Mecânica Quântica para Matemáticos em Formação

Publicações Matemáticas

Mecânica Quântica para Matemáticos em Formação

Bárbara Amaral
UFOP/UFMG

Alexandre Tavares Baraviera
UFRGS

Marcelo O. Terra Cunha
UFMG

impa

28º Colóquio Brasileiro de Matemática

Impresso no Brasil / Printed in Brazil

Capa: Noni Geiger / Sérgio R. Vaz

28º Colóquio Brasileiro de Matemática

- Cadenas de Markov y Teoría de Potencial - Johel Beltrán
- Cálculo e Estimação de Invariantes Geométricos: Uma Introdução às Geometrias Euclidiana e Afim - M. Andrade e T. Lewiner
- De Newton a Boltzmann: o Teorema de Lanford - Sérgio B. Volchan
- Extremal and Probabilistic Combinatorics - Robert Morris e Roberto Imbuzeiro Oliveira
- Fluxos Estrela - Alexander Arbieto, Bruno Santiago e Tatiana Sodero
- Geometria Aritmética em Retas e Cônicas - Rodrigo Gondim
- Hydrodynamical Methods in Last Passage Percolation Models - E. A. Cator e L. P. R. Pimentel
- Introduction to Optimal Transport: Theory and Applications - Nicola Gigli
- Introdução à Aproximação Numérica de Equações Diferenciais Parciais Via o Método de Elementos Finitos - Juan Galvis e Henrique Versieux
- Matrizes Especiais em Matemática Numérica - Licio Hernanes Bezerra
- Mecânica Quântica para Matemáticos em Formação - Bárbara Amaral, Alexandre Tavares Baraviera e Marcelo O. Terra Cunha
- Multiple Integrals and Modular Differential Equations - Hossein Movasati
- Nonlinear Equations - Gregorio Malajovich
- Partially Hyperbolic Dynamics - Federico Rodriguez Hertz, Jana Rodriguez Hertz e Raúl Ures
- Random Process with Variable Length - A. Toom, A. Ramos, A. Rocha e A. Simas
- Um Primeiro Contato com Bases de Gröbner - Marcelo Escudeiro Hernandes

ISBN: 978-85-244-327-9

Distribuição: IMPA
Estrada Dona Castorina, 110
22460-320 Rio de Janeiro, RJ
E-mail: ddic@impa.br
http://www.impa.br

Para Thales e Tshabalala (o cão), pelo carinho, pela lealdade, pelo companheirismo e também pelas bochechas.

Para Áurea, Dirceu, Flávia e Pedro, que agora ganha mais um livrinho para puxar da estante.

Para Mimi e Tatá, como sempre, e para o Andrey, pela primeira vez.

Programa

Abertura

O texto que você está lendo agora é o resultado de uma pequena aventura ou uma grande ambição: falar de mecânica quântica para matemáticos em formação. Da nossa experência, matemáticos se formam sem qualquer conhecimento de mecânica quântica. Quando, por interesse próprio, vão procurar tal formação, por razões históricas ou disponibilidade de textos[1], acabam esbarrando com textos que ou assumem, ou iniciam a discussão por análise funcional. Mas as últimas décadas permitiram o crescimento da chamada *teoria quântica da informação*, ou, como é mais comum, *informação quântica*. Um dos maiores méritos desta foi levar a uma revisão dos conceitos fundamentais da mecânica quântica e, em especial, permitir uma maior valorização dos espaços de estado de dimensão finita. Dessa forma, sai a análise funcional (como pré-requisito ou ponto de partida) e entra a álgebra linear, com a qual os estudantes têm contato desde o início de seus cursos. Esse é o espírito do texto: discutir a matemática da mecânica quântica, principalmente em dimensão finita.

Por escolha, o texto foi divido em três partes, usando uma metáfora musical. O *prelúdio* apenas prepara a obra. Não falamos de mecânica quântica nele, embora, naturalmente, tudo que lá se encontra ou tem aplicação na ou sustenta a nossa solista. O *interlúdio* é a essência do texto. É nele que a mecânica quântica é introduzida. A abordagem escolhida vai do particular para o geral, em busca da melhor compreensão. O *poslúdio* trata de alguns temas que gostaríamos de aprofundar mais, embora nem o formato nem os prazos tornem isso adequado.

[1] O que não são, de modo algum, razões independentes.

Ao escrever este livro, tínhamos em mente nosso público alvo: estudantes com o ciclo básico completo, e com gosto pela matemática. Não há necessidade de passar por todo o prelúdio, caso você queira "ir direto ao assunto". Ele foi escrito com vários objetivos complementares: tornar o texto razoavelmente autocontido[2], introduzir notação (caso especial da notação de Dirac, intensamente utilizada no capítulo 2) e discutir alguns conceitos (ou estratégias de apresentar os conceitos) que normalmente não encontram lugar no ciclo básico pressuposto. Uma sugestão razoável é que você corra os olhos pelo índice e escolha como se servir. Por outro lado, pensando nesse mesmo público, o texto é repleto de exercícios. Há dois tipos deles, os que se encontram em meio ao texto e os de final de capítulo. Isso não acontece por acaso. Ao encontrar um exercício no meio do texto, resolva-o; ou, ao menos, tente. Quase certamente ele será utilizado logo em seguida.

Cabe salientar que, normalmente, em um bacharelado em Física os estudantes tomam cerca de três disciplinas de Mecânica Quântica, enquanto este livro é originalmente destinado a um minicurso. Portanto, embora as definições básicas e suas consequências sejam apresentadas, há muito mais que não poderá ser discutido. Para isso, o estudante pode adotar textos que capricham na intuição, como [FLS], ou textos mais tradicionais, como [CDL], mais profundos [Per95], ou mais relacionados à informação quântica [Pre, NC].

Quando resolvemos encarar essa empreitada, já tínhamos experiências (razoavelmente) recentes complementares: uma dissertação de mestrado, [Ama], um livro sobre informação quântica, [Ter07a], e um livro mais introdutório sobre mecânica quântica, [Bar]. Não foi possível resistir a uma pequena dose de autoplagia e um leitor mais atento vai encontrar trechos previamente publicados. Em algumas outras partes, já temos dificuldade de lembrar quem fez a primeira redação de tal parágrafo. Assim, a responsabilidade pelos erros[3] encontrados no texto é compartilhada pelos três autores.

Entretanto, é um prazer agradecer a algumas pessoas que ajudaram a diminuir a quantidade destes erros e ainda contribuiram com sugestões. Nessa função, ainda que atropelados pela sobreposição

[2]Também se buscou usar a nova ortografia, mas não podemos garantir que tenhamos tido sucesso.

[3]E eventuais acertos.

de versões sempre incompletas, merecem destaque Gláucia Murta, Rodrigo Porto, Fernando Brandão, Ricardo Falcão, Pierre-Louis de Assis, Raphael Drumond, Carlos Felipe Lardizábal, Mateus Araújo Santos e Marco Túlio Coelho Quintino.

Um prazer ainda maior é agradecer ao Artur O. Lopes, que tanto incentivou os dois autores mais idosos deste texto, e, indiretamente, a mais jovem. Capes, CNPq e Fapemig também merecem reconhecimento pelo apoio dado aos autores ao longo dos anos.

Por fim, é hora e lugar de agradecermos e nos desculparmos com aqueles entes próximos e queridos, que concordaram com tantas renúncias em nome do livro que, finalmente, ganhou forma. Também agradecemos e nos desculpamos com os organizadores do Colóquio, que apoiaram essa iniciativa e gentilmente compreenderam as nossas fraquezas.

Bárbara Amaral

Alexandre T. Baraviera

Marcelo Terra Cunha

Prelúdio

Antes de realmente focarmos na mecânica quântica, vamos discutir alguns conceitos matemáticos que permeiam o restante do texto. Naturalmente, não nos cabe aprofundamento em cada um desses temas. Assim, esse prelúdio é visto como um momento para fixar notação e colecionar os conteúdos de maneira adequada a referências rápidas.

Começamos por números complexos, apenas colecionando suas principais propriedades e pedindo ao estudante que as relembre (ou eventualmente aprenda algumas) através dos exercícios.

Álgebra linear é a base da mecânica quântica. Por isso ganha papel de destaque nesse prelúdio.

Equações diferenciais e grupos também merecem atenção. E não resistimos à tentação de apresentar as álgebras C^*, que acreditamos desconhecidas da maioria de nossos leitores, mas que podem ser muito úteis na discussão da mecânica quântica, além de possuirem beleza intrínseca que nos atrai.

Capítulo 1

Números Complexos

O conjunto dos números complexos tem um universo infinito de aplicações. Em muitos casos eles podem facilitar os cálculos e abreviar a notação. A Mecânica Quântica faz uso dos números complexos, mas aqui eles não são só um atalho para simplificar a teoria. A importância deles é tamanha que alguns físicos afirmam que é impossível formulá-la utilizando apenas os números reais. Faremos aqui apenas um resumo das principais propriedades que serão necessárias ao longo do texto e para mais detalhes o leitor pode consultar [Soa].

1.1 Soma e Multiplicação

Definição 1.1. *Um corpo é um conjunto C em que podemos definir duas operações*

$$+ : C \times C \longrightarrow C$$
$$(a, b) \longmapsto a + b$$
$$\cdot : C \times C \longrightarrow C$$
$$(a, b) \longmapsto a \cdot b = ab$$

tais que para todos $a, b, c \in C$ valem

1) (Associatividade) $a + (b + c) = (a + b) + c$ e $a \cdot (b \cdot c) = (a \cdot b) \cdot c$;

2) (Comutatividade) $a + b = b + a$ e $a \cdot b = b \cdot a$;

3) (Existência de elemento neutro) existem elementos distintos $0 \in C$ *e* $1 \in C$ *tais que* $a + 0 = a$ *e* $a \cdot 1 = a$;

4) (Existência de inversos) Para todo $a \in C$ *existe* $-a \in C$ *tal que* $a + (-a) = 0$ *e se* $a \neq 0$ *existe* $a^{-1} \in C$ *tal que* $a \cdot a^{-1} = 1$;

5) (Distributividade) $a \cdot (b + c) = a \cdot b + a \cdot c$.

Exemplo 1.1. *O conjunto dos números racionais* $\mathbb{Q}$ *e o conjunto dos números reais* $\mathbb{R}$ *são corpos com as operações usuais de soma e multiplicação.*

O corpo que vai aparecer com mais frequência ao longo do texto é o corpo dos números complexos e por isso vamos fazer uma breve revisão de suas principais propriedades.

Definição 1.2. *Um número complexo é uma expressão do tipo:*

$$z = x + iy,$$

em que x *e* y *são números reais e* i*, chamado unidade imaginária, satisfaz a propriedade* $i^2 = -1$*. O número* $x = \mathrm{Re}(z)$ *é a parte real de* z *e* $y = \mathrm{Im}(z)$ *é a parte imaginária de* z*.*

Para definir a soma e a multiplicação de números complexos vamos usar as operações de soma e multiplicação de números reais e considerar cada número complexo como um polinômio em i, de modo que a soma de dois números complexos $z_1 = x_1 + iy_1$ e $z_2 = x_2 + iy_2$ é dada por

$$z_1 + z_2 = (x_1 + x_2) + i(y_1 + y_2),$$

e o produto de z_1 e z_2 é dado por

$$z_1 z_2 = x_1 x_2 + i x_1 y_2 + i x_2 y_1 + i^2 y_1 y_2 = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1). \tag{1.1}$$

Exercício 1.1. *Mostre que as operações definidas acima são comutativas e que a multiplicação se distribui sobre a adição. Mostre também que o elemento neutro para a adição é* $0 = 0 + i0$*, que o elemento neutro para a multiplicação, também chamado de identidade, é* $1 = 1 + i0$ *e que o inverso de* z *para a soma é* $-z = -x - iy$*.*

Para mostrar que $\mathbb{C}$ é um corpo, resta mostrar que existem os inversos multiplicativos.

Definição 1.3. *O conjugado de um número complexo $z = x + iy$ é o número complexo $\bar{z} = x - iy$. A norma de z é $|z| = \sqrt{z \cdot \bar{z}} = \sqrt{x^2 + y^2}$. Um número complexo z é chamado unitário se $|z| = 1$.*

Exercício 1.2. *Mostre que a norma de um número complexo é sempre um número real não negativo e temos que $|z| = 0$ se, e somente se, $z = 0$.*

Exercício 1.3. *Mostre que $z^{-1} = \frac{\bar{z}}{|z|^2}$ é o inverso multiplicativo de z e que se z é unitário, $z^{-1} = \bar{z}$.*

1.2 Representação Geométrica

Podemos representar os números complexos geometricamente usando o plano cartesiano. O número complexo $z = x+iy$ é representado pelo ponto (x, y) no plano cartesiano e $|z|$ representa a distância euclidiana entre o ponto $(0, 0)$ e (x, y).

A partir da representação geométrica podemos ver que se $r = |z|$ e ϕ é o ângulo formado entre a reta que liga os pontos (x, y) e $(0, 0)$ e o eixo x então

$$z = r(cos(\phi) + isen(\phi)).$$

Desse modo, se z é um complexo unitário então $z = cos(\phi) + isen(\phi)$ para algum $\phi \in \mathbb{R}$.

1.3 A Exponencial Complexa

Algumas funções definidas para números reais podem ser facilmente generalizadas para $\mathbb{C}$. Entre elas está a função exponencial.

Definição 1.4. *A exponencial de um número complexo z é definida por*

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}.$$

A exponencial está bem definida para todo número complexo. Isso segue do fato de que

$$\sum_{n=0}^{\infty} \frac{|z^n|}{n!} = \sum_{n=0}^{\infty} \frac{|z|^n}{n!} = e^{|z|}$$

e do seguinte resultado:

Proposição 1.1. *Se uma série de números complexos*

$$\sum_{n=0}^{\infty} z_n$$

é absolutamente convergente, ou seja, se

$$\sum_{n=0}^{\infty} |z_n|$$

converge, então ela é convergente.

Para números complexos sem parte real, chamados de imaginários puros, é possível mostrar, utilizando a definição acima, que a exponencial assume a forma

$$e^{yi} = cos(y) + isen(y),$$

de modo que a exponencial de um imaginário puro é sempre um número complexo unitário.

Voltando à representação geométrica dos números complexos, obtemos a representação polar de um número complexo:

$$z = r(cos(\phi) + isen(\phi)) = re^{i\phi}.$$

Também é possível mostrar que valem as seguintes propriedades

1. $e^{z+w} = e^z \cdot e^w$, para todos $z, w \in \mathbb{C}$;
2. $e^{-z} = \frac{1}{e^z}$;
3. $e^0 = 1$
4. $(e^z)^n = e^{nz}$, para todo $z \in \mathbb{C}$ e $n \in \mathbb{Z}$.
5. $e^z \neq 0$.

Exercício 1.4. *Prove que* $e^z = e^x(cos(y) + isen(y))$.

1.4 Limites e Derivadas

Outros conceitos do cálculo também podem ser generalizados para o caso complexo.

Definição 1.5. *Dado um número complexo z_0 dizemos que o número w_0 é o limite de uma função $f : \mathbb{C} \to \mathbb{C}$ quando z tende a z_0 e escrevemos*

$$\lim_{z \to z_0} f(z) = w_0$$

se para todo $\epsilon > 0$ é possível encontrar $\delta > 0$ tal que se $0 < |z - z_0| < \delta$ então $|f(z) - w_0| < \epsilon$.

Exercício 1.5. *Prove que se $f_1 : \mathbb{C} \to \mathbb{C}$ e $f_2 : \mathbb{C} \to \mathbb{C}$ são funções tais que $\lim_{z \to z_0} f_1(z) = w_1$ e $\lim_{z \to z_0} f_2(z) = w_2$ e se $c \in \mathbb{C}$ então*

1. $\lim_{z \to z_0}(f_1(z) + f_2(z)) = w_1 + w_2$;
2. $\lim_{z \to z_0}(cf_1(z)) = cw_1$;
3. $\lim_{z \to z_0}(f_1(z)f_2(z)) = w_1 w_2$;
4. *Se $w_1 \neq 0$,* $\lim_{z \to z_0} \frac{1}{f_1(z)} = \frac{1}{w_1}$.

Definição 1.6. *Dizemos que uma função $f : \mathbb{C} \to \mathbb{C}$ é contínua no ponto z_0 se*

$$\lim_{z \to z_0} f(z) = f(z_0).$$

Definição 1.7. *Seja $f : \mathbb{C} \to \mathbb{C}$ e $z_0 \in \mathbb{C}$. Se existir o limite*

$$f'(z_0) = \lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

dizemos que ele é a derivada de f em z_0. Usamos também a notação

$$f'(z_0) = \frac{df}{dz}(z_0).$$

Exercício 1.6. *Prove que se $f_1 : \mathbb{C} \to \mathbb{C}$ e $f_2 : \mathbb{C} \to \mathbb{C}$ são funções que possuem derivada em z_0 então*

1. $(f_1 + f_2)'(z_0) = f_1'(z_0) + f_2'(z_0)$;

2. $(cf_1)'(z_0) = cf_1'(z_0)$;

3. $(f_1 f_2)'(z_0) = f_1'(z_0)f_2(z_0) + f_1(z_0)f_2'(z_0)$;

4. *Se* $w_1 \neq 0$, $\left(\frac{1}{f_1}\right)'(z_0) = -\frac{f_1'(z_0)}{f_1(z_0)^2}$.

Exercício 1.7. *Mostre que se* $f : \mathbb{C} \to \mathbb{C}$ *possui derivada em* z_0 *e se* $g : \mathbb{C} \to \mathbb{C}$ *possui derivada em* $f(z_0)$ *então* $g \circ f$ *possui derivada em* z_0 *e*

$$(g \circ f)'(z_0) = g'(f(z_0))f'(z_0).$$

Para que uma função complexa $f(x+iy) = f(x,y) = u(x,y) + iv(x,y)$ possua derivada em um ponto, é necessário que as funções u e v satisfaçam condições chamadas condições de Cauchy-Riemann. Essa é uma maneira prática de mostrar que uma função complexa não possui derivada: basta mostrar que essas condições não são satisfeitas, o que em geral não é difícil. Infelizmente essas condições não são suficientes para a existência da derivada. No entanto, no caso especial em que as derivadas parciais de u e v são contínuas, essas condições são suficientes e podemos usá-las tanto para mostrar a existência da derivada, quanto para calculá-la.

Proposição 1.2 (Condições de Cauchy-Riemann)**.** *Seja* $z = x + iy$ *e* $f(z) = f(x,y) = u(x,y) + iv(x,y)$. *Se* f *tem derivada no ponto* $z_0 = x_0 + iy_0$ *então valem as condições de Cauchy-Riemann:*

$$\frac{\partial u}{\partial x}(x_0, y_0) = \frac{\partial v}{\partial y}(x_0, y_0)$$

$$\frac{\partial u}{\partial y}(x_0, y_0) = -\frac{\partial v}{\partial x}(x_0, y_0),$$

e além disso

$$f'(z_0) = \frac{\partial u}{\partial x}(x_0, y_0) + i\frac{\partial u}{\partial y}(x_0, y_0).$$

Por outro lado, se as derivadas parciais

$$\frac{\partial u}{\partial x}(x_0, y_0), \frac{\partial u}{\partial y}(x_0, y_0), \frac{\partial v}{\partial x}(x_0, y_0), \frac{\partial v}{\partial y}(x_0, y_0)$$

são contínuas em z_0 *e se as condições de Cauchy-Riemann são satisfeitas, então* f *possui derivada em* z_0.

Exercício 1.8. *Mostre que*

$$\frac{d(z^n)}{dz} = nz^{n-1}$$

$$\frac{d(e^z)}{dz} = e^z.$$

Terminamos aqui nosso breve resumo sobre números complexos. Há bem mais o que se estudar e o leitor interessado no cálculo em uma variável complexa pode procurar as referências.

1.5 Exercícios

Exercício 1.9. *Coloque na forma $x + iy$.*

1. $(3 - 5i)(2 + i)$;
2. $(1 - i)^2 - 6i$;
3. $\frac{(1-2i)^2}{2+2i}$.

Exercício 1.10. *Esboce no plano cartesiano os subconjuntos de $\mathbb{C}$ que satisfazem as seguines propriedades*

1. $|z| = 2$;
2. $|z| = |z + 1|$;
3. $\mathrm{Re}(z) = \mathrm{Im}(z + 1)$.

Exercício 1.11. *Calcule*

1. $e^{1+3\pi i}$;
2. $e^{\frac{3-\pi i}{2}}$;

Exercício 1.12. *Verifique que a as condições de Cauchy-Riemann são satisfeitas para a função*

$$f(x + iy) = \begin{cases} \frac{xy(1-i)}{x^2+y^2}, & z \neq 0, \\ 0 & z = 0. \end{cases}$$

em $z = 0$ mas que ela não possui derivada nesse ponto.

Exercício 1.13. *Mostre se as funções abaixo possuem derivada em todos os pontos*

1. $f(x+iy) = e^{-y}(\cos(x) + i\ \mathrm{sen}(y))$;
2. $f(x+iy) = e^{-x}(\cos(y) - i\ \mathrm{sen}(x))$.

Capítulo 2

Álgebra Linear

Neste capítulo pretendemos relembrar ao leitor algumas noções básicas sobre espaços vetoriais e produtos internos que serão muito utilizadas no decorrer do texto [NC, Lim, Vai]. Falaremos principalmente de espaços vetoriais complexos, que aparecem naturalmente em mecânica quântica.

2.1 Espaços Vetoriais

Um espaço vetorial V sobre um corpo C é um conjunto, cujos elementos chamaremos vetores e denotaremos por $|u\rangle$, munido de uma soma vetorial

$$\begin{array}{rcl} + : V \times V & \longrightarrow & V \\ (|u\rangle, |v\rangle) & \longmapsto & |u\rangle + |v\rangle \end{array}$$

e de um produto por escalar

$$\begin{array}{rcl} \cdot : C \times V & \longrightarrow & V \\ (\lambda, |u\rangle) & \longmapsto & \lambda|u\rangle \end{array}$$

tais que para todos $|u\rangle, |v\rangle, |w\rangle \in V$ e $\lambda, \nu \in C$ temos

1) (Associatividade) $|u\rangle + (|v\rangle + |w\rangle) = (|u\rangle + |v\rangle) + |w\rangle$;

2) (Comutatividade) $|u\rangle + |v\rangle = |v\rangle + |u\rangle$;

3) (Existência de zero) Existe vetor $0 \in V$ tal que $|u\rangle + 0 = |u\rangle$;

4) (Existência do vetor oposto) Dado $|u\rangle \in V$ existe vetor $-|u\rangle \in V$ tal que $|u\rangle + (-|u\rangle) = 0$;

5) (Associatividade) $\lambda(\nu|u\rangle) = (\lambda\nu)|u\rangle$;

6) (Distributividade) $\lambda(|u\rangle + |v\rangle) = \lambda|u\rangle + \lambda|v\rangle$;

7) (Distributividade) $(\lambda + \nu)|u\rangle = \lambda|u\rangle + \nu|u\rangle$;

8) $1|u\rangle = |u\rangle$ quando 1 é a unidade da multiplicação no corpo C.

A notação utilizada acima, a notação de Dirac, é bastante empregada, sobretudo pelos físicos que trabalham com a mecânica quântica. O símbolo $|u\rangle$ é chamado *ket* u. é importante ressaltar que o u que aparece na notação é apenas um rótulo arbitrário. Outra observação importante é que o símbolo $|0\rangle$ será usado com frequência e não representa o vetor nulo do espaço vetorial em questão e sim um vetor com o rótulo zero. Denotaremos o vetor nulo apenas pelo símbolo 0.

Exemplo 2.1 ($\mathbb{R}^n$ é um espaço vetorial sobre $\mathbb{R}$)**.** *$\mathbb{R}^n$ é o conjunto das n-uplas ordenadas $(x_1, \ldots, x_n)$, $x_i \in \mathbb{R}$. Podemos definir a soma e produto por escalar, respectivamente, por*

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$$

$$\lambda(x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n).$$

Exemplo 2.2 (O conjunto das funções contínuas é um espaço vetorial sobre $\mathbb{R}$)**.** *Seja $C_{\mathbb{R}}[0,1]$ o conjunto formado pelas funções contínuas do intervalo $[0,1]$ com valores em $\mathbb{R}$. A soma e o produto podem ser definidos como sendo*

$$(f + g)(x) = f(x) + g(x)$$

$$(\lambda f)(x) = \lambda f(x).$$

Exemplo 2.3 ($\mathbb{C}^n$ é um espaço vetorial sobre $\mathbb{C}$)**.** *$\mathbb{C}^n$ é o conjunto das n-uplas ordenadas $(x_1, \ldots, x_n)$, $x_i \in \mathbb{C}$. Podemos definir a soma e produto por escalar, respectivamente, por*

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$$

$$\lambda(x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n).$$

Exercício 2.1. *Mostre que os exemplos acima são de fato espaços vetoriais com as operações indicadas.*

No último exemplo o corpo usado foi $\mathbb{C}$. No restante deste texto o corpo *sempre* será o dos números complexos, salvo menção explícita em contrário. Em boa parte do texto o leitor também pode imaginar que o espaço vetorial em questão é $\mathbb{C}^n$.

2.2 Base e Dimensão

Definição 2.1. *Dizemos que uma expressão do tipo*

$$\alpha_1|v_1\rangle + \ldots + \alpha_k|v_k\rangle, \ \alpha_i \in C$$

é uma combinação linear dos vetores $|v_1\rangle, \ldots, |v_k\rangle$. Dado um conjunto de vetores dizemos que ele gera V se todo elemento de V pode ser escrito como combinação linear dos elementos desse conjunto.

Queremos encontrar conjuntos que gerem o espaço com o número mínimo de elementos.

Definição 2.2. *Dizemos que um conjunto de vetores $\{|v_1\rangle, \ldots, |v_k\rangle\} \subset V$ é linearmente independente (LI) se a equação*

$$\alpha_1|v_1\rangle + \cdots + \alpha_k|v_k\rangle = 0$$

só admite a solução trivial $(\alpha_1, \ldots, \alpha_k) = (0, \ldots, 0)$.

Caso contrário, dizemos que os vetores são linearmente dependentes (LD).

Exercício 2.2. *Mostre que um conjunto de vetores é LD se e somente se podemos expressar ao menos um dos vetores como combinação linear dos outros.*

Podemos nos perguntar se existe um conjunto de vetores LI de forma que todo elemento do espaço V possa ser escrito como combinação linear dos elementos desse conjunto. é possível mostrar que todo espaço vetorial possui um conjunto LI com essa propriedade.

Definição 2.3. *Uma base para um espaço vetorial V é um conjunto LI*

$$\mathcal{B} = \{|v_1\rangle, \ldots, |v_k\rangle\}$$

tal que todo vetor de V é combinação linear de $|v_1\rangle, \ldots, |v_k\rangle$. A dimensão de V é o número de vetores em uma base.[1]

é possível mostrar que duas bases de V devem ter necessariamente o mesmo número de elementos, de modo que a dimensão está bem definida, ou seja, não depende da base que escolhemos para V.

Da existência de uma base $\mathcal{B} = \{|v_1\rangle, \ldots, |v_n\rangle\}$ do espaço V surge a notação para vetores mais utilizada: dado um vetor $|v\rangle$ podemos escrevê-lo como $|v\rangle = a_1|v_1\rangle + \cdots + a_n|v_n\rangle$ e de forma única. De fato, se temos $|v\rangle = b_1|v_1\rangle + \ldots + b_n|v_n\rangle$ então

$$(a_1 - b_1)|v_1\rangle + \ldots + (a_n - b_n)|v_n\rangle = 0$$

e da condição LI temos $a_i = b_i$. Assim podemos representar o vetor por meio de seus coeficientes na base dada: $|v\rangle = (a_1, \ldots, a_n)_{\mathcal{B}}$. Quando não houver confusão a respeito da base que está sendo utilizada denotaremos apenas por $|v\rangle = [v]_{\mathcal{B}} = (a_1, \ldots, a_n)$.

2.3 Subespaços Vetoriais

Um subespaço vetorial S do espaço vetorial V é um subconjunto de V que é, ele mesmo, um espaço vetorial com as operações de soma e multiplicação por escalar definidas em V. Para isso, precisamos que as seguintes propriedades sejam satisfeitas

[1] A definição de dimensão acima vale para espaços vetoriais que podem ser gerados por um número finito de vetores, como é o caso dos exemplos 2.1 e 2.3. Em outros casos, como no exemplo 2.2, o espaço vetorial não pode ser gerado por nenhum conjunto finito e dizemos que a dimensão é infinita.

- $0 \in S$;
- $|x\rangle + |y\rangle \in S$ para todo par $|x\rangle$ e $|y\rangle \in S$;
- $\lambda|x\rangle \in S$ para todo $\lambda \in C$ e todo $|x\rangle \in S$.

Exercício 2.3. *Considere o subconjunto*

$$S = \{(t, 0, \ldots, 0) \in \mathbb{R}^n; t \in \mathbb{R}\}.$$

Mostre que S é um subespaço vetorial de $\mathbb{R}^n$.

Exercício 2.4. *Considere o subconjunto*

$$S = \{f \in C_{\mathbb{R}}[0,1]; f(0) = f(1) = 0\}.$$

Mostre que S é um subespaço vetorial de $C_{\mathbb{R}}[0,1]$.

2.4 Transformações Lineares

Sejam U e V espaços vetoriais. Uma aplicação $T\colon U \to V$ é dita uma transformação linear se dados $|u_1\rangle, |u_2\rangle \in U$ e $\lambda \in \mathbb{C}$ (que é o corpo que usaremos ao longo do texto) temos

- $T(\lambda|u_1\rangle) = \lambda T(|u_1\rangle)$
- $T(|u_1\rangle + |u_2\rangle) = T(|u_1\rangle) + T(|u_2\rangle)$

Fixemos uma base $\mathcal{B} = \{\, |e_1\rangle, \ldots, |e_m\rangle\}$ de V e $\mathcal{F} = \{|f_1\rangle, \ldots, |f_n\rangle\}$ de U. Podemos escrever um vetor $|v\rangle \in V$ na forma $|v\rangle = \sum_{i=1}^{m} v_i|e_i\rangle$, que também podemos representar na forma matricial

$$|v\rangle = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix}_{\mathcal{B}} = [v_1, v_2, \ldots, v_m]^t_{\mathcal{B}}.$$

Daí

$$T(|v\rangle) = T\left(\sum_{i=1}^{m} v_i|e_i\rangle\right) = \sum_{i=1}^{m} v_i T(|e_i\rangle) = \sum_{i=1}^{m}\sum_{j=1}^{n} v_i T_{ji}|f_j\rangle$$

em que os números T_{ji} são tais que $T(|e_i\rangle) = \sum_{j=1}^{n} T_{ji}|f_j\rangle$. Portanto podemos representar a transformação linear T por meio de uma matriz $T_{\mathcal{B},\mathcal{F}}$ com entradas T_{ij} de forma que

$$[T(|v\rangle)]_{\mathcal{F}} = T_{\mathcal{F}}^{\mathcal{B}}[v]_{\mathcal{B}}.$$

De forma similar, dada uma matriz $n \times m$ temos, desde que fixadas as bases, uma transformação linear associada. Essa identificação é tão forte que frequentemente nos referiremos à uma transformação linear apenas pela matriz que a representa, desde que não haja confusão acerca de quais são as bases usadas em cada caso.

Quando $U = V = \mathbb{C}^n$, as matrizes em questão são matrizes $n \times n$. O conjunto das matrizes $n \times n$ com coeficientes em $\mathbb{C}$ será denotado por $M_n(\mathbb{C})$.

Um caso particular de destaque são os funcionais lineares.

Definição 2.4. *Um funcional linear $\langle\chi|$ é uma transformação linear $\langle\chi| : V \to \mathbb{C}$. O espaço de todos os funcionais lineares de V é conhecido com o espaço dual de V e denotado por V^*.*

Os elementos de V^* serão denotados na notação de Dirac pelo símbolo $\langle\chi|$, que é chamado de *bra*.

2.5 Produto Interno

Dado um espaço vetorial V, um produto interno é uma aplicação

$$\begin{array}{rcl} \langle\cdot|\cdot\rangle : V \times V & \longrightarrow & \mathbb{C} \\ (|u\rangle, |v\rangle) & \longmapsto & \langle u|v\rangle \end{array}$$

satisfazendo as seguintes propriedades[2]: para todo $|u\rangle, |v\rangle, |w\rangle \in V$ e $\lambda, \mu \in \mathbb{C}$

1. $\langle \lambda u + \mu v|w\rangle = \bar{\lambda}\langle u|w\rangle + \bar{\mu}\langle v|w\rangle$;
2. $\langle u|v\rangle = \overline{\langle v|u\rangle}$;
3. $\langle u|u\rangle \geq 0$;

[2]Utilizaremos a notação $\langle \lambda u + \mu v|w\rangle$ para denotar o produto interno entre os vetores $\lambda|u\rangle + \mu|v\rangle$ e $|w\rangle$

4. Se $\langle u|u\rangle = 0$ então $|u\rangle = 0$.

Observação 1. *Em 2 a barra denota a operação de tomar o complexo conjugado do número; em 3, note que o lado esquerdo da expressão de fato é real como consequência de 2. Por último, note que*

$$\langle u|\lambda v\rangle = \overline{\langle \lambda v|u\rangle} = \overline{\bar{\lambda}\langle v|u\rangle} =$$

$$\lambda\overline{\langle v|u\rangle} = \lambda\langle u|v\rangle.$$

Por isso uma maneira usual de reescrever a condição 1 acima é

$$\langle u|\lambda v + \mu w\rangle = \lambda\langle u|v\rangle + \mu\langle u|w\rangle.$$

Exercício 2.5. *Considere o espaço vetorial $\mathbb{C}^n$ (sobre o corpo $\mathbb{C}$). Mostre que a aplicação*

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

é um produto interno, conhecido como produto interno canônico de $\mathbb{C}^n$.

Exercício 2.6. *Mostre que $C_{\mathbb{C}}[0,1]$, o conjunto de funções contínuas do intervalo $[0,1]$ com valores em $\mathbb{C}$, é um espaço vetorial sobre $\mathbb{C}$ com soma e produto, respectivamente, definidos como sendo*

$$(f+g)(x) = f(x) + g(x),$$

$$(\lambda f)(x) = \lambda f(x).$$

Mostre que

$$\langle f|g\rangle = \int_{[0,1]} \overline{f(x)} g(x) dx$$

é um produto interno em $C_{\mathbb{C}}[0,1]$.

O produto interno nos permite introduzir uma noção que generaliza a um espaço vetorial qualquer a ideia de perpendicularidade no espaço, com a qual já estamos familiarizados:

Definição 2.5. *Dizemos que dois vetores* $|u\rangle$ *e* $|v\rangle$ *são* **ortogonais** *se* $\langle u|v\rangle = 0$. *Dizemos que um conjunto* $E = \{|v_1\rangle, \ldots, |v_k\rangle\}$ *é* **ortogonal** *se seus elementos são dois a dois ortogonais. Dizemos que um conjunto* $E = \{|v_1\rangle, \ldots, |v_k\rangle\}$ *é* **ortonormal** *se é ortogonal e* $\langle v_i|v_i\rangle = 1$ *para todo* i.

No caso do espaço ser $\mathbb{R}^3$ com o produto interno canônico, ou seja,

$$\langle (x_1, x_2, x_3)|(y_1, y_2, y_3)\rangle = x_1y_1 + x_2y_2 + x_3y_3$$

a ortogonalidade significa exatamente perpendicularidade no sentido geométrico usual.

Com um espaço vetorial V munido de um produto interno podemos definir uma aplicação $\|\cdot\| : V \to \mathbb{R}$ escrevendo $\|v\| = \sqrt{\langle v|v\rangle}$. De fato podemos provar que essa função é uma norma sobre V, mas para isso precisamos de alguns resultados preliminares.[3]

Teorema 2.1. *Se* u *e* v *são ortogonais, então*

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2.$$

Demonstração. Temos

$$\|u + v\|^2 = \langle u + v|u + v\rangle = \langle u|u\rangle + \langle u|v\rangle + \langle v|u\rangle + \langle v|v\rangle =$$

$$\|u\|^2 + \langle u|v\rangle + \overline{\langle u|v\rangle} + \|v\|^2 = \|u\|^2 + \|v\|^2,$$

pois $\langle u|v\rangle = 0$. □

Observação 2. *Durante a prova obtivemos uma identidade conhecida como identidade polar:*

$$\|u + v\|^2 = \|u\|^2 + 2Re(\langle u|v\rangle) + \|v\|^2.$$

Corolário 2.2. *Usando o teorema acima o leitor pode provar, indutivamente, o seguinte resultado: se* $\{|v_1\rangle, \ldots, |v_k\rangle\}$ *são dois a dois ortogonais, então*

$$\|v_1 + \cdots + v_k\|^2 = \|v_1\|^2 + \cdots + \|v_k\|^2.$$

[3]Evitamos o uso da notação $\||u\rangle\|$ e optamos por denotar a norma de um vetor $|u\rangle$ por $\|u\|$ por simplicidade.

Teorema 2.3. *Seja $E = \{|v_1\rangle, \ldots, |v_k\rangle\}$ subconjunto ortonormal de V. Então, para todo $|v\rangle \in V$*

$$\|v\|^2 = \sum_{i=1}^{k} |\langle v_i|v\rangle|^2 + \left\| v - \sum_{i=1}^{k} \langle v_i|v\rangle v_i \right\|^2 .$$

Demonstração. Podemos escrever

$$|v\rangle = \underbrace{\sum_{i=1}^{k} \langle v_i|v\rangle |v_i\rangle}_{|a\rangle} + \underbrace{|v\rangle - \sum_{i=1}^{k} \langle v_i|v\rangle |v_i\rangle}_{|b\rangle}$$

Os vetores $|a\rangle$ e $|b\rangle$ são ortogonais: de fato

$$\langle a|b\rangle = \left\langle \sum_{i=1}^{k} \langle v_i|v\rangle v_i \middle| v - \sum_{i=1}^{k} \langle v_i|v\rangle v_i \right\rangle =$$

$$\left\langle \sum_{i=1}^{k} \langle v_i|v\rangle v_i \middle| v \right\rangle - \sum_{i=1}^{k} \sum_{j=1}^{k} \overline{\langle v_i|v\rangle} \langle v_j|v\rangle \langle v_i|v_j\rangle =$$

$$\sum_{i=1}^{k} \overline{\langle v_i|v\rangle} \langle v_i|v\rangle - \sum_{i=1}^{k} \overline{\langle v_i|v\rangle} \langle v_i|v\rangle = 0.$$

Mas então $\|v\|^2 = \|a+b\|^2 = \|a\|^2 + \|b\|^2$ e

$$\|a\|^2 = \left\langle \sum_{i=1}^{k} \langle v_i|v\rangle v_i \middle| \sum_{j=1}^{k} \langle v_j|v\rangle v_j \right\rangle = \sum_{i=1}^{k} |\langle v_i|v\rangle|^2$$

donde segue o resultado. □

Exercício 2.7. *Se $E = \{|v_1\rangle, \ldots, |v_k\rangle\}$ é subconjunto ortonormal de V, mostre que para todo vetor $|v\rangle \in V$ vale a desigualdade de Bessel:*

$$\|v\|^2 \geq \sum_{i=1}^{k} |\langle v|v_i\rangle|^2.$$

Corolário 2.4. *Dados u e v em V então vale a desigualdade de Cauchy-Bunyakovsky-Schwarz:*

$$|\langle u|v\rangle| \leq \|u\|\|v\|.$$

Demonstração. Se $|v\rangle = 0$ a desigualdade é claramente verdadeira; vamos então assumir que $|v\rangle$ é não nulo. Sendo assim, podemos considerar o vetor $\dfrac{|v\rangle}{\|v\|}$ que é unitário (por quê?) e o conjunto $E = \left\{ \dfrac{|v\rangle}{\|v\|} \right\}$ é subconjunto ortonormal de V. Portanto, pela desigualdade de Bessel,

$$\|u\|^2 \geq \left| \left\langle u \middle| \frac{v}{\|v\|} \right\rangle \right|^2 = \frac{1}{\|v\|^2} |\langle u|v\rangle|^2$$

e assim $|\langle u|v\rangle| \leq \|u\|\|v\|$. □

Agora estamos em condições de verificar que a função $\|\cdot\| : V \to \mathbb{R}_+$ é de fato uma norma, isto é, uma aplicação de um espaço vetorial nos reais não negativos que satisfaz as condições abaixo:

1. $\|\lambda u\| = |\lambda|\|u\|$;
2. $\|u + v\| \leq \|u\| + \|v\|$;
3. $\|u\| = 0 \Rightarrow |u\rangle = 0$.

Deixamos para o leitor as provas de 1 e 3 e passamos a prova de 2: temos

$$\|u + v\|^2 = \|u\|^2 + 2Re(\langle u|v\rangle) + \|v\|^2 \leq \|u\|^2 + 2|\langle u|v\rangle| + \|v\|^2 \leq$$
$$\|u\|^2 + 2\|u\|\|v\| + \|v\|^2 = (\|u\| + \|v\|)^2$$

e então o resultado segue.

Uma norma nos permite definir uma noção natural de distância no espaço V, isto é, uma métrica, que é uma função $d : V \times V \to \mathbb{R}_+$ tal que

1. $d(u, v) = d(v, u)$;
2. $d(u, w) \leq d(u, v) + d(v, w)$;
3. $d(u, v) = 0 \Rightarrow |u\rangle = |v\rangle$.

Podemos definir d como sendo $d(u, v) = \|u - v\|$.

2.5.1 Produto Interno e Funcionais Lineares

Quando um espaço vetorial é munido de um produto interno, é possível associar vetores à funcionais lineares.

Exercício 2.8. *Consideremos fixos um certo $|v_0\rangle \in V$, e um produto interno $\langle\ |\ \rangle$ em V. Mostre que $L\colon V \to \mathbb{C}$ definido como sendo $L_{v_0}|v\rangle = \langle v_0|v\rangle$ é um funcional linear.*

Em alguns casos, o exemplo acima é absolutamente geral: todo elemento de V^* pode ser escrito na forma de L_{v_0} para algum $|v_0\rangle$ em V. Esse é o caso quando V tem dimensão finita.

Teorema 2.5. *Dado $L \in V^*$ então existe um único $|v_0\rangle \in V$ tal que $L|v\rangle = \langle v_0|v\rangle$.*

Demonstração. Considere uma base ortonormal $\{|e_i\rangle\}_{i=1,\ldots,k}$ de V. Então

$$|v\rangle = \sum_{i=1}^{k} v_i|e_i\rangle$$

e

$$L|v\rangle = \sum_{i=1}^{k} v_i L|e_i\rangle = \sum_{i=1}^{k} v_i L|e_i\rangle\langle e_i|e_i\rangle =$$

$$= \sum_{i,j=1}^{k} v_i L|e_j\rangle\langle e_j|e_i\rangle = \sum_{i=1}^{k} v_i \sum_{j=1}^{k} \left\langle \overline{L|e_j\rangle} e_j \middle| e_i \right\rangle =$$

$$= \sum_{i=1}^{k} v_i \left\langle \sum_{j=1}^{k} \overline{L|e_j\rangle} e_j \middle| e_i \right\rangle = \left\langle \sum_{j=1}^{k} \overline{L|e_j\rangle} e_j \middle| \sum_{i=1}^{k} v_i e_i \right\rangle = \langle v_0|v\rangle$$

onde $|v_0\rangle$ fica unicamente determinado como sendo $\sum_{j=1}^{k} \overline{L(e_j)}|e_j\rangle$, o que conclui a demonstração. □

Observação 3. *Este resultado simples é a versão em dimensão finita de um resultado bem mais geral da análise funcional conhecido como teorema de Riesz.*

O teorema acima mostra que o produto interno fornece uma identificação natural entre elementos de um espaço vetorial V e elementos do seu espaço dual V^* dada por

$$|v\rangle \longleftrightarrow L_v.$$

A notação de Dirac se aproveita desse fato para denotar o funcional L_v pelo *bra* $\langle v|$ de modo que

$$L_v|w\rangle = (\langle v|)|w\rangle = \langle v|w\rangle.$$

2.6 Bases Ortonormais

Definição 2.6. *Dado um espaço vetorial V munido de um produto interno $\langle\cdot|\cdot\rangle$, dizemos que uma base é ortogonal se ela é um subconjunto ortogonal de V. De forma análoga, uma base será chamada de base ortonormal se é um subconjunto ortonormal do espaço vetorial V*

Exercício 2.9. *Mostre que o conjunto formado pelos vetores $|e_1\rangle = (1, 0, \ldots, 0)$, $|e_2\rangle = (0, 1, 0, \ldots, 0)$, ..., $|e_n\rangle = (0, 0, \ldots, 0, 1)$ é uma base ortonormal com o produto interno canônico*

$$\langle (v_1, \ldots, v_n)|(u_1, \ldots, u_n)\rangle = \bar{v}_1 u_1 + \ldots + \bar{v}_n u_n.$$

2.6.1 Ortogonalização de Gram-Schmidt

Se assumimos a existência de uma base qualquer para o espaço V então podemos nos perguntar se há uma base ortonormal de V e a resposta é afirmativa. Dada uma base qualquer $\{|v_1\rangle, |v_2\rangle, \ldots, |v_n\rangle\}$ de V então podemos obter uma base ortonormal $\{|u_1\rangle, |u_2\rangle, \ldots, |u_n\rangle\}$ por meio de um procedimento conhecido como *ortogonalização de Gram-Schmidt* que passamos a descrever. Para construir o vetor $|u_1\rangle$ basta tomarmos

$$|u_1\rangle = \frac{|v_1\rangle}{\|v_1\|}.$$

Para construir $|u_2\rangle$ devemos ter em mente duas coisas: queremos que $|u_2\rangle$ tenha norma unitária e que seja ortogonal ao vetor já construído

$|u_1\rangle$. Para verificar essa segunda condição procuramos um vetor na forma $|w_2\rangle = |v_2\rangle + \alpha_1|u_1\rangle$ (que está no subespaço gerado por $|v_1\rangle$ e $|v_2\rangle$) de forma que $\langle w_2|u_1\rangle = 0$. Então

$$\langle w_2|u_1\rangle = \left\langle v_2 \middle| \frac{v_1}{\|v_1\|} \right\rangle + \alpha_1\langle u_1|u_1\rangle = 0$$

ou seja,

$$\alpha_1 = -\frac{1}{\|v_1\|}\langle v_2|v_1\rangle$$

O vetor $|u_2\rangle$ é então definido como sendo $|u_2\rangle = \frac{|w_2\rangle}{\|w_2\|}$. Para obter $|u_3\rangle$ procederemos de forma similar: primeiro procuramos $|w_3\rangle = |v_3\rangle + \alpha_1|u_1\rangle + \alpha_2|u_2\rangle$ que deve ser ortogonal a $|u_1\rangle$ e a $|u_2\rangle$, o que determina α_1 e α_2 como sendo

$$\alpha_1 = -\langle v_3|u_1\rangle \qquad \text{e} \qquad \alpha_2 = -\langle v_3|u_2\rangle$$

Seguindo dessa maneira não é difícil ver que o vetor auxiliar $|w_k\rangle$ será dado pela expressão

$$|w_k\rangle = |v_k\rangle - \langle v_k|u_1\rangle|u_1\rangle - \cdots - \langle v_k|u_{k-1}\rangle|u_{k-1}\rangle$$

e que $|u_k\rangle = \frac{w_k}{\|w_k\|}$, com $k = 2, 3, \ldots, n$. Dessa forma podemos exibir todos os vetores $|u_1\rangle, \ldots, |u_n\rangle$; por construção eles geram o mesmo espaço que $|v_1\rangle, \ldots, |v_n\rangle$. São também ortonormais, sendo assim a base ortonormal procurada do espaço V.

2.7 Mudança de Base

Vamos agora abordar a questão de como escrever um certo vetor em bases distintas. Consideremos duas bases, $\mathcal{U} = \{|u_1\rangle, \ldots, |u_n\rangle\}$ e $\mathcal{V} = \{|v_1\rangle, \ldots, |v_n\rangle\}$. Dado um vetor $|\varphi\rangle$ podemos escrevê-lo como

$$|\varphi\rangle = a_1|u_1\rangle + \cdots + a_n|u_n\rangle$$

que denotamos como $|\varphi\rangle = (a_1, \ldots, a_n)_{\mathcal{U}}$, onde os a_i são então as coordenadas de $|\varphi\rangle$ na base $\mathcal{U}$. Por outro lado, também podemos escrever $|\varphi\rangle = b_1|v_1\rangle + \cdots + b_n|v_n\rangle$, denotado por $(b_1, \ldots, b_n)_{\mathcal{V}}$, que

são as coordenadas de $|\varphi\rangle$ na base $\mathcal{V}$, e queremos obter a relação entre a_i e b_i.

Se

$$|u_1\rangle = T_{11}|v_1\rangle + T_{21}|v_2\rangle + \cdots + T_{n1}|v_n\rangle$$

e, analogamente,

$$|u_i\rangle = T_{1i}|v_1\rangle + \cdots + T_{ni}|v_n\rangle$$

para todo $i \geq 2$, então

$$b = Ta$$

onde $b = (b_1, \ldots, b_n)$ e $a = (a_1, \ldots, a_n)$. De fato, note que

$$|\varphi\rangle = a_1|u_1\rangle + \cdots + a_n|u_n\rangle =$$

$$a_1\left(\sum_{k=1}^{n} T_{k1}|v_k\rangle\right) + a_2\left(\sum_{k=1}^{n} T_{k2}|v_k\rangle\right) + \cdots + a_n\left(\sum_{k=1}^{n} T_{kn}|v_k\rangle\right) =$$

$$(T_{11}a_1+T_{12}a_2+\cdots+T_{1n}a_n)|v_1\rangle+\cdots+(T_{n1}a_1+T_{n2}a_2+\cdots+T_{nn}a_n)|v_n\rangle$$

e portanto as coordenadas na base $\mathcal{V}$ são as componentes do vetor Ta. Esta matriz T é a matriz de mudança de base, que troca as coordenadas na base $\mathcal{U}$ pelas coordenadas na base $\mathcal{V}$.

2.8 Operadores Lineares

Vamos voltar a estudar as transformações lineares entre dois espaços vetoriais U e V sobre $\mathbb{C}$. Seja $L(U,V) = \{T\colon U \to V \ ; T \text{ é linear}\}$; se T e S são elementos de $L(U,V)$ então podemos definir as transformações $T+S$ e λT tais que

$$(T+S)|u\rangle = T|u\rangle + S|u\rangle \qquad \text{e} \qquad (\lambda T)|u\rangle = \lambda T|u\rangle$$

Exercício 2.10. *Mostre que as operações $T+S$ e λT definidas assim são lineares. Mostre também que com essas operações o espaço $L(U,V)$ é um espaço vetorial sobre $\mathbb{C}$.*

Podemos definir uma norma no espaço $L(U,V)$ da seguinte maneira:

$$\|T\| = \sup_{\|u\|_U=1} \|T|u\rangle\|_V$$

onde $\|\cdot\|_U$ é uma norma em U e $\|\cdot\|_V$ é uma norma em V.

Estamos interessados no espaço $L(V) = L(V,V)$, ou seja, nas transformações lineares de um espaço vetorial nele mesmo. Uma transformação $T \in L(V)$ é chamada de *operador linear*. Nesse caso particular podemos também usar a norma $\|\cdot\|$ como anteriormente definida e podemos mostrar que ela satisfaz uma propriedade adicional (nesse contexto esta norma é geralmente conhecida como norma de operador).

Lema 2.6. *Dados A e B em $L(V)$ então $\|AB\| \leq \|A\|\|B\|$.*

Demonstração. Em primeiro lugar, note que se $|v\rangle \neq 0$,

$$|A|v\rangle| = \left|A\frac{|v\rangle}{\||v\|}\right|\|v\| \leq \|A\|\|v\|$$

pois $\frac{|v\rangle}{\||v\|}$ é um vetor unitário. Agora

$$\|AB\| = \sup_{\|v\|=1} |AB|v\rangle| \leq \sup_{\|v\|=1} \|A\||B|v\rangle| \leq \|A\|\|B\|.$$

□

Com essa norma podemos definir uma distância em $L(V)$ da seguinte forma: $d(A,B) = \|A-B\|$.

Se $\dim(V) = n$, fixada uma base em V, cada elemento de $L(V)$ pode ser representado por uma matriz quadrada $n\times n$ com coeficientes complexos.[4] O conjunto dessas matrizes será denotado por $M(V)$.

2.9 Adjunta de uma Transformação Linear

Quando temos uma transformação linear $T\colon V \to V$ podemos procurar uma nova transformação $T^*\colon V \to V$ de tal forma que

$$\langle Tv|u\rangle = \langle v|T^*u\rangle \qquad \text{para todo } |u\rangle \text{ e } |v\rangle \text{ em } V.$$

[4] No caso em que V é um espaço vetorial sobre $\mathbb{C}$.

Essa transformação é conhecida como a adjunta de T e de fato está unicamente determinada.

Teorema 2.7. *Dada uma transformação linear $T\colon V \to V$ então existe uma única transformação linear $T^*\colon V \to V$ tal que $\langle Tv|u\rangle = \langle v|T^*u\rangle$ para todo $|u\rangle$ e $|v\rangle$ em V.*

Demonstração. Considere $|u\rangle \in V$ fixo. Vamos definir a aplicação $L_u\colon V \to \mathbb{C}$ como sendo

$$L_u(v) = \langle u|Tv\rangle.$$

Da linearidade de T segue que L_u é um funcional linear e portanto existe um único $|u_0\rangle$ (que, naturalmente, depende de u) tal que

$$L_u(v) = \langle u_0|v\rangle$$

Como $|u_0\rangle$ depende de $|u\rangle$, escreveremos $|u_0\rangle = f|u\rangle$.

Se agora trocamos $|u\rangle$ por $|w\rangle$ então podemos definir L_w e teremos, de forma similar, um único $|w_0\rangle$ tal que $L_w|v\rangle = \langle w_0|v\rangle$, $|w_0\rangle = f|w\rangle$. Considere então

$$L_{u+w}|v\rangle = \langle v|T(u+w)\rangle = \langle v|Tu+Tw)\rangle = L_u|v\rangle + L_w|v\rangle =$$

$$\langle v|fu\rangle + \langle v|fw\rangle = \langle v|fu+fw\rangle.$$

Por outro lado, podemos escrever $L_{u+w}|v\rangle = \langle v|f(u+w)\rangle$ e portanto $f(|u\rangle + |w\rangle) = f|u\rangle + f|w\rangle$; o leitor pode, sem dificuldade, verificar que $f(\lambda|u\rangle) = \lambda f|u\rangle$, logo f é uma transformação linear, que denotaremos por T^*. □

Definição 2.7. *Um operador linear tal que $T = T^*$ é chamado de auto-adjunto.*

Quando estamos trabalhando com operadores auto-adjuntos, muito comuns em mecânica quântica, utilizamos o fato de que $\langle Tu|v\rangle = \langle u|Tv\rangle$ para denotarmos

$$\langle Tu|v\rangle = \langle u|Tv\rangle = \langle u\ |T|\ v\rangle.$$

Como é comum usarmos matrizes para representarmos os operadores lineares, queremos saber como é a matriz A associada a uma

transformação linear auto-adjunta T. Fixemos de início uma base ortonormal . Então sabemos que o elemento de matriz a_{ij} é dado por

$$a_{ij} = \langle e_i | T e_j \rangle$$

Então temos

$$a_{ij} = \langle T^* e_i | e_j \rangle = \langle T e_i | e_j \rangle = \overline{\langle e_j | T e_i \rangle} = \bar{a}_{ji}$$

Ou seja, a matriz A é igual a conjugação de sua transposta: $A = \bar{A}^T$. As matrizes associadas a operadores auto-adjuntos são chamadas auto-adjuntas ou também *hermitianas*.

2.10 Projeção sobre um Subespaço

Dado $|v\rangle \in V$ podemos definir a projeção (ou o projetor) sobre o subespaço vetorial W gerado por $|v\rangle$ como sendo

$$P_v : V \longrightarrow W \tag{2.1}$$

$$|u\rangle \longmapsto \frac{\langle v|u\rangle}{\|v\|^2}|v\rangle. \tag{2.2}$$

Podemos procurar a adjunta de P_v, isto é, a transformação P_v^* tal que

$$\langle P_v x | y \rangle = \langle x | P_v^* y \rangle \qquad \text{para todo } |x\rangle \text{ e } |y\rangle \text{ em } V.$$

Temos que

$$\langle P_v x | y \rangle = \frac{1}{\|v\|^2}\langle \langle v|x\rangle v | y \rangle = \frac{1}{\|v\|^2}\overline{\langle v|x\rangle}\langle v|y\rangle =$$

$$\frac{1}{\|v\|^2}\langle x|v\rangle\langle v|y\rangle = \frac{1}{\|v\|^2}\langle x|\langle v|y\rangle v\rangle = \langle x|P_v(y)\rangle.$$

Logo $P_v^* = P_v$ e portanto a projeção é uma transformação auto-adjunta.

A projeção tem uma outra propriedade interessante: se aplicamos esta transformação duas vezes então temos

$$P_v(P_v(u)) = P_v\left(\frac{\langle v|u\rangle}{\|v\|^2}v\right) = \frac{\langle v|u\rangle}{\|v\|^2}P_v(v) =$$

$$\frac{\langle v|u\rangle}{\|v\|^4}\langle v|v\rangle|v\rangle = \frac{\langle v|u\rangle}{\|v\|^2}|v\rangle = P_v(u)$$

Normalmente isto é denotado simplesmente por $P_v^2 = P_v$ (e quando não há confusão omite-se o subíndice v).

Podemos definir a projeção sobre subespaços de dimensão maior. Se W é um subespaço de V com uma base ortonormal $\{|v_1\rangle, ...|v_n\rangle\}$, a projeção sobre W é dada por

$$P_W : V \longrightarrow W \tag{2.3}$$

$$|u\rangle \longmapsto \sum_{i=1}^{n} P_{v_i}|v\rangle. \tag{2.4}$$

Exercício 2.11. *Mostre que P_W também é um operador auto-adjunto tal que $P_W^2 = P_W$.*

2.11 Autovetores e Autovalores

Se $T\colon V \to V$ é uma transformação linear, então podemos procurar vetores *não nulos* satisfazendo a equação

$$T|v\rangle = \lambda|v\rangle \qquad \text{para algum } \lambda \in \mathbb{C}.$$

As soluções $|v\rangle$ são conhecidas como autovetores e o respectivo λ como autovalor de T.

Observação 4. *E se o vetor nulo fosse admitido? Bem, nesse caso, temos $0 = T(0) = \lambda 0$ para todo e qualquer λ complexo; assim os autovalores seriam todo o conjunto $\mathbb{C}$ para qualquer transformação linear, o que não parece muito interessante...*

Exemplo 2.4. *Consideremos o caso de uma transformação linear P tal que $PP = P$ (o leitor consegue imaginar um exemplo?). Então a equação de autovalores é*

$$P|v\rangle = \lambda|v\rangle$$

Mas

$$\lambda|v\rangle = P|v\rangle = PP|v\rangle = P(\lambda|v\rangle) = \lambda^2|v\rangle$$

e assim os autovalores desta transformação satisfazem a relação $\lambda = \lambda^2$*, equação que tem soluções* 0 *e* 1*. Portanto podemos concluir que uma projeção (que satisfaz a relação acima) só admite como autovalores* 0 *e* 1*.*

2.11.1 Autovalores e Autovetores de Transformações Hermitianas

Se uma transformação linear é hermitiana, isto é, se $T^* = T$, então os autovalores e autovetores adquirem propriedades interessantes que investigaremos aqui. Acerca dos autovalores temos o seguinte resultado:

Teorema 2.8. *Se* T *é hermitiana então seus autovalores são reais.*

Demonstração. Considere $T|v\rangle = \lambda|v\rangle$. Então

$$\langle Tv|v\rangle = \langle \lambda v|v\rangle = \bar{\lambda}\langle v|v\rangle.$$

Por outro lado,

$$\langle Tv|v\rangle = \langle v|T^*v\rangle = \langle v|Tv\rangle = \langle v|\lambda v\rangle = \lambda\langle v|v\rangle$$

e portanto, $\lambda\langle v|v\rangle = \bar{\lambda}\langle v|v\rangle$, mostrando que $\lambda = \bar{\lambda}$, donde $\lambda \in \mathbb{R}$. □

Já para os autovetores, podemos verificar ortogonalidade.

Teorema 2.9. *Seja* T *hermitiana e* $|v\rangle$ *e* $|u\rangle$ *dois autovetores associados, respectivamente, aos autovalores distintos* λ *e* μ*. Então* $|u\rangle$ *e* $|v\rangle$ *são ortogonais.*

Demonstração. Note que

$$\langle Tv|u\rangle = \langle \lambda v|u\rangle = \lambda\langle v|u\rangle.$$

Por outro lado,

$$\langle Tv|u\rangle = \langle v|Tu\rangle = \langle v|\mu u\rangle = \bar{\mu}\langle v|u\rangle = \mu\langle v|u\rangle.$$

Portanto $\lambda\langle v|u\rangle = \mu\langle v|u\rangle$. Como λ e μ são distintos então temos necessariamente $\langle v|u\rangle = 0$, ou seja, $|u\rangle$ e $|v\rangle$ são ortogonais. □

O resultado a seguir, conhecido como Teorema Espectral, mostra como podemos utilizar autovetores e autovalores de uma transformação hermitiana para reescrevê-la.

Teorema 2.10. *Dada uma transformação hermitiana T é possível encontrar uma base ortonormal $B = \{|v_1\rangle, \ldots, |v_n\rangle\}$ para o espaço vetorial formada por autovetores de T. Além disso, se λ_i é o autovalor associado ao autovetor $|v_i\rangle$ então*

$$T = \sum_{i=1}^{n} \lambda_i P_{v_i}.$$

Definição 2.8. *Dizemos que uma transformação linear T é diagonalizável se existe uma base para o espaço vetorial em que a matriz que representa T é diagonal.*

Exercício 2.12. *Mostre que quando podemos encontrar uma base $B = \{|v_1\rangle, \ldots, |v_n\rangle\}$ para o espaço vetorial formada por autovetores de uma aplicação T então ela é diagonalizável. Em particular, mostre que todo operador hermitiano é diagonalizável.*

2.12 Operadores Positivos

Definição 2.9. *Dizemos que um operador T em um espaço vetorial V com produto interno é positivo definido se, para todo $|v\rangle$ em V não nulo, vale*

$$\langle Tv|v\rangle > 0.$$

Dizemos que T é um operador positivo semi-definido se para todo $|v\rangle$ em V não nulo, vale

$$\langle Tv|v\rangle \geq 0.$$

Quando T é positivo definido, escrevemos $T > 0$ e quando T é positivo semi-definido, escrevemos $T \geq 0$.

Exercício 2.13. *Mostre que os autovalores de um operador positivo são todos positivos.*

Teorema 2.11. *São equivalentes:*

1. *T é auto-adjunto e todos os seus autovalores são números reais positivos;*

2. *T é um operador positivo.*

Demonstração. Se valer 1, existe uma base ortonormal

$$B = \{|v_1\rangle, ..., |v_n\rangle\}$$

tal que a matriz $A = [T]_B$ é diagonal e cada $A_{ii} = a_i > 0$. Dado $|v\rangle$ em V, escrevemos $|v\rangle = x_1|v_1\rangle + ... + x_n|v_n\rangle$. Podemos calcular

$$\langle Tv|v\rangle = \langle x_1Tv_1 + ... + x_nTv_n|x_1v_1 + ... + x_nv_n\rangle =$$
$$= \langle x_1a_1v_1 + ... + x_na_nv_n|x_1v_1 + ... + x_nv_n\rangle =$$
$$= a_1|x_1|^2 + ... + a_n|x_n|^2 > 0.$$

Para a recíproca, mostremos que $T = T^*$. Devemos mostrar que

$$\langle Tv|w\rangle = \langle v|Tw\rangle$$

para todo $|v\rangle, |w\rangle$ em V. O truque é primeiro notar que $\langle Tv|v\rangle = \langle v|Tv\rangle$, uma vez que o conjugado de um número real é ele mesmo. Depois, expandimos

$$\langle T(u+v)|u+v\rangle = \langle u+v|T(u+v)\rangle =$$
$$\langle Tu|u\rangle + \langle Tu|v\rangle + \langle Tv|u\rangle + \langle Tv|v\rangle =$$
$$= \langle u|Tu\rangle + \langle u|Tv\rangle + \langle v|Tu\rangle + \langle v|Tv\rangle.$$

Cancelando termos correspondentes, concluímos que

$$\langle Tu|v\rangle + \langle Tv|u\rangle = \langle u|Tv\rangle + \langle v|Tu\rangle.$$

Agora trocamos $|v\rangle$ por $i|v\rangle$ na expressão acima, o que resulta em

$$i\langle Tu|v\rangle - i\langle Tv|u\rangle = i\langle u|Tv\rangle - i\langle v|Tu\rangle.$$

Multiplicando por i e somando membro a membro obtemos

$$\langle Tu|v\rangle = \langle u|Tv\rangle$$

para quaisquer $|u\rangle, |v\rangle$ em V.

Juntando o resultado acima ao exercício 2.13 provamos que a propriedade 2 implica a propriedade 1. □

Exercício 2.14. *Mostre que a projeção é um operador positivo semi-definido.*

2.13 Traço e Determinante

Vamos agora definir dois números que podem ser naturalmente associados a uma dada matriz quadrada e relembrar algumas de suas propriedades.

2.13.1 Traço

O traço de uma matriz quadrada A de elementos a_{ij} é definido como sendo a soma dos elementos da diagonal principal, ou seja,

$$\mathrm{Tr}A := \sum_{i=1}^{n} a_{ii}.$$

Isso então definiu uma função $\mathrm{Tr}\colon M_n(\mathbb{C}) \to \mathbb{C}$; algumas de suas propriedades básicas estão condensadas no próximo

Lema 2.12. *Para todo $A, B \in M_n(\mathbb{C})$ e $\lambda \in \mathbb{C}$*

1. $\mathrm{Tr}(A+B) = \mathrm{Tr}A + \mathrm{Tr}B$;
2. $\mathrm{Tr}(\lambda A) = \lambda \mathrm{Tr}A$;
3. $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.

Demonstração. A prova dos dois primeiros é bastante simples e é deixada ao leitor. Para verificarmos 3 notemos que

$$\mathrm{Tr}(AB) = \sum_{i=1}^{n} (AB)_{ii} = \sum_{i=1}^{n}\sum_{k=1}^{n} A_{ik}B_{ki}$$

$$\sum_{i=1}^{n}\sum_{k=1}^{n} B_{ki}A_{ik} = \sum_{k=1}^{n}\sum_{i=1}^{n} B_{ki}A_{ik} = \sum_{k=1}^{n} (BA)_{kk} = \mathrm{Tr}(BA).$$

□

Os dois primeiros ítens do lema mostram que de fato o traço é um exemplo de funcional linear no espaço das matrizes quadradas.

2.13.2 Determinante

O determinante é uma função polinomial $\det : M_n(\mathbb{C}) \to \mathbb{C}$. No caso de matrizes 2×2, por exemplo, o determinante é definido como

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} := ad - bc.$$

O leitor pode encontrar uma discussão bastante completa do caso geral, por exemplo, no livro de Elon Lima [Lim]. Podemos resumir suas principais propriedades no lema abaixo:

Lema 2.13. det *é uma função tal que:*

- $\det A^T = \det A$*;*
- $\det \lambda A = \lambda^n \det A$*;*
- $\det AB = \det A \det B$*.*

Uma outra propriedade importante é a seguinte: uma matriz A admite inversa (ou seja, existe A^{-1} tal que $AA^{-1} = A^{-1}A = Id$) se, e somente se, $\det A \neq 0$.

Isso permite caracterizar autovalores de maneira razoavelmente simples: dizemos que λ é um autovalor se existe $|v\rangle \neq 0$ tal que $A|v\rangle = \lambda|v\rangle$. Esta expressão pode ser reescrita como sendo $(A - \lambda I)|v\rangle = 0$ e estamos procurando uma solução $|v\rangle$ não nula para a mesma. Já sabemos que a transformação linear $A - \lambda I$, quando aplicada em zero, também resulta no vetor nulo. Portanto, se temos $|v\rangle$ não nulo satisfazendo a equação isso significa que a transformação $A - \lambda I$ não é injetiva e portanto não admite inversa. Mas não ter inversa significa que $\det(A - \lambda I) = 0$, sendo esta uma equação polinomial em λ cujas raízes são exatamente os autovalores associados à transformação linear representada pela matriz A.

2.14 Produto Tensorial

Dados dois espaços vetoriais V_A e V_B sobre $\mathbb{C}$ de dimensões n_A e n_B respectivamente, podemos construir um espaço vetorial de dimensão $n_A n_B$ através do produto tensorial.[5]

[5] A mesma construção pode ser feita para espaços vetoriais sobre outros corpos.

Para construirmos[6] esse novo espaço, que denotaremos por $V_A \otimes V_B$, tomamos bases $|i_A\rangle$ para V_A e $|j_B\rangle$ para V_B e declaramos que os $n_A n_B$ elementos da forma

$$|i_A\rangle \otimes |j_B\rangle, \quad i_A = 0, 1, \ldots, n_A, \ j_B = 0, 1, \ldots, n_B$$

formam uma base para $V_A \otimes V_B$. As seguintes condições são impostas

1. Para um escalar arbitrário $a \in \mathbb{C}$ e elementos $|v_A\rangle$ de V_A e $|v_B\rangle$ de V_B,

$$a(|v_A\rangle \otimes |v_B\rangle) = (a|v_A\rangle) \otimes |v_B\rangle = |v_A\rangle \otimes (a|v_B\rangle);$$

2. Para $|v_A\rangle$ e $|u_A\rangle$ arbitrários em V_A e $|v_B\rangle$ em V_B,

$$(|v_A\rangle + |u_A\rangle) \otimes |v_B\rangle = |v_A\rangle \otimes |v_B\rangle + |u_A\rangle \otimes |v_B\rangle;$$

3. Para $|v_A\rangle$ arbitrário em V_A e $|u_B\rangle$ e $|v_B\rangle$ em V_B,

$$|v_A\rangle \otimes (|u_B\rangle + |v_B\rangle) = |v_A\rangle \otimes |u_B\rangle + |v_A\rangle \otimes |v_B\rangle.$$

A construção é independente das escolhas de base para V_A e V_B.

Definição 2.10. *Dizemos que um vetor $|v\rangle \in V_A \otimes V_B$ é decomponível se é da forma $|v_A\rangle \otimes |v_B\rangle$.*

é comum usarmos a notação $|v_A\rangle \otimes |v_B\rangle = |v_A v_B\rangle$.

Se V_A e V_B são espaços vetoriais com produto interno, podemos definir um produto interno em $V_A \otimes V_B$ da seguinte maneira: para vetores decomponíveis fazemos

$$\langle v_A v_B | u_A u_B \rangle = \langle v_A | u_A \rangle \langle v_B | u_B \rangle,$$

e em seguida estendemos aos outros vetores:

$$(\langle v_A v_B| + \langle w_A w_B|)|u_A u_B\rangle = \langle v_A v_B | u_A u_B \rangle + \langle w_A w_B | u_A u_B \rangle$$

$$\langle v_A v_B|(|w_A w_B\rangle + |u_A u_B\rangle) = \langle v_A v_B | w_A w_B \rangle + \langle v_A v_B | u_A u_B \rangle.$$

[6] Para uma definição mais precisa, veja [Vai, NC].

Os conjuntos $M(V_A)$ e $M(V_B)$ são também espaços vetoriais sobre $\mathbb{C}$ e por isso também podemos definir o produto tensorial $M(V_A) \otimes M(V_B)$. Podemos então definir uma ação de $M(V_A) \otimes M(V_B)$ em $V_A \otimes V_B$ da seguinte forma: para vetores decomponíveis fazemos

$$M_A \otimes M_B(|v_A\rangle \otimes |v_B\rangle) = M_A|v_A\rangle \otimes M_B|v_B\rangle,$$

e em seguida estendemos por linearidade aos outros vetores. Essa ação define um mapa de $M(V_A) \otimes M(V_B)$ em $M(V_A \otimes V_B)$, que é um isomorfismo de espaços vetoriais.

Definição 2.11. *Definimos o traço parcial em relação a V_A de uma matriz $M_A \otimes M_B$ em $M(V_A) \otimes M(V_B)$ por*

$$\mathrm{Tr}_A(M_A \otimes M_B) = \mathrm{Tr}(M_A)M_B,$$

e estendemos por linearidade às matrizes não decomponíveis. De maneira análoga definimos o traço parcial em relação a V_B.

Definição 2.12. *Definimos a transposta parcial em relação a V_A de uma matriz $M_A \otimes M_B$ em $M(V_A) \otimes M(V_B)$ por*

$$(M_A \otimes M_B)^{T_A} = (M_A)^T \otimes M_B,$$

e estendemos por linearidade às matrizes não decomponíveis. De maneira análoga definimos a transposta parcial em relação a V_B.

Proposição 2.14. *Se uma matriz M é positiva, então $\mathrm{Tr}_A(M)$ e $\mathrm{Tr}_B(M)$ também o são.*

Demonstração. Suponhamos que $M = \sum_i M_A^i \otimes M_B^i$. Seja $\{|j\rangle\}, j = 1, \ldots, \dim V_B$ uma base ortonormal para V_B. Então

$$\mathrm{Tr}_B(M) = \sum_{i,j} M_A^i \langle j|M_B^i|j\rangle$$

$$\langle v|\mathrm{Tr}_B(M)|v\rangle = \sum_{i,j} \langle v|M_A^i|v\rangle\langle j|M_B^i|j\rangle = \sum_{i,j} \langle v|\langle j|M_A^i \otimes M_B^i|j\rangle|v\rangle =$$

$$\sum_j \langle v|\langle j| \sum_i M_A^i \otimes M_B^i|j\rangle|v\rangle = \sum_j \langle v|\langle j|M|j\rangle|v\rangle \geq 0$$

sendo que a última desigualdade é válida pelo fato de que M é positiva e portanto cada termo na última soma é positivo. Segue então que $\mathrm{Tr}_B(M)$ também é uma matriz positiva.

De maneira análoga provamos que $\mathrm{Tr}_A(M)$ é positiva. □

Um resultado extremamente útil é a decomposição de Schmidt para espaços vetoriais com estrutura de produto tensorial.

Proposição 2.15 (Decomposição de Schmidt)**.** *Dado um vetor* $|\Psi\rangle \in V_A \otimes V_B$*, é possível encontrar bases ortonormais* $\{|\psi_A^n\rangle\}$ *para* V_A *e* $\{|\phi_B^m\rangle\}$ *para* V_B *tais que*

$$|\Psi\rangle = \sum_{i=1}^{d} \alpha_i \big|\psi_A^i\big\rangle \big|\phi_B^i\big\rangle, \tag{2.5}$$

em que $d = \min(\dim V_A, \dim V_B)$*, e* $\alpha_1 \geq \alpha_2 \geq \ldots \geq \alpha_d$*. Os coeficientes* α_i *são chamados coeficientes de Schmidt.*

Demonstração. Suponhamos $d = \dim V_A$. Seja $\rho_A = \mathrm{Tr}_B(|\Psi\rangle\langle\Psi|)$. A matriz $|\Psi\rangle\langle\Psi|$ é o projetor na direção de $|\Psi\rangle$ e portanto é uma matriz positiva. Assim, ρ_A também é uma matriz positiva, e portanto podemos encontrar uma base $\{|\psi_A^n\rangle\}$ para V_A formada por autovetores de ρ_A . Desse modo, dada uma base ortonormal qualquer $|m_B\rangle$ para V_B, podemos escrever

$$|\Psi\rangle = \sum_{n,m} c_{nm} |\psi_A^n\rangle |m_B\rangle,$$

uma vez que o conjunto $\{|\psi_A^n\rangle|m_B\rangle\}$ forma uma base para $V_A \otimes V_B$. Podemos supor que os números são reais, englobando a parte complexa em $\big|\psi_A^i\big\rangle$ ou $\big|\phi_B^i\big\rangle$.

Seja α_n^2 o autovalor de ρ_A associado ao autovetor $|\psi_A^n\rangle$. Definimos então

$$|\phi_B^n\rangle = \sum_m \frac{c_{nm}}{\alpha_n} |m_B\rangle,$$

de modo que

$$|\Psi\rangle = \sum_{i=1}^{d} \alpha_i \big|\psi_A^i\big\rangle \big|\phi_B^i\big\rangle.$$

Resta mostrar que o conjunto $\{|\phi_B^m\rangle\}$ pode ser estendido a uma base ortonormal. Para isso, devemos verificar que esse é um conjunto ortonormal. De fato

$$\langle \phi_B^n | \phi_B^m \rangle = \sum k, l \frac{c_{nk}^* c_{ml}}{\alpha_n \alpha_m} \langle k_B | l_B \rangle$$

$$= \sum_k \frac{c_{nk}^* c_{mk}}{\alpha_n \alpha_m} = \frac{1}{\alpha_n \alpha_m} \sum_k \langle \Psi | \psi_A^n \rangle | k_B \rangle \langle \psi_A^m | \langle k_B | \Psi \rangle$$

$$= \frac{1}{\alpha_n \alpha_m} \sum_k \langle \psi_A^m | \rho_A | \psi_A^n \rangle = \frac{\alpha_n \alpha_m \delta_{nm}}{\alpha_n \alpha_m} = \delta_{nm}.$$

O ordenamento não-crescente dos coeficientes pode ser feito reordenando os vetores da base. □

Os coeficientes de Schmidt são os autovalores das matrizes reduzidas $\rho_A = \mathrm{Tr}_B(|\Psi\rangle\langle\Psi|)$ e $\rho_B = \mathrm{Tr}_A(|\Psi\rangle\langle\Psi|)$. Por esse motivo o número de coeficientes não nulos (chamado número de Schmidt) e também os seus valores são os mesmos para toda decomposição. Além disso, se

$$|\Psi\rangle = \sum_i a_i |i\rangle_A |i\rangle_B,$$

$$|\Psi\rangle = \sum_i a_i |i'\rangle_A |i'\rangle_B$$

são duas decomposições distintas, as aplicações lineares U_A e U_B definidas nas bases por

$$|i\rangle_A \mapsto |i'\rangle_A, \qquad |i\rangle_B \mapsto |i'\rangle_B$$

são aplicações unitárias tais que

$$U_A \otimes U_B(|\Psi\rangle) = |\Psi\rangle.$$

Desse modo, duas decomposições de Schmidt distintas estão relacionadas por unitárias locais que fixam $|\Psi\rangle$.

2.15 Exponencial de uma Matriz

Considere uma transformação linear $T\colon V \to V$. Nosso objetivo é definir a transformação linear $e^T\colon V \to V$. A motivação para isso é a representação da exponencial (real ou complexa) como série de potências,

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

Podemos então tentar definir e^T como sendo

$$e^T = \sum_{k=0}^{\infty} \frac{T^k}{k!}.$$

A expressão envolve soma de operadores lineares, composições de operadores lineares e o produto por números reais, todas essas operações que estão bem definidas para elementos de $L(V)$. Mas há uma passagem ao limite quando utilizamos a série e por isso devemos investigar com algum cuidado a questão da convergência.

Nosso primeiro passo na direção de definir e^T é a procura de um critério de convergência em $L(V)$.

Teorema 2.16. *$L(V)$ com a norma de operador $\|\cdot\|$ é um espaço completo, isto é, sequências de Cauchy são convergentes.*

Demonstração. Uma sequência de Cauchy é uma sequência $\{S_n\}_{n\in\mathbb{N}} \subset L(V)$ tal que para todo $\epsilon > 0$ existe $N \in \mathbb{N}$ tal que

$$\|S_m - S_n\| < \epsilon \qquad \text{para todo } m \text{ e } n \text{ maiores ou iguais a } N.$$

Fixemos agora um certo elemento $|v\rangle \in V$. Podemos então considerar a sequência $\{S_n|v\rangle\}_{n\in\mathbb{N}} \subset V$; da definição de $\|\cdot\|$ sabemos que

$$|S_m|v\rangle - S_n|v\rangle| = |(S_m - S_n)|v\rangle| \leq \|S_m - S_n\|\|v\|$$

o que mostra que $\{S_n(v)\}_{n\in\mathbb{N}} \subset V$ é uma sequência de Cauchy em V para a norma $|\cdot|$. Como V é um espaço completo[7] essa sequência converge para um ponto de V que denotaremos por $S|v\rangle$. Repetindo

[7] Essa é uma consequência do fato de que V é um espaço vetorial de dimensão finita sobre $\mathbb{C}$.

a ideia para cada ponto do espaço V conseguimos então definir uma função $S\colon V \to V$, $|v\rangle \mapsto S|v\rangle$. Devemos agora verificar que essa é de fato uma função linear. Para isso, note que

$$S(|v\rangle + |u\rangle) = \lim S_n(|v\rangle + |u\rangle) = \lim (S_n|v\rangle + S_n|u\rangle) =$$

$$\lim S_n|v\rangle + \lim S_n|u\rangle = S|v\rangle + S|u\rangle$$

e o leitor não terá dificuldade em provar que $S(\lambda|v\rangle) = \lambda S|v\rangle$, mostrando que temos $S \in L(V)$. □

Agora consideraremos a sequência $\{S_n\}_{n\in\mathbb{N}} \subset L(V,V)$ definida pelas somas parciais

$$S_n = \sum_{k=0}^{n} \frac{T^k}{k!}.$$

Segundo o teorema 2.16 devemos apenas verificar que esta é uma sequência de Cauchy para saber que a mesma tem limite. Mas se consideramos m e n, por exemplo, com $m \geq n$, então

$$\|S_m - S_n\| = \left\| \sum_{k=n+1}^{m} \frac{T^k}{k!} \right\| \leq \sum_{k=n+1}^{m} \frac{\|T\|^k}{k!}$$

Por que este número é pequeno? Vejamos: considere agora a série da função exponencial real

$$e^{\|T\|} = \sum_{k\geq 0} \frac{\|T\|^k}{k!}$$

que é uma série convergente; séries convergentes tem a bela propriedade de que suas caudas ficam pequenas, ou, para ser mais claro, dado $\epsilon > 0$ existe $N \in \mathbb{N}$ tal que

$$\sum_{k\geq N} \frac{\|T\|^k}{k!} \leq \epsilon$$

Portanto, agora sabemos que se tomamos $m \geq n \geq N$ temos

$$\sum_{k=n+1}^{m} \frac{\|T\|^k}{k!} \leq \sum_{k\geq N} \frac{\|T\|^k}{k!} \leq \epsilon$$

e assim a sequência $\{S_n\}_n$ é uma sequência de Cauchy; sendo assim ela converge para uma transformação linear $S \in L(V,V)$, que é definida como sendo a exponencial da transformação linear T, ou seja, $e^T := S = \lim S_n$.

Lema 2.17. *Propriedades básicas da exponencial de T:*

1. *se*

$$D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} \text{ então } e^D = \begin{bmatrix} e^{\lambda_1} & 0 & \dots & 0 \\ 0 & e^{\lambda_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{\lambda_n} \end{bmatrix};$$

2.
$$e^{QDQ^{-1}} = Qe^DQ^{-1};$$

3. *Se T e S comutam, isto é, se $TS = ST$, então $e^{T+S} = e^Te^S$;*

4. $\det e^A = e^{\mathrm{Tr}A}$.

Demonstração. 1. Segue do fato de que se

$$D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

então

$$D^k = \begin{bmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n^k \end{bmatrix}.$$

2. Segue do fato de que

$$(QDQ^{-1})^k = QD^kQ^{-1}.$$

Para as provas de 3 e 4 sugerimos ao leitor o texto de Sotomayor [Sot]. □

Corolário 2.18. *Decorre facilmente de 1 que* $e^{0_{n\times n}} = I$ *(onde* $0_{n\times n}$ *é a matriz nula).*

2.16 Comutador de Matrizes

Uma característica interessante de transformações lineares e das matrizes que as representam (que é o que usaremos no que segue) é a não comutatividade: em geral, dadas duas matrizes A e B (correspondendo a duas transformações lineares no mesmo espaço vetorial) não é verdade que $AB = BA$.

Exercício 2.15. *Faça o teste com*

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \qquad e \qquad B = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$$

e demonstre as afirmações acima.

Para quantificar-se o quanto um certo par de matrizes deixa de ser comutativo há o conceito de comutador, definido como segue:

$$[A, B] := AB - BA.$$

O comutador, sendo uma diferença de produtos de matrizes é, ele mesmo, uma matriz. Obviamente, duas matrizes comutam se, e somente se, seu comutador é a matriz nula.

Desta definição decorre de maneira simples que

$$\mathrm{Tr}[A, B] = 0.$$

De fato

$$\mathrm{Tr}[A, B] = \mathrm{Tr}(AB - BA) = \mathrm{Tr}(AB) - \mathrm{Tr}(BA) = 0.$$

Também é claro que $[A, B] = -[B, A]$.

Se as matrizes A e B são simétricas então podemos mostrar que a matriz $[A, B]$ é anti-simétrica: efetivamente,

$$[A, B]_{ij} = (AB)_{ij} - (BA)_{ij} = \sum_k A_{ik}B_{kj} - \sum_l B_{il}A_{lj} =$$

$$\sum_k A_{ki}B_{jk} - \sum_l B_{li}A_{jl} = \sum_k B_{jk}A_{ki} - \sum_l A_{jl}B_{li} =$$

$$(BA)_{ji} - (AB)_{ji} = (BA - AB)_{ji} = [B, A]_{ji} = -[A, B]_{ij}$$

como desejado.

2.17 Exercícios

Exercício 2.16. *Seja S_n o subespaço vetorial de $M_n(\mathbb{C})$ formado pelas matrizes $n \times n$ simétricas, isto é, tais que $a_{ij} = a_{ji}$. Obtenha uma base para S_n e a sua dimensão.*

Exercício 2.17. *Considere o espaço vetorial $\mathcal{P}_n[-1, 1]$ dos polinômios de grau n reais definidos em $[-1, 1]$ munido do produto interno*

$$\langle f|g\rangle = \int_{-1}^{1} f(t)g(t)dt.$$

Verifique que o conjunto $\{1, x, \ldots, x^n\}$ é uma base para este espaço. é ortogonal? Se não é, procure construir uma base ortogonal usando a técnica da seção 2.6.1.

Exercício 2.18. *Mostre que o conjunto de matrizes 2×2 da forma*

$$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \quad \textit{com } a \textit{ e } b \textit{ complexos}$$

é um subespaço vetorial do espaço $M_2(\mathbb{C})$; exiba uma base para este subespaço.

Exercício 2.19. *Mostre que $[A, [B, C]] + [C, [A, B]] + [B, [C, A]] = 0$ (que é conhecida como identidade de Jacobi).*

Exercício 2.20. *Existem matrizes A e B satisfazendo a equação*

$$AB - BA = I?$$

Exercício 2.21. *Usando as propriedades do traço e da adjunta, mostre que $\langle A|B\rangle := \mathrm{Tr}A^*B$ é um produto interno no espaço de matrizes $M_n(\mathbb{C})$.*

Exercício 2.22. *Mostre que se A e B são matrizes anti-simétricas então o comutador $[A, B]$ também é anti-simétrico.*

Exercício 2.23. *Dado um espaço vetorial V, verificar que V^* é também um espaço vetorial.*

Exercício 2.24. *Considere o espaço $M_n(\mathbb{C})$ de matrizes de ordem n e coeficientes complexos. Verifique que uma base para este espaço é dada pelas matrizes E_{ij} para $(i, j) \in \{1, 2, \ldots, n\} \times \{1, 2, \ldots, n\}$, onde cada uma das E_{ij} é definida como segue: fixados i e j, todos os elementos e_{ab} da matriz E_{ij} são nulos, exceto o elemento $e_{ij} = 1$. Desta forma os primeiros vetores da base são*

$$E_{11} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, E_{12} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \cdots,$$

$$E_{nn} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Capítulo 3

Equações Diferenciais Ordinárias

Uma equação diferencial é uma equação envolvendo uma função desconhecida e suas derivadas. As equações diferenciais têm inúmeras aplicações práticas em medicina, engenharia, química, biologia e outras diversas áreas do conhecimento pois podem ser usadas para modelar problemas relacionados com taxas de variação. Equações diferenciais também aparecem naturalmente no estudo da dinâmica dos sistemas físicos, uma vez que a função matemática que representa um sistema em um dado instante de tempo deve em geral satisfazer uma equação diferencial. Neste capítulo fazemos um breve estudo de equações diferenciais ordinárias. Nossa atenção será voltada para equações diferenciais lineares, que são as mais usadas em mecânica quântica. Para um tratamento bastante completo do assunto o leitor pode consultar [Sot, DL].

3.1 Equações Diferenciais Ordinárias

Primeiramente vamos definir precisamente uma equação diferencial em $\mathbb{C}^n$ e comentar a respeito de alguns aspectos gerais de suas soluções.

Definição 3.1. *Uma equação diferencial ordinária é uma equação na forma*

$$\frac{d}{dt}x(t) = f(t, x(t)) \tag{3.1}$$

onde $x\colon \mathbb{R} \to \mathbb{C}^n$ *e* $f : \mathbb{R} \times \mathbb{C}^n \to \mathbb{C}^n$ *(eventualmente é necessário se restringir a um subconjunto de* $\mathbb{R}$ *para o domínio da função* x*, mas nesse texto essa preocupação não se faz necessária).*

Uma solução para a equação diferencial acima é uma curva $x(t)$ cuja velocidade em qualquer instante de tempo é igual a $f(t, x(t))$.

Exercício 3.1. *Considere a equação*

$$\frac{d}{dt}x(t) = ax(t)$$

com $a \in \mathbb{C}$ *e* $x\colon \mathbb{R} \to \mathbb{C}$*. Então não é difícil ver que* $x(t) = ce^{at}$ *é uma solução, para qualquer constante c escolhida.*

Em geral, ao resolver problemas envolvendo equações diferenciais conhecemos qual valor a função x assume em um dado $t_0 \in \mathbb{R}$. Queremos encontrar soluções de (3.1) que satisfaçam essa propriedade adicional.

Definição 3.2. *Um problema de valor inicial (PVI) é dado por uma equação diferencial*

$$\frac{d}{dt}x(t) = f(t, x(t))$$

e uma condição inicial, que é um ponto em $\mathbb{R} \times \mathbb{C}^n$

$$(t_0, v_0).$$

Uma solução para o PVI acima é uma função $x(t) : \mathbb{R} \to \mathbb{C}$ que satisfaz a equação diferencial 3.1 e tal que $x(t_0) = v_0$.

Dada uma função f, queremos saber se o PVI 3.2 possui alguma solução e, em caso afirmativo, se essa solução é única. Muitos matemáticos puros e aplicados se dedicam a questões desse tipo e um importante teorema da área é o teorema de existência e unicidade abaixo.

Teorema 3.1 (Teorema de Existência e Unicidade de Picard−Lindelöf)**.** *Seja $f : \mathbb{R} \times \mathbb{C}^n \to \mathbb{C}^n$ uma função tal que*

$$\|f(t, y_1) - f(t, y_2)\| \leq M\|y_1 - y_2\|$$

para algum M real e positivo. Então existe h real positivo tal que o problema de valor inicial 3.2 admite uma única solução no intervalo $[t_0 - h, t_0 + h]$.

É importante ressaltar que o teorema acima garante apenas existência local, ou seja, em torno de alguma vizinhança de t_0. Além disso, mesmo sabendo da existência de uma solução, pode não ser trivial encontrá-la. No entanto, se tivermos um função candidata a solução basta substituí-la na equação e verificar se ela é satisfeita. Em caso afirmativo, saberemos que essa é a solução que procuramos, uma vez que a solução é única.

3.2 Equações Diferenciais Lineares

Nossa atenção será para equações diferenciais em que a função f possui uma forma mais simples. Primeiro vamos exigir que f não dependa da variável t explicitamente.

Definição 3.3. *Um campo vetorial em $\mathbb{C}^n$ é uma aplicação $X\colon \mathbb{C}^n \to \mathbb{C}^n$; Uma equação diferencial ordinária autônoma é uma equação na forma*

$$\frac{d}{dt}x(t) = X(x(t)) \tag{3.2}$$

onde $x\colon \mathbb{R} \to \mathbb{C}^n$.

A segunda exigência que fazemos é que X seja linear.

Definição 3.4. *Uma equação diferencial*

$$\frac{d}{dt}x(t) = X(x(t))$$

é chamada linear se o campo $X\colon \mathbb{C}^n \to \mathbb{C}^n$ é linear, ou seja, se

$$X(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 X(v_1) + \lambda_2 X(v_2)$$

para todos $\lambda_1, \lambda_2 \in \mathbb{C}$ e $v_1, v_2 \in \mathbb{C}^n$.

Uma vez escolhida uma base, uma aplicação linear em $\mathbb{C}^n$ sempre pode ser escrita como multiplicação por uma matriz $n \times n$ com elementos complexos. Dessa forma, se X é um campo linear existe uma matriz A tal que $X(x) = Ax$, de modo que toda equação diferencial linear pode ser escrita na forma

$$\frac{d}{dt}x(t) = Ax(t). \tag{3.3}$$

Sabemos pelo teorema de existência e unicidade que um problema de valor inicial

$$\frac{d}{dt}x(t) = Ax(t), \quad x(0) = x_0 \tag{3.4}$$

possui uma única solução em uma vizinhança de $t = 0$. Equações diferenciais lineares possuem a propriedade adicional de que as soluções estão definidas para todo $t \in \mathbb{R}$.

Teorema 3.2. *A aplicação*

$$\begin{aligned} x: \quad \mathbb{R} &\to \quad \mathbb{C}^n \\ t &\longmapsto \quad e^{At}x_0, \end{aligned}$$

em que a exponencial é definida como na seção 2.15, é solução do PVI 3.4.

Para provar esse teorema basta verificar que a função $x(t)$ acima satisfaz o PVI 3.4. A primeira coisa a notar é que

$$x(0) = e^{A0}x_0 = e^0 x_0 = x_0,$$

ou seja, a solução satisfaz a condição inicial. Precisamos então verificar que $x(t)$ satisfaz a equação diferencial ordinária 3.3.

Note que

$$e^{At} = I + tA + \frac{1}{2!}t^2A^2 + \ldots = \sum_{k=0}^{\infty} \frac{1}{k!}t^k A^k.$$

Portanto, se derivamos obtemos

$$\frac{d}{dt}e^{At} = A + \frac{1}{2!}2tA^2 + \frac{1}{3!}3t^2A^3 + \cdots =$$

$$= A(I + tA + \frac{1}{2!}t^2A^2 + \cdots) = Ae^{At} = e^{At}A$$

(onde a última igualdade segue do fato simples de que a matriz A comuta com I e com qualquer outra potência de A, de forma que podemos colocar A em evidência à direita ou à esquerda).

Logo,

$$\frac{d}{dt}x(t) = \frac{d}{dt}e^{At}x_0 = Ae^{At}x_0 = Ax(t)$$

e a equação é satisfeita; logo, $x(t) = e^{At}x_0$ é a solução do PVI enunciado acima.

A derivação termo a termo na série que define e^{tA} deve ser justificada.

Proposição 3.3. *Seja*

$$f(X) = \sum_{i=0}^{\infty} c_n X^n$$

uma função definida através de uma série de potências absolutamente convergente. Então a série

$$g(X) = \sum_{i=0}^{\infty} nc_n X^{n-1}$$

também é absolutamente convergente e

$$f'(X) = g(X).$$

3.3 Exercícios

Exercício 3.2. *Obtenha as exponenciais das seguintes matrizes:*

$$A = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} \qquad B = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$$

(para B note que podemos escrever

$$B = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} =: D + N$$

onde N é tal que $N^2 = 0$ e N e D comutam; aproveite-se disso para obter e^B como sendo $e^{D+N} = e^D e^N$).

Exercício 3.3. *Resolva os seguintes problemas de valor inicial:*

a)

$$\begin{cases} \frac{d}{dt}x &= 3x \\ \frac{d}{dt}y &= 3y \end{cases} \qquad (x(0), y(0)) = (1, 7)$$

b)

$$\begin{cases} \frac{d}{dt}x &= 2x + 3y \\ \frac{d}{dt}y &= 2y \end{cases} \qquad (x(0), y(0)) = (2, 5)$$

Capítulo 4

Grupos

Neste capítulo recordamos o importante conceito de grupo e apresentamos exemplos de grupos especiais de transformações lineares, alguns dos quais serão usados mais tarde.

4.1 Grupos

Um conjunto G munido de uma operação $\cdot\colon G \times G \to G$, $\cdot(a,b) = a \cdot b = ab$, é dito um grupo se

1. $(ab)c = a(bc)$ (associatividade);
2. Existe $e \in G$ tal que $ea = ae = a$ para todo $a \in G$ (existência do elemento neutro);
3. Para todo $a \in G$ existe $a^{-1} \in G$ tal que $aa^{-1} = a^{-1}a = e$ (existência do elemento inverso).

Se o grupo satisfaz a propriedade adicional

4. $ab = ba$ (comutatividade)

então dizemos que G é comutativo ou abeliano.

Exemplo 4.1. *O conjunto $\mathbb{Z}$ munido da operação $+$, isto é, a adição usual, é um grupo comutativo.*

Exemplo 4.2. *O conjunto $\mathbb{R}_+^* = \{x \in \mathbb{R}, x > 0\}$ munido da operação produto $\cdot$ (ou seja, o produto usual da reta) é um grupo abeliano.*

Exemplo 4.3. *O conjunto $BL(V, V)$ das transformações lineares bijetivas de V munido da operação de composição é um grupo. A aplicação identidade $id_V \colon V \to V$ faz o papel de elemento neutro (ou unidade) deste grupo.*

Um subconjunto não vazio H de G, munido da mesma operação produto do grupo G, é dito um subgrupo se:

i Para todo h_1 e h_2 de H, temos $h_1 h_2 \in H$;

ii Para todo h em H temos $h^{-1} \in H$.

Exemplo 4.4. *O conjunto $P = \{2k, k \in \mathbb{Z}\} = \{\ldots, -2, 0, 2, 4, \ldots\} \subset \mathbb{Z}$, dotado da adição usual, é subgrupo do grupo aditivo $\mathbb{Z}$.*

4.2 Grupos de Matrizes

Como transformações lineares podem ser naturalmente associadas a matrizes, torna-se interessante encontrar grupos de matrizes, que obviamente representarão determinados grupos de transformações lineares com alguma característica especial.

No que segue todas as matrizes serão elementos de $M_n(\mathbb{C})$ e a operação de grupo é o produto de matrizes usual (o leitor pode verificar que este mesmo conjunto munido da adição usual de matrizes também é um grupo, mas nesse caso comutativo).

4.2.1 Matrizes Invertíveis

Uma matriz tem inversa se, e somente se, seu determinante é diferente de zero. Definimos

$$GL(n, \mathbb{C}) = \{A \in M_n(\mathbb{C}) \text{ tal que } \det A \neq 0\}$$

e afirmamos que este conjunto, com a operação usual de produto matricial, é um grupo. Com efeito, os elementos de $GL(n, \mathbb{C})$ têm

inversa, pelo que foi comentado na seção 2.13.2; se A e B são elementos de $GL(n,\mathbb{C})$ então $\det AB = \det A \det B \neq 0$, e portanto $AB \in GL(n,\mathbb{C})$. O leitor não terá dificuldade em verificar que a matriz identidade I também é um elemento de $GL(n,\mathbb{C})$ (sendo que associatividade é uma propriedade do produto matricial em geral). Desta forma esse conjunto é de fato um grupo, como desejado.

Um subgrupo interessante de $GL(n,\mathbb{C})$ é o que é constituído por matrizes cujo determinante é exatamente 1:

$$SL(n,\mathbb{C}) = \{A \in GL(n,\mathbb{C}) \text{ tal que } \det A = 1\}$$

De fato, se A e B estão em $SL(n,\mathbb{C})$ então

$$\det AB = \det A \det B = 1$$

e assim $AB \in SL(n,\mathbb{C})$, mostrando que $SL(n,\mathbb{C})$ é fechado com relação ao produto; por outro lado, se $A \in SL(n,\mathbb{C})$ então também tem uma inversa A^{-1} (pois esta em $GL(n,\mathbb{C})$) e

$$\det A^{-1} = \frac{1}{\det A} = 1$$

mostrando que a inversa de A está realmente em $SL(n,\mathbb{C})$.

4.2.2 Matrizes Unitárias

Uma matriz $U \in GL(n,\mathbb{C})$ é dita unitária se $U^*U = UU^* = I$.

Naturalmente, I é unitária; a inversa de uma matriz unitária é U^*, que também é unitária. E se U e V são unitárias, então

$$(UV)^*UV = V^*U^*UV = V^*IV = V^*V = I$$

mostrando que UV é de fato unitária. Desta forma definimos um grupo, o grupo de matrizes unitárias $U(n) \subset GL(n,\mathbb{C})$.

Uma propriedade interessante da transformação linear associada a unitária U é a seguinte: dado $|v\rangle \in \mathbb{C}^n$,

$$\||U|v\rangle\|^2 = \langle Uv|Uv\rangle = \langle U^*Uv|v\rangle = \langle v|v\rangle = \|v\|^2$$

ou seja, a transformação é uma isometria: ela preserva a norma de um vetor. Por essa razão não é difícil ver que a norma de operador de U é exatamente 1.

4.2.3 Matrizes Ortogonais

Uma matriz *real* $O \subset GL(n, \mathbb{R})$ é dita ortogonal se $O^T O = OO^T = Id$. Esta condição pode ser vista de maneira mais geométrica se notamos que

$$(O^T O)_{ij} = \sum_k (O^T)_{ik} O_{kj} = \sum_k O_{ki} O_{kj}$$

é de fato o produto interno canônico das colunas i e j da matriz O; logo, a condição de O ser ortogonal é o mesmo que dizer que as colunas são duas a duas ortogonais e cada coluna é normalizada.

O leitor pode verificar que o produto de matrizes ortogonais continua sendo ortogonal; a identidade também é ortogonal e a inversa de uma matriz ortogonal é ortogonal. Dessa forma definimos o grupo de matrizes ortogonais reais $O(n, \mathbb{R}) \subset GL(n, \mathbb{R})$.

4.3 Matrizes Especiais

Podemos agora estudar subgrupos dos grupos de matrizes ortogonais e unitárias que incluem uma condição a mais: a de que o determinante seja 1. Vamos olhar com calma o caso em dimensão 2 e depois ver o que se passa em geral.

4.3.1 $SU(2)$

Denotamos por $SU(2)$ o conjunto de matrizes de $M_2(\mathbb{C})$ unitárias com determinante 1. Da condição de ser unitária segue que as colunas (e linhas) devem ser ortogonais; além disso temos a condição a mais a respeito do determinante. Desta forma o grupo pode ser descrito como segue:

$$SU(2) = \left\{ M \in M_2(\mathbb{C}) | M = \begin{bmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{bmatrix} \text{ e } |\alpha|^2 + |\beta|^2 = 1 \right\}$$

São matrizes cujo determinante é exatamente $|\alpha|^2 + |\beta|^2 = 1$. Com o produto de matrizes esse conjunto se torna um grupo.

Observação 5. *Há uma forte ligação entre este grupo e a esfera*

$$S^3 = \{(x_1, \ldots, x_4) | \sum x_i^2 = 1\}$$

que tentaremos deixar clara. Em primeiro lugar, podemos representar S^3 como um subconjunto de $\mathbb{C}^2$ (e não do $\mathbb{R}^4$, como fizemos), escrevendo $\gamma = x_1 + ix_2$ e $\delta = x_3 + ix_4$. Então fica claro que a condição $\sum x_i^2 = 1$ é equivalente a $|\gamma|^2 + |\delta|^2 = 1$. Mas então podemos considerar uma matriz

$$M = \begin{bmatrix} \gamma & -\bar{\delta} \\ \delta & \bar{\gamma} \end{bmatrix}$$

que é exatamente um elemento de $SU(2)$. Desta forma, podemos induzir na esfera S^3 um produto: dados dois pontos p e q de S^3,

$$p = (x_1, x_2, x_3, x_4) \qquad e \qquad q = (y_1, y_2, y_3, y_4)$$

que também podem ser vistos como pontos de $\mathbb{C}^2$

$$p = (\gamma_1, \delta_1) \qquad e \qquad q = (\gamma_2, \delta_2)$$

(com $\gamma_1 = x_1 + ix_2$, $\delta_1 = x_3 + ix_4$, $\gamma_2 = y_1 + iy_2$ e $\delta_2 = y_3 + iy_4$) podemos naturalmente associar às matrizes

$$M_p = \begin{bmatrix} \gamma_1 & -\bar{\delta_1} \\ \delta_1 & \bar{\gamma_1} \end{bmatrix} \qquad e \qquad M_q = \begin{bmatrix} \gamma_2 & -\bar{\delta_2} \\ \delta_2 & \bar{\gamma_2} \end{bmatrix}.$$

O produto pq então pode ser definido como sendo o ponto $pq = (\gamma, \delta)$ onde γ e δ são tais que

$$\begin{bmatrix} \gamma & -\bar{\delta} \\ \delta & \bar{\gamma} \end{bmatrix} = M_p M_q.$$

Desta forma temos certeza de que pq está na esfera S^3.

Este exemplo mostra que certos objetos geométricos, como é o caso de S^3, também podem ser observados de um ponto de vista algébrico, e isso é um caso particular de uma estrutura conhecida como grupo de Lie.

4.3.2 $SU(n)$

O grupo $SU(n)$, como é de se esperar, é formado pelas matrizes de $M_n(\mathbb{C})$ unitárias com determinante 1. Novamente temos linhas (e colunas) ortogonais e mais uma restrição que é dada pelo valor do determinante. Uma maneira de se obter matrizes em $SU(n)$ consiste em tomar $H \in M_n(\mathbb{C})$ com traço zero e tal que $H = H^*$. Então afirmamos que $U = e^{iH}$ está em $SU(n)$. Primeiro, vamos verificar que U é unitária:

$$U^* = (e^{iH})^* = e^{-iH^*} = e^{-iH}$$

e assim

$$U^*U = e^{-iH}e^{iH} = I.$$

Agora basta verificar que o determinante de U é 1; para isso usaremos a seguinte propriedade que relaciona o traço e o determinante de uma dada matriz A (ver [Sot]):

$$\det e^A = e^{\mathrm{Tr}A}.$$

Então

$$\det U = \det e^{iH} = e^{\mathrm{Tr}iH} = e^{i\mathrm{Tr}H} = e^0 = 1$$

como desejado.

4.4 Representação de Grupos

Uma forma concreta de estudar um grupo abstrato é por meio de uma representação, isto é, de uma "cópia"do grupo formada por matrizes; de forma mais precisa, dizemos que uma representação do grupo G é uma função $\pi\colon G \to GL(n, \mathbb{C})$ que satisfaz

$$\pi(a.b) = \pi(a)\pi(b)$$

para todo par a e b em G; a operação entre a e b é a operação do grupo e a operação entre $\pi(a)$ e $\pi(b)$ é o produto de matrizes.

Desta propriedade deduzimos algumas coisas interessantes. Por exemplo, $\pi(a) = \pi(a.e) = \pi(a)\pi(e)$, e assim notamos que $\pi(e)$ é a matriz identidade em $GL(n, \mathbb{C})$. Outra propriedade que pode ser facilmente deduzida é $\pi(a^{-1}) = (\pi(a))^{-1}$.

Quando a função $\pi\colon G \to GL(n,\mathbb{C})$ é injetiva dizemos que a representação é *fiel*.

Abaixo daremos quatro exemplos de representações do mesmo grupo $G = \{a, e\}$ com o produto definido por $a^2 = e$ (quais são os outros possíveis produtos?):

Exemplo 4.5. *Tome $\pi\colon G \to GL(n,\mathbb{C})$ dada por $\pi(a) = \pi(e) = I$; esta representação obviamente não é fiel.*

Exemplo 4.6. *Consideremos $\pi\colon G \to GL(1,\mathbb{C})$ dada por $\pi(e) = 1$ e $\pi(a) = -1$. Esta representação é fiel.*

Exemplo 4.7. *Tome*

$$\pi(e) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad e \quad \pi(a) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Exemplo 4.8. *Tome*

$$\pi(e) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad e \quad \pi(a) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Os três últimos exemplos são representações fiéis. Mas o leitor pode notar que o último tem um caráter um pouco distinto dos outros: há um subespaço de $\mathbb{C}^2$ (o subespaço gerado pelo vetor $(1, 0)$) que é invariante pelas duas matrizes da representação; portanto esta componente pode ser ignorada e assim ficaríamos com uma representação mais simples, num espaço vetorial com uma dimensão a menos, que seria exatamente a segunda representação apresentada. Neste último caso dizemos então que se pode reduzir a representação a uma mais simples. Nos dois casos intermediários isso não é possível (ou seja, não podemos eliminar dimensões do espaço vetorial) e as representações são ditas *irredutíveis*.

4.5 Ação de Grupos

Considere uma função $\phi\colon G \times X \to X$ satisfazendo as condições seguintes:

1. Para cada $g \in G$, $\phi(g, \cdot)$ é uma bijeção de X;

2. $\phi(e,x) = x$ (ou seja, $\phi(e,\cdot)$ é a aplicação identidade);

3. $\phi(g,\phi(h,x)) = \phi(gh,x)$.

Nesse caso dizemos que ϕ é uma ação do grupo G sobre o conjunto X.

Exemplo 4.9. *Seja $G = (\mathbb{R},+)$ e $X = \mathbb{R}$. Tome $\phi(g,x) = g + x$; então não é difícil verificar que isso define uma ação do grupo aditivo $\mathbb{R}$ sobre o conjunto $\mathbb{R}$.*

Exemplo 4.10. *Seja $G = (\mathbb{R},+)$ e $X = \mathbb{R}^n$; se A é uma matriz $n \times n$ (que podemos pensar como sendo a que esta associada a uma transformação linear) então definindo $\phi(g,x) = e^{gA}x$ temos uma ação de $\mathbb{R}$ sobre o conjunto $\mathbb{R}^n$.*

Note que esta última ação corresponde à solução do problema de valor inicial para uma EDO linear quando o campo em $\mathbb{R}^n$ é definido por $X(x) = Ax$.

Exercício 4.1. *Mostre que o conjunto dos números complexos unitários $\mathcal{U}$ com a multiplicação usual é um grupo e que*

$$\begin{array}{rcl} \phi : \mathcal{U} \times \mathbb{C}^n & \longrightarrow & \mathbb{C}^n \\ (e^{i\phi}, |v\rangle) & \longmapsto & e^{i\phi}|v\rangle \end{array}$$

é uma ação de $\mathcal{U}$ em $\mathbb{C}^n$.

4.6 Órbitas e Classes de Equivalência

Definição 4.1. *Uma relação binária em um conjunto X é um subconjunto R de $X \times X$. Se $(x,y) \in R$ usaremos a notação $x \sim y$. Uma relação binária em um conjunto X é chamada relação de equivalência se satisfaz as seguintes propriedades*

1. *$x \sim x$ (reflexividade);*

2. *Se $x \sim y$ então $y \sim x$ (simetria);*

3. *Se $x \sim y$ e $y \sim z$ então $x \sim z$ (transitividade).*

A classe de equivalência do elemento x é o subconjunto de X definido por

$$[x] = \{y \in x \ ; \ x \sim y\}.$$

Exercício 4.2. *Mostre que duas classes de equivalência distintas são conjuntos disjuntos e que a união de todas as classes de equivalência é o conjunto X.*

Dada uma ação ϕ de um grupo G em um conjunto X podemos definir uma relação de equivalência em X dizendo que $x \sim y$ se existe $g \in G$ tal que $\phi(g, x) = y$.

Exercício 4.3. *Mostre que a relação definida acima é de fato uma relação de equivalência.*

A classe de equivalência de um elemento $x \in X$

$$[x] = \{y \in X \ ; \ \phi_g(x) = y, g \in G\}$$

é também chamada órbita de x pela ação de G.

4.7 A Fibração de Hopf

Uma bela construção matemática, a *fibração de Hopf*, aparece naturalmente na descrição dos estados de um qbit. Esta seção é dedicada a explicá-la.

Definição 4.2. *Uma fibração é definida por um mapa h que leva um espaço E em um espaço B, chamado* espaço base. *Um conjunto $F \subset E$ é chamado* fibra *se corresponde a $h^{-1}(p)$ para algum $p \in B$.*

Exemplo 4.11. *Um exemplo trivial é a projeção*

$$\begin{array}{ccc} h : \mathbb{R}^3 & \longrightarrow & \mathbb{R}^2 \\ \begin{pmatrix} a & b & c \end{pmatrix} & \longmapsto & \begin{pmatrix} a & b \end{pmatrix}. \end{array}$$

As fibras são retas paralelas ao eixo z.

No caso da fibração de Hopf, $E = S^3$, $B = S^2$ e $F = S^1$. Para definir o mapa h vamos identificar S^3 ao conjunto $\mathcal{V}$ de vetores $(z, w) \in \mathbb{C}^2$ tais que $|z|^2 + |w|^2 = 1$ como fizemos anteriormente

$$(a \ b \ c \ d) \leftrightarrow (a + ib \ c + id)$$

e $\mathbb{R}^2$ ao conjunto dos números complexos da maneira usual

$$(a\ b) \leftrightarrow a + ib.$$

O mapa h é a composição de dois mapas h_1 e h_2 definidos da seguinte forma

$$h_1 : S^3 \backsimeq \mathcal{V} \longrightarrow \mathbb{C} + \{\infty\}$$
$$\begin{pmatrix} \alpha & \beta \end{pmatrix} \longmapsto C = \alpha\bar{\beta}^{-1},$$
$$h_2 : \mathbb{C} \cup \{\infty\} \backsimeq \mathbb{R}^2 \cup \{\infty\} \longrightarrow S^2$$
$$C \longmapsto \Pi_E^{-1}(C),$$

em que $\Pi_E : S^2 \to \mathbb{R}^2 \cup \{\infty\}$ denota a projeção estereográfica

$$\Pi_E\begin{pmatrix} a & b & c \end{pmatrix} = \begin{pmatrix} \frac{a}{1-c} & \frac{b}{1-c} \end{pmatrix}.$$

Geometricamente, a projeção estereográfica tem um significado bem interessante. Tomamos um ponto $q = \begin{pmatrix} a & b & c \end{pmatrix}$ na esfera S^2 e construímos a reta que liga esse ponto ao polo norte $p = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}$

$$\begin{pmatrix} ta & tb & t(c-1)+1 \end{pmatrix},\ \ t \in \mathbb{R}.$$

A projeção estereográfica leva q na interseção dessa reta com o plano $z = 0$. O polo norte é levado ao ponto no infinito.

Exercício 4.4. *Mostre que a construção geométrica mencionada acima leva justamente ao mapa*

$$\Pi_E\begin{pmatrix} a & b & c \end{pmatrix} = \begin{pmatrix} \frac{a}{1-c} & \frac{b}{1-c} \end{pmatrix}.$$

Calcule a inversa de Π_E *e mostre que*

$$h_2\begin{pmatrix} x & y \end{pmatrix} = \begin{pmatrix} \frac{2x}{x^2+y^2+1} & \frac{2y}{x^2+y^2+1} & \frac{x^2+y^2-1}{x^2+y^2+1} \end{pmatrix}.$$

Observação 6. *Um exercício mais sofisticado é mostrar que os mapas* h_1 *e* h_2 *são contínuos com respeito às topologias adequadas. Também é possível mostrar que o mapa* h_2 *é uma realização do famoso homeomorfismo entre o plano e a esfera menos um ponto. Assim, ao passar do plano complexo para o plano mais um ponto, onde cada ponto é da forma* $C = \alpha\bar{\beta}^{-1}$*, dizemos que foi feita a* compactificação do plano complexo*, acrescentando o chamado* ponto de infinito*, correspondente a* $\beta = 0$*. Esta compactificação é normalmente chamada* esfera de Riemann*.*

Podemos escrever o mapa h em uma forma simpática se usarmos as representações polares $\alpha = r_1 e^{i\phi_1}$ e $\beta = r_2 e^{i\phi_2}$.

Exercício 4.5. *Verifique que*

$$C = h_1\begin{pmatrix} \alpha & \beta \end{pmatrix} = \frac{r_1}{r_2}\begin{pmatrix} \cos(\phi_2 - \phi_1) & \operatorname{sen}(\phi_2 - \phi_1) \end{pmatrix}.$$

Exercício 4.6. *Usando a expressão para C encontrada no exercício anterior verifique que*

$$h_2(C) = \begin{pmatrix} \frac{2r_1 r_2 \cos(\phi_2-\phi_1)}{r_1^2+r_2^2} & \frac{2r_1 r_2 \ \operatorname{sen}(\phi_2-\phi_1)}{r_1^2+r_2^2} & \frac{r_1^2+r_2^2}{r_1^2+r_2^2} \end{pmatrix}.$$

Para a aplicação da Fibração de Hopf em mecânica quântica, é útil relacionar a construção que acabamos de fazer aos operadores auto-adjuntos

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

chamados operadores de Pauli. Definindo

$$\langle \sigma_i \rangle = \langle v \ |\sigma_i| \ v \rangle$$

em que $|v\rangle = \begin{pmatrix} \alpha & \beta \end{pmatrix}$, temos

$$h_2(C) = \begin{pmatrix} 2\mathrm{Re}\big(\alpha\bar{\beta}\big) & 2\mathrm{Im}\big(\alpha\bar{\beta}\big) & |\alpha|^2 - |\beta|^2 \end{pmatrix} = \begin{pmatrix} \langle\sigma_1\rangle & \langle\sigma_2\rangle & \langle\sigma_3\rangle \end{pmatrix}.$$

As fibras do mapa h são as fibras do mapa h_1 pois h_2 é um mapa bijetivo. Essas fibras são as classes de equivalência $\{e^{i\phi}|v\rangle\}$ da ação de $\mathcal{U}$ em $\mathbb{C}^2$ mostrada no exemplo 4.1.

4.8 Exercícios

Exercício 4.7. *Verifique que o conjunto $G = \{a, b\}$ munido do produto $aa = a, ab = ba = b, bb = a$ é um grupo.*

Exercício 4.8. *Considere o conjunto $\{0, 1\}$ munido da multiplicação usual, ou seja, $00 = 0$, $01 = 10 = 0$ e $11 = 1$. Verifique as propriedades de grupo neste caso; o grupo em questão é conhecido como o grupo multiplicativo $\mathbb{Z}_2$; compare-o com o grupo da primeira questão.*

Exercício 4.9. *Verifique que o grupo $GL(n, \mathbb{C})$ não é comutativo (sugestão: procure exemplos adequados).*

Exercício 4.10. *Considere G o espaço de matrizes na forma*

$$\begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$$

com $t \in \mathbb{C}$; verifique que este conjunto, munido do produto usual de matrizes, é um grupo abeliano (comutativo).

Exercício 4.11. *Considerando o grupo do exercício anterior verifique que a aplicação*

$$\phi \colon G \times \mathbb{R} \to \mathbb{R}$$

$$\phi(\begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}, x) = x + t$$

é uma ação de G sobre $\mathbb{R}$.

Exercício 4.12. *Mostre que o conjunto*

$$\begin{bmatrix} 1 & x & y \\ 0 & 1 & z \\ 0 & 0 & 1 \end{bmatrix}$$

é um grupo para a operação do produto de matrizes (conhecido como grupo de Heisenberg). É comutativo?

Capítulo 5

Álgebras C^*

Neste capítulo formalizaremos um conceito que já estava latente nas páginas anteriores, o de álgebra C^*, e mostraremos mais alguns exemplos. O leitor que se interessar pelo assunto deve consultar o texto introdutório de Ruy Exel [Exe].

5.1 Álgebras C^*

Um espaço vetorial A munido de uma operação de produto é o que se chama de uma álgebra. Quando temos uma norma no espaço e ele é completo com relação a mesma (ou seja, é um espaço de Banach) então o chamamos de álgebra de Banach. Se, além disso, temos uma involução $*\colon A \to A$ satisfazendo

1. $(a+b)^* = a^* + b^*$;
2. $(\lambda a)^* = \bar{\lambda} a^*$;
3. $(ab)^* = b^* a^*$;
4. $(a^*)^* = a$;
5. $\|a^*\| = \|a\|$;

6. $\|a^*a\| = \|a\|^2$;

então temos uma álgebra C^*.

Exemplo 5.1. $\mathbb{C}$ *com a norma usual* $\|z\| = |z|$ *e com a involução* $z^* = \bar{z}$ *sendo a operação de tomar o complexo conjugado é uma álgebra* C^*.

Exemplo 5.2. *Seja* $C_0(\mathbb{R})$ *o conjunto das funções* $f\colon \mathbb{R} \to \mathbb{C}$ *contínuas e que se anulam no infinito, isto é, tais que para todo* $\epsilon > 0$ *o conjunto* $\{x : |f(x)| \geq \epsilon\}$ *é compacto. A norma*

$$\|f\| = \sup_{x\in\mathbb{R}} |f(x)|$$

torna esse espaço vetorial completo. Para termos uma álgebra precisamos introduzir um produto e o faremos da forma mais simples: $(fg)(x) = f(x)g(x)$*, o que nos dá uma álgebra comutativa. Podemos definir uma involução como sendo*

$$f^*(x) = \overline{f(x)}.$$

Com todos esses ingredientes temos então uma álgebra C^* *comutativa.*

Uma questão interessante é a de se saber se essa álgebra tem ou não unidade, isto é, uma função que denotaremos por $1(x)$ *tal que* $1f = f1 = f$ *para toda* $f \in C_0(\mathbb{R})$*. O leitor não terá dificuldade em verificar que nossa função só pode ser* $1(x) = 1$ *para todo* $x \in \mathbb{R}$*, mas esse* não é *um elemento de* $C_0(\mathbb{R})$ *pois não se anula no infinito. Desta forma essa é uma álgebra sem unidade.*

O leitor é convidado a repensar o exemplo acima, mas trocando $\mathbb{R}$ *por* $[0, 1]$ *para concluir que* $C[0, 1]$ *(norma, produto e involução como acima) é uma álgebra* C^* *com unidade.*

O exemplo acima pode ser repetido trocando $\mathbb{R}$ por um espaço X mais geral. É interessante notar que esse modelo básico de uma álgebra C^* comutativa na verdade é, num certo sentido, o único modelo pois uma álgebra desse tipo sempre acaba sendo isomorfa a uma álgebra $C(X)$ para um certo X (este é um resultado muito importante na área, conhecido como Teorema de Gelfand. O leitor curioso é remetido a [Exe] para uma discussão mais completa).

Exemplo 5.3. *Seja $M_n(\mathbb{C})$ o conjunto de matrizes $n \times n$ com coeficientes complexos. Este espaço vetorial tem um produto natural, o produto de matrizes, que o torna uma álgebra. Podemos definir uma norma como sendo a norma usual de operadores*

$$\|A\| = \sup_{v:|v|=1} |A(v)|.$$

$M_n(\mathbb{C})$ é completo nessa norma. A involução pode ser definida como sendo

$$A^* = \overline{A^t}, \qquad \text{ou seja,} \qquad (a_{ij})^* = \overline{a}_{ji},$$

onde (a_{ij}) são as entradas da matriz A. Então temos uma álgebra C^*.

A verificação dos detalhes é deixada ao leitor; vamos aqui nos limitar a mostrar a propriedade $\|A^*A\| = \|A\|^2$: Seja $v \in \mathbb{C}^n$ um vetor unitário, isto é, $|v| = 1$. Então

$$|Av|^2 = \langle Av|Av\rangle = \langle A^*Av|v\rangle \leq$$

$$|A^*Av||v| = |A^*Av| \leq \|A^*A\||v| = \|A^*A\|.$$

Tomando o supremo sobre v em ambos os lados podemos concluir que $\|A\|^2 \leq \|A^*A\|$; como já haviamos visto antes, nas propriedades da norma de operador, $\|A^*A\| \leq \|A^*\|\|A\| = \|A\|\|A\| = \|A\|^2$. Portanto

$$\|A^*A\| \leq \|A\|^2 \leq \|A^*A\|$$

e assim $\|A^*A\| = \|A\|^2$ como desejado.

É interessante notar que se a norma é modificada então a álgebra pode deixar de ser uma álgebra C^*. No exemplo acima, podemos nos perguntar o que ocorre se trocamos a norma por outra equivalente, definida como sendo

$$\|M\|_2 := \sqrt{\sum_{ij} |M_{ij}|^2}.$$

Então $\|Id\|_2 = \sqrt{n} \neq 1$; porém, uma unidade deve satisfazer $\|1^*1\| = \|1\|^2$, o que implica em $\|1\|$ sendo 1 ou 0. Desta forma vemos que $M_n(\mathbb{C})$ munido de $\|\cdot\|_2$ não é uma álgebra C^*.

5.2 Estados de uma Álgebra

Álgebras são espaços vetoriais, logo é interessante perguntar o que ocorre com seus funcionais lineares. No caso de algebras C^* com unidade uma classe especial de funcionais lineares merece bastante atenção, os chamados *estados*; na 9.6 veremos a ligação deste conceito de estado com os que ainda serão apresentados nesse texto.

Definição 5.1. *Seja A uma álgebra C^* com unidade; um funcional linear $f\colon A \to \mathbb{C}$ é chamado de estado se*

(a) $f(a^*a) \geq 0$ *para todo* $a \in A$*;*

(b) $f(1) = 1$.

Exemplo 5.4. *Considere a álgebra $A = \mathbb{C}$; os funcionais lineares $f\colon A \to \mathbb{C}$ são da forma $f(z) = \lambda z$ com λ sendo um elemento de $\mathbb{C}$. Para que f seja então um estado é preciso que $f(1) = \lambda 1 = \lambda = 1$ e assim o único possível estado é $f(z) = z$; para concluirmos que de fato é um estado basta verificar que $f(z^*z) = z^*z = |z|^2 \geq 0$, o que é verdade por ser uma propriedade da norma de um número complexo. Portanto concluímos que para essa álgebra existe um único estado $f(z) = z$.*

Exemplo 5.5. *Seja $A = C[0,1] = \{f\colon [0,1] \to \mathbb{C}\ ; f$ contínua$\}$. Considere agora uma função $p\colon [0,1] \to \mathbb{R}^+$, isto é, que assume valores não negativos e tal que*

$$\int_{[0,1]} p(x)dx = 1.$$

Então, não é difícil verificar que

$$f_p(a) = \int_{[0,1]} a(x)p(x)dx$$

é um estado para A; a linearidade é clara. Também é fácil provar que $f_p(1) = 1$ e $f_p(a^*a) \geq 0$. Mas o leitor pode notar que agora temos uma ampla possibilidade de escolhas para a função p, cada uma delas resultando em um funcional e assim, ao contrário do primeiro exemplo, temos uma situação com uma infinidade de estados para a álgebra (comutativa) A.

5.2.1 Estados da Álgebra $M_n(\mathbb{C})$

Como a álgebra que mais aparece nessas páginas é $M_n(\mathbb{C})$, vamos descrever precisamente seus estados. O primeiro passo é a definição de um produto interno: se a e b são elementos de $M_n(\mathbb{C})$ então

$$\langle a|b\rangle = \mathrm{Tr}(a^*b)$$

é um produto interno. De fato a verificação da linearidade não é difícil e deixamos a tarefa para o leitor.

Usando uma base qualquer de $\mathbb{C}^n$, por exemplo a base canônica, podemos agora ver que

$$\langle b|a\rangle = \mathrm{Tr}(b^*a) = \sum_{i=1}^{n}\langle e_i|b^*ae_i\rangle = \sum_{i=1}^{n}\langle a^*be_i|e_i\rangle =$$

$$= \sum_{i=1}^{n}\overline{\langle e_i|a^*be_i\rangle} = \overline{\sum_{i=1}^{n}\langle e_i|a^*be_i\rangle} = \overline{\mathrm{Tr}(a^*b)} = \overline{\langle a|b\rangle}.$$

Além disso,

$$\langle a|a\rangle = \mathrm{Tr}(a^*a) = \sum_{i=1}^{n}\langle e_i|a^*ae_i\rangle = \sum_{i=1}^{n}\langle ae_i|ae_i\rangle =$$

$$= \sum_{i=1}^{n}\|ae_i\|^2 \geq 0.$$

Portanto, se $\langle a|a\rangle = 0$ temos obrigatoriamente que $\|ae_i\| = 0$, donde $ae_i = 0$ para todo i, o que implica que a é a matriz nula, $a = 0$. Desta forma concluímos que $\langle\cdot|\cdot\rangle$ é de fato um produto interno para o espaço vetorial $M_n(\mathbb{C})$.

Consideremos agora um funcional linear $f\colon A \to \mathbb{C}$. Pelo teorema 2.5 sabemos que f pode ser escrito como

$$f(x) = \langle V_f|x\rangle = \mathrm{Tr}(V_f^*x)$$

para um único elemento $V_f \in A$, ou seja, para uma matriz V_f que é $n \times n$ e cujos elementos são complexos.

Para que f seja um estado devemos ter $f(1) = 1$, logo

$$f(1) = \mathrm{Tr}(V_f^* 1) = \mathrm{Tr}(V_f^*) = 1,$$

que é a primeira condição que obtemos sobre V_f.

Sendo f um estado, temos também que $f(a^*a) \geq 0$ para todo $a \in A$; esta expressão contém na verdade duas informações: a primeira é que $f(a^*a)$ é *real* (lembre-se de que f assume valores em $\mathbb{C}$). A segunda é que, sendo real, é um número não-negativo. Desta propriedade podemos deduzir que $V_f^* = V_f$ e $V_f \geq 0$. De fato, para todo $a \in A$

$$0 \leq f(a^*a) = \mathrm{Tr}(V_f^* a^* a) = \mathrm{Tr}(a V_f^* a^*) = \sum_{i=1}^{n} \big\langle e_i \big| a V_f^* a^* e_i \big\rangle =$$

$$= \sum_{i=1}^{n} \big\langle a^* e_i \big| V_f^* a^* e_i \big\rangle.$$

Mas dado um vetor $v \in \mathbb{C}^n$, podemos escrever uma transformação linear a^* tal que

$$a^*(e_1) = v \qquad \text{e} \qquad a^*|_{e_1^\perp} = 0.$$

Desta forma, temos que

$$0 \leq f(a^*a) = \big\langle v \big| V_f^* v \big\rangle \qquad \text{para qualquer } v \in \mathbb{C}^n$$

e, assim, $V_f^* \geq 0$. Também verificamos que $\left\langle v \middle| V_f^* v \right\rangle$ é real (de fato não-negativo) e portanto

$$\big\langle v \big| V_f^* v \big\rangle = \langle V_f v | v \rangle = \overline{\langle v | V_f v \rangle} = \langle v | V_f v \rangle.$$

Como esta igualdade vale para todo $v \in \mathbb{C}^n$ então concluímos que $V_f = V_f^*$, como desejado: de fato

$$\big\langle v \big| (V_f - V_f^*) v \big\rangle = 0$$

para todo vetor v; podemos então trocar v por $v + w$ e por $v + iw$. De

$$\big\langle v + w \big| (V_f - V_f^*)(v + w) \big\rangle = 0 \text{ e } \big\langle v + iw \big| (V_f - V_f^*)(v + iw) \big\rangle = 0$$

obtemos, respectivamente,

$$\langle w|(V_f - V_f^*)v\rangle = -\langle v|(V_f - V_f^*)w\rangle \text{ e}$$

$$\langle w|(V_f - V_f^*)v\rangle = \langle v|(V_f - V_f^*)w\rangle.$$

Logo $\left\langle v\middle|(V_f - V_f^*)w\right\rangle = \left\langle w\middle|(V_f - V_f^*)v\right\rangle = -\left\langle v\middle|(V_f - V_f^*)w\right\rangle$ e assim $\left\langle v\middle|(V_f - V_f^*)w\right\rangle = 0$ para qualquer escolha de vetores v e w, o que implica que devemos ter $V_f = V_f^*$, como afirmamos.

Portanto o espaço de estados da álgebra C^* definida por $M_n(\mathbb{C})$ corresponde ao espaço de elementos de $M_n(\mathbb{C})$ hermitianos, positivos e de traço unitário. Voltaremos a encontrar estes estados no capítulo 9.

5.3 Espectro de Elementos da Álgebra

Considere uma álgebra C^*, A. que tem uma unidade, denotada por 1. Para cada elemento a de A podemos definir um conjunto bastante importante que é chamado de espectro de a. Para defini-lo vamos de início introduzir um outro conjunto, o resolvente de a, denotado por $\rho(a)$ e definido como sendo

$$\rho(a) = \{\lambda \in \mathbb{C} \ : \ \text{existe } (a - \lambda 1)^{-1}\}$$

(em geral escrevemos apenas $a - \lambda$ e não $a - \lambda 1$). O espectro de a, denotado por $\sigma(a)$, então é definido como sendo o conjunto complementar de $\rho(a)$ em $\mathbb{C}$, isto é, $\sigma(a) = \mathbb{C} \setminus \rho(a)$. Ele é então o conjunto de números complexos λ tais que $(a-\lambda)$ não tem um elemento inverso na álgebra.

Exemplo 5.6. *Seja A a álgebra (com unidade)*

$$A = C[0,1] = \{f\colon [0,1] \to \mathbb{C},\ f \textit{ contínua}\}$$

e $a \in A$ o elemento que é a função $a(x) = x^2$. Para obtermos o espectro de a devemos procurar os números complexos λ tais que $(a-\lambda)$ não tem inverso; mas então devemos saber quando não se pode inverter (no sentido da álgebra) um elemento do tipo $x^2 - \lambda$ (onde x

varia entre 0 *e* 1*). Se* λ *é algum elemento de* $[0,1]$ *então a função* $x^2 - \lambda$ *se anula em algum ponto e assim não pode ser invertida; caso contrário, se* λ *não é um elemento de* $[0,1]$*, então a função nunca se anula e sempre admite inversa. Portanto concluímos que o espectro de* a *é* $[0,1]$*. O leitor pode pensar um pouco mais no assunto para concluir que o espectro de* $a \in A$ *é de fato a imagem da função* $a(x)$ *no intervalo.*

Voltemos agora à álgebra que mais aparece nessas páginas

Exemplo 5.7. *Considere* $A = M_n(\mathbb{C})$*. Dada uma matriz* $a \in A$ *seu espectro* $\sigma(a)$ *é composto pelos* $\lambda \in \mathbb{C}$ *tais que* $a - \lambda 1$ *não admite inverso. Mas já sabemos que para matrizes a existência de um inverso está intimamente ligada ao determinante: de fato* $(a - \lambda 1)$ *não tem inverso se, e somente se,*

$$\det(a - \lambda 1) = 0,$$

o que é uma equação algébrica de grau n *em* λ *cujas soluções são o espectro de* a*; já encontramos esse objeto antes e as soluções, nesse caso, também já ganharam o nome de autovalores da matriz* a*. Portanto acabamos de concluir que o espectro de* a *é formado pelos seus autovalores.*

5.4 Exercícios

Exercício 5.1. *Considere uma álgebra* C^* A *com unidade* 1*, isto é, com um elemento* 1 *tal que* $1a = a1 = a$ *para todo* $a \in A$*. Mostre que* $1^* = 1$*; mostre que* $\|1\| = 1$*.*

Exercício 5.2. *Considere a álgebra* C^* $A = M_2(\mathbb{C})$*. Se* $\varphi\colon A \to \mathbb{C}$ *é um funcional linear e que satisfaz a igualdade* $\varphi(a)\varphi(b) = \varphi(ab)$ *para todo par* a *e* b *em* A *então mostre que* φ *é o funcional nulo, ou seja,* $\varphi(a) = 0$ *para todo* $a \in A$ *(obs.: o resultado continua verdadeiro se trocamos* $M_2(\mathbb{C})$ *por* $M_n(\mathbb{C})$*).*

Interlúdio

Agora sim começaremos com a mecânica quântica.

Nesta parte do texto, a menos que o leitor já conheça boa parte do assunto, saltos não são recomendados. A parte principal de cada capítulo trabalha com os conceitos e ferramentas da mecânica quântica, mas sem nunca descer aos detalhes de como implementar estas discussões em laboratórios. Por não conseguir resistir à tentação de falar de física, o final de cada capítulo tem esse enfoque[1].

Como sempre, tão ou mais difícil do que escrever foi escolher sobre o que não escrever. Se você discordar das nossas escolhas, pode nos contactar e comentar. Mas antes, tente seguir a música desse interlúdio...

[1] Para assim permitir que o leitor com o gosto complementar pule tais secções, passando ao capítulo seguinte.

Capítulo 6

Um Bit de Mecânica Quântica

Vamos começar a tratar a mecânica quântica por seu exemplo mais simples: sistemas de dois níveis, também chamados *bits quânticos*, ou simplesmente *qbits*. Deliberadamente, vamos fugir da estratégia de apresentar uma definição geral e depois descrever exemplos especiais. Vamos, ao longo do texto, redefinindo alguns conceitos de modo a torná-los mais e mais gerais. Assim, as definições apresentadas neste capítulo são precisas apenas quando restritas a este capítulo. Ainda que pareça inconsistente, acreditamos ser didaticamente acertado.

6.1 Mecânica Quântica em Dimensão Dois

Vamos introduzir a Mecânica Quântica partindo de seu exemplo não-trivial mais simples: o bit quântico. Um bit clássico é uma variável aleatória que pode assumir dois valores, por exemplo 0 ou 1. O bit quântico, porém, declara os estados extremais 0 e 1 uma base ortogonal para o *espaço de estados* do sistema. Essa frase simples inclui várias afirmações nas entrelinhas. Vamos detalhá-las.

6.1.1 Estados e Medições

Todo sistema quântico possui um *espaço de estados*, E, que é um espaço vetorial complexo com produto escalar. Neste capítulo, $\dim(E) = 2$. Na descrição mais simples[1] de mecânica quântica, o estado de um sistema é definido por um vetor unitário em seu espaço de estados. Toda e qualquer predição sobre o sistema pode ser feita a partir do conhecimento de seu estado. Para uso nesse capítulo, adotemos:

Definição 6.1. *O* estado *de um sistema é um vetor normalizado em seu espaço de estados.*

O leitor não deve se esquecer que o espaço de estados é um espaço vetorial sobre os complexos. Assim, o espaço de estados de um qbit é isomorfo a $\mathbb{C}^2$. Uma base para o espaço de estados será dada por dois vetores linearmente independentes, $\{|e_1\rangle, |e_2\rangle\}$. Como as alternativas clássicas de um bit costumam ser denotadas 0 e 1 e a notação de Dirac prescinde de uma letra para designar o vetor (a própria figura do ket já nos indica sua presença), é comum utilizarmos a base $\{|0\rangle, |1\rangle\}$. O leitor deve ter muito cuidado para não confundir $|0\rangle$ com a origem do espaço vetorial. Claramente este não é o caso, pois $|0\rangle$ e $|1\rangle$ são linearmente independentes. Como tais vetores correspondem a alternativas clássicas[2] distintas, temos ainda que esta base é *ortonormal*. Chegamos assim à importante noção de *teste*, apresentada aqui para qbits:

Definição 6.2. *Um* teste com alternativas clássicas a e b *é associado a uma base ortonormal, denotada* $\{|a\rangle, |b\rangle\}$. Aplicar um teste *pode ser visto como decompor o vetor com relação a esta base, para em seguida selecionar apenas uma das alternativas.*

Definida uma base, todo vetor do espaço de estados pode ser escrito como combinação linear destes elementos. Para um qbit, então, seu estado será descrito por

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \tag{6.1}$$

[1] Consequentemente, mais restrita. Mas, como descrito acima, usaremos essa estratégia didática, com a promessa que, ao final, o leitor terá uma axiomatização bastante geral.

[2] Ao longo do texto, usaremos o termo *alternativas clássicas* com o sentido ainda mais restritivo de *alternativas clássicas e exclusivas*, ou seja, uma e apenas uma acontece.

onde α e β são números complexos, e a normalização exige $|\alpha|^2 + |\beta|^2 = 1$.

Exercício 6.1. *Lembrando que* $\||\psi\rangle\|^2 = \langle\psi|\psi\rangle$*, obtenha a condição de normalização apresentada acima.*

Uma das grandes novidades da mecânica quântica aparece na sua regra sobre como relacionar o estado $|\psi\rangle$ à medição das alternativas clássicas. Em benefício da clareza, vamos continuar com sistemas de dimensão 2, mas o leitor já pode tentar generalizar esta definição para dimensões arbitrárias[3].

Postulado 6.1. *Se um sistema quântico no estado* $|\psi\rangle$ *da eq.* (6.1) *é sujeitado a um* teste *com alternativas clássicas* 0 *e* 1*, a probabilidade de obter o resultado correspondente a* 0 *é dada por* $|\alpha|^2$*, enquanto a de obter* 1 *é dada por* $|\beta|^2$*.*

Os coeficientes α e β da expansão do estado $|\psi\rangle$ com respeito à base $\{|0\rangle, |1\rangle\}$ são números complexos que permitem calcular probabilidades. Feynman batizou tais coeficientes *amplitudes de probabilidades*, ou simplesmente *amplitudes*. Uma vasta gama de efeitos da mecânica quântica está ligada ao fato de podermos somar amplitudes não-nulas e obter um resultado nulo (ou muito pequeno). Este é o chamado fenômeno de interferência destrutiva, já conhecido nos fenômenos ondulatórios, mas impossível para probabilidades, que são números reais não-negativos. É o caso, por exemplo, no experimento de dupla fenda, onde regiões "escuras" aparecem quando as duas fendas estão abertas, onde haveria contagens para cada uma das fendas abertas isoladamente.

Mesmo sem querer desviar para discussões sobre fundamentos de mecânica quântica, é necessário dizer que esta foi a primeira vez que uma teoria científica se assumiu probabilística *a priori*. Mesmo que conheçamos o estado $|\psi\rangle$ de uma partícula, o resultado de observações será, em geral, probabilístico. O leitor pode comparar esta situação com a da mecânica estatística. Nesta, o conceito de probabilidades foi introduzido com a justificativa que, *na prática*, não podemos dar uma descrição precisa para um sistema macroscópico. De certa forma, é uma concessão que mentes determinísticas fizeram à dificuldade

[3] Assunto do próximo capítulo.

de trabalhar com 10^{23} coordenadas, ou mais. Mas mantinha-se a convicção que *em princípio* poderia se descrever microscopicamente um gás, por exemplo. Na mecânica quântica não; exceto se α ou β for zero, a mais completa descrição microscópica é incapaz de prever, senão probabilisticamente, o resultado do teste 0 ou 1.

Esta descrição probabilística da mecânica quântica tem uma consequência fundamental: embora gostemos muito de tratar de um sistema quântico específico, as previsões desta teoria só podem ser testadas quando preparamos igualmente um grande número de cópias do sistema, e agimos igualmente sobre todas elas (e assim poderemos comparar as frequências obtidas com as probabilidades previstas). Neste sentido, é comum pensar que o estado de um sistema é a descrição de um *ensemble*[4] e que um sistema isolado deve ser pensado como um elemento aleatório deste ensemble.

6.1.2 Depois das Medições

Como relacionamos as alternativas clássicas 0 e 1 com a base ortonormal $\{|0\rangle, |1\rangle\}$, é natural introduzir o seguinte

Postulado 6.2. *Após a realização de um teste para discriminar entre as alternativas clássicas* 0 *e* 1*, se o resultado obtido foi* 0*, o sistema passa a ser descrito pelo estado* $|0\rangle$*; se o resultado obtido foi* 1*, o sistema passa a ser descrito pelo estado* $|1\rangle$*.*

Este postulado está naturalmente associado à noção de *reprodutibilidade de testes*. Ou seja, se um teste é realizado e se obtém um resultado, repetições deste mesmo teste no mesmo sistema corroborarão o resultado obtido[5]. É importante distinguir aqui entre "agir novamente no mesmo sistema" e "realizar o teste em outro elemento do ensemble". Por construção da ideia de ensemble, seus elementos são independentes. Assim, embora sigam a mesma distribuição de

[4]Ensemble é a palavra francesa para *conjunto*. Ganhou destaque e uso próprio na mecânica estatística e na mecânica quântica correspondendo a esta noção de conjunto infinito de realizações de um certo estado.

[5]Ainda não falamos sobre evolução temporal de estados. Neste momento, adotamos tacitamente uma lei de inércia: se nada for feito, o sistema continua no mesmo estado.

probabilidade, seus resultados são independentes[6]. Agir novamente no mesmo sistema é repetir o mesmo teste duas vezes, *no mesmo representante do ensemble*. O que a definição 6.2 diz é que se fizermos esta repetição do teste, o ensemble original será dividido em apenas dois subensembles: aquele onde as duas aplicações do teste resultaram 0 e aquele onde ambas resultaram 1. E se repetirmos N vezes, ainda assim só obteremos dois subensembles: aquele onde as N repetições do teste resultaram 0 e aquele em que os N resultados foram 1.

Vale notar que submeter um sistema a um certo teste e selecionar apenas os resultados "favoráveis" pode ser entendido como uma *preparação*: se queremos preparar o estado $|0\rangle$, submetemos o sistema a um teste que discrimina 0 e 1 e descartamos todos os sistemas em que o resultado 1 for obtido.

Exercício 6.2. *Redescreva o parágrafo acima usando a ideia de subensemble.*

É ainda importante insistir que esta distinção entre agir novamente no mesmo sistema e realizar o mesmo experimento em um elemento independente do ensemble *não é* uma das peculiaridades da mecânica quântica. Vamos então discutir um exemplo clássico: o sorteio da mega-sena. A melhor maneira que temos para descrever o resultado do sorteio do concurso N da mega-sena é uma distribuição uniforme sobre todas as combinações de números permitida nesta modalidade de loteria[7]. Porém, uma vez escolhido N, a situação muda um pouco de figura. Se N corresponde a um sorteio já realizado, mas não dispomos do resultado, nossa melhor descrição continua sendo dada pela distribuição uniforme. Porém, uma vez conhecido o resultado, passamos a descrevê-lo, probabilisticamente, por uma distribuição concentrada no resultado conhecido. Em particular, se estamos preocupados com o concurso 1000 da mega-sena ($N = 1000$), as dezenas sorteadas foram $29, 38, 39, 49, 53$ e 58. Assim, se repetirmos o teste (clássico) de "sortear" o concurso 1000 da mega-sena, devemos obter o mesmo resultado, diferentemente do caso de

[6] Correspondendo à situação típica de textos de probabilidade e estatística das variáveis *i. i. d.*, ou seja, variáveis independentes e identicamente distribuídas.

[7] Acreditamos, a priori, que cada dezena é equiprovável - o que leva a equiprobabilidade das combinações - , e que os diferentes concursos são independentes.

realizar um sorteio de um outro "concurso". Insistindo uma última vez: se denotamos por M_N a variável aleatória que dá o resultado do concurso N da mega-sena, M_N é independente de $M_{N'}$ se, e somente se, $N \neq N'$.

6.1.3 O que os bits clássicos não têm

A noção de *teste* não é exclusiva da mecânica quântica. A ideia de reprodutibilidade também não (sempre ignorada a evolução temporal do sistema). O que realmente distingue a mecânica quântica da sua contrapartida clássica é a existência de testes *incompatíveis*.

Definição 6.3. *Um teste B é dito compatível com um teste A se a realização de B entre duas repetições de A não afeta a reprodutibilidade do teste A.*

Classicamente, o único teste (não-trivial) que podemos fazer com um bit é verificar se ele vale 0 ou 1. Lembremos que sua versão quântica está associada a uma base ortonormal $\{|0\rangle, |1\rangle\}$ do espaço de estados E. Mas podemos escolher livremente outra base para E. A exigência de serem alternativas classicamente distinguíveis impõe ortonormalidade.

Como um exemplo, podemos definir os vetores:

$$\begin{aligned} |+\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle); \\ |-\rangle &= \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \end{aligned} \tag{6.2}$$

Exercício 6.3. *Mostre que $\{|+\rangle, |-\rangle\}$ é uma base ortonormal.*

Podemos aplicar o teste + ou −, que corresponde a esta base. Devemos aplicar a este teste as mesmas regras que antes usávamos para 0 e 1, com sua correspondente base. Chamemos o teste 0 ou 1 de Z e o teste + ou − de X, devido a uma convenção que ficará clara na 6.1.5.

Exercício 6.4. *Relação entre os testes X e Z.*

1. *Considere o estado inicial $|0\rangle$. Quais as probabilidades de cada alternativa para o teste Z? E para o teste X?*

2. *Suponha que foi realizado o teste X e obtido o resultado $+$. Qual a probabilidade de obter* 0 *em uma realização subsequente do teste Z?*

O que o exercício acima mostra é que os testes X e Z não são compatíveis! Se fizermos sequencialmente os testes Z, X e Z, é possível obter, respectivamente, as respostas 0, $+$ e 1. Se não fosse realizado o teste X entre as duas realizações de Z, jamais poderíamos obter 0 e 1 como respostas, devido à reprodutibilidade dos testes.

Vamos discutir essa situação em mais detalhe. Feito o primeiro teste Z, se obtido o resultado 0, sabemos que devemos passar a descrever o sistema pelo estado $|0\rangle$. Neste estado, o teste X terá o resultado $+$ ou $-$ de maneira equiprovável. Com isso, a melhor descrição do sistema será dada por $|+\rangle$ no primeiro caso e $|-\rangle$ no segundo. Em ambas as alternativas, o novo teste Z também terá os resultados 0 ou 1 de maneira equiprovável. Como $|0\rangle = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle)$, o que percebemos aqui é que *não* podemos pensar em $|0\rangle$ como uma simples mistura equiprovável das alternativas $+$ e $-$. Se um teste Z é aplicado ao $|0\rangle$ a resposta é 0, sempre. Essa é uma manifestação do fenômeno de *interferência*: as alternativas clássicas $+$ e $-$ não se misturam de maneira equiprovável, mas se combinam *coerentemente*. Já nesse caso temos a manifestação do que chamamos *interferência construtiva*, para o caso do resultado 0 (pois a "soma clássica") das alternativas levaria ao resultado 0 com probabilidade $\frac{1}{2}$, e quanticamente o resultado é maior (nesse caso, 1), bem como da *interferência destrutiva*, para o caso do resultado 1.

Exercício 6.5. *Bases mutuamente neutras.*

1. *Descreva um teste com resultados a e b, onde o estado $|0\rangle$ dá probabilidades p e $1-p$.*

2. *Seja $|a\rangle$ o estado correspondente à alternativa a do teste anterior. Qual a probabilidade de obter* 0 *se um teste Z for aplicado a este estado?*

3. *Duas bases $\mathcal{B} = \{|b_0\rangle, |b_1\rangle\}$ e $\mathcal{C} = \{|c_0\rangle, |c_1\rangle\}$ são ditas mutuamente neutras se $|\langle b_i|c_j\rangle|$ é independente de i e j. Mostre que as bases $\mathcal{Z} = \{|0\rangle, |1\rangle\}$ e $\mathcal{X} = \{|+\rangle, |-\rangle\}$ são bases ortonormais mutuamente neutras.*

4. *Obtenha uma nova base, $\mathcal{Y}$, mutuamente neutra tanto com $\mathcal{X}$ quanto com $\mathcal{Z}$.*

5. *Mostre que não existe outra base mutuamente neutra com $\mathcal{X}$, $\mathcal{Y}$ e $\mathcal{Z}$.*

6.1.4 Quando perder é ganhar

Algumas tarefas muito simples do ponto de vista abstrato podem ser muito difíceis na prática[8]. Por exemplo, gerar números aleatórios. Um pensamento inocente diz que lançar uma moeda para cada bit (cara ou coroa) seria o suficiente. Mas não! Como garantir que a moeda é realmente honesta? Ou ainda, que seu lançamento é honesto?

Novamente atingimos o paradigma teórico onde aleatoriedade não surge a priori, mas da dificuldade de definir as condições iniciais com precisão, e de uma dinâmica muito sensível a tais condições. Os geradores de números "aleatórios" mais utilizados são sofisticações deste lançamento da moeda. Computadores calculam funções determinísticas mas extremamente sensíveis às condições iniciais, e estas condições iniciais envolvem dados razoavelmente aleatórios, como os últimos dígitos do relógio interno do computador, ou bits escolhidos dentro de um arquivo do qual nada se sabe... O que se obtém daí são números "suficientemente aleatórios" para a imensa maioria das aplicações: jogos de computador, simulações de Monte Carlo, geração de números primos muito grandes...

Mas a noção de "suficientemente aleatórios" é sutil. O que é suficientemente aleatório para quem só quer gerar números primos para criar uma chave RSA [Cou, Sin] e usar na sua correspondência eletrônica privada pode não ser suficientemente aleatório para um banco que opera pela internet. O que é suficientemente aleatório para quem só quer se divertir com um jogo pode não ser suficientemente aleatório para uma empresa de jogos de azar *on line*! Pode parecer estranho, mas uma interessante aplicação da mecânica quântica[9] é aproveitar a existência de testes incompatíveis para produzir números "suficientemente aleatórios".

[8] E vice-versa.

[9] Já com algum sucesso comercial[.com].

Exercício 6.6. *Usando o que você já aprendeu até o presente momento, proponha uma máquina quântica de gerar bits aleatórios*[10].

De fato, já há trabalhos na linha de considerar aleatoriedade como um *recurso*, tão valioso quanto outros que ainda discutiremos nesse texto.

6.1.5 Estados Físicos e Esfera de Bloch

Um primeiro ponto a ser levantado é que, em mecânica quântica, o vetor de estado (correspondente a uma preparação) permite calcular todas as probabilidades dos possíveis resultados de testes realizados naquele sistema. Cada teste é associado a uma base ortonormal e as probabilidades são dadas pelos módulos ao quadrado das amplitudes de probabilidade, ou seja, dos coeficientes da expansão do vetor com respeito àquela base ortonormal específica.

Exercício 6.7. *Dois vetores $|\psi\rangle$ e $e^{i\phi}|\psi\rangle$, com $\phi \in \mathbb{R}$, representam estados equivalentes, no sentido que as mesmas probabilidades são previstas para todos os testes realizados.*

Vamos, a seguir, explorar as consequências desta identificação apontada pelo exercício 6.7, no caso de um qbit. Antes, um pouco de nomenclatura: tanto um número complexo unitário $e^{i\phi}$ quanto o número real ϕ são comumente chamados de *fase*. O exercício acima é normalmente fraseado como "uma fase global é irrelevante".

Como discutido na 4.7, vetores unitários de $\mathbb{C}^2$ formam uma esfera S^3, mas a identificação do exercício 6.7 faz com que cada ponto possua uma fibra S^1 (as possíveis fases globais) e o espaço topológico formado pelos estados fisicamente distintos corresponde a uma esfera S^2. Ainda que esta construção corresponda à fibração de Hopf, em mecânica quântica costumamos associar outro nome à esfera S^2 que corresponde aos estados fisicamente distintos: trata-se da *Esfera de Bloch*.

Cada classe de equivalência pode ser representada por um vetor de estado

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

[10] Mais precisamente, gerar bits independentes, identicamente distribuídos, com distribuição equiprovável.

É comum utilizar a seguinte parametrização

$$|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\varphi}\,\mathrm{sen}\frac{\theta}{2}|1\rangle, \tag{6.3}$$

com $\theta \in [0,\pi]$ e $\varphi \in [0,2\pi]$. Naturalmente[11], esta parametrização possui singularidades. Neste caso, correspondendo aos valores $\theta = 0,\pi$, aos quais estão associados os vetores da base $\mathcal{Z}$. O exercício 6.8 mostra a conveniência de tal convenção.

Exercício 6.8. *Esfera de Bloch*

1. *Verifique que a parametrização* (6.3) *cobre todas as classes de vetores de estado fisicamente distintos.*
2. *Interprete os ângulos θ e φ de um ponto arbitrário e verifique que todos os pontos da esfera foram utilizados na parametrização.*
3. *Calcule o produto escalar $\langle 0|\psi\rangle$ e discuta a diferença entre os vetores da esfera de Bloch serem ortogonais e a posição de vetores de estado ortogonais na esfera de Bloch.*

Aproveitemos esta discussão para introduzir outras ferramentas bastante úteis na discussão de um qbit, as chamadas *matrizes de Pauli*. Estas são matrizes de automorfismos de $\mathbb{C}^2$, que escritas com respeito à base $\mathcal{Z}$ tomam a forma

$$\sigma_z = \quad Z = \quad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \tag{6.4a}$$

$$\sigma_x = \quad X = \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \tag{6.4b}$$

$$\sigma_y = \quad Y = \quad \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}. \tag{6.4c}$$

Note que as bases $\mathcal{X}$, $\mathcal{Y}$ e $\mathcal{Z}$ são as respectivas bases de autovetores dos operadores descritos acima. É comum (pelo menos como artifício de notação) considerar que estas matrizes formam um vetor de matrizes $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$, de modo que, para um vetor $\vec{v} = (v^x, v^y, v^z) \in \mathbb{R}^3$, o

[11] Qual resultado matemático está por trás deste "naturalmente"?

produto $\vec{v} \cdot \vec{\sigma}$ representa a matriz $\sum_i v^i \sigma_i$. A dupla notação utilizada (*e.g.*: σ_x e X) se deve a uma ser a notação padrão em textos de mecânica quântica, a outra a notação padrão em textos de informação quântica. Vamos utilizar ambas.

Exercício 6.9. *1. Obtenha autovalores e autovetores para X, Y e Z.*

2. Para um vetor unitário $\vec{u} \in S^2$, diagonalize $\vec{u} \cdot \vec{\sigma}$. Represente seus autovetores na esfera de Bloch.

3. Qual a relação entre os autovetores de $\vec{u} \cdot \vec{\sigma}$ e de $-\vec{u} \cdot \vec{\sigma}$?

O exercício acima mostra uma maneira canônica de relacionar um operador a cada base ortonormal de $\mathbb{C}^2$. Sendo mais preciso, relacionamos um operador a cada decomposição de $\mathbb{C}^2$ em dois subespaços unidimensionais ortogonais. Por sua vez, nos ajuda a entender melhor todos os possíveis testes a serem realizados com um qbit e a visualizá-los na esfera de Bloch: cada teste corresponde à escolha de um eixo, com seus pontos antípodas sendo os vetores da base correspondente. A nomenclatura para as bases $\mathcal{X}$, $\mathcal{Y}$ e $\mathcal{Z}$ também deve estar mais clara agora.

6.1.6 Evolução Temporal

Até agora tratamos de estados e medições e até já arriscamos uma visualização geométrica para ambos. Mas entre uma preparação e uma medição o estado do sistema pode variar. No mesmo espírito desse capítulo, vamos tratar agora do caso mais simples de evolução temporal em mecânica quântica.

Para um sistema isolado, a evolução temporal de um estado inicial será ditada pela equação de Schrödinger:

$$\frac{d}{dt}|\psi\rangle = \frac{H}{i\hbar}|\psi\rangle, \tag{6.5}$$

onde $H : E \to E$ é um operador linear, chamado *hamiltoniano* do sistema, i a unidade imaginária e $\hbar$ a famosa constante de Planck (dividida por 2π).

No capítulo 3 já vimos que

$$|\psi(t)\rangle = \exp\left(\frac{-iHt}{\hbar}\right)|\psi_0\rangle \tag{6.6}$$

é a solução da equação (6.5) com a condição inicial $|\psi(0)\rangle = |\psi_0\rangle$. De fato, como queremos manter a norma do vetor $|\psi(t)\rangle$, segue que devemos trabalhar com $H = H^*$, ou seja, o hamiltoniano deve ser autoadjunto.

Vale notar uma importante propriedade:

Exercício 6.10. *Mostre que os operadores autoadjuntos em $\mathbb{C}^2$ formam um espaço vetorial real*[12]*. Mostre ainda que, fixada uma base para $\mathbb{C}^2$, a matriz identidade e as matrizes de Pauli* (6.4) *formam uma base para este espaço.*

Explicitamente, isso significa que todo operador autoadjunto A é descrito por quatro números reais (a_I, a_x, a_y, a_z), de modo que a matriz que representa A seja dada por

$$\begin{bmatrix} a_I + a_z & a_x - ia_y \\ a_x + ia_y & a_I - a_z \end{bmatrix}.$$

Se estamos dispostos a identificar vetores que descrevem estados equivalentes (ou seja, se queremos descrever a evolução temporal na esfera de Bloch, e não em $\mathbb{C}^2$), a componente H_I não terá qualquer efeito:

Exercício 6.11. *Mostre que se $H = h_I I$, a evolução temporal de qualquer estado é dada pelo acúmulo de fase global, deixando seu vetor de Bloch fixo.*

Assim, sabendo também que a identidade comuta com qualquer operador[13], podemos nos concentrar em hamiltonianos da forma $H = h_x X + h_y Y + h_z Z$. Comecemos pelo mais fácil. Seja $H = h_z Z$. Queremos calcular

$$\exp\left(\frac{-iHt}{\hbar}\right) = \exp\left(\frac{-ih_z Zt}{\hbar}\right).$$

[12] Ou seja, sobre o corpo $\mathbb{R}$.
[13] Por que isso é importante?

Mas conhecemos um operador quando sabemos como ele atua em uma base. E, para a base $\mathcal{Z} = \{|0\rangle, |1\rangle\}$ temos

$$\begin{aligned}\exp\left(\frac{-ih_z Zt}{\hbar}\right)|0\rangle &= e^{\frac{-ih_z t}{\hbar}}|0\rangle,\\ \exp\left(\frac{-ih_z Zt}{\hbar}\right)|1\rangle &= e^{\frac{ih_z t}{\hbar}}|1\rangle,\end{aligned}$$

de onde, se definirmos $\omega = \frac{2h_z}{\hbar}$, teremos para

$$|\psi_0\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\phi}\ \mathrm{sen}\frac{\theta}{2}|1\rangle$$

a solução dada por

$$\begin{aligned}|\psi(t)\rangle &= e^{-i\frac{\omega}{2}t}\cos\frac{\theta}{2}|0\rangle + e^{i\frac{\omega}{2}t}e^{i\phi}\ \mathrm{sen}\frac{\theta}{2}|1\rangle\\ &\equiv \cos\frac{\theta}{2}|0\rangle + e^{i(\phi+\omega t)}\ \mathrm{sen}\frac{\theta}{2}|1\rangle,\end{aligned}$$

o que nos permite interpretar a evolução temporal dada por este hamiltoniano como a rotação da esfera de Bloch em torno de seu eixo z, com velocidade angular ω.

O exercício seguinte generaliza esta discussão:

Exercício 6.12. *Obtenha os autovalores e autovetores de $H = h_x X + h_y Y + h_z Z$, redefina ω como a diferença entre os autovalores e escreva a evolução temporal de um estado arbitrário (sugestão: use a base de autovetores). Por fim, interprete tal evolução temporal em termos do vetor de Bloch.*

Com isso, interpretamos, em termos do vetor de Bloch, todas as possíveis evoluções temporais de um sistema de dois níveis: os autovetores de seu hamiltoniano definem um eixo, enquanto os autovalores definem a velocidade angular com que a esfera de Bloch rotaciona em torno deste eixo. Naturalmente, há dois, e exatamente dois, pontos fixos ao longo de tal evolução.

6.2 Um pouco de Física

Ao longo deste capítulo descrevemos a mecânica quântica de um qbit, sujeito às chamadas medições projetivas e a evolução hamiltoniana.

Se, por um lado, são várias restrições (dimensão do espaço de estados, tipo de medição e tipo de evolução temporal), por outro se trata de algo bastante geral, com uma grande coleção de exemplos.

Um olhar mais atento pode ter reparado na discussão anterior que $\hbar$ tem dimensão de energia vezes tempo[14]. A grandeza mecânica que possui tal dimensão é o *momentum angular*. Um exemplo natural e importante de sistema de dois níveis é o momentum angular intrínseco de algumas partículas; o chamado *spin* das partículas de spin $\frac{1}{2}$ (partículas com spins maiores terão espaços de estado com dimensão maior). Para fixar ideias, pensemos que tal partícula é um elétron, portanto, uma partícula com carga elétrica. Uma carga elétrica com momentum angular tem certa similaridade com uma corrente elétrica, gerando momentum de dipolo magnético. Um dipolo magnético interage com campos magnéticos e uma maneira de muitos físicos tratarem a discussão da evolução temporal de um qbit é usando o chamado modelo de "pseudo-spin". O sistema de dois níveis, seja ele qual for, pode ser pensado como um spin $\frac{1}{2}$. E o hamiltoniano que vai ditar sua evolução temporal (autônoma) pode sempre ser associado a um campo magnético constante. Assim, a direção do spin vai precessionar em torno do campo e a velocidade de tal precessão será ditada pela intensidade do campo.

Tal imagem gera uma linguagem interessante, típica do contexto de ressonância nuclear magnética[15], mas que ganhou espaço também em outras comunidades. Em especial, se um campo magnético é aplicado em direção perpendicular à do vetor de spin, a trajetória descrita na evolução temporal será dada por grandes círculos. Se o tempo de interação for ajustado de forma a metade desse círculo ser percorrido, chamaremos essa evolução de um pulso π. Note que um pulso π essencialmente inverte a direção em que aponta o vetor de spin[16]. Da mesma forma, se a interação se der por um quarto de

[14]Para sermos justos, atenção não seria suficiente. O leitor teria que saber, por outras fontes, que um hamiltoniano tem unidades de energia, ou tirar a mesma conclusão da equação de Einstein $E = \hbar\omega$, que não discutimos aqui.

[15]Sim, a mesma presente em exames clínicos. RNM, para sua sigla em português, NMR em inglês.

[16]Um resultado interessante, mas que só trataremos mais adiante, é que não existe uma evolução quântica capaz de inverter o vetor de spin, qualquer que seja ele. Você consegue conciliar essa última afirmação com a discussão desse parágrafo?

volta, diremos que foi aplicado um pulso $\frac{\pi}{2}$. Caso se buscasse um linguajar mais preciso, deveria-se dizer em que direção foi feito tal pulso $\frac{\pi}{2}$, mas normalmente isso fica subentendido no contexto. Vale notar que um pulso $\frac{\pi}{2}$ é uma excelente forma de passar de um estado da base $\mathcal{Z}$ para um estado da base $\mathcal{X}$, por exemplo.

Um outro sistema quântico que pode ser bem entendido nesta discussão de qbits é o tradicional experimento de fenda dupla, reportado pela primeira vez por Young, em 1800, utilizando luz. De fato, o experimento de Young foi a maior evidência experimental a favor do caráter ondulatório da luz. Pouco mais de cem anos depois, passou-se a entender que, com relação ao experimento de fenda dupla, luz e matéria se comportam da mesma forma.

Nessa descrição, pode-se entender os estados da base $\mathcal{Z}$ como os "estados de fenda", ou seja, como seria descrito o sistema caso apenas uma das fendas estivesse aberta. Já o sistema com as duas fendas abertas será descrito pelo estado $|+\rangle$, de superposição das duas fendas. Conforme o ponto de observação em uma tela[17] adequadamente afastada do anteparo com as fendas, cada componente ($|0\rangle$ ou $|1\rangle$) acumula diferentes fases, correspondendo a uma evolução temporal onde o pseudospin precessaria devido a um campo aplicado na direção Z.

Para ser mais exato, o caso da fenda dupla não corresponde tão precisamente assim a uma evolução de pseudospin, visto que conforme nos deslocamos na tela, estaremos mais próximos a uma das duas fendas, aumentando sua participação no estado correspondente. Outro sistema físico segue esta descrição mais de perto: o interferômetro de Ramsey.

Apresentado em 1949, o interferômetro de Ramsey lhe rendeu o Prêmio Nobel de Física em 1989, sendo uma generalização (em termos de sistema) e um aperfeiçoamento (em termos da ideia central de separar os pulsos) da ressonância magnética, criada por Rabi em 1939, também lhe rendendo o Prêmio Nobel de Física em 1947.

A ideia de Rabi é aproveitar a existência de níveis de energia distintos e interagir com uma transição entre dois deles, utilizando para isso a noção clássica de ressonância. No caso de Rabi, o sistema era um núcleo e a "força externa" um campo magnético, daí ser uma

[17] Ou detector, dependendo do regime em que o experimento é realizado [Ter05].

ressonância nuclear magnética. No caso de Ramsey, o sistema é um átomo e a força externa um campo eletromagnético. Esta é a origem do relógio atômico, que nos fez inclusive rever a definição de um segundo utilizando para isso a frequência da radiação emitida por uma transição específica do átomo de Césio.

Capítulo 7

Sistemas de d níveis

Devidamente explorado o caso mais simples, vamos passar ao caso "um pouco menos simples". Este capítulo é dedicado aos sistemas quânticos de d níveis, ou seja, à mecânica quântica em espaço de estados com dimensão finita.

7.1 Mecânica Quântica em Dimensão d

Agora precisamos refazer a discussão da secção 6.1. A principal diferença advém do fato que, agora, um teste pode distinguir entre menos alternativas que a dimensão do espaço.

7.1.1 Estados e Medições

Como já afirmamos, todo sistema quântico possui um *espaço de estados* que é um espaço vetorial complexo com produto escalar, E. Neste capítulo, a única exigência é que $\dim(E) = d < \infty$. Ainda na descrição mais simples e restrita da mecânica quântica, o estado de um sistema é definido por um vetor unitário em seu espaço de estados. Toda e qualquer predição sobre o sistema pode ser feita a partir do conhecimento de seu estado. Para uso nesse capítulo, essencialmente repetimos a definição 6.1:

Definição 7.1. *O* estado *de um sistema é um vetor normalizado em seu espaço de estados.*

Assim como o espaço de estados de um qbit é isomorfo a $\mathbb{C}^2$, o espaço de estados para um sistema de dimensão d (por analogia, um *qdit*) será $E \cong \mathbb{C}^d$.

Agora devemos generalizar a definição 6.2, que é onde as diferenças aparecem.

Definição 7.2. *Seja E um espaço de estados. Um teste com alternativas distintas indexadas por i corresponde a uma decomposição ortogonal $E = \bigoplus_i E_i$.*

Que é complementada pelo

Postulado 7.1. *Sejam E um espaço de estados, $|\psi\rangle \in E$ um estado e $E = \bigoplus_i E_i$ um teste. Sejam ainda $P_i : E \to E$ os projetores ortogonais sobre cada E_i. A probabilidade de obter o resultado i é dada por $p_i = \langle\psi|P_i|\psi\rangle$.*

Exercício 7.1. *Projetores ortogonais e notação de Dirac*

1. *Seja $|\phi\rangle$ um vetor normalizado. O que faz o operador $|\phi\rangle\langle\phi|$?*
2. *Seja $\{|v_i\rangle\}_{i=1}^d$ uma base ortonormal. Defina $P_i = |v_i\rangle\langle v_i|$. Mostre que*[1] *$P_iP_j = \delta_{ij}P_j$.*
3. *Para $J \subset \{1, \ldots, d\}$ defina $P_J = \sum_{i\in J} P_i$. Mostre que $P_JP_K = P_{J\cap K}$. Em particular, $P_J^2 = P_J$.*
4. *Qual a forma diagonal de P_J? Interprete $\mathrm{Tr}P_J$, o traço de P_J.*

Exercício 7.2. *Mostre que a definição 6.2 é um caso particular da 7.2.*

A definição 7.2 e o exercício 7.1 podem ser unidos para chegar à forma mais comum de se descrever tais medições. Para cada E_i, escolha uma base ortonormal $\left\{\left|v_i^k\right\rangle\right\}$, onde o índice k corre de 1 até $d_i = \dim E_i$. Temos então uma base ortonormal para E, $\left\{\left|v_i^k\right\rangle\right\}$. Se escrevemos o vetor de estado $|\psi\rangle$ com respeito a essa base, temos $|\psi\rangle = \sum_i \sum_{k=1}^{d_i} \alpha_i^k \left|v_i^k\right\rangle$.

[1] $\delta_{ij} = 1$, se $i = j$; $\delta_{ij} = 0$, se $i \neq j$.

Exercício 7.3. *1. Mostre que p_i, a probabilidade de obter a alternativa i, é dada por $\sum_{k=1}^{d_i} \left|\alpha_i^k\right|^2$.*

2. Refaça esta discussão para o caso não-degenerado (i.e.: $d_i = 1, \forall\ i$) e compare com a definição 6.2.

Deve ficar claro porque esse tipo de medição é normalmente chamada uma *medição projetiva*. Medições mais gerais que estas serão discutidas no capítulo 9, juntamente com uma noção mais geral de estado.

Uma base para o espaço de estados será dada por d vetores linearmente independentes, $\{|e_1\rangle, |e_2\rangle, \ldots, |e_d\rangle\}$. Justamente pela associação de testes a bases ortonormais, é bastante comum que no contexto de mecânica quântica, salvo menção em contrário, bases sejam sempre ortonormais.

O Teorema Espectral permite associar esta noção de teste a uma outra noção, muito presente nos textos de mecânica quântica do século XX: a de *observável*[2]. Seja A um operador auto-adjunto. O teorema espectral nos diz que ele pode ser escrito como

$$A = \sum_i a_i P_i,$$

onde a_i são seus autovalores (reais) e P_i projetores sobre os respectivos auto-espaços. Assim, é comum, no chamado processo de *quantização canônica*, associar a cada grandeza da mecânica clássica um *observável* A, que é um operador auto-adjunto. A definição 7.2 passa a ser lida como: *os resultados possíveis para cada medição são dados pelos autovalores de A*, com as probabilidades previamente associadas. Podemos então calcular a esperança de A (também chamada valor médio, ou valor esperado), em um estado $|\psi\rangle$, dada por

$$\langle A\rangle = \sum_i a_i p_i = \sum_i a_i \langle\psi|P_i|\psi\rangle = \langle\psi| \sum_i a_i P_i |\psi\rangle = \langle\psi|A|\psi\rangle.$$

7.1.2 Depois das Medições

Aqui também teremos mudanças significativas em relação à situação de um qbit. É fácil entender a razão. Se submetíamos um qbit a

[2] Na secção 9.6 voltaremos ao conceito de observáveis, mas, naturalmente, em outro contexto.

um teste com duas alternativas clássicas, a decomposição imposta ao espaço de estados era "completa", no sentido que cada subespaço da soma direta tinha dimensão 1. Em um subespaço de dimensão 1, todos os vetores representam o mesmo estado físico, assim o estado após a medição não dependia do estado pré-medição e podíamos (arbitrariamente) escolher um vetor de estado pós-medição dependendo apenas do resultado de tal processo.

Postulado 7.2. *Considere um teste com alternativas clássicas i, dado pela decomposição $E = \bigoplus_i E_i$, com respectivos projetores ortogonais P_i. Se o teste foi aplicado ao estado $|\psi\rangle$ e a alternativa i foi obtida, após o teste o sistema será descrito pelo estado* $|\psi_i\rangle = \dfrac{P_i|\psi\rangle}{\|P_i|\psi\rangle\|}$.

O postulado 7.2 retém a principal propriedade do postulado 6.2: a reprodutibilidade dos testes.

Exercício 7.4. *Demonstre a afirmação acima.*

Por outro lado, traz a diferença marcante (natural e já comentada): o estado após a medição, $|\psi_i\rangle$, depende do estado antes da medição, $|\psi\rangle$.

Exercício 7.5. *Mostre que o postulado 6.2 pode ser visto como caso particular do postulado 7.2 se acrescentarmos a noção de equivalência de estados do exercício 6.7.*

A noção de compatibilidade continua presente aqui. De fato, toda a área de pesquisa associada à *lógica quântica* [Pit, Coh] nasce aqui. Dois testes serão compatíveis se, e somente se, existir uma decomposição ortogonal que é um refinamento[3] comum a ambos. Neste caso, estes dois testes podem ser realizados de maneira simultânea (ou, em outras palavras, a ordem em que são realizados não é importante) e tal realização simultânea é descrita pelo refinamento comum dado pelas intersecções dos subespaços associados a cada teste.

Exercício 7.6. *Considere dois testes $E = \bigoplus_i E_i$ e $E = \bigoplus_j F_j$ e sejam A e B observáveis associados a estes. Mostre que os teste são compatíveis se, e somente se, $[A, B] = 0$.*

[3]Uma decomposição ortogonal $E = \bigoplus_j F_j$ é um refinamento de $E = \bigoplus_i E_i$ se para todo j, F_j é subespaço de algum E_i.

Dessa forma, um conjunto de testes será mutuamente compatível quando existir um refinamento comum a todos e um teste será *completo* quando não pode mais ser refinado, ou seja, todos os subespaços envolvidos na decomposição são unidimensionais.

7.1.3 Geometria

Da mesma forma que para os qbits, a fase global é irrelevante quando tratamos das probabilidades dos resultados de testes. E isso traz riqueza para a geometria do problema.

Para qualquer sistema quântico com espaço de estados de dimensão finita d, os possíveis vetores de estado são vetores de norma 1 em $E \cong \mathbb{C}^d$, um conjunto naturalmente identificado com a esfera S^{2d-1} (lembrando que neste caso, a dimensão indicada é com respeito aos reais). O conjunto das classes de equivalência $[|\psi\rangle]$ pode ser visto como o conjunto de todos os subespaços unidimensionais (complexos) de $E \cong \mathbb{C}^d$. Mas esta é exatamente a definição do *espaço projetivo complexo* $\mathbb{CP}^{d-1}$. Em particular, o conjunto dos vetores de estado fisicamente distintos para um qbit é homeomorfo a $\mathbb{CP}^1$, a chamada *linha projetiva complexa*. É bem entendido, e nossa discussão sobre a fibração de Hopf deve ter deixado claro, que $\mathbb{CP}^1$ pode ser visto como a esfera de Riemann (ou de Bloch, dependendo do contexto). Assim, sua dimensão complexa é 1, por isso linha[4], enquanto sua dimensão real é 2, condizente com esfera.

O conjunto dos estados fisicamente distintos pode ser visto como

$$S^{2d-1}/S^1 \cong \mathbb{C}^d/\mathbb{C}^* \cong \mathbb{CP}^{d-1}.$$

No próximo capítulo teremos consquências interessantes desta geometria.

7.1.4 Evolução Temporal

Na secção 6.1.6 já apresentamos a equação de Schrödinger

$$\frac{d}{dt}|\psi\rangle = \frac{H}{i\hbar}|\psi\rangle,$$

[4]Interessante notar que linhas projetivas são compactas.

resposável pela evolução temporal de um sistema quântico isolado.

No capítulo 4 vimos o conceito de ação de grupo. Na evolução temporal ditada pela equação de Schrödinger, temos um exemplo onde o grupo $\mathbb{R}$ age sobre E, como no exemplo 4.10.

Vamos aproveitar para ver essa mesma discussão com outros olhos. Já concluímos que temos uma ação de grupo:

$$\begin{array}{ccc} \mathbb{R} \times E & \longrightarrow & E \\ (t, |\psi_0\rangle) & \longmapsto & |\psi(t)\rangle \end{array}$$

e que sua restrição para cada tempo t será dada por

$$U(t) = \exp\left(\frac{-iHt}{\hbar}\right),$$

que é um operador unitário, chamado *operador de evolução temporal* (por um tempo t). Devemos notar que

$$U(t_1)U(t_2) = U(t_1 + t_2),$$

para todo $t_1, t_2 \in \mathbb{R}$. Assim, esta família de operadores unitários forma um *subgrupo a um parâmetro* do grupo $U(d)$ correspondente (d a dimensão complexa de E). Ainda com outros olhos, este subgrupo a um parâmetro pode ser visto como uma curva (diferenciável) em $U(d)$, assim como

$$t \longmapsto U(t)|\phi\rangle$$

pode ser vista como uma curva em E (ou mesmo, nos vetores unitários de E) para cada $|\phi\rangle$ (unitário), ou ainda, se passarmos ao quociente, uma curva em $\mathbb{CP}^{d-1}$.

7.2 Um exemplo: o Laplaciano discreto

Um sistema de d níveis pode, fisicamente, ser interpretado de muitas formas. Uma delas é imaginar uma partícula quântica (por exemplo, um elétron) que se move em um material composto de exatamente d átomos e no qual admitimos que essa partícula só pode estar próxima destes átomos e não em um lugar qualquer. Sendo assim estamos idealizando a situação e admitindo que a posição da partícula é exatamente um sistema de d níveis, que correspondem às d posições

dos átomos do material. Esse modelo, com toda a ingenuidade que aparenta, é um ponto de partida razoável para entender, por exemplo, as propriedades de transporte de eletricidade e de calor em um cristal [AM].

Considere os operadores lineares em $\mathbb{C}^d$ definidos por

$$N_+ = \begin{bmatrix} 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & \\ 0 & 0 & 0 & 1 & \dots \\ \vdots & & & \ddots & \\ 1 & 0 & 0 & \dots & 0 \end{bmatrix} \quad \text{e} \quad N_- = \begin{bmatrix} 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & 0 & \\ 0 & 1 & 0 & 0 & \dots \\ \vdots & & & \ddots & \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}.$$

Neste caso, é fácil verificar que $N_+^* = N_-$ e $N_-^* = N_+$, logo estes operadores não correspondem a observáveis. Mas não é dificil obter seus autovetores e autovalores. Note que

$$N_+ \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_0 \end{bmatrix} \quad \text{e} \quad N_- \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{bmatrix} = \begin{bmatrix} a_{n-1} \\ a_0 \\ \vdots \\ a_{n-2} \end{bmatrix}.$$

Podemos então definir $b_{k,l} = e^{i\frac{2\pi}{d}lk}$ para $l = 1, 2, \dots, d$ e $k = 0, 1, \dots, d-1$. Definimos assim os vetores (já normalizados)

$$|B_l\rangle = \frac{1}{\sqrt{d}} \begin{bmatrix} b_{0,l} \\ b_{1,l} \\ \vdots \\ b_{n-1,l} \end{bmatrix}.$$

Não é difícil verificar que $N_+|B_l\rangle = e^{i\frac{2\pi}{d}l}|B_l\rangle$ e $N_-|B_l\rangle = e^{-i\frac{2\pi}{d}l}|B_l\rangle$.

Vamos agora definir o operador $\Delta = N_+ + N_- - 2I$; este de fato é hermitiano e, portanto, um observável. Seus autovetores são os mesmos $|B_l\rangle$ já definidos, e os autovalores são obtidos como segue:

$$\begin{aligned} \Delta|B_l\rangle &= N_+|B_l\rangle + N_-|B_l\rangle - 2I|B_l\rangle \\ &= \left(e^{i\frac{2\pi}{d}l} + e^{-i\frac{2\pi}{d}l} - 2\right)|B_l\rangle = 2(\cos\frac{2\pi}{d}l - 1)|B_l\rangle \end{aligned}$$

Exercício 7.7. *Para a e b reais, obtenha os autovalores e os autovetores do operador $a(N_+ + N_-) + bI$.*

Os vetores $|B_l\rangle$ são uma base de $\mathbb{C}^d$ e portanto um estado inicial $|\psi\rangle$ qualquer pode ser expresso como combinação linear $|\psi\rangle = \sum_{l=1}^{d} c_l(0)|B_l\rangle$. Para obter a evolução temporal deste estado inicial, se considerarmos Δ como hamiltoniano do sistema, devemos resolver a equação de Schrödinger

$$\frac{d}{dt}|\psi(t)\rangle = -i\Delta|\psi(t)\rangle.$$

Supondo que cada c_l é uma função do tempo, obtemos uma família de equações

$$\frac{d}{dt}c_l(t) = -i\lambda_l c_l(t)$$

cuja solução é

$$c_l(t) = e^{-i\lambda_l t}c_l(0).$$

7.2.1 Operador Posição

Podemos definir um outro operador como sendo

$$X = \begin{bmatrix} 0 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & \\ 0 & 0 & 2 & 0 & \dots \\ \vdots & & & \ddots & \\ 0 & 0 & 0 & \dots & n-1 \end{bmatrix}.$$

Nesse caso, é claro que os autovalores são $0, 1, \dots, d-1$ e correspondem aos autovetores

$$|v_k\rangle := e_{k+1} \qquad \text{para} \quad k = 0, 1, \dots, d-1,$$

onde os e_i são os vetores da base canônica de $\mathbb{C}^d$.

Agora considere os autovetores $|B_l\rangle$ da seção anterior. Se temos um estado $|\psi\rangle = |B_l\rangle$ (para algum l fixo) então podemos perguntar

qual é a probabilidade de se obter o valor k (k entre 0 e $d-1$) numa medição do observável X. Mas

$$|\psi\rangle = |B_l\rangle = \sum_{j=0}^{n-1} b_{j,l}|v_j\rangle.$$

Portanto, a probabilidade de se obter a posição k, que é o módulo ao quadrado do coeficiente de $|v_k\rangle$, é dada por $1/d$, de maneira independente de k. Logo todas as posições são equiprováveis num estado descrito por $|\psi\rangle = |B_l\rangle$. Porém, agora note que esta probabilidade também não depende do l escolhido! Assim, qualquer que seja o auto-estado de Δ temos que a posição tem uma distribuição equiprovável.

O operador Δ pode ser interpretado como sendo associado à energia de uma partícula num cristal, sendo que X está relacionado a sua posição nessa rede cristalina. Os autovalores de Δ são os possíveis valores da energia e os de X, os possíveis valores da posição. O que constatamos acima é que quando uma partícula está num estado que é auto-estado de Δ, e portanto tem uma energia bem definida, então temos enorme desconhecimento sobre sua posição, pois há igual probabilidade de encontrá-la em todas as posições possíveis.

Exercício 7.8. *Considerando-se um estado $|\psi\rangle$ que é auto-estado de X, qual é a probabilidade de que tenha um determinado valor de energia (isto é, um determinado autovalor do operador Δ)?*

O que encontramos aqui é mais um exemplo das chamadas *Bases Mutuamente Neutras*[5], que tanto aparecem naturalmente, como na discussão aqui apresentada, como podem ser utilizadas, por exemplo, para aplicações em criptografia. De fato, é um problema interessante, e apenas parcialmente resolvido, encontrar, para dimensão d, o número máximo de bases mutuamente neutras para aquele espaço.

7.3 A Relação de Incerteza

Considere um estado $|\Psi\rangle$ (normalizado); vamos assumir que temos dois observáveis A e B, ambos com média zero para este estado (isto

[5] Do inglês, *Mutually Unbiased Basis*.

não é tão restritivo quanto parece: sempre se pode redefinir um observável como sendo

$$\tilde{A} = A - \langle \Psi | A\Psi \rangle$$

que tem média zero no estado dado). Para uma variável aleatória qualquer, X, definimos sua variância como

$$\mathrm{Var}(X) = \left\langle X^2 \right\rangle - \langle X \rangle^2.$$

Para observáveis quânticos, os valores esperados serão calculados segundo sua prescrição. Ao consideramos apenas observáveis com média nula, $\mathrm{Var}(A) = \left\langle A^2 \right\rangle$. Veremos que

$$[A, B] = AB - BA \neq 0$$

tem consequências bastante interessantes.

Teorema 7.1 (Relação de Incerteza de Heisenberg)**.** *Sejam A e B dois observáveis de média zero e tais que $[A, B] \neq 0$. Então*

$$\mathrm{Var}(A)\,\mathrm{Var}(B) \geq \frac{1}{4} |\langle \Psi | [A, B] \Psi \rangle|^2.$$

Demonstração. Sabemos que

$$\begin{aligned} \mathrm{Var}(A)\,\mathrm{Var}(B) &= \left\langle \Psi \middle| A^2 \Psi \right\rangle \left\langle \Psi \middle| B^2 \Psi \right\rangle \\ &= \langle A\Psi | A\Psi \rangle \langle B\Psi | B\Psi \rangle \\ &= \|A\Psi\|^2 \, \|B\Psi\|^2 \\ &\geq |\langle A\Psi | B\Psi \rangle|^2 \end{aligned}$$

onde, na última passagem, foi utilizada a desigualdade de Cauchy-Bunyakovsky-Schwarz. Note que

$$\begin{aligned} \langle A\Psi | B\Psi \rangle &= \langle \Psi | AB\Psi \rangle \\ &= \langle \Psi | ([A, B] + BA)\Psi \rangle \\ &= \langle \Psi | [A, B]\Psi \rangle + \overline{\langle \Psi | BA\Psi \rangle} \\ &= \langle \Psi | [A, B]\Psi \rangle + \overline{\langle BA\Psi | \Psi \rangle} \\ &= \langle \Psi [A, B] | \Psi \rangle + \overline{\langle A\Psi | B\Psi \rangle}. \end{aligned}$$

Portanto temos

$$\langle \Psi | [A, B]\Psi \rangle = \langle A\Psi | B\Psi \rangle - \overline{\langle A\Psi | B\Psi \rangle} = 2i \,\mathrm{Im}(\langle A\Psi | B\Psi \rangle).$$

Logo,

$$\mathrm{Im}(\langle A\Psi|B\Psi\rangle) = \frac{1}{2i}\langle\Psi|[A,B]\Psi\rangle.$$

Assim,

$$\mathrm{Var}(A)\,\mathrm{Var}(B) \geq |\langle A\Psi|B\Psi\rangle|^2 \geq |\mathrm{Im}(\langle A\Psi|B\Psi\rangle)|^2 = \frac{1}{4}|\langle\Psi|[A,B]\Psi\rangle|^2.$$

□

A consequência deste resultado matemático é profunda: significa que ao se medir duas quantidades distintas, associadas a observáveis que não comutam, então o produto de suas respectivas dispersões não pode ser feito menor do que uma certa quantidade (assumindo que o valor esperado do comutador de A e B naquele estado é não nulo); se o estado tem a dispersão de A pequena, por exemplo, então a de B deve ser suficientemente grande, o que torna grande a incerteza sobre o valor desse observável no estado em questão. Isso justifica o nome pelo qual esse resultado é conhecido.

Por outro lado, estamos falando de dispersão e isso implica em uma quantidade que só pode ser obtida com muitas medições efetuadas em diversos sistemas identicamente preparados. Em princípio não está proibido conhecer com precisão arbitrária os valores dos observáveis A ou B num determinado estado $|\psi\rangle$. Muito menos as relações de incerteza exigem que as medições de A e B sejam realizadas no mesmo sistema.

Exercício 7.9. *Escolha um par de observáveis, A e B, satisfazendo as condições do teorema 7.1 e um estado $|\psi\rangle$ tais que a variância de um deles seja nula. Verifique explicitamente que nesse caso $\langle\Psi|[A,B]\Psi\rangle = 0$.*

7.4 Mais um pouco de Física

Se o capítulo anterior tratou das partículas de spin $\frac{1}{2}$, este trata das partículas com qualquer spin, assim como problemas envolvendo momentum angular orbital (aquele que classicamente é dado por $\vec{r}\times\vec{p}$).

Se um qbit permitia entender uma transição entre dois níveis atômicos, agora podemos trabalhar com processos onde vários níveis desempenham papel relevante. Se o qbit bem representava o experimento de fenda dupla, agora podemos trabalhar com fendas múltiplas.

Todos os exemplos citados acima são importantes e interessantes, mas o que acontece se tratarmos de um experimento de fendas múltiplas com partículas de spin $\frac{1}{2}$, por exemplo? Esse é um primeiro exemplo onde queremos tratar um sistema quântico composto, o assunto do capítulo 8.

Capítulo 8

Sistemas Quânticos Compostos

Agora poderemos discutir um dos aspectos mais interessantes da mecânica quântica. Assim como um par de variáveis aleatórias podem ser considerado uma nova variável aleatória em um espaço produto, um par de sistemas quânticos também pode ser visto como um novo sistema quântico, em um espaço de estados produto. Mas as coisas são um pouquinho diferentes...

8.1 Dois Qbits

8.1.1 Estados e Medições

Dois bits clássicos podem assumir quatro valores: 00, 01, 10 e 11. Deve ser claro da própria maneira de escrever que os dois bits trabalhados são distintos: existem o primeiro bit e o segundo bit, ou ainda o bit A e o bit B. Portanto, dois bits clássicos correspondem a uma variável aleatória com quatro possíveis valores. Dois bits quânticos corresponderão a um sistema quântico de 4 níveis, com uma base para seu espaço de estados dada por $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. Conforme apresentado na 2.14, podemos ver esta base como a base produto $\mathcal{Z} \otimes \mathcal{Z}$, o que permite reconhecermos um isomorfismo $\mathbb{C}^4 \cong \mathbb{C}^2 \otimes \mathbb{C}^2$.

Explicitamente, isso significa que, dentro da descrição que estamos trabalhando até o momento, qualquer estado de dois qbits se escreve

$$|\psi\rangle = \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle.$$

Se submetermos o sistema a um teste que distingue entre essas quatro alternativas clássicas, a probabilidade de obter o par ij é $|\alpha_{ij}|^2$. Note que este teste pode ser entendido como medições na base $\mathcal{Z}$ em cada qbit. Alguns outros testes relacionados vêm a seguir:

Exercício 8.1. *Considere ainda* $|\psi\rangle = \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle$.

1. *Quais as probabilidades dos possíveis resultados de um teste que apenas distingue* 0 *de* 1 *no primeiro qbit? E no segundo? Em cada caso, qual será o estado após a medição?*

2. *E para um teste que verifica se os dois resultados são iguais ou diferentes?*

Há um fato bastante sutil no exercício 8.1. Os testes envolvidos podem todos ser refinados pela decomposição $E = \bigoplus_{ij} E_{ij}$, onde $E_{ij} = \mathrm{Im}(|ij\rangle\langle ij|)$ (onde Im() denota a imagem da transformação linear em questão), $i, j \in \{0, 1\}$, sendo portanto compatíveis. Esta última corresponde a um teste completo, onde o número de possíveis respostas coincide com a dimensão do espaço de estados, sendo o único refinamento comum aos dois testes do item 1. Já o teste do item 2, com apenas duas respostas possíveis, corresponde à decomposição $E = E_{=} \oplus E_{\neq}$, onde cada subespaço envolvido é bidimensional. É fácil verificar que $E_{=} = E_{00} \oplus E_{11}$ e $E_{\neq} = E_{01} \oplus E_{10}$. Assim, cada resultado das duas medições compatíveis do item 1 só é consistente com um resultado do item 2. Porém, um resultado do item 2 não determina o resultado do teste mais fino. Esta distinção será essencial no argumento apresentado na 14.3.3. Vamos explorá-la em mais detalhes no próximo exercício.

Exercício 8.2. *Os quatro vetores abaixo são chamados* estados de

Bell[1]*:*

$$|\Phi_\pm\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle), \tag{8.1a}$$

$$|\Psi_\pm\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle). \tag{8.1b}$$

1. *Calcule a probabilidade de cada possível resultado em um teste* $E = E_= \oplus E_\neq$ *aplicado a cada estado de Bell, bem como os respectivos estados após a medição.*

2. *Agora para o estado inicial* $|01\rangle$*, quais os possíveis resultados e qual o estado após a obtenção de cada um, para o mesmo teste? E se, após a realização deste, fizermos um teste* 0 *ou* 1 *no primeiro bit, qual a probabilidade de obter cada resposta? Os dois testes envolvidos neste item são compatíveis?*

3. *Sendo* $|+-\rangle$ *o correspondente elemento da base produto* $\mathcal{X} \otimes \mathcal{X}$ *(ver capítulo 6), responda as mesmas perguntas do item anterior.*

A partir da ideia que os dois qbits em questão *podem*[2] estar espacialmente afastados, testes como do item 1 do exercício 8.1 são chamados *locais*, em um caso agindo apenas na *parte A*, no outro na *parte B*. A discussão anterior pode ser resumida dizendo que existe uma maneira local de obter a resposta do teste associado à decomposição $E = E_= \oplus E_\neq$, mas esta não é a forma mais geral de implementar tal teste. De fato, existe um refinamento local para tal teste, mas o teste propriamente dito não é local.

Adotando agora esta interpretação que cada parte do sistema composto pode estar em um laboratório diferente, vemos que os estados quânticos se dividem naturalmente em duas classes:

- Aqueles estados que podem ser preparados apenas com a utilização de operações *locais* e comunicação entre os laboratórios (utilizaremos a sigla em inglês: *LOCC* para *Local Operations and Classical Communication*);

[1] A notação utilizada também é razoavelmente padrão.

[2] É uma possibilidade, não uma exigência. Ainda assim, o linguajar se mantém.

- Aqueles que não podem ser preparados de tal forma, ou seja, exigem operações conjuntas que não podem ser decompostas em operações locais e comunicação clássica.

Esta discussão será aprofundada no capítulo 10, quando já teremos em mãos uma noção mais geral de estado, a ser apresentada no capítulo 9. Com o cenário que temos no momento, os estados que podem ser preparados por LOCC são da forma $|\alpha\rangle \otimes |\beta\rangle$, ou seja, representado por vetores decomponíveis[3] de $\mathbb{C}^2 \otimes \mathbb{C}^2$. Já vetores não decomponíveis corresponderão ao segundo caso, sendo chamados *estados emaranhados*. Sendo mais explícito:

Definição 8.1. *Um estado representado por um vetor de $\mathbb{C}^2 \otimes \mathbb{C}^2$ será dito:*

1. *Fatorável, quando representado por um vetor decomponível;*
2. *Emaranhado, caso contrário.*

Exercício 8.3.
1. *Mostre que para um estado $|\alpha\rangle \otimes |\beta\rangle$ as probabilidades de um teste local em A e de outro teste local em B são independentes. Obtenha ainda, para cada resultado destes testes, uma forma para o estado do sistema após tal intervenção.*
2. *Mostre que isso não ocorre, necessariamente, se o estado inicial do sistema for emaranhado.*
3. *Em especial, considere um teste local completo na parte A (ou seja, uma decomposição ortonormal não trivial do $\mathbb{C}^2$ correspondente à parte A) e mostre que o estado após a medição é sempre decomponível, mas que o vetor correspondente à parte B depende do resultado do teste realizado em A.*

A discussão do exercício 8.3 deve se tornar mais natural se você utilizar a decomposição de Schmidt (2.5).

Exercício 8.4. *Considere o estado $|\Psi_-\rangle$, da eq.* (8.1b).

1. *Obtenha as probabilidades dos possíveis resultados do teste associado à base $\mathcal{Z} \otimes \mathcal{Z}$.*

[3] No contexto de mecânica quântica é comum chamá-los de *vetores produto*.

2. Faça o mesmo para os testes associados a $\mathcal{X} \otimes \mathcal{X}$ e a $\mathcal{Y} \otimes \mathcal{Y}$.

Cada um dos resultados que você obteve acima mostra que os bits gerados pelas respostas de cada teste aplicado aos qbits estão correlacionados. Cada um destes resultados sozinho não é surpreendente. Exemplos assim acontecem em nosso "mundo clássico" frequentemente. Considere que uma moeda foi cortada ao meio, de modo que uma semi-moeda só tem cara e a outra coroa. Você põe cada uma em um envelope e manda cada envelope para um amigo, mas sem saber qual semi-moeda foi colocada em cada um. Os bits gerados por este teste clássico têm o mesmo tipo de correlação que os bits obtidos por cada um dos testes do exercício 8.2. Porém, os dois qbits preparados em $|\Psi_-\rangle$ possuem algo que as semi-moedas não possuem: a possibilidade de realização de testes diferentes (medir com respeito a outras bases). Para realçar ainda mais esta situação, lembremos que um dado padronizado possui seis faces numeradas de 1 a 6 e que faces opostas sempre somam 7. Inspirados no exemplo da moeda, podemos considerar a possibilidade de cortar um dado destes paralelamente a um par de faces, colocar cada metade em um envelope aleatório e mandar para dois amigos. Conhecendo a regra da brincadeira, após abrir seu envelope, cada amigo sabe o que o outro recebeu. Mas note que se o corte foi realizado paralelamente às faces 2 e 5, nenhum amigo pode receber a face 4 completa. O que os qbits nos permitem, de certo modo, é enviar os semi-dados para cada amigo antes de fazer o corte! De posse dos seus envelopes, eles podem decidir sobre qual corte fazer. E, se fizerem os mesmos cortes, obterão bits complementares, da mesma forma que no exemplo da moeda.

Exercício 8.5. *Ainda com o estado $|\Psi_-\rangle$, quais as probabilidades se for feita uma medição na base $\mathcal{X} \otimes \mathcal{Z}$?*

Exercício 8.6. *Adapte a situação da moeda cortada para obter outro sistema clássico que pode replicar as correlações aqui representadas pelo corte do "dado quântico".*

O que o exercício 8.6 nos diz é que, ainda que a historinha do dado pareça convincente, suas correlações podem ser obtidas com sistemas clássicos (portanto, não deve figurar entre as *surpresas quânticas*). No capítulo 14 voltaremos a esse tema, apresentando lá sim, resultados

quânticos surpreendentes, no sentido que nenhum sistema clássico será capaz de imitá-los.

8.1.2 Estados Fisicamente Distintos

Na 7.1.3 apontamos que o conjunto dos estados fisicamente distintos de um qdit é identificado com $\mathbb{CP}^{d-1}$ e que há consequências interessantes da geometria do espaço de estados para sistemas compostos. Vamos começar a explorá-la neste caso de dois qbits.

Uma boa maneira de trabalhar em $\mathbb{CP}^m$ é usar as chamadas *coordenadas homogêneas*. Assim, uma classe é definida por coordenadas $[x_0 : x_1 : \ldots : x_m]$, entendido que $[\lambda x_0 : \lambda x_1 : \ldots : \lambda x_m]$ representa a mesma classe, para todo $\lambda \neq 0$. As componentes de um vetor de estado podem então ser vistas como coordenadas homogêneas que definem um ponto em $\mathbb{CP}^m$, mesmo que isso não seja normalmente dito em livros de mecânica quântica.

Entendido que os estados fisicamente distintos de dois qbits formam um $\mathbb{CP}^3$, enquanto os estados de um qbit formam um $\mathbb{CP}^1$ cada, uma pergunta natural é: onde se encontram os estados fatoráveis neste $\mathbb{CP}^3$? Esta pergunta pode ser respondida de maneira construtiva. Em termos de kets, considere os estados $|\alpha\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle$ para o qbit A e $|\beta\rangle = \beta_0|0\rangle + \beta_1|1\rangle$ para B. Temos então o estado produto $|\alpha\rangle \otimes |\beta\rangle = \alpha_0\beta_0|00\rangle + \alpha_0\beta_1|01\rangle + \alpha_1\beta_0|10\rangle + \alpha_1\beta_1|11\rangle$. Todo vetor de estado produto (de dois qbits) é desta forma. Em termos de coordenadas homogênas, aproveitando a mesma notação, temos

$$\begin{array}{rcl} \mathbb{CP}^1 \times \mathbb{CP}^1 & \to & \mathbb{CP}^3 \\ ([\alpha_0 : \alpha_1], [\beta_0 : \beta_1]) & \mapsto & [\alpha_0\beta_0 : \alpha_0\beta_1 : \alpha_1\beta_0 : \alpha_1\beta_1] \end{array} \qquad (8.2)$$

que é conhecido como *mergulho de Segre*. Do ponto de vista da geometria algébrica, o mergulho de Segre é uma maneira de tornar um produto cartesiano de espaços projetivos em uma subvariedade de um espaço projetivo maior, usando para isso uma aplicação algébrica (expressada por polinômios homogêneos).

Exercício 8.7. *Calcule a dimensão sobre os reais do conjunto dos estados fatoráveis de dois qbits e do conjunto dos estados emaranhados de dois qbits. Com isso, responda a pergunta: se você sortear*

aleatoriamente (com distribuição uniforme) um estado em $\mathbb{CP}^3$, qual a probabilidade de ele representar um estado emaranhado?

Exercício 8.8. *Se você conhece o* Teorema de Bézout*[Har], deve conseguir demonstrar o seguinte resultado*[4]*: em todo subespaço bidimensional de $\mathbb{C}^2 \otimes \mathbb{C}^2$ há vetor decomponível.*
Sugestão: Passe o problema para $\mathbb{CP}^3$, descreva o subespaço bidimensional e a subvariedade dos estados produto (a imagem do mergulho de Segre) e obtenha a intersecção destas.

8.1.3 Dois spins $\frac{1}{2}$

Já apresentamos os sistemas de spin $\frac{1}{2}$ como importante exemplo de qbit, utilizado inclusive para melhor entender as possíveis evoluções temporais destes. Agora vamos explorar mais uma propriedade, presente quando mais de um spin é considerado. Por enquanto, mais de um significa dois.

As matrizes de Pauli (6.4) estão intimamente relacionadas aos observáveis de spin. De fato, cada uma dessas três matrizes representa o observável associado à medição da respectiva componente do spin. Em particular,

$$S_x = \frac{\hbar}{2}\sigma_x, \quad S_y = \frac{\hbar}{2}\sigma_y, \quad S_z = \frac{\hbar}{2}\sigma_z, \tag{8.3a}$$

onde S_u é a componente u do spin. Cabe notar que tais observáveis não são compatíveis, não havendo um estado com as três componentes de spin definidas.

De maneira mais geral, se $\vec{u} = (u^x, u^y, u^z)$ é um vetor unitário, usamos a notação

$$\vec{u} \cdot \vec{\sigma} = u^x\sigma_x + u^y\sigma_y + u^z\sigma_z$$

para representar o operador acima definido. Com ela, temos

$$S_{\vec{u}} = \frac{\hbar}{2}\vec{u} \cdot \vec{\sigma} \tag{8.3b}$$

que representa a componente do spin na direção do vetor $\vec{u}$.

[4] Nada intuitivo, sem essa caracterização geométrica.

Exercício 8.9. *Quais os autovalores e autovetores dos observáveis de spin* (8.3a)*? E para $S_{\vec{u}}$, da eq.* (8.3b)*? Mostre que se $\vec{u}$ é vetor da base canônica, não há inconsistência na notação.*

Quando consideramos dois spins $\frac{1}{2}$, faz sentido pensarmos em observáveis relacionados a uma componente do spin de uma das partículas. Estes serão dados por

$$S_{\vec{u}} \otimes I \text{ ou } I \otimes S_{\vec{v}},$$

respectivamente para a componente $\vec{u}$ do primeiro spin ou para componente $\vec{v}$ do segundo.

Como $\vec{u}$ e $\vec{v}$ são vetores de $\mathbb{R}^3$, sem qualquer vinculação com a dimensão do espaço de estados, é natural definir[5]

$$S_{\vec{u}} = S_{\vec{u}} \otimes I + I \otimes S_{\vec{u}}, \tag{8.4}$$

e estudarmos seus autovalores e autovetores. O mais simples é começar por $\vec{u} = (0, 0, 1)$.

Exercício 8.10. *Com respeito à base $\mathcal{Z} \otimes \mathcal{Z}$, obtenha as matrizes que representam os operadores $S_z \otimes I$, $I \otimes S_z$ e S_z.*

Com o exercício 8.10 você deve ter obtido três autovalores para a componente z do spin do sistema composto: $\hbar$, 0 e $-\hbar$ e deve ter notado que o autovalor 0 é degenerado.

Exercício 8.11. *Com respeito à mesma base, obtenha matrizes que representam $S_x = S_x \otimes I + I \otimes S_x$ e $S_y = S_y \otimes I + I \otimes S_y$.*

Agora você pode verificar que o estado de Bell $|\Psi_-\rangle$ é autovetor comum a todo $S_{\vec{u}}$ do sistema composto (note que, também para o sistema composto, $S_{\vec{u}} = \vec{u} \cdot \vec{\sigma}$).

Exercício 8.12. *Explique, tanto com a linguagem de decomposições ortogonais, quanto com a liguagem de operadores, por que não há contradição entre os fatos dos observáveis S_x, S_y e S_z não comutarem e terem <u>um</u> autovetor comum.*

[5]Esperamos que o contexto deixe sempre claro onde age cada operador. A mesma notação $S_{\vec{u}}$ está sendo usada para o observável associado à componente $\vec{u}$ do spin de cada partícula e do sistema composto. Caso o leitor prefira uma notação mais clara, porém carregada, sugerimos: $S^A_{\vec{u}} = S_{\vec{u}} \otimes I$, $S^B_{\vec{u}} = I \otimes S_{\vec{u}}$ e $S^{AB}_{\vec{u}} = S^A_{\vec{u}} + S^B_{\vec{u}}$, que também utilizaremos eventualmente.

O que acabamos de obter é uma decomposição bastante interessante, razoavelmente óbvia em termos de dimensões, mas com consequências profundas na mecânica quântica:

$$\mathbb{C}^2 \otimes \mathbb{C}^2 \cong \mathbb{C} \oplus \mathbb{C}^3, \tag{8.5}$$

onde $\mathbb{C}$ se refere ao espaço vetorial gerado por $|\Psi_-\rangle$ e $\mathbb{C}^3$ seu complemento ortogonal. Em especial, você deve reexaminar os resultados do exercício 8.4 para buscar compreendê-los mais profundamente.

Exercício 8.13. *Considere a ação de grupo $U(2) \times \mathbb{C}^4 \to \mathbb{C}^4$, definida em vetores decomponíveis por $(U, |\alpha\rangle \otimes |\beta\rangle) \mapsto U|\alpha\rangle \otimes U|\beta\rangle$ e estendida por linearidade.*

1. *Descreva geometricamente as órbitas de cada estado de Bell obtidas por esta ação. Em particular, quantas são e qual a dimensão de cada uma?*
2. *Mostre que esta ação "passa ao quociente", isto é, induz uma ação de $U(2)$ sobre $\mathbb{CP}^3$. Como são as órbitas dos estados de Bell nessa descrição?*
3. *Como isso se relaciona com a decomposição* (8.5)*?*

8.1.4 Evolução Temporal

A evolução temporal continua sendo ditada pela Equação de Schrödinger (6.5). O ponto de discussão agora serão os diferentes hamiltonianos que determinam tal evolução.

Se $H^{AB} = H^A + H^B$, onde H^A denota algum operador auto-adjunto da forma $H_A \otimes I$, enquanto H^B é da forma $I \otimes H_B$, teremos $\left[H^A, H^B\right] = 0$ e, consequentemente[6],

$$U^{AB}(t) = \exp\left(-iH^{AB}t\right) = \exp\left(-iH^At\right)\exp\left(-iH^Bt\right) = U_A(t) \otimes U_B(t).$$

Isso significa que cada base produto será levada por $U^{AB}(t)$ em uma outra base produto. Portanto, uma evolução temporal assim propaga as correlações, sem criá-las nem destruí-las.

[6] A partir daqui adotamos o hábito de escolher unidades de forma que $\hbar = 1$. Pode ser um interessante exercício o leitor identificar onde estão estes $\hbar$ escondidos.

Exercício 8.14. *Mostre que os coeficientes de Schmidt de* $|\psi\rangle$ *e de* $U^{AB}(t)|\psi\rangle$*, para* U^{AB} *como acima, são os mesmos.*

Podemos entender este resultado sob a óptica das ações de grupo. Já vimos que a evolução temporal pode ser entendida como a ação do subgrupo a um parâmetro $\left\{U^{AB}(t); t \in \mathbb{R}\right\}$ do grupo $U(4)$ de todas as unitárias 4×4. Este subgrupo age em $\mathbb{C}^4 \cong \mathbb{C}^2 \otimes \mathbb{C}^2$ e tal ação "passa ao projetivo" (exercício 8.13). O ponto central é que, neste caso de $U^{AB} = U_A \otimes U_B$, a ação dinâmica "respeita" o mergulho de Segre, ou seja, a órbita de cada ponto da imagem do mergulho (correspondente aos estados produto) está inteiramente contida nesta mesma subvariedade[7].

De forma mais resumida, um hamiltoniano da forma

$$H^{AB} = H_A \otimes I + I \otimes H_B \tag{8.6}$$

gera dinâmicas independentes nas partes A e B que estão sendo consideradas conjuntamente, como um sistema composto.

A coisa muda de figura quando a forma (8.6) não pode ser alcançada, ou seja, quando não temos geradores independentes para as dinâmicas de cada parte.

Exercício 8.15. *Tome como exemplo o hamiltoniano* $H^{AB} = \omega\sigma_z \otimes \sigma_z$*.*

1. *Mostre que existe estado produto que se mantém produto pela evolução temporal;*

2. *Mostre que também existe estado produto que se torna emaranhado pela evolução temporal;*

3. *Podemos escrever este* H^{AB} *na forma* (8.6)*?*

Neste caso, é comum escrever-se o hamiltoniano do sistema composto na forma (não única)

$$H^{AB} = H_A \otimes I + I \otimes H_B + H_{\text{int}}, \tag{8.7}$$

[7]Pode-se dizer mais: os elementos de $U(4)$ que respeitam o mergulho de Segre ou são da forma $U_A \otimes U_B$, ou seu produto com $U_{\text{Swap}} : |\alpha\rangle \otimes |\beta\rangle \mapsto |\beta\rangle \otimes |\alpha\rangle$, fato este demonstrado na ref. [Dru].

onde H_{int} é chamado *hamiltoniano de interação*, sendo o responsável por criar (ou destruir) correlações entre as partes. Dizemos assim que os dois qubits *interagem*.

Em geral, os autovetores de um sistema interagente são emaranhados (veja exercício 8.7) e os autovalores são incomensuráveis. Com isso, genericamente um estado inicial produto é levado a um estado emaranhado, para todo $t > 0$. Neste sentido é correto dizer que, em geral, interação cria emaranhamento em sistemas quânticos[8].

Exercício 8.16. *Obtenha as condições para que um hamiltoniano com termo de interação permita que $U(t)$ leve algum estado produto em estado produto, para algum $t > 0$. Justifique por que, genericamente, isso não acontece.*

Exercício 8.17. *Escreva um hamiltoniano para dois qbits tal que o autovetor associado ao menor autovalor seja produto, mas os demais autovetores não.*

8.2 Sistemas de Duas Partes

Sistematizando e generalizando a discussão anterior, podemos enunciar o seguinte:

Postulado 8.1. *Se tratamos conjuntamente dois sistemas, aos quais estão associados, respectivamente, os espaços de estados E e F, o espaço de estados do sistema composto é $E \otimes F$.*

Conceitos como base produto, medição local, LOCC e emaranhamento passam imediatamente para o cenário bipartido $E \otimes F$. A decomposição de Schmidt nos indica que, no que diz respeito ao emaranhamento de estados puros, o espaço de menor dimensão é o mais importante.

Um resultado central para a teoria do emaranhamento em estados puros é devido a Nielsen [Nie]: para determinar se um estado puro $|\psi\rangle$ de um sistema bipartido pode ser levado por operações locais e comunicação clássica a outro estado $|\phi\rangle$, basta comparar seus vetores de Schmidt.

[8] E, pelo mesmo motivo, é interessante entender como se pode obter estados emaranhados sem lançar mão da interação direta entre as partes.

Se a decomposição de Schmidt de $|\psi\rangle$ é $|\psi\rangle = \sum_i \psi_i |\alpha_i\rangle|\beta_i\rangle$, com a convenção que os coeficientes são reais, não-negativos e escritos em ordem decrescente, chamamos $\vec{\psi} = \left(\psi_i^2\right)_i$ de vetor de Schmidt do estado $|\psi\rangle$. Note que a normalização de $|\psi\rangle$ implica que o vetor de Schmidt é um vetor de probabilidades, ordenado.

Para dois vetores de probabilidade, $\vec{p} = (p_i)_i$ e $\vec{q} = (q_i)_i$, escritos em ordem decrescente, dizemos que $\vec{p}$ majora $\vec{q}$, e denotamos $\vec{p} \succ \vec{q}$, quando

$$\sum_{i=1}^{k} p_i \geq \sum_{i=1}^{k} q_i, \quad \forall k. \tag{8.8}$$

O resultado de Nielsen [Nie] é que se $\vec{\psi} \succ \vec{\phi}$, então existe uma estratégia de LOCC capaz de converter $|\phi\rangle$ em $|\psi\rangle$. Se não é permitida a utilização de outros sistemas quântico auxiliares[9], o critério é ainda mais restritivo: se a majoração for estrita (quer dizer, para algum k a desigualdade em (8.8) é estrita), não apenas existe estratégia de LOCC para converter $|\phi\rangle$ em $|\psi\rangle$, como não existe estratégia de LOCC capaz de converter $|\psi\rangle$ em $|\phi\rangle$.

Interessante entender que a relação de majoração impõe uma *ordem parcial* nos vetores de probabilidades e que o resultado discutido acima mostra que essa ordem parcial é levada ao emaranhamento dos estados quânticos de duas parte. A melhor forma de entender por que o ordenamento é parcial (e quando ele é total) parece ser resolver o seguinte:

Exercício 8.18. *Mantendo a notação $\vec{p}$ e $\vec{q}$ para vetores de probabilidade, $\vec{\psi}$ e $\vec{\phi}$ para vetores de Schmidt dos estados $|\psi\rangle$ e $|\phi\rangle$, respectivamente:*

1. *Obtenha $\vec{p}$ e $\vec{q}$ de forma que nem $\vec{p} \succ \vec{q}$, nem $\vec{q} \succ \vec{p}$;*

2. *Mostre que se $\vec{p} = (p_1, p_2)$ e $\vec{q} = (q_1, q_2)$, necessariamente ou $\vec{p} \succ \vec{q}$, ou $\vec{q} \succ \vec{p}$; Este item pode ser enunciado como: distribuições de probabilidade de Bernoulli[10] são completamente ordenadas pela relação de majoração;*

[9] E a proópria definição de LOCC os descarta; aqui estamos apenas sendo enfáticos e o leitor curioso pode encontrar na ref. [JP] o motivo.

[10] Aquelas onde o espaço amostral tem apenas dois elementos.

3. *Use $\vec{\psi} = \vec{p}$ e $\vec{\phi} = \vec{q}$ do item 1 para exibir estados quânticos de sistemas de duas partes que não podem ser conectados por LOCC em nenhum sentido;*

4. *Mostre que os estados puros de dois qubits são completamente ordenados com respeito ao emaranhamento.*

Com relação à geometria dos estados fisicamente distintos de sistemas bipartidos, sugerimos o exercício a seguir.

Exercício 8.19. *Considere agora dois espaços projetivos complexos, $\mathbb{CP}^m$ e $\mathbb{CP}^n$. Construa o mergulho de Segre destes dois espaços, ou seja, construa uma aplicação semelhante à* (8.2) *no espaço projetivo com a dimensão adequada e faça a relação deste com os estados produto de um sistema quântico de duas partes.*

Exercício 8.20. *Releia a subsecção sobre evolução temporal de dois qbits, 8.1.4, fazendo sua generalização para sistemas bipartidos quaisquer.*

É claro que, ao especificar como considerar dois sistemas conjuntamente, estamos também dando a receita para considerar qualquer quantidade de sistemas como partes de um sistema maior.

8.3 Mais Qbits

Seguindo com a estratégia de fixar conceitos com os exemplos mais simples, podemos passar ao caso onde juntamos mais qbits.

Se tivéssemos três bits clássicos, teríamos $2^3 = 8$ configurações possíveis:

$$000, 001, 010, 011, 100, 101, 110, 111.$$

Aqueles acostumados com a notação binária[11] perceberam que estas configurações correspondem a "contar" de 0 a 7, sempre usando três algarismos binários.

Quanticamente, estas configurações se tornam uma base ortonormal para o espaço de estados, que pode ser identificado com

[11] E há 10 tipo de pessoas no mundo: as que entendem binários e as outras.

$\mathbb{C}^8 \cong \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$, para o qual também usamos a seguinte notação[12]: $\left(\mathbb{C}^2\right)^{\otimes 3}$.

A generalização é imediata e o espaço de estados para n qbits será isomorfo a $\left(\mathbb{C}^2\right)^{\otimes n} \cong \mathbb{C}^{2^n}$.

8.3.1 Emaranhamento: W vs GHZ

Em vários sentidos, há vários emaranhamentos quando temos mais que dois qbits. Começando pelo caso de três qbits, onde chamamos as partes de A, B e C, podemos reconhecer três bipartições: $\{\{A,B\},\{C\}\}$, $\{\{A,C\},\{B\}\}$ e $\{\{B,C\},\{A\}\}$, além da tripartição $\{\{A\},\{B\},\{C\}\}$. É justo perguntar a cada estado se ele é emaranhado ou fatorável com respeito a cada uma dessas possíveis partições. É claro que se um estado for fatorável com respeito à "partição completa", $\{\{A\},\{B\},\{C\}\}$, também será com respeito a todas as demais partições, mas a recíproca só é verdeira se entendida com cuidado (veja exercício 8.22).

Mas também há mais de um emaranhamento de uma forma mais sutil. Para dois qbits, os estados de Bell e seus equivalentes locais[13] são maximamente emaranhados. Em particular, se tivermos uma fonte de estados de Bell, é possível gerar qualquer outro estado utilizando esta fonte e LOCC. Isso muda completamente quando mais partes são envolvidas. Há dois estados (além de seus equivalentes locais) que podem, com bastante justiça, ser chamados de *maximamente emaranhados*. Apesar de tal justiça, nenhum deles retém a propriedade que basta uma fonte deles para podermos gerar qualquer estado de três qbits aplicando LOCC [DVC]. Em particular, tendo uma fonte de um deles, não é possível obter o outro. Seus exemplos típicos são:

$$|GHZ\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle), \tag{8.9a}$$

$$|W\rangle = \frac{1}{\sqrt{3}}(|001\rangle + |010\rangle + |100\rangle), \tag{8.9b}$$

cujos nomes são homenagens a Greenberger, Horne e Zeilinger [GHZ] e a Wootters [CKW].

[12] Uma espécie de *potência tensorial*.

[13] Ou seja, aqueles que podem ser obtidos aplicando unitárias locais a eles.

Exercício 8.21. *Diferença entre $|GHZ\rangle$ e $|W\rangle$*

1. *Mostre que cada qbit de $|GHZ\rangle$ está emaranhado com os demais.*

2. *Qual o estado dos qbits A e B após cada possível resultado de um teste $\mathcal{Z}$ no qbit C? Há emaranhamento nestes estados?*

3. *Mostre que cada qbit de $|W\rangle$ está emaranhado com os demais.*

4. *Qual o estado dos qbits A e B após cada possível resultado de um teste $\mathcal{Z}$ no qbit C? Há emaranhamento nestes estados?*

Para um número maior de qbits teremos ainda mais partições possíveis e pode-se falar de emaranhamento com respeito a cada uma delas. Naturalmente, se uma partição $\mathcal{R}$ é um refinamento[14] de uma partição $\mathcal{P}$, um estado $\mathcal{R}$-fatorável será também $\mathcal{P}$-fatorável; usando a contrapositiva, um estado $\mathcal{P}$-emaranhado é também $\mathcal{R}$-emaranhado. Além disso, os estados (8.9) são imediatamente generalizados, além de ganharem companhia de outras famílias também interessantes.

8.3.2 Geometria

Aumentando o número de partes, aumenta a riqueza das construções geométricas encontradas. Começando por três qbits, deve ser claro que os estados fisicamente distintos formam um $\mathbb{CP}^7$. Para ganhar intuição, vale se concentrar no seguinte:

Exercício 8.22. 1. *Mostre que os estados $\{\{A,B\},\{C\}\}$-fatoráveis correspondem à imagem do mergulho de Segre $\mathbb{CP}^3\times\mathbb{CP}^1\to\mathbb{CP}^7$.*

2. *Mostre que todo estado simultaneamente fatorável com respeito às partições $\{\{A,B\},\{C\}\}$ e $\{\{A,C\},\{B\}\}$ é também fatorável com respeito às partições $\{\{A\},\{B,C\}\}$ e $\{\{A\},\{B\},\{C\}\}$*

3. *Interprete o item anterior em termos das posições relativas das imagens dos diferentes mergulhos de Segre envolvidos.*

[14] Cada conjunto da partição $\mathcal{P}$ é união de conjuntos da partição $\mathcal{R}$.

4. Obtenha a dimensão de cada conjunto envolvido nos itens anteriores.

Exercício 8.23. *Pense um pouco nos diversos mergulhos de Segre envolvidos no caso de quatro qubits.*

8.3.3 Vários spins $\frac{1}{2}$

Vamos agora retomar a discussão da subsecção 8.1.3. Para entender melhor o processo, vamos passar a discussão para três partículas de spin $\frac{1}{2}$. É importante destacar que estamos sempre considerando partículas *distinguíveis*[15].

Nosso problema é entender como o sistema se comporta perante a ação de operadores coletivos. Se nossos três spins são rotulados A, B e C, queremos generalizar a equação (8.4), ou seja, vamos considerar

$$S_{\vec{u}} = S_{\vec{u}}^A + S_{\vec{u}}^B + S_{\vec{u}}^C, \tag{8.10}$$

onde $S_{\vec{u}}^A = S_{\vec{u}} \otimes I \otimes I$, $S_{\vec{u}}^B = I \otimes S_{\vec{u}} \otimes I$ e $S_{\vec{u}}^C = I \otimes I \otimes S_{\vec{u}}$, e procederíamos de maneira análoga[16] para mais spins.

O que pretendemos mostrar é que a decomposição dada pela expressão (8.5) para dois spins $\frac{1}{2}$, terá a forma

$$\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2 \cong \mathbb{C}^2 \oplus \mathbb{C}^2 \oplus \mathbb{C}^4. \tag{8.11}$$

A melhor maneira de entender tal decomposição (e a formação dos chamados multipletos) é definir os operadores de levantamento e abaixamento. Para cada spin $\frac{1}{2}$ eles são dados por $\sigma_+ = \sigma_x + i\sigma_y$ e $\sigma_- = \sigma_x - i\sigma_y$. Matricialmente, temos

$$\sigma_+ = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{e} \quad \sigma_- = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Os operadores coletivos serão dados por

$$J_+ = \sigma_+^A + \sigma_+^B + \sigma_+^C \quad \text{e} \quad J_- = \sigma_-^A + \sigma_-^B + \sigma_-^C.$$

[15] Outras questões, também interessantes, aparecem quando consideramos partículas indistinguíveis em mecânica quântica, mas não vamos abordar estas questões aqui.

[16] Chamamos $S_{\vec{u}}^P$ de extensão trivial de $S_{\vec{u}}$, agindo na parte P.

Exercício 8.24. *Mostre que* $J_-^* = J_+$.

Agora vamos explorar os operadores de levantamento e abaixamento para verificar explicitamente a decomposição (8.11):

Exercício 8.25. *Não vamos nos preocupar com normalização neste exercício. Sejam* $|\psi_1\rangle = |111\rangle$, $|\psi_2\rangle = |011\rangle - |101\rangle$, $|\psi_3\rangle = |011\rangle + |101\rangle - 2|110\rangle$.

1. *Mostre que* $J_-|\psi_i\rangle = 0$, $i = 1, 2, 3$;
2. *Calcule* ${J_+}^k|\psi_i\rangle$;
3. *Verifique que todos os vetores obtidos neste exercício são ortogonais e explique a relação dos cálculos que você fez com a decomposição* (8.11).

Ou ainda, de uma maneira mais simétrica:

Exercício 8.26. *Sejam* $|\phi_1\rangle = |111\rangle$, $|\phi_+\rangle = |011\rangle + \gamma|101\rangle + \gamma^2|110\rangle$, $|\phi_-\rangle = |011\rangle + \gamma^2|101\rangle + \gamma|110\rangle$, *onde* $\gamma^3 = 1$.

1. *Mostre que* $J_-|\phi_i\rangle = 0$, $i = 1, 2, 3$;
2. *Calcule* ${J_+}^k|\phi_i\rangle$;
3. *Verifique que todos os vetores obtidos neste exercício são ortogonais e explique a relação dos cálculos que você fez com a decomposição* (8.11) *e com o exercício 8.25.*

Você pode agora tentar generalizar o que foi apresentado nos exercícios 8.25 e 8.26 e, em especial, mostrar que

$$\left(\mathbb{C}^2\right)^{\otimes 4} \cong \mathbb{C} \oplus \mathbb{C} \oplus \mathbb{C}^3 \oplus \mathbb{C}^3 \oplus \mathbb{C}^3 \oplus \mathbb{C}^5. \tag{8.12}$$

8.4 Compondo ou Decompondo?

Até agora usamos uma abordagem "de baixo para cima", ou seja, começamos com dois sistemas e resolvemos a questão de como tratá-los conjuntamente. Mas também cabe encarar a situação "de cima para baixo", começando por um único sistema e perguntando como poderemos dividi-lo em subsistemas.

Para responder essa questão fazemos uma exigência de consistência: se tratarmos as partes conjuntamente, devemos reobter o todo. E assim, como a dimensão do produto tensorial de dois espaços vetoriais é o produto de suas dimensões, as possíveis decomposições devem respeitar a decomposição em fatores primos da dimensão do espaço de estados do sistema "grande". Dessa forma, para alguns casos há, nesse sentido, uma única decomposição:

$$\begin{aligned} \mathbb{C}^4 &\cong \mathbb{C}^2 \otimes \mathbb{C}^2, \\ \mathbb{C}^6 &\cong \mathbb{C}^2 \otimes \mathbb{C}^3, \\ &\vdots \\ \mathbb{C}^{pq} &\cong \mathbb{C}^p \otimes \mathbb{C}^q, \end{aligned}$$

para primos (não necessariamente distintos) p e q. Mas para inteiros "mais compostos", já temos algo mais rico, como

$$\mathbb{C}^{12} \cong \mathbb{C}^2 \otimes \mathbb{C}^6 \cong \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^3 \cong \mathbb{C}^4 \otimes \mathbb{C}^3,$$

enquanto sistemas com espaços de estado de dimensão prima são, a esse respeito, atômicos[17].

É interessante notar que essa condição relacionada às dimensões restringe as possíveis "fatorações tensoriais", mas apenas as determina a menos de isomorfismos. É comum chamarmos duas fatorações distintas (mesmo que em espaços de mesmas dimensões) de diferentes *estruturas de produto tensorial* (do inglês *tensor product structures*, TPS), como definido na referência [ZLL]. Primeiro devemos entender melhor o que consideramos duas estruturas distintas e com isso podemos apresentar uma receita de como obter estruturas de produto tensorial em um espaço vetorial de dimensão composta.

Em mecânica quântica, podemos restringir nossa discussão a bases ortonormais. Vamos fazer a discussão nesses termos, em benefício do leitor acostumado. Se é dada uma estrutura de produto tensorial da forma

$$\mathbb{C}^{mn} \cong \mathbb{C}^m \otimes \mathbb{C}^n, \tag{8.13}$$

onde não necessariamente m e n são primos[18], podemos escolher

[17] No sentido original da palavra: indivisíveis.

[18] Ou seja, estamos preocupados em como fazer uma separação; se m ou n não for primo, o processo pode ainda ser continuado.

bases ortonormais para cada fator e teremos a base produto, também ortonormal, para $\mathbb{C}^{mn}$. Podemos inverter este processo e assim obter diferentes TPS: escolhidos uma base ortonormal para $\mathbb{C}^{mn}$ e um ordenamento para essa base, seus vetores poderão ser numerados $\{|e_{ij}\rangle\}$, com $i = 1, \ldots, m$ e $j = 1, \ldots, n$. Podemos então declarar que $|e_{ij}\rangle = |\alpha_i\rangle \otimes |\beta_j\rangle$, com $\{|\alpha_i\rangle\}$ uma base ortonormal para um fator $\mathbb{C}^m$ e $\{|\beta_j\rangle\}$ uma base ortonormal para o outro fator $\mathbb{C}^n$. Duas escolhas assim feitas gerarão estruturas de produto tensorial equivalentes se a unitária, U, de C^{mn}, que leva uma base ordenada em outra, for decomponível, *i.e.*: $U = U_A \otimes U_B$, com respeito a uma das estruturas[19].

É interessante notar que as propriedades de emaranhamento só são definidas quando uma estrutura de produto tensorial é apresentada. Assim, estados produto em uma TPS podem ser emaranhados em outra e vice-versa. De fato, a construção acima mostra que para todo vetor de estado existe uma TPS com respeito à qual ele é produto. Se, além disso, reinterpretarmos o fato de um vetor genérico ser emaranhado (exercícios 8.7 e 8.19), como com respeito a uma TPS genérica aquele vetor é emaranhado, somos levados a concluir que emaranhamento não é uma propriedade intrínseca de estados quânticos, mas dependem da TPS subentendida [TDV].

Exercício 8.27. *Defina uma TPS a partir da base de Bell, eq.* (8.1)*. Mostre que, com respeito a essa TPS, os vetores* $|ij\rangle$ *são maximamente emaranhados.*

8.5 Um pouquinho mais de Física

Vamos seguir Einstein e Feynman. Feynman afirma que o experimento de fenda dupla contém *o único mistério* da mecânica quântica [FLS], enquanto Einstein tem uma citação famosa: "você sempre deve fazer as coisas da maneira mais simples possível"[20].

Assim, vamos voltar ao interferômetro de fenda dupla, mas agora considerando experimentos com um sistema auxiliar. Este sistema auxiliar tem como objetivo registrar "por qual fenda" passa a partí-

[19] Por que é suficiente ser decomponível com respeito a uma das estruturas?
[20] A citação continua: "Nunca mais simples que isso".

cula interferométrica. Para ser o mais simples possível, consideraremos os dois "estados de fenda", $|d\rangle$ e $|e\rangle$, enquanto o "ponteiro" que registra por qual fenda a partícula passou terá seu espaço de estados gerado por $|\nearrow\rangle$ e $|\nwarrow\rangle$. A dinâmica deste sistema será considerada de forma ideal: o estado inicial do "ponteiro" será

$$|\uparrow\rangle = \frac{1}{\sqrt{2}}(|\nearrow\rangle + |\nwarrow\rangle),$$

com a evolução temporal sendo condicionada ao estado de fenda:

$$|d\rangle \otimes |\uparrow\rangle \quad \longmapsto \quad |d\rangle \otimes |\nearrow\rangle, \tag{8.14a}$$
$$|e\rangle \otimes |\uparrow\rangle \quad \longmapsto \quad |e\rangle \otimes |\nwarrow\rangle. \tag{8.14b}$$

Ao considerar que a partícula em superposição de igual peso dos dois estados de fenda interagiu com o "discriminador de alternativas", teremos a evolução:

$$\frac{1}{\sqrt{2}}(|d\rangle + |e\rangle) \otimes |\uparrow\rangle \longmapsto \frac{1}{\sqrt{2}}(|d, \nearrow\rangle + |e, \nwarrow\rangle), \tag{8.14c}$$

onde o estado final é emaranhado e já usamos uma notação mais compactada.

Já interpretamos anteriormente que o padrão de inferferência é visto quando consideramos uma medição que depende de algum parâmetro. Por exemplo, uma medição projetiva com

$$\Pi(\varphi) = |\varphi\rangle\langle\varphi|, \quad \text{onde} \quad |\varphi\rangle = \frac{1}{\sqrt{2}}\big(|d\rangle + e^{i\varphi}|e\rangle\big),$$

realizada no estado $\frac{1}{\sqrt{2}}(|d\rangle + |e\rangle)$ terá sucesso com probabilidade

$$p(\varphi) = \frac{1}{2}(\langle d| + \langle e|)\Pi(\varphi)(|d\rangle + |e\rangle) = \frac{1}{2}(1 + \cos\varphi) = \cos^2\frac{\varphi}{2}, \tag{8.15}$$

típica de um padrão de interferência.

Se agora consideramos o projetor $\Pi(\varphi) \otimes I$ e o estado final de (8.14), obtemos

$$p(\varphi) = \frac{1}{2}(\langle d, \nearrow| + \langle e, \nwarrow|)\,\Pi(\varphi) \otimes I\,(|d, \nearrow\rangle + |e, \nwarrow\rangle) = \frac{1}{2}, \tag{8.16}$$

e o padrão de interferência se foi.

Uma boa maneira de interpretar esse resultado é que o padrão de interferência presente em (8.15) é fruto da impossibilidade de se distinguir entre as alternativas interferométricas (nas palavras de Feynman, somam-se amplitudes, para depois obter probabilidades), enquanto (8.14) permite esta discriminação e com isso perde-se o padrão de interferência, restando uma soma de alternativas clássicas (8.16).

Alguns resultados interessantes como a teleportação de estados quânticos, a distribuição quântica de chaves e os algoritmos de Deutsch, Shor e Grover podem ser entendidos com o tanto de mecânica quântica já estudado até aqui [Ter07a].

Outro assunto que permeou este capítulo foi a adição de momentum angular. Vimos como considerar operações conjuntas sobre várias partículas de spin $\frac{1}{2}$ faz com que o espaço de estados se decomponha naturalmente em vários multipletos. Em especial, a interpretação normalmente dada à decomposição (8.5) é que a soma de dois spins $\frac{1}{2}$ dá origem a um spin 0 (o singleto) e um spin 1 (o tripleto, associado ao $\mathbb{C}^3$ na fórmula). Da mesma forma, (8.11) será lida como a soma de três spins $\frac{1}{2}$ dá origem a dois spins $\frac{1}{2}$ coletivos, e mais um spin $\frac{3}{2}$, associado ao $\mathbb{C}^4$ presente na decomposição. Por fim, a expressão (8.12) se lê como na soma de quatro spins $\frac{1}{2}$ gera dois singletos (spins 0) distintos, mais três sistemas de três níveis (spins 1) e um sistema de cinco níveis (spin 2).

De maneira mais geral, ao somar um spin j a um spin l, obtemos os possíveis spins entre $|j-l|$ e $j+l$, respeitando a paridade dos multipletos (ou seja, ou são todos sistemas com um número par de níveis, ou todos com um número ímpar). E essa soma é associativa, ou seja, podemos reobter o resultado de (8.12) somando os multipletos obtidos na (8.5).

Capítulo 9

Operador Densidade

Formulamos a Mecânica Quântica usando a linguagem de vetores de estado. Uma formulação alternativa e mais geral é possível usando uma ferramenta conhecida como operador densidade ou matriz densidade [CDL].

O operador densidade é em geral usado para indicar que nosso conhecimento é incompleto devido às imperfeições na preparação dos estados, ou devido à impossibilidade de conhecimento completo do estado quântico do sistema, o que acontece quando o estado de um sistema composto é emaranhado.

9.1 Operador Densidade como Ponto de Partida

Nessa seção vamos generalizar um pouco mais a definição de estado de um sistema físico.

Postulado 9.1 (Estados do sistema)**.** *A cada sistema quântico está associado um espaço vetorial sobre $\mathbb{C}$, que denotaremos por E. Os estados do sistema são representados por operadores positivos semi-definidos de traço um em E, que chamaremos de operadores densidade.*

Novamente, nos preocuparemos apenas com os casos de dimensão

finita. O conjunto de todos os operadores densidade de um sistema físico será denotado por $D(E)$. Veremos em breve que a definição de estado dada em 7.1 é um caso particular da definição 9.1 acima.

Uma característica importante do conjunto de operadores densidade que o torna adequado para ser o conjunto de estados de um sistema é a convexidade.

Definição 9.1. *Um conjunto C em um espaço vetorial real V é chamado convexo se dados dois vetores $v, u \in C$ os pontos*

$$w = \lambda v + (1 - \lambda)u \ , \quad \lambda \in [0, 1],$$

também pertencem a C. O ponto w é chamado combinação convexa de u e v.

Geometricamente, um conjunto C é convexo se dados dois pontos em C o segmento de reta que os liga está contido em C.

Exemplo 9.1. *Um triângulo e um quadrado são conjuntos convexos, assim como uma pirâmide e um cubo. Cada elemento pode ser escrito como soma covexa dos vértices.*

Exemplo 9.2. *O intervalo aberto (a, b) e o intervalo fechado $[a, b]$ em $\mathbb{R}$ são conjuntos convexos. Também são conjuntos convexos os discos abertos e fechados em $\mathbb{R}^2$. Mais geralmente, a bola fechada e a bola aberta de raio r em $\mathbb{R}^n$ são conjuntos convexos.*

Exemplo 9.3. *O quadrante em $\mathbb{R}^n$ formado por todos os pontos cujas coordenadas são positivas é um conjunto convexo.*

Exemplo 9.4. *Uma estrela é um exemplo de um conjunto que não é convexo: os segmentos de reta que ligam as pontas da estrela estão fora dela.*

Teorema 9.1. *O conjunto de operadores densidade de um sistema físico é um conjunto convexo.*

Demonstração. Dados $\rho_1, \rho_2 \in D(E)$ e $p \in [0, 1]$ temos que

$$\rho = p\rho_1 + (1 - p)\rho_2$$

é um operador positivo pois

$$\langle\psi\ |\rho|\ \psi\rangle = p\langle\psi\ |\rho_1|\ \psi\rangle + (1-p)\langle\psi\ |\rho_2|\ \psi\rangle \geq 0.$$

Além disso, ρ possui traço um, uma vez que

$$\mathrm{Tr}\rho = p\mathrm{Tr}\rho_1 + (1-p)\mathrm{Tr}\rho_2 = 1.$$

Logo ρ também é um operador densidade, o que mostra que $D(E)$ é convexo. □

Alguns conjuntos convexos possuem pontos especiais que não podem ser escritos como soma convexa de outros pontos.

Definição 9.2. *Um elemento de um conjunto convexo é chamado extremal se não pode ser escrito como soma convexa de outros elementos de C.*

Exemplo 9.5. *Os vértices são pontos extremais do quadrado, do cubo e do triângulo, e os pontos na esfera de raio r são os pontos extremais da bola fechada de raio r. A bola aberta de raio r e o quadrante em $\mathbb{R}^n$ formado por todos os pontos cujas coordenadas são positivas são exemplos de conjuntos convexos sem pontos extremais.*

Teorema 9.2. *Os pontos extremais de $D(E)$ são os projetores sobre subespaços unidimensionais.*

Demonstração. Todo operador ρ positivo de traço um pode ser escrito em decomposição espectral

$$\rho = \sum_i p_i P_i, \quad p_i \geq 0, \quad \sum_i p_i = 1,$$

em que cada P_i é um projetor sobre um subespaço de dimensão um. Desse modo, todo operador densidade pode ser escrito como soma convexa de projetores. Por outro lado, um projetor em um subespaço unidimensional nunca pode ser escrito como soma convexa de outros. Logo os pontos extremais de $D(E)$ são os projetores P_i. □

Definição 9.3. *Os pontos extremais do conjunto $D(E)$ são chamados estados puros do sistema quântico.*

Os estados que não são puros são chamados *mistos* e sempre podem ser escritos como soma convexa de estados puros. Essa decomposição, no entanto, não é única, e existem muitas maneiras diferentes de escrever um estado misto como soma convexa de estados puros.

Para recuperarmos a definição 7.1, observamos que a cada projetor unidimensional está associado de maneira única uma direção em E. Desse modo, podemos identificar os estados puros de um sistema quântico com as classes de equivalência de vetores unitários em E pela relação

$$|\psi\rangle \sim e^{i\phi}|\psi\rangle, \quad \phi \in \mathbb{R},$$

uma vez que $|\psi\rangle$ e $e^{i\phi}|\psi\rangle$ geram o mesmo subespaço e portanto o projetor associado a eles é o mesmo. Assim, um estado puro do sistema é caracterizado por uma reta complexa passando pela origem em E que como já vimos, são os pontos do espaço projetivo $\mathbb{CP}^{d-1}$. Essa é uma das vantagens de se representar estados usando operadores densidade, pois a cada estado físico corresponde apenas um operador.

Os estados puros, pontos extremais de $D(E)$, são os estados que foram considerados nos capítulos anteriores. Existem muitos operadores densidade em $D(E)$ que não são puros. Em breve veremos por que precisamos deles.

9.1.1 Testes e Operadores Densidade

Por enquanto ainda não vamos alterar a definição de teste 7.2, mas devemos modificar o postulado 7.1 para ajustá-lo à nova definição de estado.

Postulado 9.2. *Sejam $D(E)$ o conjunto de operadores densidade de um sistema físico, $\rho \in D(E)$ um estado e $E = \bigoplus_i E_i$ um teste. Sejam ainda $P_i : E \to E_i$ os projetores ortogonais sobre cada E_i. A probabilidade de obter o resultado i é dada por $p_i = \mathrm{Tr}(\rho P_i)$ e se a alternativa i for obtida, após o teste o sistema será descrito pelo estado $\rho_i = \dfrac{P_i \rho P_i}{\mathrm{Tr}(P_i \rho P_i)}$.*

O postulado acima concorda com o postulado 7.1: quando ρ representa um estado puro, ou seja, quando $\rho = |\psi\rangle\langle\psi|$ temos que $p_i = \mathrm{Tr}(\rho P_i) = \langle\psi \ |P_i|\ \psi\rangle$, e

$$\rho_i = |\psi_i\rangle\langle\psi_i|,$$

em que

$$|\psi_i\rangle = \frac{P_i|\psi\rangle}{\|P_i|\psi\rangle\|}.$$

Exercício 9.1. *Demonstre as afirmações acima.*

Exercício 9.2. *Mostre que a reprodutibilidade dos testes também vale com o postulado 9.2.*

Como vimos no capítulo 7, testes estão associados à decomposição espectral de operadores auto-adjuntos. Dado um operador auto-adjunto A, os resultados possíveis para cada medição são dados pelos autovalores de A, que serão denotados por a_i. As probabilidades continuam iguais: $p_i = \mathrm{Tr}(\rho P_i)$. O valor esperado de A em um estado ρ é

$$\langle A\rangle = \sum_i p_i a_i = \sum_i a_i \mathrm{Tr}(\rho P_i) = \mathrm{Tr}\left(\rho\left(\sum_i a_i P_i\right)\right) = \mathrm{Tr}(\rho A). \tag{9.1}$$

Exercício 9.3. *Mostre que se ρ representa um estado puro, $\langle A\rangle$ dado pela equação acima concorda com o que foi provado no capítulo 7 para um estado $|\psi\rangle$:*

$$\langle A\rangle = \langle\psi|A|\psi\rangle.$$

9.1.2 Estados Mistos de um Qbit

Um estado geral de um qbit é representado por um operador densidade agindo em $\mathbb{C}^2$. O conjunto dos operadores hermitianos é um espaço vetorial real e uma base para esse espaço é formado pelos operadores de Pauli juntamente com o operador identidade I. Desse modo, um operador densidade de um qbit pode ser sempre escrito na forma

$$\rho = \frac{1}{2}(I + a\sigma_1 + b\sigma_2 + c\sigma_3). \tag{9.2a}$$

O coeficiente de I deve ser $1/2$ porque ela é a única matriz da base que tem traço não nulo, igual a dois, e $\mathrm{Tr}(\rho) = 1$. Agora devemos impor condições ao vetor $\begin{pmatrix} a & b & c \end{pmatrix}$ para que o operador seja positivo.

Em forma matricial temos

$$\rho = \frac{1}{2}\begin{bmatrix} 1+c & a-ib \\ a+ib & 1-c \end{bmatrix}. \tag{9.2b}$$

Para que ρ seja uma matriz positiva é necessário e suficiente que $det(\rho) \geq 0$, uma vez que $\mathrm{Tr}(\rho) \geq 0$. Essa condição é equivalente a

$$a^2 + b^2 + c^2 \leq 1. \tag{9.2c}$$

Logo podemos fazer uma associação bijetiva entre operadores densidade de um qbit e pontos na bola de raio um em $\mathbb{R}^3$, comumente chamada *bola de Bloch*. Os pontos na esfera S^2 correspondem aos operadores que possuem determinante igual a zero, que nesse caso são exatamente os estados puros. Essa associação coincide com a que fizemos utilizando a fibração de Hopf na seção 6.1.5.

9.2 Operador Densidade como fruto da Ignorância Clássica

Vamos entender agora porque é necessário aumentarmos o espaço de estados para incluir estados mistos. Suponhamos que um aparato prepara vários exemplares de um sistema físico cujo espaço de estados é E. Suponhamos também que a preparação pode ser feita em dois estados puros distintos: com probabilidade p o sistema é preparado no estado $\rho_1 = |\psi_1\rangle\langle\psi_1|$ e com probabilidade $1-p$ o sistema é preparado no estado $\rho_2 = |\psi_2\rangle\langle\psi_2|$. Quando um dos exemplares é liberado pelo aparato não sabemos em qual estado ele foi preparado. Nesse cenário, a descrição do estado liberado pelo aparato é feita de acordo com o seguinte postulado

Postulado 9.3. *Se um sistema físico foi preparado no estado ρ_1 com probabilidade p ou no estado ρ_2 com probabilidade $1-p$ então a matriz densidade que o descreve é*

$$\rho = p\rho_1 + (1-p)\rho_2.$$

Se realizarmos um teste com alternativas clássicas i no sistema considerado acima devemos obter a resposta i com probabilidade $p_i =$

$pp_i^1 + (1-p)p_i^2$ em que p_i^j é a probabilidade de obtermos i no estado ρ_j, $j = 1, 2$. De fato

$$p_i = \mathrm{Tr}(\rho P_i) = p\mathrm{Tr}(\rho_1 P_i) + (1-p)\mathrm{Tr}(\rho_2 P_i) = pp_i^1 + (1-p)p_i^2.$$

9.3 Operador Densidade como fruto da Ignorância Quântica

A seção anterior mostra que nos casos em que não possuímos informação completa sobre o sistema ele será representado por um estado misto. Existe outra situação em que, mesmo começando com um estado puro, somos levados a considerar estados mistos: quando temos acesso a apenas um dos subsistemas de um sistema composto.

Se o espaço de estados do sistema A é E_A e o espaço de estados do sistema B é E_B então os estados do sistema composto AB são representados por operadores densidade em $E_A \otimes E_B$. Podemos usar o isomorfismo

$$\Psi : L(E_A) \otimes L(E_B) \longrightarrow L(E_A \otimes E_B)$$

definido para vetores decomponíveis da forma

$$\Psi(O_A \otimes O_B)|v_A\rangle \otimes |v_B\rangle = O_A|v_A\rangle \otimes O_B|v_B\rangle$$

e estendido por linearidade para os outros vetores. Esse é um isomorfismo que preserva traço e positividade, de modo que os operadores densidade em $E_A \otimes E_B$ podem ser vistos como operadores positivos de traço um em $L(E_A) \otimes L(E_B)$.

Seria adequado associar um estado, e portanto um operador densidade, a cada sistema simples, especialmente em um cenário onde as partes A e B estejam separadas espacialmente. Para isso vamos precisar da definição de traço parcial 2.11.

Postulado 9.4 (Operadores densidade reduzidos)**.** *Dado o operador densidade ρ que descreve um sistema quântico composto AB, o operador densidade ρ_A que descreve o sistema A é dado por*

$$\rho_A = \mathrm{Tr}_B(\rho),$$

e o operador ρ_B que descreve o sistema B é dado por

$$\rho_B = \mathrm{Tr}_A(\rho).$$

O operador ρ_A é chamado operador densidade reduzido do sistema A e ρ_B é chamado operador densidade reduzido do sistema B.

Exercício 9.4. *Seja ρ o operador densidade associado a um sistema composto AB e ρ_A o operador densidade reduzido associado à parte A. Mostre que as probabilidades para o teste local associado ao operador $O \otimes I$ realizado em ρ são iguais às probabilidades para o teste associado ao operador O realizado em ρ_A.*

Exercício 9.5. *Encontre os operadores densidade reduzidos ρ_A e ρ_B de um sistema de dois qbits que se encontra em um dos estados de Bell. Verifique que o estado do sistema composto não é o produto tensorial $\rho_A \otimes \rho_B$.*

O exercício acima mostra que, apesar de podermos associar operadores densidade ρ_A e ρ_B a um sistema composto descrito pelo estado ρ, não é sempre verdade que $\rho = \rho_A \otimes \rho_B$. Além disso, mesmo que ρ represente um estado puro, ρ_A e ρ_B podem não o ser! De acordo com Schrödinger [Sch], uma outra maneira de expressarmos essa situação peculiar da mecânica quântica é: "*The best possible knowledge of a whole does not necessarily include the best possible knowledge of all its parts*". Essa é mais uma surpresa que aparece como consequência do emaranhamento: para estados emaranhados, mesmo puros, os operadores densidade reduzidos são sempre mistos.

Teorema 9.3. *Um estado puro é fatorável se e somente se os operadores densidade reduzidas ρ_A e ρ_B correspondem a estados puros.*

Demonstração. Basta utilizarmos a decomposição de Schmidt

$$|\psi\rangle = \sum_i a_i |ii\rangle.$$

Se $|\psi\rangle$ é fatorável então apenas um coeficiente de Schmidt a_j pode ser não nulo de modo que $\rho_A = |j\rangle\langle j|$ e $\rho_B = |j\rangle\langle j|$ são estados puros. Por outro lado, se dois ou mais coeficientes de Schmidt são não nulos então temos que

$$\rho_A = \sum_i a_i^2 |i\rangle\langle i|$$

é um estado misto. □

Corolário 9.4. *(da demonstração) Para um estado puro dos sistema de duas partes, os autovalores não nulos de ρ_A e ρ_B são os mesmos, com as mesmas multiplicidades.*

9.4 Medições Generalizadas

Agora que já generalizamos a noção de estado, podemos também propor medições generalizadas [NC].

Definição 9.4. *Uma medição generalizada será dada por um conjunto $\{M_i\}$ de* operadores de medição *no espaço de estados tais que $\sum_i M_i^\dagger M_i = I$, onde I denota o operador identidade. Se o estado do sistema antes da medição é ρ, a probabilidade de obter o resultado i é dada por $p_i = \mathrm{Tr}\left(M_i \rho M_i^\dagger\right)$, e caso o resultado i seja obtido, o estado do sistema após a medição será $\rho_i = \dfrac{M_i \rho M_i^\dagger}{\mathrm{Tr} M_i \rho M_i^\dagger}$.*

Os operadores $M_i^\dagger M_i$ são positivos semi-definidos. Desta forma, uma medição generalizada está associada a uma partição do operador identidade em soma de operadores positivos semi-definidos. Por este motivo, esta definição está ligada ao conceito de *medida a valores em operadores positivos*, com a sigla em inglês POVM - medida aqui tendo seu sentido matemático usual (e não o sentido físico de uma medição). De fato, a definição 9.4 pede um pouco mais que uma POVM, uma vez que os operadores M_i são dados. O conhecimento da POVM permite obter as probabilidades dos possíveis resultados posteriores, mas não permite definir o estado do sistema após a medição.

Exercício 9.6. *Mostre que as medições generalizadas incluem as medições projetivas da definição 7.2.*

Exemplo 9.6 (Processo de medição para qbits)**.** *Sejam $\{\vec{v_i}\}$ um conjunto de vetores unitários em $\mathbb{R}^3$ e α_i constantes tais que $0 < \alpha_i < 1$, $\sum_i \alpha_i^2 = 1$ e*

$$\sum_i \alpha_i^2 \vec{v_i} = 0.$$

Então os operadores

$$M_i = \frac{\alpha_i}{\sqrt{2}}(I + \vec{v_i} \cdot \vec{\sigma})$$

definem um processo de medição para o sistema de um qbit [LBe].

Exemplo 9.7. *Particularizando o exemplo anterior,*

$$\begin{aligned} \vec{v_1} &= (0,0,1), \\ \vec{v_2} &= \left(-\frac{\sqrt{3}}{2}, 0, -\frac{1}{2}\right), \\ \vec{v_3} &= \left(\frac{\sqrt{3}}{2}, 0, -\frac{1}{2}\right) \end{aligned}$$

e $\alpha_1 = \alpha_2 = \alpha_3 = \frac{1}{\sqrt{3}}$ *satisfazem as condições acima e portanto os operadores associados*

$$M_1 = \frac{1}{\sqrt{6}}(I + \sigma_z),$$

$$M_2 = \frac{1}{\sqrt{6}}\left(I - \frac{\sqrt{3}}{2}\sigma_x - \frac{1}{2}\sigma_z\right),$$

$$M_3 = \frac{1}{\sqrt{6}}\left(I + \frac{\sqrt{3}}{2}\sigma_x - \frac{1}{2}\sigma_z\right)$$

definem um processo de medição para um qbit.

Exercício 9.7. *Encontre o POVM relacionado ao processo de medição descrito no exemplo 9.7.*

Exercício 9.8.

1. *Encontre a probabilidade de encontrarmos o valor* 1 *se realizarmos a medição descrita no exemplo 9.7 no estado* $|+\rangle$ *e o estado do sistema após a medição caso esse resultado seja obtido.*

2. *Seja* ρ_1 *o operador densidade obtido no item anterior. Mostre que se repetirmos o processo de medição nesse estado a probabilidade de obtermos os resultados* 2 *e* 3 *é não nula. A condição de reprodutibilidade continua valendo para medições generalizadas?*

9.5 Evolução Temporal

Nos capítulos anteriores abordamos a evolução temporal de um sistema isolado, dada pela equação de Schrödinger. Agora vamos considerar o caso mais geral de evolução temporal [BŻ].

Vamos estudar os *mapas quânticos*, que são mapas que levam o conjunto de matrizes densidade nele próprio, de uma maneira que faça sentido do ponto de vista físico, o que explicaremos melhor mais a frente.

Dado um sistema físico com espaço de estados E de dimensão d, vamos fixar uma base ortonormal em E e representar um operador densidade em $D(E)$ por sua matriz em relação a essa base, que chamaremos matriz densidade.[1] Vamos começar com mapas

$$\begin{array}{rcl} \Phi : D(E) & \longrightarrow & M(E) \\ \rho & \longmapsto & \rho', \end{array}$$

tais que $\Phi(D(E)) \subset D(E)$.

A primeira condição que exigimos de um mapa desse tipo é que ele seja linear. A justificativa para tal restrição é que não queremos que o resultado da operação dependa de como escrevemos uma matriz densidade como soma convexa de outras. Desse modo temos:

$$\Phi(p_1\rho_1 + p_2\rho_2) = p_1\Phi(\rho_1) + p_2\Phi(\rho_2).$$

O mapa Φ pode ser representado por uma matriz que age em um espaço vetorial de dimensão d^2, ou seja, uma matriz $d^2 \times d^2$. Usaremos dois índices para indicar as componentes de uma matriz densidade $(d \times d)$ e quatro índices para indicar as componentes de um mapa agindo no espaço de matrizes densidade $(d^2 \times d^2)$. Assim temos:

$$\rho'_{m\mu} = \sum_{n\nu} \Phi^{m\mu}_{n\nu} \rho_{n\nu}.$$

O mapa Φ deve levar matrizes densidade em matrizes densidade, ou seja, ρ' deve ser uma matriz positiva de traço um. Isso implica que

[1] Também utilizaremos a notação ρ para uma matriz densidade e $D(E)$ para o conjunto de todas as matrizes densidade do sistema.

1. $\Phi(\rho)$ deve ser autoadjunta:

$$\rho' = (\rho')^\dagger : \rho'_{m\mu} = \overline{\rho'_{\mu m}} \Rightarrow$$

$$\sum_{n\nu} \Phi_{\substack{m\mu \\ n\nu}} \rho_{n\nu} = \sum_{n\nu} \overline{\Phi_{\substack{\mu m \\ \nu n}}}\,\overline{\rho_{\nu n}} = \sum_{n\nu} \overline{\Phi_{\substack{\mu m \\ \nu n}}} \rho_{n\nu} \Rightarrow \Phi_{\substack{m\mu \\ n\nu}} = \overline{\Phi_{\substack{\mu m \\ \nu n}}}.$$

   A última implicação é óbvia quando consideramos Φ como um mapa do espaço $M(E)$ em $M(E)$. No entanto, se consideramos Φ como um mapa de $D(E)$ em $D(E)$, ela continua válida. Para vermos isso, basta usarmos matrizes com apenas um elemento não nulo, igual a um, na diagonal, e matrizes com apenas um bloco 2×2 não nulo na diagonal, dos tipos

$$\frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad \frac{1}{2}\begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}.$$

2. $\mathrm{Tr}(\rho') = 1$:

$$\sum_m \rho'_{mm} = \sum_m \sum_{n\nu} \Phi_{\substack{mm \\ n\nu}} \rho_{n\nu} = 1.$$

   Como essa equação deve valer para todo ρ, podemos usar $\rho = |i\rangle\langle i|$, caso em que $\rho_{n\nu} = \delta_{n\nu}\delta_{ni}$, para concluir que $\sum_m \Phi_{\substack{mm \\ n\nu}} = 1$, se $n = \nu$. Para concluir que $\sum_m \Phi_{\substack{mm \\ n\nu}} = 0$, se $n \neq \nu$, utilizamos novamente as matrizes com os blocos mostrados acima.

3. A matriz ρ' deve ser positiva, ou seja, Φ deve levar matrizes positivas em matrizes positivas.

Definição 9.5. *Um mapa $\Phi : M(E) \longrightarrow M(E)$ é chamado positivo se $\Phi(\rho)$ é positiva para toda matriz positiva ρ.*

Para estudarmos melhor que restrições essas propriedades impõem ao mapa Φ, vamos definir a *matriz dinâmica* associada a Φ:

$$D_{\substack{mn \\ \mu\nu}} = \Phi_{\substack{m\mu \\ n\nu}}.$$

Em termos da matriz dinâmica as condições acima podem ser dadas por:

1. $\rho' = (\rho')^\dagger \Leftrightarrow D = D^\dagger$.

2. $\mathrm{Tr}(\rho') = 1 \Leftrightarrow \sum_m D^{mn}_{m\nu} = \delta_{n\nu}$.

Resta estudar qual é a condição imposta a D pela positividade de Φ. Vejamos inicialmente o que acontece para estados puros $\rho = |z\rangle\langle z|$, $\rho_{n\nu} = z_n\overline{z_\nu}$. Se Φ for positivo então ρ' é positiva, o que implica que:

$$\begin{aligned} 0 \leq \langle x|\rho'|x\rangle &= \sum_{m\mu} \overline{x_m}\rho'_{m\mu}x_\mu = \sum_{m\mu}\overline{x_m}\left(\sum_{n\nu} z_n D^{mn}_{\mu\nu}\overline{z_\nu}\right)x_\mu = \\ &= \langle w|\langle x|D|x\rangle|w\rangle, \end{aligned}$$

em que $|w\rangle$ é o vetor cuja cordenada w_n é igual a $\overline{z_n}$, $|x\rangle|w\rangle = |x\rangle\otimes|w\rangle$ e $\langle w|\langle x|$ é o elemento de $(E\otimes E)^*$ associado a $|x\rangle|w\rangle$. Logo, se Φ é um mapa positivo, D deve satisfazer a condição $\langle w|\langle x|D|x\rangle|w\rangle \geq 0$ para todos $|w\rangle, |x\rangle \in E$. Essa propriedade é chamada *positividade por blocos*.

Para ver que essa condição além de necessária é também suficiente, devemos mostrar que ela implica que ρ' é positiva também quando ρ é um estado misto, o que segue por convexidade. Tomamos $\rho = \sum_i p_i|z_i\rangle\langle z_i|$, $\rho_{n\nu} = \sum_i p_i(z_i)_n\overline{(z_i)_\nu}$. Nesse caso,

$$\begin{aligned} \langle x|\rho'|x\rangle &= \sum_{m\mu}\overline{x_m}\rho'_{m\mu}x_\mu = \sum_{m\mu}\overline{x_m}\sum_{n\nu} D^{mn}_{\mu\nu}\left(\sum_i p_i(z_i)_n\overline{(z_i)_\nu}\right)x_\mu \\ &= \sum_i p_i \sum_{m\mu}\overline{x_m}\left(\sum_{n\nu}(z_i)_n D^{mn}_{\mu\nu}\overline{(z_i)_\nu}\right)x_\mu \geq 0. \end{aligned}$$

Isso prova o seguinte teorema:

Teorema 9.5 (Jamiołkowski)**.** *Um mapa linear $\Phi : M(E) \longrightarrow M(E)$ é positivo se e somente se a matriz dinâmica é positiva por blocos.*

No entanto, a positividade do mapa Φ não é suficiente para que ele represente uma operação fisicamente permitida. Suponhamos que nosso sistema seja apenas um subsistema de um sistema maior cujo

espaço de estados é $E \otimes E'$, em que E é o espaço de estados associado ao nosso sistema de interesse e E' é o espaço de estados de um sistema adicional. Gostaríamos que um mapa fisicamente permitido não só levasse a matriz densidade do nosso sistema em uma matriz densidade, mas que também o fizesse se considerarmos a operação agindo em $E \otimes E'$. Isso quer dizer que não só Φ deve ser um mapa positivo, mas também deve ser positiva toda extensão da forma $\Phi \otimes I$, em que I é o operador identidade em $M(E')$.

Definição 9.6. *Se o mapa $\Phi \otimes I$ agindo em $M(E \otimes E')$ é positivo, em que E' é um espaço vetorial de dimensão k, dizemos que Φ é um mapa k-positivo. Se Φ é um mapa k-positivo para todo $k \in \mathbb{N}$ então Φ é chamado um mapa completamente positivo.*

A exigência que impomos agora em Φ é que ele seja um mapa completamente positivo. Vejamos que implicação essa propriedade tem sobre a matriz dinâmica correspondente. Como ela é uma matriz $d^2 \times d^2$, podemos visualizá-la como uma matriz agindo em um espaço vetorial de dimensão d^2, que pode ser identificado com $E \otimes E$. Como ela é hermitiana, podemos escrevê-la em decomposição espectral:[2]

$$D = \sum_i d_i |\chi^i\rangle\langle\chi^i|, \quad \underset{\mu\nu}{D_{mn}} = \sum_i d_i \chi^i_{mn} \overline{\chi^i_{\mu\nu}}.$$

Tomamos um estado puro em um espaço de estados estendido,

$$\rho \in M(E \otimes E'), \quad \rho_{mm'\mu\mu'} = z_{mm'} \overline{z_{\mu\mu'}}.$$

e aplicamos o mapa estendido $\Phi \otimes I$ a ρ:

$$\begin{aligned}
\rho'_{mm'\mu\mu'} &= \sum_{nn'\nu\nu'} (\Phi \otimes I)_{\substack{mm'\mu\mu' \\ nn'\nu\nu'}} \rho_{nn'\nu\nu'} \\
&= \sum_{nn'\nu\nu'} \Phi_{\substack{m\mu \\ n\nu}} I_{\substack{m'\mu' \\ n'\nu'}} \rho_{nn'\nu\nu'} \\
&= \sum_{nn'\nu\nu'} \Phi_{\substack{m\mu \\ n\nu}} \delta_{m'n'} \delta_{\mu'\nu'} z_{nn'} \overline{z_{\nu\nu'}} \\
&= \sum_{n\nu} \Phi_{\substack{m\mu \\ n\nu}} z_{nm'} \overline{z_{\nu\mu'}} = \sum_{n\nu} \sum_i d_i \chi^i_{mn} z_{nm'} \overline{\chi^i_{\mu\nu}} \overline{z_{\nu\mu'}}.
\end{aligned}$$

[2] Escrevemos $|\chi^i\rangle$ com dois índices pois estamos usando a estrutura tensorial de $E \otimes E$.

Agora tomamos um outro vetor $|x\rangle \in E' \otimes E$ e testamos se $\langle x|\rho'|x\rangle \geq 0$:

$$\begin{aligned}
\langle x|\rho'|x\rangle &= \sum_{mm'\mu\mu'} \overline{x_{mm'}}\rho'_{mm'\mu\mu'}x_{\mu\mu'} \\
&= \sum_{mm'\mu\mu'} \overline{x_{mm'}}\left(\sum_{n\nu}\sum_i d_i\chi^i_{mn}z_{nm'}\overline{\chi^i_{\mu\nu}}\overline{z_{\nu\mu'}}\right)x_{\mu\mu'} \\
&= \sum_i d_i\left(\sum_{mm'n}\chi^i_{mn}z_{nm'}\overline{x_{mm'}}\right)\left(\sum_{\mu\mu'\nu}\overline{\chi^i_{\mu\nu}}\overline{z_{\nu\mu'}}x_{\mu\mu'}\right) \\
&= \sum_i d_i\left|\sum_{mm'n}\chi^i_{mn}z_{nm'}\overline{x_{mm'}}\right|^2.
\end{aligned}$$

Essa quantidade deve ser não-negativa para todo $|z\rangle$ e todo $|x\rangle$ que escolhermos. Isso só acontece se cada um dos d_i for um número não-negativo, ou seja, se D for uma matriz positiva semi-definida.

Por outro lado, se D é uma matriz positiva e $\rho = \sum_j p_j|z^j\rangle\langle z^j|$, então vale:

$$\begin{aligned}
&\sum_{mm'\mu\mu'} \overline{x_{mm'}}\rho'_{mm'\mu\mu'}x_{\mu\mu'} = \\
&= \sum_i\sum_j d_ip_j\left(\sum_{mm'n}\chi^i_{mn}z^j_{nm'}\overline{x_{mm'}}\right)\left(\sum_{\mu\mu'\nu}\overline{\chi^i_{\mu\nu}}\overline{z^j_{\nu\mu'}}x_{\mu\mu'}\right) \\
&= \sum_i\sum_j d_ip_j\left|\sum_{mm'n}\chi^i_{mn}z^j_{nm'}\overline{x_{mm'}}\right|^2 \geq 0.
\end{aligned}$$

Com isso, acabamos de provar o seguinte teorema:

Teorema 9.6 (Choi)**.** *Um mapa linear Φ é completamente positivo se e somente se a matriz dinâmica correspondente é positiva semi-definida.*

Uma forma muito útil de caracterizar os mapas completamente positivos é através da representação de Kraus [Kra].

Teorema 9.7 (Representação de Kraus)**.** *Um mapa linear Φ é completamente positivo se, e somente se, é da forma*

$$\rho \longmapsto \rho' = \sum_i A_i \rho A_i^\dagger,$$

em que cada A_i é uma matriz quadrada da mesma dimensão de ρ. Além disso, Φ preserva o traço se, e somente se, as matrizes A_i satisfazem

$$\sum_i A_i^\dagger A_i = I.$$

Demonstração. Suponhamos que Φ seja completamente positivo e seja D a matriz dinâmica associada. Como D é positiva, pode ser escrita em decomposição espectral

$$D = \sum_i d_i |\chi^i\rangle\langle\chi^i| \ , \quad d_i > 0.$$

Definindo $|A^i\rangle = \sqrt{d_i}|\chi^i\rangle$, temos que

$$D = \sum_i |A^i\rangle\langle A^i|, \quad \underset{\mu\nu}{D^{mn}} = \sum_i A^i_{mn}\overline{A^i_{\mu\nu}}.$$

Cada vetor $|A^i\rangle \in E\otimes E$ possui d^2 coordenadas que indexamos usando dois índices para deixar evidente a estrutura de produto tensorial. Assim podemos identificar cada $|A^i\rangle$ com um operador A_i agindo em E da forma $(A_i)_{mn} = A^i_{mn}$. Daí temos:

$$\rho'_{m\mu} = \sum_{n\nu} \underset{n\nu}{\Phi^{m\mu}}\rho_{n\nu} = \sum_{n\nu} \underset{\mu\nu}{D^{mn}}\rho_{n\nu} =$$

$$\sum_{n\nu}\sum_i A^i_{mn}\overline{A^i_{\mu\nu}}\rho_{n\nu} = \sum_{n\nu}\sum_i (A_i)_{mn}\rho_{n\nu}(A_i)^\dagger_{\nu\mu} = \sum_i (A_i\rho A_i^\dagger)_{m\mu}$$

$$\Rightarrow \rho' = \sum_i A_i \rho A_i^\dagger.$$

Se Φ preservar o traço, temos também:

$$\begin{aligned}\delta_{\nu n} &= \sum_m \underset{m\nu}{D^{mn}} = \sum_i\sum_m (A_i)_{mn}\overline{(A_i)_{m\nu}} \\ &= \sum_i\sum_m (A_i)^\dagger_{\nu m}(A_i)_{mn} = \sum_i (A_i^\dagger A_i)_{\nu n},\end{aligned}$$

ou seja,

$$\sum_i A_i^\dagger A_i = I.$$

Por outro lado, se $\Phi(\rho) = \sum_i A_i \rho A_i^\dagger$, então

$$\Phi \otimes I(\sigma) = \sum_i A_i \otimes I(\sigma) A_i^\dagger \otimes I,$$

que é claramente um mapa positivo. □

Na demonstração acima, usamos o fato de que $L(E,F) \equiv E^* \otimes F$, em que E^* denota o espaço dual de E. Como estamos trabalhando em dimensão finita, vale $E^* \equiv E$ de modo que $L(E,F) \equiv E \otimes F$. Assim temos $L(E) \equiv E \otimes E$ e podemos identificar cada vetor $|A^i\rangle$ com um operador A_i.

9.6 Uma Axiomatização Alternativa

Seguindo o caminho de Walter Thirring [Thi], vamos apresentar uma outra axiomatização para a mecânica quântica, onde os conceitos centrais são os observáveis, enquanto estados são apenas as ferramentas que levam estes objetos a seus valores esperados. Para bem apreciar este capítulo, assumimos que o leitor já tem uma familiaridade mínima com álgebras C^*, como aqui apresentado no capítulo 5.

9.6.1 Mecânica Quântica e Álgebras de Operadores

Até este momento a mecânica quântica foi apresentada com ênfase no conceito de estado, visto inicialmente como um vetor de um espaço vetorial complexo, e depois considerado como um operador densidade. Esta passagem de vetor para operador dá origem a uma visão da mecânica quântica baseada fundamentalmente em operadores e não mais em vetores.

Nessa visão, os observáveis são os elementos hermitianos de uma álgebra C^* com unidade, denotada por A; os estados (no sentido da mecânica quântica) são os estados da álgebra A, ou seja, funcionais

lineares f tais que $f(a^*a) \geq 0$ e $f(1) = 1$. Uma medição do observável a tem seus resultados contidos no espectro do elemento a, denotado por $\sigma(a)$.

Para deixar o parágrafo acima menos misterioso, vamos identificar esses elementos no caso de um qbit. A álgebra A em questão é a álgebra $M_2(\mathbb{C})$ das matrizes 2×2 com coeficientes complexos. Os observáveis, que são os elementos hermitianos de A, correspondem a matrizes tais que $(\overline{a})^T = a$, ou seja, tais que as entradas satisfazem $a_{i,j} = \overline{a}_{j,i}$. Os estados são funcionais positivos e tais que $f(1) = 1$. Como visto no final do capítulo sobre álgebras, esses funcionais podem ser escritos como $f(x) = \mathrm{Tr}\Big(V_f^* x\Big)$ onde V_f é uma matriz de traço unitário e positiva, ou seja, os estados correspondem exatamente a matrizes densidade. Um observável como σ_x, ao ser medido, produz resultados que estão no espectro do elemento σ_x, que correspondem ao seus autovalores (que, nesse caso, sabemos ser -1 e 1).

9.6.2 Mas nem é tão novo assim...

Porém nos cabe lembrar que a visão acima, embora o conceito de álgebra de operadores tenha sido efetivamente criado após o surgimento da mecânica quântica, não é exclusividade do mundo quântico: na verdade podemos representar a mecânica clássica da mesma forma. Por exemplo, considere uma partícula que se move na reta, descrita pela hamiltoniana

$$H = \frac{p^2}{2m} + V(q).$$

Os observáveis típicos nesse caso são posição (q) e momento (p, que está ligado a velocidade), mas podemos pensar em qualquer função dessas variáveis como sendo também um observável. A energia cinética, $K = \frac{p^2}{2m}$ é um exemplo. Portanto o conjunto de observáveis é na verdade o conjunto de funções contínuas reais $C(\mathbb{R}^2) = \{f \colon \mathbb{R}^2 \to \mathbb{R}\}$; este conjunto é uma álgebra C^* com produto definido por $(f.g)(x) = f(x)g(x)$. Mas note então que essa álgebra é *comutativa*, ao contrário da álgebra de matrizes que está associada à descrição quântica. Portanto podemos dizer que a novidade de fato na passagem do mundo

clássico para o quântico é a troca de uma álgebra de observáveis comutativa por uma não-comutativa.

9.7 Mais um bocadinho de Física

Naturalmente, toda essa discussão encontra aplicações diversas. Operadores densidade são usados, por exemplo, para descrever os estados de equilíbrio térmico, fazendo a fronteira da mecânica quântica com a *mecânica estatística* e também com a *termodinâmica*.

Mas também podemos encontrar aplicações do postulado 9.3 em áreas como a criptografia. Para ser mais preciso, podemos utilizá-lo para interpretar o protocolo de distribuição quântica de chaves criptográficas BB84 [BB84]. Nesse protocolo Ana prepara estados, que são enviados para Bernardo, que faz um teste. Até aí, nada demais. O interessante é que Ana prepara sempre um de quatro estados de um qbit: $\{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$, de maneira equiprovável. Já Bernardo faz sempre um de dois testes: $\mathcal{X}$ ou $\mathcal{Z}$. Se considerarmos que o sorteio de Ana corresponde a um par de bits clássicos: o primeiro definindo qual base ela irá usar: $\mathcal{X}$ ou $\mathcal{Z}$, e o segundo dizendo qual dos dois estados dessa base ela deve preparar, a mecânica quântica estudada no capítulo 6 é suficiente para dizer que quando o primeiro bit de Ana coincide com o bit que Bernardo sorteia para definir o teste que irá usar, o segundo bit de Ana estará completamente correlacionado com o bit que Bernardo irá extrair deste teste. Por outro lado, se o primeiro bit de Ana for distinto, o resultado da medição é independente da preparação. O protocolo segue com Bernardo divulgando, já de posse do resultado, qual dos dois testes ele realizou e Ana, após comparar com sua preparação, decide pela aceitação ou descarte do bit obtido. Outras estratégias clássicas de amplificação de privacidade e reconciliação de informação são adotadas de forma a gerar uma chave privada, utilizando um canal quântico público.

O ponto central, não para a criptografia, mas para a física que queremos discutir aqui, está em considerar todo o processo de Ana como uma preparação de estado. De maneira bem geral (depois vamos incluir a equiprobabilidade), o qbit enviado por Ana pode ser

descrito por

$$\rho = p(\mathcal{X},+)|+\rangle\langle+| + p(\mathcal{X},-)|-\rangle\langle-| + p(\mathcal{Z},0)|0\rangle\langle 0| + p(\mathcal{Z},1)|1\rangle\langle 1|, \tag{9.3}$$

que pode ser reescrito como

$$\begin{aligned} \rho &= p(\mathcal{X})\,(p(+|\mathcal{X})|+\rangle\langle+| + p(-|\mathcal{X})|-\rangle\langle-|) + \\ &+ p(\mathcal{Z})\,(p(0|\mathcal{Z})|0\rangle\langle 0| + p(1|\mathcal{Z})|1\rangle\langle 1|) \qquad (9.4a) \\ &= p(\mathcal{X})\,\rho_{\mathcal{X}} + p(\mathcal{Z})\,\rho_{\mathcal{Z}}. \qquad (9.4b) \end{aligned}$$

Usando a imagem da bola do Bloch, $\rho_{\mathcal{X}}$ é um ponto no segmento que une $|+\rangle\langle+|$ e $|-\rangle\langle-|$, ambos no equador da esfera, enquanto $\rho_{\mathcal{Z}}$ é um ponto no eixo que une os polos $|0\rangle\langle 0|$ e $|1\rangle\langle 1|$.

Se usássemos $\rho_{\mathcal{X}}$ ou $\rho_{\mathcal{Z}}$, exclusivamente, no protocolo descrito anteriormente, não haveria segredo algum, pois conhecedor da estratégia utilizada, qualquer espião poderia fazer o teste correspondente, para depois enviar o estado que ele tivesse após a medição para Bernardo. O interessante é que, ao exigirmos também equiprobabilidade, estaremos na intersecção dos dois segmentos, ou seja

$$\rho_{\mathcal{X}} = \rho_{\mathcal{Z}},$$

e consequentemente

$$\rho = \rho_{\mathcal{X}} = \rho_{\mathcal{Z}} = \frac{1}{2}I.$$

Moral da história, do ponto de vista de descrição de estado, ou ainda, se alguém fosse usar os bits que Ana prepara, sem nunca mais voltar a se comunicar com ela, teria o estado maximamente misto em mãos. Ainda mais interessante: a discussão acima mostra três maneiras distintas de Ana "preparar o estado maximamente misto": $\rho_{\mathcal{X}}$, $\rho_{\mathcal{Z}}$ e ρ. Existem ainda várias outras. O interessante é que a preparação (9.3), aliada ao conhecimento que Bernardo tem dela e à possibilidade dele se comunicar com Ana, permite o estabelecimento da chave.

Aqui fizemos toda essa discussão em termos de ignorância clássica (sec. 9.2), mas você pode refraseá-lo em termos de ignorância quântica (sec. 9.3) e ver que, nesse caso, o emaranhamento entre Ana e Bernardo (antes que ele fizesse a medição) desempenharia um papel interessante.

Capítulo 10

Sistemas Quânticos Compostos - bis

Agora que já temos uma definição mais geral de estados e medições quânticos, podemos examinar com outros olhos as correlações presentes em sistemas compostos. Mantendo o espírito do texto, vamos abordar vários assuntos, alguns deles sob intensa investigação atual, começando pelo caso mais simples e introduzindo generalidade e complexidade posteriormente. Não poderemos[1] nos aprofundar em todos esses assuntos. Vemos isso como um convite ao leitor para cuidar de seu próprio aprofundamento, tornando-se assim um pesquisador do assunto[2].

10.1 Dois Qbits

Já sabemos que o espaço de estados para dois qbits é isomorfo a $\mathbb{C}^4 \cong \mathbb{C}^2 \otimes \mathbb{C}^2$, com cada $\mathbb{C}^2$ correspondendo ao espaço de estados de um qbit. Da mesma forma, sabemos os estados, propriamente ditos, são dados por $D\big(\mathbb{C}^2 \otimes \mathbb{C}^2\big)$, um conjunto convexo, fechado, de dimensão real 15, contido no espaço vetorial real dos operadores auto-adjuntos em $\mathbb{C}^4$.

[1] Por limitação de espaço, de tempo e mesmo de conhecimento.

[2] Seja para saciar sua curiosidade, seja como atividade profissional.

Exercício 10.1. *Prove as afirmações acima.*

Da mesma forma que no capítulo 8 se mostrava importante entender os estados de sistemas compostos que os tornavam independentes, ou seja, aqueles onde os resultados de qualquer medição em uma parte eram (estatisticamente) independentes de qualquer medição na outra, queremos entender, com a visão mais geral de estados, aqueles que retêm esta propriedade.

Exercício 10.2. *Mostre que um estado ρ^{AB} de um sistema de dois qbits gera resultados independentes para medições locais se, e somente se, for decomponível, i.e.: $\rho^{AB} = \sigma^A \otimes \tau^B$. (Sugestão: os resultados de qualquer medição local na parte A são descritos por $\rho^A = \mathrm{Tr}_B \rho^{AB}$. Pense no estado reduzido de uma das partes condicionado ao resultado de uma medição na outra.)*

Agora a convexidade de $D(E)$ desempenha um papel importante:

Exercício 10.3. *Mostre que $\rho = \frac{1}{2}(|01\rangle\langle 01| + |10\rangle\langle 10|)$ não é decomponível. (Sugestão: novamente, pense no estado de uma parte condicionado ao resultado de uma medição na outra.)*

O estado descrito no exercício 10.3 é um estado quântico, mas as correlações que ele descreve não. Podemos pensar vários sistema clássicos com correlações equivalentes. Se nos restringirmos a medições projetivas na base $\mathcal{Z}$, temos dois bits com a condição de soma 1, distribuídos de maneira equiprovável. Podemos considerar que o segundo bit é o resultado de aplicar a operação NOT ao primeiro[3]. Se fizermos medições em outras bases, ou mesmo medições generalizadas, o resultado será ainda menos correlacionado.

Geometricamente, o exercício 10.3 mostra que o conjunto dos estados produto (aqueles descritos por operadores densidade decomponíveis) não é convexo. O Postulado 9.3 implica que um operador

[3]Uma forma lúdica de descrever essas correlações é pensar no "mundo das figuras de baralho", onde os habitantes não têm pés, mas sim cabeças simétricas. Se um cara ou coroa é disputado, com a moeda caindo sobre o tampo de vidro de uma mesa, situada no plano "equatorial" (aquele que corta a "cintura" das figuras), teremos uma das cabeças vendo o resultado cara, a outra coroa, de forma equiprovável. Claro que outra historinha que pode ser recordada é a das semi-moedas e das faces do semi-dado, apresentada no capítulo 8.

densidade da forma

$$\rho^{AB} = \sum_i p_i \rho_i^A \otimes \rho_i^B \tag{10.1}$$

gera probabilidades conjuntas para medições em A e B que podem ser descritas classicamente. Sendo mais explícito: se $\{M_j\}$ são operadores de medição no sistema A, $\{N_k\}$ são operadores de medição no sistema B, então $\{M_j \otimes N_k\}$ são operadores de medição no sistema composto (verifique!). Se tal processo de medição é aplicado ao estado (10.1), teremos:

$$\begin{aligned}
p(j,k) &= \mathrm{Tr}\Big(M_j \otimes N_k \ \rho^{AB} \ M_j^\dagger \otimes N_k^\dagger\Big) \\
&= \mathrm{Tr}\left(M_j \otimes N_k \sum_i p_i \rho_i^A \otimes \rho_i^B \ M_j^\dagger \otimes N_k^\dagger\right) \\
&= \sum_i p_i \mathrm{Tr}\Big(M_j \otimes N_k \ \rho_i^A \otimes \rho_i^B \ M_j^\dagger \otimes N_k^\dagger\Big) \\
&= \sum_i p_i \mathrm{Tr}\Big(M_j \rho_i^A M_j^\dagger \ \otimes \ N_k \rho_i^B N_k^\dagger\Big) \\
&= \sum_i p_i \mathrm{Tr}\Big(M_j \rho_i^A M_j^\dagger\Big) \mathrm{Tr}\Big(N_k \rho_i^B N_k^\dagger\Big) = \sum_i p_i \, p(j|i) \, p(k|i),
\end{aligned}$$

que é uma maneira clássica[4] de descrever probabilidades conjuntas correlacionadas.

Esta discussão sugere a seguinte:

Definição 10.1. *Estados de dois qubits são classificados em:*

1. *Fatorável, se ρ^{AB} for decomponível;*
2. *Separável, se ρ^{AB} pode ser escrito como na equação* (10.1)*;*
3. *Emaranhado, se ρ^{AB} não pode ser escrito como na equação* (10.1)*.*

[4] Já antecipando que poderemos encontrar estados quânticos e processos de medição tais que nossa descrição clássica cai por terra.

É claro que todo estado fatorável é também separável. Também é verdade que todo estado puro emaranhado pela definição 8.1 também é emaranhado pela definição 10.1, mas isso não é tão evidente. O ponto central é que a definição de emaranhamento é tão boa quanto a de convergência: se conseguimos mostrar que vale certa coisa, sabemos que o estado é separável. Mas mostrar explicitamente que pode ser obtida uma decomposição como na equação (10.1) não é simples. Isso justifica a procura de critérios de separabilidade, que permitam garantir separabilidade (ou emaranhamento – nem sempre um critério é conclusivo em ambas as direções) sem precisar explicitar a forma (10.1) (assim como um critério de convergência garante a convergência, sem necessariamente calcular o limite, muito menos provar sua existência).

Antes, porém, umas palavrinhas sobre o termo *separável*, sem nenhuma relação com seu significado em análise funcional, por exemplo. Aqui o termo é obtido da teoria quântica da informação, onde as partes A e B normalmente são associadas a laboratórios distantes, onde personagens como Ana e Bernardo[5] atuam. Se considerarmos que não há restrição para a preparação de um estado ρ qualquer em um laboratório, a equação (10.1) pode ser interpretada como um algoritmo: por algum processo, a variável aleatória i é realizada, com distribuição p_i. Obtido o resultado i, Ana e Bernardo são comunicados e devem preparar, respectivamente, ρ_i^A e ρ_i^B. Sem o conhecimento da i, a equação (10.1) é a melhor descrição possível para o estado do sistema composto. Como os laboratórios são espacialmente separados, justifica-se a nomenclatura.

10.1.1 Critérios de Separabilidade

Há muitos critérios e não nos cabe ser completos aqui. Vamos apresentar alguns, seja por importância histórica, seja por facilidade de aplicação, ou ainda por nos ensinar algo sobre o conjunto dos estados quânticos.

[5] Em textos de língua inglesa tais personagens são sempre *Alice* e *Bob*.

Transposição Parcial

Uma propriedade simples e importante dos estados separáveis foi percebida por Asher Peres [Per96]. O ponto de partida é que, definida uma base[6], a operação de transpor uma matriz leva um operador densidade em outro. Em símbolos:

$$\begin{array}{rccc} T: & L(E) & \longrightarrow & L(E) \\ & A & \longmapsto & A^t \end{array} \tag{10.2a}$$

é tal que $T(D(E)) \subseteq D(E)$.

Exercício 10.4. *Prove essa afirmação.*

Com isso, caso se faça a transposição apenas em uma das partes de um sistema composto, teremos uma operação definida em $L(E_A \otimes E_B)$ da maneira usual: define-se nos operadores decomponíveis, estendendo por linearidade (veja definição 2.12). Em símbolos:

$$\begin{array}{rccc} T_A: & L(E_A \otimes E_B) & \longrightarrow & L(E_A \otimes E_B) \\ & A \otimes B & \longmapsto & A^t \otimes B \end{array} \tag{10.2b}$$

e

$$\begin{array}{rccc} T_B: & L(E_A \otimes E_B) & \longrightarrow & L(E_A \otimes E_B) \\ & A \otimes B & \longmapsto & A \otimes B^t, \end{array} \tag{10.2c}$$

que são chamadas *transposição parcial*, respectivamente com respeito à primeira ou à segunda parte. Segue da observação anterior que se aplicarmos a transposição parcial a um estado separável, obteremos um novo estado (separável) igualmente válido. Com efeito:

$$T_A\left(\sum_i p_i \rho_i^A \otimes \rho_i^B\right) = \sum_i p_i \left(\rho_i^A\right)^t \otimes \rho_i^B.$$

[6] É importante ser explícito com relação a um fato: não existe uma operação de transposição canônica para operadores. A transposição é feita com respeito a uma base, visto que o que é naturalmente definido é a transposição de matrizes. Ainda assim, as propriedades que usaremos não dependem da escolha da base, por isso em vários pontos vamos nos referir à transposição sem nos preocupar com a base escolhida para identificar operadores e matrizes.

O ponto importante é que não é verdade que $T_A(D(E_A \otimes E_B)) \subseteq D(E_A \otimes E_B)$. Para perceber isso, vamos recorrer ao nosso velho conhecido $|\Psi_-\rangle$. Note que

$$\begin{aligned} T_A|\Psi_-\rangle\langle\Psi_-| &= \frac{1}{2}T_A(|01\rangle\langle 01| - |10\rangle\langle 01| - |01\rangle\langle 10| + |10\rangle\langle 10|) \\ &= \frac{1}{2}(|01\rangle\langle 01| - |00\rangle\langle 11| - |11\rangle\langle 00| + |10\rangle\langle 10|). \end{aligned}$$

Exercício 10.5. *Agora mostre que $T_A|\Psi_-\rangle\langle\Psi_-|$ não é um operador densidade.*

Essa é uma demonstração, por contradição, que $|\Psi_-\rangle\langle\Psi_-|$ é emaranhado. Acabamos de deduzir e aplicar o chamado *critério de Peres*:

Critério 10.1. *Um estado ρ^{AB} tal que $T_A\,\rho^{AB}$ não é positivo semi-definido é, necessariamente, um estado emaranhado.*

Exercício 10.6. *Use o critério de Peres e a decomposição de Schmidt para mostrar que todo estado emaranhado pela definição 8.1 é também emaranhado pela definição 10.1.*

Exercício 10.7. *Mostre que $T_A\,\rho^{AB}$ tem os mesmos autovalores que $T_B\,\rho^{AB}$. Enuncie a propriedade que decorre daí com respeito ao critério de Peres.*

Mapas Positivos

De fato, Peres conjecturou que seu critério fosse não apenas necessário, mas também suficiente para detectar emaranhamento. Veremos adiante que, em geral, esse não é o caso. Mas para dois qbits é! E quem entendeu isso foi a família Horodecki [H$^{\otimes 3}$96], colocando a discussão em termos mais gerais.

Os termos mais gerais em questão são os chamados mapas positivos:

$$\Lambda : L(E) \longrightarrow L(E)$$

tais que para todo $\pi \in L(E)$ positivo (semi-definido), $\Lambda\,\pi$ também é positivo (semi-definido). A transposição (com respeito a alguma base escolhida) é um exemplo de mapa positivo. O fato interessante, talvez não intuitivo, é que extensões triviais de mapas positivos podem não

ser positivos. Ou seja, $\Lambda \otimes I$, que atua em $L(E) \otimes L(F) \cong L(E \otimes F)$, pode não ser positivo, ainda que Λ o seja. A transposição é novamente o exemplo. Um mapa tal que toda extensão trivial é positiva chama-se *completamente positivo*. A família Horodecki generalizou o critério de Peres da seguinte forma:

Critério 10.2. *Um estado ρ^{AB} é emaranhado se, e somente se, existe um mapa positivo, mas não completamente positivo, Λ, tal que $\Lambda \otimes I\, \rho^{AB}$ não é positivo semi-definido.*

O ponto interessante é que já era um resultado conhecido que, se nos restringirmos a extensões triviais de mapas positivos $\Lambda : L\big(\mathbb{C}^2\big) \to L\big(\mathbb{C}^2\big)$ da forma[7] $\Lambda \otimes I : L\big(\mathbb{C}^2 \otimes \mathbb{C}^2\big) \to L\big(\mathbb{C}^2 \otimes \mathbb{C}^2\big)$, a transposição é essencialmente o único mapa positivo e não completamente positivo. Para ser mais preciso, todo mapa positivo $\Lambda : L\big(\mathbb{C}^2\big) \to L\big(\mathbb{C}^2\big)$ pode ser escrito na forma:

$$\Lambda = \Lambda_1 + \Lambda_2 \circ T,$$

onde Λ_1 e Λ_2 são mapas 2-positivos, ou seja, tais que $\Lambda_i \otimes I : L\big(\mathbb{C}^2 \otimes \mathbb{C}^2\big) \to L\big(\mathbb{C}^2 \otimes \mathbb{C}^2\big)$ são positivos. Assim, $\Lambda \otimes I\, \rho$ só pode não ser positivo se $T \otimes I\, \rho$ não for positivo.

Testemunhas de Emaranhamento

Outro fato importante é geométrico. Por construção, o conjunto dos estados separáveis é convexo e fechado. Vamos denotá-lo, em geral, $S(E)$. Assim, qualquer ponto exterior a $S\big(\mathbb{C}^2 \otimes \mathbb{C}^2\big)$ pode ser separado dele por um hiperplano. Aproveitando ainda que $D\big(\mathbb{C}^2 \otimes \mathbb{C}^2\big)$ está contido no hiperplano afim definido por $\mathrm{Tr}\rho = 1$, o hiperplano separador referido acima pode ser dado na forma $w(\sigma) = 0$, onde

$$w : L\big(\mathbb{C}^2 \otimes \mathbb{C}^2\big) \longrightarrow \mathbb{R} \tag{10.3a}$$

é um funcional linear no espaço dos operadores auto-adjuntos de $\mathbb{C}^2 \otimes \mathbb{C}^2$. Agora dualidade e o teorema de representação de Riesz entram em cena, para dizer que tal funcional w pode ser representado

[7] Este resultado também é válido para extensões $\Lambda \otimes I : L\big(\mathbb{C}^2 \otimes \mathbb{C}^3\big) \to L\big(\mathbb{C}^2 \otimes \mathbb{C}^3\big)$, mas apenas. Para qualquer dimensão maior são conhecidos contra-exemplos.

utilizando o produto escalar do espaço em questão e um elemento do mesmo espaço. Ou seja

$$w(\rho) = \mathrm{Tr}(W\rho), \tag{10.3b}$$

para algum $W \in L\big(\mathbb{C}^2 \otimes \mathbb{C}^2\big)$ auto-adjunto.

Com isso, usando $D = D\big(\mathbb{C}^2 \otimes \mathbb{C}^2\big)$ e $S = S\big(\mathbb{C}^2 \otimes \mathbb{C}^2\big)$, podemos enunciar o critério das testemunhas de emaranhamento:

Critério 10.3. *Um estado $\rho^{AB} \in D$ é emaranhado se, e somente se, existe um operador auto-adjunto W tal que* $\mathrm{Tr}(W\rho) < 0$*, enquanto* $\mathrm{Tr}(W\sigma) \geq 0$ *para todo $\sigma \in S$.*

Uma vantagem adicional do critério 10.3 é que, pela equação (9.1), W pode ser visto como uma grandeza mensurável, tornando a detecção do emaranhamento uma tarefa realizável em laboratório [CT].

10.1.2 Quantificadores de Emaranhamento

A quantificação de emaranhamento também é um problema interessante, para o qual há apenas soluções parciais[8].

Entre as abordagens possíveis, algumas dependem da otimização entre protocolos LOCC [BDSW], outras impõem condições que devem ser obedecidas por quantificadores [VPRK, Vid], outras transformam critérios de separabilidade como os vistos em quantificadores, casos que vamos apresentar com algum detalhe. Por fim, mas não menos importantes, há aquelas que buscam inspiração em propriedades geométricas [VT] ou informacionais [VP]. As referências são citadas apenas como um ponto de partida, não sendo adequado tentar ser completo neste tema, aqui.

Negatividade

A ideia de transposição parcial levou a um quantificador chamado negatividade [LK, VW]. Para dois qbits, foi mostrado que um estado pode ter, no máximo, um autovalor negativo [VADM, Ama]. O módulo deste autovalor pode ser tomado como definição desse quantificador.

[8] E muitas soluções parciais, pelo qual não nos cabe discuti-las aqui.

Concorrência

Um outro quantificador nasceu da intenção de tornar o *emaranhamento de formação* [BDSW] uma quantidade diretamente computável. Acabou ganhando "vida própria" e hoje em dia é considerado como um outro quantificador [Woo].

Emaranhamento Testemunhado

Uma grande família de quantificadores nasce quando passamos a otimizar as testemunhas do emaranhamento de um estado, sujeitas a certas restrições [EBA]. Neste caso, o módulo do valor obtido pelo funcional calculado no estado também serve como quantificador. É interessante que vários outros quantificadores previamente definidos por outros caminhos, podem ser incluídos nesta família de quantificadores, dependendo apenas do tipo de restrição que se impõe às possíveis testemunhas.

10.1.3 Geometria

Uma boa forma de ganhar intuição sobre a geometria do conjunto de estados $D\big(\mathbb{C}^2 \otimes \mathbb{C}^2\big)$ é generalizar a noção de vetor de Bloch. Utilizando as matrizes de Pauli (6.4), podemos escrever

$$\rho = \frac{1}{4}\left(I \otimes I + \vec{r}\cdot\vec{\sigma} \otimes I + I \otimes \vec{s}\cdot\vec{\sigma} + \sum_{jk} t_{jk}\, \sigma_j \otimes \sigma_k \right), \qquad (10.4)$$

onde os 15 parâmetros necessários ganham a forma de dois vetores, $\vec{r}$ e $\vec{s}$, e uma matriz $t = [t_{jk}]$, com os índices j, k assumindo os valores x, y, z. Para melhor interpretá-los, devemos lembrar que as três matrizes de Pauli têm traço nulo, que a identidade tem traço 2, e com isso obter

$$\rho^A = \mathrm{Tr}_B \rho = \frac{1}{2}(I + \vec{r}\cdot\vec{\sigma}), \qquad (10.5a)$$

$$\rho^B = \mathrm{Tr}_A \rho = \frac{1}{2}(I + \vec{s}\cdot\vec{\sigma}), \qquad (10.5b)$$

que permite reconhecer que $\vec{r}$ é o vetor de Bloch do sistema A, assim como $\vec{s}$ do sistema B. Se os sistemas forem independentes, ou seja,

se $\rho = \rho^A \otimes \rho^B$, teremos $t_{jk} = r_j s_k$. Qualquer desvio disso indica correlações do sistema.

Há algumas formas canônicas para estes parâmetros. Caso estejamos interessados em entender o emaranhamento do estado, é natural considerar que a ação de unitárias locais não trará efeitos. Mais precisamente, as órbitas da ação

$$\begin{aligned} \Phi : \quad (SU(2) \times SU(2)) \times D\big(\mathbb{C}^2 \otimes \mathbb{C}^2\big) \quad &\longrightarrow D\big(\mathbb{C}^2 \otimes \mathbb{C}^2\big) \\ ((U_A, U_B), \rho) \quad &\longmapsto U_A \otimes U_B \, \rho \, U_A^\dagger \otimes U_B^\dagger \end{aligned}$$

são compostas por estados equivalentes, com respeito ao emaranhamento. Podemos usar esta liberdade para diagonalizar a matriz t e com isso passar a trabalhar com um estado $\tilde{\rho}$ caracterizado por três "vetores": $\vec{r}$, $\vec{s}$ e $\vec{t}$, este último definido pelos elementos da diagonal da matriz t correspondente a um elemento da órbita de ρ que tem a matriz t diagonal.

Exercício 10.8. *Obtenha o efeito da ação Φ sobre os coeficientes $\vec{r}$, $\vec{s}$ e $[t_{jk}]$, de modo a justificar o parágrafo anterior.*

É fácil notar que I, σ_x e σ_z são matrizes simétricas, enquanto σ_y é anti-simétrica. Dessa forma, a transposição troca o sinal da componente y do vetor de Bloch. Da mesma forma, a transposição parcial, digamos no sistema A, troca o sinal da componente y de $\vec{r}$ e de $\vec{t}$. Isso pode ser usado para visualizar algumas propriedades [H$^{\otimes 2}$]. Um caso particularmente bonito e importante envolve estados com $\vec{r} = \vec{s} = \vec{0}$. Por motivos razoavelmente claros, tais estados são conhecidos como estados T. Pelo que já foi discutido, um estado T será um estado produto se, e só se, $\vec{t} = \vec{0}$, caso em que ρ corresponde ao estado maximamente misturado, a órbita de Φ é completamente degenerada e o estado é invariante por qualquer transposição parcial. Mas e para $\vec{t} \neq \vec{0}$, o que podemos afirmar? É o que o exercício a seguir vai trabalhar.

Exercício 10.9. *Esse exercício vai trabalhar com estados da forma*

$$\rho = \frac{1}{4}(I \otimes I + t_x \, \sigma_x \otimes \sigma_x + t_y \, \sigma_y \otimes \sigma_y + t_z \, \sigma_z \otimes \sigma_z). \tag{10.6}$$

1. *Mostre que combinações convexas de estados da forma* (10.6) *também são da mesma forma e descreva o que acontece com o vetor $\vec{t}$;*

2. *Mostre que os quatro estados de Bell* (8.1) *são estados T, correspondendo a diferentes vetores $\vec{t}$;*

3. *Que região em $\mathbb{R}^3$ corresponde a todas as combinações convexas dos estados de Bell? Vamos denotá-la por $\mathcal{T}$;*

4. *Mostre que $\mathcal{T}$ corresponde a todos os estados T;*

5. *O que acontece com $\mathcal{T}$ quando fazemos a transposição parcial na primeira parte? Vamos denotá-la $T_A\mathcal{T}$;*

6. *Qual o significado de $\mathcal{T} \cap T_A\mathcal{T}$? E qual região de $\mathbb{R}^3$ ela representa?*

Para mais detalhes o leitor pode consultar [Ama]. Para uma abordagem distinta à mesma questão, pode consultar o capítulo 4 de [Ara].

Exercício 10.10. *Use os estados T do exercício 10.9 para mostrar que a transformação $\Lambda : \frac{1}{2}\Big(I + \vec{b}\cdot\vec{\sigma}\Big) \mapsto \frac{1}{2}\Big(I - \vec{b}\cdot\vec{\sigma}\Big)$ não é uma evolução quântica permitida. (Sugestão: reveja a discussão sobre evoluções quânticas do capítulo 9.)*

Voltando ao conjunto $D\big(\mathbb{C}^2 \otimes \mathbb{C}^2\big)$ de todos os estados de dois qbits, há um outro resultado bastante importante, por nos permitir formar uma imagem mais adequada deste. Na referência [ŻHLS] é mostrado que existe uma bola fechada centrada no estado mais misturado toda formada de estados separáveis. A consequência importante disso é que o conjunto os separáveis, $S\big(\mathbb{C}^2 \otimes \mathbb{C}^2\big)$ possui a mesma dimensão que $D\big(\mathbb{C}^2 \otimes \mathbb{C}^2\big)$, e tem um volume que é uma fração positiva do volume de $D\big(\mathbb{C}^2 \otimes \mathbb{C}^2\big)$. Este resultado sobre o volume dos estados separáveis, aliado à visão geométrica do conjunto de estados, permite entender como natural e esperado um fenômeno tido algumas vezes como surpreendente: a morte do emaranhamento em tempo finito. Nos casos em que a dinâmica possui um atrator no interior do conjunto dos estados separáveis, o destino de qualquer emaranhamento é morrer em tempo finito. A história pode ser diferente se o atrator tocar a fronteira do conjunto dos separáveis e será diferente se tal atrator for composto apenas de estados emaranhados [Ter07b].

Por outro lado, ainda que caiba a descrição de um conjunto convexo, S, contido em outro conjunto convexo[9], D, sabemos que suas

[9]Como a gema dentro de um ovo.

fronteiras não são completamente regulares. Não há, todavia, uma descrição completa delas. Sabemos dizer que são subvariedades diferenciáveis por partes, ou seja, que tais fronteiras são uniões de subvariedades diferenciáveis com bordo, coladas de maneira menos regular. Um efeito interessante desta irregularidade foi descrito e observado na referência [CSC+].

10.2 Sistemas Bipartites

A maior parte do que falamos para dois qbits vale para dois sistemas de dimensão finita. Para um espaço de estados $\mathbb{C}^m \otimes \mathbb{C}^n$, $D(\mathbb{C}^m \otimes \mathbb{C}^n)$ será um conjunto convexo, compacto de dimensão real $m^2n^2 - 1$. As definições de separabilidade e emaranhamento são rigorosamente as mesmas já apresentadas.

Exercício 10.11. *Releia a secção 10.1 com a preocupação de identificar quais resultados dependem de serem dois qbits e quais se generalizam diretamente, fazendo a generalização onde adequado.*

O critério da transposição parcial, conforme enunciado, continua válido; o que não vale, exceto se $m = 2, n = 3$, é sua recíproca, e essa é a grande novidade quando passamos a sistemas bipartidos em dimensão maior. Já são conhecidos exemplos, tanto para $m = 2, n = 4$, quanto para dois *qtrits*, *i.e.*: $m = n = 3$, de estados emaranhados cuja transposta parcial também é um estado possível [H$^{\otimes 3}$98]. Estes estados são chamados *PPT-emaranhados*, da sigla, em inglês, para Transposta Parcial Positiva. Isso dá origem a um interessante problema em aberto na área. Autovetores associados a autovalores negativos da transposta parcial dão origem tanto a testemunhas de emaranhamento, quanto a estratégias para *destilar* tal emaranhamento: ou seja, uma maneira de atuar conjuntamente (mas de maneira local: LOCC é o paradigma adotado) sobre vários representantes deste estado e obter alguma outra quantidade de pares de Bell (pelo menos de maneira aproximada - a definição precisa envolve o limite assintótico). Na sua ausência, não há receita para destilar emaranhamento e a conjectura é a equivalência entre emaranhamento PPT e emaranhamento que não pode ser destilado (o chamado *emaranhamento preso*, do inglês *bound entanglement*).

Quantificadores

A discussão geral de quantificadores fica mais rica, mas a maioria das ideias usadas para dois qbits encontra contra-partida em sistemas bipartites de dimensão finita.

Em especial, dos quantificadores citados na 10.1.2, somente o emaranhamento testemunhado já foi feito de maneira bastante geral.

A negatividade pode ser redefinida[10] como a soma dos módulos dos autovalores negativos da transposta parcial de ρ. Pela discussão anterior, fica claro que existem estados emaranhados com negatividade zero, violando uma das exigências para ser um (bom) quantificador de emaranhamento (ser zero para todo estado separável, e apenas para eles). Ainda assim, a negatividade quantifica alguma coisa, relacionada ao emaranhamento (possivelmente associada ao *emaranhamento destilável*).

Já a concorrência, depois de ganhar status de quantificador por si só, também ganhou generalizações para sistemas maiores.

Geometria

A geometria dos conjuntos $D(\mathbb{C}^m \otimes \mathbb{C}^n)$ e $S(\mathbb{C}^m \otimes \mathbb{C}^n)$, literalmente, ganha mais espaço. Não há uma visão pictórica tão agradável quanto os vetores de Bloch ou os estados T, mas continua válida a noção que, se estamos preocupados em entender o emaranhamento, devemos nos concentrar nas órbitas da ação (veja, por exemplo, [SHK])

$$\begin{aligned} \Phi: \quad (SU(m) \times SU(n)) \times D(\mathbb{C}^m \otimes \mathbb{C}^n) \quad &\longrightarrow D(\mathbb{C}^m \otimes \mathbb{C}^n) \\ ((U_A, U_B), \rho) \quad &\longmapsto U_A \otimes U_B \, \rho \, U_A^\dagger \otimes U_B^\dagger. \end{aligned}$$

Com relação ao volume dos estados separáveis, também segue verdadeiro o resultado que existe uma bola de separáveis centrada no estado maximamente misto [Ż]. Por sua vez, a razão entre o volume de tal esfera e o volume de todos os estados decresce fortemente com a dimensão.

Para muito mais sobre geometria de estados quânticos, recomendamos a referência [BŻ].

[10] Há alguma discordância, bem justificada, sobre um fator de escala na definição da negatividade. Portanto, ao utilizá-la ou encontrá-la em um texto, é bom verificar a definição adotada.

TPS

A discussão sobre diferentes estruturas de produto tensorial, iniciada na secção 8.4, também encontra eco aqui. Já vimos que, para vetores de estado (*i.e.*: estados puros), sempre existem TPS tais que um dado estado é fatorável e outras em que ele é emaranhado. Será que isso se repete para operadores densidade?

É fácil concluir que a resposta é não. Basta considerarmos o estado maximamente misturado, que será separável para qualquer TPS. Em seguida, podemos usar o resultado que para qualquer TPS existe uma bola de estados separáveis centrada na máxima mistura, para um argumento de continuidade e compacidade[11] permitir concluir que há um raio mínimo. Ou seja: existe um conjunto com medida positiva de estados que são separáveis com respeito a qualquer estrutura de produto tensorial. Podemos chamá-los de *absolutamente separáveis*.

10.3 Sistemas Multipartites

Quando o número de partes aumenta temos ainda mais espaço para encontrar estruturas interessantes. Consequentemente, temos mais problemas e menos se conhece sobre suas respostas. Vamos explorar apenas a ponta de um *iceberg*, para dar o gosto do problema.

Deixando de lado a discussão (interessante) sobre diferentes estruturas de produto tensorial, vamos considerar um espaço de estados já decomposto em N fatores (*i.e.*: partes):

$$E = \bigotimes_{i=1}^{N} E_i, \tag{10.7}$$

onde cada E_i é um espaço de estados de dimensão finita. Com respeito a esta N-partição, é natural definirmos como estados produto aqueles da forma

$$\rho = \bigotimes_{i=1}^{N} \rho^i, \tag{10.8}$$

onde $\rho^i \in D(E_i)$. Para estes estados, medições em partes distintas serão estatisticamente independentes. O passo seguinte é definir os

[11] Dimensão finita é importante aqui.

estados separáveis como combinações convexas de estados produto e chamarmos de emaranhados aqueles que estão no complemento deste conjunto.

Embora faça sentido, tal estratégia tem um grande inconveniente. Por exemplo, se temos três partes, usualmente denominadas A, B e C, somos levados a dizer que um estado da forma $\rho^A \otimes \rho^{BC}$, onde ρ^{BC} é um estado emaranhado de duas partes, é também um estado emaranhado (visto que não é separável). Claramente o ponto é que ainda temos várias partições para o conjunto $\{1, 2, \ldots, N\}$, correspondendo a "desfazer" separações entre certas partes, ou seja, considerá-las conjuntamente. O exemplo específico trata de um estado separável (até fatorável) quando consideramos a partição $\{\{A\}, \{B, C\}\}$.

Para contornar esse inconveniente basta lembrarmos que separabilidade (consequentemente emaranhamento) é sempre definida com respeito a uma partição dada[12]. Dessa forma, a equação (10.7) determina a partição mais fina que estamos dispostos a considerar, ou seja,

$$\{1, 2, \ldots, N\} = \bigcup_{i=1}^{N} \{i\},$$

mas outras partições mais grossas que esta são permitidas. Dada uma partição $\mathcal{P}$ de $\{1, 2, \ldots, N\}$, com P_i denotando os conjuntos da partição e $q = q(\mathcal{P})$ a quantidade de conjuntos da partição (naturalmente $1 \leq q \leq N$), vamos definir estados $\mathcal{P}$-produto como aqueles da forma

$$\rho = \bigotimes_{i=1}^{q} \rho^{P_i}, \tag{10.9}$$

onde $\rho^{P_i} \in D\left(\bigotimes_{j \in P_i} E_j\right)$, ou seja, ρ^{P_i} é um estado conjunto das partes relacionadas em P_i. Da mesma forma que antes, as combinações convexas dos estados $\mathcal{P}$-produto serão ditas $\mathcal{P}$-separáveis, enquanto estados que não são $\mathcal{P}$-separáveis são chamados $\mathcal{P}$-emaranhados.

[12] De maneira mais geral, com respeito a estrutura de produto tensorial considerada, mas deixemos isso de lado.

Exercício 10.12. *Mostre que se $\mathcal{R}$ é um refinamento de $\mathcal{P}$, todo estado $\mathcal{R}$-separável é também $\mathcal{P}$-separável. Enuncie e compreenda a contrapositiva dessa afirmação.*

Exercício 10.13. *Se $\mathcal{P}_1$ e $\mathcal{P}_2$ são duas partições de $\{1, 2, \ldots, N\}$, o que podemos afirmar sobre um estado que é $\mathcal{P}_1$- e $\mathcal{P}_2$-produto? E $\mathcal{P}_1$- e $\mathcal{P}_2$-separável? E $\mathcal{P}_1$- e $\mathcal{P}_2$-emaranhado?*

O exercício 10.12 define uma hierarquia (não-completa) de emaranhamentos, a partir das possíveis partições. Podemos ainda definir uma nova estratificação a partir do número q de conjuntos da partição. Seja Υ_q o conjunto de todas as partições de $\{1, 2, \ldots, N\}$ em q conjuntos. Um estado será dito q-separável se puder ser escrito como

$$\rho = \sum_{\mathcal{P}\in\Upsilon_q} p_{\mathcal{P}} \rho^{\mathcal{P}}, \tag{10.10}$$

onde $\rho^{\mathcal{P}} \in S\Big(\bigotimes_{j\in\mathcal{P}} E_j\Big)$, ou seja $\rho^{\mathcal{P}}$ é um estado $\mathcal{P}$-separável e $(p_{\mathcal{P}})$ é um vetor de probabilidades.

Exercício 10.14. *Mostre que para todo $p \in (0, 1)$ o estado*

$$p|0\rangle\langle 0| \otimes |\Psi_-\rangle\langle\Psi_-| + (1-p)|\Psi_-\rangle\langle\Psi_-| \otimes |0\rangle\langle 0|$$

é 2-separável, sem ser $\mathcal{P}$-separável para nenhuma 2-partição do conjunto $\{A, B, C\}$.

Exercício 10.15. *Mostre que se $r \geq q$, todo estado r-separável é também q-separável. Em particular, verifique que todo estado é 1-separável e que o primeiro conceito de separabilidade apresentada corresponde a N-separabilidade.*

Nesta direção que estamos indo, poderíamos discutir vários tópicos interessantes, mas que serão deixados para uma outra oportunidade. Poderíamos usar o conceito de *diâmetro de uma partição*, correspondendo à cardinalidade do maior conjunto, e definir separabilidade com respeito a partições de diâmetro máximo dado. Além disso, poderíamos investir em dois conceitos interessantes: estados reduzidos (aqueles que obtemos quando ignoramos alguma parte) e estados condicionais (aqueles que consideramos depois de realizar algum teste, neste caso em uma das partes, e obter algum resultado).

Também há bastante interesse em considerar uma situação do tipo grafo em estrela, onde uma parte especial tem contato com várias partes "similares" e descrever como essas partes similares restringem propriedades do estado da parte especial...

10.4 Um tantinho mais de Física

Os conceitos apresentados até aqui também trazem consequências e interpretações muito interessantes. Uma compreensão errônea das correlações de um par EPR[13], por exemplo, faz com que se pense que é possível "mandar mensagens" diretamente por essas correlações, afinal "quando Ana mede na base $\mathcal{Z}$ e obtém 1, ela sabe que Bernardo obterá 0 caso meça na mesma base $\mathcal{Z}$." Vamos explicar por que, embora não haja nada de errado nessa frase, ela não permite concluir pela utilidade para comunicação deste "conhecimento" de Ana.

A afirmação em questão é condicional: "quando Ana mede na base $\mathcal{Z}$ e obtém 1...". Ana não tem como escolher o resultado de sua medição. Você pode fazer a objeção: "mas mecânica quântica não trata do conceito de preparação de estado, em geral fazendo um teste e descartando as alternativas indesejáveis?" Sim, novamente uma frase correta. Mas que supõe que quem atua sobre o sistema e descarta os casos indesejáveis tem acesso ao sistema todo. Para ser aplicada a este caso, Ana deveria dizer a Bernardo se deve manter a sua parte do par, ou descartá-la. E, para isso, ela precisou usar comunicação. Mais precisamente, precisou enviar um bit de informação (descartar ou manter), para que o par seja capaz de "comunicar" um bit: o resultado da possível medição de Bernardo na base $\mathcal{Z}$.

Esta discussão pode ser refraseada em termos de estados reduzidos e estados condicionais. Para um estado de Bell (8.1), os estados reduzidos são sempre maximamente misturados. Ou seja, qualquer teste que Ana ou Bernardo decidam fazer possui resultados equiprováveis. O ponto interessante e importante é que, ainda que localmente equiprováveis, esses resultados estão muito longe de serem independentes, já que os estados condicionais são sempre puros: feito um teste local, digamos por Ana, o estado condicional do sistema é puro e fatorável;

[13] Se preferir, chame de par de Bell. Mas aqui é justo fazer uma homenagem a quem mais se incomodou com o que parecia *spooky action at a distance*.

existe um teste local de Bernardo com resultado certo. Ela, de fato, sabe o resultado que Bernardo obterá, caso faça o referido teste. Mas isso não é mais que a discussão da semi-moeda feita no capítulo 8.

Poslúdio

Passado o principal e já nos aproximando do fim do curso, nos propomos agora a dar um rápido passeio por temas, em algum sentido, mais avançados.

A mecânica quântica na reta é mais avançada por exigir espaços vetoriais de dimensão infinita. Por outro lado, não conseguiríamos reagir a um crítico que reclamasse de um livro sobre mecânica quântica que não tratasse do problema fisicamente mais básico: quantização de uma partícula sujeita a um potencial, incluído aí o onipresente oscilador harmônico.

A versão quântica dos sistemas de funções iteradas é avançado por ser assunto de pesquisa recente, com o mérito adicional de ter tornado dois dos autores deste livro co-autores[14]. Também não poderíamos evitar as críticas e acusações de ingratidão se não incluíssemos tal assunto na nova etapa desta parceria.

A questão de bem entender em que a mecânica quântica difere do pensamento clássico é avançada em vários sentidos. Aqui, mais uma vez, só conseguiremos tocar a ponta de um iceberg. Ainda assim, será possível apresentar algumas demonstrações de como falham algumas hipóteses aparentemente naturais.

A sensação é que o curso e o livro já estão perto do fim, mas as notas finais devem convidar o estudante a seguir buscando conhecimento.

[14] E, de certa forma, serem parte da origem deste livro.

Capítulo 11

Um Pouco de Mecânica Quântica na Reta

Neste capítulo falaremos sobre a mecânica quântica num intervalo da reta, ou na própria reta. Usaremos agora um espaço de estados que é mais sofisticado que os já descritos até aqui, por isso pedimos licença para uma certa informalidade e ainda alguma confiança do leitor pois a justificativa de algumas passagens é mais sofisticada e será omitida. Acreditamos, no entanto, que a intuição obtida com os exemplos estudados até agora será suficiente para tornar ao menos palatáveis os resultados que serão expostos.

11.1 Partícula Clássica na Reta

A melhor descrição clássica feita pela mecânica de uma partícula na reta envolve basicamente o conhecimento, em cada instante, de duas coisas: sua posição, representada por um ponto na reta, e sua velocidade. A posição será representada pela variável x e o momentum da partícula (que, também nos casos mais simples, vem a ser o produto de sua massa pela velocidade) é representado por p.

Quando a partícula está sujeita à ação de um campo de forças, $F\colon \mathbb{R} \to \mathbb{R}$, podemos descrever seu movimento por meio da lei de

Newton:

$$F = \frac{dp}{dt} = m\frac{d^2}{dt^2}x = ma, \tag{11.1}$$

com condições iniciais $x(0) = x_0$ e $v(0) = \frac{d}{dt}x(0) = v_0$.

Um campo de forças pode ser convenientemente representado por um potencial, uma função $V\colon \mathbb{R} \to \mathbb{R}$ tal que

$$-\frac{d}{dx}V(x) = F(x).$$

Não é difícil ver que podemos obter uma função V satisfazendo essa propriedade se definirmos

$$V(x) = -\int_0^x F(s)ds.$$

O leitor pode então se perguntar o porque da escolha do ponto 0 como extremo inferior da integral e a resposta é que isso é apenas uma convenção; se 0 for trocado por qualquer outro ponto será obtida uma nova função V que continua satisfazendo a condição acima. De fato a diferença entre essas funções será uma constante (pois ambas têm a mesma derivada).

Com essa função potencial podemos reescrever a equação de Newton numa versão conhecida como mecânica hamiltoniana, que consiste essencialmente em se definir uma função (a função de Hamilton)

$$H(p, x) = \frac{p^2}{2m} + V(x) \tag{11.2}$$

e tomar como equações de movimento as equações de Hamilton

$$\begin{cases} \dfrac{d}{dt}x &= \dfrac{\partial}{\partial p}H, \\ \dfrac{d}{dt}p &= -\dfrac{\partial}{\partial x}H. \end{cases} \tag{11.3}$$

O leitor não terá dificuldades em ver que o sistema acima equivale à lei de Newton. A mudança essencial é de interpretação. Enquanto na versão newtoniana buscamos a função $x(t)$ utilizando uma EDO de segunda ordem, na mecânica hamiltoniana queremos entender o par $(x(t), p(t))$, governado por uma equação de primeira ordem; $\left(\frac{\partial H}{\partial p}, -\frac{\partial H}{\partial x}\right)$ é um campo vetorial no chamado *espaço de fase* do sistema.

11.2 Partícula Quântica

Devemos agora procurar descrever uma versão quântica do problema. Para isso a primeira coisa a se fazer é identificar qual o espaço de estados adequado para isso.

Precisamos descrever uma partícula na reta e temos um espaço vetorial naturalmente associado a ela que é o espaço $L^2(\mathbb{R})$. O produto interno é definido como sendo

$$\langle f|g\rangle = \int_{\mathbb{R}} \overline{f(x)} g(x) dx.$$

Este parece ser um espaço bastante conveniente para representar a posição da partícula.

Precisamos então compreender como representar os operadores de posição e momentum, pois essas são as quantidades básicas que desejamos obter em medições na mecânica. Seguindo o procedimento usado até aqui, esses devem ser operadores auto-adjuntos em L^2. Para a posição, o operador natural é considerar

$$\hat{x} := x, \tag{11.4}$$

cuja ação sobre uma função é a seguinte: $\hat{x}f := xf(x)$.

O operador momentum, por sua vez, é dado por

$$\hat{p} := -i\frac{d}{dx}, \tag{11.5}$$

ou seja, $\hat{p}f = -i\frac{df}{dx}(x)$.

Podemos verificar a comutatividade (ou não) dos operadores $\hat{x}$ e $\hat{p}$; para isso usaremos uma função auxiliar ϕ (que assumimos diferenciável):

$$[\hat{x},\hat{p}]\phi = (\hat{x}\hat{p} - \hat{p}\hat{x})\phi =$$

$$\hat{x}\left(-i\frac{d}{dx}\phi\right) + i\frac{d}{dx}(x\phi) = -ix\frac{d}{dx}\phi + i\phi + ix\frac{d}{dx}\phi = i\phi,$$

ou seja,

$$[\hat{x},\hat{p}] = i; \tag{11.6}$$

sendo assim os operadores de posição e momentum *não comutam* e para esse par vale também a relação de incerteza do teorema 7.1

(é bom lembrar, no contexto de espaços de dimensão finita; mas a generalização pode ser feita sem problemas com o uso de algumas ferramentas mais avançadas):

$$\mathrm{Var}(\hat{x})\,\mathrm{Var}(\hat{p}) \geq \frac{1}{4}|\langle\psi|[\hat{x},\hat{p}]\psi\rangle|^2 = \frac{1}{4}|\langle\psi|\psi\rangle|^2 = \frac{1}{4}, \tag{11.7}$$

pois assumimos que ψ é um vetor normalizado. Essa é a conhecida relação de incerteza momentum-posição que foi originalmente encontrada por Heisenberg. Em particular, como o comutador é proporcional à identidade, a relação de incerteza é a mesma para todo estado , o que significa que, não podemos ter a dispersão das medidas de posição e a dispersão das medidas de momentum arbitrariamente pequenas.

11.3 O Operador Hamiltoniano e a Equação de Schrödinger

O operador hamiltoniano é uma versão operatorial da função hamiltoniana que foi mostrada no início. Usando a notação $\hat{H}$ vamos defini-lo como

$$\hat{H} = \frac{1}{2m}(\hat{p})^2 + V(\hat{x}) = -\frac{1}{2m}\frac{d^2}{dx^2} + V(x) \tag{11.8}$$

ou seja, temos um operador diferencial.

A equação de Schrödinger, que descreve a evolução temporal de um estado, é dada por

$$\hat{H}\Psi = i\frac{d}{dt}\Psi, \tag{11.9}$$

onde Ψ é visto como vetor no espaço de estados apropriado. Ao lembrarmos que este é um espaço de funções na variável x, teremos de fato uma equação diferencial parcial

$$-\frac{1}{2m}\frac{\partial^2}{\partial x^2}\Psi + V\Psi = i\frac{\partial}{\partial t}\Psi. \tag{11.10}$$

Para resolvermos equações como esta (note que o operador $\hat{H}$ é linear) um método bastante empregado é o da separação de variáveis,

que é usado para se obter um candidato a solução (e depois é preciso usar algumas técnicas um pouco mais cuidadosas para verificar que o candidato a solução é de fato uma solução da equação em questão). A separação de variáveis consiste em se procurar soluções da equação na forma de produto de funções de apenas uma variável, ou seja, $\Psi(x,t) = \psi(x)T(t)$. Depois, usando-se a linearidade, pode-se combinar estas soluções para então tentar produzir a solução do problema original, que inclui condições adicionais.

Usando a hipótese de que $\Psi = \psi T$ na equação de Schrödinger, temos

$$-\frac{1}{2m}\frac{d^2}{dx^2}\psi T + V(x)\psi T = i\psi\frac{d}{dt}T.$$

Dividindo ambos os lados por ψT obtemos

$$\frac{-\frac{1}{2m}\frac{d^2}{dx^2}\psi + V(x)\psi}{\psi} = \frac{i\frac{d}{dt}T}{T}.$$

Note que o lado esquerdo depende apenas da variável x e o lado direito apenas da variável t. A única situação em que estas duas funções de variáveis distintas podem ser iguais é se ambas são constantes e a constante, obviamente, é a mesma; esta costuma ser chamada de constante de separação e será denotada por E. Desta maneira obtemos duas equações diferenciais ordinárias lineares:

$$-\frac{1}{2m}\frac{d^2}{dx^2}\psi + V\psi = E\psi, \tag{11.11a}$$

$$\frac{d}{dt}T = -iET. \tag{11.11b}$$

A primeira equação é conhecida como equação de Schrödinger independente do tempo, e sua solução pode ser mais ou menos difícil dependendo do potencial $V(x)$ que se utiliza. A segunda equação tem uma solução simples, a função

$$T(t) = e^{-iEt}.$$

Isso mostra que para potenciais independentes do tempo, a dificuldade de se encontrar soluções está concentrada na obtenção de $\psi(x)$ pois a parte temporal tem uma solução simples. No que segue

abordaremos alguns casos simples onde é possível obter ψ de maneira explícita.

A constante de separação E também merece algumas palavras. De fato ela corresponde aos autovalores do operador $\hat{H}$, pois satisfaz

$$\hat{H}\psi = E\psi. \tag{11.12}$$

A função de Hamilton na mecânica clássica é uma constante de movimento associada a energia mecânica do sistema; os autovalores de $\hat{H}$ na mecânica quântica correspondem à energia do sistema quântico em questão. Muitas vezes (como no exemplo que daremos a seguir), esses autovalores formam um conjunto discreto e portanto a energia não pode assumir um contínuo de valores, como habitualmente acontece no caso clássico, mas apenas um conjunto discreto, sendo então quantizada. Esse é um dos aspectos chave da teoria quântica[1]. Quanto à função ψ (que é um vetor, um elemento de um espaço de estados conveniente), sua interpretação é a seguinte: a probabilidade de encontrar uma partícula descrita pelo estado ψ (atenção, estamos usando a palavra estado novamente no sentido de vetor!) num intervalo I da reta é dada por

$$\mathbb{P}(x \in I) = \int_I |\psi(s)|^2 ds.$$

11.4 A Partícula em uma Caixa Unidimensional

11.4.1 Caso Clássico

Queremos obter o comportamento de uma partícula livre que se movimenta dentro de uma caixa unidimensional; ou seja, sua posição é representada como sendo um número real no intervalo $[0, L]$, onde L é o comprimento da caixa. Se a partícula é livre então ela se move sem a influência de uma força exterior dentro da caixa e assim sua velocidade é constante pois pela lei de Newton

$$ma = F = 0.$$

[1] E origem do seu nome.

Logo a aceleração (que é a variação de velocidade) é nula. Porém, quando a partícula colide com as paredes da caixa (situadas em $x = 0$ e $x = L$), ela sofre a ação de uma força que tende a fazê-la continuar dentro da caixa. Estamos assumindo que essa colisão é perfeitamente elástica e que a parede é um objeto sólido com massa infinitamente maior que a da partícula: nesse caso o efeito da colisão é o de simplesmente trocar o sentido do movimento, fazendo com que a velocidade da partícula troque de v para $-v$ logo após a colisão. Em resumo, temos um movimento no qual a partícula tem velocidade com módulo constante, mas com o sinal (isto é, o sentido do movimento) que troca a cada colisão, o que não deve surpreender o leitor.

Podemos agora fazer uma pergunta mais divertida: se fixamos um intervalo qualquer $[a, b]$ dentro da caixa, qual é a fração de tempo, em média, gasta pela nossa partícula dentro desse intervalo? Formulando a questão de maneira mais precisa: fixando um instante $T > 0$, qual a parcela de tempo entre 0 e T na qual a partícula esteve em $[a, b]$, ou seja, qual o comprimento do conjunto

$$\{t \in [0, T] : x(t) \in [a, b]\}?$$

Como a velocidade é constante (em intensidade), esse tempo de fato é proporcional ao comprimento do intervalo e será então $|b-a|/L$ (pois dessa forma a fração de tempo de ficar em $[0, L]$ será exatamente 1, como poderíamos esperar). Podemos interpretar essa razão de forma probabilista: esse número é a probabilidade de se observar esse sistema clássico e encontrar a partícula no intervalo $[a, b]$.

11.4.2 Caso Quântico

Agora devemos fazer uma descrição quântica do sistema e, dada a sua relativa simplicidade, investigar se há alguma relação facilmente visível entre o clássico e o quântico.

Começamos por encontrar o espaço de estados adequado; como o nome do capítulo indica, pensamos inicialmente na reta. Mas o que significa o fato da partícula estar na caixa? Como comparamos com as paredes clássicas discutidas anteriormente? Significa que queremos probabilidade zero de encontrar a partícula fora da caixa e isso é consistente com pensar que, no intervalo $[0, L]$ temos um potencial constante, enquanto fora desse intervalo temos um outro valor, muito

maior, com a diferença entre esses valores bem maior que qualquer parâmetro[2] de interesse no problema. É o que os físicos resumem por "infinito". Com isso, vamos trabalhar com funções de $L^2(\mathbb{R} \to \mathbb{C})$ que se anulam fora de $[0, L]$. Mais ainda, é natural pedirmos que essas funções tenham certas regularidades (se V fosse $\mathcal{C}^\infty$ exigiríamos Ψ de classe $\mathcal{C}^2$, mas como V não é sequer contínua, exigimos apenas $\Psi \in \mathcal{C}^0$). Com isso, é razoável considerarmos como espaço de estados para o problema da partícula na caixa o subespaço de $L^2([0, L])$ composto pelas funções duas vezes diferenciáveis em $(0, L)$ e que se anulam na fronteira. Como a partícula é livre, o potencial é nulo e estamos usando a equação de Schrödinger independente do tempo

$$-\frac{d^2}{dx^2}\psi(x) = E\psi(x)$$

(onde, por simplicidade, assumimos que a massa $m = 1/2$) com condições de fronteira $\psi(0) = \psi(L) = 0$. Devemos então resolver esse problema de autovalores.

Nesse caso a solução não é difícil: as funções $\psi_1 = \cos(\sqrt{E}x)$ e $\psi_2 = \operatorname{sen}(\sqrt{E}x)$ claramente satisfazem a equação acima; como esta é linear então as combinações lineares de ψ_1 e ψ_2 também são soluções, o que nos dá a forma geral de uma solução como sendo

$$\psi(x) = A_1\psi_1 + A_2\psi_2,$$

com constantes A_1 e A_2 que devem ser encontradas de forma que a condição de fronteira seja satisfeita (e também a condição de normalização, uma vez que o significado dessa função é expresso em termo de probabilidades):

$$0 = \psi(0) = A_1$$

e

$$0 = \psi(L) = A_2 \operatorname{sen}\sqrt{E}L.$$

Desta forma notamos que os valores possíveis para E, ou seja, os autovalores, obedecem $\operatorname{sen}\Big(\sqrt{E}L\Big) = 0$; rotulando-os por $n = 1, 2, \ldots$, temos

$$E_n = n^2(\pi/L)^2.$$

[2]Comparável.

Observação 7. *O leitor consegue imaginar uma boa razão para não incluirmos $n = 0$ nas soluções acima? Afinal, se $E = 0$ a equação é claramente satisfeita... Bem, note que para $E = 0$ a autofunção correspondente é $\psi(x) = 0$, a função nula. E esta função, multiplicada por uma constante, continua sendo nula. Desta maneira temos um vetor nulo, que não vai nos ajudar a gerar nenhuma solução interessante e por isso não o incluímos na lista de soluções.*

Estes são os autovalores do operador (e há uma infinidade deles, ao contrário do que se passava até aqui, quando considerávamos apenas espaços vetoriais de dimensão finita); as autofunções correspondentes (já normalizadas) são, respectivamente,

$$\psi_n(x) = \sqrt{\frac{2}{L}}\,\mathrm{sen}\frac{n\pi x}{L}, \qquad n = 1, 2, \ldots$$

Um estado em geral é, então, dado por uma série (que é uma espécie de combinação linear, mas com infinitas parcelas) que envolve as autofunções encontradas acima:

$$\psi = \sum_{n=1}^{\infty} c_n \psi_n.$$

(Naturalmente deve-se pensar no problema da convergência num caso deste tipo, mas preferimos deixar esta questão de lado neste texto).

11.4.3 Um Exemplo de Limite Clássico

Vamos usar o exemplo da partícula na caixa para tentar entender como a mecânica quântica se relaciona com a mecânica clássica. Mais uma vez, ficaremos apenas com um exemplo bastante simples, que é o seguinte: queremos compreender como a probabilidade de se encontrar uma partícula num certo intervalo $[a, b]$ varia quando consideramos as autofunções da subseção anterior para energias cada vez maiores, ou seja, no limite quando o número n tende a infinito. Se temos uma partícula no estado ψ_n, a probabilidade de encontrá-la no intervalo $[a, b]$ é dada por

$$\mathbb{P}_n(x \in [a, b]) = \int_{[a,b]} |\psi_n(s)|^2 ds = \frac{2}{L}\int_{[a,b]} \mathrm{sen}^2\frac{n\pi s}{L} ds =$$

$$\frac{2}{L}\left[\int_{[a,b]}\left(\frac{1}{2}-\frac{1}{2}\cos\left(2\frac{n\pi s}{L}\right)\right)\right]ds = \frac{|b-a|}{L} - \frac{1}{2\pi n}\,\mathrm{sen}\left(2\frac{n\pi s}{L}\right)\Big|_a^b.$$

Quando tomamos o limite de $n \to \infty$ o segundo termo tende a zero (pois o seno é uma função limitada) e portanto

$$\mathbb{P}_n(x \in [a,b]) \to \frac{|b-a|}{L},$$

ou seja, para os estados descritos por números n elevados (que correspondem fisicamente a situações de energias bem elevadas) obtem-se que a probabilidade de encontrar uma partícula no intervalo $[a,b]$ está cada vez mais próxima da probabilidade que já havíamos calculado no caso clássico.

Exercício 11.1. *Refaça a discussão acima com n fixo e $L \to \infty$.*

11.5 O Oscilador Harmônico

Vamos agora considerar um exemplo bastante interessante de sistema quântico, o *oscilador harmônico*. Trata-se de uma partícula que se move na reta e está ligada à origem por uma força do tipo $F = -kx$. Ou seja, quando a partícula está na região de x positivo a força é negativa e quando a posição é negativa temos uma força positiva. Dessa forma a força sempre tende a levar a partícula de volta à origem. O exemplo típico da mecânica é uma mola, que sempre tende a restaurar o equilíbrio[3].

Para descrevermos essa situação no contexto de uma partícula quântica devemos começar por encontrar o potencial que corresponde à força acima:

$$V(x) = -\int_0^x F(s)ds = \frac{kx^2}{2}.$$

[3]Importante entender a onipresença de osciladores harmônicos em física: se descrevermos qualquer sistema mecânico por um potencial, da mesma forma que estamos fazendo aqui em dimensão 1, seus mínimos (classicamente) serão pontos de equilíbrio estáveis. Genericamente, podemos aproximar tais mínimos por funções quadráticas a partir deste ponto, a chamada *aproximação harmônica*.

Seguindo a prescrição já usada neste capítulo agora temos de considerar o operador hamiltoniano

$$\hat{H} = \frac{1}{2m}\hat{p}^2 + V(\hat{x}),$$

onde $\hat{p}$ é o operador momentum, $\hat{p} = -i\frac{\partial}{\partial x}$ e $V(\hat{x})$ é o operador que multiplica uma função $\psi(x)$ por $V(x)$.

A equação de Schrödinger independente do tempo, (11.12), agora fica sendo[4]

$$\big(\hat{p}^2 + \hat{x}^2\big)\psi(x) = E\psi(x).$$

Nosso objetivo é encontrar as funções ψ e os respectivos valores de E que satisfazem a equação acima.

Esta é uma equação diferencial ordinária que pode ser resolvida pelo método das séries de potências: essa técnica consiste em se supor que $\psi(x)$ pode ser escrita na forma

$$\psi(x) = \sum_{n=0}^{\infty} a_n x^n,$$

substituir na equação diferencial e obter uma relação de recorrência envolvendo os coeficientes a_n. O leitor pode encontrar essa abordagem em diversos livros, por isso não prosseguiremos nessa direção.

Vamos usar uma outra técnica, mais algébrica, que consiste em definir o operador

$$\hat{a} = \hat{x} + i\hat{p}; \tag{11.13}$$

note que

$$\hat{a}^* = \hat{x}^* + (i\hat{p})^* = \hat{x} - i\hat{p} \neq \hat{a}.$$

Dessa forma, esse operador não é auto-adjunto e portanto não representa um observável. No entanto, note que

$$\hat{a}^*\hat{a} = (\hat{x} - i\hat{p})(\hat{x} + i\hat{p}) = \hat{x}^2 + \hat{p}^2 + i[\hat{x}, \hat{p}] = \hat{H} - 1.$$

Ou seja, $\hat{H} = \hat{a}^*\hat{a} + 1$.

Exercício 11.2. *Mostre que o operador $\hat{a}^*\hat{a}$ é auto-adjunto.*

[4] Por simplicidade, adotamos $m = \frac{1}{2}$ e $k = 2$. Veja o exercício 11.10.

Denotaremos por $\hat{N}$ o operador auto-adjunto $\hat{a}^*\hat{a}$. Então podemos verificar que $\hat{N}$ é um operador positivo: de fato

$$\langle \psi|\hat{a}^*\hat{a}\psi\rangle = \langle \hat{a}\psi|\hat{a}\psi\rangle = \|\hat{a}\psi\|^2 \geq 0,$$

para qualquer vetor ψ. Podemos nos perguntar se existe algum vetor ψ_0 tal que $\hat{a}\psi_0 = 0$ (e portanto, $\hat{N}\psi_0 = 0$). A resposta é sim, e não é difícil obter tal vetor: a equação $\hat{a}\psi_0 = 0$ corresponde á equação diferencial

$$x\psi_0(x) + \frac{d}{dx}\psi_0(x) = 0. \tag{11.14}$$

Esta é uma equação diferencial separável e o leitor não terá dificuldade em verificar que a solução geral é dada por

$$\psi_0(x) = Ae^{-\frac{x^2}{2}},$$

onde o módulo da constante A pode ser determinado usando-se a normalização de ψ_0:

$$1 = \langle \psi_0|\psi_0\rangle = \int_{\mathbb{R}} |A|^2 e^{-x^2} dx = |A|^2\sqrt{\pi}$$

e, com a fase escolhida real positiva,

$$\psi_0(x) = \frac{1}{\pi^{1/4}}e^{-x^2}. \tag{11.15}$$

Agora podemos ver que

$$\hat{H}\psi_0 = (\hat{a}^*\hat{a} + 1)\psi_0 = \hat{a}^*\hat{a}\psi_0 + 1\psi_0 = 1\psi_0,$$

ou seja, ψ_0 é autofunção de $\hat{H}$ correspondente ao autovalor 1. E de fato 1 é o menor autovalor possível para $\hat{H}$, pois se temos $\hat{H}\phi = \lambda\phi$ para algum $\phi \neq 0$ então

$$\hat{H}\phi = \hat{N}\phi + 1\phi = \lambda\phi \Rightarrow \hat{N}\phi = (\lambda - 1)\phi.$$

Mas $\hat{N}$ é positivo, ou seja

$$0 \leq \left\langle \phi \middle| \hat{N}\phi \right\rangle = \langle \phi|(\lambda - 1)\phi\rangle = (\lambda - 1)\langle \phi|\phi\rangle$$

e assim $\lambda \geq 1$, ou seja, o menor autovalor possível para $\hat{H}$ é 1. Desta forma já encontramos, com ψ_0, a autofunção associada ao estado de menor energia do oscilador harmônico, muitas vezes chamado de estado fundamental.

O leitor pode perguntar nesse momento sobre a possibilidade de existência de outras autofunções linearmente independentes associadas ao autovalor 1; a preocupação é legítima.

Exercício 11.3. *Use o Teorema de Existência e Unicidade para a equação* (11.14) *para concluir que o auto-espaço associado ao autovalor* 1 *de* $\hat{H}$ *é unidimensional.*

Queremos agora encontrar outros autovalores e suas respectivas autofunções. Para isso note que

$$\hat{a}\hat{a}^* = \hat{H} + 1.$$

Agora considere $\tilde{\psi}_1 = \hat{a}^*\psi_0$ (você consegue obter explicitamente a função $\tilde{\psi}_1(x)$?). Para este vetor,

$$\hat{H}\tilde{\psi}_1 = \hat{H}\hat{a}^*\psi_0 = (\hat{a}^*\hat{a} + 1)\hat{a}^*\psi_0 = \hat{a}^*\hat{a}\hat{a}^*\psi_0 + \tilde{\psi}_1 =$$

$$\hat{a}^*\Big(\hat{H} + 1\Big)\psi_0 + \tilde{\psi}_1 = \hat{a}^*(\psi_0 + \psi_0) + \tilde{\psi}_1 = 3\tilde{\psi}_1.$$

Ou seja, $\tilde{\psi}_1$ é autovetor de $\hat{H}$ com autovalor 3. Mas note que

$$\left\|\tilde{\psi}_1\right\|^2 = \langle \hat{a}\psi_0 | \hat{a}\psi_0 \rangle = \langle \psi_0 | \hat{a}\hat{a}^*\psi_0 \rangle = 2\|\psi_0\|^2 = 2. \tag{11.16}$$

Assim, podemos escolher o autovetor normalizado como sendo

$$\psi_1 = \frac{1}{\sqrt{2}}\tilde{\psi}_1.$$

De forma análoga, podemos definir $\tilde{\psi}_2 = \hat{a}^*\tilde{\psi}_1 = (\hat{a}^*)^2\psi_0$ e o leitor pode verificar que teremos $\hat{H}\tilde{\psi}_2 = 5\tilde{\psi}_2$, e repetindo o procedimento de normalização, eq. (11.16), podemos definir

$$\psi_2 = \frac{1}{\sqrt{3!}}\tilde{\psi}_2;$$

de maneira geral,

$$\psi_n = \frac{1}{\sqrt{n!}}(\hat{a}^*)^n\psi_0 \tag{11.17}$$

será autofunção normalizada de $\hat{H}$, com o autovalor associado $(2n+1)$.

Exercício 11.4. *Use o fato de $\hat{H}$ ser positivo e o exercício 11.3 para mostrar que com o procedimento descrito geramos todos os autoespaços de $\hat{H}$. (Sugestão: suponha uma outra autofunção, φ, e aplique o operador $\hat{a}$ a ela.)*

Podemos usar agora essas soluções para obter o valor esperado de determinados observáveis. Como exemplo, vamos calcular o valor esperado de $\hat{x}^2$ no estado ψ_n. Já sabemos que esse valor esperado é dado pela expressão $\langle \psi_n | \hat{x}^2 \psi_n \rangle$. Agora note que, em função dos operadores $\hat{a}$ e $\hat{a}^*$ podemos escrever

$$\hat{x} = \frac{\hat{a} + \hat{a}^*}{2} \tag{11.18}$$

e então

$$\hat{x}^2 = \frac{1}{4}(\hat{a} + \hat{a}^*)(\hat{a} + \hat{a}^*) = \frac{1}{4}\Big(\hat{a}^2 + \hat{a}\hat{a}^* + \hat{a}^*\hat{a} + (\hat{a}^*)^2\Big).$$

Logo

$$\langle \psi_n | \hat{x}^2 \psi_n \rangle =$$

$$\frac{1}{4}\Big(\langle \psi_n | \hat{a}^2 \psi_n \rangle + \langle \psi_n | \hat{a}\hat{a}^* \psi_n \rangle \langle \psi_n | \hat{a}^* \hat{a} \psi_n \rangle + \Big\langle \psi_n \Big| (\hat{a}^*)^2 \psi_n \Big\rangle\Big).$$

Analisemos cada termo da expressão acima com calma:

$$\Big\langle \psi_n \Big| (\hat{a}^*)^2 \psi_n \Big\rangle \propto \langle \psi_n | \psi_{n+2} \rangle = 0,$$

visto que os vetores ψ_n são autovetores associados a autovalores distintos de $\hat{H}$. De forma similar podemos ver que $\langle \psi_n | \hat{a}^2 \psi_n \rangle$ também é igual a zero. Para os outros termos note que $\hat{a}^*\hat{a} = \hat{H} - 1$ e $\hat{a}\hat{a}^* = \hat{H} + 1$. Desta forma

$$\langle \psi_n | \hat{a}^* \hat{a} \psi_n \rangle = \Big\langle \psi_n \Big| \Big(\hat{H} - 1\Big) \psi_n \Big\rangle = \Big\langle \psi_n \Big| \hat{H} \psi_n \Big\rangle - \langle \psi_n | \psi_n \rangle =$$

$$(2n+1) - 1 = 2n$$

e

$$\langle \psi_n | \hat{a}\hat{a}^* \psi_n \rangle = \Big\langle \psi_n \Big| \Big(\hat{H} + 1\Big) \psi_n \Big\rangle = \Big\langle \psi_n \Big| \hat{H} \psi_n \Big\rangle + \langle \psi_n | \psi_n \rangle =$$

$$(2n+1)+1 = 2n+2.$$

Portanto temos

$$\langle \psi_n | \hat{x}^2 \psi_n \rangle = \frac{1}{4}[(2n+2)+2n] = n + \frac{1}{2}.$$

Exercício 11.5. *Calcule* $\langle \psi_n | \hat{p}^2 \psi_n \rangle$ *e verifique o que acontece com a relação de incerteza de Heisenberg nesses estados.*

Note que mesmo para o estado de mais baixa energia, $n = 0$, temos incerteza associada ao observável **x**. Este fato gera bastante discussão, alguma confusão e uma nomenclatura interessante: tratam-se das *flutuações de ponto zero*, ou ainda *flutuações quânticas*.

11.6 Exercícios

Exercício 11.6. *Reflita um pouco sobre a influência do tamanho da caixa nos níveis de energia da partícula no caso da seção 11.4.*

Exercício 11.7. *Para as autofunções da partícula na caixa obtenha os valores esperados da posição. Lembre que*

$$\langle \hat{x} \rangle = \langle \psi | \hat{x} \psi \rangle = \int_0^L x |\psi(x)|^2 dx.$$

Procedendo de forma similar, obtenha os valores esperados do momentum.

Exercício 11.8. *O leitor deve tentar resolver o problema da partícula em uma caixa considerando agora que esta está entre* $-L/2$ *e* $L/2$*. Como são os autovalores? E os autovetores?*

Exercício 11.9. *Obtenha expressões fechadas para* $\hat{a}^* \psi_n$ *e* $\hat{a} \psi_n$*.*

Exercício 11.10. *Refaça a discussão do oscilador harmônico mantendo as constantes* m*,* k *e* $\hbar$*. Lembre-se que para definir operadores* $\hat{a}$ *e* $\hat{a}^*$ *é necessário somar objetos de mesma dimensão*[5]*. Provavelmente você gostará de definir* $\omega = \sqrt{\dfrac{k}{m}}$*.*

[5] No sentido físico da palavra: comprimentos só podem ser somados a comprimentos, não a velocidades ou grandezas de outras dimensões.

Capítulo 12

Sistema de Funções Iteradas Quântico

Neste capítulo desejamos introduzir o interessante conceito de sistema de funções iteradas quântico, alvo de estudos recentes na literatura. Para entender esse objeto devemos rapidamente ver o que é um sistema dinâmico, um sistema iterado de funções e por fim ver como esse aparece de forma natural no contexto da mecânica quântica.

12.1 Sistemas Dinâmicos

Por *sistema dinâmico* entendemos o seguinte: um conjunto X (em geral um espaço métrico) e uma aplicação $f\colon X \to X$, que pode ou não ter uma inversa f^{-1} e com algum grau de regularidade (continuidade, diferenciabilidade). Um dos objetivos é tentar entender o que ocorre quando se aplica a função f a um ponto $x \in X$ por diversas vezes: em suma, caracterizar o conjunto

$$\{x, f(x), f(f(x)), f(f(f(x))), \ldots\}$$

(conhecido como órbita do ponto x) seus pontos de acumulação, como esse conjunto varia quando variamos o ponto inicial x, dentre ou-

tras perguntas. Nesse contexto é comum denotar $f(f(x))$ por $f^2(x)$, para dizer que a função f foi iterada duas vezes; de forma similar $f(f(f(x)))$ é denotada simplesmente por $f^3(x)$ e assim sucessivamente. Portanto $f^n(x)$ representa a composição de f n vezes, e não a n-ésima potência de x (que pode nem mesmo estar definida pois x não precisa estar em um conjunto numérico).

Exemplo 12.1. *Considere $X = [0, 1]$ e $f(x) = \sqrt{x}$, que é uma função contínua e invertível nesse intervalo. Se começamos com $x = 0$, então é claro que $f(0) = 0 = f^2(0) = \cdots = f^n(0)$ para todo n natural; o mesmo ocorre se começamos com $x = 1$: temos $f(1) = 1 = f^2(1) = \ldots = f^n(1)$. Estes dois pontos, por razões óbvias, são pontos fixos para f e nesses casos a descrição da órbita e dos pontos de acumulação da mesma é imediata. E para um $x \in (0, 1)$? Note que $f(x) > x$ para pontos nesse intervalo e o que de fato ocorre é que, nesse caso, $f^n(x) \to 1$ quando n cresce.*

Observação 8. *O leitor pode ilustrar o que é descrito no último exemplo usando uma calculadora e apertando diversas vezes a tecla de raiz quadrada.*

Exemplo 12.2. *Seja $X = \mathbb{Z}$ e $f(x) = x + 1$. Nesse caso temos uma dinâmica que claramente vai para $+\infty$ qualquer que seja o ponto x inicial.*

12.2 Sistema de Funções Iteradas

Sem tentar ser o mais geral possível, podemos ver um sistema de funções iteradas como sendo formado pelos seguintes elementos: um conjunto X, aplicações $f_i \colon X \to X$, $i = 1, 2, \ldots, k$, e números reais não negativos $p_i, i = 1, \ldots, k$ tais que $p_1 + p_2 + \cdots + p_k = 1$, o que permite interpretá-los como sendo uma probabilidade sobre o conjunto $\{1, 2, \ldots, k\}$.

Nesse caso, no lugar de iterar apenas uma função f, devemos escolher um índice em $\{1, 2, \ldots, k\}$, digamos j, com probabilidade p_j e então iterar f_j. Dessa forma, a evolução de um ponto x sob a dinâmica é de fato uma evolução aleatória. Podemos então perguntar o que ocorre com a evolução de um certo ponto para diferentes sorteios.

Exemplo 12.3. *Seja* $X = [0,1]$, $f_1(x) = x/3$ *e* $f_2(x) = (2+x)/3$, *com* $p_1 = p_2 = 1/2$. *Onde está contida a dinâmica limite nesse caso? Vejamos: após o primeiro iterado a imagem estará contida na imagem de* f_1 *ou na imagem de* f_2, *respectivamente, os conjuntos* $[0, 1/3]$ *e* $[2/3, 1]$. *Após o segundo iterado, a imagem estará contida na imagem, por* f_1 *ou* f_2, *dos dois intervalos anteriores, que são os conjuntos* $[0, 1/9]$, $[2/9, 3/9]$, $[6/9, 7/9]$ *e* $[8/9, 1]$. *Na próxima etapa ficaremos com oito conjuntos, com comprimentos iguais a* $1/27$ *e assim sucessivamente. Esta construção, como o leitor já deve ter percebido, é exatamente a do conjunto de Cantor: de início é retirado o intervalo* $(1/3, 2/3)$*; depois, nos intervalos fechados restantes, retira-se o terço central. Desta forma, a dinâmica desse Sistema tem como conjunto limite exatamente o conjunto de Cantor* K.

Exemplo 12.4. *Seja* $X = \mathbb{Z}$, $f_1(x) = x+1$ *e* $f_2(x) = x-1$, *com* $p_1 = p_2 = 1/2$. *Nesse caso podemos interpretar a dinâmica da seguinte forma: temos probabilidade* $1/2$ *de iterar* f_1, *que representa dar um passo de comprimento* 1 *para a direita e probabilidade* $1/2$ *de iterar* f_2, *ou seja, dar um passo de comprimento* 1 *para a esquerda. Esse é um modelo bastante conhecido e estudado, conhecido como* passeio aleatório. *Supondo que começamos em* $x = 0$ *e fazemos* N *iterações, é fácil obter a probabilidade de estar em um certo* $n \in \mathbb{Z}$.

12.3 Sistema de Funções Iteradas Quântico

Imaginemos agora a seguinte situação: temos um sistema quântico cuja evolução está sujeita a algum tipo de ruído ou flutuação que é aleatória. Uma forma de modelar este caso é pensar que temos não uma evolução temporal (descrita por um operador unitário U) mas sim um conjunto de operadores unitários $U_j, j \in \{1, 2, \ldots, k\}$ e probabilidades p_j.

O estado do sistema quântico pode ser descrito, como previamente, por uma matriz densidade ρ. A evolução temporal do sistema então é dada por

$$\Phi(\rho) = \sum_{j=1}^{k} p_j U_j \rho U_j^*. \tag{12.1a}$$

Não é difícil verificar que se ρ é uma matriz densidade então $\Phi(\rho)$ também é matriz densidade, ou seja, Φ é uma aplicação no espaço de matrizes densidade. Este é um espaço interessante porém não é um espaço vetorial (lembre-se, as matrizes densidade tem traço um, mas a soma de matrizes de traço um tem traço igual a dois, e portanto não está no espaço), o que nos impede de usar as técnicas bem conhecidas da álgebra linear. Para contornar este problema lidaremos com uma extensão de Φ para matrizes $d \times d$ quaisquer, que continuaremos denotando por Φ:

$$\Phi(X) = \sum_{j=1}^{k} p_j U_j X U_j^*. \tag{12.1b}$$

Desta forma note que

$$\Phi(X+\lambda Y) = \sum_{j=1}^{k} p_j U_j (X+\lambda Y) U_j^* = \sum_{j=1}^{k} p_j U_j X U_j^* + \lambda \sum_{j=1}^{k} p_j U_j Y U_j^*$$

e portanto Φ é uma aplicação linear de $M_d(\mathbb{C})$ que então pode ser representada por uma matriz (qual a dimensão dessa matriz?).

Nesse contexto torna-se totalmente natural procurar soluções em X e λ da equação

$$\Phi(X) = \lambda X,$$

que nada mais é do que uma equação para autovalores e autovetores. Dessa maneira a existência de soluções é algo já garantido por resultados básicos de álgebra linear.

O espaço $M_d(\mathbb{C})$ admite um produto interno bastante natural que é definido como sendo

$$\langle A|B\rangle := \mathrm{Tr}(A^*B).$$

Este produto interno induz uma norma que é simplesmente

$$\|A\| = \sqrt{\langle A|B\rangle} = \sqrt{\mathrm{Tr}(A^*A)}.$$

Para transformações U unitárias podemos verificar que

$$\|UAU^*\| = \sqrt{\mathrm{Tr}(UA^*U^*\,UAU^*)} = \sqrt{\mathrm{Tr}(UA^*AU^*)} =$$

$$\sqrt{\mathrm{Tr}(A^*AU^*U)} = \sqrt{\mathrm{Tr}(A^*A)} = \sqrt{\|A\|^2} = \|A\|.$$

Sendo assim, temos

$$\|\Phi(X)\| \leq \sum_{j=1}^{k} p_j \|U_j X U_j^*\| = \sum_{j=1}^{k} p_j \|X\| = \|X\|$$

e portanto os autovalores da aplicação linear Φ são tais que $|\lambda| \leq 1$, ou seja, estão todos no disco unitário. Por outro lado, não é difícil verificar que $\Phi(I) = I$, e assim $\lambda = 1$ está de fato no espectro do operador.

Uma pergunta interessante que pode ser feita nesse contexto é a de como caracterizar um estado limite, ou seja, dado um estado inicial ρ_0 saber como será, após longo tempo, o estado descrito pela evolução temporal Φ definida acima. Em outras palavras, caracterizar

$$\lim_{n \to \infty} \Phi^n(\rho_0). \tag{12.2}$$

Considerando a extensão de Φ para o espaço de todas as matrizes e a equação de autovalores $\Phi(X) = \lambda X$, vamos aceitar, por hipótese, que temos a seguinte situação: *o subespaço associado ao autovalor* 1 *tem dimensão* 1*, e portanto é gerado por apenas uma matriz (que já sabemos ser* I*); os demais autovalores são todos estritamente menores (em norma) do que* $\lambda_0 \in (0, 1)$. Nesse caso, como a dinâmica de Φ é linear, podemos decompor o estado inicial ρ_0 em uma combinaçao de autovetores do tipo

$$\rho_0 = a_0 I + \sum_{i=1}^{D} a_i X_i,$$

onde os X_i são autovetores associados aos autovalores de norma menor do que 1. Desta forma, é fácil ver que

$$\Phi^n(\rho_0) = a_0 I + \sum_{i=1}^{D} a_i \lambda_i^n X_i$$

e assim o vetor limite é $a_1 I$; porém este vetor não corresponde a um estado, pois não tem traço unitário, mas obviamente podemos normalizá-lo para que isso ocorra e assim temos a dinâmica assintótica nessa situação particular.

Exemplo 12.5. *Temos um exemplo da situação acima quando a dinâmica é dada por*

$$\Phi(X) = \frac{1}{2}X + \frac{1}{2}UXU^*,$$

onde

$$U = \frac{1}{2}\begin{bmatrix} 1 & -\sqrt{3} \\ \sqrt{3} & 1 \end{bmatrix}.$$

O leitor interessado em uma descrição mais precisa de estados limite deve consultar [LP, NAJ]. O problema também pode ser generalizado pela consideração de uma transformação Φ que não é linear em ρ (e para a qual o raciocínio acima não pode ser aplicado), como por exemplo

$$\Phi(\rho) = \sum_{j=1}^{k} p_j(\rho) U_j \rho U_j^*, \tag{12.3}$$

onde $\sum_j p_j(\rho) = 1$ para todo ρ; nesse caso as probabilidades de ocorrência de cada uma das dinâmicas U_i dependem do estado em consideração. Um problema dessa natureza pode ser abordado com ferramentas um pouco mais sofisticadas do que as usadas aqui e o leitor curioso pode consultar, por exemplo, [BLLT] para uma abordagem deste caso.

Capítulo 13

Desigualdades de Bell

A mecânica quântica é uma teoria muito diferente da mecânica clássica em vários aspectos e um deles é o fato de que tudo que podemos saber sobre uma medição são as probabilidades dos resultados possíveis. Probabilidades também aparecem na física clássica, mas como fruto do conhecimento parcial a respeito do sistema em questão. Em mecânica quântica, as probabilidades parecem ser intrínsecas à teoria e isso causa um certo desconforto. Será que o mundo é realmente probabilístico ou falta alguma coisa na teoria?

13.1 EPR e os Elementos de Realidade

Uma bola de tênis viajando em sua trajetória entre as raquetes de dois jogadores tem posição e velocidade definidos em cada instante de tempo. Se não podemos determiná-los é por não possuirmos instrumentos adequados para realizar cada teste com precisão. Em 1935, Einstein, Podolsky e Rosen (EPR) publicaram o famoso artigo "*Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?*" em que eles argumentam que uma teoria completa não deveria ser intrinsecamente probabilística como a mecânica quântica [EPR]. A ideia central é que um elétron deve ser parecido com a bola de tênis: possui velocidade e posição bem definidos e como a mecânica quântica não é capaz de prevê-los deve ser uma teoria incompleta.

EPR começam definindo o que são *elementos de realidade*: existe um elemento de realidade associado a um observável físico se esse observável pode ser determinado com precisão sem que o sistema seja perturbado. Eles afirmam que em uma teoria completa todo elemento de realidade deve ter valor bem definido. A realização de um teste apenas revela esse valor. Vamos ver o que acontece nos exemplos que tratamos nesse texto.

Suponhamos que um sistema esteja associado a um espaço de estados E de dimensão d. Sabemos que, se os operadores A_1 e A_2 não comutam, os testes associados a eles não são compatíveis e existem estados puros do sistema nos quais não podemos prever o resultado de ambos, com precisão arbitrária. Com esse argumento, EPR concluem que ou a mecânica quântica não é completa ou que operadores que não comutam não podem estar ambos associados a elementos de realidade.

Para eliminar a segunda opção eles propõe uma situação parecida com a seguinte: consideremos dois qbits no estado emaranhado

$$|\Psi_-\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}} = \frac{|+-\rangle - |-+\rangle}{\sqrt{2}} \tag{13.1}$$

e que estejam distantes um do outro. Se realizarmos o teste associado ao observável σ_z no primeiro qbit, podemos obter 0 ou 1 com probabilidade 0.5. Suponhamos que 0 seja o resultado. Então o estado após a medição é $|01\rangle$, o estado do segundo qbit é $|1\rangle$ e se testarmos σ_z no segundo qbit obteremos 1. Se realizarmos a medição do observável σ_x no primeiro qbit, podemos obter $+$ ou $-$, também com probabilidade 0.5. Suponhamos que $+$ seja o resultado. Então o estado após a medição é $|+-\rangle$, o estado do segundo qbit é $|-\rangle$ e se testarmos σ_x no segundo qbit obteremos $-$ como resposta. Agora entra o ponto central do argumento de EPR: se os qbits estão distantes então uma medição no primeiro não pode afetar o segundo. Assim, escolhendo medir σ_x ou σ_z no primeiro qbit, podemos determinar o valor de σ_x ou σ_z no segundo qbit, sem perturbá-lo. Logo ambos os observáveis podem ser associados a elementos de realidade.

O argumento acima elimina a possibilidade de que dois observáveis que não comutam não podem estar ambos associados a elementos de realidade. Desse modo, segundo EPR, a mecânica quântica deve ser uma teoria incompleta.

O artigo de EPR iniciou uma longa discussão. Seria possível encontrar uma teoria em que os elementos de realidade de EPR possuíssem valores definidos? Teorias desse tipo ficaram conhecidas como *Teorias de Variáveis Ocultas*[1] (TVO).

13.2 Bell

Em 1964, John Bell propôs uma maneira de testar a existência de variáveis ocultas [Bel64]. Ele mostrou que, em uma teoria de variáveis ocultas obedecendo hipóteses razoáveis[2], os valores esperados de alguns observáveis deveriam satisfazer uma inequação. Esta é a primeira *desigualdade de Bell* da história. Se em algum experimento essa desigualdade fosse violada, poderíamos concluir que tais teorias não poderiam ser verdadeiras.

Bell derivou sua desigualdade no contexto de um teste em um par de partículas de spin $\frac{1}{2}$ no estado $|\Psi_-\rangle$. Vamos supor que o estado do sistema seja descrito por uma variável λ, que faz o papel de variável oculta e que determina qual será o valor obtido quando realizamos uma medição da componente de spin em uma direção $\vec{u}$. A variável λ pode ser contínua ou discreta, pode ter uma componente ou várias. Em geral, temos apenas que $\lambda \in \Lambda$, com Λ um certo espaço de parâmetros para variáveis ocultas.

Vamos supor que um teste é realizado em cada parte do par: na parte A vamos medir a componte de spin na direção do vetor $\vec{a}$ e na parte B vamos medir a componente de spin na direção $\vec{b}$. O valor obtido em A, $v(\vec{a}, \lambda)$, depende da direção $\vec{a}$ escolhida e de λ. Analogamente o valor obtido em B, $v(\vec{B}, \lambda)$, depende da direção $\vec{b}$ escolhida e de λ. Sabemos apenas que os valores possíveis para ambos os resultados são ± 1.

Suponhamos que $p(\lambda)$ seja a distribuição de probabilidade de λ. Então o valor esperado de $v(\vec{a}, \lambda)v(\vec{b}, \lambda)$ é

$$E\left(\vec{a}, \vec{b}\right) = \int_\lambda p(\lambda) v(\vec{a}, \lambda) v(\vec{b}, \lambda) d\lambda.$$

[1] Do inglês, *hidden-variable theories*.
[2] Antecipando a conclusão, *aparentemente* razoáveis.

Depois de alguns cálculos, Bell mostra que dadas três direções $\vec{a}$, $\vec{b}$ e $\vec{c}$ vale a desigualdade

$$1 + E\Big(\vec{b}, \vec{c}\Big) \geq \Big|E(\vec{a}, \vec{b}) - E(\vec{a}, \vec{c})\Big|. \tag{13.2}$$

Para que não haja contradição com a mecânica quântica, $E\Big(\vec{a}, \vec{b}\Big)$ deve concordar com o valor esperado de $\vec{a} \cdot \vec{\sigma} \otimes \vec{b} \cdot \vec{\sigma}$.

Exercício 13.1. *Mostre que*

$$\langle \Psi_- | \vec{a} \cdot \vec{\sigma} \otimes \vec{b} \cdot \vec{\sigma} | \Psi_- \rangle = -\vec{a} \cdot \vec{b}.$$

Exercício 13.2. *Use o exercício 13.1 e a escolha* $\vec{a} = (1, 0, 0)$, $\vec{b} = (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$ *e* $\vec{c} = (0, 0, 1)$, *para obter uma violação da desigualdade de Bell* (13.2).

Desde então várias outras desigualdades foram demonstradas, todas elas com o mesmo espírito: provar que restrições válidas para teorias de variáveis ocultas não são satisfeitas por todos os estados da mecânica quântica. Essas desigualdades também ficaram conhecidas como *desigualdades de Bell*. Algumas delas são de demonstração simples, e uma das mais famosas é a desigualdade CHSH.

13.3 A Desigualdade CHSH

A desigualdade CHSH, das iniciais de Clauser, Horne, Shimony e Holt, que a provaram em [CHSH], é uma desigualdade de Bell simples e operacional. Essa desigualdade é obtida quando consideramos a situação em que dois laboratórios compartilham um sitema composto AB. Em cada laboratório é possível realizar dois testes[3]: na parte A podem ser realizados os testes correspondentes aos observáveis A_1 e A_2, enquanto na parte B podem ser realizados os testes

[3] Estamos mantendo a linguagem de testes e observáveis, pois este é um livro de mecânica quântica; mas para entender a derivação da desigualdade CHSH, e de desigualdades de Bell em geral, é importante lembrar justamente que elas *não* tratam de mecânica quântica. A mecânica quântica só entra nessa história por ser a única teoria com algum respaldo experimental que permite violar desigualdades de Bell.

correspondentes aos observáveis B_1 e B_2. As respostas possíveis para todos os testes envolvidos são 1 e -1. Vamos supor que a escolha de qual teste é realizado em uma das partes é independente da escolha na outra parte[4].

Em uma TVO, cada observável O deve possuir valor bem definido, que denotaremos por $v(O)$. Queremos verificar quais são os valores possíveis de

$$\begin{aligned} CHSH &= v(A_1)v(B_1) + v(A_2)v(B_1) + v(A_2)v(B_2) - v(A_1)v(B_2) \\ &= (v(A_1) + v(A_2))v(B_1) + (v(A_2) - v(A_1))v(B_2). \quad (13.3) \end{aligned}$$

Os resultados possíveis para todos os testes são ± 1, de modo que ou $v(A_1) + v(A_2) = 0$, ou $v(A_1) - v(A_2) = 0$. Em todo caso, temos $CHSH = \pm 2$.

Podemos calcular a esperança dessa quantidade, que por ser uma combinação convexa de seus possíveis valores, obedecerá

$$|\langle CHSH \rangle| \leq 2, \quad (13.4)$$

a famosa desigualdade CHSH.

Para verificações experimentais, é importante usarmos a linearidade da esperança, para escrever

$$\begin{aligned} \langle CHSH \rangle &= \langle v(A_1)v(B_1) + v(A_2)v(B_1) + v(A_2)v(B_2) - v(A_1)v(B_2) \rangle \\ &= \langle v(A_1)v(B_1) \rangle + \langle v(A_2)v(B_1) \rangle \\ &\quad + \langle v(A_2)v(B_2) \rangle - \langle v(A_1)v(B_2) \rangle, \end{aligned}$$

o que permite que o valor esperado de $CHSH$ seja obtido fazendo medições independentes de A_i em uma parte, B_j em outra, para depois colecionar os resultados e saber para qual dos quatro valores esperados ele contribui (de acordo com o par (i, j)). Feito esse trabalho de pós-processamento, obtém-se $|\langle CHSH \rangle|$ e pode-se verificar se (13.4) é satisfeita.

Até agora a mecânica quântica não entrou na brincadeira. A dedução foi feita supondo que os resultados dos testes são governados por uma teoria de variáveis ocultas.

[4] O que pode ser interpretado como a hipótese que não há comunicação entre as partes, até que as medições sejam realizadas.

Exercício 13.3. *Use novamente o exercício 13.1 e os observáveis*

$$A_1 = \sigma_z, \quad A_2 = \sigma_x,$$

$$B_1 = \frac{-\sigma_x - \sigma_z}{\sqrt{2}}, \quad B_2 = \frac{-\sigma_x + \sigma_z}{\sqrt{2}}$$

para obter uma violação de (13.4).

Desse modo, ou a mecânica quântica não está correta, ou a natureza não pode ser descrita através de uma teoria de variáveis ocultas.

O fato da desigualdade CHSH não ser satisfeita para todos os estados de dois qbits e escolhas de observáveis locais significa que alguma das exigências que foram feitas na demonstração da desigualdade não pode ser satisfeita. Duas suposições cruciais que aparecem nos cálculos são[5]:

- *Realismo*: os observáveis físicos A_1, A_2, B_1, B_2 possuem valores definidos, independentes da realização ou não de suas medições;
- *Localidade*: os testes realizados na parte A não alteram os resultados dos testes realizados na parte B.

A violação da desigualdade CHSH mostra que não vale o *realismo local*, ou seja, as duas suposições anteriores não podem ser feitas ao mesmo tempo. Tanto realimo quanto localidade são propriedades aparentemente válidas no nosso dia a dia, mas na descrição do mundo microscópico, pelo menos uma delas deve ser descartada.

Há vários experimentos de violações de desigualdades de Bell, mas também há várias discussões sobre porque cada experimento já realizado ainda não cumpre todas as exigências necessárias para eliminar a possibilidade das variáveis ocultas descreverem uma *realidade* microscópica consistente com nossos preconceitos clássicos.

Dois dos *loopholes* mais famososos são os de *detecção* e de *localidade*, ou *sinalização*. O *loophole* de detecção se origina no fato de não haver detector com eficiência total: sempre há uma parcela dos sistemas preparados que não são detectados, seja por uma ou por outra

[5]Por vezes também se considera como outra suposição o *livre arbítrio*, no sentido de cada experimentador poder escolher livremente qual dos observáveis irá medir em cada rodada do experimento.

parte. Mas, mentes criativas alegam, esse efeito, supostamente aleatório, de detecção ou não detecção deve ser determinístico em uma TVO. E podem ser justamente esses dados "faltantes" os responsáveis pela violação das desigualdades, ou seja, se eles também fossem detectados e incluídos na estatística das contagens, não haveria violação. Já o *loophole* de localidade se refere à necessidade de garantir que as escolhas independentes de A_i e B_j realmente o sejam. A primeira vez que um outro princípio físico pôde ser experimentalmente invocado para fechar este *loophole* foi quando Alain Aspect e colaboradores [Asp] fizeram um experimento onde as escolhas dependiam de circuitos eletrônicos independentes, localizados em laboratórios suficientemente afastados para que um sinal enviado por um laboratório não fosse capaz de atingir o outro se viajasse à velocidade da luz, antes que a outra decisão fosse tomada; em linguajar de *teoria da relatividade*, esses eram eventos com separação *tipo espaço*, portanto fora dos cones de causalidade.

Para mais aprofundamento em língua portuguesa, sugerimos as referências [QA] e [Rab].

Capítulo 14

Contextualidade

No capítulo anterior mostramos que não é possível encontrar uma teoria realista local que concorde com a mecânica quântica, uma vez que a primeira deve satisfazer a desigualdade CHSH, que é violada para alguns estados quânticos. Podemos nos perguntar se existe uma maneira de demonstrar essa impossibilidade encontrando alguma contradição que seja independente do estado do sistema. Como nenhum experimento até hoje mostrou alguma contradição com as previsões da mecânica quântica, incluindo os experimentos que testam a desigualdade CHSH, assumimos que essas teorias devem ser compatíveis com ela. Teorias de variáveis ocultas com essa propriedade serão chamadas *Teorias de Variáveis Ocultas Compatíveis* (TOVC). Um estudo bem completo sobre o assunto pode ser encontrado em [Cab].

14.1 von Neumann

Um dos primeiros a tentar mostrar a impossibilidade de variáveis ocultas compatíveis foi von Neumann em [vNe]. A ideia é mais ou menos a seguinte: sejam A e B as matrizes que representam dois observáveis em um sistema quântico. Sabemos que, se ρ é a matriz densidade que representa um estado desse sistema, então

$$\langle A\rangle = \mathrm{Tr}(\rho A), \quad \langle B\rangle = \mathrm{Tr}(\rho B),$$

$$\langle A+B\rangle = \mathrm{Tr}(\rho(A+B)) = \mathrm{Tr}(\rho A) + \mathrm{Tr}(\rho B) = \langle A\rangle + \langle B\rangle.$$

Em uma teoria de variáveis ocultas, um observável A tem valor definido em cada estado do sistema[1]. Denotaremos esse valor por $v(A)$. Para que essa teoria seja compatível com a mecânica quântica, $v(A)$ deve ser um dos autovalores de A. Em sua tentativa de refutar a existência de tais teorias, von Neumann assumiu que a expressão

$$\langle A+B\rangle = \langle A\rangle + \langle B\rangle \tag{14.1}$$

também deveria ser válida para uma TVOC. Pelo fato de que A possui valor definido temos $v(A) = \langle A\rangle$ e portanto

$$v(A+B) = v(A) + v(B). \tag{14.2}$$

Com a restrição (14.2), não é difícil mostrar que não existe uma TVOC.

Exemplo 14.1. *Em verdade, um contra-exemplo. Vamos considerar um sistema de um qbit e os observáveis* $A = \sigma_x$ *e* $B = \sigma_y$*. Em mecânica quântica os resultados possíveis para uma medição desses observáveis são* ± 1 *e portanto* $v(A) = \pm 1$ *e* $v(B) = \pm 1$*. Assim temos que para uma TVOC satisfazendo a equação* (14.2)

$$v(A+B) = -2,\ 0 \text{ ou } 2.$$

No entanto $A + B = \sigma_x + \sigma_y$ *possui autovalores* $\pm\sqrt{2}$ *e por isso* $v(A+B) = \pm\sqrt{2}$ *o que é uma contradição.*

14.1.1 A Falha na Demonstração de von Neumann

A crítica feita ao argumento mostrado acima se deve ao fato de que assumimos a relação (14.2) entre valores de observáveis não compatíveis, que não podem ser medidos simultaneamente em mecânica quântica [Mer]. O fato de valer a relação (14.1) para os valores esperados em mecânica quântica não é suficiente para exigirmos que isso seja válido em teorias de variáveis ocultas, ou seja, o fato de não valer (14.2) para observáveis não compatíveis não contradiz a mecânica

[1] Em uma TVOC o estado do sistema é determinado pelo vetor de estado da mecânica quântica mais uma outra variável que pode ser um número real, um vetor real, etc...

quântica em ponto algum. Por outro lado, se os observáveis são compatíveis, (14.2) deve ser satisfeita. Mais geralmente, se $A_1, \ldots, A_n$ é um conjunto de observáveis compatíveis em mecânica quântica que obedecem uma relação do tipo

$$f(A_1, \ldots, A_n) = 0$$

em que f é uma aplicação qualquer, então a mesma relação deve ser satisfeita pelos valores assumidos pelos observáveis correspondentes em uma TVOC:

$$f(v(A_1), \ldots, v(A_n)) = 0.$$

Se os observáveis são incompatíveis, não podemos assumir que uma relação válida em mecânica quântica também seja válida em uma TVOC. A contradição que aparece no exemplo 14.1 veio justamente ao fazermos uma restrição desse tipo e por isso o argumento de von Neumann é falho.

14.1.2 Um Modelo de Variáveis Ocultas Compatível em Dimensão Dois

Podemos construir um modelo de variáveis ocultas bem simples para um qbit [Bel66]. Seja A um operador no espaço de estados correspondente E. Sabemos que A pode ser escrito na forma

$$A = a_0 I + a_1 \sigma_x + a_2 \sigma_y + a_3 \sigma_z,$$

em que cada a_i é um número real.

Fazendo $\vec{a} = (a_1, a_2, a_3)$ temos que os autovalores de A, e portanto os possíveis valores de $v(A)$, são

$$v(A) = a_0 \pm \|\vec{a}\|.$$

Seja $|\phi\rangle$ um vetor em E e $\vec{n}$ o ponto na esfera de Bloch correspondente a $|\phi\rangle$. Então

$$\langle A \rangle = \langle \phi \, |A| \, \phi \rangle = a_0 + \vec{a} \cdot \vec{n}.$$

Além do vetor de estado $|\phi\rangle$, vamos supor que o estado do sistema também seja descrito por um vetor $\vec{m} \in S^2$. O vetor $\vec{m}$ faz o papel de variável oculta de maneira que o estado completo do sistema é dado

pelo par $(|\phi\rangle, \vec{m})$. Esse par determina o valor a ser obtido no teste de acordo com a regra:

$$\begin{cases} v(A,\vec{m}) = a_0 + \|\vec{a}\| & \text{se } (\vec{m}+\vec{n})\cdot\vec{a} \geq 0, \\ v(A,\vec{m}) = a_0 - \|\vec{a}\| & \text{se } (\vec{m}+\vec{n})\cdot\vec{a} < 0, \end{cases}$$

em que $v(A,\vec{m})$ denota o valor atribuído ao teste A no estado $(|\phi\rangle, \vec{m})$.

Esse modelo é compatível com mecânica quântica, uma vez que

$$\int_{S^2} v(A,\vec{m})\, d\vec{m} = \langle A\rangle, \quad \forall\, |\phi\rangle.$$

14.2 Gleason

Gleason não estava preocupado com teorias de variáves ocultas. Ele estava interessado em estudar medidas no conjunto de subespaços fechados de um espaço vetorial [Gle].

Definição 14.1. *Seja E um espaço vetorial e $\mathcal{F}$ o conjunto dos subespaços fechados de E. Uma medida em $\mathcal{F}$ é uma função $\mu : \mathcal{F} \to \mathbb{R}_+$ tal que se $\{E_i\}$ é uma coleção enumerável de subespaços de E mutuamente ortogonais que geram o subespaço E_I então*

$$\mu(E_I) = \sum_i \mu(E_i). \tag{14.3}$$

Uma medida desse tipo é chamada uma medida de probabilidade se $\mu : \mathcal{F} \to [0,1]$, $\mu(\{0\}) = 0$ e $\mu(E) = 1$.

Teorema 14.1 (Gleason)**.** *Seja E um espaço de Hilbert separável*[2] *de dimensão maior ou igual a três, sobre $\mathbb{R}$ ou $\mathbb{C}$. Então toda medida de probabilidade em $\mathcal{F}$ é da forma*

$$\mu(E_i) = \mathrm{Tr}(\rho P_i), \tag{14.4}$$

em que P_i é o projetor sobre o subespaço E_i, para algum ρ que é operador positivo semi-definido de traço um.

[2]Um espaço de Hilbert é um espaço vetorial com produto interno completo com a norma gerada por ele. Um espaço é dito separável se possui um subconjunto denso enumerável. Os espaços $\mathbb{C}^n$ são espaços de Hilbert separáveis, e como são os exemplos considerados nesse texto, não precisamos nos preocupar muito com essas exigências.

Apesar de não estar interessado em TVO's, na demonstração de seu famoso teorema Gleason prova o seguinte resultado:

Lema 14.2 (Lema de Gleason)**.** *Seja* $\{|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle\}$ *LI em um espaço de Hilbert e*

$$R_3 = \{c_1|\psi_1\rangle + c_2|\psi_2\rangle + c_3|\psi_3\rangle \ ; \ c_i \in \mathbb{R}\}.$$

Então qualquer medida μ *em* R_3 *deve ser uma função contínua de* c_1, c_2, c_3*.*

Suponhamos que o espaço de estados de um sistema físico seja um espaço vetorial E de dimensão finita maior ou igual a três. Em uma teoria de variáveis ocultas, o resultado do teste correspondente a um projetor unidimensional $P_\phi = |\phi\rangle\langle\phi|$ em um dado estado deve ter um valor definido. Se essa teoria é compatível com a mecânica quântica os valores possíveis são 0 e 1 (que são os autovalores de P_ϕ).

Suponhamos que $\{|\phi_1\rangle, |\phi_2\rangle, \ldots, |\phi_n\rangle\}$ seja uma base ortonormal para H. Então

$$P_{\phi_1} + \ldots + P_{\phi_n} = I \tag{14.5}$$

e portanto

$$v(P_{\phi_1}) + \ldots + v(P_{\phi_n}) = v(I) = 1 \tag{14.6}$$

uma vez que os observáveis P_{ϕ_i} são compatíveis para vetores $|\phi_i\rangle$ ortogonais.

Como $v(P_{\phi_i})$ vale 0 ou 1, a equação (14.6) implica que $v(P_{\phi_{i_0}}) = 1$ para algum i_0 e $v(P_{\phi_i}) = 0$ se $i \neq i_0$. Fixemos $|\phi\rangle$ e suponhamos que $v(P_\phi) = 1$. Para qualquer $|\psi\rangle$ ortogonal a $|\phi\rangle$ temos que $v(P_\phi) = 0$ uma vez que o conjunto $\{|\phi\rangle, |\psi\rangle\}$ pode ser estendido a uma base para a qual vale a equação (14.6). Agora vamos mostrar que isso não pode acontecer pelo lema de Gleason.

Podemos criar uma medida no conjunto de subespaços de E definindo para os subespaços unidimensionais

$$\begin{cases} \mu(E_\Phi) = 0, & \text{se } P_\Phi \text{ possui o valor 0 para aquele estado,} \\ \mu(E_\Phi) = 1, & \text{se } P_\Phi \text{ possui o valor 1 para aquele estado,} \end{cases}$$

em que E_ϕ é o subespaço gerado por $|\phi\rangle$. Estendemos a outros subespaços utilizando a propriedade (14.3).

Criamos então uma medida μ que vale 1 em E_ϕ e vale 0 para todo subespaço ortogonal a E_ϕ. Se considerarmos a restrição dessa medida a um subconjunto como o conjunto R_3 que aparece no lema de Gleason, ela deve ser contínua, mas é impossível construir uma medida contínua com tais propriedades.

Um maneira geométrica de vizualizarmos a impossibilidade dessa construção é atribuindo cores aos vetores de $R_3 \subset E$ de acordo com o valor da medida no subespaço gerado por esse vetor. Atribuiremos a cor vermelha se $v(P_\phi) = 1$ e verde se $v(P_\phi) = 0$. O problema de construir uma medida contínua com as propiedades desejadas é equivalente ao problema de colorir a esfera continuamente com as cores vermelha e verde de maneira que se um ponto é vermelho o círculo no plano ortogonal a ele deve ser todo verde.

Exercício 14.1. *Mostre que tal coloração da esfera é impossível.*

14.2.1 A Crítica de Bell

No argumento acima, assumimos que se $v(P_\phi) = 1$ então $v(P_\psi) = 0$ para todo $|\psi\rangle$ ortogonal a $|\phi\rangle$. Essa é a hipótese de *não-contextualidade*: o valor que um observável assume não depende do conjunto de observáveis compatíveis[3] que é testado com ele. A princípio, nada garante que podemos assumir não-contextualidade, mas é mais uma exigência que parece natural para uma teoria realista.

Em algumas situções podemos usar a hipótese de localidade para garantir não-contextualidade. Suponhamos que $\{A, B_1, \ldots, B_n\}$ seja um conjunto de observáveis compatíveis e que $\{A, C_1, \ldots, C_n\}$ também seja um conjunto de observáveis compatíveis (os observáveis B_i não são necessariamente compatíveis com os observáveis C_j). Suponhamos também que o teste A seja realizado por uma parte do aparato, enquanto outra parte pode escolher entre realizar os testes relacionados a $B_1, \ldots, B_n$, ou os testes relacionados a $C_1, \ldots, C_n$.

Nesse cenário, e com a hipótese de localidade, ou seja, que não há ação a distância, esperamos que as mudanças na parte do aparato que mede $B_1, \ldots, B_n$ ou $C_1, \ldots, C_n$ não afetem o resultado do teste A. Logo $v(A)$ não deve depender do conjunto de observáveis que vamos

[3]Cuidado para não confundir compatibilidade de observáveis com compatibilidade de TVO com a mecânica quântica.

testar na outra parte do aparato e a hipótese de não-contextualidade pode ser justificada.

14.3 Bell, Kochen e Specker

A ideia de Kochen e Specker é em alguns aspectos parecida com a da seção anterior. O argumento mostrado por eles em [KS] também descarta TVOC não-contextuais. A grande diferença entre a demonstração deles e a que aparece como consequência do lema de Gleason é que na segunda precisamos de todas as combinações lineares reais de três vetores LI no espaço de estados do sistema. Na demonstração de Kochen e Specker eles usam um conjunto finito de vetores. Mais precisamente, 117 vetores. Outros trabalhos apresentam demonstrações semelhantes com um número menor de vetores.

A prova de Kochen-Specker é feita em um espaço de estados de dimensão três. Novamente vamos assumir que se $v(P_\phi) = 1$ então $v(P_\psi) = 0$ para todo $|\psi\rangle$ ortogonal a $|\phi\rangle$. A ideia é encontrar um conjunto finito de vetores $\{|\phi_1\rangle, \ldots, |\phi_n\rangle\}$ tal que não seja possível atribuir valores 0 ou 1 a $v(P_{\phi_i})$ de maneira que essa restrição seja satisfeita.

Podemos representar um conjunto de vetores utilizando um diagrama de Kochen-Specker: cada vetor $|\phi_i\rangle$ no conjunto corresponde a um vértice em um grafo e dois vértices estarão ligados por uma aresta se os vetores correspondentes forem ortogonais. Vamos colorir os vértices do grafo de acordo com os valores associados a $v(P_{\phi_i})$. Se $v(P_{\phi_i}) = 1$ o vértice correspondente a $|\phi_i\rangle$ será colorido de vermelho; se $v(P_{\phi_i}) = 0$ o vértice correspondente a $|\phi_i\rangle$ será colorido de verde.

Pelo fato de estarmos em um espaço de dimensão três e pela condição de exclusividade, (14.6), se colorirmos um vértice de vermelho, então os vértices ligados a ele devem ser coloridos de verde e se em um triângulo dois vértices são coloridos de verde então o terceiro deve ser colorido de vermelho.

Agora o que devemos fazer é encontrar um diagrama de Kochen-Specker que não possa ser colorido dessa maneira. Isso prova o resultado que ficou conhecido como Teorema de Bell-Kochen-Specker[4].

[4] A demonstração do teorema foi feita por Kochen e Specker em [KS], mas a hipótese de não-contextualidade foi apontada por Bell em [Bel66] e por isso o

Teorema 14.3 (Teorema de Bell-Kochen-Specker)**.** *Não existe uma teoria de variáveis ocultas não-contextual compatível com a mecânica quântica.*

Não vamos entrar em detalhes da prova original, devido a sua complexidade. Exibiremos uma prova mais econômica em dimensão três e duas provas bem simples, uma em dimensão quatro e uma em dimensão oito com um número bem menor de vetores.

14.3.1 Uma Demonstração Econômica em Dimensão Três

Uma das provas mais simples do teorema de Bell-Kochen-Specker em dimensão três utiliza trinta e três vetores [Per91]. Para simplificar a notação, sejam $m = -1$ e $s = \sqrt{2}$. As trinta e três direções[5] desejadas, são definidas pelos seguinte vetores:

$$(1,0,0),\quad (0,1,1),\quad (0,1,s),\quad (s,1,1),$$

$$(0,m,1),\quad (0,m,s),\quad (s,m,1),\quad (s,m,m),$$

bem como as permutações das suas coordenadas.

Exercício 14.2. *Mostre que são definidas trinta e três direções no processo que acabamos de descrever.*

O conjunto acima possui duas propriedades importantes: é invariante por permutações dos eixos e por troca de sentido dos eixos. Isso permite que associemos o valor 1 a algumas direções arbitrariamente, sem perda de generalidade, uma vez que uma escolha diferente seria equivalente a uma troca de eixos ou de sentido em um dos eixos.

A tabela a seguir resume a demonstração do teorema BKS utilizando os trinta e três vetores. Para simplificar a notação, um vetor (a,b,c) será representado apenas por abc. Os vetores em cada linha da tabela são ortogonais. Aos vetores da primeira coluna é atribuído o valor 1 e, por consequência, aos vetores que aparecem nas outras

teorema ganha o nome dos três.

[5] O importante é o projetor sobre o subespaço gerado pelo vetor. Se dois vetores são múltiplos, eles geram o mesmo subespaço e por isso o projetor é o mesmo.

colunas deve ser atribuído o valor 0. A justificativa para a atribuição do valor 1 ao vetor da primeira coluna aparece na última coluna.

Trio			Vetores	$\perp$ ao 1°	Justificativa
001	100	010	110	$1m0$	Escolha do eixo z
101	$m01$	010			Escolha de sentido em x
011	$0m1$	100			Escolha de sentido em y
1ms	$m1s$	110	$s0m$	$0s1$	Troca entre x e y
10s	$s0m$	010	smm		O 2° e o 3° valem zero
s11	$01m$	smm	$m0s$		O 2° e o 3° valem zero
s01	010	$10s$	mms		O 2° e o 3° valem zero
11s	$1m0$	$11s$	$0sm$		O 2° e o 3° valem zero
01s	100	$0sm$	$1s1$		O 2° e o 3° valem zero
1s1	$10m$	$0sm$	msm		O 2° e o 3° valem zero
100	$0s1$	$01s$			CONTRADIÇÃO

Na tabela acima não são usados os trinta e três vetores. No entanto não podemos descartar os vetores que não apareceram. Eles são necessários porque devemos ter um conjunto invariante por troca de eixos e de sentido nos eixos para que as escolhas nos quatro primeiros passos possam ser feitas sem perda de generalidade.

14.3.2 Propriedades das Matrizes de Pauli

As duas próximas demonstrações vão depender fortemente de propriedades das matrizes de Pauli, (6.4). Vamos indicá-las a seguir:

Exercício 14.3. *1. Mostre que $\sigma_a\,\sigma_b = i\varepsilon^{abc}\sigma_c$, onde $a,b,c = x,y,z$ e $\varepsilon^{abc} = 1$, se abc é uma permutação par de xyz, $\varepsilon^{abc} = -1$, se abc é uma permutação ímpar de xyz e $\varepsilon^{abc} = 0$, se a,b,c não são distintos dois a dois;*

2. Conclua que $[\sigma_a, \sigma_b] = 2i\varepsilon^{abc}\sigma_c$;

3. Mostre que $[\sigma_a \otimes \sigma_a, \sigma_b \otimes \sigma_b] = 0$;

4. Da mesma forma, $[\sigma_a \otimes \sigma_b, \sigma_b \otimes \sigma_a] = 0$;

5. E, se a,b,c são dois a dois distintos, $[\sigma_a \otimes \sigma_a, \sigma_b \otimes \sigma_c] = 0$;

6. Mostre ainda que, se $[A_1, A_2] = 0$ e $[B_1, B_2] = 0$, então

$$[A_1 \otimes B_1, A_2 \otimes B_2] = 0.$$

14.3.3 Uma Demonstração Simples em Dimensão Quatro

Seja E o espaço de estados de um sistema de dois qbits. Vamos considerar os testes em E correspondentes aos nove operadores abaixo

$$\begin{array}{lll} A_1 = \sigma_x \otimes I & A_2 = I \otimes \sigma_x & A_3 = \sigma_x \otimes \sigma_x \\ A_4 = I \otimes \sigma_y & A_5 = \sigma_y \otimes I & A_6 = \sigma_y \otimes \sigma_y \\ A_7 = \sigma_x \otimes \sigma_y & A_8 = \sigma_y \otimes \sigma_x & A_9 = \sigma_z \otimes \sigma_z \end{array} \tag{14.7}$$

Vamos mostrar que não é possível atribuir valores definidos $v(A_i)$ que sejam independentes do conjunto de operadores compatíveis que são testados juntamente com A_i. Os operadores acima satisfazem as seguintes propriedades

1. Os três operadores em cada linha e em cada coluna comutam;
2. O produto dos operadores na coluna da direita é $-I$. O produto dos operadores nas outras duas colunas é I. O produto dos operadores em cada linha é I.

Como os valores atribuídos por uma TVOC a operadores que comutam devem satisfazer as mesmas identidades que os operadores satisfazem, a propriedade 2 requer que

$$P_1 = v(A_1)v(A_2)v(A_3) = 1 \tag{14.8a}$$
$$P_2 = v(A_4)v(A_5)v(A_6) = 1 \tag{14.8b}$$
$$P_3 = v(A_7)v(A_8)v(A_9) = 1 \tag{14.8c}$$
$$P_4 = v(A_1)v(A_4)v(A_7) = 1 \tag{14.8d}$$
$$P_5 = v(A_2)v(A_5)v(A_8) = 1 \tag{14.8e}$$
$$P_6 = v(A_3)v(A_6)v(A_9) = -1 \tag{14.8f}$$

Assim temos que

$$1 = P_1P_2P_3 = P_4P_5P_6 = -1$$

o que é uma contradição. Logo, não pode haver uma teoria de variáveis ocultas não-contextual compatível com a mecânica quântica.

Vale lembrar que nessa demonstração a não-contextualidade aparece ao assumirmos que $v(A_i)$ não muda se testamos A_i com os operadores que aparecem na mesma linha ou na mesma coluna.

14.3.4 Uma Demonstração Simples em Dimensão Oito

Dessa vez vamos trabalhar com o espaço de estados E de um sistemas de três qbits. Vamos considerar os testes em E correspondentes aos dez operadores abaixo

$$A_1 = \sigma_y \otimes I \otimes I$$

$$A_2 = \sigma_x \otimes \sigma_x \otimes \sigma_x \quad A_3 = \sigma_y \otimes \sigma_y \otimes \sigma_x \quad A_4 = \sigma_y \otimes \sigma_x \otimes \sigma_y \quad A_5 = \sigma_x \otimes \sigma_y \otimes \sigma_y$$

$$A_6 = I \otimes I \otimes \sigma_x \qquad A_7 = I \otimes I \otimes \sigma_y$$

$$A_8 = \sigma_x \otimes I \otimes I$$

$$A_9 = I \otimes \sigma_y \otimes I \qquad A_{10} = I \otimes \sigma_x \otimes I$$

Os operadores estão dispostos em cinco linhas de quatro operadores: $A_1A_3A_6A_9$, $A_1A_4A_7A_{10}$, $A_2A_3A_4A_5$, $A_2A_6A_8A_{10}$ e $A_5A_7A_8A_9$. Essas linhas formam uma estrela de cinco pontas. Valem as seguintes propriedades:

1. Os observáveis em cada linha da estrela comutam;
2. O produto dos observáveis em cada linha da estrela é I, exceto para a linha horizontal $A_2A_3A_4A_5$ em que o produto vale $-I$.

As propriedades acima implicam que

$$\begin{aligned}
P_1 = &\quad v(A_1)v(A_3)v(A_6)v(A_9) &&= 1, && (14.9a)\\
P_2 = &\quad v(A_1)v(A_4)v(A_7)v(A_{10}) &&= 1, && (14.9b)\\
P_3 = &\quad v(A_2)v(A_6)v(A_8)v(A_{10}) &&= 1, && (14.9c)\\
P_4 = &\quad v(A_5)v(A_7)v(A_8)v(A_9) &&= 1, && (14.9d)\\
P_5 = &\quad v(A_2)v(A_3)v(A_4)v(A_5) &&= -1. && (14.9e)
\end{aligned}$$

Como consequência, temos a contradição

$$-1 = P_1P_2P_3P_4P_5 = \prod_i v(A_i)^2 = 1.$$

14.4 Um Modelo de Variáveis Ocultas Contextual

Para construir um modelo de variáveis ocultas basta definir uma regra para encontrar os valores $v(P_\phi)$ atribuídos aos testes que correspondem a projetores P_ϕ. Isso ocorre porque todo operador auto-adjunto pode ser escrito como combinação de projetores que comutam:

$$A = \sum_i \lambda_i P_{\phi_i},$$

em que λ_i é o autovalor de A correspondente ao autovetor $|\phi_i\rangle$. Como podemos escolher os $|\phi_i\rangle$ ortogonais, podemos supor $[P_{\phi_i}, P_{\phi_j}] = 0$ e por isso

$$v(A) = \sum_i \lambda_i v(P_{\phi_i}).$$

Em [Bel66], é apresentado um modelo de variáveis ocultas contextual. Suponhamos que um aparato em questão testa os projetores $P_{\phi_1}, \ldots, P_{\phi_n}$ cujos valores esperados sejam $a_1, a_2 - a_1, a_3 - a_2, \ldots, a_n - a_{n-1}$. Como variável oculta tomamos um número real λ entre zero e um. O valor $v(P_{\phi_i}, \lambda)$ será dado pela regra

$$\begin{cases} v(P_{\phi_i}, \lambda) = 1 & \text{se } a_{i-1} < \lambda \leq a_i, \\ v(P_{\phi_i}, \lambda) = 0 & \text{caso contrário.} \end{cases}$$

Observe que os valores atribuídos a cada a_i dependem do conjunto de projetores em questão e não apenas de λ. É por essa razão que esse modelo é contextual.

Para mostrar que esse modelo é compatível com a mecânica quântica basta ver que

$$\langle P_{\phi_i} \rangle = \int_0^1 v(P_{\phi_i}, \lambda) d\lambda = a_i - a_{i-1}.$$

Apesar de artificial, o modelo acima mostra que é possível criarmos teorias de variáveis ocultas, desde que sejam contextuais. A discussão sobre variáveis ocultas surgiu quando algumas pessoas se sentiram incomodadas com o fato de que a mecânica quântica se comportava de maneira contra-intuitiva. O objetivo era recuperar a noção que temos em mecânica clássica de que todo observável físico possui um valor pré-definido, que existe independente do processo de medição e que é apenas revelado por ele. No entanto, para recuperar essa propriedade, devemos aceitar a contextualidade: o valor que um observável assume depende do conjunto

de observáveis que é testado com ele. Ficamos com um cobertor curto: se puxamos de um lado, perdemos do outro. Isso mostra que é impossível recuperar para a mecânica quântica as propriedades intuitivas do mundo clássico em que vivemos. Se a mecânica quântica estiver correta, e até agora não há nenhum indício que aponte o contrário, o comportamento do mundo microscópico é bem estranho[6], e não há nada que possamos fazer.

[6]Ou estranhos somos nós, que generalizamos uma série de preconcepções a partir de uma intuição moldada pela experiência clássica e tentamos aplicá-las a um domínio alheio.

Bibliografia

[.com] `http://www.idquantique.com` e `http://www.magiqtech.com` são bons exemplos.

[Ama] B. Amaral, *"Emaranhamento em sistemas de dois qubits,"* dissertação de mestrado, UFMG (2010). Disponível em `http://www.mat.ufmg.br/~tcunha/DisBarbara.pdf`

[Ara] M. Araújo Santos, *"Fundamentos matemáticos da separabilidade quântica,"* monografia de iniciação científica (2010). Disponível em `http://www.mat.ufmg.br/~tcunha/MonografiaMateus.pdf`

[AM] N.W. Ashcroft e N.D. Mermin, *"Solid State Physics,"* Brooks Cole (1976).

[Asp] A. Aspect *et al.*, "Experimental Tests of Realistic Local Theories via Bell's Theorem," *Phys. Rev. Lett.* **47**, 460 (1981).

[Bar] A. T. Baraviera, *"Introdução à Mecânica Quântica,"* 1º Colóquio de Matemática da Região Sul (2010). Disponível em `http://www.mat.ufmg.br/~tcunha/Baravi-ColSul.html`

[BLLT] A. Baraviera, C. F. Lardizabal, A. O. Lopes e M. Terra Cunha, "A Thermodynamic Formalism for Density Matrices in Quantum Information," *App. Math. Res. eXpress* **1**, 63 (2010).

[Bel64] J.S. Bell, *"On the Einstein Podolsky Rosen Paradox," Physics* **1**, 195 (1964). Reimpresso em [Bel87].

[Bel66] J. S. Bell, "On the problem of hidden variables in quantum mechanics," *Rev. Mod. Phys.* **38**, 447 (1966). Reimpresso em [Bel87].

[Bel87] J.S. Bell, *"Speakable and unspeakable in quantum mechanics,"* Cambridge University Press (1987).

[BŻ] I. Bengtsson e K. Życzkowski, *"Geometry of Quantum States. An Introduction to Quantum Entanglement,"* Cambridge University Press (2006).

[BB84] C.H. Bennett e G. Brassard, *Procceedings of International Conference on Computers, Systems, and Signal Processing*, Bangalore, India (1984).

[BDSW] C.H. Bennett, D.P. DiVincenzo, J.A. Smolin e W.K. Wootters, "Mixed-state entanglement and quantum error correction," *Phys. Rev. A* **54**, 3824 (1996).

[Cab] A. Cabello, *"Pruebas algebraicas da imposibilidad de variables ocultas en mecánica cuántica,"*, tese de doutorado, Madrid (1996). Disponível em `http://www.adancabello.com` ou em `http://www.mat.ufmg.br/~tcunha/Tese-Adan.html`

[CT] D. Cavalcanti e M. O. Terra Cunha, "Estimating entanglement on unknown quantum states," *App. Phys. Lett.* **89**, 084102 (2006).

[CSC+] D. Cavalcanti, P. L. Saldanha, O. Cosme *et al.*, "Geometrically induced singular behavior of entanglement," *Phys. Rev. A* **78**, 012318 (2008).

[CHSH] J.F. Clauser, M.A. Horne, A. Shimony e R.A. Holt, " Proposed Experiment to Test Local Hidden-Variable Theories," *Phys. Rev. Lett.* **23**, 880 (1969).

[CKW] V. Coffman, J. Kundu e W.K. Wootters, "Distributed entanglement," *Phys. Rev. A* **61**, 052306 (2000).

[CDL] C. Cohen-Tannoudji, B. Diu e F. Lalöe, *"Quantum Mechanics,"* Wiley-Interscience (2006).

[Coh] D. W. Cohen, *"An Introduction to Hilbert Space and Quantum Logic,"* Springer (1989).

[Cou] S.C. Coutinho, *"Números Inteiros e Criptografia RSA,"* IMPA (1997).

[DL] C. I. Doering, A.O. Lopes, *"Equações Diferenciais Ordinárias,"* Coleção Matemática Universitária, IMPA (2008).

[Dru] R.C. Drumond, *"Dinâmica de Emaranhamento e Geometria de Estados Quânticos,"* tese de doutorado, UFMG (2011). Disponível em `http://www.mat.ufmg.br/~tcunha/TeseRCDrumond.pdf`

[DVC] W. Dür, G. Vidal e J. I. Cirac, "Three qubits can be entangled in two inequivalent ways," *Phys. Rev. A* **62**, 062314 (2000).

[EPR] A. Einstein, B. Podolsky e N. Rosen, "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete," *Phys. Rev.* **47**, 777 (1935).

[EBA] J. Eisert, F.G.S.L. Brandão e K.M.R. Audenaert, "Quantitative entanglement witnesses," *New J. Phys.* **9**, 46 (2007).

[Exe] R. Exel, *"Uma introdução às C^*-álgebras."* Disponível em `www.mat.ufmg.br/~tcunha/RuyExel.html`

[FLS] R.P. Feynman, R.B. Leighton e M. Sands, *"The Feynman Lectures on Physics,"* vol. 3, Addison-Wesley publishing company (1965).

[Gle] A.M. Gleason, "Measures on the Closed Subspaces of a Hilbert Space," *J. Math. Mech.* **6**, 885 (1957).

[GHZ] D.M. Greenberger, M.A. Horne, A. Shimony e A. Zeilinger, "Bell's theorem without inequalities," *Am. J. Phys.* **58**, 1131 (1990).

[Har] J. Harris, *"Algebraic Geometry - A First Course,"* Springer-Verlag (1992).

[H$^{\otimes 2}$] R. e M. Horodecki, "Information-theoretic aspects of inseparability of mixed states," *Phys. Rev. A* **54**, 1838 (1996).

[H$^{\otimes 3}$96] M., P., e R. Horodecki, "Separability of mixed states: necessary and sufficient conditions," *Phys. Lett. A* **223**, 1 (1996).

[H$^{\otimes 3}$98] M., P., e R. Horodecki, "Mixed-state entanglement and distillation: is there a "bound" entanglement in nature?," *Phys. Rev. Lett.* **80**, 5239 (1998).

[JP] D. Jonathan e M.B. Plenio, "Entanglement-Assisted Local Manipulation of Pure Quantum State," *Phys. Rev. Lett.* **83**, 3566 (1999).

[KS] S. Kochen e E. Specker, "The Problem of Hidden Variables in Quantum Mechanics," *J. Math. Mech.* **17**, 59 (1967).

[Kra] K. Kraus, *"States, Effects and Operators: Fundamental Notions of Quantum Theory,"* Springer-Verlag (1983).

[LBe] Michel Le Bellac, *"Quantum Physics,"* Cambridge University Press (2006).

[LK] J. Lee e M.S. Kim, "Entanglement teleportation via Werner states," *Phys. Rev. Lett.* **84**, 4236 (2000).

[Lim] E.L. Lima, *"Álgebra Linear,"* Coleção Matemática Universitária, IMPA (2008).

[LP] C. Liu e N. Petulante, "On limiting distributions of quantum Markov chains," arXiv:1010.0741.

[Mer] D. Mermim, "Hidden variables and the two theorems of John Bell," *Rev. Mod. Phys.* **65**, 803 (1993).

[Nie] M.A. Nielsen, "Conditions for a Class of Entanglement Transformations," *Phys. Rev. Lett.* **83**, 436 (1999).

[NC] M.A. Nielsen e I.L. Chuang, *"Quantum Computation and Quantum Information,"* Cambridge University Press (2000).

[NAJ] J. Novotny, G. Alber e I. Jex, "Asymptotic Evolution of Random Unitary Operations," *Cent. Eur. J. Phys.* **8**, 1001 (2010).

[Per91] A. Peres, "Two simple proofs of the Kochen-Specker theorem," *J. Phys. A: Math. Gen.* **24**, L175 (1991).

[Per95] A. Peres, *"Quantum Theory: Concepts and Methods,"* Kluwer Academic Publishers (1995).

[Per96] A. Peres, "Separability criterion for density matrices," *Phys. Rev. Lett.* **76**, 1413 (1996).

[Pit] I. Pitowsky, "Quantum Probability - Quantum Logic," *Lect. Notes Phys.* **321**, 1 (1989).

[Pre] J. Preskill, *"A course on quantum computation,"* notas de aula. Disponíveis em
`http://www.mat.ufmg.br/~tcunha/Preskill.html`

[QA] M.T.C. Quintino e M. Araújo Santos, *"Desigualdades de Bell: Uma introdução à não-localidade quântica,"* (2010). Disponível em
`http://www.mat.ufmg.br/~tcunha/Bell-Mateus-MTulio.pdf`

[Rab] R.L.S. Rabelo, *"Não-localidade quântica: matemática e fundamentos,"* dissertação de mestrado, UFMG (2010). Disponível em
`http://www.mat.ufmg.br/~tcunha/DisRafael.pdf`

[SHK] A. Sawicki, A. Huckleberry e M. Kuś, "Symplectic Geometry of Entanglement," *Commun. Math. Phys.* **305**, 441 (2011).

[Sch] E. Schrödinger, "Discussion of Probability Relations between Separated Systems," *Math. Proc. Camb. Phil. Soc.* **31**, 555 (1935).

[Sin] S. Singh, *"The Code Book: the science of secrecy from ancient egypt to quantum cryptography,"* Anchor Books (1999).

[Soa] M.G. Soares, *"Cálculo de uma Variável Complexa,"* Coleção Matemática Universitária, IMPA (2009).

[Sot] J. Sotomayor, *"Lições de Equações Diferenciais Ordinárias,"* Projeto Euclides, IMPA (1979).

[Ter05] M.O. Terra Cunha, *"Emaranhamento: caracterização, manipulação e conseqüências,"* tese de doutorado, UFMG (2005). Disponível em
`http://www.mat.ufmg.br/~tcunha/TeseMTerraCunha.pdf`

[Ter07a] M. Terra Cunha, *"Noções de Informação Quântica,"* Monografias de Matemática, IMPA-SBM (2007).

[Ter07b] M.O. Terra Cunha, "The Geometry of Entanglement Sudden Death," *New J. Phys.* **9**, 237 (2007).

[TDV] M.O. Terra Cunha, J.A. Dunningham e V. Vedral, "Entanglement in single-particle systems," *Proc. Royal Soc. A* **463**, 2277 (2007).

[Thi] W. Thirring, *"Quantum Mathematical Physics: Atoms, Molecules and Large Systems,"* Springer (2002).

[Vai] I. Vainsencher, *"Notas de Aula de Álgebra Linear II,"* disponíveis em `http://www.mat.ufmg.br/~tcunha/Israel-AlgLin.html`

[VPRK] V. Vedral, M.B. Plenio, M.A. Rippin e P. L. Knight, "Quantifying Entanglement," *Phys. Rev. Lett.* **78**, 2275 (1997).

[VP] V. Vedral e M.B. Plenio, "Entanglement measures and purification procedures," *Phys. Rev. A* **57**, 1619 (1998).

[VADM] F. Verstraete, K. Audenaert, J. Dehaene e B. De Moor, "A comparison of the entanglement measures negativity and concurrence," *J. Phys. A: Math. Gen.* **34**, 10327 (2001).

[VT] G. Vidal e R. Tarrach, "Robustness of entanglement," *Phys. Rev. A* **59**, 141 (1999).

[Vid] G. Vidal, "Entanglement monotones," *J. Mod. Opt.* **47**, 355 (2000).

[VW] G. Vidal e R.F. Werner, "Computable measure of entanglement," *Phys. Rev. A* **65**, 032314 (2002).

[vNe] J. von Neumann, *"Mathematische Grundlagen der Quantenmechanik,"* Springer (1932); English translation: *"Mathematical Foundations of Quantum Mechanics,"* Princeton University Press (1955).

[WF] W.K. Wootters, B.D. Fields, "Optimal state-determination by mutually unbiased measurements," *Ann. Phys.* **191**, 363 (1989).

[Woo] W.K. Wootters, "Entanglement of Formation of an Arbitrary State of Two Qubits," *Phys. Rev. Lett.* **80**, 2245 (1998).

[ZLL] P. Zanardi, D. A. Lidar e S. Lloyd, "Quantum tensor product structures are observable induced," *Phys. Rev. Lett.* **92**, 060402 (2004).

[ŻHLS] K. Życzkowski, P. Horodecki, M. Lewenstein e A. Sanpera, "Volume of the set of separable states," *Phys. Rev. A* **58**, 883 (1998).

[Ż] K. Życzkowski, "Volume of the set of separable states. II," *Phys. Rev. A* **60**, 3496 (1999).