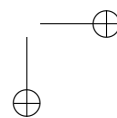
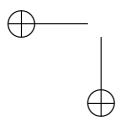
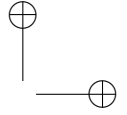
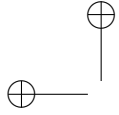


Introdução à Mecânica Quântica

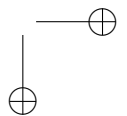
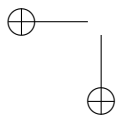
A. T. Baraviera
Instituto de Matemática - UFRGS



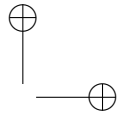
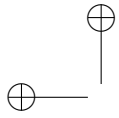


Conteúdo

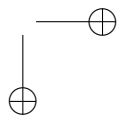
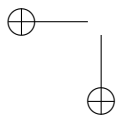
Conteúdo	i
Prefácio	1
1	2
1.1 Alguns aspectos históricos	2
1.2 Relações entre os mundos clássico e quântico	3
2	5
2.1 Espaços vetoriais	5
2.2 Subespaços vetoriais	6
2.3 Produto interno	7
2.4 Bases ortonormais	11
2.4.1 Ortogonalização de Gram-Schmidt	13
2.5 Mudança de base	13
2.6 Exercícios	14
3	16
3.1 Transformações lineares	16
3.2 Funcionais lineares	17
3.3 Operadores lineares	19
3.4 Adjunta de uma transformação linear	19
3.5 Projeção sobre um subespaço	21
3.6 Autovetores e autovalores	21
3.6.1 Autovalores e autovetores de transformações hermitianas	22

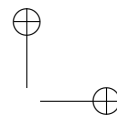
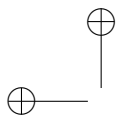


3.7	Traço e determinante	23
3.7.1	Traço	23
3.7.2	Determinante	24
3.8	Exponencial de uma matriz	25
3.9	Comutador de matrizes	28
3.10	Exercícios	29
4		31
4.1	Notação de Dirac	31
4.2	Usando esta notação	32
4.3	Exercícios	32
5		34
5.1	Equações diferenciais ordinárias lineares	34
5.2	Exercícios	35
6		37
6.1	Grupos	37
6.2	Grupos de matrizes	38
6.2.1	Matrizes unitárias	38
6.2.2	Matrizes ortogonais	39
6.3	Matrizes especiais	39
6.3.1	$SU(2)$	39
6.3.2	$SU(n)$	40
6.4	Ação de grupos	41
7		43
7.1	Sistema de dois níveis clássico	43
7.2	Sistema de dois níveis quântico	43
7.3	O significado de probabilidade	45
7.4	Algumas palavras sobre estatística	46
7.5	Observáveis e seu valor esperado	47
7.6	O princípio da incerteza	48
7.7	Evolução temporal	50
7.8	Evolução temporal após uma medição	51
7.9	Operadores hermitianos em \mathbb{C}^2	52
7.10	Exponenciais de σ_i	53
7.11	O spin de uma partícula num campo magnético	53



<i>CONTEÚDO</i>	iii
7.12 Exemplos	55
7.12.1 Exemplo 1	55
7.12.2 Exemplo 2	56
7.12.3 Exemplo 3	56
8	59
8.1 Aquecimento: sistemas com 3 e 4 níveis	59
8.2 N níveis	61
8.3 Operador posição	63
9	65
9.1 Dois sistemas de três níveis	65
9.2 Dois sistemas	66
9.3 Dois spins	67
10	68
10.1 Álgebras C^*	68
10.2 Estados de uma álgebra	71
10.3 Exercícios	72
11	74
11.1 Introdução	74
11.2 Matriz densidade	75
11.3 Redefinindo estado	76
11.3.1 Relação com os estados de uma álgebra C^*	77
11.4 Sistema de dois níveis	77
11.4.1 Caracterização geométrica	77
11.4.2 Um exemplo	78
11.4.3 Evolução temporal	79
12	81
12.1 Partícula clássica na reta	81
12.2 Partícula quântica	82
12.3 O operador Hamiltoniano e a equação de Schrödinger	84
12.4 A partícula em uma caixa unidimensional	86
12.4.1 Caso clássico	86
12.4.2 Caso quântico	87
12.4.3 Um exemplo de limite clássico	89

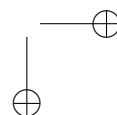
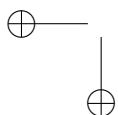


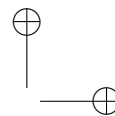
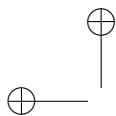


iv

CONTEÚDO

12.5 Exercícios	89
13	91
Bibliografia	93





Prefácio

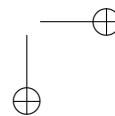
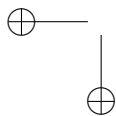
Estas notas foram escritas com o objetivo de apoiar o mini-curso de introdução à mecânica quântica que será ministrado no colóquio de matemática da região sul. A meta é pouco ambiciosa: pretendo apenas dar um pequeno primeiro passo para convidar alunos dos cursos de matemática a aprender um pouco mais sobre esta área tão importante e que pode, quando certa maturidade é atingida, servir como enorme estímulo e motivação para problemas matemáticos bem sofisticados. Ao mesmo tempo gostaria que servissem também para atrair alguns alunos dos cursos de física para o estudo de um pouco mais de matemática. A interação entre físicos e matemáticos sempre foi produtiva para ambos e espero que continue cada vez mais intensa.

Agradeço aos organizadores do colóquio pela oportunidade de oferecer este mini-curso e, especialmente, a Artur Lopes: pela insistência que fez estas notas verem a luz do dia e também por todas as conversas ao longo dos últimos anos sobre problemas relacionados ao conteúdo deste texto, que seguramente estão presentes nas entrelinhas.

Minha incompetência seguramente deixou inúmeros errinhos ao longo do texto; o leitor, como exercício, deve procurá-los e pode incomodar este autor pelo endereço

`baravi@mat.ufrgs.br`

Ficarei feliz se tiver pelo menos um leitor atento.

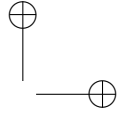
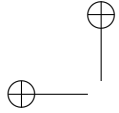


Capítulo 1

1.1 Alguns aspectos históricos

No final do século XIX a física vivia um momento de certa tranquilidade e alguns chegavam a propor que a tarefa dos futuros pesquisadores seria apenas a de efetuar medições mais precisas das constantes fundamentais das teorias que descreviam o mundo. Reconheciam que havia apenas duas pequenas lacunas, mas essas certamente seriam preenchidas em breve: uma dizia respeito a compreensão do experimento de Michelson e Morley, sobre a velocidade de propagação da luz em direções distintas (uma sendo a da translação da Terra e outra perpendicular à primeira) e a outra sobre a dedução teórica da curva de intensidade da radiação de corpo negro: a expressão de como um objeto escuro e aquecido emite radiação eletromagnética. Para surpresa e espanto de muitos, dessas duas pequenas lacunas surgiram, respectivamente, a teoria da relatividade e a teoria quântica, os dois pilares da física atual, conceitualmente muito distintos das teorias anteriores.

A primeira descrição teórica bem sucedida da radiação de corpo negro é devida ao físico Max Planck. Ele apresentou sua teoria em dezembro de 1900 e a principal inovação que permitiu o sucesso de sua formulação estava na hipótese de que a emissão ou absorção de luz (usando essa palavra como sinônimo de radiação eletromagnética) não se dava de forma contínua mas sim discreta, em pequenos pedaços chamados por ele de “quanta”(palavra que é o plural em latim de “quantum”).



1.2. RELAÇÕES ENTRE OS MUNDOS CLÁSSICO E QUÂNTICO³

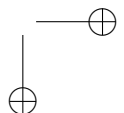
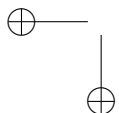
Esse sucesso encorajou seguidores, e Einstein então aplica a idéia para compreender o problema do efeito fotoelétrico, surgimento de corrente elétrica em determinados metais quando são expostos a luz. Essa explicação é um dos famosos artigos do ano maravilhoso de Einstein, 1905, e foi essencialmente por ela que ele ganhou o prêmio Nobel em 1921.

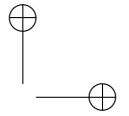
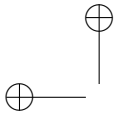
Anos depois Niels Bohr aplicou a hipótese quântica a um modelo atômico e assim conseguiu uma boa descrição de como um determinado elemento emite frequências específicas de luz quando devidamente excitado, fenômeno que já era bem conhecido no século XIX, mas que até então carecia de uma explicação teórica.

Esses trabalhos ajudaram a consolidar a hipótese quântica como algo efetivamente necessário para a compreensão do mundo físico, mas ainda não existia verdadeiramente uma mecânica quântica. Essa só começa a surgir com os trabalhos de Schrödinger, Heisenberg, Born e Dirac que desenvolvem as equações fundamentais da teoria e descobrem a base de sua formulação matemática. Algum tempo depois, J. von Neumann ajuda a fundamentar melhor o aspecto matemático dessas idéias e introduz, motivado por questões dessa teoria, o importante conceito de álgebra de operadores.

1.2 Relações entre os mundos clássico e quântico

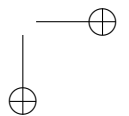
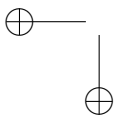
Dois são os rumos possíveis para tentar compreender as relações entre a mecânica quântica e a mecânica clássica: do macro para o micro, devemos tentar começar com a mecânica clássica e a partir dela tentar fabricar a mecânica para os objetos pequenos, ou seja, a mecânica quântica; essencialmente esse foi o caminho seguido pelos criadores da teoria e é o que normalmente se chama de quantização. O outro caminho é a passagem do micro para o macro: começando com a descrição quântica tentar recuperar, em alguma espécie de limite, o comportamento clássico (pois não é demais lembrar que um sistema clássico nada mais é que um aglomerado de sistemas quânticos em interação); isso é o que se chama de limite semi-clássico. Durante a fase pioneira da teoria Niels Bohr propõe o que se chama de “princípio da correspondência”: para encontrar-se as leis corretas da mecânica

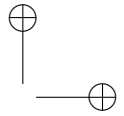
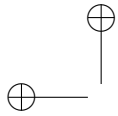




quântica é preciso procurar por aquelas que, no limite clássico, reproduzem o comportamento já conhecido pela mecânica. É preciso ressaltar que essas relações são sofisticadas e ainda são tema de muitas investigações.

Nessas notas tentaremos ser mais breves e pragmáticos: apresentaremos a quantização de sistemas bem simples, alguns que em geral não têm um bom análogo clássico, evitando as dificuldades dos caminhos mencionados acima. Com exceção de um capítulo, todos os espaços vetoriais que aparecem no texto são de dimensão finita e automaticamente são espaços de Hilbert, o que evitará a passagem por sutilezas da análise funcional. Esperamos que o leitor fique motivado e empreenda seu caminho por esse mundo mais rico assim que adquirir alguma maturidade.





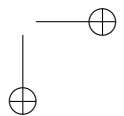
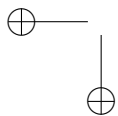
Capítulo 2

Neste capítulo pretendemos relembrar ao leitor algumas noções básicas sobre espaços vetoriais e produtos internos que serão muito utilizadas no decorrer do texto.

2.1 Espaços vetoriais

Um espaço vetorial sobre um corpo é um par (V, C) munido de duas operações, a soma vetorial $V \times V \rightarrow V$ e o produto por escalar $C \times V \rightarrow V$, que satisfazem as propriedades: para todo $x, y, z \in V$ e $\lambda, \nu \in C$ temos

- 1) (associatividade) $x + (y + z) = (x + y) + z$
- 2) (comutatividade) $x + y = y + x$
- 3) (existência de zero) Existe vetor $0 \in V$ tal que $x + 0 = 0$
- 4) (existência do vetor oposto) Dado $x \in V$ existe vetor $-x \in V$ tal que $x + (-x) = 0$
- 5) (associatividade) $\lambda(\nu x) = (\lambda\nu)x$
- 6) (distributividade) $\lambda(x + y) = \lambda x + \lambda y$
- 7) (distributividade) $(\lambda + \nu)x = \lambda x + \nu x$



8) $1x = x$ quando 1 é a unidade da multiplicação no corpo C

O leitor pode verificar sem muita dificuldade que os exemplos seguintes são de fato espaços vetoriais.

Exemplo 2.1.1. $(\mathbb{R}^n, \mathbb{R})$ onde \mathbb{R}^n é o conjunto das n -uplas ordenadas (x_1, \dots, x_n) , $x_i \in \mathbb{R}$, com a soma e produto por escalar definidos, respectivamente, por

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

$$\lambda(x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n)$$

Exemplo 2.1.2. $(C_{\mathbb{R}}[0, 1], \mathbb{R})$ onde $C_{\mathbb{R}}[0, 1]$ é o conjunto de funções contínuas do intervalo $[0, 1]$ com valores em \mathbb{R} , com soma e produto, respectivamente, definidos como sendo

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

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

Exemplo 2.1.3. $(\mathbb{C}^n, \mathbb{C})$ onde \mathbb{C}^n é o conjunto das n -uplas ordenadas (x_1, \dots, x_n) , $x_i \in \mathbb{C}$, com a soma e produto por escalar definidos, respectivamente, por

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

$$\lambda(x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n)$$

Neste 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 Subespaços vetoriais

Um subespaço vetorial S do espaço vetorial V é um subconjunto de V que é, ele mesmo, um espaço vetorial: para isso, precisamos que

$$0 \in S$$

2.3. PRODUTO INTERNO

7

$$u + v \in S \quad \text{para todo par } u \text{ e } v \text{ em } S$$

e, para todo λ no corpo C e todo $u \in S$ temos

$$\lambda v \in S$$

Exemplo 2.2.1. Considere o subconjunto de \mathbb{R}^n formado por todos os vetores da forma $S = \{(t, 0, \dots, 0) \mid t \in \mathbb{R}\}$. S então é um subespaço vetorial de \mathbb{R}^n .

Exemplo 2.2.2. Considere o subconjunto de ${}_{\mathbb{R}}C[0, 1]$ definido como sendo

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

Então também não é difícil ver que S é um subespaço vetorial de $C_{\mathbb{R}}[0, 1]$.

2.3 Produto interno

Dado um espaço vetorial V , um produto interno é uma aplicação $\langle \cdot \rangle : V \times V \rightarrow \mathbb{C}$ satisfazendo as seguintes propriedades: para todo $u, v, w \in V$ e $\lambda, \mu \in \mathbb{C}$

$$(i) \quad \langle \lambda u + \mu v, w \rangle = \bar{\lambda} \langle u, w \rangle + \bar{\mu} \langle v, w \rangle$$

$$(ii) \quad \langle u, v \rangle = \overline{\langle v, u \rangle}$$

$$(iii) \quad \langle u, u \rangle \geq 0$$

$$(iv) \quad \text{Se } \langle u, u \rangle = 0 \text{ então } u = 0.$$

Observação 2.3.1. Em (ii) a barra denota a operação de tomar o complexo conjugado do número; em (iii), note que o lado esquerdo da expressão de fato é real como consequência de (ii). Por último, note que

$$\begin{aligned} \langle u, \lambda v \rangle &= \overline{\langle \lambda v, u \rangle} = \overline{\lambda \langle v, u \rangle} = \\ &= \lambda \overline{\langle v, u \rangle} = \lambda \langle u, v \rangle \end{aligned}$$

Por isso uma maneira usual de reescrever a condição (i) acima é como

$$(i') \quad \langle u, \lambda v + \mu w \rangle = \lambda \langle u, v \rangle + \mu \langle u, w \rangle$$

Exemplo 2.3.1. Considere o espaço vetorial \mathbb{C}^n (sobre o corpo \mathbb{C}). Então a aplicação

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

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

Exemplo 2.3.2. $(C_{\mathbb{C}}[0, 1], \mathbb{C})$ onde $C_{\mathbb{C}}[0, 1]$ é o conjunto de funções contínuas do intervalo $[0, 1]$ com valores em \mathbb{C} , com soma e produto, respectivamente, definidos como sendo

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

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

Então podemos definir um produto interno como segue:

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

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

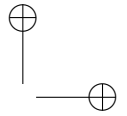
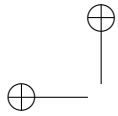
Definição 2.3.2. Dizemos que dois vetores u e v são **ortogonais** se $\langle u, v \rangle = 0$. Dizemos que um conjunto $E = \{v_1, \dots, v_k\}$ é **ortonormal** se seus elementos são, dois a dois, ortogonais e $\langle v_i, v_i \rangle = 1$ para todo i .

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

$$\langle (x_1, x_2, x_3), (y_1, y_2, y_3) \rangle = x_1 y_1 + x_2 y_2 + x_3 y_3$$

a ortogonalidade significa exatamente perpendicularidade no sentido geométrico usual.

Com um espaço vetorial V munido de um produto interno podemos definir uma aplicação $\| \cdot \|: V \rightarrow \mathbb{R}$ 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.



2.3. PRODUTO INTERNO

9

Teorema 2.3.3. *Se u e v são ortogonais, então*

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

Prova: Temos

$$\begin{aligned} \|u + v\|^2 &= \langle u + v, u + v \rangle = \langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle = \\ &= \|u\|^2 + \langle u, v \rangle + \overline{\langle u, v \rangle} + \|v\|^2 = \|u\|^2 + \|v\|^2 \end{aligned}$$

pois $\langle u, v \rangle = 0$

□

Observação 2.3.4. *Durante a prova obtivemos uma identidade conhecida como identidade polar:*

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

Corolário 2.3.5. *Usando o teorema acima o leitor pode provar, indutivamente, o seguinte resultado: se $\{v_1, \dots, v_k\}$ são dois a dois ortogonais, então*

$$\|v_1 + \dots + v_k\|^2 = \|v_1\|^2 + \dots + \|v_k\|^2$$

Teorema 2.3.6. *Seja $E = \{v_1, \dots, v_k\}$ subconjunto ortonormal de V . Então, para todo $v \in V$*

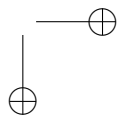
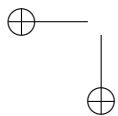
$$\|v\|^2 = \sum_{i=1}^k |\langle v, v_i \rangle|^2 + \|v - \sum_{i=1}^k \langle v, v_i \rangle v_i\|^2$$

Prova: Podemos escrever

$$v = \underbrace{\sum_{i=1}^k \langle v, v_i \rangle v_i}_a + \underbrace{v - \sum_{i=1}^k \langle v, v_i \rangle v_i}_b$$

Os vetores a e b são ortogonais: de fato $\langle a, b \rangle =$

□



Corolário 2.3.7. *Se $E = \{v_1, \dots, v_k\}$ é subconjunto ortonormal de V então para todo vetor $v \in V$ vale a desigualdade de Bessel:*

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

Corolário 2.3.8. *Dados u e v em V então vale a desigualdade de Cauchy-Bunyakovsky-Schwarz:*

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

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

$$\|u\|^2 \geq |\langle u, \frac{v}{\|v\|} \rangle|^2 = \frac{1}{\|v\|^2} |\langle u, v \rangle|^2$$

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

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

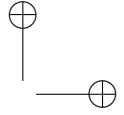
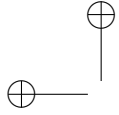
- i $\|\lambda u\| = |\lambda| \|u\|$
- ii $\|u + v\| \leq \|u\| + \|v\|$
- iii $\|u\| = 0 \Rightarrow u = 0$

Deixamos para o leitor as provas de (i) e (iii) e passamos a prova de (ii): temos

$$\begin{aligned} \|u + v\|^2 &= \|u\|^2 + 2\operatorname{Re}(\langle u, v \rangle) + \|v\|^2 \leq \|u\|^2 + 2|\langle u, v \rangle| + \|v\|^2 \leq \\ &\|u\|^2 + 2\|u\| \|v\| + \|v\|^2 = (\|u\| + \|v\|)^2 \end{aligned}$$

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 $d: V \times V \rightarrow \mathbb{R}_+$ tal que



2.4. BASES ORTONORMAIS

11

- i $d(u, v) = d(v, u)$
- ii $d(u, w) \leq d(u, v) + d(v, w)$
- iii $d(u, v) = 0 \Rightarrow u = v$

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

2.4 Bases ortonormais

Dizemos que uma expressão do tipo $\alpha_1 v_1 + \dots, \alpha_k v_k$ é uma combinação linear dos vetores v_1, \dots, v_k .

Chegamos então ao importante conceito de independência linear.

Definição 2.4.1. Dizemos que um conjunto de vetores $\{v_1, \dots, v_k\} \subset V$ é linearmente independente (LI) se a equação

$$\alpha_1 v_1 + \dots + \alpha_k v_k = 0$$

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

Caso contrário, se os vetores são linearmente dependentes (LD) então podemos expressar ao menos um dos vetores como combinação linear dos outros.

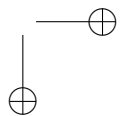
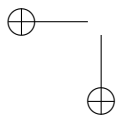
Podemos nos perguntar se existe um conjunto de vetores LI de forma que todo elemento do espaço V possa ser escrito como combinação linear dos elementos desse conjunto. Esse conjunto LI, se existe, é o que se chama de base de V .

Da existência de uma base $\mathcal{B} = \{v_1, \dots, v_n\}$ do espaço V surge a notação para vetores que usaremos em todo o restante do texto: dado um vetor v podemos escrevê-lo como $v = a_1 v_1 + \dots + a_n v_n$ e de forma única: de fato, se temos $v = b_1 v_1 + \dots + b_n v_n$ então

$$(a_1 - b_1)v_1 + \dots + (a_n - b_n)v_n = 0$$

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

Dado um espaço vetorial V munido de um produto interno \langle, \rangle então dizemos que uma base é ortogonal se ela é um subconjunto ortogonal de V .



Exemplo 2.4.1. A base canônica de \mathbb{C}^n : esse é o nome que se dá a base formada pelos vetores $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots, 0)$, \dots , $e_n = (0, 0, \dots, 0, 1)$.

Não é difícil notar que estes vetores são linearmente independentes: de fato a equação

$$0 = (0, 0, \dots, 0) = \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_n e_n = (\alpha_1, \alpha_2, \dots, \alpha_n)$$

de onde facilmente vem $\alpha_i = 0$ para todo $i = 1, \dots, n$.

Por outro lado também é fácil notar que qualquer vetor pode ser escrito como combinação dos elementos acima:

$$v \in \mathbb{C}^n \Rightarrow v = (v_1, v_2, \dots, v_n) = v_1 e_1 + \dots + v_n e_n$$

o que mostra que de fato o conjunto $\{e_1, e_2, \dots, e_n\}$ é uma base de \mathbb{C}^n .

Quando se adota o produto interno canônico

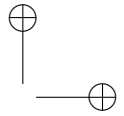
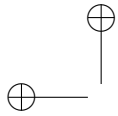
$$\langle (v_1, \dots, v_n), (u_1, \dots, u_n) \rangle := v_1 \bar{u}_1 + \dots + v_n \bar{u}_n$$

também não é difícil ver que esta base é ortonormal.

Exemplo 2.4.2. Consideremos o espaço \mathcal{P}_n dos polinômios de coeficientes complexos e grau menor ou igual a n . O leitor não terá dificuldade em verificar que o conjunto $\{1, x, \dots, x^n\}$ é uma base para este espaço.

Exemplo 2.4.3. Considere o espaço $M_n(\mathbb{C})$ de matrizes de ordem n e coeficientes complexos. Uma base para este espaço é dada pelas matrizes E_{ij} para $(i, j) \in \{1, 2, \dots, n\} \times \{1, 2, \dots, n\}$, 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

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



2.5. MUDANÇA DE BASE

13

2.4.1 Ortogonalização de Gram-Schmidt

Agora 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 $\{v_1, v_2, \dots, v_n\}$ de V então podemos obter uma base ortonormal $\{u_1, u_2, \dots, u_n\}$ por meio de um procedimento conhecido como *ortogonalização de Gram-Schmidt* que passamos a descrever. Para construir o vetor u_1 basta tomarmos

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

Para construir u_2 devemos ter em mente duas coisas: queremos que u_2 tenha norma unitária e que seja ortogonal ao vetor já construído u_1 . Para verificar essa segunda condição procuramos um vetor na forma $w_2 = v_2 + \alpha_1 u_1$ (que está no subespaço gerado por v_1 e v_2) de forma que $\langle w_2, u_1 \rangle = 0$. Então

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

ou seja,

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

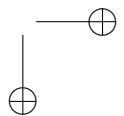
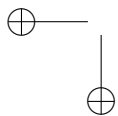
O vetor u_2 é então definido como sendo $u_2 = \frac{w_2}{\|w_2\|}$. Para obter u_3 procederemos de forma similar: primeiro procuramos $w_3 = v_3 + \alpha_1 u_1 + \alpha_2 u_2$ que deve ser ortogonal a u_1 e a u_2 , o que determina α_1 e α_2 como sendo

$$\alpha_1 = -\langle v_3, u_1 \rangle \quad \text{e} \quad \alpha_2 = -\langle v_3, u_2 \rangle$$

Seguindo dessa maneira podemos construir todos os vetores u_1, \dots, u_n ; por construção eles geram o mesmo espaço que v_1, \dots, v_n . São também ortonormais, sendo assim a base procurada do espaço V .

2.5 Mudança de base

Vamos agora abordar a questão de como escrever um certo vetor em bases distintas. Consideremos duas bases, $\mathcal{U} = \{u_1, \dots, u_n\}$ e



$\mathcal{V} = \{v_1, \dots, v_n\}$. Dado um vetor φ podemos escrevê-lo como

$$\varphi = a_1 u_1 + \dots + a_n u_n$$

que denotamos como $\varphi = [a_1, \dots, a_n]_{\mathcal{U}}$, onde os a_i são então as coordenadas de φ na base \mathcal{U} . Mas também podemos escrever $\varphi = b_1 v_1 + \dots + b_n v_n$, denotado por $[b_1, \dots, b_n]_{\mathcal{V}}$, que são as coordenadas de φ na base \mathcal{V} , e queremos obter a relação entre a_i e b_i .

Se

$$u_1 = T_{11}v_1 + T_{21}v_2 + \dots + T_{n1}v_n$$

e, analogamente,

$$u_i = T_{1i}v_1 + \dots + T_{ni}v_n$$

para todo $i \geq 2$, então

$$b = Ta$$

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

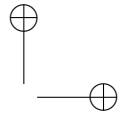
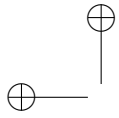
$$\begin{aligned} \varphi &= a_1 u_1 + \dots + a_n u_n = \\ &= a_1 \left(\sum_{k=1}^n T_{k1} v_k \right) + a_2 \left(\sum_{k=1}^n T_{k2} v_k \right) + \dots + a_n \left(\sum_{k=1}^n T_{kn} v_k \right) = \\ &= (T_{11}a_1 + T_{12}a_2 + \dots + T_{1n}a_n)v_1 + \dots + (T_{n1}a_1 + T_{n2}a_2 + \dots + T_{nn}a_n)v_n = \end{aligned}$$

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} .

2.6 Exercícios

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

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



2.6. EXERCÍCIOS

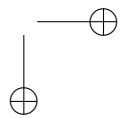
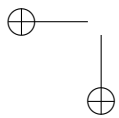
15

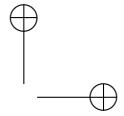
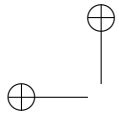
Verifique que o conjunto $\{1, x, \dots, x^n\}$ é uma base para este espaço. É ortogonal? Se não é, procure construir uma base ortogonal usando a técnica da última seção.

3- Mostre que o conjunto de matrizes 2×2 da forma

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

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





Capítulo 3

Neste capítulo recordamos uma classe especial de funções entre espaços vetoriais, as transformações lineares.

3.1 Transformações lineares

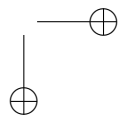
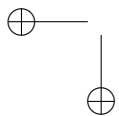
Sejam U e V espaços vetoriais. Uma aplicação $T: U \rightarrow V$ é dita uma transformação linear se satisfaz as seguintes condições: para todo $u, v \in U$ e $\lambda \in \mathbb{C}$ (que é o corpo que usaremos ao longo do texto) temos

- i $T(\lambda u) = \lambda T(u)$
- ii $T(u + v) = T(u) + T(v)$

Seja $L(U, V) = \{T: U \rightarrow V \text{ tal que } T \text{ é linear}\}$; se T e S são elementos de $L(U, V)$ então podemos definir as transformações $T + S$ e λT tais que

$$(T + S)(u) = T(u) + S(u) \quad \text{e} \quad (\lambda T)(u) = \lambda T(u)$$

Não é difícil verificar que $T + S$ e λT são ambas lineares, o que mostra que dessa forma dotamos o espaço $L(U, V)$ de uma estrutura de espaço vetorial.



3.2. FUNCIONAIS LINEARES

17

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

$$\|T\| = \sup_{|u|_1=1} |Tu|_2$$

onde $|\cdot|_1$ é uma norma em U e $|\cdot|_2$ é uma norma em V .

Quando fixamos uma base \mathcal{B} de V e \mathcal{F} de U podemos escrever um vetor $v \in V$ como sendo $v = \sum_{i=1}^k v_i e_i$, que também representamos na forma matricial

$$v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_k \end{bmatrix}_{\mathcal{B}} = [v_1, v_2, \dots, v_k]_{\mathcal{B}}^t$$

Então

$$T(v) = T\left(\sum_{i=1}^k v_i e_i\right) = \sum_{i=1}^k v_i T(e_i) = \sum_{i=1}^k \sum_j v_i T_{ji} \tilde{f}_j$$

Portanto podemos representar a transformação linear T por meio de uma matriz $T_{\mathcal{B}, \mathcal{F}}$ com entradas $\{T_{ij}\}_{ij}$ de forma que

$$[T(v)]_{\mathcal{F}} = T_{\mathcal{B}, \mathcal{F}} [v]_{\mathcal{B}}$$

De forma similar, dada uma matriz $k \times k$ temos, desde que fixadas as bases, uma transformação linear associada. Essa identificação é tão forte que frequentemente nos referiremos às transformações lineares apenas pela matriz que a representa, desde que não haja confusão acerca de qual é a base usada em cada caso.

3.2 Funcionais lineares

Um funcional linear L é uma transformação linear $L: V \rightarrow \mathbb{C}$. O espaço de todos os funcionais lineares de V é conhecido com o espaço dual de V e denotado por V^* . O leitor não terá dificuldade para verificar que V^* é também um espaço vetorial com a soma e produto por escalar previamente definidos.

Exemplo 3.2.1. Consideremos fixos um certo $v_0 \in V$, e um produto interno $\langle \cdot, \cdot \rangle$ em V . Então $L: V \rightarrow \mathbb{C}$ definido como sendo $L_{v_0}(v) = \langle v_0, v \rangle$ é um funcional linear.

De fato o exemplo acima é absolutamente geral: na verdade todo elemento de V^* pode ser escrito na forma de L_{v_0} para algum v_0 em V .

Teorema 3.2.1. Dado $L \in V^*$ então existe um único $v_0 \in V$ tal que $L(v) = \langle v_0, v \rangle$.

Prova: Considere uma base ortonormal $\{e_i\}_{i=1, \dots, k}$ de V . Então

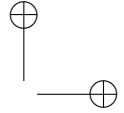
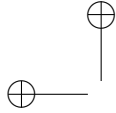
$$v = \sum_{i=1}^k v_i e_i$$

e

$$\begin{aligned} L(v) &= \sum_{i=1}^k v_i L(e_i) = \sum_{i=1}^k v_i L(e_i) \langle e_i, e_i \rangle = \\ &= \sum_{i,j=1}^k v_i L(e_j) \langle e_j, e_i \rangle = \sum_{i=1}^k v_i \sum_{j=1}^k \overline{\langle L(e_j) e_j, e_i \rangle} = \\ &= \sum_{i=1}^k v_i \left\langle \sum_{j=1}^k \overline{L(e_j)} e_j, e_i \right\rangle = \left\langle \sum_{j=1}^k \overline{L(e_j)} e_j, \sum_{i=1}^k v_i e_i \right\rangle = \\ &\qquad \qquad \qquad \langle v_0, v \rangle \end{aligned}$$

onde v_0 fica unicamente determinado como sendo $\sum_{j=1}^k \overline{L(e_j)} e_j$, o que conclui o teorema. \square

Observação 3.2.2. Este resultado simples é a versão em dimensão finita de um resultado bem mais geral da análise funcional conhecido como teorema de Riesz.



3.3. OPERADORES LINEARES

19

3.3 Operadores lineares

Estamos interessados no espaço $L(V, V)$, ou seja, nas transformações lineares de um espaço vetorial nele mesmo. Uma transformação $T \in L(V, V)$ 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 3.3.1. *Dados A e B em $L(V, V)$ então $\|AB\| \leq \|A\|\|B\|$*

Prova: Em primeiro lugar, note que se $v \neq 0$,

$$|Av| = |A \frac{v}{|v|}| |v| \leq \|A\| |v|$$

pois $v/|v|$ é um vetor unitário. Agora

$$\|AB\| = \sup_{|v|=1} |ABv| \leq \sup_{|v|=1} \|A\| |Bv| \leq \|A\| \|B\|$$

□

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

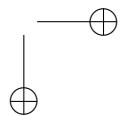
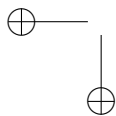
3.4 Adjunta de uma transformação linear

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

$$\langle Tv, u \rangle = \langle v, T^*u \rangle \quad \text{para todo } u \text{ e } v \text{ em } V$$

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

Teorema 3.4.1. *Dada uma transformação linear $T: V \rightarrow V$ então existe uma única transformação linear $T^*: V \rightarrow V$ tal que $\langle Tv, u \rangle = \langle v, T^*u \rangle$ para todo u e v em V .*



Prova: Considere $u \in V$ fixo. Vamos definir a aplicação $L: V \rightarrow \mathbb{C}$ 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 (que, naturalmente, depende de u) tal que

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

Como u_0 depende de u , escreveremos $u_0 = f(u)$.

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

$$\begin{aligned} L_{u+w}(v) &= \langle v, u + w \rangle = L_u(v) + L_w(v) = \\ &= \langle v, f(u) \rangle + \langle v, f(w) \rangle = \langle v, f(u) + f(w) \rangle \end{aligned}$$

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

Definição 3.4.2. Um operador linear tal que $A = A^*$ é chamado de *auto-adjunto*.

Nesse contexto essas matrizes também são chamadas de *hermitianas*.

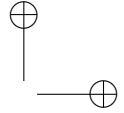
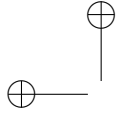
Como é a matriz associada a uma transformação linear auto-adjunta? Fixemos de início uma base, que pode ser a base canônica de \mathbb{C}^n . Então sabemos que o elemento de matriz a_{ij} é dado por

$$a_{ij} = \langle e_i, Ae_j \rangle$$

Então temos

$$a_{ij} = \langle A^* e_i, e_j \rangle = \langle Ae_i, e_j \rangle = \overline{\langle e_j, Ae_i \rangle} = \bar{a}_{ji}$$

Ou seja, a matriz A é tal que é igual a conjugação de sua transposta: $A = \bar{A}^T$.



3.5. PROJEÇÃO SOBRE UM SUBESPAÇO

21

3.5 Projecção sobre um subespaço

Dado $v \in V$ podemos definir a projecção sobre o subespaço vetorial gerado por v como sendo

$$P_v: V \rightarrow V \quad P_v(u) = \frac{\langle v, u \rangle}{\|v\|} v$$

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

$$\langle Px, y \rangle = \langle x, P^*y \rangle \quad \text{para todo } x \text{ e } y \text{ em } V$$

Mas então

$$\begin{aligned} \langle Px, y \rangle &= \frac{1}{\|v\|} \langle \langle v, x \rangle v, y \rangle = \frac{1}{\|v\|} \langle v, x \rangle \langle v, y \rangle = \\ &= \frac{1}{\|v\|} \langle x, v \rangle \langle y, v \rangle = \frac{1}{\|v\|} \langle x, \langle v, y \rangle v \rangle = \langle x, P_v(y) \rangle \end{aligned}$$

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

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

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

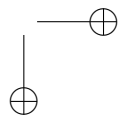
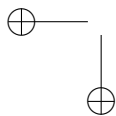
Normalmente isto é denotado simplesmente por $P_v^2 = P_v$ (e quando não ha confusão omite-se o subíndice v).

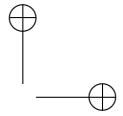
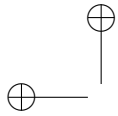
3.6 Autovetores e autovalores

Se $T: V \rightarrow V$ é uma transformação linear, então podemos procurar vetores não nulos satisfazendo a equação

$$T(v) = \lambda v \quad \text{para algum } \lambda \in \mathbb{C}$$

As soluções são conhecidas como autovetores (e, o respectivo λ , como autovalor) de T .





Observação 3.6.1. *E se vetores nulos fossem admitidos? 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 3.6.1. *Consideremos o caso de uma transformação linear P tal que $PP = P$ (o leitor consegue imaginar um exemplo?). Então a equação de autovalores é*

$$Pv = \lambda v$$

Mas

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

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.

3.6.1 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 3.6.2. *Se T é hermitiana então seus autovalores são reais.*

Prova: Considere $T(v) = \lambda v$. Então

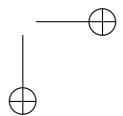
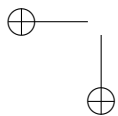
$$\langle Tv, v \rangle = \langle \lambda v, v \rangle = \bar{\lambda} \langle v, v \rangle$$

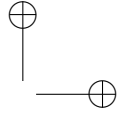
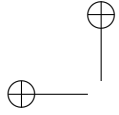
Por outro lado,

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

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

Podemos também investigar o que ocorre com os autovetores dessas transformações lineares.





3.7. TRAÇO E DETERMINANTE

23

Teorema 3.6.3. *Se T é hermitiana e v e u são dois autovetores associados, respectivamente, aos autovalores distintos λ e μ . Então u e v são ortogonais.*

Prova: 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 e v são ortogonais. \square

3.7 Traço e determinante

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

3.7.1 Traço

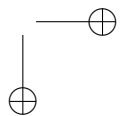
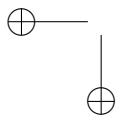
O traço de uma matriz quadrada A de elementos a_{ij} é definido como sendo a soma dos elementos da diagonal principal, ou seja,

$$\text{tr}(A) := \sum_{i=1}^n a_{ii}$$

Isso então definiu uma função $\text{tr}(\cdot) : M_n(\mathbb{C}) \rightarrow \mathbb{C}$; algumas de suas propriedades básicas estão condensadas no próximo

Lema 3.7.1. *Para todo $A, B \in M_n(\mathbb{C})$ e $\lambda \in \mathbb{C}$*

- a) $\text{tr}(A + B) = \text{tr}(A) + \text{tr}(B)$
- b) $\text{tr}(\lambda A) = \lambda \text{tr}(A)$
- c) $\text{tr}(AB) = \text{tr}(BA)$



Prova: A prova dos dois primeiros é bastante simples e é deixada ao leitor. Para verificarmos c) notemos que

$$\begin{aligned} \operatorname{tr}(AB) &= \sum_{i=1}^n (AB)_{ii} = \sum_{i=1}^n \sum_{k=1}^n A_{ik} B_{ki} \\ \sum_{i=1}^n \sum_{k=1}^n B_{ki} A_{ik} &= \sum_{k=1}^n \sum_{i=1}^n B_{ki} A_{ik} = \sum_{k=1}^n (BA)_{kk} = \operatorname{tr}(BA) \end{aligned}$$

□

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

3.7.2 Determinante

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

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

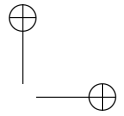
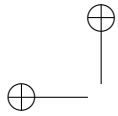
O leitor pode encontrar uma discussão bastante completa do caso geral, por exemplo, no livro de Elon Lima [7]. Podemos resumir suas principais propriedades no lema abaixo:

Lema 3.7.2. *det é uma função tal que:*

- a) $\det A^T = \det A$
- b) $\det \lambda A = \lambda^n \operatorname{tr}(A)$
- c) $\det AB = \det A \det B$

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 $v \neq 0$ tal que $Av = \lambda v$. esta expressão pode ser reescrita como sendo $(A - \lambda I)v = 0$



3.8. EXPONENCIAL DE UMA MATRIZ

25

e estamos procurando uma solução v 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 não nulo satisfazendo a equação isso significa que a transformação $A - \lambda I$ não é injetiva e portanto não admite inversa. Mas para não ter inversa isso significa que $\det A - \lambda I = 0$, que é uma equação polinomial em λ cujas raízes são exatamente os autovalores associados a transformação linear representada pela matriz A .

3.8 Exponencial de uma matriz

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

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

Podemos então tentar definir e^T como sendo

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

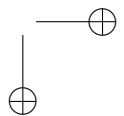
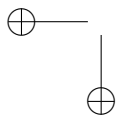
A expressão envolve soma de operadores lineares, composições de operadores lineares e o produto por numeros reais, todas essas operacoes que estao bem definidas para elementos de $L(V, V)$. Mas há uma passagem ao limite quando utilizamos a série e por isso devemos investigar com algum cuidado a questão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

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

Teorema 3.8.1. $L(V, V)$ com a norma de operador $\|\cdot\|$ é um espaço completo, isto é, sequencias de Cauchy são convergentes.

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

$$\|T_m - T_n\| < \epsilon \quad \text{para todo } m \text{ e } n \text{ maiores ou iguais a } N$$



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

$$|T_m(v) - T_n(v)| = |(T_m - T_n)(v)| \leq \|T_m - T_n\| |v|$$

o que mostra que $\{T_n(v)\}_{n \in \mathbb{N}} \subset V$ é uma sequência de Cauchy em V para a norma $|\cdot|$. Como V é um espaço completo essa sequência converge para um ponto de V que denotaremos por $T(v)$. Repetindo a idéia para cada ponto do espaço V conseguimos então definir uma função $T: V \rightarrow V$, $v \mapsto T(v)$. Devemos agora verificar que essa é de fato uma função linear. Para isso note que

$$T(v + u) = \lim T_n(v + u) = \lim T_n(v) + T_n(u) =$$

$$\lim T_n(v) + \lim T_n(u) = T(v) + T(u)$$

e o leitor não terá dificuldade em provar que $T(\lambda v) = \lambda T(v)$, mostrando que temos $T \in L(V, V)$. \square

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

$$T_n = e^T = \sum_{k=0}^n \frac{T^k}{k!}$$

Segundo o teorema 3.8.1 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

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

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,

3.8. EXPONENCIAL DE UMA MATRIZ

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 seqüência $\{T_n\}_n$ é uma seqüência de Cauchy, como desejado.

Lema 3.8.2. *Propriedades básicas da exponencial de T:*

a) se

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

b)

$$e^{QDQ^{-1}} = Qe^DQ^{-1}$$

c) Se T e S comutam, isto é, se $TS = ST$, então $e^{T+S} = e^T e^S$

d) $\det e^A = e^{\text{tr}(A)}$

Prova:

a) Note que se

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

então não é difícil ver que

$$DD = \begin{bmatrix} \lambda_1^2 & 0 & \dots & 0 \\ 0 & \lambda_2^2 & 0 & \dots \\ \vdots & \vdots & \ddots & 0 \end{bmatrix}$$

e,

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

Assim

$$e^D = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots \\ \vdots & \vdots & \ddots & 1 \end{bmatrix} + \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots \\ \vdots & \vdots & \ddots & 0 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} \lambda_1^2 & 0 & \cdots & 0 \\ 0 & \lambda_2^2 & 0 & \cdots \\ \vdots & \vdots & \ddots & 0 \end{bmatrix} + \cdots$$

$$+ \frac{1}{k!} \begin{bmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & 0 & \cdots \\ \vdots & \vdots & \ddots & 0 \end{bmatrix} + \cdots = \begin{bmatrix} e^{\lambda_1} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2} & 0 & \cdots \\ \vdots & \vdots & \ddots & 0 \end{bmatrix}$$

b)

$$e^{QDQ^{-1}} = I + QDQ^{-1} + \cdots + \frac{1}{k!} [QDQ^{-1}]^k + \cdots =$$

$$= QQ^{-1} + QDQ^{-1} + \cdots + \frac{1}{k!} QD^kQ^{-1} + \cdots = Q(I + D + \cdots + \frac{1}{k!} D^k + \cdots)Q^{-1} =$$

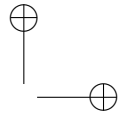
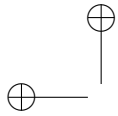
$$= Qe^DQ^{-1}$$

Para as provas de c) e d) sugerimos ao leitor o texto de Sotomayor [8]. \square

3.9 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 $AB = BA$; por exemplo, faça o teste com

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



3.10. EXERCÍCIOS

29

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

$$\text{tr}([A, B]) = 0$$

De fato

$$\text{tr}([A, B]) = \text{tr}(AB - BA) = \text{tr}(AB) - \text{tr}(BA) = 0$$

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,

$$\begin{aligned} [A, B]_{ij} &= (AB)_{ij} - (BA)_{ij} = \sum_k A_{ik}B_{kj} - \sum_l B_{il}A_{lj} = \\ &= \sum_k A_{ki}B_{jk} - \sum_l B_{li}A_{jl} = \sum_k B_{jk}A_{ki} - \sum_l A_{jl}B_{li} = \\ &= (BA)_{ji} - (AB)_{ji} = (BA - AB)_{ji} = [B, A]_{ji} = -[A, B]_{ij} \end{aligned}$$

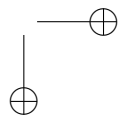
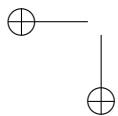
como desejado.

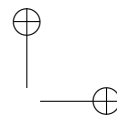
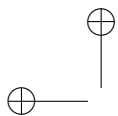
3.10 Exercícios

- 1- Mostre que $[A, [B, C]] + [C, [A, B]] + [B, [C, A]] = 0$ (que é conhecida como identidade de Jacobi).
- 2- Existem matrizes A e B satisfazendo a equação

$$AB - BA = I?$$

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

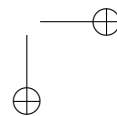
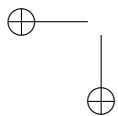


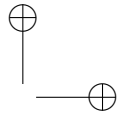
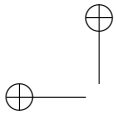


30

CAPÍTULO 3.

- 4- Mostre que se A e B são matrizes anti-simétricas então o comutador $[A, B]$ também é anti-simétrico.
- 5- Dado um espaço vetorial V , verificar que V^* é também um espaço vetorial.





Capítulo 4

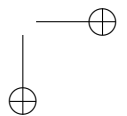
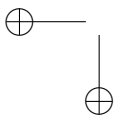
Neste capítulo introduzimos uma notação bastante empregada, sobretudo pelos físicos, para trabalhar com a mecânica quântica. Neste livro ela será usada apenas em algumas ocasiões, mas como é muito comum acreditamos que o leitor ganha algo tornando-se familiarizado com isso o mais cedo possível, o que certamente facilitará o seu acesso a literatura.

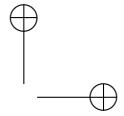
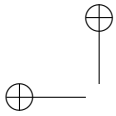
4.1 Notação de Dirac

Os objetos básicos da mecânica quântica são um espaço vetorial completo, ou seja, um espaço de Hilbert \mathcal{H} e os operadores lineares de \mathcal{H} em \mathcal{H} . A notação de Dirac tem por objetivo representar esses objetos: vetores, funcionais lineares e operadores. Uma das vantagens é que as propriedades algébricas se tornam praticamente óbvias pelo que é sugerido na notação.

Vamos então estabelecer o dicionário entre a notação usual e a de Dirac.

$|\psi\rangle$, chamado de ket, representa, na notação de Dirac, um vetor no espaço \mathcal{H} . É importante dizer que o ψ que aparece na notação é apenas um rótulo arbitrário. Uma notação comum, por exemplo, é $|0\rangle$ que não significa necessariamente o vetor nulo no espaço, mas sim um vetor com o rótulo 0, como por exemplo, v_0 .





$\langle\psi|$, chamado de bra, representa o elemento do dual de \mathcal{H} , o funcional linear $F_\psi(\cdot) = \langle\psi, \cdot\rangle$. Já sabemos que há uma identificação entre funcionais lineares e vetores, e esta notação reforça este fato.

Assim, $\langle\psi, \varphi\rangle$ representa a ação do funcional F_ψ sobre o vetor φ , que é o número complexo $\langle\psi, \varphi\rangle$.

Por fim é possível escrever $|\psi\rangle\langle\varphi|$ que representa um operador de \mathcal{H} em \mathcal{H} , o operador $T_{\psi,\varphi}(\cdot) = \langle\varphi, \cdot\rangle\psi$. No caso particular de $\varphi = \psi$ e com ψ unitário, ou seja, $\langle\psi, \psi\rangle = 1$, o operador de fato é uma projeção sobre o subespaço gerado por ψ e será denotado simplesmente por $E_\psi = T_{\psi,\psi}$.

4.2 Usando esta notação

A praticidade dessa notação fica evidente em muitos dos cálculos efetuados na mecânica quântica. Um exemplo simples é o seguinte: assumamos que temos dois vetores $|0\rangle$ e $|1\rangle$ que são ortogonais, ou seja, $\langle 0, 1\rangle = 0$ e estão normalizados de forma que $\langle 0, 0\rangle = \langle 1, 1\rangle = 1$. Então considere

$$(|0\rangle\langle 0| + |1\rangle\langle 1|)|0\rangle$$

O termo entre parenteses representa um operador linear e está sendo aplicado no vetor $|0\rangle$. Das regras operacionais que vimos temos então que a expressão acima é o mesmo que

$$|0\rangle\langle 0|0\rangle + |1\rangle\langle 1|0\rangle = |0\rangle$$

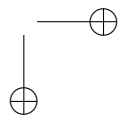
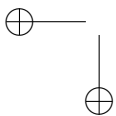
Esta operação é a equivalente, na notação de Dirac, a aplicar-se a operação linear representada pela matriz

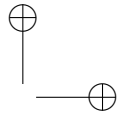
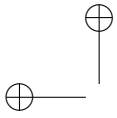
$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

ao vetor $(1, 0)$.

4.3 Exercícios

- 1- Você consegue imaginar o porque de se chamar $\langle\phi|$ de bra e $|\phi\rangle$ de ket? (Pense em $\langle\phi, \phi\rangle$ e procure um dicionário de inglês).





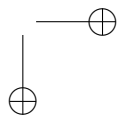
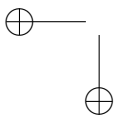
4.3. EXERCÍCIOS

33

2- Verifique as seguintes expressões:

$$\langle \alpha | (|\psi\rangle + |\varphi\rangle) = \langle \alpha, \psi \rangle + \langle \alpha, \varphi \rangle$$

$$(\lambda|\psi\rangle)^* = \bar{\lambda}\langle\psi| \quad \text{para todo } \lambda \in \mathbb{C}$$



Capítulo 5

Neste capítulo fazemos um breve estudo de equações diferenciais ordinárias lineares. Para um tratamento bastante completo do assunto o leitor deve consultar o livro de Sotomayor [8].

5.1 Equações diferenciais ordinárias lineares

Um campo vetorial em \mathbb{C}^n é uma aplicação $X: \mathbb{C}^n \rightarrow \mathbb{C}^n$; se X é uma aplicação linear então dizemos que o campo é linear e sabemos que podemos reescrever esta por meio de uma matriz A $n \times n$ com coeficientes complexos, de forma a ter $X(x) = Ax$.

Uma equação diferencial ordinária é uma equação na forma

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

onde $x: \mathbb{R} \rightarrow \mathbb{C}^n$ (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). Escrita dessa maneira essa equação pode ter mais de uma solução, mesmo quando se assume que o campo é bastante regular. Por exemplo, considere a equação

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

5.2. EXERCÍCIOS

35

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

Para eliminar essa ambiguidade assumiremos duas coisas: a linearidade do campo X (que não é fundamental para a unicidade mas simplificará nossa vida) e uma condição inicial $x(0) = x_0$ para a curva x . Desta forma, estamos interessados em resolver o problema de valor inicial (PVI) para a equação linear:

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

Afirmamos que $x(t) = e^{At}x_0$ (onde a exponencial é definida como na seção anterior) é solução do PVI. De fato, a primeira coisa a notar é que $x(0) = e^{A0}x_0 = e^0x_0 = x_0$, ou seja, a solução satisfaz a condição inicial. Precisamos então verificar que $x(t)$ satisfaz a equação diferencial ordinária. Note que

$$e^{At} = I + tA + \frac{1}{2!}t^2A^2 + \dots = \sum_{k \geq 0} \frac{1}{k!}t^kA^k$$

Portanto, se derivamos obtemos

$$\begin{aligned} \frac{d}{dt}e^{At} &= A + \frac{1}{2!}2tA^2 + \frac{1}{3!}3t^2A^3 + \dots = \\ &= A(I + tA + \frac{1}{2!}t^2A^2 + \dots) = Ae^{At} = e^{At}A \end{aligned}$$

(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 e 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.

5.2 Exercícios

1- Obtenha as exponenciais das seguintes matrizes:

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

(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 $e^{D+N} = e^D e^N$).

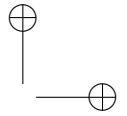
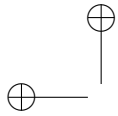
2- Resolva os seguintes problemas de valores iniciais:

a)

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

b)

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



Capítulo 6

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

6.1 Grupos

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

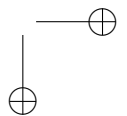
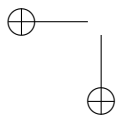
- i $ab \in G$
- ii $(ab)c = a(bc)$
- iii Existe $e \in G$ tal que $ea = ae = a$ para todo $a \in G$
- iv Para todo $a \in G$ existe $a^{-1} \in G$ tal que $aa^{-1} = a^{-1}a = e$

Se o grupo satisfaz a propriedade adicional

- v $ab = ba$

então dizemos que G é comutativo ou abeliano.

Exemplo 6.1.1. *O conjunto \mathbb{Z} munido da operação $+$, isto é, a adição usual, é um grupo comutativo.*



Exemplo 6.1.2. 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 6.1.3. O conjunto $BL(V, V)$ das transformações lineares bijetivas de V munido da operação de composição é um grupo. A aplicação identidade $id_V: V \rightarrow V$ faz o papel de unidade deste grupo.

6.2 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).

6.2.1 Matrizes unitárias

Uma matriz U é dita unitária se $U^*U = I$.

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

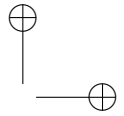
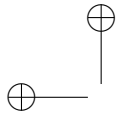
$$(UV)^*UV = V^*U^*UV = V^*IV = V^*V = I$$

mostrando que UV é de fato unitária.

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

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

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.



6.3. MATRIZES ESPECIAIS

39

6.2.2 Matrizes ortogonais

Uma matriz *real* O é dita ortogonal se $O^T O = Id$. 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.

6.3 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.

6.3.1 $SU(2)$

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:

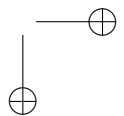
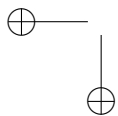
$$SU(2) = \{M \in M_2(\mathbb{C}) \mid M = \begin{bmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{bmatrix} \text{ e } |\alpha|^2 + |\beta|^2 = 1\}$$

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

Observação 6.3.1. *Há uma forte ligação entre este grupo e a esfera*

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

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),



escevendo $\gamma = x_1 + ix_2$ 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 ,

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

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

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

podemos naturalmente associar as matrizes

$$M_p = \begin{bmatrix} \gamma_1 & -\bar{\delta}_1 \\ \delta_1 & \bar{\gamma}_1 \end{bmatrix} \quad e \quad M_q = \begin{bmatrix} \gamma_2 & -\bar{\delta}_2 \\ \delta_2 & \bar{\gamma}_2 \end{bmatrix}$$

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

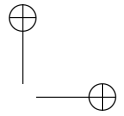
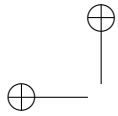
$$\begin{bmatrix} \gamma & -\bar{\delta} \\ \delta & \bar{\gamma} \end{bmatrix} = M_p M_q$$

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.

6.3.2 $SU(n)$

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



6.4. AÇÃO DE GRUPOS

41

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 [8]):

$$\det e^A = e^{\text{tr}(A)}$$

Então

$$\det U = \det e^{iH} = e^{\text{tr}(iH)} = e^{i\text{tr}(H)} = e^0 = 1$$

como desejado.

6.4 Ação de grupos

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

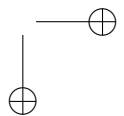
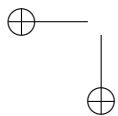
- 1 Para cada $g \in G$, $\phi(g, \cdot)$ é uma bijeção de X .
- 2 $\phi(e, x) = x$ (ou seja, $\phi(e, \cdot)$ é a aplicação identidade).
- 3 $\phi(g, \phi(h, x)) = \phi(gh, x)$.

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

Exemplo 6.4.1. *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 6.4.2. *Seja $G = (\mathbb{R}, +)$ e $X = \mathbb{R}^n$; se A é uma matriz $n \times n$ (que podemos pensar como sendo a representação de 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 a solução do problema de valor inicial para uma EDO linear quando o campo em \mathbb{R}^n é definido por $X(x) = Ax$.



Exercícios

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

2- 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).

3- Considerando o grupo do exercício anterior verifique que a aplicação

$$\phi: G \times \mathbb{R} \rightarrow \mathbb{R}$$

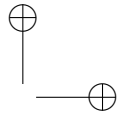
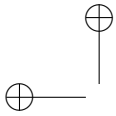
$$\phi\left(\begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}, x\right) = x + t$$

é uma ação de G sobre \mathbb{R} .

4- 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 7

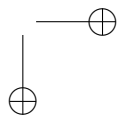
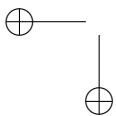
Neste capítulo vamos introduzir a mecânica quântica de sistemas de dois níveis.

7.1 Sistema de dois níveis clássico

Um sistema de dois níveis clássico é algum tipo de sistema que pode assumir, ao longo do tempo, apenas dois estados. Na mecânica estamos acostumados com variáveis contínuas e o leitor terá dificuldade em imaginar exemplos, mas podemos pensar em outras situações para ilustrar um comportamento desse tipo: por exemplo, podemos pensar na evolução do estado de uma lâmpada que pode apenas estar em uma das seguintes situações: $\{0, 1\}$, onde 0 denota a lâmpada apagada e 1, como é de se esperar, a lâmpada acesa. Habitualmente $\{0, 1\}$ é chamado de conjunto de estados. Sua evolução temporal é então uma função $x: \mathbb{R} \rightarrow \{0, 1\}$ e obter essa função é, tipicamente, o que deseja alguém que pretende descrever matematicamente esse sistema.

7.2 Sistema de dois níveis quântico

A descrição quântica de um sistema de dois níveis reúne duas características interessantes: é bastante simples e, por outro lado, já revela



muitas das diferenças de tratamento com relação ao caso clássico. Por essa razão nos deteremos longamente em exemplos desse tipo.

Mas há razões mais profundas: se no caso clássico o leitor pode ter ficado decepcionado com o exemplo da lâmpada e o considera muito artificial (tendo, se isto serve de consolo, a companhia deste autor), no caso quântico sistemas de dois níveis aparecem com muito mais naturalidade e são importantes na descrição, por exemplo, do comportamento dos elétrons em cada átomo.

Nossa primeira preocupação é encontrar o espaço de estados e já aqui surge uma grande diferença. Podemos começar da seguinte maneira: para cada um dos dois estados clássicos, que por simplicidade denotamos por 0 e 1, associamos um vetor, respectivamente v_0 e v_1 ; estes são uma base de um espaço vetorial complexo. Bem, um espaço vetorial complexo de dimensão 2 é \mathbb{C}^2 . Então a primeira tentativa é a de se chamar de estado um vetor $\psi \in \mathbb{C}^2$, que pode ser escrito como

$$\psi = \alpha v_0 + \beta v_1 \quad \alpha, \beta \in \mathbb{C}$$

Porém é preciso fornecer uma interpretação para isso e aqui também encontramos algo bem diferente do clássico: quando se efetua uma medição no sistema com o intuito de saber se ele se encontra no estado 0 ou 1, a probabilidade de se encontrar 0 é dada por $|\alpha|^2$ e a probabilidade de se encontrar 1 é, como se pode imaginar, $|\beta|^2$. Por exemplo, se temos o estado

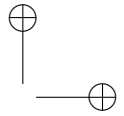
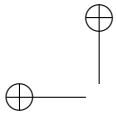
$$\omega = \frac{1}{\sqrt{5}}v_0 + \frac{2}{\sqrt{5}}v_1$$

então uma medição desse tipo pode encontrar o estado 0 com probabilidade $1/5$.

Portanto essa interpretação já impõe um vínculo ao que podemos chamar de estado, pois é preciso que $|\alpha|^2 + |\beta|^2 = 1$. Se escrevemos $\alpha = x_1 + ix_2$ e $\beta = x_3 + ix_4$, com todos os x_i reais, então essa condição é equivalente a $x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1$, que é um subconjunto de \mathbb{R}^4 conhecido como S^3 , a versão da esfera com três dimensões (e que vive dentro de um espaço de quatro dimensões).

Uma outra maneira prática, em geral bem mais usada, de reescrever ψ de forma a obedecer a restrição acima é a seguinte:

$$\psi = \left(e^{i\gamma} \cos\left(\frac{\theta}{2}\right), e^{i(\gamma+\phi)} \sin\left(\frac{\theta}{2}\right) \right) \quad (7.1)$$



7.3. O SIGNIFICADO DE PROBABILIDADE

45

com γ, ϕ e θ reais.

Seria bonito termos o conjunto S^3 como o espaço de estados, mas podemos ir além. Se agora trocamos ψ por

$$\varphi = e^{i\theta}\psi = e^{i\theta}\alpha v_0 + e^{i\theta}\beta v_1 \quad \theta \in \mathbb{R}$$

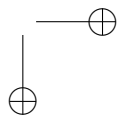
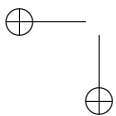
o que ocorre? Segundo a interpretação acima, a probabilidade de se obter 0 numa medição do estado φ será $|e^{i\theta}\alpha|^2 = |\alpha|^2$ e, de forma análoga, a probabilidade de se obter 1 será $|e^{i\theta}\beta|^2 = |\beta|^2$. Ou seja, as probabilidades serão as mesmas tanto para ψ quanto para φ , o que torna esses dois vetores fisicamente indistinguíveis e portanto ambos representam o *mesmo estado*. Logo, nosso espaço de estados precisa identificar dois vetores de \mathbb{C}^2 quando um é múltiplo de outro por um fator $e^{i\theta}$. De maneira mais concreta, isso significa que podemos ignorar o fator comum $e^{i\gamma}$ na expressão 7.1 e escrever um estado como sendo simplesmente

$$\psi = \left(\cos\left(\frac{\theta}{2}\right), e^{i\phi}\sin\left(\frac{\theta}{2}\right) \right) \quad (7.2)$$

Portanto o espaço de estados de um sistema de dois níveis pode ser completamente descrito por meio de dois ângulos θ e ϕ , $\theta \in [0, \pi]$, $\phi \in [0, 2\pi]$, que de fato podem ser vistos como latitude e longitude sobre a esfera usual S^2 . Esta esfera é conhecida como esfera de Bloch. O leitor então pode constatar que o espaço de estados, que de início parecia ser um objeto de quatro dimensões (reais, pois para descrevermos \mathbb{C}^2 precisamos de quatro números reais) acaba por ser algo bidimensional; no entanto isto já é infinitamente mais rico que a descrição clássica de um sistema de dois níveis, que comporta apenas um espaço com dois *pontos*.

7.3 O significado de probabilidade

O fato de somente se poder prever as probabilidades de obtenção de um determinado resultado torna a mecânica quântica incômoda para muitas pessoas, por ser um abandono da visão determinista que permeava a mecânica clássica (e de fato, toda a ciência). Se o leitor também sente este mal-estar talvez console saber que está



em companhia de ninguém menos que Albert Einstein, que espessou dessa forma sua insatisfação: ”Deus não joga dados!”.

Bem, não queremos entrar em discussões teológicas, mas sim pensar melhor o que significa dizer que a probabilidade de se encontrar 0 ao efetuar uma medição sobre o estado

$$\omega = \frac{1}{\sqrt{5}}v_0 + \frac{2}{\sqrt{5}}v_1$$

é de $1/5$. Uma primeira tentativa seria dizer que se efetuarmos diversas medições sobre esse objeto então, a longo prazo, a frequência de se obter 0 é de $1/5$. Essa é, por exemplo, a interpretação que podemos dar a uma frase como ”a probabilidade de sair cara num lançamento de uma moeda honesta é de $1/2$ ”. No entanto no caso quântico essa interpretação não é possível pela seguinte razão: ao se efetuar a medição o estado do sistema muda. Por exemplo, se a medição encontrou o valor 0 então o estado mudou de ω para v_0 , logo não faz sentido repetir as medições com o mesmo objeto já que ele não está mais no mesmo estado.

Por isso a forma de interpretar é a seguinte: consideramos uma grande quantidade de objetos, todos no estado ω . Efetuando medições separadas em cada um deles (que é medido apenas uma vez) então a frequência de medições com resultado 0 se aproxima de $1/5$.

7.4 Algumas palavras sobre estatística

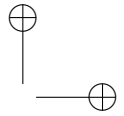
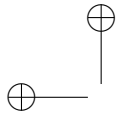
Quando temos uma variável aleatória x que pode assumir os valores $\{x_1, \dots, x_n\}$ com probabilidades, respectivamente, $\{p_1, \dots, p_n\}$ podemos definir o valor médio de x , ou valor esperado, como sendo

$$\langle x \rangle = \sum_{i=1}^n p_i x_i$$

Ou seja, uma média ponderada dos possíveis valores da variável x .

Dado $\lambda \in \mathbb{R}$, podemos considerar a variável aleatória λx , que assume os valores $\{\lambda x_1, \dots, \lambda x_n\}$ com probabilidades $\{p_1, \dots, p_n\}$. Nesse caso não é difícil ver que $\langle \lambda x \rangle = \lambda \langle x \rangle$.

Podemos também definir a variância, que é uma maneira de saber se os sorteios de x ficam todos próximos de um certo valor (que é o



7.5. OBSERVÁVEIS E SEU VALOR ESPERADO

47

valor médio de x) ou não. Uma maneira natural de avaliar isto é fazendo a seguinte média:

$$\Delta x = \sqrt{\langle (x - \langle x \rangle)^2 \rangle}$$

Mas de fato temos

$$\begin{aligned} \langle (x - \langle x \rangle)^2 \rangle &= \langle x^2 - 2x\langle x \rangle + \langle x \rangle^2 \rangle = \langle x^2 \rangle - 2\langle x \rangle \langle x \rangle + \langle x \rangle^2 = \\ &= \langle x^2 \rangle - \langle x \rangle^2 \end{aligned}$$

e assim

$$\Delta x = \sqrt{\langle x^2 \rangle - \langle x \rangle^2}$$

Este número é sempre maior ou igual a zero (e está sempre bem definido, pois estamos calculando a média de um quadrado que é algo não negativo) e também é chamado de dispersão da variável aleatória.

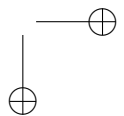
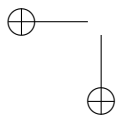
7.5 Observáveis e seu valor esperado

Um observável é uma característica do sistema em estudo que pode ser registrada por meio de uma medição. Cada observável está associado a um operador hermitiano em $M_2(\mathbb{C})$: os estados que podem ser encontrados por meio de uma medição do observável $A \in M_2(\mathbb{C})$, $A = A^*$, são os autovetores de A . Os valores da medição são os autovalores correspondentes, que no caso de operadores hermitianos, são reais. Se supomos que esses autovalores são distintos, então já sabemos que os respectivos autovetores são ortogonais, e portanto são uma base ortonormal (lembre que sempre podemos supor os autovetores normalizados) de \mathbb{C}^2 . Dessa forma, podemos representar um estado qualquer como combinação linear dos autovetores de A , que denotaremos v_1 e v_2 , com autovalores λ_1 e λ_2 . Um estado ψ então pode ser escrito na forma

$$\psi = \alpha_1 v_1 + \alpha_2 v_2$$

Podemos então perguntar qual o valor esperado em uma medição do observável A no estado ψ : bem, podemos obter o valor λ_i com probabilidade $|\alpha_i|^2$. Logo, o valor esperado sera

$$\langle A \rangle_\psi = \lambda_1 |\alpha_1|^2 + \lambda_2 |\alpha_2|^2 = \bar{\alpha}_1 \lambda_1 \alpha_1 + \bar{\alpha}_2 \lambda_2 \alpha_2 =$$



$$\begin{aligned} \bar{\alpha}_1 \lambda_1 \alpha_1 \langle v_1, v_1 \rangle + \bar{\alpha}_2 \lambda_2 \alpha_2 \langle v_2, v_2 \rangle &= \langle \alpha_1 v_1, \lambda_1 \alpha_1 v_1 \rangle + \langle \alpha_2 v_2, \lambda_2 \alpha_2 v_2 \rangle = \\ &= \langle \alpha_1 v_1, A(\alpha_1 v_1) \rangle + \langle \alpha_2 v_2, A(\alpha_2 v_2) \rangle = \langle \psi, A\psi \rangle \end{aligned}$$

Desta forma temos uma expressão simples para obter o valor esperado de um observável num determinado estado e já sabemos que o mesmo será real, pois

$$\langle A \rangle_\psi = \langle \psi, A\psi \rangle = \langle A\psi, \psi \rangle = \overline{\langle \psi, A\psi \rangle} = \overline{\langle A \rangle_\psi}$$

7.6 O princípio da incerteza

Considere um estado Ψ (normalizado); vamos assumir que temos dois observáveis A e B , ambos com média zero para este estado (isto 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). Logo, $\Delta A = \langle (A^2) \rangle$. Vamos também assumir que

$$[A, B] = AB - BA \neq 0$$

ou seja, nossos observáveis não comutam.

Teorema 7.6.1 (Princípio da Incerteza de Heisenberg). *Sejam A e B dois observáveis de média zero e tais que $[A, B] \neq 0$. Então*

$$\Delta A \Delta B \geq \frac{1}{2} |\langle \Psi, [A, B] \Psi \rangle|.$$

Prova:

$$\begin{aligned} \Delta A^2 \Delta B^2 &= \langle \Psi, A^2 \Psi \rangle \langle \Psi, B^2 \Psi \rangle \\ &= \langle A\Psi, A\Psi \rangle \langle B\Psi, B\Psi \rangle \\ &= \|A\Psi\|^2 \|B\Psi\|^2 \\ &\geq |\langle A\Psi, B\Psi \rangle|^2 \end{aligned}$$

onde, na última passagem, foi utilizada a desigualdade de Cauchy-Bunyakovsky-Schwarz; logo, temos $\Delta A \Delta B \geq |\langle A\psi, B\psi \rangle|$. Note que

7.6. O PRINCÍPIO DA INCERTEZA

49

$$\begin{aligned}
 \langle A\Psi, B\Psi \rangle &= \langle \Psi, AB\Psi \rangle \\
 &= \langle \Psi, ([A, B] + BA)\Psi \rangle \\
 &= \langle \Psi, [A, B]\Psi \rangle + \langle \Psi, BA\Psi \rangle \\
 &= \langle \Psi, [A, B]\Psi \rangle + \overline{\langle BA\Psi, \Psi \rangle} \\
 &= \langle \Psi[A, B], \Psi \rangle + \overline{\langle A\Psi, B\Psi \rangle}
 \end{aligned}$$

Portanto temos

$$\langle A\Psi, B\Psi \rangle - \overline{\langle A\Psi, B\Psi \rangle} = 2Im(\langle A\Psi, B\Psi \rangle) = \langle \Psi, [A, B]\Psi \rangle.$$

Logo,

$$Im(\langle A\Psi, B\Psi \rangle) = \frac{1}{2} \langle \Psi, [A, B]\Psi \rangle.$$

Assim,

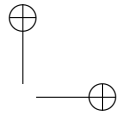
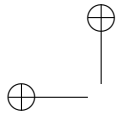
$$\Delta A \Delta B \geq |\langle A\Psi, B\Psi \rangle| \geq |Im(\langle A\Psi, B\Psi \rangle)| = \frac{1}{2} |\langle \Psi, [A, B]\Psi \rangle|.$$

□

Observação 7.6.2. *O leitor deve notar que absolutamente nada de bidimensional apareceu na prova deste resultado: a demonstração vale para operadores em espaços vetoriais de dimensão qualquer, desde que, claro, sejam válidas as mesmas hipóteses.*

Observação 7.6.3. *Por razões históricas preferimos manter a terminologia “princípio da incerteza” para este conhecido resultado, ainda que nesse contexto ele nada tenha de princípio: é uma consequência de propriedades prévias dos objetos que estamos discutindo. Por este motivo muitos autores defendem a denominação “relação de incerteza”.*

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; se diminuimos muito a dispersão de A , por exemplo, então a de B aumenta muito e



será grande a incerteza sobre o valor desse observável no estado em questão, o que 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 e B num determinado estado ψ .

7.7 Evolução temporal

A evolução temporal é uma descrição de como o estado de um certo sistema varia em função do tempo. A mecânica quântica tem uma evolução temporal relativamente simples, sendo dada por uma família de operadores lineares que depende do tempo da seguinte forma:

$$U(t) = e^{-itH}$$

onde H é um operador hermitiano. Como já observado antes, $U^*(t) = e^{itH}$ e então segue que este operador U é unitário para qualquer t real.

O operador H (e, em ultima instancia, U) traduz matematicamente a situação física do sistema em estudo, que é o que determina a sua evolução no tempo. Mais tarde trataremos da situação do spin, onde poderemos identificar H de maneira bastante concreta num contexto físico.

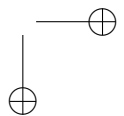
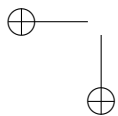
De posse de $U(t)$ podemos então escrever o estado em um instante qualquer:

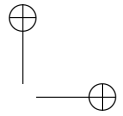
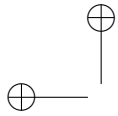
$$\psi(t) = U(t)\psi = e^{-itH}\psi_0 \tag{7.3}$$

Podemos também descrever a evolução temporal do valor esperado de um certo observável A :

$$\begin{aligned} \langle A \rangle(t) &= \langle \psi(t), A\psi(t) \rangle = \langle e^{-itH}\psi_0, Ae^{-itH}\psi_0 \rangle = \\ &= \langle \psi_0, e^{itH}Ae^{-itH}\psi_0 \rangle = \langle \psi_0, A(t)\psi_0 \rangle \end{aligned}$$

onde denotamos por $A(t) = e^{itH}Ae^{-itH}$ uma família de operadores que coincide com A para $t = 0$. A expressão $\langle \psi(t), A\psi(t) \rangle$ é conhecida como a representação de Schrödinger do valor esperado, e a ênfase é dada a evolução temporal do estado ψ sendo que o operador A





7.8. EVOLUÇÃO TEMPORAL APÓS UMA MEDIÇÃO 51

está fixo; a expressão $\langle \psi_0, A(t)\psi_0 \rangle$ é conhecida como representação de Heisenberg, e a ênfase é dada a evolução temporal do operador A sendo que o estado é que está fixo agora. Naturalmente ambas se equivalem, mas dependendo do problema em consideração uma pode ser mais conveniente do que a outra.

Na representação de Schrödinger, podemos verificar qual é a equação obedecida por $\psi(t)$. Se derivamos 7.3 obtemos

$$\frac{d}{dt}\psi(t) = -iH\psi(t)$$

Esta equação diferencial poderia ter sido usada como ponto de partida, e de fato foi isso que fez Schrödinger.

Já na representação de Heisenberg devemos encontrar uma equação diferencial para a evolução do operador A . Derivando a expressão para $A(t)$ obtemos o seguinte:

$$\begin{aligned} \frac{d}{dt}A(t) &= iHe^{itH}Ae^{-itH} + e^{itH}A(-iH)e^{-itH} = \\ & iHe^{itH}Ae^{-itH} - ie^{itH}Ae^{-itH}H = \quad (\text{pois } H \text{ comuta com } e^{-itH}) \\ & iHA(t) - iA(t)H = i[H, A(t)] \end{aligned}$$

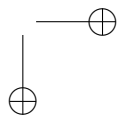
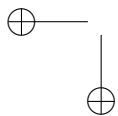
(onde $[X, Y] = XY - YX$ é o comutador das matrizes X e Y). Ou seja, nessa representação podemos usar como ponto de partida a equação diferencial para o operador

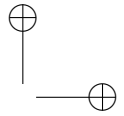
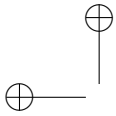
$$\frac{d}{dt}A(t) = i[H, A(t)]$$

que foi introduzida por Heisenberg, mais ou menos ao mesmo tempo em que Schrödinger introduzia sua versão da mecânica quântica.

7.8 Evolução temporal após uma medição

A evolução temporal contínua que apareceu na seção anterior não é, infelizmente, a única a aparecer na mecânica quântica. O processo de medição interfere com o objeto que está sendo medido e produz uma evolução bem distinta e descontínua.





Por exemplo, considere um estado $\psi = \frac{1}{\sqrt{2}}v_0 + \frac{1}{\sqrt{2}}v_1$, onde v_0 e v_1 são os auto-estados de um operador A , que corresponde a um observável. Ao efetuarmos uma medição de A só podemos obter como resultado um dos autovalores de A e nesse caso o estado deverá ser, imediatamente após a medição, o autovetor correspondente; ou seja, o processo de medição transforma instantaneamente o estado ψ em v_0 ou v_1 , o que é claramente descontínuo e também é uma operação irreversível no tempo, pois não é possível retroceder e recuperar o estado original ψ apenas a partir do conhecimento de v_0 ou v_1 .

Este comportamento é conhecido, de maneira mais dramática do que deveria, como o ”colapso do estado” e é algo presente como um postulado dentro da teoria. Se é possível trocá-lo por algo menos drástico como por exemplo a evolução temporal da seção anterior, que é suave, mas em um contexto bem mais completo que incluía também a descrição do aparato de medição (que é um sistema quântico com uma enorme quantidade de partículas) é algo ainda em investigação, como muitos outros aspectos da teoria quântica.

7.9 Operadores hermitianos em \mathbb{C}^2

Vamos agora nos concentrar no objeto básico que representa os observáveis de sistemas de dois níveis, os operadores hermitianos. Fixando uma base, que evidentemente será a base canônica, podemos escrever esses operadores como matrizes, cujas entradas são

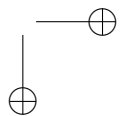
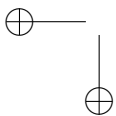
$$a_{ij} = \langle e_i, Ae_j \rangle$$

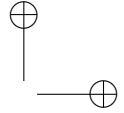
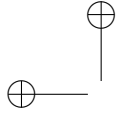
A condição de ser hermitiano implica em $A = A^*$ e portanto

$$a_{ij} = \langle e_i, Ae_j \rangle = \langle Ae_i, e_j \rangle = \overline{\langle e_j, Ae_i \rangle} = \bar{a}_{ji}$$

Logo a_{11} e a_{22} são reais e $a_{12} = a + ib = \bar{a}_{21}$. Isto nos diz que a matriz associada ao operador A (que, por abuso, continuamos denotando por A) tem a forma

$$A = \begin{bmatrix} w + z & x - iy \\ x + iy & w + z \end{bmatrix} = xI + x\sigma_x + y\sigma_y + z\sigma_z$$





7.10. EXPONENCIAIS DE σ_I

53

onde

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

são conhecidas como matrizes de Pauli. Em resumo, o que acabamos de mostrar é que o conjunto $\{I, \sigma_x, \sigma_y, \sigma_z\}$ é uma base do espaço vetorial de matrizes que representam transformações hermitianas.

7.10 Exponenciais de σ_i

O cálculo das exponencias das matrizes de Pauli é facilitado pelo fato de que suas potências são fáceis de se obter. Por exemplo, $\sigma_x^2 = I$. Desta forma, podemos calcular

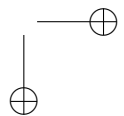
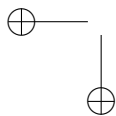
$$\begin{aligned} e^{-it\sigma_x} &= I - it\sigma_x + \frac{1}{2!}(-it\sigma_x)^2 + \frac{1}{3!}(-it\sigma_x)^3 + \dots = \\ &= I - it\sigma_x - \frac{1}{2!}t^2I + \frac{1}{3!}it^3\sigma_x + \dots = \\ &= I\left(1 - \frac{1}{2!}t^2 + \dots\right) - i\sigma_x\left(t - \frac{1}{3!}t^3 + \dots\right) = \\ &= I \cos t - i\sigma_x \sin t = \begin{bmatrix} \cos t & -i \sin t \\ -i \sin t & \cos t \end{bmatrix} \end{aligned}$$

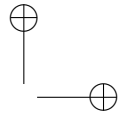
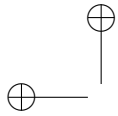
O leitor pode notar que, para qualquer t real, esta matriz está no grupo $SU(2)$.

7.11 O spin de uma partícula num campo magnético

A discussão desta seção de fato contém a origem histórica de um observável que é um importante exemplo de sistema de dois níveis na física.

Para simplificar a discussão vamos considerar um experimento feito da seguinte maneira: elétrons (que são nossas partículas quânticas por excelência) são preparados em uma câmara e enviados por um tubo até um aparelho de medição, que pode ser um aparelho





fotográfico. O tubo está sujeito a um campo magnético, um campo vetorial que afeta de alguma forma o movimento de partículas com carga elétrica, que em geral os físicos denotam por B .

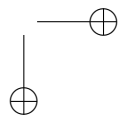
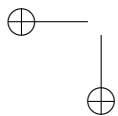
A câmara prepara os elétrons e os envia em um ritmo bem lento, de maneira que podemos supor que um não interfere no outro.

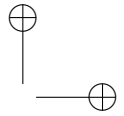
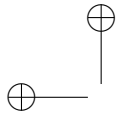
Fazendo o experimento, um físico do início do século XX certamente esperaria encontrar uma mancha na hora de revelar o filme, mostrando que elétrons distintos se comportam de maneiras diversas ao passar pelo mesmo campo magnético. Porém, para surpresa dos que efetivamente fizeram isso (não exatamente com elétrons no experimento original de Stern e Gerlach, mas isso não afeta nossa discussão) a mancha na verdade ficava reduzida a dois pontos! Isso mostrava um comportamento bastante novo, pois indicava que a maneira como os elétrons reagiam ao campo magnético essencialmente podia ser descrita pelo que estamos chamando nesse capítulo de um sistema de dois níveis. A grandeza física associada a esse comportamento ficou conhecida pelo nome de spin, palavra em inglês que pode ser traduzida como “girar” e que está de fato associada a primeira tentativa de interpretação do experimento que descrevemos, que consistia em imaginar o elétron como uma pequena esfera com movimento de rotação em torno de seu próprio eixo (como uma miniatura da Terra), o que geraria sua reação ao campo magnético exterior. Porém essa visão não só entrava em conflito com muito do que era conhecido do eletromagnetismo, como também ainda não explicava o misterioso fato de que o eixo de rotação parecer ter uma preferência no espaço.

O campo magnético, que supomos constante no tempo, é um campo vetorial em geral denotado por $B = (B_x, B_y, B_z)$. A evolução temporal de um estado (do spin) sob a ação deste campo é dada pelo operador hamiltoniano

$$H = B_x \sigma_x + B_y \sigma_y + B_z \sigma_z$$

Por exemplo, se o campo é vertical, ou seja $B = (0, 0, 1)$ então temos um operador simples, $H = \sigma_z$; os seus autovalores são $+1$ e -1 , correspondendo aos autovetores $(1, 0)$ e $(0, 1)$. Um estado qualquer então será $\psi = (\alpha, \beta)$, com $|\alpha|^2 + |\beta|^2 = 1$ e medições nesse estado podem produzir dois resultados (que correspondem às partículas que formavam os dois pontos na palca fotográfica mencionada acima), $+1$





7.12. EXEMPLOS

55

e -1 com probabilidade, respectivamente, $|\alpha|^2$ e $|\beta|^2$. Desta forma os físicos encontraram pela primeira vez na natureza um sistema cuja descrição quântica é exatamente como sendo um sistema de dois níveis, e se o leitor ainda não está convencido da importância desta grandeza vale comentar que muitas das propriedades químicas dos materiais que conhecemos são consequência da interação dos spins dos elétrons em cada um de seus átomos.

7.12 Exemplos

7.12.1 Exemplo 1

Vamos considerar um sistema de dois níveis cuja evolução temporal é dada pelo operador

$$H = \begin{bmatrix} V & 1 \\ 1 & V \end{bmatrix}$$

Os autovalores e autovetores são

$$\lambda_- = V-1, \psi_- = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{e} \quad \lambda_+ = V+1, \psi_+ = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Vamos agora imaginar a medição de um observável

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

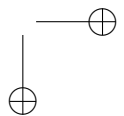
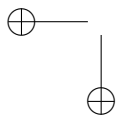
cujos autovalores e autovetores são:

$$\lambda_0 = 0, \varphi_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{e} \quad \lambda_1 = 1, \varphi_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Precisamos, é claro, dizer em qual estado efetuaremos a medição de A e escolheremos, por exemplo, o estado ψ_- obtido acima. Não é difícil ver que

$$\psi_- = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Então o resultado da medição de A neste estado será 0 com probabilidade $1/2$ e 1 com probabilidade também $1/2$; portanto o valor esperado deste observável é $1/2$, o que também pode ser obtido com a expressão $\langle \psi_-, A\psi_- \rangle$ (verifique!). O leitor pode agora calcular também o valor esperado de A no estado ψ_+ e constatar que também será $1/2$.



7.12.2 Exemplo 2

Agora consideramos o seguinte hamiltoniano:

$$H = \begin{bmatrix} a & -1 \\ -1 & -a \end{bmatrix}$$

onde $a \geq 0$, cujos autovalores e autovetores são

$$\lambda_+ = \sqrt{1+a^2}, \psi_+ = \begin{bmatrix} 1 \\ a - \sqrt{1+a^2} \end{bmatrix} \quad \text{e}$$

$$\lambda_- = -\sqrt{1+a^2}, \psi_- = \begin{bmatrix} 1 \\ a + \sqrt{1+a^2} \end{bmatrix}$$

(note que os autovetores não estão normalizados; para normalizá-los devemos multiplicá-los por constantes convenientes). Se o estado inicial é $\psi = \psi_-$ (que corresponde ao menor autovalor de H então uma medição de A (o observável definido no exemplo anterior) resulta em 0 com probabilidade $|\langle(1,0), \gamma\psi_-\rangle|^2$, e em 1 com probabilidade $|\langle(0,1), \gamma\psi_-\rangle|^2$ (onde γ é a constante de normalização do vetor ψ_-). Portanto temos que

$$\frac{P(0)}{P(1)} = \frac{|\langle(1,0), \gamma\psi_-\rangle|^2}{|\langle(0,1), \gamma\psi_-\rangle|^2} = \frac{1}{(a + \sqrt{1+a^2})^2} < 1$$

(note que a constante de normalização desaparece quando consideramos a razão) pois $a + \sqrt{1+a^2} \geq \sqrt{1+a^2} > 1$, e assim temos que $P(0) < P(1)$ mostrando que há uma maior probabilidade de se obter 1 do que 0 quando se mede A no estado ψ_- .

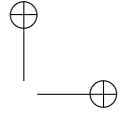
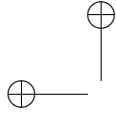
7.12.3 Exemplo 3

Consideramos agora a evolução temporal de um sistema descrito pelo operador

$$H = \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Então temos

$$\psi(t) = e^{-it\sigma_z} \psi(0) = \begin{bmatrix} e^{-it} & 0 \\ 0 & e^{it} \end{bmatrix} \psi(0)$$



7.12. EXEMPLOS

57

Podemos então tentar saber como será a evolução temporal do valor esperado de um certo observável, por exemplo,

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Sabemos que

$$\begin{aligned} \langle A \rangle(t) &= \langle \psi(t), A\psi(t) \rangle = \langle e^{-it\sigma_z}\psi(0), e^{-it\sigma_z}\psi(0) \rangle = \\ &= \langle \psi(0), e^{it\sigma_z} A e^{-it\sigma_z}\psi(0) \rangle \end{aligned}$$

Logo devemos calcular $e^{it\sigma_z}\sigma_x e^{-it\sigma_z}$:

$$\begin{bmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} e^{-it} & 0 \\ 0 & e^{it} \end{bmatrix} = \begin{bmatrix} 0 & e^{2it} \\ e^{-2it} & 0 \end{bmatrix}$$

Se $\psi(0) = (a, b)$ então podemos fazer o cálculo e obtemos

$$\begin{aligned} \langle \sigma_x \rangle(t) &= \bar{a}e^{2it}b + \bar{b}e^{-2it}a = \\ &= \bar{a}e^{2it}b + \overline{\bar{a}e^{2it}b} = 2\text{Re}\bar{a}e^{2it}b \end{aligned}$$

Vejamos em que isso resulta em dois casos particulares:

Se $\psi(0) = (1, 0)$ então $\langle \sigma_x \rangle(t) = 0$.

Se $\psi(0) = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ então $\langle \sigma_x \rangle(t) = \cos 2t$.

Exercícios

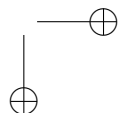
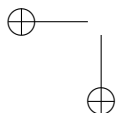
1- Quais, dentre os vetores abaixo, representam o mesmo estado físico?

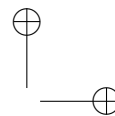
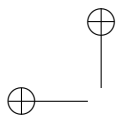
$$a) \frac{1}{\sqrt{2}}(1, -1) \quad b) \frac{1}{\sqrt{2}}(-i, i) \quad c) \frac{1}{\sqrt{2}}(i, 0)$$

2- Obtenha a evolução temporal de sistemas descritos pelos seguintes hamiltonianos:

$$H = \sigma_x \quad H = \sigma_y$$

(note que $\sigma_x^2 = \sigma_y^2 = I$)

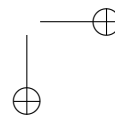
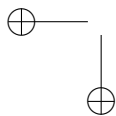


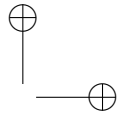
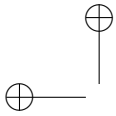


3- Considere o operador

$$H = B_x \sigma_x + B_y \sigma_y + B_z \sigma_z$$

com B_x, B_y e B_z reais. Então mostre que e^{-itH} é elemento de $SU(2)$.





Capítulo 8

Neste capítulo vamos estender a mecânica quântica para sistemas com N níveis.

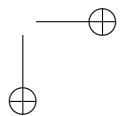
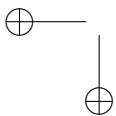
8.1 Aquecimento: sistemas com 3 e 4 níveis

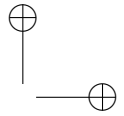
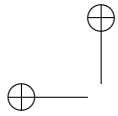
Para evitar um salto muito abrupto vamos tentar primeiro compreender como trabalhar com um sistema que tem 3 níveis, ou seja, que pode assumir 3 estados. De forma análoga ao caso de 2 níveis, devemos considerar um espaço vetorial de dimensão 3, ou seja, \mathbb{C}^3 . Neste espaço os estados serão vetores de norma 1 e também identificaremos dois vetores ψ e ϕ se $\psi = e^{i\theta}\phi$ para algum θ real. A evolução temporal e os observáveis serão representados, novamente, por operadores hermitianos, ou seja, por operadores representados por matrizes hermitianas.

Exemplo 8.1.1. *Tome*

$$H = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Este observável tem como autovalores 0, 1 e 2, correspondendo aos autovetores $(1, 0, 0)$, $(0, 1, 0)$ e $(0, 0, 1)$. Se consideremos o estado $\psi =$





(α, β, γ) (com $|\alpha|^2 + |\beta|^2 + |\gamma|^2 = 1$ então a probabilidade de uma medição de H resultar em 1 é dada por $|\beta|^2$).

Exemplo 8.1.2. Tome

$$H = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

Os autovalores são -1 (com multiplicidade 2) e 2. Esta matriz é hermitiana, e portanto representa um observável (ou pode ser o hamiltoniano, o operador que nos dá a evolução temporal).

Exemplo 8.1.3. Tome

$$H = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

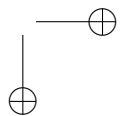
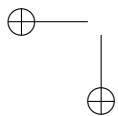
O leitor pode verificar que $\lambda = 0$ é um autovalor de H associado ao autovetor

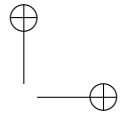
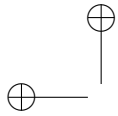
$$\psi = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

Podemos também considerar o caso de um sistema com 4 níveis, de maneira completamente análoga.

Exemplo 8.1.4. (Este exemplo é baseado no capítulo 12 de Feynman [4]; recomendo ao leitor a consulta aquele texto para uma belíssima discussão da física do problema)

Estamos interessados em entender um modelo bastante simplificado do átomo de hidrogênio, que contém apenas duas partículas, um próton e um elétron. Seguindo Feynmann, uma maneira de modelar este sistema é levar em conta apenas duas informações: tanto o próton como o elétron tem spin, e portanto cada um deles pode ser pensado como um sistema de dois níveis, com estados que denominaremos simplesmente de $|+\rangle$ e $|-\rangle$ (e assim usamos um pouco a notação de Dirac para vetores). O átomo, composto das duas partículas, pode então ser descrito por meio de quatro estados, onde o primeiro símbolo se refere ao spin do próton e o segundo ao do





8.2. N NÍVEIS

61

elétron: $(++)$, $(+-)$, $(-+)$ e $(--)$. Usando estes quatro vetores como base temos então nosso sistema de quatro níveis cuja evolução temporal é dada pelo operador

$$H = \begin{bmatrix} A & 0 & 0 & 0 \\ 0 & -C & B & 0 \\ 0 & B & -C & 0 \\ 0 & 0 & 0 & D \end{bmatrix}$$

O cálculo dos autovalores não é difícil e obtemos $A, D, -C - B$ e $-C + B$. Um caso particular é tomar $C = D = A$ e $B = 2A$, com $A > 0$. Assim os autovalores passam a ser A com multiplicidade 3 e $-3A$ com multiplicidade 1. Podemos então procurar os autovetores; para $\lambda = -3A$, o menor deles, temos

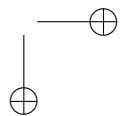
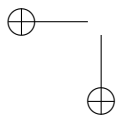
$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}$$

e o leitor pode então continuar o processo e obter os demais autovetores.

8.2 N níveis

Passamos agora ao caso de N níveis. Como o leitor já pode esperar, nosso espaço vetorial agora é \mathbb{C}^N , novamente consideramos vetores normalizados e com a identificação de dois vetores ψ e ϕ se $\psi = e^{i\theta}\phi$ para definir os estados e os observáveis são representados pelos operadores hermitianos de \mathbb{C}^N . No que segue descreveremos um exemplo interessante desse tipo de situação.

Um sistema de N níveis pode, fisicamente, ser interpretado de muitas formas. Uma delas é imaginar uma partícula quântica (por exemplo, um elétron) que se move um material composto de exatamente N 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 N níveis, que correspondem às N posições



dos átomos do material. Esse modelo, com toda a ingenuidade que aparenta, é um ponto de partida razoável para entender, por exemplo, o comportamento de partículas em um cristal.

Considere os operadores lineares em \mathbb{C}^n 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 difícil 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 $b_{k,l} = e^{i\frac{2\pi}{n}lk}$ para $l = 1, 2, \dots, n$ e $k = 0, 1, \dots, n-1$. Definimos assim os vetores (já normalizados)

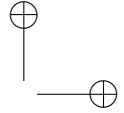
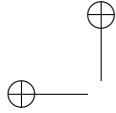
$$B_l = \frac{1}{\sqrt{n}} \begin{bmatrix} b_{0,l} \\ b_{1,l} \\ \vdots \\ b_{n-1,l} \end{bmatrix}$$

Não é difícil verificar que $N_+ B_l = e^{i\frac{2\pi}{n}l} B_l$ e $N_- B_l = e^{-i\frac{2\pi}{n}l} B_l$.

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 já definidos, e os autovalores são obtidos como segue:

$$\begin{aligned} \Delta B_l &= N_+ B_l + N_- B_l - 2I B_l = (e^{i\frac{2\pi}{n}l} + e^{-i\frac{2\pi}{n}l} - 2) B_l = \\ &= 2\left(\cos \frac{2\pi}{n}l - 1\right) B_l \end{aligned}$$

Os vetores B_l são uma base de \mathbb{C}^n e portanto um estado inicial ψ qualquer pode ser expresso como combinação linear $\psi = \sum c_l B_l$.



8.3. OPERADOR POSIÇÃO

63

Para obter a evolução temporal deste estado inicial devemos resolver a equação de Schrödinger

$$\frac{d}{dt}\psi(t) = -iH\psi(t)$$

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

$$c_l(t) = e^{-i\lambda_l t}$$

8.3 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 $0, 1, \dots, n-1$ e correspondem aos autovetores

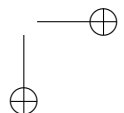
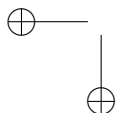
$$v_k := e_{k+1} \quad \text{para } k = 0, 1, \dots, n-1$$

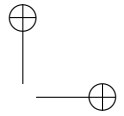
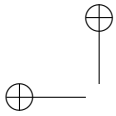
onde os e_i são os vetores da base canônica de \mathbb{C}^n .

Agora considere os autovetores B_l da seção anterior. Se temos um estado $\psi = B_l$ (para algum l fixo) então podemos perguntar qual é a probabilidade de se obter o valor k (k entre 0 e $n-1$) numa medição do observável X . Mas

$$\psi = B_l = \sum_{j=0}^{n-1} b_{j,l} v_j$$

Portanto a probabilidade de se obter a posição k , que é o módulo ao quadrado do coeficiente de v_k é dada por $1/n$, de maneira independente de k . Logo todas as posições são equiprováveis num estado descrito por $\psi = B_l$. Porém agora note que esta probabilidade também





não depende do l escolhido! Assim, qualquer que seja o autoestado de Δ temos que a posição tem uma distribuição equiprovavel.

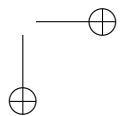
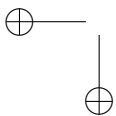
O operador Δ pode ser interpretado como sendo associado a energia de uma partícula num cristal, sendo que X está relacionado a sua posição nessa rede cristalina. Os autovalores de Δ são os possíveis valores da energia e os de X , os possíveis valores da posição. O que constatamos acima é que quando uma 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.

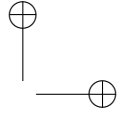
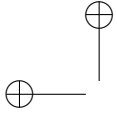
Exercícios

- 1- Para a e b reais, obtenha os autovalores e os autovetores do operador

$$a(N_+ + N_-) + bI$$

- 2- Considerando-se um estado ψ que é auto-estado de X qual é a probabilidade de que tenha um determinado valor de energia (isto é, um determinado auto-valor do operador Δ)?





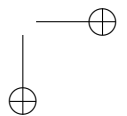
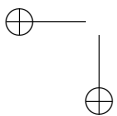
Capítulo 9

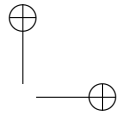
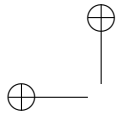
Nesse capítulo faremos a descrição de sistemas compostos na mecânica quântica, o que requererá o conceito de produto tensorial exposto no início do texto.

9.1 Dois sistemas de três níveis

Para começar a discussão vamos considerar uma situação bastante concreta, com dois sistemas de três níveis; a descrição de um sistema como esse, com uma base de estados que vamos denotar por (usando a notação de Dirac) $|1\rangle$, $|2\rangle$ e $|3\rangle$, usa o espaço de Hilbert \mathbb{C}^3 . Ingentuamente podemos pensar que a descrição de um sistema conjunto contendo dois subsistemas de três níveis sera feita com o uso do produto cartesiano $\mathbb{C}^3 \times \mathbb{C}^3$ (onde a primeira componente representa o primeiro subsistema e a segunda representa o segundo); isso significaria usar como espaço de Hilbert \mathbb{C}^6 , mas essa abordagem não está correta!

Para encontrarmos o espaço mais adequado vamos proceder com mais calma: precisamos, antes de mais nada, saber qual é a base do espaço que descreve o sistema composto; precisamos de informação sobre o primeiro subsistema e também sobre o segundo. Uma maneira simples de fazer isso é considerar que a base do novo espaço é formada pelos seguintes vetores: $|1\rangle \otimes |1\rangle$, $|2\rangle \otimes |1\rangle$, $|3\rangle \otimes |1\rangle$, $|1\rangle \otimes |2\rangle$, $|2\rangle \otimes |2\rangle$, $|3\rangle \otimes |2\rangle$, $|1\rangle \otimes |3\rangle$, $|2\rangle \otimes |3\rangle$ e $|3\rangle \otimes |3\rangle$. Dessa forma precisamos de nove





vetores na base, ou seja, estamos usando o espaço \mathbb{C}^9 , que de fato é o produto tensorial de \mathbb{C}^3 por \mathbb{C}^3 , isto é, $\mathbb{C}^3 \otimes \mathbb{C}^3$ (o leitor percebe o porque de nossa escolha de dois sistemas de três níveis, e não de dois sistemas de dois? Nesse caso teríamos o produto tensorial de \mathbb{C}^2 por \mathbb{C}^2 que é o \mathbb{C}^4 , assim como o produto cartesiano $\mathbb{C}^2 \times \mathbb{C}^2$ e nesse caso não notaríamos a diferença... E de fato este já caso já apareceu, sem muito alarde, num exemplo do capítulo sobre sistemas de dois níveis que tratava de um modelo imensamente simplificado de um átomo de hidrogênio, que é um sistema composto de dois subsistemas, um próton e um elétron).

Muitas vezes a brevidade faz com que seja usada a notação $|ij\rangle$ no lugar de $|i\rangle \otimes |j\rangle$ e adotaremos esse procedimento no que segue.

9.2 Dois sistemas

Generalizando o procedimento da última seção vemos que a descrição de um sistema composto de dois subsistemas descritos, respectivamente, pelos espaços de Hilbert \mathcal{H}_1 e \mathcal{H}_2 , é feita usando-se o espaço de Hilbert $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. Evitando maiores abstrações, este espaço (o produto tensorial de \mathcal{H}_1 por \mathcal{H}_2) pode ser definido como segue: Se $\{v_i\}_{i=1,\dots,d_1}$ é uma base de \mathcal{H}_1 e $\{u_j\}_{j=1,\dots,d_2}$ é uma base de \mathcal{H}_2 então a base de $\mathcal{H}_1 \otimes \mathcal{H}_2$ é o conjunto $\{v_i \otimes u_j\}_{(i,j) \in \{1,\dots,d_1\} \times \{1,\dots,d_2\}}$ e assumimos as seguintes propriedades: para $v, w \in \mathcal{H}_1$, $u, t \in \mathcal{H}_2$ e $\lambda \in \mathbb{C}$,

$$1- (\lambda v) \otimes u = \lambda v \otimes u = v \otimes (\lambda u)$$

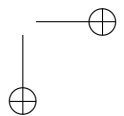
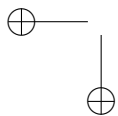
$$2- (v + w) \otimes u = v \otimes u + w \otimes u$$

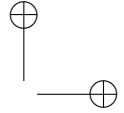
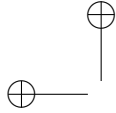
$$3- v \otimes (u + t) = v \otimes u + v \otimes t$$

Por fim, o produto interno de \mathcal{H} é definido da seguinte forma: para os elementos da base

$$\langle v_i \otimes u_j, v_l \otimes u_m \rangle := \langle v_i, v_l \rangle_1 \langle u_j, u_m \rangle_2$$

e estendemos, pela linearidade de $\langle \cdot, \cdot \rangle$, para um vetor qualquer de \mathcal{H} .





9.3 Dois spins

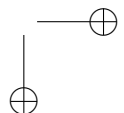
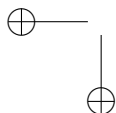
Um tipo de sistema composto bastante estudado é o que se compõe de dois sistemas de dois níveis, por exemplo, os spins de duas partículas que interagiram e se separaram. Esses pares são bastante importantes em discussões sobre a fundamentação conceitual da mecânica quântica e ficaram conhecidos como pares de Bell em referência a um conhecido físico chamado John Bell, que retomou um debate iniciado com um artigo de Einstein, Podolsky e Rosen (no qual a discussão de pares de partículas ocupa papel essencial) sobre a completude da mecânica quântica. Resumindo, estes autores tentavam mostrar que existiriam aspectos da realidade física que *não* encontravam um correspondente dentro da teoria quântica; sendo assim, esta seria incompleta. Não pretendemos entrar nessa discussão mas remetemos o leitor ao capítulo 20 de [1] para uma exposição cuidadosa do tópico.

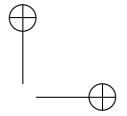
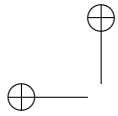
Deixando os aspectos conceituais de lado, os estados de Bell são os seguintes:

$$|\Phi_{\pm}\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle)$$

$$|\Psi_{\pm}\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle)$$

Por exemplo, o estado Ψ_+ representa uma superposição dos estados $|01\rangle$ e $|10\rangle$, que correspondem, respectivamente, a primeira partícula no estado $|0\rangle$ e a segunda no estado $|1\rangle$ para $|01\rangle$ e a primeira partícula no estado $|1\rangle$ e a segunda no estado $|0\rangle$ para $|10\rangle$. Estes são estados que estão no espaço $\mathbb{C}^2 \otimes \mathbb{C}^2 = \mathbb{C}^4$.





Capítulo 10

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 [3].

10.1 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 $*$: $A \rightarrow A$ satisfazendo

i $(a + b)^* = a^* + b^*$

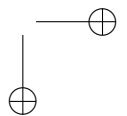
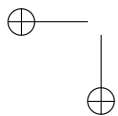
ii $(\lambda a)^* = \bar{\lambda} a^*$

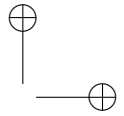
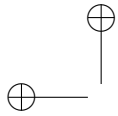
iii $(ab)^* = b^* a^*$

iv $(a^*)^* = a$

v $\|a^*\| = \|a\|$

vi $\|a^* a\| = \|a\|^2$





10.1. ÁLGEBRAS C^*

69

então temos uma álgebra C^* .

Exemplo 10.1.1. \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 10.1.2. Seja $C_0(\mathbb{R})$ o conjunto das funções $f: \mathbb{R} \rightarrow \mathbb{C}$ 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

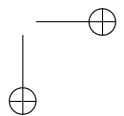
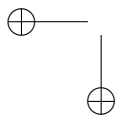
$$f^*(x) = \overline{f(x)}$$

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 [3] para uma discussão mais completa).



Exemplo 10.1.3. *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

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

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

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

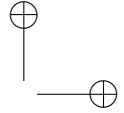
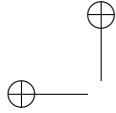
$$|A^*Av||v| = |A^*Av| \leq \|A^*A\||v| = \|A^*A\|$$

Então tomando o supremo sobre v em ambos os lados podemos concluir que $\|A\|^2 \leq \|A^*A\|$; como já havíamos visto antes nas propriedades da norma de operador, $\|A^*A\| \leq \|A^*\| \|A\| = \|A\| \|A\| = \|A\|^2$ e portanto temos a propriedade desejada.

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^* .



10.2 Estados de uma álgebra

Álgebras são espaços vetoriais, logo é interessante perguntar o que ocorre com seus funcionais lineares. No caso de álgebras C^* com unidade uma classe especial de funcionais lineares merece bastante atenção, e é chamada de estado; no momento oportuno tentaremos fazer a conexão deste conceito de estado com aquele que já vem sendo usado desde o início do texto.

Definição 10.2.1. *Seja A uma álgebra C^* com unidade; um funcional linear $f: A \rightarrow \mathbb{C}$ é chamado de estado se*

- (a) $f(a^*a) \geq 0$ para todo $a \in A$.
- (b) $f(1) = 1$.

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

$$\langle a, b \rangle = \text{tr}(a^*b)$$

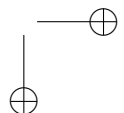
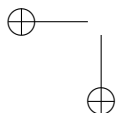
é um produto interno. De fato a verificação da linearidade não é difícil e deixamos a tarefa para o leitor.

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

$$\begin{aligned} \langle b, a \rangle &= \text{tr}(b^*a) = \sum_{i=1}^n \langle e_i, b^*a e_i \rangle = \sum_{i=1}^n \langle a^*b e_i, e_i \rangle = \\ &= \sum_{i=1}^n \overline{\langle e_i, a^*b e_i \rangle} = \overline{\sum_{i=1}^n \langle e_i, a^*b e_i \rangle} = \overline{\text{tr}(a^*b)} = \overline{\langle a, b \rangle} \end{aligned}$$

Alem disso,

$$\begin{aligned} \langle a, a \rangle &= \text{tr}(a^*a) = \sum_{i=1}^n \langle e_i, a^*a e_i \rangle = \sum_{i=1}^n \langle a e_i, a e_i \rangle = \\ &= \sum_{i=1}^n \|a e_i\|^2 \geq 0 \end{aligned}$$



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, $a = 0$.

Consideremos agora um funcional linear $f: A \rightarrow \mathbb{C}$. Pelo teorema 3.2.1 sabemos que f pode ser escrito como

$$f(x) = \langle x, V_f \rangle = \text{tr}(x^* V_f)$$

para um único elemento $V_f \in A$.

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

$$f(1) = \text{tr}(1^* V_f) = \text{tr}(1 V_f) = \text{tr}(V_f) = 1$$

que é a primeira condição que obtemos sobre V_f .

A segunda condição é que $V_f \geq 0$. De fato, para todo $a \in A$

$$\begin{aligned} 0 \leq f(a^* a) &= \text{tr}(a^* a V_f) = \text{tr}(a V_f a^*) = \sum_{i=1}^n \langle e_i, a V_f a^* e_i \rangle = \\ &= \sum_{i=1}^n \langle a^* e_i, V_f a^* e_i \rangle \end{aligned}$$

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

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

Desta forma, temos que

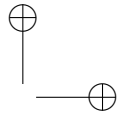
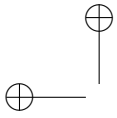
$$0 \leq f(a^* a) = \langle v, V_f v \rangle \quad \text{para qualquer } v \in \mathbb{C}^n$$

e desta forma $V_f \geq 0$.

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

10.3 Exercícios

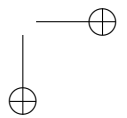
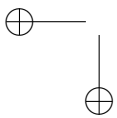
- 1- Considere uma álgebra C^* A com unidade 1, isto é, com um elemento 1 tal que $1a = a1 = a$ para todo $a \in A$. Mostre que $1^* = 1$; mostre que $\|1\| = 1$.

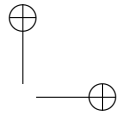
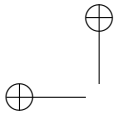


10.3. EXERCÍCIOS

73

- 2- Considere a álgebra C^* $A = M_2(\mathbb{C})$. Se $\varphi: A \rightarrow \mathbb{C}$ 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 $M_n(\mathbb{C})$).





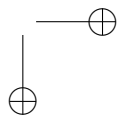
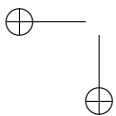
Capítulo 11

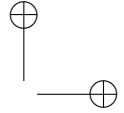
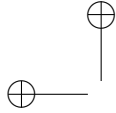
Neste capítulo vamos descrever situações quânticas com alguma aleatoriedade clássica.

11.1 Introdução

Vamos nesse capítulo considerar situações como a seguinte: imaginemos duas partículas, uma no estado φ e a outra no estado ψ . Pretendemos efetuar uma medição de um observável, digamos A , mas não sabemos a priori qual das partículas será medida, sabemos apenas que a probabilidade de estarmos com a partícula de estado φ é p e a de estarmos com ψ é $(1 - p)$. Este é o tipo de incerteza com a qual já estamos acostumados, do mesmo tipo que a incerteza sobre qual será a cor de uma bola retirada de uma urna ou qual serão os próximos números da mega-sena. A esta é acrescida uma incerteza tipicamente quântica pois mesmo a certeza de termos um determinado estado (por exemplo, φ) não nos permite nada além de prever apenas valores esperados de um certo observável.

Onde podemos encontrar uma situação desse tipo? Bem, uma possibilidade é quando se tenta descrever um sistema com uma grande quantidade de partículas (por exemplo, todo o ar dentro de uma sala); não parece razoável tentar descrever o estado quântico de cada uma, mas eventualmente é possível saber quais são os estados presentes e qual a probabilidade de cada um. Quando se efetua uma medição





11.2. MATRIZ DENSIDADE

75

de algum observável numa partícula escolhida ao acaso estamos exatamente na situação sugerida acima, na qual não sabemos qual o estado quântico exato. Essa é a situação descrita no âmbito do que se chama de mecânica estatística quântica.

11.2 Matriz densidade

Dado um elemento não nulo $v \in V$ podemos definir a projeção sobre a direção de v (que assumimos normalizado) como sendo o operador

$$P_v: V \rightarrow V \quad P_v(u) = \langle v, u \rangle v$$

Já sabemos que este operador é hermitiano; também podemos provar que a projeção é um operador positivo, isto é,

$$\langle u, P_v(u) \rangle \geq 0 \text{ para todo vetor } u$$

De fato

$$\begin{aligned} \langle u, P_v(u) \rangle &= \langle u, \langle v, u \rangle v \rangle = \langle v, u \rangle \langle u, v \rangle = \\ &= \langle v, u \rangle \overline{\langle v, u \rangle} = |\langle v, u \rangle|^2 \geq 0 \end{aligned}$$

Na situação descrita na introdução podemos representar o sistema por meio do seguinte operador:

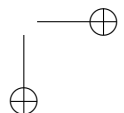
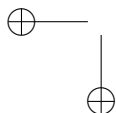
$$\rho = pP_\varphi + (1-p)P_\psi$$

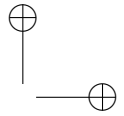
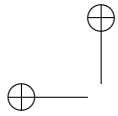
Note que $\rho = \rho^*$, pois as projeções são hermitianas. E

$$\text{tr}(\rho) = p \text{tr}(P_\varphi) + (1-p) \text{tr}(P_\psi)$$

Mas o traço de uma projeção sobre um vetor normalizado é

$$\begin{aligned} \text{tr}(P_v) &= \sum_{i=1}^n \langle e_i, P_v(e_i) \rangle = \sum_{i=1}^n \langle e_i, \langle v, e_i \rangle v \rangle = \\ &= \sum_{i=1}^n |\langle v, e_i \rangle|^2 = \|v\|^2 = 1 \end{aligned}$$





e portanto $\text{tr}(\rho) = 1$. Por fim, notamos que $\rho \geq 0$, ou seja, ρ é um operador positivo: de fato

$$\langle u, \rho u \rangle = p \langle u, P_\varphi(u) \rangle + (1 - p) \langle u, P_\psi(u) \rangle \geq 0$$

pois cada uma das projeções é positiva.

Baseado nesta situação definimos um operador ρ , que é uma matriz em $M_n(\mathbb{C})$, como sendo um operador-densidade (ou matriz-densidade) quando $\rho = \rho^*$, $\text{tr}(\rho) = 1$ e $\rho \geq 0$.

Podemos obter o valor esperado do observável A no estado ρ como segue:

$$\langle A \rangle = \text{tr}(A\rho)$$

11.3 Redefinindo estado

A palavra estado até esse momento estava sendo usada com o seguinte significado: uma classe de equivalência entre vetores unitários de um certo espaço de Hilbert, que identifica dois vetores ψ e $e^{i\theta}\psi$ ($\theta \in \mathbb{R}$), ou seja, dois vetores que são múltiplos por um fator que é um número complexo de norma unitária.

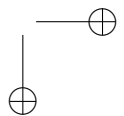
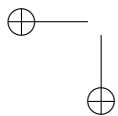
No início deste capítulo transferimos a tarefa de representar um estado associado a um sistema com aleatoriedade a um operador ρ , conhecido como matriz densidade. Deste momento em diante esse será o sentido no qual usaremos a palavra estado. Os estados de antes, ou seja, classes de vetores unitários, são representados nessa nova linguagem de operadores por meio do operador projeção:

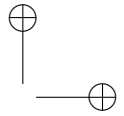
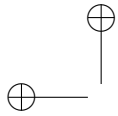
$$\text{vetor unitário } \psi \quad P_\psi(v) = \langle \psi, v \rangle \psi$$

Essa classe especial de estados, que são operadores de projeção, é conhecida como a classe dos *estados puros*.

Porém devemos verificar algo importante: já sabemos que os vetores ψ e $e^{i\theta}\psi$ representam o mesmo estado. Se ambos estiverem relacionados a dois operadores diferentes então nossa representação acima terá uma severa falha. Vamos então verificar quais são as projeções associadas a estes vetores:

$$P_\psi(v) = \langle \psi, v \rangle \psi$$





11.4. SISTEMA DE DOIS NÍVEIS

77

e

$$P_{e^{i\theta}\psi}(v) = \langle e^{i\theta}\psi, v \rangle e^{i\theta}\psi = e^{-i\theta} \langle \psi, v \rangle e^{i\theta}\psi = \langle \psi, v \rangle \psi = P_\psi(v)$$

mostrando que de fato os dois vetores que representam o mesmo estado estão associados ao mesmo estado puro definido pelo operador densidade $\rho = P_\psi = P_{e^{i\theta}\psi}$.

11.3.1 Relação com os estados de uma álgebra C^*

No final do capítulo anterior caracterizamos os estados da álgebra C^* das matrizes $M_n(\mathbb{C})$ como estando associados a matrizes V_f positivas e de traço 1. Os observáveis da mecânica quântica estão, por sua vez, associados ao conjunto formado pelas matrizes hermitianas. Já os estados estão representados por operadores ρ hermitianos, positivos e de traço 1. Então um estado (no sentido quântico) ρ está associado a um estado (no sentido da álgebra C^*) que é o funcional $f(x) = \text{tr}(x^*\rho)$, mostrando que estes conceitos estão intimamente relacionados.

11.4 Sistema de dois níveis

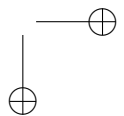
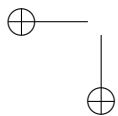
11.4.1 Caracterização geométrica

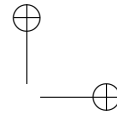
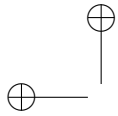
Nesse momento retomamos os sistemas de dois níveis para poder observá-los do ponto de vista da matriz-densidade. Nesse caso a caracterização geométrica do espaço de estados e dos estados puros é relativamente simples e por isso vamos mostrá-la.

Um estado agora corresponde a uma matriz-densidade que é uma matriz complexa ρ 2×2 e que satisfaz as seguintes condições: $\rho \geq 0$, $\rho = \rho^*$ e $\text{tr}(\rho) = 1$.

Da segunda condição obtemos que as entradas da diagonal precisam ser reais, e as que estão fora da diagonal são conjugadas; usando também a terceira condição temos então o aspecto de geral da matriz como sendo

$$\rho = \begin{bmatrix} 1/2 + x & y + iz \\ y - iz & 1/2 - x \end{bmatrix}$$





com x, y e z reais. Agora podemos usar a primeira condição, a positividade de ρ : um operador positivo precisa ter autovalores maiores ou iguais a zero. Como o determinante de uma matriz é dado pelo produto de seus autovalores, então podemos concluir que $\det \rho \geq 0$ e então

$$\frac{1}{4} - x^2 - y^2 - z^2 \geq 0, \quad \text{ou seja,} \quad x^2 + y^2 + z^2 \leq \frac{1}{4}$$

de forma que cada estado (isto é, cada matriz-densidade) corresponde a um ponto dentro da esfera de raio $1/2$.

E quanto aos estados puros? Bem, estes correspondem a operadores de projeção sobre uma dada direção. Mas como já vimos, uma projeção P satisfaz a propriedade $P^2 = P$ e portanto só admite como autovalores 0 e 1, sendo que toda projeção sobre um subespaço de dimensão menor do que a do o espaço ambiente tem necessariamente o autovalor 0. Portanto para os estados puros um dos autovalores de ρ é zero e assim o determinante também é zero; logo, para os estados puros temos

$$0 = \det \rho = \frac{1}{4} - x^2 - y^2 - z^2 \Rightarrow x^2 + y^2 + z^2 = \frac{1}{4}$$

ou seja, cada estado puro corresponde a um ponto na superfície da esfera de raio $1/2$.

Desta forma temos uma caracterização completa dos estados de um sistema de dois níveis como sendo uma esfera e dos correspondentes estados puros como sendo a superfície dessa esfera.

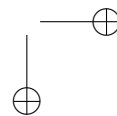
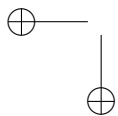
11.4.2 Um exemplo

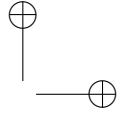
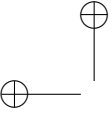
Para exemplificar o cálculo de valores esperados por meio da matriz densidade vamos considerar a situação descrita pelo estado

$$\rho = \frac{1}{4}|0\rangle\langle 0| + \frac{3}{4}|1\rangle\langle 1|$$

Usaremos a notação de Dirac no cálculo que segue e convidamos o leitor a reproduzir o resultado usando a notação matricial usual. Queremos obter o valor esperado do observável σ_x : note que este pode ser reescrito como

$$\sigma_x = |0\rangle\langle 1| + |1\rangle\langle 0|$$





11.4. SISTEMA DE DOIS NÍVEIS

79

Precisamos então obter $\text{tr}(\sigma_x \rho)$; note que

$$\begin{aligned} \sigma_x \rho &= (|0\rangle\langle 1| + |1\rangle\langle 0|) \left(\frac{1}{4}|0\rangle\langle 0| + \frac{3}{4}|1\rangle\langle 1| \right) = \\ & \frac{1}{4}|1\rangle\langle 0| + \frac{3}{4}|0\rangle\langle 1| \end{aligned}$$

Para o cálculo do traço devemos então fazer

$$\begin{aligned} \text{tr}(\sigma_x \rho) &= \langle 0| \left(\frac{1}{4}|1\rangle\langle 0| + \frac{3}{4}|0\rangle\langle 1| \right) |0\rangle + \\ & \langle 1| \left(\frac{1}{4}|1\rangle\langle 0| + \frac{3}{4}|0\rangle\langle 1| \right) |1\rangle = 0 \end{aligned}$$

Se agora queremos o valor esperado de σ_z devemos começar por ver que este operador, na notação de Dirac, toma a forma

$$\sigma_z = |0\rangle\langle 0| - |1\rangle\langle 1|$$

Então

$$\begin{aligned} \sigma_z \rho &= (|0\rangle\langle 0| - |1\rangle\langle 1|) \left(\frac{1}{4}|0\rangle\langle 0| + \frac{3}{4}|1\rangle\langle 1| \right) = \\ & \frac{1}{4}|0\rangle\langle 0| - \frac{3}{4}|1\rangle\langle 1| \end{aligned}$$

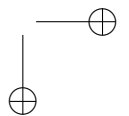
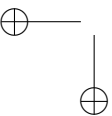
E assim

$$\begin{aligned} \text{tr}(\sigma_z \rho) &= \langle 0| \left(\frac{1}{4}|0\rangle\langle 0| - \frac{3}{4}|1\rangle\langle 1| \right) |0\rangle + \\ & \langle 1| \left(\frac{1}{4}|0\rangle\langle 0| - \frac{3}{4}|1\rangle\langle 1| \right) |1\rangle = \frac{1}{4} - \frac{3}{4} = -\frac{1}{2} \end{aligned}$$

11.4.3 Evolução temporal

A evolução temporal de uma matriz densidade é de fato a evolução temporal de um operador e é dada pela seguinte equação, conhecida como equação de Landau-von Neumann:

$$\frac{d}{dt} \rho_t = i[\rho_t, H]$$



onde H é o operador hamiltoniano associado ao sistema em questão. Esta equação é satisfeita por

$$\rho_t = e^{-itH} \rho_0 e^{itH}$$

A verificação dessa afirmação é deixada a cargo do leitor; como sugestão, consulte a seção 7.7 do capítulo 7. Note que ρ_t é realmente uma matriz densidade em qualquer instante. De fato,

$$\text{tr}(\rho_t) = \text{tr}(e^{-itH} \rho_0 e^{itH}) = \text{tr}(e^{itH} e^{-itH} \rho_0) = \text{tr}(\rho_0) = 1$$

se, é claro, assumimos que ρ_0 é uma matriz densidade e assim tem traço igual a 1. O fato de que ρ_t é autoadjunto não é muito difícil (leitor, verifique!). Por fim, devemos verificar a positividade de ρ_t :

$$\langle v, e^{-itH} \rho_0 e^{itH} v \rangle = \langle e^{itH} v, \rho_0 e^{itH} v \rangle = \langle u, \rho_0 u \rangle \geq 0$$

pois ρ_0 é operador positivo.

Exercícios

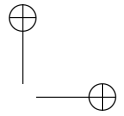
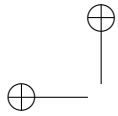
- 1- Verifique a afirmação de que os vetores ψ e $e^{i\theta}\psi$ correspondem ao mesmo operador de projeção $P_\psi(v) = \langle \psi, v \rangle \psi$ (sugestão: lembre que o operador de projeção se escreve como $|\psi\rangle\langle\psi|$).
- 2- Sejam ρ_1 e ρ_2 matrizes densidade (de mesma dimensão). Então mostre que $\rho = a\rho_1 + (1-a)\rho_2$ também é uma matriz densidade para todo $a \in [0, 1]$ (ρ é o que se chama de uma combinação convexa de ρ_1 e ρ_2).
- 3- Considere os operadores W_1, W_2, \dots, W_n tais que

$$\sum_{k=1}^n W_k^* W_k = 1$$

e ρ uma matriz densidade. Mostre que

$$\sum_{k=1}^n W_k^* \rho W_k$$

também é uma matriz densidade.



Capítulo 12

Neste capítulo falaremos sobre a mecânica quântica num intervalo da reta. Usaremos um espaço de Hilbert que é mais sofisticado que os já descritos até aqui, por isso pedimos paciência e 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.

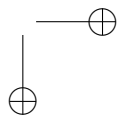
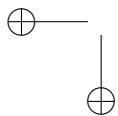
12.1 Partícula clássica na reta

A 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 momento da partícula (que vem a ser o produto de sua massa pela velocidade) é representado por p .

Quando a partícula está sujeita a ação de uma força $F: \mathbb{R} \rightarrow \mathbb{R}$ podemos descrever seu movimento por meio da lei de Newton:

$$F = ma = m \frac{d^2}{dt^2} x$$

com condições iniciais $x(0) = x_0$ e $v(0) = \frac{d}{dt} x(0) = v_0$.



Uma força pode ser convenientemente representada por um potencial, uma função $V: \mathbb{R} \rightarrow \mathbb{R}$ tal que

$$-\frac{d}{dx}V(x) = F(x)$$

Não é difícil ver que podemos obter uma função V satisfazendo essa propriedade se definimos

$$V(x) = -\int_0^x F(s)ds$$

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 tem 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)

$$H(p, x) = \frac{p^2}{2m} + V(x)$$

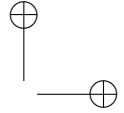
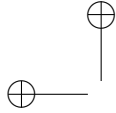
e tomar como equações de movimento as equações de Hamilton

$$\begin{cases} \frac{d}{dt}x &= \frac{\partial}{\partial p}H \\ \frac{d}{dt}p &= -\frac{\partial}{\partial x}H \end{cases}$$

O leitor não terá dificuldades em ver que o sistema acima equivale a lei de Newton.

12.2 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 Hilbert adequado para isso.



12.2. PARTÍCULA QUÂNTICA

83

Em primeiro lugar, qual é o espaço de estados? Precisamos descrever uma partícula na reta e temos um espaço de Hilbert naturalmente associado a ela que é o espaço $L^2(\mathbb{R})$. O produto interno é definido como sendo

$$\langle f, g \rangle = \int_{\mathbb{R}} f(x)\overline{g(x)}dx$$

Este parece ser um espaço bastante adequado para representar a posição da partícula.

Precisamos então compreender como representar os operadores de posição e momento, pois essas são as quantidades básicas que desjamos 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

$$\hat{x} := x$$

O operador momento, por sua vez, é dado por

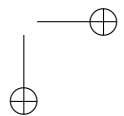
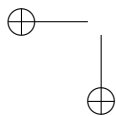
$$\hat{p} := -i\frac{d}{dx}$$

Podemos verificar a comutatividade (ou não) dos operadores \hat{x} e \hat{p} ; para isso usaremos uma função auxiliar ϕ :

$$\begin{aligned} [\hat{x}, \hat{p}]\phi &= (\hat{x}\hat{p} - \hat{p}\hat{x})\phi = \\ \hat{x}(-i\frac{d}{dx}\phi) + i\frac{d}{dx}(x\phi) &= -ix\frac{d}{dx}\phi + i\phi + ix\frac{d}{dx}\phi = \\ &= i\phi \end{aligned}$$

ou seja, $[\hat{x}, \hat{p}] = i$; sendo assim esses operadores não comutam e para esse par vale também a relação de incerteza que já deduzimos antes (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):

$$\Delta x \Delta p \geq \frac{1}{2} |\langle \psi, [\hat{x}, \hat{p}] \psi \rangle| = \frac{1}{2} |\langle \psi, \psi \rangle| = \frac{1}{2}$$



pois assumimos que ψ é um vetor normalizado. Essa é a conhecida relação de incerteza momento-posição que foi originalmente encontrada por Heisenberg, e que significa que, em medições efetuadas simultaneamente em um determinado estado, não podemos ter a dispersão das medidas de posição e a dispersão das medidas de momento arbitrariamente pequenas.

12.3 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} podemos escrevê-lo como sendo

$$\hat{H} = \frac{1}{2m} \frac{d^2}{dx^2} + V(x) = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x)$$

ou seja, temos um operador diferencial.

A equação de Schrödinger, que descreve a evolução temporal de um estado, é dada por

$$\hat{H}\Psi = -i\hbar \frac{d}{dt} \Psi$$

onde a função Ψ depende de x e t . Desta forma, temos de fato uma equação diferencial parcial

$$-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \Psi + V\Psi = -i\hbar \frac{\partial}{\partial t} \Psi$$

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.

12.3. O OPERADOR HAMILTONIANO E A EQUAÇÃO DE SCHRÖDINGER⁸⁵

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

$$\frac{-\frac{1}{2m} \frac{d^2}{dx^2} \psi T + V(x) \psi}{\psi} = \frac{-i \frac{d}{dt} T}{T}$$

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:

$$-\frac{1}{2m} \frac{d^2}{dx^2} \psi + V\psi = E\psi$$

$$\frac{d}{dt} T = iET$$

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 um 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 elas correspondem aos autovalores do operador \hat{H} , pois satisfazem

$$\hat{H}\psi = E\psi$$

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 a 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.

12.4 A partícula em uma caixa unidimensional

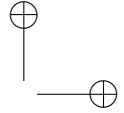
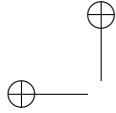
12.4.1 Caso clássico

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

$$ma = F = 0$$

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 $x = L$) então 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 intensidade 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



12.4. A PARTÍCULA EM UMA CAIXA UNIDIMENSIONAL 87

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) então 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 $[a, b]$.

12.4.2 Caso quântico

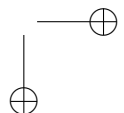
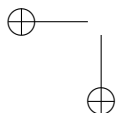
Agora devemos fazer uma descrição quântica do sistema e, dada a sua relativa simplicidade, investigar se há alguma relação facilmente visível entre o clássico e o quântico.

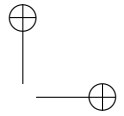
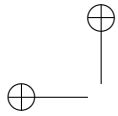
Começamos por encontrar o espaço de Hilbert adequado; como estamos interessados em posições que assumem valores entre 0 e L parece natural considerar o espaço $L^2[0, L]$. Como a partícula não pode sair da caixa, então sua probabilidade de ser encontrada do lado de fora (ou seja, em posições fora de $[0, L]$) é nula e então é natural assumir que a função $\psi(x)$ deverá ser nula fora da caixa e, por continuidade, precisa se anular em 0 e em L (aqui o leitor mais familiarizado com os espaços L^2 poderá se perguntar: mas esses espaços não são de funções apenas mensuráveis? A resposta é sim, mas é preciso levar em conta que estamos utilizando o domínio do operador de Schrödinger, que exige também funções diferenciáveis, e portanto, contínuas); Como a partícula é livre o potencial é nulo e estamos usando a equação de Schrödinger independente do tempo

$$-\frac{d^2}{dx^2}\psi(x) = E\psi(x)$$

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 $\psi_2 = \sin(\sqrt{E}x)$ 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

$$0 = \psi(L) = A_2 \sin \sqrt{E}L$$

Desta forma notamos que os valores possíveis para \sqrt{E} , ou seja, os autovalores, são múltiplos inteiros de π : vamos rotulá-los por $n = 1, 2, \dots$

$$E_n = n^2(\pi/L)^2$$

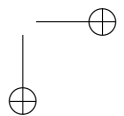
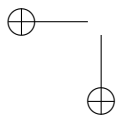
Observação 12.4.1. *O leitor consegue imaginar uma boa razão para não incluirmos $n = 0$ nas soluções acima? Afinal, se $E = 0$ a equação é claramente satisfeita... Bem, note que para $E = 0$ a autofunção correspondente é $\psi(x) = 0$, a função nula. E esta função, multiplicada por uma constante, continua 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.*

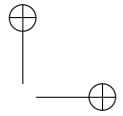
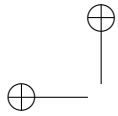
Estes são os autovalores do operador (e são infinitos, ao contrário do que se passava até aqui, quando considerávamos apenas espaços vetoriais de dimensão finita); as autofunções correspondentes (já normalizadas) são, respectivamente,

$$\psi_n(x) = \sqrt{\frac{2}{L}} \sin \frac{n\pi x}{L} \quad n = 1, 2, \dots$$

Um estado em geral é, então, dado por uma combinação linear das autofunções acima:

$$\psi = \sum_{n=1}^{\infty} c_n \psi_n$$





12.5. EXERCÍCIOS

89

12.4.3 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 então a probabilidade de encontrá-la no intervalo $[a, b]$ é dada por

$$\mathbb{P}_n(x \in [a, b]) = \int_{[a,b]} \psi_n^*(s)\psi_n(s)ds = \frac{2}{L} \int_{[a,b]} \sin^2 \frac{n\pi s}{L} ds =$$

$$\frac{2}{L} \left[\int_{[a,b]} (1/2 - 1/2 \cos 2 \frac{n\pi s}{L}) ds \right] = |b - a|/2 - \frac{1}{2L} \frac{L \sin 2 \frac{n\pi s}{L}}{2\pi n} \Big|_a^b$$

Quando tomamos o limite de $n \rightarrow \infty$ o segundo termo tende a zero (pois o seno é uma função limitada) e portanto

$$\mathbb{P}_n(x \in [a, b]) \rightarrow |b - a|/L$$

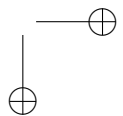
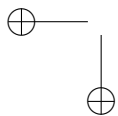
ou seja, para os estados descritos por numeros n elevados (que correspondem fisicamente a situações de energias bem elevadas) obtem-se que a probabilidade de se encontrar uma partícula no intervalo $[a, b]$ está cada vez mais próxima da probabilidade que já havíamos calculado no caso clássico.

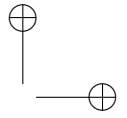
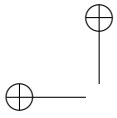
12.5 Exercícios

- 1- Reflita um pouco sobre a influência do tamanho da caixa nos níveis de energia da partícula no caso da seção 12.4.
- 2- Para as autofunções da partícula na caixa obtenha os valores esperados da posição. Lembre que

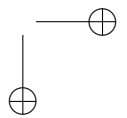
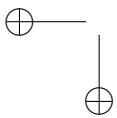
$$\langle x \rangle = \langle \psi, x\psi \rangle = \int_0^L x |\psi(x)|^2 dx$$

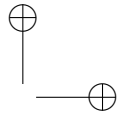
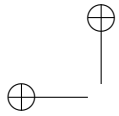
Procedendo de forma similar, obtenha os valores esperados do momento.





- 3- 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?



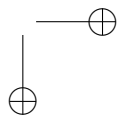
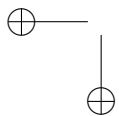


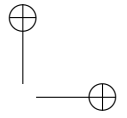
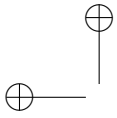
Capítulo 13

Estas notas tem apenas o modesto objetivo de ser um primeiro olhar sobre esta area fascinante que é a mecânica quântica. Se o leitor se sentiu atraído pelo tema então poderá aproveitar algumas das sugestões de leitura que concentramos aqui.

Livros de divulgação científica, a despeito de sua linguagem em geral mais simplificada, podem conter idéias e discussões conceituais muito interessantes e que são certamente úteis para alguém que de fato pretende compreender melhor os fenômenos quânticos, para além de sua pura e simples formulação matemática. Nesse sentido, vale muito a pena consultar “The fabric of reality”, do físico David Deutsch [2]. Outro belo livro é “The fabric of the cosmos”, de Brian Greene [5]; ele discute diversos aspectos da física moderna, incluindo aí a mecânica quântica, de maneira simplificada mas muito interessante.

Para ter uma visão mais completa e num nível matemático em grande parte compreensível por quem seguiu este texto deve-se olhar “Lectures on quantum theory”, um belíssimo livro-texto de Chris Isham

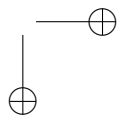
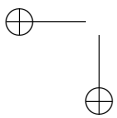




[6]. Também bastante interessante e motivador é o livro "Quantum mechanics: a modern development", de Leslie E. Ballentine [1].

A mecânica quântica fornece um novo paradigma que pode modificar radicalmente o que chamamos de computação; uma discussão bonita e muito esclarecedora deste tema pode ser encontrada no livro "Noções de informação quântica" de Marcelo Terra Cunha [9].

Por último, não custa repetir: as conhecidas "Feynman lectures" [4] são brilhantes e sua leitura é obrigatória e enriquecedora.



Bibliografia

- [1] L. E. Ballentine. **Quantum Mechanics: a modern development** World Scientific, 1998
- [2] D. Deutsch. **The fabric of reality** Penguin Books, 1997
- [3] R. Exel. **Uma introdução às C^* -álgebras**
<http://www.mtm.ufsc.br/~exel/papers/intro.pdf>
- [4] R. Feynman, R.B. Leighton e M. Sands. **The Feynmann lectures on physics, vol.3** Addison-Wesley, 1965
- [5] B. Greene. **The fabric of the cosmos** Penguin Books, 2004
- [6] C. J. Isham. **Lectures on quantum theory** Imperial College Press, Londres 1995
- [7] E. Lima. **Álgebra Linear**(segunda edição) Coleção Matemática Universitária - IMPA, Rio de Janeiro
- [8] J. Sotomayor. **Lições de equações diferenciais ordinárias** Projeto Euclides, Rio de Janeiro 1979
- [9] M. O. Terra Cunha. **Noções de informação quântica** publicações matemáticas do IMPA, Rio de Janeiro 2007.