auto-vetores são não nulos, devemos ter pelo menos um $x_i \neq 0$, o que dará $\lambda = d_i$, e assim não há outros auto-valores além dos indicados. Reciprocamente, se os vetores característicos de A gerarem K^n , pode-se extrair

uma base X_1, \dots, X_n para K^n , ou seja, tais que $AX_i^T = \lambda_i X_i^T$, sendo os λ_i os auto-valores. Seja M matriz cujas colunas são os vetores X_i^T , que é inversível pela hipótese de que esses vetores formam uma base. Pondo

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix},$$

temos $AM = A(X_1^T \dots X_n^T) = (\lambda_1 X_1^T \dots \lambda_n X_n^T) = MD$, ou seja, $M^{-1}AM = D$. \square

A demonstração deste teorema nos dá o modo de achar a matriz M que efetua a diagonalização de A , desde que A seja diagonalizável: as colunas de M são formadas por auto-vetores linearmente independentes de A .

Exercício 7.1. Verifique que os auto-valores de A abaixo são 1 e 2 (duplo). Ache os auto-valores correspondentes. Verifique se a matriz é diagonalizável e, em caso afirmativo, ache a matriz diagonal D semelhante a A :

$$A = \begin{pmatrix} 3 & -1 & 1 \\ 2 & 0 & 2 \\ 2 & -1 & 2 \end{pmatrix}$$

Exercício 7.2. Idem para a matriz

$$A = \begin{pmatrix} 5 & -1 & 3 \\ -6 & 4 & -6 \\ -6 & 2 & 4 \end{pmatrix}$$

O teorema (7.3) acima nos diz algo importante. Mesmo que o operador (ou a matriz) seja, degenerado, ou seja, que seus auto-valores não sejam todos distintos, caso ele seja diagonalizável, é possível encontrar uma base para o espaço em questão constituída de auto-vetores do operador (ou da matriz). Isso tem relevância para algumas expressões em física. Por exemplo, admita (usando a notação de Dirac) que q_1, \dots, q_n sejam auto-valores de um operador T (falaremos somente de operadores) associados aos auto-vetores $|q_1\rangle, \dots, |q_n\rangle$. Pode ser que nem todos os q_i sejam distintos. Se quisermos somar somente aqueles auto-valores que são iguais a q_i , escrevemos

$$\sum_{j|q_j=q_i} q_j, \tag{3.18}$$

podendo ainda adicionar (se bem que a notação fique carregada) que j varia de 1 a n .

Evidentemente, se todos os auto-valores forem distintos (T é não-degenerado, ou maximal), esta soma terá somente um termo. Esta notação será importante para a formulação *algoritmo estatístico* da mecânica quântica [Red.87, p.8].

8. Matrizes e operadores ortogonais e unitários

Nesta seção, suporemos que \mathcal{V} é um espaço com produto interno.

Definição 8.1 (Isometria). Uma **isometria** sobre \mathcal{V} é um operador linear T sobre \mathcal{V} tal que $\|T(\alpha)\| = \|\alpha\|$ para todo $\alpha \in \mathcal{V}$.

Intuitivamente, uma isometria é uma aplicação que preserva comprimentos (iso-metria), como fica claro pelo teorema seguinte:

Teorema 8.1. Se T é uma isometria sobre \mathcal{V} , então, para todos $\xi, \eta \in \mathcal{V}$:

1. T preserva distâncias, ou seja, $\|\xi - \eta\| = \|T(\xi) - T(\eta)\|$.
2. T preserva produtos internos, ou seja, $\langle \xi | \eta \rangle = \langle T(\xi) | T(\eta) \rangle$.
3. T preserva conjuntos ortonormais.
4. T preserva medidas angulares.

Demonstração. Imediata, tendo em vista a definição e as propriedades da norma. \square

Exercício 8.1. Prove o teorema anterior.

Definição 8.2 (Operadores ortogonal e unitário). Uma isometria sobre um espaço vetorial complexo é chamado de **operador unitário**. Uma isometria sobre um espaço vetorial real é chamado de **operador ortogonal**.

Exercício 8.2. (a) Mostre que $T(x, y, z) = (x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta)$ é uma isometria sobre o \mathbb{R}^2 (operador ortogonal); (b) Idem para $T(x, y) = (x, -y)$ (reflexão em relação ao eixo Y).

Repare que a matriz do primeiro operador na base canônica, que aparece em muitos textos, é

$$[T] = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}.$$

Definição 8.3 (Matriz ortogonal, matriz unitária). Uma matriz real (complexa) de ordem n $A = [a_{ij}]$ é **ortogonal** (respec. **unitária**) se $A^T A = I$ (respec. $A^* A = I$).

Exercício 8.3. Mostre que as matrizes de Pauli são unitárias.

A definição implica que, no primeiro caso, $\sum_{k=1}^n a_{ki} a_{kj} = \delta_{ij}$, e no segundo caso, $\sum_{k=1}^n a_{ki}^* a_{kj} = \delta_{ij}$, onde a^* é o conjugado de a . Ou seja, as colunas de A são vetores ortonormais relativamente ao produto interno canônico (de cada espaço). Pode-se ainda demonstrar que uma matriz $n \times n$ A real (complexa) representa um operador ortogonal (unitário) T relativamente a uma base ortonormal se e somente se for ortogonal (unitária).

Outros fatos importantes são os seguintes, aqui somente enunciados:

1. Se A for ortogonal ou unitária, $\det(A) = \pm 1$
2. $A^{-1} = A^T$ (respect., $A^{-1} = A^*$, onde A^* é a transposta conjugada de A)
3. As linhas de A são vetores ortonormais relativamente ao produto interno canônico.

Exercício 8.4. Ache a, b, c, d para que A seja unitária e mostre que B é ortogonal:

$$A = \begin{pmatrix} 1+i & -1-i \\ a+bi & c+di \end{pmatrix}$$

$$B = \begin{pmatrix} 1/3 & 2/3 & 2/3 \\ 2/3 & -2/3 & 1/3 \\ -2/3 & -1/3 & 2/3 \end{pmatrix}$$

Definição 8.4 (Adjunto de um operador). Seja T operador sobre \mathcal{V} . O adjunto de T , quando existe, é um operador T^\dagger tal que, para todos $\alpha, \beta \in \mathcal{V}$, satisfaz

$$\langle \alpha | T^\dagger \beta \rangle = \langle T \alpha | \beta \rangle.$$

Se o espaço vetorial (obviamente, com produto interno) for de dimensão finita (digamos, n), resulta que

Teorema 8.2. *Todo operador linear T sobre um espaço de dimensão finita admite um adjunto T^\dagger .*

Demonstração. Ver [HofKun.79, pp.374-5].

□

Exercício 8.5. Mostre que o adjunto do operador $T(x, y) = (x - y, x + y)$ sobre o \mathbb{R}^2 é $T(x, y) = (x + y, -x + y)$. Ache as matrizes desses dois operadores na base canônica e verifique que elas são as transpostas uma da outra.

No caso de dimensão infinita, nem sempre existe o adjunto de um operador, mas quando o adjunto existe, é único. Pode-se ainda mostrar (ibid., p. 375-6) que relativamente a qualquer base ortonormal, a matriz representativa do operador é a transposta conjugada da matriz representativa de seu adjunto (na mesma base). Ademais, quando o operador não é limitado, por vezes é complicado definir o seu adjunto já que, não sendo definido em todo o espaço, deve-se levar em conta os seu domínio de definição, mas não exploraremos isso aqui. Exemplos de operadores não limitados são os que representam os observáveis posição e momento de uma partícula.

Teorema 8.3. *Valem as seguintes propriedades, em sentido óbvio:*

1. $(T^\dagger)^\dagger = T$
2. $(T + U)^\dagger = T^\dagger + U^\dagger$
3. $(TU)^\dagger = U^\dagger T^\dagger$
4. $(cT)^\dagger = c^* T^\dagger$, sendo c^* o conjugado de c .

Demonstração. É um bom exercício. □

Operadores não-limitados são importantes em mecânica quântica. Os observáveis posição e momento de uma partícula são não-limitados, mas não lidaremos com eles exceto em poucos exemplos. Lembramos que, em espaços de dimensão finita, todos os operadores são limitados. Ademais, relativamente a operadores não-limitados, distingue-se entre operadores auto-adjuntos e operadores hermitianos, conforme abaixo, mas isso não desempenhará um papel relevante aqui, uma vez que não desenvolveremos a mecânica quântica em detalhes. Assim, o leitor pode tomar os dois conceitos como expressando a mesma ideia.

Definição 8.5. Um operador é **densamente definido** se para todo número real $\varepsilon > 0$ e para qualquer vetor β do espaço, existe sempre um vetor α no domínio do operador tal que $\|\beta - \alpha\| < \varepsilon$, ou seja, um vetor no domínio arbitrariamente próximo do vetor dado).

Definição 8.6 (Operadores auto-adjuntos ou hermitianos). Se T é densamente definido e se $T = T^\dagger$, então T é dito **auto-adjunto**. Um operador é **hermitiano** se for auto-adjunto e limitado (cf. definição 2.2)

Conseqüentemente, devido à definição de adjunto de um operador, tem-se que T é auto-adjunto (hermitiano) se e somente se

$$\langle \alpha | T \beta \rangle = \langle T \alpha | \beta \rangle.$$

No caso de dimensão finita, os dois conceitos (auto-adjunto e hermitiano) se confundem. Se o operador é auto-adjunto, seu domínio é todo o espaço.

Os físicos escrevem T^* para T^\dagger em virtude de que a operação de tomar o adjunto tem semelhanças com a de tomar o conjugado de um número complexo, e do seguinte fato: a matriz de T auto-adjunto relativamente a uma base ortonormal é hermitiana. Suas colunas são vetores ortonormais relativamente ao produto interno canônico.

Algumas propriedades dos operadores auto-adjuntos e hermitianos

Vamos enfatizar que há várias razões para nos interessarmos pelos operadores auto-adjuntos ou hermitianos. Dentre elas, as seguintes (a despeito da redundância, achamos por bem enunciar várias das propriedades):

1. Esses operadores são sempre diagonalizáveis, logo, existe sempre uma base para o espaço formada por auto-vetores do operador.
2. Logo, em particular, existe uma base ortonormal para o espaço formada por auto-vetores do operador.
3. Todos os auto-valores de um operador auto-adjunto são reais, e isso tem importância capital para o formalismo da mecânica quântica. A demonstração deste fato é simples: se T é auto-adjunto, então $\langle \alpha | T \beta \rangle = \langle T \alpha | \beta \rangle$. Se por outro lado γ é autovetor de T , ou seja, existe um escalar λ tal que (escrevendo sem os parênteses) $T\gamma = \lambda\gamma$, então $\lambda \|\gamma\|^2 = \lambda \langle \gamma | \gamma \rangle = \langle \gamma | \lambda \gamma \rangle = \langle \gamma | T \gamma \rangle = \langle T \gamma | \gamma \rangle = \langle \lambda \gamma | \gamma \rangle = \lambda^* \langle \gamma | \gamma \rangle = \lambda^* \|\gamma\|^2$. Conseqüentemente, $\lambda = \lambda^*$, o que ocorre se e somente se λ for real.
4. Auto-vetores associados a auto-valores distintos são ortogonais.

5. Esses operadores representarão os observáveis físicos no formalismo da mecânica quântica, e seus auto-valores serão os valores possíveis da medição de um observável.

Qual a relação entre operadores auto-adjuntos e unitários? A resposta é o seguinte

Teorema 8.4. *Um operador linear T é auto-adjunto se e somente se o auto-adjunto existe (sempre existe no caso finito) e*

$$TT^\dagger = T^\dagger T = I.$$

Ou seja, o inverso de um operador auto-adjunto é o seu adjunto (ele mesmo portanto). Ademais, se T é unitário, é um automorfismo do espaço vetorial (um isomorfismo do espaço sobre si mesmo).

SOMAS, SOMAS DIRETAS E PROJEÇÕES

CONCEITOS IMPORTANTES e fundamentais para o formalismo usual da mecânica quântica são apresentadas neste capítulo. Em particular, os operadores de projeção e a soma direta de espaços, de vetores e de operadores devem ser conhecidos do estudante.

1. Soma de sub-espaços

Definição 1.1 (Soma de sub-espaços). Sejam $\mathcal{W}_1, \dots, \mathcal{W}_k$ sub-espaços vetoriais de \mathcal{V} . Dizemos que \mathcal{W} é **soma** dos \mathcal{W}_i , e escrevemos

$$\mathcal{W} = \mathcal{W}_1 + \dots + \mathcal{W}_k,$$

se \mathcal{W} é o sub-espaço (como se provará abaixo) gerado pela união dos \mathcal{W}_i .

É imediato que um vetor α pertence a \mathcal{W} se e somente se pode ser escrito na forma $\alpha = \alpha_1 + \dots + \alpha_k$, com $\alpha_i \in \mathcal{W}_i$ ainda que esta decomposição possa não ser única. O caso da unicidade é distintivo, e será comentado abaixo.

Exemplo 1.1. (Usaremos aqui o exemplo 3.4) Seja \mathcal{V} o espaço \mathbb{R}^3 , e sejam $\mathcal{W}_1 = \{(x, 0, 0) : x \in \mathbb{R}\}$, $\mathcal{W}_2 = \{(x, y, 0) : x, y \in \mathbb{R}\}$, e $\mathcal{W}_3 = \{(0, y, 0) : y \in \mathbb{R}\}$, e $\mathcal{W}_4 = \{(0, 0, z) : z \in \mathbb{R}\}$. Então $\mathbb{R}^3 = \mathcal{W}_1 + \mathcal{W}_2 + \mathcal{W}_4 = \mathcal{W}_1 + \mathcal{W}_3 + \mathcal{W}_4$, $\mathcal{W}_2 = \mathcal{W}_1 + \mathcal{W}_3$.

Exercício 1.1. Verifique os detalhes do exemplo anterior.

Teorema 1.1. *A soma de sub-espços é um sub-espço.*

Demonstração. Claro que o teorema deveria ser enunciando com maior precisão, mas acreditamos que é claro para o leitor. Faremos a demonstração para o caso de dois sub-espços, a qual pode ser estendida por indução para um número arbitrário (e finito) de sub-espços. Para o caso geral, algo a mais seria necessário, mas deixaremos esta situação de lado. Assim, sejam \mathcal{W}_1 e \mathcal{W}_2 sub-espços de \mathcal{V} . Sejam α e β vetores de $\mathcal{W}_1 + \mathcal{W}_2$. Logo, pelo que se viu acima, existem vetores α_1 e α_2 em \mathcal{W}_1 e β_1 e β_2 em \mathcal{W}_2 tais que $\alpha = \alpha_1 + \beta_1$ e $\beta = \alpha_2 + \beta_2$. Logo, $\alpha + \beta = (\alpha_1 + \beta_2) + (\alpha_2 + \beta_1)$ devido às propriedades dos vetores. Mas $\alpha_1 + \beta_2 \in \mathcal{W}_1$, pois é um sub-espço, assim como $\alpha_2 + \beta_1 \in \mathcal{W}_2$. Consequentemente, $\alpha + \beta \in \mathcal{W}_1 + \mathcal{W}_2$. Analogamente, mostramos que se $\alpha \in \mathcal{W}_1 + \mathcal{W}_2$ e se k é um escalar, então $k\alpha \in \mathcal{W}_1 + \mathcal{W}_2$ (exercício). Portanto, pelo teorema 3.1, segue o teorema. \square

Exercício 1.2. Aceite a provocação feita no início da demonstração: enuncie corretamente o teorema.

Definição 1.2 (Soma direta). Seja $\mathcal{W} = \mathcal{W}_1 + \dots + \mathcal{W}_k$ nas condições da definição precedente. Dizemos que a soma é direta, e escrevemos

$$\mathcal{W} = \mathcal{W}_1 \oplus \dots \oplus \mathcal{W}_k$$

se, para todo $2 \leq j \leq k$, tivermos

$$\mathcal{W}_j \cap (\mathcal{W}_1 + \dots + \mathcal{W}_{j-1}) = \{0\}. \quad (4.1)$$

A expressão (4.1) é fácil de entender. Para $j = 2$, ela fica $\mathcal{W}_2 \cap \mathcal{W}_1 = \{0\}$, ou seja, os dois sub-espços devem ter interseção trivial. Para $j = 3$, fica $\mathcal{W}_3 \cap (\mathcal{W}_1 + \mathcal{W}_2) = \{0\}$, ou seja, o sub-espço \mathcal{W}_3 deve interceptar apenas trivialmente o sub-espço gerado por \mathcal{W}_1 e \mathcal{W}_2 , e assim sucessivamente. Em outras palavras, na medida em que j cresce, vamos ‘saindo’ dos sub-espços que íamos obtendo nos passos anteriores.

Pode-se agora demonstrar que a soma é direta se e somente se um vetor $\alpha \in \mathcal{W}$ pode ser escrito de modo único como $\alpha = \alpha_1 + \dots + \alpha_k$, com $\alpha_i \in \mathcal{W}_i$.

Exercício 1.3. A partir do exemplo anterior, mostre que $\mathcal{W}_2 = \mathcal{W}_1 \oplus \mathcal{W}_3$, $\mathbb{R}^3 = \mathcal{W}_1 \oplus \mathcal{W}_3 \oplus \mathcal{W}_4$, mas que a soma $\mathbb{R}^3 = \mathcal{W}_1 + \mathcal{W}_2 + \mathcal{W}_4$ não é direta.

Definição 1.3 (sub-espço invariante). Seja T operador linear sobre \mathcal{V} e \mathcal{W} sub-espço de \mathcal{V} . Então \mathcal{W} é **invariante** sob T se para cada $\xi \in \mathcal{W}$, tem-se que $T(\xi) \in \mathcal{W}$.

2. Projeções

Definição 2.1 (Projeção). Uma **projeção** sobre um espaço vetorial \mathcal{V} é um operador E sobre \mathcal{V} tal que $E^2 = E$.

Considere o \mathbb{R}^2 com o produto interno canônico, e seja $\mathcal{B} = \{\alpha_1, \alpha_2\}$ base ortonormal para este espaço. Então para todo vetor α , existem escalares (números reais) x_1 e x_2 tais que $\alpha = x_1\alpha_1 + x_2\alpha_2$ (veja a Figura 4.1). Portanto, levando em conta que x_1 e x_2 são os coeficientes de Fourier da decomposição, temos

$$\begin{aligned}\langle \alpha_1 | \alpha \rangle &= \langle \alpha_1 | x_1\alpha_1 + x_2\alpha_2 \rangle = \langle \alpha_1 | x_1\alpha_1 \rangle + \langle \alpha_1 | x_2\alpha_2 \rangle = \\ &= x_1\langle \alpha_1 | \alpha_1 \rangle + x_2\langle \alpha_1 | \alpha_2 \rangle = x_1.\end{aligned}$$

Da mesma forma, $\langle \alpha_2 | \alpha \rangle = x_2$. Portanto,

$$\alpha = \langle \alpha_1 | \alpha \rangle \alpha_1 + \langle \alpha_2 | \alpha \rangle \alpha_2.$$

Consideremos agora o operador $E_{|\alpha_1\rangle}$, projeção sobre o vetor $|\alpha_1\rangle$ unitário, definido por

$$E_{|\alpha_1\rangle}|\alpha\rangle := \langle \alpha_1 | \alpha \rangle |\alpha_1\rangle, \quad (4.2)$$

É fácil ver

$$\begin{aligned}E_{|\alpha_1\rangle}^2|\alpha\rangle &= E_{|\alpha_1\rangle}(E_{|\alpha_1\rangle}|\alpha\rangle) = E_{|\alpha_1\rangle}(\langle \alpha_1 | \alpha \rangle |\alpha_1\rangle) = \langle \alpha_1 | \alpha \rangle E_{|\alpha_1\rangle}|\alpha_1\rangle = \langle \alpha_1 | \alpha \rangle \cdot |\alpha_1\rangle = \\ &= E_{|\alpha_1\rangle}|\alpha\rangle,\end{aligned}$$

ou seja, $E_{|\alpha_1\rangle}$ é de fato uma projeção. Deste modo, obtemos em particular que o vetor projeção de $|\alpha\rangle$ sobre $|\alpha_1\rangle$ (normalizado) é obtido fazendo-se

$$E_{|\alpha_1\rangle}|\alpha\rangle = \langle \alpha_1 | \alpha \rangle \cdot |\alpha_1\rangle.$$

Se α_1 não fosse unitário, teríamos que dividi-lo pela sua norma, obtendo

$$E_{|\alpha_1\rangle}|\alpha\rangle = \left\langle \frac{\alpha_1}{\|\alpha_1\|} \middle| \alpha \right\rangle \cdot \frac{|\alpha_1\rangle}{\|\alpha_1\|} = \frac{\langle \alpha_1 | \alpha \rangle}{\|\alpha_1\|^2} \cdot |\alpha_1\rangle,$$

expressão que já conhecemos do Processo de Gram-Schmidt. Semelhante desenvolvimento aplica-se a α_2 .

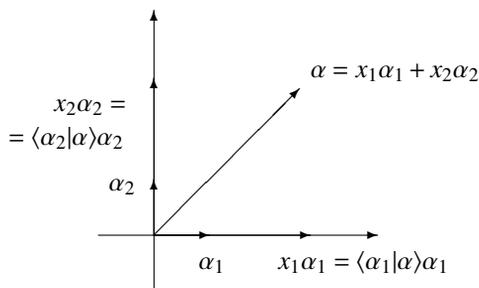


Figura 4.1: Projeções ortogonais de α sobre α_1 e α_2 unitários.

Exercício 2.1. No \mathbb{R}^3 , encontre o vetor projeção do vetor $\beta = (1, 3, 5)$ sobre o vetor $\alpha = (-1, 2, 0)$. E agora sobre o vetor $\gamma = (0, -1, 2)$. Finalmente, encontre a projeção do vetor β sobre o sub-espaço gerado por α e γ .

Recorde que $E^2 = E.E$, sendo a operação indicada a composição (produto) de operadores. Suponhamos que $\mathcal{V} = \mathcal{W}_1 \oplus \cdots \oplus \mathcal{W}_k$, e $\alpha = \alpha_1 + \cdots + \alpha_k$ conforme acima. Para cada $i = 1, \dots, k$ definimos um operador projeção do seguinte modo: $E_i(\alpha) = \alpha_i$. Assim, podemos escrever $\alpha = E_1(\alpha) + \cdots + E_k(\alpha)$, e tendo em vista que $I(\alpha) = \alpha$ (operador identidade), vem que

$$I = E_1 + \cdots + E_k. \quad (4.3)$$

Teorema 2.1. *Seja T operador linear sobre \mathcal{V} , $\mathcal{W}_1, \dots, \mathcal{W}_k$ sub-espaços de \mathcal{V} , tais que $\mathcal{V} = \mathcal{W}_1 \oplus \cdots \oplus \mathcal{W}_k$, e sejam E_i projeções associadas aos sub-espaços \mathcal{W}_i . Então, uma condição necessária e suficiente para cada \mathcal{W}_i seja invariante sob T é que T comute com cada E_i .*

Demonstração. Não será feita aqui. O leitor pode consultar [HofKun.79, pp.274-5]. \square

Resultado importante é o seguinte, conhecido como Teorema Espectral, que em suma afirma que **todo operador pode ser escrito como uma combinação linear de projeções**, e vale em particular para operadores hermitianos (ou ortogonais, no caso real):

Teorema 2.2 (Teorema Espectral). *Se $\lambda_1, \dots, \lambda_k$ são auto-valores distintos de T diagonalizável e E_i os operadores projeção sobre os espaços característicos dos λ_i ($i = 1, \dots, k$), então*

$$T = \lambda_1 E_1 + \lambda_2 E_2 + \dots + \lambda_k E_k.$$

Demonstração. Suponha T diagonalizável e $\lambda_1, \dots, \lambda_k$ são auto-valores distintos de T . Sejam \mathcal{W}_i ($i = 1, \dots, k$) os sub-espaços característicos associados aos auto-valores correspondentes. Então sabemos que podemos escrever $\mathcal{V} = \mathcal{W}_1 \oplus \dots \oplus \mathcal{W}_k$. Sejam E_i as projeções associadas a cada um dos espaços desta decomposição, respeitando-se os índices. Então, para $\alpha \in \mathcal{V}$, podemos escrever $\alpha = E_1(\alpha) + \dots + E_k(\alpha)$, logo $T(\alpha) = T(E_1(\alpha) + \dots + E_k(\alpha)) = TE_1(\alpha) + \dots + TE_k(\alpha)$, ou seja, tendo em vista que os \mathcal{W}_i são invariantes pela ação de T , então as projeções E_i comutam com T , resultando $T(\alpha) = E_1 T(\alpha) + \dots + E_k T(\alpha) = E_1 c_1 \alpha + \dots + E_k c_k \alpha = c_1 E_1(\alpha) + \dots + c_k E_k(\alpha)$, logo $T = c_1 E_1 + \dots + c_k E_k$. \square

Exercício 2.2. (i) Mostre que os operadores E_i definidos acima são de fato projeções. (ii) Mostre que para projeções, vale a expressão $E_i E_j = 0$ se $i \neq j$. (iii) Mostre que a cada operador de projeção E está associado um sub-espaço vetorial de \mathcal{V} , e reciprocamente. (iv) Mostre que os únicos auto-valores possíveis para um operador de projeção são 0 e 1.¹

2.1. Notação para projeções

Uma notação útil no caso da mecânica quântica é a seguinte. Seja \mathcal{W} sub-espaço de \mathcal{V} e $\{\alpha_1, \dots, \alpha_k\}$ base ortonormal para \mathcal{W} . Então, para $\alpha \in \mathcal{V}$, o operador

$$E_{\mathcal{W}}(\alpha) = \sum_i \langle \alpha_i | \alpha \rangle \alpha_i$$

¹Este é o motivo pelo qual eles são escolhidos na mecânica quântica para representar propriedades dos sistemas físicos cujas únicas medidas possíveis resultem em algo como SIM ou NÃO.

ou, na notação de Dirac,

$$E_{\mathcal{W}}|\alpha\rangle = \sum_i \langle \alpha_i | \alpha \rangle |\alpha_i\rangle$$

é um operador de projeção sobre \mathcal{W} (exercício). Se a dimensão de \mathcal{W} for 1, o vetor projetado (suposto não nulo) é uma base para \mathcal{W} , de sorte que podemos normalizá-lo e chamá-lo de $|\psi\rangle$ para simplificar, de sorte podemos escrever $E_{\mathcal{W}}|\psi\rangle = \langle \psi | \alpha \rangle |\psi\rangle = |\psi\rangle \langle \psi | \alpha \rangle$ e então resulta a notação conveniente

$$E_{\mathcal{W}} = |\psi\rangle \langle \psi|. \quad (4.4)$$

Assim, para acharmos a projeção de $|\alpha\rangle$ sobre \mathcal{W} , basta obter

$$E_{\mathcal{W}}|\alpha\rangle = |\psi\rangle \langle \psi | \alpha \rangle,$$

que é um modo de reescrever $E_{\mathcal{W}}|\alpha\rangle = \langle \psi | \alpha \rangle |\psi\rangle$ da forma como estamos acostumados, ou seja, com o escalar precedendo o vetor.

Exemplo 2.1. Considere novamente \mathcal{V} como o espaço \mathbb{R}^3 bem como seus sub-espacos $\mathcal{W}_1 = \{(x, 0, 0) : x \in \mathbb{R}\}$ e $\mathcal{W}_2 = \{(x, y, 0) : x, y \in \mathbb{R}\}$. Seja $\alpha = (3, -2, 4)$ um vetor de \mathcal{V} , $\mathcal{A} = \{\beta_1\} = \{(1, 0, 0)\}$ e $\mathcal{B} = \{\alpha_1, \alpha_2\} = \{(1, 0, 0), (0, 1, 0)\}$ bases ortonormais ordenadas de \mathcal{W}_1 e de \mathcal{W}_2 respectivamente. As projeções de α sobre cada um dos espacos podem ser obtidas da seguinte maneira:

$$E_{\mathcal{W}_1}|\alpha\rangle = |\beta_1\rangle \langle \beta_1 | \alpha \rangle = 3(1, 0, 0) = (3, 0, 0)$$

Sobre o segundo sub-espaco,

$$E_{\mathcal{W}_2}|\alpha\rangle = \sum_{i=1}^2 |\alpha_i\rangle \langle \alpha_i | \alpha \rangle = |\alpha_1\rangle \langle \alpha_1 | \alpha \rangle + |\alpha_2\rangle \langle \alpha_2 | \alpha \rangle = (3, -2, 0).$$

Exercício 2.3. Encontre a projeção do vetor $\alpha = (2, -3, -1)$ sobre os sub-espacos do \mathbb{R}^3 gerados respectivamente por $\{(1, -1, 0), (1, 1, 1)\}$ e $\{(-1, 0, -1), (0, 1, -1)\}$. Repare que os vetores geradores não estão normalizados.

3. Resolução da identidade

Iniciemos com um exemplo. Usando convenientemente a notação de Dirac, suponha que temos um espaço de Hilbert \mathcal{H} de dimensão 3 e que $\{|\alpha\rangle, |\beta\rangle, |\gamma\rangle\}$ seja uma base ortonormal para tal espaço. Seja \mathcal{U} um sub-espaço de \mathcal{V} gerado por $|\alpha\rangle$ e $|\beta\rangle$. Então o operador

$$P_{\mathcal{U}} := |\alpha\rangle\langle\alpha| + |\beta\rangle\langle\beta|$$

é uma projeção sobre \mathcal{U} (exercício). Ou seja,

$$P_{\mathcal{U}}(x|\alpha\rangle + y|\beta\rangle + z|\gamma\rangle) = x|\alpha\rangle + y|\beta\rangle.$$

Denotaremos por $\mathbf{1}$ o operador identidade sobre um espaço de Hilbert \mathcal{H} . A situação mais geral sugerida pelo que se viu acima é a de que o operador identidade pode ser escrito como soma de projeções (lembre da equação (4.3)):

$$\mathbf{1} = \sum_i |\alpha_i\rangle\langle\alpha_i|, \quad (4.5)$$

onde $\{|\alpha_i\rangle\}$ é uma base de \mathcal{H} . Uma tal expressão para $\mathbf{1}$ é denominada de **resolução da identidade**.

Claro que, para qualquer vetor $|\alpha\rangle$, tem-se que $\mathbf{1}|\alpha\rangle = |\alpha\rangle$, o que sugere que podemos utilizar a resolução da identidade sempre que desejarmos.

4. A função traço

A função traço, que já encontramos no capítulo 3, associa cada matriz quadrada a um escalar que é a soma dos elementos de sua diagonal principal. Ou seja, se $A = [a_{ij}]$ é de ordem n , então

$$\text{Tr}(A) = \sum_{i=1}^n a_{ii}.$$

No formalismo da física quântica, há um modo de representar a função traço que é importante, o qual ainda mostra o significado da expressão "traço de um operador", e resulta do seguinte

Teorema 4.1. *Seja T operador linear sobre \mathcal{V} com produto interno. Se a série $\sum_i \langle \psi_i | T(\psi_i) \rangle$ converge e tem o mesmo limite independentemente da base $\{|\psi_i\rangle\}$, então*

$$\text{Tr}(T) = \sum_i \langle \psi_i | T\psi_i \rangle = \sum_i \langle \psi_i | T|\psi_i \rangle. \quad (4.6)$$

Repare que para passar para o último termo, simplesmente mudamos a notação, escrevendo $T|\psi_i\rangle$ ao invés de $T\psi_i$. Aceitando-se o teorema precedente, seja A a matriz de T na base canônica (como o teorema se refere a algo invariante pela troca de bases, basta que consideremos a canônica). Suponha por simplicidade que o espaço seja o \mathbb{R}^2 e que

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Como os vetores da base são $\psi_1 = (1, 0)$ e $\psi_2 = (0, 1)$, que na forma de matriz (ou de suas coordenadas em relação à base canônica) vem que

$$\sum_{i=1}^2 \langle \psi_i | A | \psi_i \rangle = (1\ 0) \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + (0\ 1) \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = a_{11} + a_{22},$$

que coincide com a definição dada acima.

5. Complemento ortogonal

Principalmente devido às aplicações da Álgebra Linear à mecânica quântica, é conveniente chamarmos a atenção para o complemento ortogonal de um sub-espaço.

Definição 5.1 (Complemento ortogonal). Seja \mathcal{W} sub-espaço de \mathcal{V} . Chama-se de **complemento ortogonal** de \mathcal{W} ao sub-espaço (como se pode provar, veja exercício abaixo) \mathcal{W}^\perp , dito \mathcal{W} -perpendicular, formado por todos os vetores de \mathcal{V} que são ortogonais a todos os vetores de \mathcal{W} . Em símbolos,

$$\mathcal{W}^\perp := \{\beta \in \mathcal{V} : \langle \beta | \alpha \rangle = 0, \text{ para } \alpha \in \mathcal{W}\}.$$

Exercício 5.1. Mostre que \mathcal{W}^\perp é sub-espaço de \mathcal{V} .

Exemplo 5.1. Vamos considerar alguns de nossos sub-espaços favoritos para dar alguns exemplos e propor exercícios. Considere novamente os seguintes sub-espaços do \mathbb{R}^3 :

1. $X = \{(x, 0, 0) : x \in \mathbb{R}\}$ ("eixo X").
2. $Y = \{(0, y, 0) : y \in \mathbb{R}\}$ ("eixo Y").
3. $Z = \{(0, 0, z) : z \in \mathbb{R}\}$ ("eixo Z").
4. $XY = \{(x, y, 0) : x, y \in \mathbb{R}\}$ ("plano XY").
5. $XX = \{(x, 0, z) : x, z \in \mathbb{R}\}$ ("plano XZ").
6. $YZ = \{(0, y, z) : y, z \in \mathbb{R}\}$ ("plano YZ").

Verifica-se facilmente que (tente ao menos se convencer disso) $X^\perp = YZ$, $Y^\perp = XZ$, $Z^\perp = XY$.

Algumas propriedades que podem ser facilmente derivadas são:

Teorema 5.1. *Tem-se que, para todo sub-espço \mathcal{W} do espaço \mathcal{V} (com a notação anterior);*

1. $(\mathcal{W}^\perp)^\perp = \mathcal{W}$
2. $\{0\}^\perp = \mathcal{V}$
3. $\mathcal{V} = \mathcal{W} \oplus \mathcal{W}^\perp$
4. *Obtém-se uma base para \mathcal{V} unindo-se uma base de \mathcal{W} com uma base de \mathcal{W}^\perp .*

O último item, devido às propriedades da soma direta, acarreta que para todo vetor $\alpha \in \mathcal{V}$, existem **únicos** $\beta \in \mathcal{W}$ e $\gamma \in \mathcal{W}^\perp$ tais que $\alpha = \beta + \gamma$.

O vetor γ da expressão acima, que podemos denotar por β^\perp , é denominado sem surpresa de **complemento ortogonal de β relativamente a α** . Suas propriedades são relevantes, principalmente em mecânica quântica:

Teorema 5.2. *Se $\mathcal{V} = \mathcal{W} \oplus \mathcal{W}^\perp$, e usando-se a notação anterior $\alpha = \beta + \beta^\perp$, tem-se:*

1. *Os vetores β e β^\perp são linearmente independentes.*
2. *Se $\alpha = \beta + \beta^\perp$, então $(\beta^\perp)^\perp = \beta$. Este condicional pode parecer estranho, por isso vamos explicar. Tome a contrapositiva: $(\beta^\perp)^\perp \neq \beta \rightarrow \alpha \neq \beta + \beta^\perp$. É fácil ver que esta sentença tem modelo. Tome $\alpha = (1, 1, 0)$ e sejam $\beta = (0, 1, 0) \in$*

XY e $\beta^\perp = (0, 0, 1) \in Z$ (na notação dos exemplos anteriores). Portanto $\alpha = \beta + \beta^\perp$. Porém, podemos considerar um outro vetor de XY ortogonal a β , digamos $\gamma = (1, 0, 0) \in XY$, e claramente $\alpha \neq \beta + \gamma$. Portanto, $(\beta^\perp)^\perp$ não é qualquer vetor ortogonal a β , mas exatamente aquele que, somado com ele, fornece o α dado.

Exercício 5.2 (Importante). Considere o conjunto \mathcal{F} de todos os sub-espacos fechados de um espaco de Hilbert \mathcal{H} , munido das operações seguintes, para \mathcal{W} e \mathcal{U} sub-espacos: $\mathcal{W} \sqcap \mathcal{U} := \mathcal{W} \cap \mathcal{U}$ (interseção conjuntista); $\mathcal{W} \sqcup \mathcal{U} := [\mathcal{W} \cup \mathcal{U}]$ (espaco gerado pela união); $\overline{\mathcal{W}} := \mathcal{W}^\perp$, e os elementos distinguidos $\{O\}$ (sub-espaco trivial) e o próprio \mathcal{F} . Mostre que a estrutura $\langle \mathcal{F}, \sqcap, \sqcup, \overline{}, \{O\}, \mathcal{H} \rangle$ não é uma álgebra de Boole, mas um **reticulado ortomodular** (detalhes em [DalGiuGre.04], [Kra.16]).

PRODUTO TENSORIAL

GRASSMANN introduziu também os princípios da hoje chamada Álgebra Multilinear, que trata de estruturas que não envolvem unicamente um determinado ‘tipo’ de vetor, mas uma multiplicidade de vetores. Aqui, não desenvolveremos nada nesse sentido, mas apenas alguns traços de um conceito envolvendo multiplicidade de vetores que tem a ver com a mecânica quântica, o **produto tensorial** de vetores e de espaços vetoriais, em especial de espaços de Hilbert. Mesmo assim, ficaremos restritos aos casos mais simples, visando unicamente fornecer ao leitor algo com o que ele possa adentrar aos detalhes formais mínimos da mecânica quântica, como entender como um estado de emaranhamento é expresso no formalismo.

1. Produto tensorial

Definição 1.1 (Produto Tensorial de dois espaços). Sejam \mathcal{V} e \mathcal{W} dois espaços vetoriais sobre um mesmo corpo de dimensões n e m respectivamente. Associamos a eles um espaço vetorial $\mathcal{V} \otimes \mathcal{W}$ de dimensão $n.m$ cujos vetores são denotados $\alpha \otimes \beta$, satisfazendo a seguintes condições:

1. Se $\alpha, \alpha_1, \alpha_2 \in \mathcal{V}$ e $\beta, \beta_1, \beta_2 \in \mathcal{W}$, então valem as leis distributivas

$$(a) \quad \alpha \otimes (\beta_1 + \beta_2) = \alpha \otimes \beta_1 + \alpha \otimes \beta_2$$

$$(b) \quad (\alpha_1 + \alpha_2) \otimes \beta = \alpha_1 \otimes \beta + \alpha_2 \otimes \beta$$

2. Se k é um escalar qualquer, tem-se que

$$(k\alpha) \otimes \beta = \alpha \otimes (k\beta) = k(\alpha \otimes \beta).$$

3. Se $\{\alpha_i\}$ é uma base para \mathcal{V} e $\{\beta_j\}$ é uma base para \mathcal{W} , então

$$\{\alpha_i \otimes \beta_j\}$$

é uma base para $\mathcal{V} \otimes \mathcal{W}$, que portanto tem dimensão $n.m.$

Se denotarmos os produtos internos de \mathcal{V} e \mathcal{W} (supondo que existam) por $\langle | \rangle_{\mathcal{V}}$ e $\langle | \rangle_{\mathcal{W}}$ respectivamente, a seguinte aplicação define um produto interno sobre $\mathcal{V} \otimes \mathcal{W}$, como se comprova facilmente:

$$\langle \alpha_1 \otimes \beta_1 | \alpha_2 \otimes \beta_2 \rangle := \langle \alpha_1 | \alpha_2 \rangle_{\mathcal{V}} \cdot \langle \beta_1 | \beta_2 \rangle_{\mathcal{W}}. \quad (5.1)$$

Repare que a multiplicação do segundo membro faz sentido tendo em vista que os dois espaços são espaços vetoriais sobre o mesmo corpo. A definição acima pode ser generalizada para mais de dois espaços (sobre um mesmo corpo).

Pode-se provar que se \mathcal{V} e \mathcal{W} forem espaços de Hilbert, então $\mathcal{V} \otimes \mathcal{W}$ é por sua vez um espaço de Hilbert.

Produto tensorial de operadores Se T_1 é operador linear sobre \mathcal{V} e T_2 é operador linear sobre \mathcal{W} , definimos $T_1 \otimes T_2$ sobre $\mathcal{V} \otimes \mathcal{W}$ do seguinte modo. Se α e β são vetores de \mathcal{V} e \mathcal{W} respectivamente, pomos

$$(T_1 \otimes T_2)(\alpha \otimes \beta) := T_1(\alpha) \otimes T_2(\beta). \quad (5.2)$$

Se T_1 e T_2 são lineares, resulta que $T_1 \otimes T_2$ é também linear, e esta característica se estende para outras propriedades dos operadores, como se eles forem unitários, hermitianos, auto-adjuntos. Importante mencionar que o produto tensorial, seja de vetores, de espaços ou de operadores, *não é comutativo*, ou seja, em geral temos

$$|\alpha\rangle \otimes |\beta\rangle \neq |\beta\rangle \otimes |\alpha\rangle,$$

o mesmo se dando nos demais casos.

Notação de Dirac Na notação de Dirac, o vetor $|\alpha\rangle \otimes |\beta\rangle$ é muitas vezes escrito como $|\alpha\rangle|\beta\rangle$ ou simplesmente por $|\alpha\beta\rangle$. Isso pode ser generalizado. Deste modo, empregando uma notação mais útil para a mecânica quântica,¹

$$|x_1, x_2, \dots, x_n\rangle, \quad (5.3)$$

que deve ser lido como um vetor do espaço $\mathcal{H} = \mathcal{H}_1 \otimes \dots \otimes \mathcal{H}_n$.

Utilizamos o produto tensorial de espaços de Hilbert para encontrar vetores que representem os estados de sistemas compostos, com mais de uma partícula por exemplo. Assim, se temos um elétron e^- e um pósitron e^+ , associamos a eles os espaços de Hilbert \mathcal{H}_1 e \mathcal{H}_2 respectivamente, e os estados do sistema composto pelos dois *quanta*² será dado por vetores de $\mathcal{H}_1 \otimes \mathcal{H}_2$. Mais à frente veremos exemplos.

Porém, quando os sistemas físicos que estamos considerando são indiscerníveis, para representar os estados do sistema conjunto empregamos um produto tensorial de um espaço (que representa os estados de um dos sistemas) por ele mesmo exatamente o número de vezes que corresponde ao cardinal da coleção de sistemas. Assim, se a σ associamos o espaço de Hilbert \mathcal{H} , a um sistema contendo três sistemas indiscerníveis de σ associamos o espaço $\mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H}$, cujos vetores são da forma $|x_1, x_2, x_3\rangle$.

Um pouco mais sobre isso é relevante para esclarecer alguns pontos que normalmente passam despercebidos. Suponha que temos duas partículas indiscerníveis (ou *idênticas*, como dizem os físicos) rotuladas 1 e 2.³

Esses conceitos não são triviais e são problemáticos em MQ, mas não trataremos da questão detalhadamente aqui; o leitor interessado pode consultar [FreKra.06]. Primeiramente, observe a necessidade de nossa linguagem: quando falamos de entidades, nomeamo-las, damos-lhes nomes ou as descrevemos de algum modo (como ‘o elétron que está no nível mais externo do átomo tal’) — os filósofos chamam isso de uma *descrição definida*. Isso faz com

¹Aqui é um dos casos (talvez o nosso único) em que vetores não são representados por letras gregas latinas. É que seguimos a ortodoxia em física.

²É comum a referência deste modo aos sistemas quânticos.

³Em matemática e em filosofia, ‘idênticos’ significa que são o mesmo, que não há dois. Assim, se $a = b$ é o caso, então isso diz que há um só objeto que pode tanto ser nomeado por a quanto por b . Diz-se que a e b têm o mesmo *referente*. Mas em física, os físicos costumam dizer que dois elétrons são *idênticos* mas que podem ser discerníveis ou indiscerníveis. Usamos aqui a terminologia matemática.

que a sua identidade se torne indubitável.⁴ Se queremos preservar a indiscernibilidade, o que é necessário que se faça em várias situações, por exemplo em química, esses rótulos terão que ‘desaparecer’ de algum modo, deixando de conferir identidade ao objeto, e isso é conseguido por meio de um truque matemático de considerar que os únicos vetores (ou funções de onda) que representam estados ‘reais’ são simétricos (bósons) ou anti-simétricos (férmions) relativamente à permutação de dois desses rótulos, como x_i e x_j . Para captar esta ideia, os físicos introduzem um princípio por meio do seguinte postulado:

Seja $|\psi\rangle = |x_1, x_2, \dots, x_n\rangle$ vetor que descreve o estado de um sistema de n partículas indistinguíveis. Seja P_{ij} **operador de permutação** que troca os índices das partículas x_i e x_j , ou seja,

$$P_{ij}|x_1, x_2, \dots, x_i, \dots, x_j, \dots, x_n\rangle = \pm|x_1, x_2, \dots, x_j, \dots, x_i, \dots, x_n\rangle.$$

O sinal positivo é usado quando as partículas são bósons e o negativo quando são férmions. Seja A um observável que é representado por um operador hermitiano \hat{A} . O postulado diz o seguinte, em notação simplificada (escrevendo-se $P_{ij}\psi$ ao invés de $P_{ij}|\psi\rangle$):

Postulado da Indistinguibilidade Nas condições acima, temos:

$$\langle\psi|\hat{A}|\psi\rangle = \langle P_{ij}\psi|\hat{A}|P_{ij}\psi\rangle. \quad (5.4)$$

Em palavras, o **valor esperado** da medida do observável A para o sistema no estado $|\psi\rangle$ é o mesmo antes e depois de uma permutação de partículas (ou, no formalismo, de seus rótulos). Como já aprendemos antes, $\langle\psi|\hat{A}|\psi\rangle$ pode ser lida como o produto interno do vetor $|\psi\rangle$ pelo vetor $|P_{ij}\psi\rangle$ ou como a imagem deste vetor pelo funcional linear caracterizado por $|\psi\rangle$ (ou seja, por $\langle\psi|$).

Isto significa que a permutação *não produz qualquer efeito físico detetável*, o que caracteriza a indiscernibilidade das partículas. O fato de que para férmions a permutação muda o sinal não atrapalha porque no cômputo das probabilidades, o que entra em jogo é o *quadrado* da função de onda (ou do vetor de estado), de acordo com a **regra de Born**, temos que $\| |\psi_{ij}\rangle \|^2 = \| |\psi_{ji}\rangle \|^2$, como veremos.

⁴Logo, na suposição de que existam sistemas indiscerníveis, nem nomes próprios e nem descrições definidas podem fazer sentido nesse domínio. Sobre isso, ver [Kra16a, Cap.6].

Os físicos perceberam que para preservar estes fatos, os sistemas envolvendo múltiplas partículas deverem ter seus estados descritos por vetores *simétricos* ou *anti-simétricos* relativamente à permutação de índices. Se denotarmos por P um operador de permutação e por $|\psi_{ij}\rangle$ o vetor de estado do sistema obtido do vetor $|\psi_{ji}\rangle$ pela permutação das partículas i e j , a função será simétrica se $P|\psi_{ij}\rangle = |\psi_{ji}\rangle$, e anti-simétrica se $P|\psi_{ij}\rangle = -|\psi_{ji}\rangle$. Isso casa perfeitamente com a existência de duas classes de partículas, os *bósons* e os *férmions*. Todas as partículas conhecidas caem sob uma dessas categorias. Grosso modo, férmions são partículas de matéria, e os bósons fazem a interação entre elas. Uma outra classificação diz que férmions têm spin semi-inteiro: $\pm 1/2, \pm 3/2, \dots$, enquanto que os bósons têm spin inteiro $0, \pm 1, \pm 2, \dots$.

Vetores simétricos e anti-simétricos são os seguintes, onde por exemplo $|\psi_i^A\rangle$ deve ser lido como indicando que a partícula i está na situação A ; o mesmo para os demais:

$$|\psi_{ij}\rangle = \frac{1}{\sqrt{2}}(|\psi_i^A\rangle \otimes |\psi_j^B\rangle \pm |\psi_j^A\rangle \otimes |\psi_i^B\rangle) \quad (5.5)$$

O fator $\frac{1}{\sqrt{2}}$ é dito **fator de normalização** e faz com que os vetores sejam unitários (produto interno canônico). Estes vetores são ditos indicarem estados *singlete(s)*. O sinal $+$ é usado no caso de bósons e o $-$ no caso de férmions. Repare o que eles indicam: trocando-se i por j e vice-versa, o primeiro vetor (o simétrico com o sinal $+$) não se altera, enquanto que o segundo troca de sinal. Este truque faz com que possamos fazer de conta que os rótulos dados antes, i e j , não identificam as partículas porque os estados que interessam são descritos pelos vetores acima; vetores como $|\psi_i\rangle \otimes |\psi_j\rangle$, assumem os físicos, *não descrevem estado algum*. Porém, em $|\psi_i\rangle \otimes |\psi_j\rangle$, ambos os sistemas estão no mesmo estado, o que pode acontecer somente para bósons.⁵

2. Emaranhamento

Os estados compostos mais simples são aqueles que podem ser escritos por meio de um produto tensorial de vetores, com cada vetor pertencente a um dos espaços que participam do produto. Por exemplo, considerando o nosso já investigado caso do spin, suponha que o primeiro vetor seja $|\psi_1\rangle = a_u|u\rangle_{\mathcal{H}_1} +$

⁵Férmions não podem partilhar o mesmo estado quântico devido ao Princípio de Exclusão de Pauli.

$a_d|d\rangle_{\mathcal{H}_1}$ (equação 3.13) no espaço \mathcal{H}_1 , onde foram colocadas referências a este espaço. O segundo vetor está por hipótese em um outro espaço, digamos que seja $|\psi_2\rangle = b_u|u\rangle_{\mathcal{H}_2} + b_d|d\rangle_{\mathcal{H}_2}$ em \mathcal{H}_2 , supostos normalizados, ou seja, $|a_u|^2 + |a_d|^2 = a_u^* a_u + a_d^* a_d = 1$ e $|b_u|^2 + |b_d|^2 = b_u^* b_u + b_d^* b_d = 1$.

O estado conjunto, denominado de **estado produto**, ou **estado separável**, (ou **estado fatorável**) seria uma soma como a seguinte:

$$|\psi_{12}\rangle_{\mathcal{H}_1 \otimes \mathcal{H}_2} = (a_u|u\rangle_{\mathcal{H}_1} + a_d|d\rangle_{\mathcal{H}_1}) \otimes (b_u|u\rangle_{\mathcal{H}_2} + b_d|d\rangle_{\mathcal{H}_2}).$$

Podemos dizer que em uma tal situação cada sistema comporta-se independentemente dos demais, ou seja, trata-se de um vetor cuja expansão em uma base qualquer pode ser fatorada em um produto tensorial de dois termos, um em cada espaço (claro que isso pode ser generalizado para mais de dois espaços). Por exemplo, deixemos $\{|\beta_i\rangle_1\}$ representar uma base para o espaço de Hilbert \mathcal{H}_1 e $\{|\gamma_j\rangle_2\}$ representar uma base para um espaço \mathcal{H}_2 . Assim, $\{|\beta_i\rangle_1|\gamma_j\rangle_2\}$ denota uma base para o produto tensorial $\mathcal{H}_1 \otimes \mathcal{H}_2$ (omitimos o símbolo ‘ \otimes ’ nos vetores da base, assim como os sub-índices). Por exemplo, podemos ter um estado descrito pelo vetor

$$\begin{aligned} |\psi\rangle &= \frac{1}{\sqrt{2}}(|\beta_1\rangle|\gamma_3\rangle + |\beta_4\rangle|\gamma_3\rangle) + \frac{1}{\sqrt{2}}(|\beta_4\rangle|\gamma_3\rangle + |\beta_4\rangle|\gamma_1\rangle) = \\ &= \frac{1}{\sqrt{2}}(\underbrace{|\beta_1\rangle + |\beta_4\rangle}_{\mathcal{H}_1}) \otimes \frac{1}{\sqrt{2}}(\underbrace{|\gamma_3\rangle + |\gamma_1\rangle}_{\mathcal{H}_2}), \end{aligned}$$

separável.

A situação mais complexa é quando isso não acontece, ou seja, quando o vetor produto não pode ser decomposto em um produto de vetores cada um em um espaço, de forma que não se pode dizer que eles agem independentemente, mas um dos estados tornam-se intrinsecamente **emaranhados** um com o outro (supondo somente dois deles, mas isso pode ser generalizado). A palavra ‘emaranhamento’ (*entanglement*) foi cunhada por Schrödinger em 1935, tendo ele destacado que é um caso típico da mecânica quântica:

“I would not call that *one* but rather *the* characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought.” (E. Schrödinger)⁶

⁶“Eu não chamaria de *um* mas em vez disso de *a* característica peculiar da mecânica quântica, aquela que acarreta sua total discrepância para com as linhas clássicas de pensamento.”

O mais típico caso de emaranhamento é dado pelos nossos exemplos acima pelos vetores em (5.6) ou, usando o nosso exemplo, poderia ser

$$|\psi_{12}\rangle_{\mathcal{H}_1 \otimes \mathcal{H}_2} = \frac{1}{\sqrt{2}} \left((|u\rangle \otimes |d\rangle)_{\mathcal{H}_1 \otimes \mathcal{H}_2} \pm (|d\rangle \otimes |u\rangle)_{\mathcal{H}_1 \otimes \mathcal{H}_2} \right) \quad (5.6)$$

A prova de que este vetor não pode ser escrito como um produto de vetores um em cada espaço é tediosa, por isso rogamos ao leitor que nos acredite.

2.1. Posições definidas ?

Vamos enunciar uma regra básica da mecânica quântica, que chamaremos de RAA:

[**Regra dos auto-valores e dos auto-vetores**] A medida de um observável A , representado por um operador auto-adjunto \hat{A} , para o sistema em certo estado $|\psi\rangle$ possui o valor λ para uma propriedade representada por A se e somente se $|\psi\rangle$ é auto-vetor de \hat{A} e λ é um autovalor associado a $|\psi\rangle$. Ou seja, se e somente se

$$\hat{A}|\psi\rangle = \lambda|\psi\rangle.$$

Vejam agora uma aplicação desta regra.

Vamos supor que A seja o observável **posição**, de uma partícula 1, cujos estados estão representados em um espaço de Hilbert \mathcal{H}_1 e B o operador posição de uma partícula 2, cujos estados são vetores de \mathcal{H}_2 , e que o sistema composto por essas duas partículas (isso pode ser generalizado) esteja no estado emaranhado (em $\mathcal{H}_1 \otimes \mathcal{H}_2$)

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left(|\beta_1\rangle_1 |\gamma_1\rangle_2 + |\beta_2\rangle_1 |\gamma_2\rangle_2 \right), \quad (5.7)$$

de forma que por hipótese os $|\beta_j\rangle$ são auto-vetores de \hat{A} e os $|\gamma_k\rangle$ são auto-vetores de \hat{B} .

Repare uma coisa importante: não temos outro jeito, com nossa linguagem, exceto a nos referirmos às partículas como ‘partícula 1’ e ‘partícula 2’. Nossa linguagem é *uma linguagem feita para falar de objetos distintos uns dos outros*. No entanto, em situações como a em análise, queremos que esses rótulos não distingam as partículas. Por este motivo, selecionamos vetores simétricos ou anti-simétricos por permutações dos rótulos.

Para medirmos a posição da partícula 1 sem medir a da partícula 2, usamos o operador

$$\hat{A} \otimes I_2,$$

onde I_2 é o operador identidade sobre \mathcal{H}_2 . Analogamente, para medir a posição somente da partícula 2, usamos o operador

$$I_1 \otimes \hat{B},$$

onde I_1 é o operador identidade sobre \mathcal{H}_1 .

Então, verifica-se que $|\psi\rangle$ é um autovetor do operador

$$I_1 \otimes \hat{B} - \hat{A} \otimes I_2,$$

que representa a *diferença* entre as posições. No entanto, $|\psi\rangle$ não é autovetor de qualquer dos operadores $\hat{A} \otimes I_2$ ou $I_1 \otimes \hat{B}$, assim, por RAA, as partículas 1 e 2 *não têm posições definidas*.

Isso traz um complicador para a filosofia da mecânica quântica, que é a seguinte. As noções de espaço e de tempo da mecânica quântica não relativista são as mesmas da mecânica clássica, dito **espaço-tempo Newtoniano** ou **espaço-tempo de Newton-Galileo** (ver à frente a definição 3.2). São noções *absolutas*, e mais ou menos coincidem com o que pensa a pessoa comum, não versada em ciência: o espaço é o local onde os eventos ocorrem, não muda, está fixo e se estende até onde podemos imaginar, e o tempo é universal, variando igualmente para todos e em todas as situações, podendo ser marcada por uma espécie de ‘relógio universal’. Matematicamente, isso é descrito por um conceito matemático denominado *variedade diferenciável*, mas faremos uma simplificação, identificando o espaço-tempo com o espaço vetorial \mathbb{R}^4 com o produto interno canônico, que é denominado de **espaço euclidiano**, que aparecerá novamente no capítulo seguinte.

Ora, em \mathbb{R}^4 , dois pontos distintos $A = (x_1, y_1, z_1, t)$ (a quarta coordenada representa o tempo) e $B = (x_2, y_2, z_2, t)$ (repare que o tempo é o mesmo), são necessariamente distintos. Logo, se temos por hipótese duas partículas, uma em A e outra em B , elas são distintas (pois ocupam posições distintas) e teriam em princípio posições definidas. O problema é que se considerarmos uma partícula isolada, nunca podemos dizer que temos a partícula 1 exatamente em A devido ao princípio de indeterminação; quanto mais aproximamos a posição, mais tornamos o momento indeterminado e vice-versa, não obstante a

partícula em A (isso é uma maneira de dizer) seja descrita por uma função de onda ψ_A . Em mecânica quântica, não há trajetórias de partículas. Mais grave é o caso de considerarmos o sistema conjunto ‘partícula 1 e partícula 2’. Neste caso, o sistema é descrito por uma função de onda simétrica (equação 5.7 para bósons) ou anti-simétrica (para férmions, usando o sinal $-$ na mesma equação) e pelos motivos expostos acima não se pode dizer que elas tenham posições definidas. Mas então, como fica a ‘determinação’ das posições imposta quando adotamos o espaço-tempo newtoniano? Bom, este é o imbróglio. Não creio que haja qualquer resposta satisfatória.

Exercício 2.1. Mostre que de fato $|\psi\rangle$ não é autovetor nem de $\hat{A} \otimes I_2$ e nem de $I_1 \otimes \hat{B}$. (Basta ver que o vetor não se transforma, por qualquer desses operadores, em um múltiplo escalar dele mesmo.)

2.2. O gato de Schrödinger

Uma outra aplicação interessante da regra RAA é a seguinte, devida a Erwin Schrödinger (1935). Trata-se de uma **experiência de pensamento** destinada a mostrar as incongruências trazidas pela mecânica quântica se tentássemos levar os seus resultados ao mundo macroscópico. Saliente-se desde já que até o presente momento não se constataram efeitos quânticos dessa natureza em objetos como um gato, mas que tais resultados são hoje em dia comuns na escala quântica. Trata-se, mais uma vez, do emaranhamento.

Schrödinger imaginou um gato em um container completamente fechado junto com uma amostra de material radioativo que pode decair (emitir radiação) dentro de uma hora, e um frasco de veneno. Se a radiação for emitida, ela aciona um ‘dispositivo maligno’ (em suas palavras) que quebra o frasco e o veneno mata o gato. Caso contrário, nada ocorre e o gato continua vivo. A pergunta é: qual a situação do gato antes de se passar uma hora? A resposta dada pela mecânica quântica é que ele estará em um estado de superposição entre os estados $|\text{gato vivo}\rangle \otimes |\text{não decaiu}\rangle$ e $|\text{gato morto}\rangle \otimes |\text{decaiu}\rangle$, ou seja, simplificando de modo óbvio a notação,

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|v\rangle|n\rangle + |m\rangle|d\rangle). \quad (5.8)$$

O que ocorre quando, depois de uma hora, o cientista abre a porta para ver o que ocorreu? Segundo a mecânica quântica, neste momento (por um motivo

ainda muito discutido) a função de onda (o vetor de estado) (5.8) **colapsa** em um dos componentes, e cada um deles tem igual probabilidade ($= |\frac{1}{\sqrt{2}}|^2 = \frac{1}{2}$) de ocorrer. Assim, o cientista constata o óbvio, que houve decaimento, uma partícula foi emitida, o dispositivo foi acionado, o veneno liberado e o gato morreu, ou nada disso aconteceu e o gato continua vivo.

No entanto, antes de que o cientista abra a porta, o estado do sistema é dado pelo vetor (5.8), e nada se pode dizer sobre o que está ocorrendo lá dentro exceto pelo que ‘diz’ o vetor (5.8). Sejam A e B os observáveis cujas medidas descrevem as situações do gato e do material radioativo respectivamente, e sejam \mathcal{H}_1 e \mathcal{H}_2 os espaços de Hilbert respectivos. Ora, o vetor $|\psi\rangle$ não é auto-vetor nem do operador $\hat{A} \otimes I_2$ e nem de $I_2 \otimes \hat{B}$, no sentido já visto antes. Assim, simplesmente olhando para o estado (5.8), não podemos concluir, como muitos textos afirmam, que antes do cientista olhar, o gato está vivo e morto, pois, digamos assim, ele não pode assumir esses estados, ou ter essas propriedades pela regra RAA.

Em resumo, de acordo com a mecânica quântica, e contrariamente a muitos textos, **não há qualquer contradição envolvida** com a superposição de estados. Ademais, poderíamos dizer para reforçar este ponto, se uma contradição é tomada como sendo a conjunção de de uma sentença com sua negação, ou seja, uma fórmula da forma $A \wedge \neg A$, é preciso que se especifiquem os significados da conjunção e da negação, pois o que é uma contradição depende do sistema lógico utilizado. No caso quântico, o que temos é uma soma vetorial, que não tem as propriedades de uma conjunção (por exemplo a da lógica clássica – o leitor pode consultar um livro de lógica para isso) e o complemento ortogonal (os vetores $|v\rangle$ e $|m\rangle$ são ortogonais um ao outro) não tem as propriedades da negação (clássica, pelo menos). O complemento ortogonal é uma **involução** [DalGiuGre.04]. Finalmente, para que tenhamos uma contradição, com o gato estando vivo e morto previamente à medida, seria necessário que assumíssemos que o gato *possui* (em sentido *atual*) as duas propriedades, que elas pudessem assumir valores ao mesmo tempo, o que é impedido pela mecânica quântica. O emaranhamento, como salientou Schrödinger, é a novidade, o que faz a mecânica quântica realmente diferente da física clássica e não tem uma ‘explicação’ dentro desta. É a novidade. O gato, se dele podemos falar, antes da medição, não está nem vivo e nem morto, mas em uma terceira situação, que é descrita pelo vetor (5.8), e que não conseguimos explicar em palavras (que são calcadas em nossa experiência com a física clássica).

Há muita, realmente muita, literatura tratando do assunto. Mas insista-se que é preciso cuidado para não se extrair da mecânica quântica mais do que ela pode fornecer. A fala (nossa inclusive) livre sobre o gato e sua situação (o seu estado) previamente à medição, ao cientista ter feito a observação, como fizemos acima. Segundo Bohr, a mecânica quântica não diz nada sobre isso: “é errado pensar que a tarefa da física é desvendar como é a Natureza. A física concerne [apenas] com o que podemos falar acerca da Natureza.” E é dele também uma outra frase célebre:⁷ “se a mecânica quântica não produziu em você um choque profundo, é porque você ainda não a entendeu.”

É no entanto desejo do filósofo associar uma ontologia à teoria física. Neste caso, devemos reconhecer que, antes da medida ser realizada, não temos o gato propriamente dito, mas uma entidade em estado de superposição e estados. De que entidade estamos falando? Esta é a questão. J. -M. Levy-Leblond e F. Balibar sugerem que se tratam de ‘entidades novas’, às quais eles sugerem o nome **quanton**. Dizem eles

“Devemos abandonar a ideia de que todo objeto físico é uma onda ou uma partícula. Nem é possível dizer, como feito algumas vezes, que partículas ‘tornam-se’ ondas no domínio quântico e conversamente, que ondas são ‘transformadas’ em partículas. Nem deveria ser dito que os objetos quânticos têm uma natureza dual, que é simultaneamente ondulatória e corpuscular (algo que é um absurdo lógico, desde que os dois conceitos são mutuamente exclusivos).

“É portanto necessário reconhecer que temos aqui uma espécie diferente de entidade, uma que é especificamente quântica. Por esta razão nós as nomeamos *quantons*,⁸ mesmo sendo esta nomenclatura não utilizada universalmente. Esses quantons comportam-se de uma maneira muito específica . . . ” [LevBal.96, p.69]

2.3. O fascínio pelo emaranhamento

Como salientam Susskind e Friedman [SusFri.14, p.167], há dois fatos que fazem com que o emaranhamento seja tão “fascinante”(aqui com adendos):

⁷Um google em "Bohr quotes" apresenta muitas outras de suas citas.

⁸esta terminologia já havia sido utilizada por Mario Bunge anteriormente como sinônimo para *microsistema*. Ver [Bun.73, p.94].

1. O estado emaranhado é uma descrição completa do sistema, no sentido de que *nada mais pode ser conhecido sobre ele*. Este é um pressuposto que vai contra a ‘visão realista’ de Einstein, mas era adotado por Bohr e outros.
2. É um estado que fornece informação sobre o estado conjunto, e *nada pode ser conhecido sobre os sistemas individuais*. A isso denomina-se *holismo*. Em outros termos, aparentemente a mecânica quântica permite a existência de entidades *A* e *B* que, mesmo estando em locais separados espacialmente, têm estados que não podem ser descritos um sem referência ao outro. Deste modo, esses sistemas não fariam parte de ‘realidades independentes’, mas sim de um composto unindo (emaranhando) os estados de *A* e de *B* em uma totalidade.

Esta segunda situação, a de que há um vínculo (ou **correlação**) entre os dois (ou mais) estados, é importante e constitui o que de mais estranho e relevante há em mecânica quântica, sendo fundamental em experimentos e em informação quântica. Para dar uma ideia, vamos apontar para mais um exemplo de uma situação envolvendo correlação quântica. Primeiramente, devemos constatar que correlações existem mesmo em sistemas clássicos, descritos pela física clássica, como supostamente é o Dr. Bertlmann, amigo do grande físico John Stewart Bell. Bell imaginou a seguinte experiência de pensamento. O Dr. Bertlmann sempre veste meias com cores diferentes e sempre de modo imprevisível, de sorte que, sem olhar, jamais seremos capazes de saber a cor de suas meias. Mas fica claro que se a do pé direito é verde, a do esquerdo é de uma outra cor. Portanto, as cores das meias saem de casa com o Dr. Bertlmann correlacionadas. Se ao nos depararmos com ele constatamos que a meia do pé direito é rosa, *sabemos imediatamente* (ou seja, predizemos com probabilidade igual a 1) que a do pé esquerdo não é. Isso não tem mistério algum. Poderíamos ser levados a pensar que o emaranhamento quântico seria algo do mesmo tipo. O problema da mecânica quântica, no entanto, não é assim simples. Em um artigo célebre [Bel.87], John Bell explicou as razões disso.

Em linhas gerais, e a descrição aqui só pode ser assim informal, pois se trata de um dos fenômenos quânticos mais surpreendentes, é que no caso quântico não há nada nem ninguém dizendo ao Dr. Bertlmann que meias vestir pela manhã. Na verdade, para simular o caso quântico, devemos assumir que nem ele sabe a cor de suas meias, pois as veste no escuro, mas os pares de meias foram ‘preparados’ para que cada par contenha uma de cada cor, se depois

foram colocados aleatoriamente em sua gaveta de meias. A única coisa que sabemos, e que ele mesmo sabe, é que suas meias partilham um estado de superposição rosa-outra cor, verde-outra cor, e assim por diante, mas sem saber que cores são. E é somente quando a meia é observada, ou seja, quando se faz uma medição que a cor surge; antes, pelo menos segundo a interpretação padrão, dita Interpretação de Copenhague (devida principalmente a Bohr e Heisenberg, Jordan, Pauli e Born), nada pode ser dito sobre a cor das meias; é como se *no ato da medida* a cor da meia fosse estabelecida. Pascual Jordan, saiu-se com esta:

“Observações não somente disturbam o que está para ser medido, mas produzem [o resultado da] a medida ... Compelimos [digamos, um elétron] a assumir uma posição definida ... [Mas] nós mesmos produzimos os resultados das medidas.” (citado por [Mer.85])

Einstein, que era um *realista* (veja mais abaixo), nunca aceitou isso, sempre acreditando que Bertlmann ou alguém em sua casa haviam determinado as cores das meias desde o início e nos competiria simplesmente em estabelecê-las. Ou seja, ao ir à universidade, as meias de Bertlmann ‘levariam consigo’ as suas cores pré-estabelecidas (esta analogia é algo complicada, pois no caso das meias e de outros objetos ‘clássicos’, é difícil conceber que isso possa não ser exatamente como einstein pensava. Mas este é o modo ‘clássico’, realista, de pensar, que é contestado pela mecânica quântica). Com efeito, no mundo quântico isso não é assim; segundo a escola de Copenhague, as meias não têm cor antes de serem observadas. Estranho? Sim de fato, porém experimentos recentes (e principalmente os realizados depois de 1982, quando se estabeleceu um procedimento rigoroso e padronizado para se medir essas coisas), parece que não pode haver mais dúvidas: Bohr estava certo e Einstein errado.

Vamos ver isso de um modo um pouco mais técnico com uma característica tipicamente quântica, o spin. Temos duas partículas de spin-1/2, digamos um elétron (partícula 1) e um pósitron (partícula 2), obtidas (por decaimento) de uma partícula de spin 0 em um ponto O , conforme a figura 5.1. Daqui para a frente, nesta seção ao menos, abandonaremos o símbolo \otimes de produto tensorial.

Digamos que o eixo escolhido para medir o spin seja o eixo z . Então, em O , o sistema composto pelas duas partículas é dado pelo vetor abaixo, a menos do

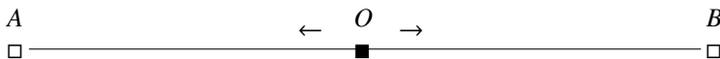


Figura 5.1: Um típico caso de emaranhamento.

fator de normalização $\frac{1}{\sqrt{2}}$, que omitiremos, e escolhemos o sinal menos porque tanto elétrons como pósitrons são férmions (o pósitron é a anti-partícula do elétron), lembrando que $|u\rangle$ e $|d\rangle$ denotam os estados UP e DOWN na direção z :

$$|\psi_{12}\rangle = |u_1\rangle|d_2\rangle - |d_1\rangle|u_2\rangle. \quad (5.9)$$

Assim, se medimos o spin do elétron na direção z e constatamos que ele está UP, o estado **colapsa** imediatamente (e de forma não determinista, pois tudo o que podemos avaliar são probabilidades) no estado $|u_1\rangle|d_2\rangle$ pelo Lema do Colapso que veremos mais à frente. Deste modo, podemos constatar que o spin do pósitron está no estado DOWN. Até aí, aparentemente, nada de mais. O problema é que podemos escolher uma outra direção, digamos x , de forma que teríamos (lembrando que $|r\rangle$ e $|l\rangle$ denotam os estados LEFT e RIGHT na direção x):

$$|r_1\rangle = |u_1\rangle + |d_1\rangle, \quad |r_2\rangle = |u_2\rangle + |d_2\rangle, \quad (5.10)$$

e

$$|l_1\rangle = |u_1\rangle - |d_1\rangle, \quad |l_2\rangle = |u_2\rangle - |d_2\rangle, \quad (5.11)$$

Na direção x , encontramos portanto

$$\begin{aligned} |r_1\rangle|l_2\rangle - |l_1\rangle|r_2\rangle &= (|u_1\rangle + |d_1\rangle) \cdot (|u_2\rangle - |d_2\rangle) - (|u_1\rangle - |d_1\rangle) \cdot (|u_2\rangle + |d_2\rangle) = \\ &= |u_1\rangle|u_2\rangle - |u_1\rangle|d_2\rangle + |d_1\rangle|u_2\rangle - |d_1\rangle|d_2\rangle - |u_1\rangle|u_2\rangle - |u_1\rangle|d_2\rangle + |d_1\rangle|u_2\rangle + |d_1\rangle|d_2\rangle = \\ &= -2(|u_1\rangle|d_2\rangle - |d_1\rangle|u_2\rangle) = -2|\psi_{12}\rangle. \end{aligned}$$

Ora, já sabemos desde o início que um múltiplo escalar de um vetor de estado denota o mesmo estado pois estão no mesmo *ray*. Assim, os estados

$|\psi_{12}\rangle$ e $-2|\psi_{12}\rangle$. representam os mesmos estados, mas agora podemos imaginar que o sistema está em uma superposição do elétron ‘girando’ (*spining*) para a direita e o pósitron para a esquerda na direção x .

Isso indica que, se agora A quiser medir o spin do elétron na direção x sem avisar B e achar RIGHT, saberá que se B medir achará LEFT necessariamente. E isso é comprovado em laboratório. Vamos ser um pouco mais explícitos, pois com isso você aprenderá um pouco mais de mecânica quântica através do formalismo que vimos empregando. O interessante é que você pode mudar a direção da medida, assim como Bertlmann mudava de meias, e os valores de spin (ou a cor da outra meia) muda em consonância, sem que seja preciso *preparar* nada (contar a cor colocada em um dos pés para saber que a outra será diferente). E isso é verificado em laboratório.

A pergunta a se fazer é: por que não podemos supor, como queria Einstein em sua **hipótese realista**, que as partículas já tenham todas as suas propriedades e seus valores pré-determinados quando produzidos em O , como por exemplo os valores de spin nas direções, de forma que tudo o que necessitamos fazer é medir em uma direção para saber a da outra instantaneamente sem precisar medi-la? Einstein ainda argumentava a favor da **localidade**, uma hipótese que diz, resumidamente, que os eventos acontecem *localmente*, ou seja, se realizamos um experimento em um ponto P_1 , para que algo se reflita em um outro ponto P_2 , é preciso que algum tipo de informação seja transmitida de P_1 para P_2 , e isso, de acordo com a relatividade restrita, não pode ser feito a uma velocidade maior do que a da luz. Em particular, não pode ser feita instantaneamente.

O estado emaranhado diz unicamente que os spins das duas partículas de nosso exemplo continuam correlacionados, não importando quão distantes elas possam estar uma da outra; se uma é medida UP, a outra terá spin DOWN devido a esta correlação e a nada mais (não há velocidade supra lumina para informação, não há telepatia, nada!).

A resposta técnica a Einstein, que na literatura é denominada de **realismo local**, não é simples de ser dada e não a repetiremos aqui pois isso nos desviaria bastante, ainda que constitua o cerne da mecânica quântica atual, e foi alcançada a partir dos anos 1960 principalmente devido aos trabalhos de John S. Bell (1928-1990). O que resulta, e hoje em dia se tratando de um fato indubitável,⁹ é que a hipótese realista local de Einstein não pode ser mantida.

⁹Experimentos recentes sugerem que todos os *loopholes* (falhas) dos experimentos

Tem-se como certo que a mecânica quântica é não local, ou seja, que de fato acontecem coisas 'instantaneamente'. No entanto, a relatividade restrita não é afetada, porque nenhuma informação está sendo transmitida. O que se sabe em B a partir do que se realizou em A se deve à correlação quântica, um fenômeno que não tem contra-parte na física clássica e que tem chamado a atenção dos cientistas pelas suas possíveis aplicações, como em informação quântica, conforme já dito antes. Nas nossas referências bibliográfica, encontram-se discussões sobre esses temas.

Pelo menos você está vendo como a Álgebra Linear nos ajuda a encontrar uma descrição matemática para essa "confusão quântica".

ÁLGEBRA LINEAR E MECÂNICA QUÂNTICA

V EREMOS neste capítulo como a Álgebra Linear, mais especificamente a teoria dos espaços de Hilbert,¹ pode ser utilizada para uma formulação da mecânica quântica padrão (não-relativista). Não iremos muito longe, pois o assunto é amplo e necessitaria de uma discussão profunda, enquanto que aqui o que tencionamos é dar ao leitor uma ideia de como os instrumentos da Álgebra Linear são utilizados. Para um tratamento mais detalhado, recomendamos [Pru.81].

A sugestão dessa abordagem é devida a von Neumann. Dissemos *uma formulação* porque não há um único modo de apresentar a teoria, e escolhemos uma enfatizando o papel estrutural das teorias físicas. Detalhes sobre a abordagem à mecânica quântica envolvendo espaços de Hilbert podem ser encontrados na literatura; outros textos excelentes são [Hug.89], [Ish.95], [Red.87], [Gri.11].

Os postulados da MQ implicam que se conhecemos o estado do sistema em um tempo t , conheceremos o seu estado em qualquer outro tempo t' , anterior a t ou a ele posterior. Ou seja, o estado do sistema evolui **deterministicamente** pelas leis dinâmicas do sistema, centralizadas na célebre **equação**

¹A generalização da Álgebra Linear para o tratamento de espaços de dimensões infinitas requer um aparato matemático mais sofisticado, que aqui não será apresentado, como já alertado antes. Nesses casos, manteremos nossa descrição informal.

de Schrödinger. Esta equação, em uma de suas formas, pode ser apresentada como segue.

Há um operador unitário e limitado $U(t)$ que age sobre o vetor de estado em um tempo t , de forma que $|\psi(t)\rangle$ fica determinado desde que saibamos seu estado em um tempo inicial t_0 pela (uma das versões da) equação de Schrödinger

$$|\psi(t)\rangle = U(t)|\psi(t_0)\rangle. \quad (6.1)$$

Susskind e Friedman fazem uma observação importante, que citamos:

“...conhecer o estado quântico [de um sistema] não significa que possamos prever com certeza o resultado de um experimento. Por exemplo, saber que o estado de spin é $|r\rangle$ pode nos dizer do resultado de uma medição de σ_x , mas não nos diz nada acerca da medição de σ_z ou de σ_y . Por esta razão, a equação [(6.1) precedente] não corresponde ao determinismo clássico. O determinismo clássico nos permite prever os resultados dos experimentos. A evolução quântica dos estados permite-nos computar probabilidades dos resultados de experimentos *posteriores*.” [SusFri.14, p.98]

E eles continuam:

“Esta é uma das diferenças centrais entre as mecânicas clássica e quântica. Ela remonta à distinção entre estados e medidas (...). Na mecânica clássica, não há diferença real entre estados e medições. Na mecânica quântica, a diferença é profunda.” (ibid.)

Com efeito, é bom que o leitor tenha em mente uma das diferenças fundamentais entre as mecânicas clássica e quântica. Naquela, assume-se como um postulado que qualquer sistema físico tem suas propriedades assumindo valores bem determinados em qualquer instante de tempo. Todas elas. Se não conhecemos algum desses valores, isso se deve a problemas nossos ou de nossos aparelhos de medida, mas eles existem! É justamente aí que entram as probabilidades, como medida de nossa ignorância. Na quântica, como a frase citada acima sugere, isso não é assim: podemos saber o estado do spin porque o medimos σ_x , mas isso não implica que possamos saber σ_y ; aliás, *não podemos saber!* Pois não podemos conhecer valores de spin em direções distintas simultaneamente, uma vez que os operadores que representam os observáveis correspondentes

não comutam. Somente temos probabilidades, que entram na teoria não como medida da nossa ignorância, mas de modo essencial. Análises mais detalhadas, no entanto, exigem mais do que o exposto neste texto.

1. Ignorância do estado: operador de densidade

Acima consideramos unicamente os chamados **estados puros**, que desempenham o papel de concentrar toda a informação disponível sobre o sistema em um dado momento. No entanto, há também situações em que somos ignorantes sobre o estado do sistema, e o máximo que podemos fazer é atribuir uma probabilidade a cada um dos estados possíveis. Digamos que $|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_n\rangle$ sejam os estados possíveis e que temos uma fonte de partículas que pode oferecer como *output* esses estados, cada um com uma probabilidade w_i . Obtemos como estado do sistema algo do tipo

$$|\psi\rangle = w_1|\psi_1\rangle + w_2|\psi_2\rangle + \dots + w_n|\psi_n\rangle, \quad (6.2)$$

onde os w_j representam probabilidades (logo, $\sum_{i=1}^n w_i = 1$).

Tais estados são **misturas estatísticas** ou simplesmente **misturas**.² Tanto estados puros quanto misturas podem ser descritos por meio de **operadores densidade**. Vejamos do que se trata, ainda que sem os detalhes mais técnicos (neste ponto, basicamente seguimos a exposição em [Tor.99, p.345-7]).

Omitiremos a variável temporal t , e supomos que o sistema está em um estado puro representado por um vetor unitário (ou normalizado, como preferem os físicos) $|\psi\rangle$. O operador estatístico representando este estado é simplesmente o operador de projeção $E_{|\psi\rangle}$:

$$\rho = E_{|\psi\rangle} = |\psi\rangle\langle\psi|. \quad (6.3)$$

Seja agora $\{|\psi_i\rangle\}$ uma base ortonormal tal que $|\psi\rangle = \sum_i c_i|\psi_i\rangle$. Claro que sabemos que os c_i são os coeficientes de Fourier $c_i = \langle\psi_i|\psi\rangle$. Assim, a matriz representativa de ρ na base ortonormal dada tem os elementos (conforme página

²É interessante notar como agem os físicos. Essas probabilidades w_i que atribuímos são **probabilidades subjetivas**, ao passo que o cômputo das probabilidades pela Regra de Born é (supostamente) de uma probabilidade objetiva. Na verdade, a natureza da **probabilidade quântica** é algo ainda em aberto — ninguém sabe direito do que se trata.

53) $\rho_{ij} = \langle \psi_i | \rho | \psi_j \rangle = \langle \psi_i | \psi \rangle \langle \psi | \psi_j \rangle = c_i c_j^*$. Assim,

$$\text{Tr}(\rho) = \sum_i \rho_{ii} = \sum_i |c_i|^2 = 1, \quad (6.4)$$

tendo em vista que $|\psi\rangle$ é unitário. O valor esperado da medida de um observável A para o sistema no estado $|\psi\rangle$ pode ser expresso em termos do operador ρ como segue:

$$\begin{aligned} \langle A \rangle_{|\psi\rangle} &= \langle \psi | \hat{A} | \psi \rangle = \sum_{i,j} \langle \psi | \psi_i \rangle \langle \psi_i | \hat{A} | \psi_j \rangle \langle \psi_j | \psi \rangle = \\ &= \sum_{i,j} \langle \psi_j | \psi \rangle \langle \psi | \psi_i \rangle \langle \psi_i | \hat{A} | \psi_j \rangle = \sum_{i,j} \langle \psi_j | \rho | \psi_i \rangle \langle \psi_i | \hat{A} | \psi_j \rangle = \\ &= \sum_j \langle \psi_j | \rho \hat{A} | \psi_j \rangle = \text{Tr}(\rho \hat{A}). \end{aligned}$$

Ou seja,

$$\langle A \rangle_{|\psi\rangle} = \text{Tr}(\rho \hat{A}). \quad (6.5)$$

A descrição dada por meio do operador densidade, ou pela matriz densidade, o que resulta no mesmo, é mais geral e abrange todos os casos. O caso particular dos estados puros pode ser obtido quando se considera apenas um termo na decomposição 6.2. Na verdade, há até um critério para sabermos se o estado é puro: o traço de ρ^2 deve ser igual a 1; se for menor que 1, trata-se de uma mistura (mas não entraremos aqui nesses detalhes).

Vamos considerar um caso particular de duas partículas em estados possíveis $|\psi_1\rangle$ e $|\psi_2\rangle$ com probabilidades w_1 e w_2 respectivamente, e seja

$$|\psi\rangle = w_1 |\psi_1\rangle + w_2 |\psi_2\rangle.$$

Seja A um observável. Se o estado $|\psi\rangle$ for puro sabemos que $\langle A \rangle_{|\psi\rangle} = \langle \psi | \hat{A} | \psi \rangle$, e que no caso de uma mistura, temos $\langle A \rangle_{|\psi\rangle} = w_1 \langle A \rangle_{\psi_1} + w_2 \langle A \rangle_{\psi_2}$. Logo

$$\langle A \rangle_{|\psi\rangle} = w_1 \langle \psi_1 | \hat{A} | \psi_1 \rangle + w_2 \langle \psi_2 | \hat{A} | \psi_2 \rangle.$$

Mas também temos que $\langle \psi | A | \psi \rangle = \text{Tr}(\rho \hat{A})$, sendo $\rho = w_1 |\psi_1\rangle \langle \psi_1| + w_2 |\psi_2\rangle \langle \psi_2|$. O operador de densidade descreve toda a estatística da situação em que algumas medidas podem dar $|\psi_1\rangle$ e outras $|\psi_2\rangle$. Se o sistema envolve duas partículas

de spin- $1/2$ (como dois elétrons) e se podemos ter metade das vezes spin $1/2$ e metade $-1/2$ em uma direção, o operador de densidade será (com u para UP e d para DOWN)

$$\rho = \frac{1}{2}(|u\rangle\langle u|) + \frac{1}{2}(|d\rangle\langle d|) = \frac{1}{2}\mathbf{1},$$

expressão esta que carrega uma informação importante: $|u\rangle\langle u|$ é UP na direção (suponha z) e $|d\rangle\langle d|$ é DOWN *na mesma direção*.

No entanto, olhando para $\frac{1}{2}\mathbf{1}$, não obtemos informação alguma sobre a direção a ser medido o spin. Deste modo, podemos dizer que é possível tomar o spin em qualquer direção que se deseje, e o resultado será sempre o mesmo. Finalmente, observamos que a mistura

$$\rho = \frac{1}{2}(|u\rangle\langle u|) + \frac{1}{2}(|d\rangle\langle d|)$$

é completamente diferente do estado puro $\frac{1}{\sqrt{2}}(|u\rangle + |d\rangle)$. Frequentemente, diz-se que este último indica que metade das vezes o spin é UP e metade das vezes é DOWN, mas isso é equivocado.

Na verdade, como vimos na expressão (3.14), temos por exemplo

$$|u\rangle_x = |r\rangle = \frac{1}{\sqrt{2}}(|u\rangle_z + |d\rangle_z).$$

Assim, enquanto que $|u\rangle_z$ tem uma direção bem definida, o mesmo não se pode dizer de $\frac{1}{2}\mathbf{1}$.

2. O Princípio da Indeterminação

Nesta seção, vamos fazer uma digressão sobre um dos principais conceitos da mecânica quântica, introduzido por Werner Heisenberg em 1927, o Princípio da Indeterminação (antes, mas ainda muitas vezes, denominado de Princípio de Incerteza). Isso mostrará o uso de alguns conceitos introduzidos antes, como o comutador (seção 1.2), o valor esperado, e compatibilidade e incompatibilidade de observáveis.

Na mecânica quântica, contrariamente ao que ocorre na mecânica clássica, há observáveis que *não comutam*, ditos *incompatíveis*. Isso quer dizer, em linguagem informal, que eles não podem ser medidos simultaneamente. Exemplos típicos são a posição e o momento de uma partícula. Na mecânica

clássica, por outro lado, é um pressuposto que todos os observáveis podem assumir valores simultaneamente, e a medida de um deles não afeta as medidas dos demais. Mas, na mecânica quântica, se medimos a posição de uma partícula, depois medimos seu momento e depois novamente a posição, a segunda medida da posição pode fornecer um resultado diferente da primeira medida. Isso se deve ao fato de que a medida do momento coloca a partícula em um *auto-estado de momento* (ou seja, em um estado que é descrito por um auto-vetor do operador que corresponde ao momento), e isso deleta a informação antes obtida pela medida da posição. Isso ocorre também com o spin, por exemplo. Se medimos o spin em uma direção e depois em outra, uma nova medida do spin na primeira direção pode fornecer um resultado diferente da primeira, o que se atesta experimentalmente pelo chamado **experimento de Stern-Gerlach** [Alb.94, Cap.1], [Hug.89, Int.]. Mas voltemos à posição e ao momento.

No caso em que a medida de uma grandeza afeta a medida de uma outra, dizemos que elas são *incompatíveis*, e que são *compatíveis* em caso contrário. Matematicamente, temos a seguinte caracterização; as três afirmações a seguir são equivalentes:³

1. Os observáveis A e B são compatíveis.
2. Os operadores correspondentes \hat{A} e \hat{B} têm uma base de auto-vetores em comum.
3. Os operadores \hat{A} e \hat{B} comutam, isto é, $[\hat{A}, \hat{B}] = 0$.

O Princípio de Indeterminação entra justamente para dar uma ideia de *quanto* a medida de um observável pode influenciar a medida de outro. Colocando em outros termos, em vez de considerarmos uma sucessão de medidas dos observáveis A e B , ‘preparamos’ um grande número de sistemas de mesmo tipo (ou ‘idênticos’ no linguajar do físico) e medimos A para metade deles e B para a outra metade, o que fornece valores esperados $\langle A \rangle$ e $\langle B \rangle$, que terão *desvio padrão* σ_A e σ_B respectivamente.⁴ Então podemos falar em termos de certos estados para os quais A e B são ambos aproximadamente definidos, a saber, quando seu desvio padrão é pequeno. Os valores σ_A e σ_B podem ser nulos, no

³Para um argumento sobre a equivalência dessas três afirmativas, ver [Sch.15].

⁴O desvio padrão fornece a informação de quanto a medida se ‘desvia’ do valor esperado. Veja mais abaixo, equação 6.14.

caso em que o estado em questão é um auto-vetor do operador correspondente ao observável que está sendo medido. Se A e B são compatíveis, podemos encontrar auto-vetores para os quais tanto σ_A quanto σ_B são nulos (devido ao fato 2 acima), mas se eles são incompatíveis, então deve valer a seguinte forma geral do Princípio de Indeterminação:

$$\sigma_A \cdot \sigma_B \geq \left| \frac{1}{2i} \langle [\hat{A}, \hat{B}] \rangle \right|, \quad (6.6)$$

o qual diz que *não há* estado para o qual o produto dos desvios padrão seja menor do que a quantidade que está no lado direito da desigualdade acima. Para o caso especial de posição e momento, como $[X, P] = i\hbar$ (conforme a seção 1.2), a desigualdade anterior assume a conhecida forma

$$\sigma_X \cdot \sigma_P \geq \frac{\hbar}{2}. \quad (6.7)$$

No caso geral (6.6), o lado direito da desigualdade envolve o valor esperado do comutador dos observáveis A e B . Para o caso de eles serem compatíveis, $[\hat{A}, \hat{B}] = 0$. Para objetos macroscópicos, o valor \hbar torna-se negligenciável de forma que o Princípio de Indeterminação não coloca obstáculos, e assim até mesmo posição e momento podem ser medidos simultaneamente, como ocorre na mecânica clássica.

3. Uma formulação

Na linha de trabalho proposto por Patrick Suppes, (ver [KraAre.16]), apresentamos nesta seção uma versão axiomática para a mecânica quântica não relativista que leva em conta um conjunto de sistemas físicos.

Definição 3.1 (Mecânica quântica). Uma **mecânica quântica** não relativista é uma estrutura da forma

$$Q = \langle S, \{\mathcal{H}_i\}, \{\hat{A}_{ij}\}, \{\hat{U}_{ik}\}, \mathcal{B}(\mathbb{R}) \rangle, \text{ com } i \in I, j \in J, k \in K$$

onde:

1. S é uma coleção⁵ cujos elementos são chamados de *objetos físicos* ou *sistemas físicos*.

⁵Questiona-se de essa coleção pode ser considerada como sendo um *conjunto* no

2. $\{H_i\}$ é uma coleção de espaços de Hilbert separáveis cuja cardinalidade depende da particular aplicação em análise.
3. $\{\hat{A}_{ij}\}$ é uma coleção de operadores auto-adjuntos (ou hermitianos) sobre um espaço particular H_i .
4. $\{U_{ik}\}$ é uma coleção de operadores unitários sobre um espaço de Hilbert particular H_i .
5. $\mathcal{B}(\mathbb{R})$ é a coleção de borelianos da reta real.
6. Os postulados abaixo devem ser satisfeitos.

Por simplicidade, o leitor pode assumir que $\mathcal{B}(\mathbb{R})$ encerra uma coleção de intervalos da reta real nos quais podem ser encontrados os valores da medida dos observáveis, que são auto-valores de operadores auto-adjuntos (hermitianos), logo números reais.

Mais tecnicamente, ‘boreliano’ é o termo utilizado para designar subconjuntos de \mathbb{R} , ditos **conjuntos de Borel**, que têm as seguintes características, algo técnicas. Considere um conjunto qualquer S , e considere a coleção $\mathcal{B}(S)$ de subconjuntos de S tais que: (i) $S \in \mathcal{B}(S)$; (ii) se $A \in \mathcal{B}(S)$, seu complemento $S - A \in \mathcal{B}(S)$; (iii) se A_1, A_2, \dots é qualquer coleção contável⁶ de subconjuntos em $\mathcal{B}(S)$, então sua união $\bigcup_{i=1}^{\infty} A_i \in \mathcal{B}(S)$. Neste caso, $\mathcal{B}(S)$ é um *campo de Borel* de S . Se X é uma coleção qualquer de sub-conjuntos de S , o menor campo de Borel de S (no sentido de estar contido em todos os outros) que contém X é o *campo de Borel gerado por X* . Considere agora o campo de Borel do conjunto dos números reais, $\mathcal{B}(\mathbb{R})$, gerado pelos conjuntos abertos de \mathbb{R} . Os conjuntos de $\mathcal{B}(\mathbb{R})$ são chamados de **borelianos** de \mathbb{R} . As suas propriedades são úteis à matemática e à física quântica, mas não nos interessam aqui.

O formalismo dos espaços de Hilbert, cujos postulados veremos em momentos, não fala do espaço e nem do tempo. No entanto, para qualquer aplicação, esses conceitos devem constar de alguma maneira. Nosso modo de considerá-los é dado pela seguinte definição.

sentido usual da palavra. Um conjunto, como usualmente qualificado, é uma coleção de objetos *distintos* uns dos outros. No entanto, em se tratando de objetos quânticos, devido à possibilidade de eles serem indiscerníveis, isso pode ser problemático do ponto de vista dos fundamentos; para uma análise desta problemática, ver [FreKra.o6].

⁶Finito ou enumerável.

Definição 3.2. A cada sistema físico $s \in S$ associamos uma 4-upla da forma

$$\langle \mathbb{E}^4, \psi(\mathbf{x}, t), \Delta, P \rangle.$$

Aqui, \mathbb{E}^4 é o espaço-tempo de Newton-Galileo, o mesmo da mecânica clássica,⁷ mas o leitor que desconhece este conceito pode pensar no espaço vetorial \mathbb{R}^4 munido do produto interno canônico. Cada ponto deste espaço é descrito por uma quádrupla ordenada $(\mathbf{x}, t) = (x, y, z, t)$, onde $\mathbf{x} = (x, y, z)$ denota a tripla das coordenadas ‘espaciais’ do ponto (relativamente à base canônica) e t é um parâmetro que percorre a reta real ou um intervalo desta reta, e representa o tempo. O negrito em \mathbf{x} indica que se trata de um vetor. $\psi(\mathbf{x}, t)$ é uma função com domínio \mathbb{E}^4 , dita **função de onda** do sistema. Mais abaixo veremos como ela pode ser obtida. Por outro lado, $\Delta \in \mathcal{B}(\mathbb{R})$ é um boreliano, enquanto que P é uma função definida, para algum i (determinado pelo sistema s), em $\mathcal{H}_i \times \{\hat{A}_{ij}\} \times \mathcal{B}(\mathbb{R})$ e assumindo valores em $[0, 1]$. Os valores de $P(\psi, \hat{A}, \Delta) \in [0, 1]$ representam a probabilidade de que a medida de um observável A (representado por um operador auto-adjunto \hat{A}) para o sistema no estado $\psi(\mathbf{x}, t)$ esteja no boreliano Δ .

Podemos ver mais de perto a relação entre o vetor de estado e a função de onda do seguinte modo. Vamos considerar a situação de uma única partícula movendo-se ao longo do eixo x e seja \hat{X} o operador de posição, que como já dissemos tem um espectro contínuo.⁸ Deixemos (\mathbf{x}, t) representar a posição do sistema no tempo t . Subentendendo t , a posição é dada simplesmente por \mathbf{x} . Seja $\{|x_i\rangle\}$ uma base para o espaço de Hilbert dos estados de posição, que neste caso tem dimensão infinita — os físicos dizem que o sistema tem *infinitos graus de liberdade*. Ora, o estado $|\psi\rangle$ do sistema em um dado instante é combinação linear dos vetores da base, e como estamos em dimensão infinita, o somatório com o qual estamos acostumados dá lugar a uma integral, mas a ideia é a

⁷Na relatividade restrita, o espaço-tempo é denominado de *espaço-tempo de Minkowski*; na relatividade geral, a descrição é distinta e bem mais complexa.

⁸Muitas vezes, quando temos operadores com espectro contínuo, temos que fazer alguns truques matemáticos para ainda dizer que os operadores têm auto-vetores de modo a formarem uma base, o que nos exige trabalhar em espaços de Hilbert ‘equipados’, ditos **rigged Hilbert spaces** pelos físicos, que são obtidos adicionando-se à teoria dos espaços de Hilbert o que se denominam **distribuições**, como a **função delta de Dirac**. Mas pularemos esses detalhes mais técnicos aqui. O leitor interessado pode consultar [Mad.05], ou então as explicações dadas em [Hug.89, §1.5].

mesma:

$$|\psi\rangle = \int_i x_i |x_i\rangle dx,$$

onde $x_i = \langle x_i | \psi \rangle$ são os coeficientes de Fourier da expansão dada. Esses coeficientes são definidos como sendo as funções de onda (agora sem a notação de Dirac):

$$\psi(\mathbf{x}, t) = \langle x_i | \psi(t) \rangle.$$

Situações mais gerais envolvendo a função de onda relativa a outros operadores podem ser vistas em [SusFri.14, p.134].

Como veremos abaixo, o quadrado da função de onda é o mais importante: $|\langle x_i | \psi(t) \rangle|^2$ nos dá a probabilidade de encontrarmos a partícula em x_i . Como já deve ter ficado claro, isso é tudo o que temos: probabilidades.

3.1. Os postulados

Postulado 1 A cada sistema $s \in S$ associamos um espaço de Hilbert complexo separável $\mathcal{H} \in \{\mathcal{H}_i\}$. Sistemas compostos são associados a espaços de Hilbert que são produtos tensoriais dos espaços de cada um dos sistemas componentes.

Postulado 2 Os vetores dos sistemas unidimensionais de \mathcal{H} denotam os *estados* em que o sistema pode se apresentar. Esses espaços são chamados de *rays*. Para simplificar a notação, cada ray é representado por um vetor unitário que o gera, de modo que, por abuso de linguagem, podemos dizer, como fazem a grande parte dos textos de física, que os estados do sistema são descritos por vetores unitários do espaço de Hilbert. Esses vetores são denominados de **estados puros**. Estados de **mistura** foram mencionados acima, os quais são dados por matrizes (ou operadores) densidade.

Se $|\psi\rangle$ e $|\varphi\rangle$ denotam estados de um sistema σ , então qualquer combinação linear desses estados, ou seja, expressões da forma $a|\psi\rangle + b|\varphi\rangle$, para a e b números complexos, também representam estados do sistema. Tais estados são **superposições** dos estados dados. Esta hipótese (de que os vetores superpostos também denotam estados) é denominada de **Postulado da Superposição**.

Postulado 3 Cada observável físico A , como posição, momento, carga elétrica, spin, etc. é representado por um operador auto-adjunto \hat{A} sobre o espaço dos estados. O **Postulado da Quantização** diz que os valores possíveis da medida de um desses observáveis pertence ao seu **espectro**, a coleção de seus autovalores. Assim, para fazer jus à nossa estrutura introduzida acima, podemos dizer que a cada observável A associamos um operador $\hat{A} \in \{\hat{A}_{ij}\}$.⁹

Postulado 4 [Regra de Born] Dado um sistema $s \in S$, a ele associamos um espaço de Hilbert \mathcal{H} . Seja A um observável a ser medido quando o sistema está em um estado $|\psi\rangle$. Seja $\{|\alpha_n\rangle\}$ ($n = 1, 2, \dots$) uma base ortonormal para \mathcal{H} formada por auto-vetores do operador auto-adjunto \hat{A} que representa A . Denotemos por a_n os respectivos auto-valores, ou seja, $\hat{A}|\alpha_n\rangle = a_n|\alpha_n\rangle$. Então existem escalares (em geral, números complexos) c_j tais que $|\psi\rangle = \sum_j c_j |\alpha_j\rangle$, com $\sum_j |c_j|^2 = 1$. Além disso, sabemos o que são os c_j ; eles são os coeficientes de Fourier da expansão, ou seja, $c_j = \langle \alpha_j | \psi \rangle$. A **Regra de Born**, ou **Algoritmo Estatístico** nas palavras de Redhead [Red.87, p.8] (que estamos seguindo neste postulado), diz que a probabilidade de que a medida de A para o sistema no estado $|\psi\rangle$ seja um autovalor a_n (lembre que pelo **Postulado da Quantização** os valores possíveis estão no espectro de \hat{A}) é

$$\text{Prob}(a_n)_A^{|\psi\rangle} = \sum_{j/a_j=a_n} |c_j|^2 = \sum_{j/a_j=a_n} |\langle \alpha_j | \psi \rangle|^2. \quad (6.8)$$

O somatório corresponde à soma dos quadrados de todos os coeficientes iguais a a_n . Se há mais de um, isso quer dizer que o operador tem auto-valores repetidos, que são então considerados, e é dito ser um **operador degenerado**. Caso todos os auto-valores sejam distintos, o operador é **não degenerado** e neste caso haveria um único termo a ser considerado, e poderíamos escrever

$$\text{Prob}(a_n)_A^{|\psi\rangle} = |c_j|^2 = |\langle \alpha_j | \psi \rangle|^2. \quad (6.9)$$

Como já vimos antes, **misturas**, ou **mesclas estatísticas** aparecem quando somos ignorantes acerca do estado do sistema, ou seja, em qual estado entre os vetores do conjunto $\{|\psi_i\rangle\}$ ($i = 1, \dots, k$) está o sistema. Para dar uma melhor

⁹Repare o que os índices estão dizendo: há uma classe de operadores auto-adjuntos indexada por j que tem vínculo com o espaço de Hilbert dado, que foi indexado por i . O mesmo vai acontecer abaixo com os operadores unitários.

descrição, não custa dizer o que acontece nesses casos, mesmo que saltemos as explicações detalhadas. Expressamos nossa ignorância pela introdução de ‘pesos probabilísticos’ w_i (que são probabilidades subjetivas) para indicar a probabilidade de encontrarmos o sistema em um dos estados suspeitos. Claro que devemos ter $\sum_{i=1}^k w_i = 1$. Suponha que $\Delta = \{a_{m_1}, a_{m_2}, \dots\}$ é um subconjunto de auto-valores de \hat{A} . Então a **Regra de Born** vai dizer que a probabilidade da medida de A para o sistema no estado $|\psi\rangle$ é dada por

$$\text{Prob}(\Delta)_A^{|\psi\rangle} = \sum_{i=1}^k w_i \cdot \sum_{j/a_n=a_{m_1}, a_{m_2}, \dots} |\langle \alpha_j | \psi_j \rangle|^2. \quad (6.10)$$

Da mesma forma obtemos a probabilidade da medida estar em Δ para o caso de estados puros (não misturas), simplesmente ignorando os pesos w_i acima, ou seja,

$$\text{Prob}(\Delta)_A^{|\psi\rangle} = \sum_{j/a_n=a_{m_1}, a_{m_2}, \dots} |\langle \alpha_j | \psi_j \rangle|^2. \quad (6.11)$$

Este postulado permite ainda que indiquemos a forma de se calcular o **valor esperado** da medida de A para o sistema no estado $|\psi\rangle$, a saber, a expressão que já é nossa conhecida,

$$\langle A \rangle_{|\psi\rangle} = \sum_j a_j |c_j|^2 = \langle \psi | \hat{A} | \psi \rangle. \quad (6.12)$$

A notação é uma simplificação para $\langle \psi | \hat{A} | \psi \rangle$. O que esta expressão está nos dizendo? Se arrumarmos a notação, pondo

$$\hat{A}|\psi\rangle = \sum_{j=1}^n c_j A|q_j\rangle = \sum_{j=1}^n c_j a_j |\alpha_j\rangle,$$

vem que

$$\langle \psi | \hat{A} | \psi \rangle = \left\langle \sum c_j |\alpha_j\rangle \left| \sum c_j a_j |\alpha_j\rangle \right\rangle = \sum c_j^* c_j a_j \langle \alpha_j | \alpha_j \rangle = \sum |c_j|^2 a_j.$$

Portanto,

$$\langle A \rangle_{|\psi\rangle} = \sum_{j=1}^n a_j |c_j|^2. \quad (6.13)$$

Resulta da definição acima que podemos definir uma quantidade que representa a *incerteza* na medida de um observável A , ou o *desvio padrão* de A ,¹⁰ a saber,

$$\sigma_A = \sqrt{\langle A^2 \rangle - \langle A \rangle^2}, \quad (6.14)$$

onde $A^2 = \hat{A}\hat{A}$. O desvio padrão de A indica o quanto a medida varia ('dispersão estatística') em relação ao valor esperado. Se o valor é baixo, indica-se que há uma 'tendência' a que os valores obtidos estejam próximos à média.¹¹

Postulado 5 [Postulado da Dinâmica] Se o sistema em um instante t_0 está em um estado $|\psi(t_0)\rangle$, então em um tempo distinto t ele evolui (evolução unitária) para um estado $|\psi(t)\rangle$ de acordo com a **equação de Schrödinger**

$$\psi(t) = \hat{U}(t)\psi(t_0), \quad (6.15)$$

onde $\hat{U} \in \{\hat{U}_{ik}\}$ é um operador unitário.

Se $\{U(t)\}$ é uma família de operadores unitários sobre um espaço de Hilbert \mathcal{H} , dados em função de um parâmetro real t relativamente ao qual são contínuos (que entenderemos como representando o tempo), então se esses operadores satisfazem a condição $U(t_1 + t_2) = U(t_1)U(t_2)$ para todos $t_1, t_2 \in \mathbb{R}$, pode-se mostrar que existe um único operador H satisfazendo

$$U(t) = e^{-iHt},$$

para todo $t \in \mathbb{R}$, sendo $e^{iHt} = \cos(tH) - i \sin(tH)$. O operador H é o **operador hamiltoniano**, e representa a energia do sistema. Uma das formulações da equação de Schrödinger pode ser dada agora em função do hamiltoniano, da seguinte forma:

$$-i\hbar \frac{\partial |\psi\rangle}{\partial t} = H|\psi\rangle, \quad (6.16)$$

onde i é a nossa conhecida unidade imaginária e $\hbar = \frac{h}{2\pi}$ é a **constante reduzida de Planck** (a constante de Planck é h).

¹⁰Intuitivamente, representa o quanto os valores das medidas realizadas —recorde que em física praticamente nunca se realiza uma só medição— se distribuem em torno do valor $\langle A \rangle_{|\psi\rangle}$.

¹¹Em estatística, o desvio padrão de uma *variável aleatória* X é definido como $\sigma_X = \sqrt{\langle X^2 \rangle - \langle X \rangle^2}$.

Há casos particulares desta equação, como exemplificaremos a seguir, mas antes vamos recordar de algo já dito antes à página 37 e mais acima. Na versão da mecânica quântica que estamos delineando, que é denominada de **mecânica de ondas** (devida a Schrödinger), o estado de um sistema físico é descrito, como vimos à página 37, por uma entidade matemática chamada **função de onda**, $\psi(\mathbf{x}, t)$, sendo $\mathbf{x} = (x, y, z)$ uma variável que representa as coordenadas espaciais do sistema, enquanto que t representa o tempo, a qual satisfaz a equação mais geral

$$\int_{-\infty}^{\infty} |\psi(\mathbf{x}, t)|^2 dx = 1. \quad (6.17)$$

Repare que o somatório foi aqui substituído por uma integral porque o observável físico em análise é a *posição* do sistema (que pode ser qualquer ponto sobre o eixo real), logo trata-se de um operador com **espectro contínuo**. Ademais, nesta mecânica, como se vê pela equação de Schrödinger, é o *estado* do sistema que evolui, pois é ele que está sendo derivado relativamente ao tempo. Isso é relevante porque há (pelo menos) uma outra interpretação devida a Heisenberg, na qual a evolução é descrita em termos da evolução dos *operadores*.¹² Não abordaremos a versão de Heisenberg aqui, mas é relevante observar que Schrödinger mostrou, como se acredita em geral, que as duas mecânicas são equivalentes.¹³

Para t fixado, $\psi(\mathbf{x}, t)$ é um elemento de um espaço de Hilbert separável isomorfo a \mathcal{L}^2 visto acima, e que pode ser tomado como sendo ele mesmo. A equação 6.17 é denominada de **condição de normalização**. Eliminado as variáveis espacial e temporal por simplicidade, podemos reescrevê-la assim:

$$\|\psi\|^2 = \int_{-\infty}^{\infty} \psi^* \psi dx = \int_{-\infty}^{\infty} |\psi|^2 dx = 1.$$

Na teoria, $\psi(\mathbf{x}, t)$ satisfaz a **equação de Schrödinger**, dependente do tempo (dada a variável temporal que estamos utilizando). A interpretação dessa função, devida a Max Born diz que dado um intervalo qualquer Δ da reta real¹⁴

¹²Os livros de mecânica quântica, em geral escritos em inglês, denominam essas duas abordagens de *Schrödinger's picture* e *Heisenberg's picture* respectivamente.

¹³Na verdade, a equivalência de duas teorias requer que os axiomas de uma sejam provados como teoremas da outra. O procedimento de Schrödinger foi outro; ver [Cas.og].

¹⁴À página 37 foi denominado de intervalo $[a, b]$.

(um **boreliano**, o quadrado da função representa a **probabilidade** de encontrarmos o sistema dentro de Δ , o que podemos escrever assim:

$$\text{Prob}(\Delta) = \int_{\Delta} |\psi(\mathbf{x}, t)|^2 dx. \quad (6.18)$$

O fato da equação 6.17 ser igualada à unidade indica que, em algum ponto da reta real (tomada agora toda ela como sendo Δ), o sistema está *com certeza* (probabilidade igual a 1). Finalmente, observa-se que o fato da equação 6.18 não se alterar quando substituirmos $\psi(\mathbf{x}, t)$ por $k \cdot \psi(\mathbf{x}, t)$, com $k \in \mathbb{C}$, desde que $|k| = 1$, mostra que a probabilidade é conservada no tempo.

O caso particular que dissemos que iríamos exemplificar diz respeito a um só sistema quântico, que pode ser uma partícula de massa m , que se desloca por força de um potencial V na direção do eixo dos x . Como temos uma só coordenada, não precisaremos utilizar a notação vetorial $\mathbf{x} = (x, y, z)$, assumindo que temos simplesmente a coordenada x dependendo do tempo, logo, se quisermos, $x(t)$, que representa a posição da partícula no tempo t . O modo de determinar $x(t)$ é resumidamente o seguinte [Gri.11, p.1].

Na mecânica clássica, aplicamos a segunda lei de Newton, $F = ma$, a força F pode ser dada em função do potencial por $F = -\frac{\partial V}{\partial x}$, de modo que obtemos

$$-\frac{\partial V}{\partial x} = m \cdot \frac{d^2 x}{dt^2}.$$

Com as condições iniciais que fornecem posição e velocidade em um instante t_0 , obtemos $x(t)$. E lembre que a derivada segunda da posição (relativamente ao tempo fornece a aceleração, enquanto que a derivada primeira fornece a velocidade).

Na mecânica quântica, o estado é descrito pela função de onda $\psi(x, t)$ (já considerando uma só coordenada de posição), obtida resolvendo-se uma equação engenhosamente criada por Schrödinger (a história, ou talvez estória, desta equação é bem interessante; ver [Kum.09, Cap.9]),¹⁵ que é um caso particular das equações do mesmo nome apresentadas acima (e escrevendo somente ψ para $\psi(x, t)$):

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2} + V\psi. \quad (6.19)$$

¹⁵Não vou apresentar a ‘dedução’ desta equação aqui, porque há vários sites no YouTube que fazem isso muito bem.

Como na equação de Newton, se soubermos as condições iniciais $\psi(x, t_0)$, podemos determinar $\psi(x, t)$ para qualquer tempo t anterior ou posterior a t_0 . Como fica a interpretação estatística?

Já sabemos que o quadrado da função ψ nos dá uma probabilidade; mais precisamente, para este caso,

$$\int_a^b |\psi(x, t)|^2 dx$$

nos dá a **probabilidade de encontrar a partícula entre a e b no instante t** . Deste modo, a probabilidade é a área abaixo do gráfico de $|\psi|^2$ entre os pontos a e b no eixo x .

Uma ‘pegadinha’ comum é perguntar como a equação de Schrödinger se aplicaria a fótons, que são partículas *sem massa*, ou sejam com massa igual a zero. Deste modo, $m = 0$ e não podemos ter uma quantidade nula no denominador da equação 6.19. O problema se resolve atentando-se para o fato de que os fótons não são tratados pela mecânica quântica não relativista, mas pelas teorias quânticas de campos, e o que é zero é sua *massa de repouso*. Fótons são absorvidos e emitidos livremente, não havendo conservação de seu número, e a mecânica quântica não relativista trata somente de sistemas com um número determinado e fixo de sistemas físicos. Os fótons têm uma *massa relativista*, que resulta da equação relativista $E^2 = (m_0 c^2)^2 + p^2 c^2$, onde m_0 é a massa de repouso, p é o momento, c a velocidade da luz no vácuo e E a energia. Se $m_0 = 0$ no caso do fóton, isso se reduz a $E^2 = p^2 c^2$. Tendo em vista a célebre equação de Einstein $E = mc^2$, temos então que $mc^2 = pc$ e portanto $m = \frac{p}{c}$, onde m é a *massa relativista*. Assim, temos que a massa (relativista) de um fóton é proporcional ao seu momento. A equação de Schrödinger aplica-se a partículas não relativistas e com massa não nula.

Postulado 6 [Postulado do Colapso] Sendo A um observável e $|\psi\rangle = \sum_j c_j |\alpha_j\rangle$ o estado do sistema, se a medida fornece o autovalor a_n , imediatamente após a medida o sistema entra (colapsa para) no estado $|\alpha_n\rangle$ com probabilidade $|c_n|^2 = |\langle \psi_n | \psi \rangle|^2$.

Importante observar o caráter **não determinista** do colapso. Há no entanto interpretações que evitam falar no colapso, encontrando outras formas de explicação. Uma delas, de grande repercussão na atualidade, é a **interpretação dos muitos mundos** que prefere dizer que, quando da medida, o mundo se

bifurca, criando-se dois mundos reais mas incomunicáveis. Por mais estranha que esta interpretação possa parecer, tem muitos adeptos.

Para sistemas compostos, envolvendo mais de um sistema físico, o espaço de Hilbert adotado é um produto tensorial dos espaços de Hilbert de cada um dos sistemas componentes, e adota-se o mesmo espaço de Hilbert no caso de partículas 'idênticas', ou seja, se temos três partículas de mesma espécie, como três elétrons, e se \mathcal{H} é o espaço de Hilbert cujos vetores representam o estado de uma delas, então os estados do sistema composto pelas três é $\mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H}$. Usando um tal espaço, os postulados acima são assumidos da mesma forma.

Daqui para a frente, a mecânica quântica não-relativista pode ser desenvolvida, como os textos que indicamos na Bibliografia deixam claro. Tal desenvolvimento, no entanto, não é nossa finalidade aqui. Deixamos então a sequência a cargo do leitor, com a esperança de que o que se viu acima possa servir como uma introdução ao assunto.

Referências Bibliográficas

- [Alb.94] Albert, D. (1994), *Quantum Mechanics and Experience*. Cambridge and Massachusetts: Harvard Un. Press.
- [Ari.53] Aristotle, (1853), *Organon*, Vol.I. Trad. Octavius Freire Owen. London: H. G. Bohn, pp. 244-356.
<https://archive.org/details/organonorlogica00porpgoog>
- [Bar.76] Barsotti, L. (1976), *Álgebra Linear*. Curitiba: A. M. Cavalvante & Cia. 2a. ed.
- [Bel.87] Bell, J. S. (1987), Bertlmann's socks and the nature of reality. In Bell, J. S., *Speakable and Unsayable in Quantum Mechanics*. Cambridge: Cambridge Un. Press., Chap. 16.
- [BirNeu.36] Birkhoff, G. and von Neumann, J. (1936), 'The logic of quantum mechanics', *Annals of Mathematics*, vol. 37, pp. 823-843.
- [Bun.73] Bunge, M. (1973), *Philosophy of Physics*. Dordrecht and Boston: D. Reidel.
- [Cas.08] Casado, C. M. M. (2008), A brief history of the equivalence between the two mechanics. *Lat. Am. J. Phys. Educ.* 2 (2): 152-5.
- [Cor.92] Corry, L. (1992), Nicolas Bourbaki and the concept of mathematical structure. *Synthese* 92 (3): 315-48.
- [DalGiuGre.04] Dalla Chiara, M.L., Giuntini, R. and Greechie, R. (2004), *Reasoning in Quantum Theory: Sharp and Unsharp Quantum Logics*. Dordrecht/Boston/London: Kluwer Ac. Pu. (Trends in Logic - *Studia Logica Library* Vol. 22).

- [RonDomFre.16] de Ronde, C., Domenech, G. and Freytes, H. (2016), Quantum logic in historical and philosophical perspective. *Internet Encyclopedia of Philosophy*.
- [FreKra.06] French, S. and Krause, D., (2006), *Identity in Physics: a Historical, Philosophical and Formal Analysis*, Oxford: Oxford Un. Press.
- [Gri.11] Griffiths, D. J. (2011), *Mecânica Quântica*. 2a. ed. São Paulo: Pearson Prentice Hall.
- [HeiZim.12] Heinosaari, T. and Ziman, M. (2012), *The Mathematical Languages of Quantum Theory: From Uncertainty to Entanglement*. Cambridge: Cambridge Un. Press.
- [Hen.79] Henkin, L. (1979), 'Verdade e demonstrabilidade', in S. Morgenbasser (org.), *Filosofia da Ciência*, São Paulo: Cultrix, pp. 53-64.
- [Hug.89] Hughes, R. I. G. (1989), *The Structure and Interpretation of Quantum Mechanics*. Cambridge, MA; London, UK.
- [HofKun.79] Hoffman, K. e Kunze, R., (1979), *Álgebra Linear*. Rio de Janeiro: LTC, 2a. ed.
- [Ish.95] Isham, C. (1995), *Lectures on Quantum Theory*. London: Imperial College Press.
- [Jec.77] Jech, T. (1977), About the axiom of choice. In Barwise, J. (ed.), *Handbook of Mathematical Logic*. North-Holland, pp. 345-70.
- [Kle.52] Kleene, S. C. (1952), *Introduction to Metamathematics*. New York: North-Holland.
- [Kro.12] Kronz, F. and Lupher, T. (2012), Quantum Theory: von Neumann vs. Dirac, *The Stanford Encyclopedia of Philosophy* (Summer 2012 Edition).
- [Kra.02] Krause, D. (2002), *Introdução aos Fundamentos Axiomáticos da Ciência*. S. Paulo, EPU.
- [Kra.16] Krause, D. (2016), *Introdução à Lógica Quântica*, em preparo.
- [Kra16a] Krause, D. (2016), *Tópicos em Ontologia Analítica*. S. Paulo: Ed.UNESP.

- [KraAre.16] Krause, D. and Arenhart, J. R. B. (2016), *The Logical Foundations of Scientific Theories: Languages, Structures, and Models*. London: Routledge (no prelo).
- [Kum.09] Kumar, M. (2009), *Quantum*. London: Icon Books.
- [LevBal.96] Levy-Leblond, J. -M. and Balibar, F. (1996), *Quantics: Rudiments of Quantum Physics*. Amsterdam: North-Holland.
- [Mad.05] de la Madrid, R. (2005), The role of the rigged Hilbert space in quantum mechanics. <http://arxiv.org/abs/quant-ph/0502053v1>
- [Mer.85] Mermin, N. D. (1985), Is the moon there when nobody looks? Reality and the quantum theory. *Physics Today* Abril: 38-47.
- [PavMeg.99] Pavičić, M. and Megill, N. D. (1999), Non-orthomodular models for both standard quantum logic and standard classical logic: repercussions for quantum computers. <http://lanl.arxiv.org/abs/quant-ph/9906101v3>
- [Pru.81] Prugovečki, E., *Quantum Mechanics in Hilbert Space*. 2nd. ed., New York: Academic Press.
- [Red.87] Redhead, M. (1987), *Incompleteness, Nonlocality and Realism: A Prolegomenon to the Philosophy of Quantum Mechanics*. Oxford: Clarendon Press.
- [SapAlc.15] Sapunaru, R. A. e Alcaraz, J. (2015), Correspondência entre Leibniz e Varignon sobre o cálculo e a álgebra. *Peri* 7 (2): 202-10.
- [Sch.15] Schroeder, D. V. (2015), Compatible and incompatible observables. Em <http://physics.weber.edu/schroeder/quantum/Compatible.pdf>
- [Sty.02] Styer, D. F. et al. (2002), 'Nine formulations of quantum mechanics'. *Am. J. Phys.* 70 (3): 288-97.
- [Sup.57] Suppes, P. (1957), *Introduction to Logic*. Van Nostrand Co.
- [SusFri.14] Suskind, L. and Friedman, A. (2014), *Quantum Mechanics: A Theoretical Minimum*. New York: Basic Books.

[Sze.04] Szekeres, P. (2004), *A Course in Modern Mathematical Physics*. Cambridge: Cambridge Un. Press.

[Tor.99] Torretti, R. (1999), *The Philosophy of Physics*. Cambridge: Cambridge Un. Press.

Índice Remissivo

- K*-espaço vetorial, 3
- quod erat demonstrandum*, 4
- rays, 38
- quanton, 100

- adição de vetores, 2
- adjunto de um operador, 76
- Algoritmo da Quantização, 69
- Algoritmo Estatístico, 116
- Aristóteles
 - ciência de princípios, 28
- auto-adjunto, 78
- auto-valores, 67
- auto-vetores, 67
- automorfismo, 79
- axioma da escolha, 14, 21, 27, 60

- Balibar, F., 100
- barra de Halmos, \square , 4
- base canônica, 14
- base de Hamel, 27
- base ortonormal, 41
- bases, 13
 - existência de, 21
 - ortonormais, 22
- Bell, J.S., 101, 104
- Bertlmann, R., 101
- Birkhoff, G., 41
- Bohr, N., 101, 102
 - e a mecânica quântica, 100
- Boole, G.
 - álgebra de, 39

- boreliano, 113, 119
- Born, M., 102, 116, 119
 - interpretação probabilista, 45
- Bourbaki, N., 1
- Bunge, M., 100
- bósons, 93, 94

- Cantor, G., 27
- ciência de princípios, 28
- colapso, 103
- colapso do vetor de estado, 99
- combinação linear, 8
- combinação linear infinita, 16
- complemento ortogonal, 87, 88
- computação quântica, 28
- computador, 57, 110
- condição de normalização, 119
- conjunto escolha, 27
- conjuntos de Borel, 113
- constante de Planck, 58, 118
- contradição, 99
 - o que é, 99
- coordenadas de um vetor, 17
- coordenadas polares, 35
- Copenhague
 - escola de, 102
- corpo, 2
 - postulados de, 2
- correlação quântica, 101, 105

- delta de Kronecker, 43
- densidade de probabilidade, 46

- dependência linear, 8
- descrição definida, 92
- desigualdade de Cauchy-Schwarz, 33
- desvio padrão, 111
 - definição, 117
- determinismo, 106
- diagonalização, 72
- dimensão, 14
- Dirac, P.A.M., 47
 - debate sobre o rigor, 25
 - função delta, 25, 114
 - importância da notação de, 61
 - notação de, 61
- dispersão estatística
 - veja* desvio padrão 117
- distribuições, 114
- distância, 34

- Einstein, A., 104
 - equação de, 121
 - visão realista, 101
- emaranhamento, 94, 95, 98
 - fascínio do, 100
- equação de Schrödinger, 106, 118, 119
- espaço de Hilbert, x, 35
- espaço de operadores, 53
- espaço dos polinômios, 17
- espaço dual, 61
- espaço euclidiano, 97
- espaço gerado, 13
- espaço linear
 - veja* espaço vetorial 2
- espaço separável, 36
- espaço vetorial
 - definição, 1, 2
- espaço-tempo, 113
- espaço-tempo Newtoniano, 97
- espaços de dimensão infinita, 16, 26
- espaços de Hilbert *rigged*, 114
- espaços vetoriais
 - isomorfismo entre, 23

- espectro contínuo, 114, 118
- espectro de um operador, 67, 115
- estado emaranhado, 103, 104
- estado fatorável, 95
- estado produto, 95
- estado puro
 - critério para saber, 109
- estado separável, 95
- estado singlete, 94, 96
- estados puros, 108, 115
- estrutura, 1
- Euler, L.
 - fórmula de, 35
- evolução unitária, 118
- experimento de Stern-Gerlach, 111
- experiência de pensamento, 98

- fator de fase, 38
- fator de normalização, 94
- formalismo, 37
- Fourier, J.
 - coeficientes de, 42–44, 65, 82, 116
 - série de, 42
- franjas de interferência, 34
- funcional linear, 47, 53, 59
- função
 - bijetiva, 23, 25, 57
 - descontínua de primeira espécie, 41
 - descontínua de segunda espécie, 41
 - escolha, 28
 - injetiva, 25, 56, 57
 - seccionalmente contínua, 41
 - sobrejetiva, 25, 56, 57
- função de onda, 45, 114, 118, 120
- função delta, 25
- função seccionalmente contínua, 42
- função traço, 31, 59, 86
- funções
 - ortonormais, 42
- funções de Kronecker, 37, 44
- funções quadrado integráveis, 37, 45

- férmions, 93, 94
- gato de Schrödinger, 98
- Goudsmit, S., 63
- Grassmann, H., 1
- grupo, 5, 32
- grupo comutativo, 5
- Heisenberg's picture, 119
- Heisenberg, W., 40, 65, 102
 - mecânica de matrizes, 119
 - Princípio de Indeterminação, 59, 110
 - relação de comutação, 58
- hipótese realista, 104
- holismo, 101
- igual por definição, 5
- independência linear, 8
- infinito
 - diferentes tipos de, 27
- informação quântica, 101
- interpretação dos muitos mundos, 121
- inverso de um operador, 54
- involução, 99
- isometria, 75
- isomorfismo, 23
- Jordan, P., 102
- kernel, 54
- kets, 2
- Kronecker, L.
 - delta de, 43
 - funções de, 44
- lei de composição externa, 3
- lei distributiva
 - não validade da, 28, 40
- lema de Tukey, 22
- Lema de Zorn, 21, 22, 28, 45
- Levy-Leblond, J. -M., 100
- localidade, 104
- loopholes, 104
- lógica da mecânica quântica, 28
- lógica clássica, 28, 39
- lógica intuicionista, 4
- lógica quântica, 28, 38, 39, 41
- lógica subjacente, 27
- massa de repouso, 121
- massa relativista, 121
- matriz de mudança de base
 - veja* matriz de mudança de coordenadas 18
- matriz de mudança de coordenadas, 18
- matriz diagonalizável, 73
- matriz hermitiana, 78
- matriz inversível, 19
 - condição necessária e suficiente, 19
- matriz ortogonal, 75, 76
- matriz representativa, 53
- matriz simétrica, 16
- matriz singular, 68
- matriz transposta, 16, 31
- matriz unitária, 75, 76
- matrizes de Pauli, 69, 70, 72
- matrizes semelhantes, 69, 72
- mecânica de ondas, 118
- mecânica quântica, 106
 - caracterização, 112
 - postulados, 115
- mecânicas clássica e quântica
 - diferenças, 107
- medida de um observável
 - valor possível, 69
- medição, 102
- mistura, *veja* misturas estatísticas
- misturas estatísticas, 108, 109, 116
- Newton, I.
 - segunda lei de, 120
- norma, 32
- notação de Dirac, 2, 47, 92
- noções absolutas, 97

- núcleo de um operador, 54
- o espaço $\ell^2(\mathbb{N})$, 37
- o espaço \mathcal{L}^2 , 45
- o reticulado da mecânica quântica, 40
- observáveis, 58
 - compatibilidade, 110
 - incompatibilidade, 110
 - que não comutam, 110
- observáveis compatíveis
 - critérios, 111
- observável posição, 96
- operador
 - densamente definido, 77
- operador auto-adjunto, 76, 78, 79
- operador contínuo, 59
- operador de densidade, 108, 109
- operador de permutação, 93
- operador de projecção, 82
- operador degenerado, 69, 116
- operador densidade, 108
- operador diagonalizável, 73
- operador hamiltoniano, 118
- operador hermitiano, 78
- operador identidade, 58
- operador inversível, 54
- operador limitado, 59
- operador linear, 48
 - representação matricial, 50
- operador maximal, 69
- operador não limitado, 77
- operador não-degenerado, 69
- operador ortogonal, 75
- operador unitário, 75
- operadores auto-adjuntos, 58
- operadores de posição e momento, 77
- operadores de projecção, 80
- operadores hermitianos, 58
- operação binária, 2
- oposto
 - unicidade do, 4
- paradoxo de Skolem, 27
- partículas idênticas, 92
- Pauli, W., 63, 102
 - Princípio de Exclusão, 94
- piloto automático
 - a matemática como, 18
- polinômio característico, 68
- posição e momento, 58
- Postulado da Indistinguibilidade, 93
- Postulado da Quantização, 115, 116
- Postulado da Superposição, 115
- Postulado do Colapso, 121
- postulados da mecânica quântica, 115
- Princípio de Incerteza
 - veja* Princípio de Indeterminação 110
- Princípio de Indeterminação, 40, 59, 97, 110, 111
 - caso geral, 111
 - para posição e momento, 112
- Princípio Maximal, 22
- probabilidade, 46, 119
- probabilidade quântica, 108
- probabilidade subjetiva, 108
- Processo de Gram-Schmidt, 42, 44, 83
- produto de matrizes, 31
- produto de operadores, 54
- produto interno, 29
- produto interno canônico, 30
- produto tensorial, 90
- produto tensorial de espaços vetoriais, 90
- produto tensorial de operadores, 91
- produto tensorial de vetores, 90
- produtos internos, 29
- projecção ortogonal, 44
- projecção sobre uma coordenada, 59
- projeções, *veja* operador de projecção, 82
 - notação para, 84, 85
- pré-espaco de Hilbert, 29
- quanton, 100
- RAA, 97–99

- Regra auto-valores e auto-vetores, 96
 realidades independentes, 101
 realismo, 104
 realismo local, 104
 Redhead, M., 69
 redução ao absurdo, 4
 redução ao absurdo clássica, 5
 redução ao absurdo intuicionista, 5
 Regra de Born, 93, 108, 116
 relatividade geral, 38
 relatividade restrita, 38
 relação de equivalência, 23
 relógio universal, 97
 resolução da identidade, 86
 restrição de uma operação, 10
 reticulado ortomodular, 40, 41, 89
 Schrödinger's picture, 119
 Schrödinger, E., 95
 e a principal característica da MQ, 95
 equação de, 106
 gato de, 98
 mecânica de ondas, 118
 Schwartz, L., 25
 sequência convergente, 35
 sequência de Cauchy, 35
 sesquilinear, 30
 sistemas indiscerníveis, 92
 soma de sub-espacos, 80
 soma direta, 80, 81
 spin, 40, 63, 102
 sub-espaco fechado, 36
 sub-espaco gerado pela união, 12
 sub-espaco invariante, 82
 sub-espacos
 condição necessária e suficiente para, 10
 sub-espacos fechados, 38
 sub-espacos vetoriais, 10
 superposição, 29, 34
 veja combinação linear 8
 superposições, 115
 Suppes, P., 1, 112
 série de Fourier, 41
 Teorema de Löwenheim-Skolem, 27
 Teorema de Riesz, 60
 significado do, 61
 Teorema de Steinitz, 15
 Teorema do Núcleo e da Imagem, 55
 Teorema dos Eixos Principais, 73
 Teorema Espectral, 84
 teoria da informação quântica, 28
 teoria quântica de campos, 28, 38, 121
 termo de interferência, 34
 trajetórias de partículas, 98
 transformação linear, 48
 traço de um operador, 53
 Uhlenbeck, G., 63
 valor esperado, 53, 93, 109–111, 117
 variedade diferenciável, 97
 variáveis ocultas, 64
 variável aleatória, 117
 vetor nulo, O , 3
 vetores ortogonais, 41
 $\alpha \perp \beta$, 41
 vetores ortonormais, 41
 von Neumann, J., 37, 41, 106
 debate sobre o rigor, 24
 Zermelo-Fraenkel
 teoria de conjuntos, 27