

O Bit Quântico

A matemática necessária para descrever o sistema quântico mais simples que existe.

BÁRBARA AMARAL

UNIVERSIDADE FEDERAL DE SÃO JOÃO DEL-REI

MARCELO TERRA CUNHA

UNIVERSIDADE ESTADUAL DE CAMPINAS

26 de abril de 2017

Para Thales e Tshabalala, como sempre, e para a Sucata, pela primeira vez.

Para Mimi e Tatá, como sempre, para o Andrey, novamente, e para Rudá, Antony e Ravi pela primeira vez.

Sumário

1	Números Complexos	3
1.1	Soma e Multiplicação	3
1.2	Representação Geométrica	5
1.3	A Exponencial Complexa	5
1.4	Exercícios	6
2	O espaço $\mathbb{R}^2$	7
2.1	Definição e Propriedades	7
2.2	A geometria de $\mathbb{R}^2$	9
2.3	Combinações Lineares	10
2.4	Independência Linear	13
2.5	Subespaços Vetoriais	16
2.5.1	Subespaços gerados	19
2.5.2	Base e Dimensão	20
2.6	Produto Interno	21
3	O espaço $\mathbb{C}^2$	27
3.1	Definição e Propriedades	27
3.2	Subespaços Vetoriais	28
3.3	Combinações Lineares	29
3.4	Independência Linear	30
3.4.1	Subespaços gerados	31
3.4.2	Base e Dimensão	31
3.5	Produto Interno	32
4	Um Bit de Mecânica Quântica	35
4.1	Estados e Medições	36
4.2	Depois das Medições	38
4.3	O que os bits clássicos não têm	39
4.4	Quando perder é ganhar	41
5	Um pouquinho de Física	43
5.1	Ondas eletromagnéticas	43
5.1.1	A dualidade onda-partícula	43

5.1.2	Polarização da Luz	45
5.2	Spin	52
6	Matrizes	55
6.1	Definição	55
6.2	Operações envolvendo matrizes	57
6.2.1	Multiplicação de um escalar por uma matriz	57
6.2.2	Adição e subtração de matrizes	58
6.2.3	Multiplicação de matrizes	59
6.2.4	Transposição	62
6.2.5	Matriz Hermitiana Conjugada	63
6.3	Algumas matrizes especiais	64
6.3.1	Matrizes Diagonais	64
6.3.2	Matrizes triangulares	65
6.3.3	Matrizes Nulas	66
6.3.4	Matrizes Identidade	66
6.3.5	Matrizes Simétricas	67
6.3.6	Matrizes Anti-simétricas	67
6.3.7	Matrizes Hermitianas	67
6.3.8	Matrizes Anti-hermitianas	68
6.3.9	Matrizes Ortogonais	68
6.3.10	Matrizes Unitárias	69
6.4	Representação matricial de $\mathbb{R}^2$ e $\mathbb{C}^2$	70
6.5	Transformações Lineares	70
6.6	Algumas definições especiais para matrizes quadradas	72
6.6.1	Traço	72
6.7	Matriz Inversa	73
6.7.1	Definição e propriedades	73
6.7.2	Determinação da inversa	74
6.8	Determinantes	76
6.8.1	Cálculo do determinante	76
7	Generalização para dimensões maiores	81
7.1	O Espaço $\mathbb{R}^n$	81
7.2	Subespaços Vetoriais	82
7.3	Combinações Lineares	85
7.4	Dependência e Independência linear	88
7.4.1	Subespaços gerados	90
7.4.2	Base e Dimensão	91
7.5	Produto Interno	95
7.6	O Espaço $\mathbb{C}^n$	98
7.7	Definição e Propriedades	98
7.8	Combinações Lineares	99
7.9	Subespaços Vetoriais	101

SUMÁRIO

7

7.10	Dependência e Independência linear	102
7.10.1	Subespaços gerados	103
7.10.2	Base e Dimensão	104
7.10.3	Produto Interno	105
8	Evoluções Unitárias	109
8.1	O grupo $U(2)$	109
8.2	Subgrupos de $U(2)$	110
8.2.1	Subgrupos a um parâmetro	111
8.2.2	A exponencial	111
8.2.3	O grupo $SU(2)$	113
8.3	Alguns outros grupos interessantes	113
9	Aproximação de Unitárias	115
9.1	Operações Quânticas Universais	115
9.2	Subconjuntos densos	116
9.3	Operações Universais de um qbit	118
9.4	E para mais qbits?	119

Prefácio

Primeiramente, Fora Temer!

Em 2011, iniciamos nossa empreitada de apresentar a mecânica quântica começando por seus exemplos mais simples com o livro *Mecânica Quântica para Matemáticos em Formação*, escrito em parceria com o Professor Alexandre Baraviera. O livro foi destinado a um curso ministrado no 28º Colóquio Brasileiro de Matemática realizado naquele mesmo ano e repetido algumas vezes em outras situações. Como mencionamos lá, aquele texto foi o resultado de uma pequena aventura ou uma grande ambição: falar de mecânica quântica para matemáticos em formação. Da experiência bem sucedida, concluímos que, se tomado o devido cuidado, é possível apresentar os principais aspectos da teoria sem assumir pré-requisitos matemáticos muito avançados. O texto que você está lendo agora é fruto de uma grande aventura e uma ambição ainda maior: falar de mecânica quântica para alunos do Ensino Médio.

De maneira geral, os primeiros sistemas estudados em cursos introdutórios de Física Quântica são escolhidos do ponto de vista histórico, o que não necessariamente os torna mais adequados do ponto de vista didático. Para fazer uma apresentação matematicamente consistente, são necessários vários conceitos de Álgebra Linear avançada e Análise Funcional, que trata do estudo de espaços de funções. Essa é uma área belíssima da matemática, mas sua complexidade faz com que tais textos sejam muito duros para estudantes nos primeiros anos da graduação e inacessíveis para alunos do Ensino Médio. As últimas décadas, no entanto, permitiram o crescimento da chamada *teoria quântica da informação*, ou, como é mais comum, *informação quântica*. Um dos maiores méritos desta foi levar a uma revisão dos conceitos fundamentais da mecânica quântica e, em especial, permitir uma maior valorização dos espaços de estado de dimensão finita. Dessa forma, sai a análise funcional (como pré-requisito ou ponto de partida) e entra a álgebra linear. Esse é o espírito do texto: discutir a matemática da mecânica quântica através do seu exemplo mais simples, o que nos leva naturalmente aos $\mathbb{R}^2$ e $\mathbb{C}^2$. Assim, são suficientes conceitos básicos de Álgebra Linear, o que torna o conteúdo acessível para alunos ainda no ensino médio.

Essa grande aventura começou com o *Programa Mentores* da Olimpíada Brasileira de Matemática das Escolas Públicas, para o qual foi preparado o curso à distância *Fundamentos Matemáticos da Física Quântica*, dividido em dois módulos. Nesse curso, é apresentado o conteúdo contido nesse texto e algumas generaliza-

ções. Desse modo, toda a organização do programa e principalmente os alunos que vieram se aventurar conosco durante os três semestres em que o curso foi oferecido merecem nosso agradecimento especial. Agradecemos também ao Professor Mário Jorge Dias Carneiro por se lembrar de nós em sua busca por atividades interessantes para a VIII Biental da Matemática. É uma grande honra. Seus exemplos seguem nos norteando!

Por fim, é hora e lugar de agradecermos e nos desculparmos, mais uma vez, com aqueles entes próximos e queridos, que concordaram com tantas renúncias em nome de mais um projeto que, finalmente, ganhou forma.

Bárbara Amaral & Marcelo Terra Cunha

Capítulo 1

Números Complexos

O conjunto dos números complexos tem um universo infinito de aplicações. Em muitos casos eles podem facilitar os cálculos e abreviar a notação e permitir cálculos que seriam impossíveis utilizando apenas números reais. A Mecânica Quântica faz uso dos números complexos, mas aqui eles não são só um atalho para simplificar a teoria. A importância deles é tamanha que alguns físicos afirmam que é impossível formulá-la utilizando apenas os números reais. Faremos aqui apenas um resumo das principais propriedades que serão necessárias ao longo do texto e para mais detalhes o leitor pode consultar [1, 2, 3].

1.1 Soma e Multiplicação

Definição 1. Um *corpo* é um conjunto C em que podemos definir duas operações

$$\begin{aligned} + : C \times C &\longrightarrow C \\ (a, b) &\longmapsto a + b \end{aligned} \quad (1.1)$$

$$\begin{aligned} \cdot : C \times C &\longrightarrow C \\ (a, b) &\longmapsto a \cdot b = ab \end{aligned} \quad (1.2)$$

chamadas, respectivamente, de *soma* e *multiplicação*, tais que para todos $a, b, c \in C$ valem

- 1) (Associatividade) $a + (b + c) = (a + b) + c$ e $a \cdot (b \cdot c) = (a \cdot b) \cdot c$;
- 2) (Comutatividade) $a + b = b + a$ e $a \cdot b = b \cdot a$;
- 3) (Existência de elemento neutro) existem elementos distintos $0 \in C$ e $1 \in C$ tais que $a + 0 = a$ e $a \cdot 1 = a$;
- 4) (Existência de inversos) Para todo $a \in C$ existe $-a \in C$ tal que $a + (-a) = 0$ e se $a \neq 0$ existe $a^{-1} \in C$ tal que $a \cdot a^{-1} = 1$;
- 5) (Distributividade) $a \cdot (b + c) = a \cdot b + a \cdot c$.

Exercício 1. O conjunto dos números racionais $\mathbb{Q}$ e o conjunto dos números reais $\mathbb{R}$ são corpos com as operações usuais de soma e multiplicação.

Os corpos que vão aparecer com mais frequência ao longo do texto são o corpo dos números reais e o corpo dos números complexos. Vamos assumir que o leitor já está familiarizado com as operações definidas em $\mathbb{R}$. Para os números complexos faremos aqui uma breve introdução de suas principais propriedades.

Definição 2. Um *número complexo* é uma expressão do tipo:

$$z = x + iy,$$

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

O conjunto formado por todos os números complexos será denotado por $\mathbb{C}$. Para definir a soma e a multiplicação de números complexos vamos usar as operações de soma e multiplicação de números reais e considerar cada número complexo como um polinômio em i .

Definição 3. A *soma* de dois números complexos $z_1 = x_1 + iy_1$ e $z_2 = x_2 + iy_2$ é dada por

$$z_1 + z_2 = (x_1 + x_2) + i(y_1 + y_2). \quad (1.3)$$

Definição 4. O *produto* de z_1 e z_2 é dado por

$$z_1 z_2 = x_1 x_2 + ix_1 y_2 + ix_2 y_1 + i^2 y_1 y_2 = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1). \quad (1.4)$$

Exercício 2. Mostre que as operações definidas acima são comutativas e que a multiplicação se distribui sobre a adição. Mostre também que o elemento neutro para a adição é $0 = 0 + i0$, que o elemento neutro para a multiplicação, também chamado de identidade, é $1 = 1 + i0$ e que o inverso de z para a soma é $-z = -x - iy$.

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

Definição 5. O *conjugado* de um número complexo $z = x + iy$ é o número complexo

$$\bar{z} = x - iy. \quad (1.5)$$

O *módulo* de z é definido por

$$|z| = \sqrt{z \cdot \bar{z}} = \sqrt{x^2 + y^2}. \quad (1.6)$$

Um número complexo z é chamado *unitário* se $|z| = 1$.

Exercício 3. Mostre que a módulo um número complexo é sempre um número real não negativo e temos que $|z| = 0$ se, e somente se, $z = 0$.

Exercício 4. Mostre que

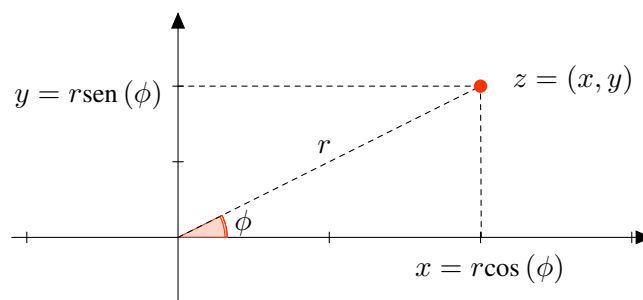
$$z^{-1} = \frac{\bar{z}}{|z|^2} \quad (1.7)$$

é o inverso multiplicativo de z e que se z é unitário, $z^{-1} = \bar{z}$.

Teorema 1. *O conjunto $\mathbb{C}$ é um corpo com a soma e a multiplicação definidas, respectivamente, pelas equações (1.3) e (1.4). O elemento neutro para a adição é $0 = 0 + i0$, o elemento neutro para a multiplicação, é $1 = 1 + i0$, o inverso de z para a soma é $-z = -x - iy$ e o inverso de $z \neq 0$ para a multiplicação é $z^{-1} = \frac{\bar{z}}{|z|^2}$.*

1.2 Representação Geométrica

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



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

$$z = r (\cos(\phi) + i \sin(\phi)). \quad (1.8)$$

Essa é a chamada *forma geométrica* do número complexo z . Desse modo, se z é um complexo unitário então

$$z = \cos(\phi) + i \sin(\phi) \quad (1.9)$$

para algum $\phi \in \mathbb{R}$.

1.3 A Exponencial Complexa

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

Definição 6. A exponencial de um número complexo $z = x + iy$ é definida por

$$e^z = e^{x+iy} = e^x (\cos(y) + i \operatorname{sen}(y)). \quad (1.10)$$

Exercício 5. Os números complexos da forma $z = iy$, que possuem parte real igual a zero, são chamados *imaginários puros*. Mostre que a exponencial e^{iy} de um imaginário puro é um número complexo unitário.

É possível mostrar que valem as seguintes propriedades

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

Utilizando a exponencial complexa, é possível reescrever a representação geométrica de um número complexo, o que nos leva à chamada *representação polar* de um número complexo:

$$z = r (\cos(\phi) + i \operatorname{sen}(\phi)) = re^{i\phi}. \quad (1.11)$$

1.4 Exercícios

Exercício 6. Coloque os números complexos abaixo na forma $x + iy$.

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

Exercício 7. Esboce no plano cartesiano os subconjuntos de $\mathbb{C}$ que satisfazem as seguintes propriedades

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

Exercício 8. Calcule

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

Capítulo 2

O espaço $\mathbb{R}^2$

Em vários conjuntos podemos definir uma operação envolvendo os elementos do próprio conjunto, que chamaremos genericamente de *soma*, e uma operação envolvendo elementos do conjunto e números reais, que chamaremos de *multipliação por escalar*.

Em vários desses conjuntos essas operações satisfazem certas propriedades que as tornam relevantes tanto do ponto de vista puramente matemático quando do ponto de vista prático. Quando isso acontece o conjunto em questão é chamado de *espaço vetorial sobre $\mathbb{R}$* . Estudar esse tipo de conjunto e suas propriedades é importante para várias áreas da ciência. Apresentaremos nesse capítulo um dos exemplos mais simples de espaço vetorial: o espaço vetorial $\mathbb{R}^2$. O leitor interessado pode também consultar as referências [4, 5, 6, 7].

2.1 Definição e Propriedades

Definição 7. Representamos por $\mathbb{R}^2$ o conjunto de todas as duplas de números reais, que denotaremos por

$$|u\rangle = (x, y). \quad (2.1)$$

Essas duplas de números reais podem ser representadas por pontos no plano cartesiano, como mostra a figura 2.1, mas podem também ser representadas por setas que ligam a origem $\mathbf{0} = (0, 0)$ ao ponto de coordenadas (x, y) , como mostra a figura 2.2. Os elementos de $\mathbb{R}^2$ serão chamados *vetores*.

Definição 8. Dados $|u\rangle = (x_1, y_1)$ e $|v\rangle = (x_2, y_2)$ em $\mathbb{R}^2$ definimos a *soma*

$$|u\rangle + |v\rangle = (x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2) \quad (2.2)$$

e o *produto* de $|u\rangle$ por um escalar $\lambda \in \mathbb{R}$

$$\lambda |u\rangle = \lambda (x_1, y_1) = (\lambda x_1, \lambda y_1). \quad (2.3)$$

Teorema 2. Para todos $|u\rangle, |v\rangle, |w\rangle \in \mathbb{R}^2$ e $\lambda, \nu \in \mathbb{R}$ as operações de soma e produto por escalar satisfazem as propriedades:

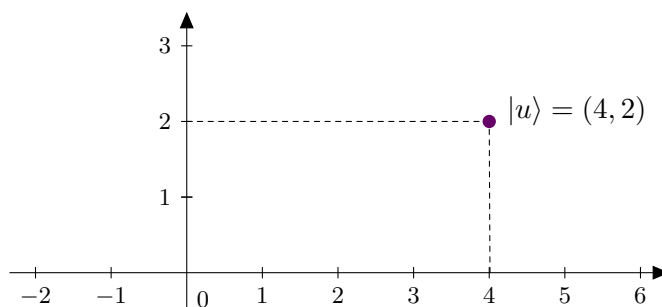


Figura 2.1: Representação geométrica do espaço $\mathbb{R}^2$ no plano cartesiano, em que é marcado o ponto $P = (4, 2)$, correspondente ao vetor $|u\rangle = (4, 2)$.

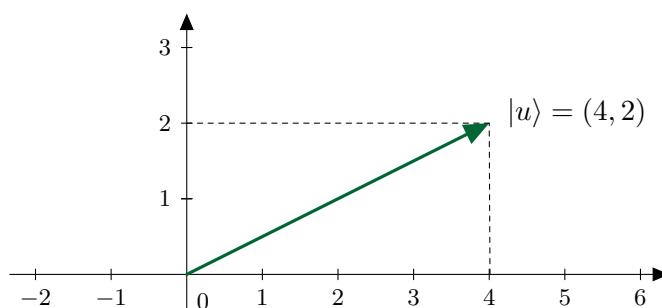


Figura 2.2: O elemento $(4, 2) \in \mathbb{R}^2$ também pode ser representado pela seta que liga a origem do plano cartesiano $(0, 0)$ ao ponto de coordenadas $P = (4, 2)$.

1. (Associatividade) $|u\rangle + (|v\rangle + |w\rangle) = (|u\rangle + |v\rangle) + |w\rangle$;
2. (Comutatividade) $|u\rangle + |v\rangle = |v\rangle + |u\rangle$;
3. (Existência de zero) O elemento $\mathbf{0} = (0, 0) \in \mathbb{R}^2$ é tal que $|u\rangle + \mathbf{0} = |u\rangle$;
4. (Existência de inverso aditivo) Dado $|u\rangle = (x, y) \in \mathbb{R}^2$, o vetor $-|u\rangle = (-x, -y) \in \mathbb{R}^2$ é tal que $|u\rangle + (-|u\rangle) = \mathbf{0}$;
5. (Associatividade) $\lambda(\nu |u\rangle) = (\lambda\nu) |u\rangle$;
6. (Distributividade) $\lambda(|u\rangle + |v\rangle) = \lambda |u\rangle + \lambda |v\rangle$;
7. (Distributividade) $(\lambda + \nu) |u\rangle = \lambda |u\rangle + \nu |u\rangle$;
8. $1 |u\rangle = |u\rangle$.

Exercício 9. Mostre que as operações de soma e produto escalar definidas em $\mathbb{R}^2$ pelas Equações (2.2) e (2.3), respectivamente, de fato satisfazem todas as propriedades enumeradas no Teorema 2.

2.2 A geometria de $\mathbb{R}^2$

As operações de soma e multiplicação por escalar em $\mathbb{R}^2$ podem ser representadas geometricamente, o que possibilita que as definições e resultados que veremos ao longo das próximas seções possam ser visualizados através da geometria.

Em $\mathbb{R}^2$, cada vetor $|u\rangle = (x, y)$ pode ser representado por um segmento orientado que liga a origem do plano cartesiano ao ponto com coordenadas (x, y) . A figura 2.3 mostra a representação geométrica do vetor $|u\rangle = (3, 4) \in \mathbb{R}^2$.

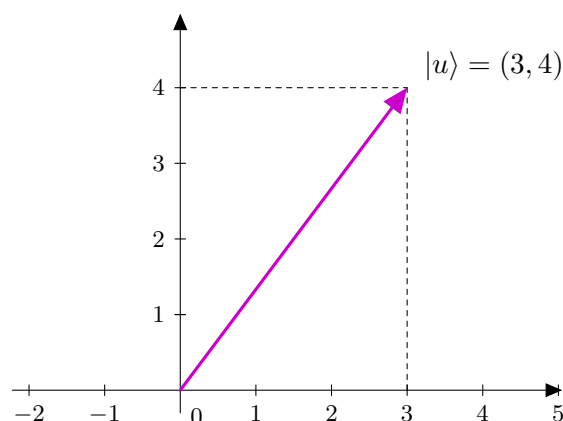


Figura 2.3: Representação geométrica do vetor $\vec{u} = (3, 4) \in \mathbb{R}^2$.

A soma de dois vetores também pode ser representada geometricamente. Dados dois vetores $|u\rangle$ e $|v\rangle$, sua soma é igual ao vetor obtido da seguinte forma: tomamos o segmento orientado que representa $|u\rangle$; em seguida, tomamos o segmento orientado que representa $|v\rangle$ com origem na extremidade de $|u\rangle$; o vetor $|u\rangle + |v\rangle$ é representado pelo segmento orientado que vai da origem até a extremidade de $|v\rangle$. A soma dos vetores $|u\rangle = (3, 4)$ e $|v\rangle = (6, 2)$ é ilustrada na figura 2.4.

A multiplicação de um vetor por um escalar também pode ser representada geometricamente. Dado um vetor $|v\rangle \in \mathbb{R}^2$ e $\lambda \in \mathbb{R}$, o vetor $\lambda |v\rangle$ é encontrado da seguinte forma: se $\lambda = 0$ então $\lambda |v\rangle = \mathbf{0}$; caso contrário, $\lambda |v\rangle$ tem comprimento $|\lambda|$ vezes o comprimento de $|v\rangle$ e mesma direção de $|v\rangle$ (dizemos que eles são *paralelos*); $\lambda |v\rangle$ tem o mesmo sentido de $|v\rangle$ se $\lambda > 0$ e sentido oposto se $\lambda < 0$. A figura 2.5 ilustra os vetores $|v\rangle = (3, 4)$, $2|v\rangle$, $\frac{|v\rangle}{2}$, $-|v\rangle$, $-2|v\rangle$, $-\frac{|v\rangle}{2}$.

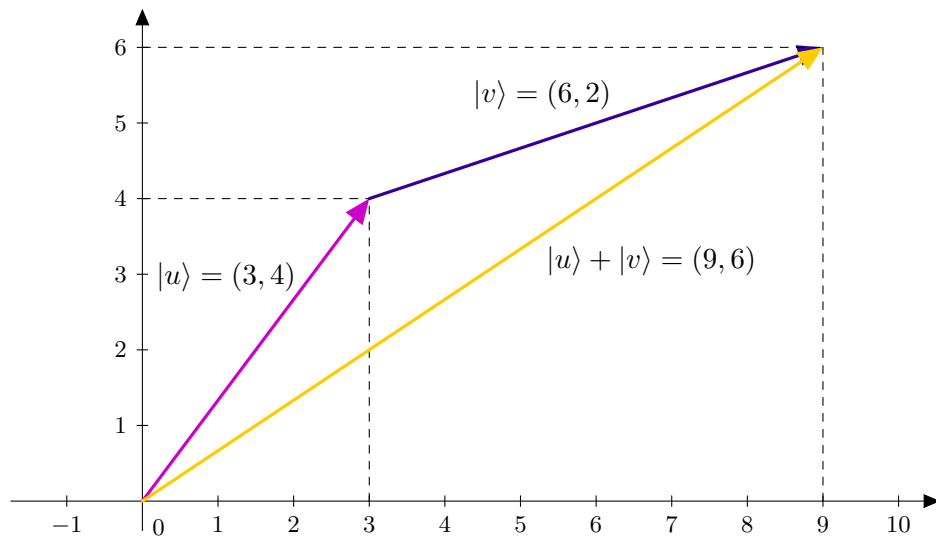


Figura 2.4: Representação geométrica dos vetores $|u\rangle = (3, 4)$ e $|v\rangle = (6, 2)$ e de sua soma $|u\rangle + |v\rangle = (9, 6)$.

2.3 Combinações Lineares

Em $\mathbb{R}^2$, os elementos podem ser combinados livremente através da soma e da multiplicação por escalar, gerando outros vetores de $\mathbb{R}^2$.

Definição 9. Dizemos que $|v\rangle$ é uma *combinação linear* de $|u_1\rangle, \dots, |u_k\rangle$ se existem constantes $a_1, \dots, a_k \in \mathbb{R}$ tais que

$$|v\rangle = a_1 |u_1\rangle + \dots + a_k |u_k\rangle.$$

Exemplo 1. O vetor $(7, 2) \in \mathbb{R}^2$ é uma combinação linear dos vetores $(2, 1)$ e $(1, 0)$ porque

$$(7, 2) = 2 \times (2, 1) + 3 \times (1, 0).$$

Exemplo 2. O vetor nulo $\mathbf{0}$ é sempre combinação linear de qualquer conjunto de vetores $|v_1\rangle, \dots, |v_k\rangle$. De fato, temos que

$$\mathbf{0} = 0 \times |v_1\rangle + \dots + 0 \times |v_k\rangle.$$

Exemplo 3. Considere os vetores $|e_1\rangle = (1, 0)$, $|e_2\rangle = (0, 1)$. Qualquer vetor em $\mathbb{R}^2$ é combinação linear desses dois vetores. De fato, dado $|v\rangle = (a_1, a_2)$ podemos escrever

$$\begin{aligned}(a_1, a_2) &= (a_1, 0) + (0, a_2) \\ &= a_1 \times (1, 0) + a_2 \times (0, 1) \\ &= a_1 |e_1\rangle + a_2 |e_2\rangle.\end{aligned}$$

Exemplo 4. Se um vetor $|v\rangle$ é combinação linear de um vetor $|u\rangle$ então

$$|v\rangle = \lambda |u\rangle$$

ou seja, $|v\rangle$ é múltiplo de $|u\rangle$.

Exemplo 5. Considere os vetores $|u_1\rangle = (1, 0)$ e $|u_2\rangle = (1, 1)$. O vetor $|v\rangle = (4, 2)$ é combinação linear de $|u_1\rangle$ e $|u_2\rangle$. Para provar esse fato, suponhamos inicialmente que $|v\rangle$ possa ser escrito como combinação de $|u_1\rangle$ e $|u_2\rangle$. Isso quer dizer que existem $a_1, a_2 \in \mathbb{R}$ tais que

$$\begin{aligned} |v\rangle &= a_1 |u_1\rangle + a_2 |u_2\rangle \\ (4, 2) &= a_1 (1, 0) + a_2 (1, 1) \\ (4, 2) &= (a_1 + a_2, a_2). \end{aligned} \tag{2.4}$$

Isso nos leva ao sistema de equações envolvendo a_1 e a_2

$$\begin{cases} a_1 + a_2 = 4 \\ a_2 = 2 \end{cases}$$

cuja solução é $a_1 = 2$ e $a_2 = 2$. De fato,

$$(4, 2) = 2 \times (1, 0) + 2 \times (1, 1).$$

Exemplo 6. Considere os vetores $|u_1\rangle = (1, 1)$ e $|u_2\rangle = (2, 2)$. O vetor $|v\rangle = (1, 2)$ não é combinação linear de $|u_1\rangle$ e $|u_2\rangle$. Para provar esse fato, suponhamos inicialmente que $|v\rangle$ possa ser escrito como combinação de $|u_1\rangle$ e $|u_2\rangle$. Isso quer dizer que existem $a_1, a_2 \in \mathbb{R}$ tais que

$$\begin{aligned} |v\rangle &= a_1 |u_1\rangle + a_2 |u_2\rangle \\ (1, 2) &= a_1 (1, 1) + a_2 (2, 2) \\ (1, 2) &= (a_1 + 2a_2, a_1 + 2a_2), \end{aligned}$$

o que nos leva ao sistema de equações

$$\begin{cases} a_1 + 2a_2 = 1 \\ a_1 + 2a_2 = 2 \end{cases}$$

que claramente não podem ser satisfeitas simultaneamente. Logo esse sistema não possui solução, e portanto as constantes a_1 e a_2 com as propriedades desejadas não existem.

De fato, qualquer combinação linear desses vetores é da forma

$$a_1 \times (1, 1) + a_2 \times (2, 2) = (a_1 + 2a_2, a_1 + 2a_2)$$

sempre possui as duas coordenadas iguais, e portanto nunca pode ser igual ao vetor $(1, 2)$.

De maneira geral, dados vetores $|u_1\rangle, \dots, |u_m\rangle \in \mathbb{R}^n$, um vetor $|v\rangle$ será combinação linear de $|u_1\rangle, \dots, |u_m\rangle$ se existirem constantes $a_1, \dots, a_m$ tais que

$$|v\rangle = a_1 |u_1\rangle + \dots + a_m |u_m\rangle.$$

Para descobrir se essas constantes de fato existem e, se existem, quanto elas valem, devemos resolver um sistema de equações que envolvem essas constantes. Por enquanto, vamos trabalhar com poucos vetores de maneira que os sistemas que teremos que resolver serão simples e em geral podemos resolvê-los isolando uma das variáveis em uma das equações e substituindo nas outras. De maneira geral, no entanto, é necessário resolver sistemas mais complicados, o que pode ser feito através de ferramentas mais sofisticadas que não serão abordadas nesse curso. O leitor interessado pode encontrar uma apresentação detalhada de tais ferramentas nas referências [4, 5, 6, 7].

Exercício 10. Verifique se o vetor $|v\rangle$ é combinação linear dos vetores $|u_1\rangle, \dots, |u_n\rangle$ dados abaixo. Em caso afirmativo, encontre a combinação linear desses vetores que gera $|v\rangle$.

1. $|v\rangle = (1, 1); |u_1\rangle = (1, 0);$
2. $|v\rangle = (2, 2); |u_1\rangle = (1, 1);$
3. $|v\rangle = (2, 3); |u_1\rangle = (1, 0), |u_2\rangle = (0, 1);$
4. $|v\rangle = (2, 3); |u_1\rangle = (1, 0), |u_2\rangle = (1, 1);$
5. $|v\rangle = (2, 1); |u_1\rangle = (1, 0), |u_2\rangle = (1, 1), |u_3\rangle = (1, -1);$
6. $|v\rangle = (-1, 3); |u_1\rangle = (1, 0), |u_2\rangle = (1, 1), |u_3\rangle = (1, -1).$

2.4 Independência Linear

Queremos encontrar condições que informem se dentre um conjunto de vetores, algum deles pode ser escrito como combinação linear dos outros. Para isso definimos a noção de dependência linear.

Definição 10. Dizemos que um conjunto de vetores $\{|u_1\rangle, \dots, |u_k\rangle\} \subset \mathbb{R}^2$ é *linearmente independente* (LI) se a equação

$$a_1 |u_1\rangle + \dots + a_k |u_k\rangle = \mathbf{0}$$

só admite a solução trivial $a_1 = \dots = a_k = 0$, ou seja, se a única combinação linear dos vetores que gera o vetor nulo é aquela em que todas as constantes são nulas. Caso contrário, dizemos que os vetores são *linearmente dependentes* (LD).

Um conjunto de vetores é LD se existe uma combinação linear deles que gera o vetor nulo em que pelo menos uma das constantes é não-nula.

Exemplo 7. Um conjunto de vetores que contém o vetor nulo é sempre LD. De fato, tomando o conjunto $\{|u_1\rangle, \dots, |u_k\rangle, \mathbf{0}\}$ podemos escrever a combinação linear

$$0 \times |u_1\rangle + \dots + 0 \times |u_k\rangle + 1 \times \mathbf{0} = \mathbf{0}.$$

Temos então uma combinação linear dos vetores que gera o vetor nulo em que a última constante utilizada é diferente de zero. Isso implica que esses vetores são LD.

Exemplo 8. Um conjunto formado por um único vetor não nulo $|u\rangle$ é sempre LI. Nesse caso, as únicas combinações lineares possíveis são múltiplos de $|u\rangle$ e para que um desses múltiplos seja o vetor nulo temos que

$$\lambda |u\rangle = \mathbf{0}$$

o que implica que $\lambda = 0$, uma vez que $|u\rangle \neq \mathbf{0}$.

Exemplo 9. Os vetores $(1, 0)$, $(1, 1)$ e $(3, 2)$ são LD. De fato, fazendo a combinação linear

$$a_1 (1, 0) + a_2 (1, 1) + a_3 (3, 2) = \mathbf{0}$$

temos que

$$\begin{cases} a_1 + a_2 + 3a_3 = 0 \\ a_2 + 2a_3 = 0 \end{cases}.$$

Da segunda equação temos que $a_2 = -2a_3$ e da primeira equação temos que $a_1 - 2a_3 + 3a_3 = 0$, o que implica que $a_1 = -a_3$. Assim, esse sistema de equações possui uma infinidade de soluções e concluímos que

$$-a_3 (1, 0) - 2a_3 (1, 1) - a_3 (3, 2) = \mathbf{0}$$

para qualquer valor de a_3 . Fazendo, por exemplo, $a_3 = 1$ obtemos

$$-(1, 0) - 2(1, 1) - (3, 2) = \mathbf{0},$$

ou seja, temos uma combinação linear dos vetores $(1, 0)$, $(1, 1)$ e $(3, 2)$ que gera o vetor nulo sem que todas as constantes sejam nulas. Isso implica que esses vetores são LD.

Exemplo 10. Os vetores $(1, 1)$, $(1, 0)$ são LI. De fato, supondo que

$$a_1 (1, 1) + a_2 (1, 0) = \mathbf{0}$$

obtemos o sistema de equações

$$\begin{cases} a_1 + a_2 = 0 \\ a_1 = 0 \end{cases}$$

que possui somente a solução $a_1 = a_2 = 0$.

2.4. INDEPENDÊNCIA LINEAR

15

Três ou mais vetores em $\mathbb{R}^2$ são sempre LD. Isso acontece porque o problema de identificar se esses vetores são ou não LI leva a um sistema que sempre possui solução não trivial.

Teorema 3. *O número máximo de vetores LI em $\mathbb{R}^2$ é igual a 2.*

Demonstração. Considere o conjunto de vetores $|u_1\rangle, \dots, |u_k\rangle$ com $k > 2$. Para saber se esses vetores são LD, devemos encontrar constantes $a_1, \dots, a_k$ que tornem a equação

$$a_1 |u_1\rangle + \dots + a_k |u_k\rangle = \mathbf{0}$$

verdadeira. Essa equação nos leva a um sistema com duas equações e k incógnitas, que sempre possui solução não trivial. A demonstração desse resultado está fora do escopo desse texto e o leitor interessado pode encontrar a demonstração nas referências [4, 5, 6, 7]. $\square$

A definição de dependência linear é importante porque ela permite identificar se dentro de um conjunto de vetores, um deles pode ser escrito como combinação linear dos outros ou não. De fato, as duas propriedades são equivalentes.

Teorema 4. *Um conjunto de vetores é LD se, e somente, se podemos expressar ao menos um dos vetores como combinação linear dos outros.*

Demonstração. Suponhamos que $\{|v_1\rangle, \dots, |v_k\rangle\}$ seja um conjunto LI. Então existe uma combinação linear desses vetores que resulta no vetor nulo

$$a_1 |v_1\rangle + \dots + a_k |v_k\rangle = \mathbf{0}$$

com pelo menos uma das constantes $a_i \neq 0$. Suponhamos que $a_1 \neq 0$. Podemos então escrever

$$|v_1\rangle = -\frac{a_2}{a_1} |v_2\rangle - \dots - \frac{a_k}{a_1} |v_k\rangle,$$

o que mostra que $|v_1\rangle$ é combinação linear de $|v_2\rangle, \dots, |v_k\rangle$.

Por outro lado, se um dos vetores, digamos $|v_1\rangle$, é combinação linear dos outros, então existem constantes $b_2, \dots, b_k$ tais que

$$|v_1\rangle = b_2 |v_2\rangle + \dots + b_k |v_k\rangle,$$

o que implica que

$$|v_1\rangle - b_2 |v_2\rangle - \dots - b_k |v_k\rangle = \mathbf{0},$$

o que mostra que os vetores $\{|v_1\rangle, \dots, |v_k\rangle\}$ são LD. $\square$

Exemplo 11. Um conjunto com dois vetores é LD se, e somente se, um deles é múltiplo do outro. De fato, se dois vetores são LD, um deles pode ser escrito como combinação linear do outro.

Exercício 11. Verifique se os conjuntos de vetores abaixo são LI ou LD.

1. $(1, 1)$;
2. $(1, 2), (3, 6)$;
3. $(1, 2), (3, 1)$;
4. $(1, 2), (3, 1), (1, 1)$;
5. $(1, 1), (1, 0), (4, 6)$.

Exercício 12. Em cada item do exercício anterior em que os vetores são LD, escreva uma combinação linear dos vetores que resulta o vetor nulo em que pelo menos uma das constantes usadas seja diferente de zero. Escreva um dos vetores como combinação linear dos outros.

Exercício 13. Se os vetores $|u\rangle, |v\rangle, |w\rangle$ são linearmente dependentes então $|w\rangle$ é uma combinação linear de $|u\rangle$ e $|v\rangle$?

2.5 Subespaços Vetoriais

Definição 11. Um *subespaço vetorial* S de $\mathbb{R}^2$ é um subconjunto não vazio de $\mathbb{R}^2$ para o qual as seguintes propriedades são satisfeitas:

1. $|u\rangle + |v\rangle \in S$ para todo par $|u\rangle$ e $|v\rangle \in S$;
2. $\lambda |u\rangle \in S$ para todo $\lambda \in \mathbb{R}$ e todo $|u\rangle \in S$.

Quando S é um subespaço, ele é preservado pelas operações de soma e multiplicação por escalar, isto é, essas operações, ao serem aplicadas a elementos de S , produzem como resultado vetores que também são elementos de S . Essa é uma restrição muito forte: apenas uma classe especial de subconjuntos de $\mathbb{R}^2$ possui essa propriedade.

Exemplo 12. Considere em $\mathbb{R}^2$ o conjunto de vetores cujas extremidades pertencem ao conjunto S ilustrado na figura 2.6. Observe que ao somar dois vetores de S o resultado pode estar fora de S . Isso quer dizer que S não é um subespaço de $\mathbb{R}^2$.

Exemplo 13. Considere em $\mathbb{R}^2$ o conjunto de vetores cujas extremidades pertencem à reta S ilustrada na figura 2.7. Observe que ao somar dois vetores de S o resultado pode estar fora de S . Isso quer dizer que S não é um subespaço de $\mathbb{R}^2$.

Exemplo 14. Considere em $\mathbb{R}^2$ o conjunto de vetores cujas extremidades pertencem à reta S ilustrada na figura 2.8. Observe que ao somar dois vetores de S o resultado está sempre em S . O mesmo acontece com a multiplicação por escalar. Isso quer dizer que S é um subespaço de $\mathbb{R}^2$.

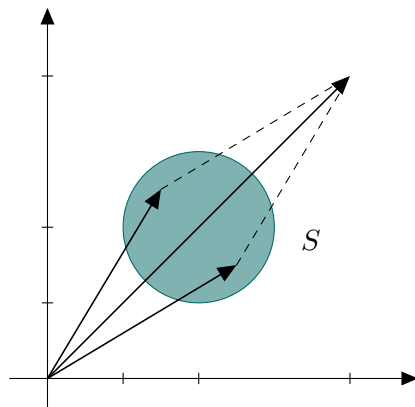


Figura 2.6: Um subconjunto $S \subset \mathbb{R}^2$ que não é um subespaço vetorial de $\mathbb{R}^2$.

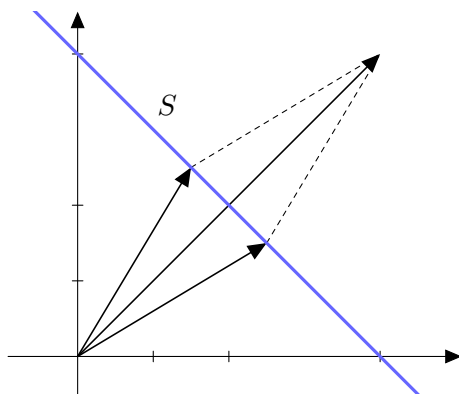


Figura 2.7: Um subconjunto $S \subset \mathbb{R}^2$ que não é um subespaço vetorial de $\mathbb{R}^2$.

Exemplo 15. O subconjunto

$$S = \{(t, 0) \in \mathbb{R}^2; t \in \mathbb{R}\}$$

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

Solução. Nos exemplos anteriores, utilizamos a geometria de $\mathbb{R}^2$ para entender se o subconjunto S é ou não um subespaço vetorial. Agora vamos utilizar a álgebra para fazer essa verificação.

A solução de um problema como esse passa sempre por três passos. O primeiro deles é verificar qual é a propriedade que define o subconjunto S . Nesse exemplo, o subconjunto S é o conjunto de todos os vetores de $\mathbb{R}^2$ que possuem a segunda coordenada igual a zero.

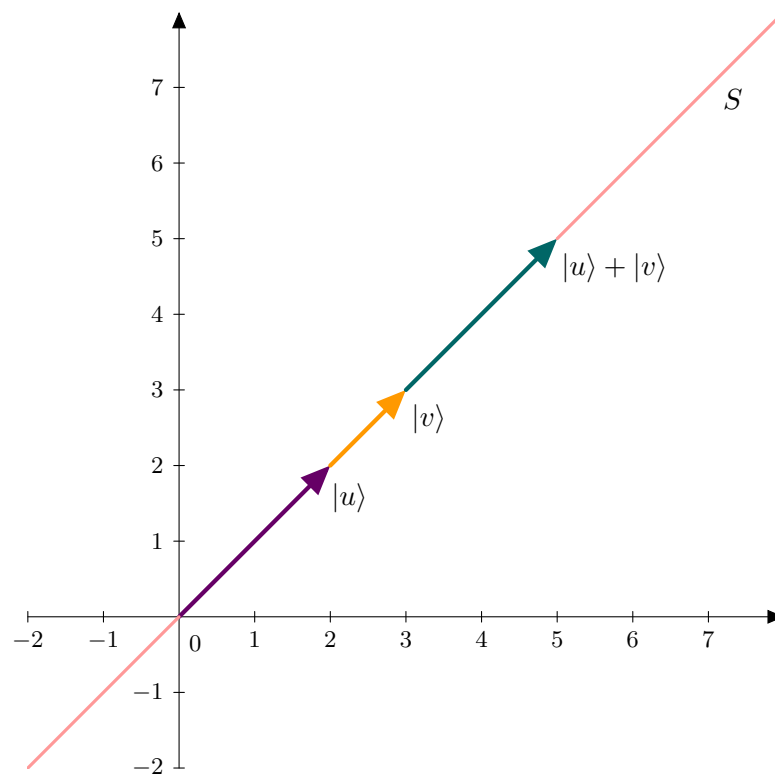


Figura 2.8: Um subconjunto $S \subset \mathbb{R}^2$ que é um subespaço vetorial de $\mathbb{R}^2$.

O segundo passo é verificar se a soma de quaisquer elementos de S permanece em S . Suponhamos então que $|u\rangle$ e $|v\rangle \in S$. Isso quer dizer que $|u\rangle = (t_1, 0)$ e $|v\rangle = (t_2, 0)$ com t_1 e $t_2 \in \mathbb{R}$. Assim

$$|u\rangle + |v\rangle = (t_1 + t_2, 0)$$

também pertence a S .

O terceiro passo é verificar se a multiplicação de elementos de S por constante permanece em S . Tomando $|u\rangle = (t_1, 0) \in S$ qualquer e $\lambda \in \mathbb{R}$ temos

$$\lambda|u\rangle = (\lambda t_1, 0)$$

que também pertence a S . Como as propriedades 1 e 2 são satisfeitas, S é um subespaço de $\mathbb{R}^2$.

Observe que esse subespaço corresponde aos vetores sobre o eixo x e portanto S é uma reta passando pela origem.

☐

Exemplo 16. Considere o subconjunto

$$S = \{(1, t) \in \mathbb{R}^2; t \in \mathbb{R}\}.$$

2.5. SUBESPAÇOS VETORIAIS

19

O subconjunto S não é um subespaço vetorial de $\mathbb{R}^2$. De fato, temos que $(1, 1)$ pertence à S , mas $2 \times (1, 1) = (2, 2)$ não pertence a S . Observe que S é uma reta que não passa pela origem.

Exemplo 17. Os subespaços de $\mathbb{R}^2$ correspondem a $S = \{\mathbf{0}\}$, retas passando pela origem ou $S = \mathbb{R}^2$. Veremos mais adiante como podemos provar esse fato.

Exercício 14. Verifique quais são, dentre os subconjuntos de $\mathbb{R}^2$ abaixo, subespaços vetoriais.

1. O subconjunto de vetores que possuem primeira coordenada igual a zero;
2. O subconjunto de vetores que possuem primeira coordenada igual a 1;
3. O subconjunto de vetores (x, y) tais que $xy = 0$;
4. O subconjunto de vetores (x, y) tais que $y + 3x = 0$;
5. O subconjunto de vetores (x, y) tais que $y + 3x = 1$.

Exercício 15. Se S_1 e S_2 são subespaços vetoriais de $\mathbb{R}^2$, mostre que $S_1 \cap S_2$ também é um subespaço vetorial de $\mathbb{R}^2$. O mesmo acontece com $S_1 \cup S_2$?

2.5.1 Subespaços gerados

Definição 12. Dado um conjunto de vetores $|u_1\rangle, \dots, |u_k\rangle$ em $\mathbb{R}^2$, o *subespaço gerado* por esses vetores é o conjunto de todas as suas combinações lineares:

$$\langle |u_1\rangle, \dots, |u_k\rangle \rangle = \{a_1 |u_1\rangle + \dots + a_k |u_k\rangle; a_i \in \mathbb{R}\}.$$

Exercício 16. Prove que $\langle |u_1\rangle, \dots, |u_k\rangle \rangle$ é de fato um subespaço de $\mathbb{R}^2$.

Exercício 17. O vetor $(7, 2) \in \mathbb{R}^2$ é uma combinação linear dos vetores $(2, 1)$ e $(1, 0)$ porque

$$(7, 2) = 2 \times (2, 1) + 3 \times (1, 0).$$

Portanto $(7, 2) \in \langle (2, 1), (1, 0) \rangle$.

Exercício 18. O subespaço gerado por um vetor não nulo $|v\rangle \in \mathbb{R}^2$ é uma reta passando pela origem. De fato

$$\langle |v\rangle \rangle = \{t |v\rangle; t \in \mathbb{R}\}$$

que corresponde justamente à reta passando pela origem na direção do vetor $|v\rangle$.

2.5.2 Base e Dimensão

Podemos nos perguntar se existe um conjunto de vetores LI de forma que todo elemento de $\mathbb{R}^2$ possa ser escrito como combinação linear dos elementos desse conjunto. É possível mostrar que $\mathbb{R}^2$ possui infinitos conjuntos LI com essa propriedade.

Definição 13. Uma base para $\mathbb{R}^2$ é um conjunto LI

$$\mathcal{B} = \{|u_1\rangle, \dots, |u_k\rangle\}$$

tal que todo vetor de $\mathbb{R}^2$ é combinação linear de $|u_1\rangle, \dots, |u_k\rangle$, ou seja

$$\langle |u_1\rangle, \dots, |u_k\rangle \rangle = \mathbb{R}^2.$$

De maneira intuitiva, uma base para $\mathbb{R}^2$ é um conjunto com o número mínimo de vetores que precisamos para gerar todos os outros vetores a partir de combinações lineares.

Exercício 19. Os vetores $|u_1\rangle = (1, 0)$ e $|u_2\rangle = (0, 1)$ formam uma base para $\mathbb{R}^2$. De fato, esses vetores são LI e qualquer vetor de $\mathbb{R}^2$ pode ser escrito como combinação linear de $|u_1\rangle$ e $|u_2\rangle$.

O Teorema 3 implica que não pode haver uma base para $\mathbb{R}^2$ com mais de dois elementos, já que tal conjunto seria LD. Por outro lado, um conjunto com apenas um vetor não pode gerar todo o $\mathbb{R}^2$, pois o Exemplo 18 nos mostra que o subespaço gerado por um vetor é uma reta passando pela origem. Isso implica que todas as bases de $\mathbb{R}^2$ possuem dois elementos. Podemos entender esse fato intuitivamente: se temos apenas um vetor, podemos andar em uma direção apenas e vamos conseguir chegar apenas aos vetores cujas extremidades estão sobre uma reta que passa pela origem. Para chegar a todos os outros vetores do plano, precisamos nos locomover em pelo menos duas direções independentes. Por outro lado, se tivermos três vetores, um deles pode ser obtido a partir de combinação linear dos outros dois. Isso quer dizer que esse vetor é “desnecessário”, ou seja, poderíamos chegar aos mesmos vetores utilizando apenas os outros dois vetores. Logo o número mínimo de vetores que precisamos para gerar todo o $\mathbb{R}^2$ fazendo combinações lineares é 2.

Definição 14. A *dimensão* de $\mathbb{R}^2$ é o número de vetores em uma base de $\mathbb{R}^2$. Logo $\mathbb{R}^2$ tem dimensão 2.

Todas as bases de $\mathbb{R}^2$ possuem dois elementos, mas nem todo conjunto com dois elementos é uma base, já que esses vetores podem ser LD. No entanto, se um conjunto com dois vetores é LI, ele é automaticamente uma base.

Teorema 5. Em $\mathbb{R}^2$, qualquer conjunto LI com dois elementos é uma base.

Demonstração. Suponhamos que $|v_1\rangle$ e $|v_2\rangle$ sejam LI. Para mostrar que eles formam uma base, devemos mostrar que eles geram $\mathbb{R}^2$. Dado $|u\rangle$ um vetor qualquer

2. $\langle u | v \rangle = \langle v | u \rangle$;
3. $\langle u | u \rangle \geq 0$;
4. Se $\langle u | u \rangle = 0$ então $|u\rangle = 0$.

Exercício 21. Prove que o produto interno em $\mathbb{R}^2$ satisfaz as propriedades listadas no Teorema 6.

O produto interno entre $|u\rangle = (x_1, y_1)$ e $|v\rangle = (x_2, y_2)$ também pode ser denotado por

$$\langle u | v \rangle = (x_1, y_1) \cdot (x_2, y_2).$$

Exemplo 18. O produto interno entre os vetores $(1, 3)$ e $(4, -2)$ em $\mathbb{R}^2$ é igual a

$$(1, 3) \cdot (4, -2) = (1)(4) + (3)(-2) = 4 - 6 = -2.$$

Definição 16. A partir do produto interno podemos definir a aplicação

$$\| \cdot \| : \mathbb{R}^2 \rightarrow \mathbb{R}$$

escrevendo

$$\|v\| = \sqrt{\langle v | v \rangle}.$$

Essa função é chamada *norma* em $\mathbb{R}^2$.

Definição 17. Dizemos que dois vetores $|u\rangle$ e $|v\rangle$ são *ortogonais* se $\langle u | v \rangle = 0$. Dizemos que um conjunto $E = \{|v_1\rangle, \dots, |v_k\rangle\}$ é *ortogonal* se seus elementos são dois a dois ortogonais. Dizemos que um conjunto $E = \{|v_1\rangle, \dots, |v_k\rangle\}$ é *ortonormal* se é ortogonal e $\langle v_i | v_i \rangle = 1$ para todo i .

Dois vetores ortogonais em $\mathbb{R}^2$ são representados geometricamente por setas perpendiculares.

Exercício 22. Mostre que o número máximo de vetores ortogonais em $\mathbb{R}^2$ é 2.

Exercício 23. Mostre que se u e v são ortogonais, então

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

Teorema 7. Se $\{|v_1\rangle, \dots, |v_k\rangle\}$ é um conjunto ortogonal, então valem as seguintes propriedades:

1. O conjunto $\{|v_1\rangle, \dots, |v_k\rangle\}$ é LI;
2. Se $|v\rangle = a_1 |v_1\rangle + \dots + a_k |v_k\rangle$, então

$$a_i = \frac{\langle v | v_i \rangle}{\|v_i\|^2}.$$

2.6. PRODUTO INTERNO

23

Demonstração.

1. Suponhamos que

$$a_1 |v_1\rangle + a_2 |v_2\rangle + \dots + a_k |v_k\rangle = 0. \quad (2.6)$$

Para provar que o conjunto $\{|v_1\rangle, \dots, |v_k\rangle\}$ é LI, devemos mostrar que a única possibilidade para as constantes a_i na equação acima é que todas elas sejam nulas.

Fazendo o produto interno por $|v_i\rangle$ em ambos os lados da equação acima, temos que

$$\begin{aligned} (a_1 |v_1\rangle + a_2 |v_2\rangle + \dots + a_k |v_k\rangle) \cdot |v_i\rangle &= 0 \cdot |v_i\rangle \\ a_1 (\langle v_1 | v_i \rangle) + a_2 (\langle v_2 | v_i \rangle) + \dots + a_k (\langle v_k | v_i \rangle) &= 0 \\ a_i (\langle v_i | v_i \rangle) &= 0 \\ a_i \|v_i\|^2 &= 0. \end{aligned}$$

Como $\|v_i\| \neq 0$, a equação acima implica que $a_i = 0$. Logo a equação (7.29) só pode ser verdadeira se todas as constantes a_i são nulas e portanto o conjunto $\{|v_1\rangle, \dots, |v_k\rangle\}$ é LI.

2. Suponhamos que

$$|v\rangle = a_1 |v_1\rangle + a_2 |v_2\rangle + \dots + a_k |v_k\rangle. \quad (2.7)$$

Fazendo o produto interno por $|v_i\rangle$ em ambos os lados da equação acima, temos que

$$\begin{aligned} (a_1 |v_1\rangle + a_2 |v_2\rangle + \dots + a_k |v_k\rangle) \cdot |v_i\rangle &= |v\rangle \cdot |v_i\rangle \\ a_1 (\langle v_1 | v_i \rangle) + a_2 (\langle v_2 | v_i \rangle) + \dots + a_k (\langle v_k | v_i \rangle) &= \langle v | v_i \rangle \\ a_i (\langle v_i | v_i \rangle) &= \langle v | v_i \rangle \\ a_i \|v_i\|^2 &= \langle v | v_i \rangle. \end{aligned}$$

A equação acima implica que $a_i = \frac{\langle v | v_i \rangle}{\|v_i\|^2}$, como queríamos provar.

□

Definição 18. Dizemos que uma base $B = \{|v_1\rangle, |v_2\rangle\}$ para $\mathbb{R}^2$ é uma *base ortogonal* se seus elementos são ortogonais. Dizemos uma base $B = \{|v_1\rangle, |v_2\rangle\}$ é uma *base ortonormal* se é ortogonal e $\langle v_i | v_i \rangle = 1$ para todo i .

Exercício 24. A base canônica $B = \{|e_1\rangle = (1, 0), |e_2\rangle = (0, 1)\}$ de $\mathbb{R}^2$ é uma base ortonormal. De fato, é fácil mostrar que o produto interno $\langle e_i | e_j \rangle$ é igual a zero se $i \neq j$ e igual a 1 se $i = j$.

Exercício 25. Mostre que o conjunto

$$B = \left\{ |v_1\rangle = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right), |v_2\rangle = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right) \right\}$$

é uma base ortonormal de $\mathbb{R}^2$.

Solução. Calculando os produtos internos temos

$$\langle v_1 | v_2 \rangle = 0, \quad \langle v_1 | v_1 \rangle = \langle v_2 | v_2 \rangle = 1$$

o que mostra que B é um conjunto ortonormal. Pelo Teorema 31, B é um conjunto LI e portanto é uma base, já que possui dois vetores. $\square$

Outra consequência do Teorema 31 é que dado um vetor qualquer $|v\rangle$ e uma base ortogonal $B = \{|v_1\rangle, |v_2\rangle\}$ de $\mathbb{R}^2$, é fácil encontrar as constantes que devemos usar para escrever $|v\rangle$ como combinação linear de $|v_1\rangle, |v_2\rangle$.

Exercício 26. Escreva $|v\rangle = (2, -1)$ como combinação linear dos vetores

$$|v_1\rangle = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), |v_2\rangle = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right).$$

Solução. Pelo teorema 31, sabemos que se

$$|v\rangle = \alpha_1 |v_1\rangle + \alpha_2 |v_2\rangle$$

então cada constante α_i é dada por

$$\alpha_i = \frac{\langle v | v_i \rangle}{\|v_i\|^2} = \langle v | v_i \rangle.$$

Observe que a expressão se simplifica ainda mais, já que os vetores tem norma igual a 1. Calculando os produtos internos, temos

$$\begin{aligned}\alpha_1 &= \langle v \mid v_1 \rangle = \frac{2}{\sqrt{2}} - \frac{1}{\sqrt{2}} = \frac{1}{\sqrt{2}} \\ \alpha_2 &= \langle v \mid v_2 \rangle = \frac{2}{\sqrt{2}} + \frac{1}{\sqrt{2}} = \frac{3}{\sqrt{2}}\end{aligned}$$

Temos então que

$$|v\rangle = \frac{1}{\sqrt{2}}|v_1\rangle + \frac{3}{\sqrt{2}}|v_2\rangle.$$

9

Exercício 27. Verifique se os conjuntos de vetores abaixo são ortogonais. Verifique quais deles são ortonormais. Verifique quais deles constituem bases de $\mathbb{R}^2$.

2.6. PRODUTO INTERNO

25

1. $\{(1, 2), (1, 0)\}$;
2. $\left\{\left(\frac{2}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(\frac{-1}{\sqrt{3}}, \frac{2}{\sqrt{3}}\right)\right\}$;
3. $\{(1, 2), (2, -1)\}$.

Exercício 28. Considere os conjuntos acima que são bases de $\mathbb{R}^2$. Escreva o vetor $(1, 1)$ como combinação linear dos elementos dessas bases.

Exercício 29. Sejam $|v\rangle = (1, 1)$ e $|u\rangle = (a, -1)$. Para quais valores da constante a os vetores $|v\rangle$ e $|u\rangle$ são ortogonais?

Exercício 30. Mostre que se $|u\rangle$ é ortogonal a $|v\rangle$ então $\lambda |u\rangle$ é ortogonal a $|v\rangle$ para todo valor de $\lambda \in \mathbb{R}$.

Exercício 31. Mostre que se $|v\rangle$ é ortogonal a $|v_1\rangle, \dots, |v_k\rangle$ então $|v\rangle$ é ortogonal a qualquer combinação linear de $|v_1\rangle, \dots, |v_k\rangle$.

Capítulo 3

O espaço $\mathbb{C}^2$

As definições feitas no capítulo anterior para o conjunto $\mathbb{R}^2$ podem ser generalizadas para o conjunto $\mathbb{C}^2$. Muitas definições e resultados são análogos ao caso real e por isso não repetiremos todas as demonstrações aqui. O leitor interessado pode também consultar as referências [4, 7].

3.1 Definição e Propriedades

Definição 19. Representamos por $\mathbb{C}^2$ o conjunto de todas as duplas de *números complexos*, que denotaremos por

$$|u\rangle = (z, w), \quad z, w \in \mathbb{C}.$$

Os elementos de $\mathbb{C}^2$ também serão chamados *vetores*. Diferentemente de $\mathbb{R}^2$, o conjunto $\mathbb{C}^2$ não pode ser representado geometricamente, pois pra isso precisaríamos de 4 eixos reais.

Definição 20. Dados $|u\rangle = (z_1, w_1)$ e $|v\rangle = (z_2, w_2)$ em $\mathbb{C}^2$, podemos a *soma*

$$|u\rangle + |v\rangle = (z_1 + z_2, w_1 + w_2) \quad (3.1)$$

e o *produto* de $|u\rangle$ por um escalar $\lambda \in \mathbb{C}^2$

$$\lambda |u\rangle = (\lambda z_1, \lambda w_1) \quad (3.2)$$

Teorema 8. Para todos $|u\rangle, |v\rangle, |w\rangle \in \mathbb{C}^2$ e $\lambda, \nu \in \mathbb{C}$ as operações de soma e produto por escalar satisfazem as propriedades:

1. (*Associatividade*) $|u\rangle + (|v\rangle + |w\rangle) = (|u\rangle + |v\rangle) + |w\rangle$;
2. (*Comutatividade*) $|u\rangle + |v\rangle = |v\rangle + |u\rangle$;
3. (*Existência de zero*) O elemento $\mathbf{0} = (0, 0) \in \mathbb{C}^2$ é tal que $|u\rangle + \mathbf{0} = |u\rangle$;

4. (Existência de inverso aditivo) Dado $|u\rangle = (x, y) \in \mathbb{C}^2$, o vetor $-|u\rangle = (-x, -y) \in \mathbb{C}^2$ é tal que $|u\rangle + (-|u\rangle) = \mathbf{0}$;
5. (Associatividade) $\lambda(\nu|u\rangle) = (\lambda\nu)|u\rangle$;
6. (Distributividade) $\lambda(|u\rangle + |v\rangle) = \lambda|u\rangle + \lambda|v\rangle$;
7. (Distributividade) $(\lambda + \nu)|u\rangle = \lambda|u\rangle + \nu|u\rangle$;
8. $1|u\rangle = |u\rangle$.

Exercício 32. Mostre que as operações de soma e produto por escalar definidas em $\mathbb{C}^2$ de fato satisfazem todas as propriedades acima.

3.2 Subespaços Vetoriais

Definição 21. Um *subespaço vetorial* S de $\mathbb{C}^2$ é um subconjunto não vazio de $\mathbb{C}^2$ que para o qual as seguintes propriedades são satisfeitas:

1. $|x\rangle + |y\rangle \in S$ para todo par $|x\rangle$ e $|y\rangle \in S$;
2. $\lambda|x\rangle \in S$ para todo $\lambda \in \mathbb{C}$ e todo $|x\rangle \in S$.

Quando S é um subespaço, ele é preservado pelas operações de soma e multiplicação por escalar, isto é, essas operações, ao serem aplicadas a elementos de S , produzem como resultado vetores que também são elementos de S .

Exemplo 19. O subconjunto

$$S = \{(t, 0) \in \mathbb{C}^2; t \in \mathbb{C}\}$$

é um subespaço vetorial de $\mathbb{C}^2$.

Solução. A solução de um problema como esse passa sempre por três passos. O primeiro deles é verificar qual é a propriedade que define o subconjunto S . Nesse exemplo, o subconjunto S é o conjunto de todos os vetores de $\mathbb{C}^2$ que possuem a segunda coordenada igual a zero.

O segundo passo é verificar se a soma de quaisquer elementos de S permanece em S . Suponhamos então que $|u\rangle$ e $|v\rangle \in S$. Isso quer dizer que $|u\rangle = (t_1, 0)$ e $|v\rangle = (t_2, 0)$ com t_1 e $t_2 \in \mathbb{C}$. Assim

$$|u\rangle + |v\rangle = (t_1 + t_2, 0)$$

também pertence a S .

O terceiro passo é verificar se a multiplicação de elementos de S por constante permanece em S . Tomando $|v\rangle = (t_1, 0) \in S$ qualquer e $\lambda \in \mathbb{C}$ temos

$$\lambda|v\rangle = (\lambda t_1, 0)$$

3.3. COMBINAÇÕES LINEARES

29

que também pertence a S . Como as propriedades 1 e 2 são satisfeitas, S é um subespaço de $\mathbb{C}^2$.

□

Exemplo 20. Considere o subconjunto

$$S = \{(i, t) \in \mathbb{C}^2; t \in \mathbb{C}\}.$$

O subconjunto S não é um subespaço vetorial de $\mathbb{C}^2$. De fato, temos que $(i, 1)$ pertence a S , mas $2 \times (i, 1) = (2i, 2)$ não pertence a S .

Exercício 33. Verifique quais são, dentre os subconjuntos de $\mathbb{C}^2$ abaixo, subespaços vetoriais.

1. O subconjunto de vetores que possuem primeira coordenada igual a zero;
2. O subconjunto de vetores que possuem primeira coordenada igual a $1 + i$;
3. O subconjunto de vetores (z, w) tais que $zw = 0$;
4. O subconjunto de vetores (z, w) tais que $iz + 3w = 0$;
5. O subconjunto de vetores (z, w) tais que $z - 3iw = 1$.

3.3 Combinações Lineares

Em $\mathbb{C}^2$, os elementos podem ser combinados livremente através da soma e da multiplicação por escalar.

Definição 22. Dizemos que $|v\rangle$ é uma *combinação linear* de $|u_1\rangle, \dots, |u_k\rangle$ se existem constantes $a_1, \dots, a_k \in \mathbb{C}$ tais que

$$|v\rangle = a_1 |u_1\rangle + \dots + a_k |u_k\rangle. \quad (3.3)$$

Exemplo 21. O vetor $(2 + i, i) \in \mathbb{C}^2$ é uma combinação linear dos vetores $(2, i)$ e $(1, 0)$ porque

$$(2 + i, i) = (2, i) + i(1, 0).$$

Exercício 34. Mostre que o vetor nulo $\mathbf{0}$ é sempre combinação linear de qualquer conjunto de vetores $|v_1\rangle, \dots, |v_k\rangle$.

Exemplo 22. Considere os vetores $|e_1\rangle = (1, 0)$, $|e_2\rangle = (0, 1)$. Qualquer vetor em $\mathbb{C}^2$ é combinação linear desses dois vetores. De fato, dado $|v\rangle = (a_1, a_2)$ podemos escrever

$$\begin{aligned} (a_1, a_2) &= (a_1, 0) + (0, a_2) \\ &= a_1 (1, 0) + a_2 (0, 1) \\ &= a_1 |e_1\rangle + a_2 |e_2\rangle. \end{aligned}$$

De maneira geral, dados vetores $|u_1\rangle, \dots, |u_m\rangle \in \mathbb{C}^2$, um vetor $|v\rangle$ será combinação linear de $|u_1\rangle, \dots, |u_m\rangle$ se existirem constantes complexas $a_1, \dots, a_m$ tais que

$$|v\rangle = a_1|u_1\rangle + \dots + a_m|u_m\rangle.$$

Para descobrir se essas constantes de fato existem e, se existem, quanto elas valem, devemos resolver um sistema de equações que envolvem essas constantes. Nesse curso vamos trabalhar com poucos vetores de maneira que os sistemas que teremos que resolver serão simples e em geral podemos resolvê-los intuitivamente.

Exercício 35. Verifique se o vetor $|v\rangle$ é combinação linear dos vetores $|u_1\rangle, \dots, |u_n\rangle$ dados abaixo. Em caso afirmativo, encontre a combinação linear desses vetores que gera $|v\rangle$.

1. $|v\rangle = (i, 1+i); |u_1\rangle = (1, 0);$
2. $|v\rangle = (2, 2); |u_1\rangle = (i, i);$
3. $|v\rangle = (2, 3+i); |u_1\rangle = (1, 0), |u_2\rangle = (0, 1);$
4. $|v\rangle = (2, 3+i); |u_1\rangle = (1, 0), |u_2\rangle = (1, 1).$
5. $|v\rangle = (2, 1); |u_1\rangle = (1, 0), |u_2\rangle = (1, i), |u_3\rangle = (1, -i).$
6. $|v\rangle = (-i, 3-2i); |u_1\rangle = (1, 0), |u_2\rangle = (1, 1), |u_3\rangle = (1, -1).$

3.4 Independência Linear

Queremos encontrar condições que informem se dentre um conjunto de vetores, algum deles pode ser escrito como combinação linear dos outros. Para isso, estendemos a definição de dependência linear também para $\mathbb{C}^2$.

Definição 23. Dizemos que um conjunto de vetores $\{|u_1\rangle, \dots, |u_k\rangle\} \subset \mathbb{C}^2$ é *linearmente independente* (LI) se a equação

$$a_1 |u_1\rangle + \cdots + a_k |u_k\rangle = \mathbf{0}$$

só admite a solução trivial $a_1 = \dots = a_k = 0$, ou seja, se a única combinação linear dos vetores que gera o vetor nulo é aquela em que todas as constantes são nulas.

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

Um conjunto de vetores é LD se existe uma combinação linear deles que gera o vetor nulo em que pelo menos uma das constantes é não-nula.

Três ou mais vetores em $\mathbb{C}^2$ são sempre LD. Isso acontece porque o problema de identificar se esses vetores são ou não LI leva a um sistema que sempre possui solução não trivial.

Teorema 9. *O número máximo de vetores LI em $\mathbb{C}^2$ é igual a 2.*

A definição de dependência linear é importante porque ela permite identificar se dentro de um conjunto de vetores, um deles pode ser escrito como combinação linear dos outros ou não. De fato, as duas propriedades são equivalentes.

Teorema 10. *Um conjunto de vetores é LD se, e somente se, podemos expressar ao menos um dos vetores como combinação linear dos outros.*

As provas são idênticas ao caso de $\mathbb{R}^2$ e não iremos repeti-las aqui.

Exercício 36. Verifique se os conjuntos de vetores abaixo são LI ou LD.

1. $(1, i)$;
2. $(1, 2i), (3i, 6)$;
3. $(1, 2i), (3, 1)$;
4. $(1, 2i), (3, 1), (1, 1)$;
5. $(1, 1 + i), (i, 0), (4, 6)$.

Exercício 37. Em cada item do exercício anterior em que os vetores são LD, escreva uma combinação linear dos vetores que resulta no vetor nulo em que pelo menos uma das constantes usadas seja diferente de zero. Escreva um dos vetores como combinação linear dos outros.

3.4.1 Subespaços gerados

Definição 24. Dado um e conjunto de vetores $|u_1\rangle, \dots, |u_k\rangle$ em $\mathbb{C}^2$, o *subespaço gerado* por esses vetores é o conjunto de todas as suas combinações lineares:

$$\langle |u_1\rangle, \dots, |u_k\rangle \rangle = \{a_1 |u_1\rangle + \dots + a_k |u_k\rangle; a_i \in \mathbb{C}\}.$$

Exercício 38. Prove que $\langle |u_1\rangle, \dots, |u_k\rangle \rangle$ é de fato um subespaço de $\mathbb{C}^2$.

3.4.2 Base e Dimensão

Podemos nos perguntar se existe um conjunto de vetores LI de forma que todo elemento de $\mathbb{C}^2$ possa ser escrito como combinação linear dos elementos desse conjunto. É possível mostrar que $\mathbb{C}^2$ também possui infinitos conjuntos LI com essa propriedade.

Definição 25. Uma base para $\mathbb{C}^2$ é um conjunto LI

$$\mathcal{B} = \{|u_1\rangle, \dots, |u_k\rangle\}$$

tal que todo vetor de $\mathbb{C}^2$ é combinação linear de $|u_1\rangle, \dots, |u_k\rangle$, ou seja

$$\langle |u_1\rangle, \dots, |u_k\rangle \rangle = V.$$

De maneira intuitiva, uma base para $\mathbb{C}^2$ é um conjunto com o número mínimo de vetores que precisamos para gerar todos os outros vetores a partir de combinações lineares.

Exemplo 23. Os vetores $|u_1\rangle = (1, 0)$ e $|u_2\rangle = (0, 1)$ formam uma base para $\mathbb{C}^2$. De fato, esses vetores são LI e qualquer vetor de $\mathbb{C}^2$ pode ser escrito como combinação linear de $|u_1\rangle$ e $|u_2\rangle$.

O Teorema 9 implica que não pode haver uma base para $\mathbb{C}^2$ com mais de dois elementos, já que tal conjunto seria LD. Por outro lado, um conjunto com apenas um vetor não pode gerar todo o $\mathbb{C}^2$. Isso implica que todas as bases de $\mathbb{C}^2$ possuem *exatamente* dois elementos.

Definição 26. A *dimensão* de $\mathbb{C}^2$ é o número de vetores em uma base de $\mathbb{C}^2$. Logo $\mathbb{C}^2$ tem dimensão 2.

Um resultado muito útil é o seguinte:

Teorema 11. *Em $\mathbb{C}^2$, qualquer conjunto LI com dois elementos é uma base.*

Assim, para verificar se um conjunto de vetores é ou não uma base para $\mathbb{C}^2$, basta verificarmos se ele possui o número correto de vetores e se esses vetores são LI. O teorema acima garante que nessas condições, o conjunto gera $\mathbb{C}^2$ e portanto é uma base. Podemos resumir esse resultado da seguinte forma:

Uma base para $\mathbb{C}^2$ é um conjunto LI com exatamente 2 vetores.

Exercício 39. Verifique se os vetores abaixo formam uma base para $\mathbb{C}^2$.

1. $(1, 1)$;
2. $(i, 2), (1, 0)$;
3. $(1, 2), (i, 0), (0, i)$;
4. $(1, -i), (0, i)$;
5. $(1, i), (0, 1), (0, i)$.

3.5 Produto Interno

A única definição que não é exatamente a mesma para $\mathbb{C}^2$ é a definição de produto interno. Quando utilizamos números complexos, para manter as propriedades importantes do caso real devemos utilizar o complexo conjugado das coordenadas do primeiro vetor.

3.5. PRODUTO INTERNO

33

Definição 27. O *produto interno* entre os vetores $|u\rangle = (z_1, w_1)$ e $|v\rangle = (z_2, w_2)$ é definido por

$$\langle u \mid v \rangle = \bar{z}_1 z_2 + \bar{w}_1 w_2.$$

Teorema 12. *O produto interno em $\mathbb{C}^2$ satisfaz as seguintes propriedades:*

1. $(\lambda \vec{u} + \mu \vec{v}) \cdot \vec{w} = \lambda (\langle u | w \rangle) + \mu (\langle v | w \rangle);$
2. $\langle u | v \rangle = \langle v | u \rangle;$
3. $\langle u | u \rangle \geq 0;$
4. *Se $\langle u | u \rangle = 0$ então $|u\rangle = 0$.*

O produto interno entre $|u\rangle = (z_1, w_1)$ e $|v\rangle = (z_2, w_2)$ também pode ser denotado por

$$\langle u \mid v \rangle = (z_1, w_1) \cdot (z_2, w_2).$$

Exemplo 24. O produto interno entre os vetores $(i, 3)$ e $(4 + i, 2 + i)$ em $\mathbb{C}^2$ é igual a

$$(i, 3) \cdot (4 + i, 2 + i) = (-i)(4) + (4 - i)(2 + i) = -4i + 9 + 2i = 9 - 2i.$$

Definição 28. A partir do produto interno podemos definir a aplicação

$$\|\cdot\|: \mathbb{C}^2 \rightarrow \mathbb{C}$$

escrevendo

$$\|v\| = \sqrt{\langle v | v \rangle}.$$

Essa função é chamada *norma* em $\mathbb{C}^2$.

Definição 29. Dizemos que dois vetores $|u\rangle$ e $|v\rangle$ são *ortogonais* se $\langle u | v \rangle = 0$. Dizemos que um conjunto $E = \{|v_1\rangle, \dots, |v_k\rangle\}$ é *ortogonal* se seus elementos são dois a dois ortogonais. Dizemos que um conjunto $E = \{|v_1\rangle, \dots, |v_k\rangle\}$ é *ortonormal* se é ortogonal e $\langle v_i | v_i \rangle = 1$ para todo i .

Exercício 40. Mostre que o número máximo de vetores ortogonais em $\mathbb{C}^2$ é 2.

Teorema 13. Se $\{|v_1\rangle, \dots, |v_k\rangle\}$ é um conjunto ortogonal, então valem as seguintes propriedades:

1. O conjunto $\{|v_1\rangle, \dots, |v_k\rangle\}$ é LI;
2. Se $V = a_1 |v_1\rangle + \dots + a_k |v_k\rangle$, então

$$a_i = \frac{\langle v | v_i \rangle}{|| |v_i\rangle ||^2}.$$

A demonstração desse resultado é idêntica a feita para $\mathbb{R}^2$ e portanto não será repetida.

Definição 30. Dizemos que uma base $B = \{|v_1\rangle, |v_2\rangle\}$ para $\mathbb{C}^2$ é uma *base ortogonal* se seus elementos são ortogonais. Dizemos uma base $B = \{|v_1\rangle, |v_2\rangle\}$ é uma *base ortonormal* se é ortogonal e $\langle v_i | v_i \rangle = 1$ para todo i .

Exemplo 25. A base canônica $B = \{|e_1\rangle = (1, 0), |e_2\rangle = (0, 1)\}$ de $\mathbb{C}^n$ é uma base ortonormal. De fato, é fácil mostrar que o produto interno $\langle e_i | e_j \rangle$ é igual a zero se $i \neq j$ e igual a 1 se $i = j$.

Exemplo 26. Mostre que o conjunto

$$B = \{|v_1\rangle = \left(\frac{i}{\sqrt{2}}, \frac{i}{\sqrt{2}}\right), |v_2\rangle = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)\}$$

é uma base ortonormal de $\mathbb{C}^2$.

Solução. Calculando os produtos internos temos

$$\langle v_1 | v_2 \rangle = 0, \quad \langle v_1 | v_1 \rangle = \langle v_2 | v_2 \rangle = 1$$

o que mostra que B é um conjunto ortonormal. Pelo teorema 31, B é um conjunto LI e portanto é uma base, já que possui dois vetores. $\square$

Exercício 41. Verifique se os conjuntos de vetores abaixo são ortogonais. Verifique quais deles são ortonormais. Verifique quais deles constituem bases de $\mathbb{C}^2$.

1. $\{(i, 2i), (1, 0)\};$
2. $\left\{\left(\frac{2i}{\sqrt{3}}, \frac{i}{\sqrt{3}}\right), \left(\frac{-i}{\sqrt{3}}, \frac{2i}{\sqrt{3}}\right)\right\};$
3. $\{(1, 2), (2i, -i)\}.$

Exercício 42. Considere os conjuntos acima que são bases de $\mathbb{C}^2$. Escreva o vetor (i, i) como combinação linear dos elementos dessas bases.

afirmações nas entrelinhas. Vamos detalhá-las.

4.1 Estados e Medições

Em nossa descrição de um *bit quântico*, o primeiro passo será encontrar um objeto matemático que represente adequadamente o *estado do sistema*. Esse objeto nos dará todas informações que podemos obter a respeito do comportamento desse sistema. O conjunto ao qual pertencem tais objetos é o que chamamos de *espaço de estados* do sistema. Assim, todo sistema quântico possui um *espaço de estados* E .

Para o bit quântico, temos que $E = \mathbb{C}^2$. O estado de um bit quântico é definido por um vetor unitário em seu espaço de estados. Toda e qualquer predição sobre o sistema pode ser feita a partir do conhecimento de seu estado. Para uso nesse texto, vamos adotar:

Postulado 1. *O estado de um bit quântico é um vetor unitário em seu espaço de estados $E = \mathbb{C}^2$.*

Uma base para o espaço de estados será dada por dois vetores linearmente independentes, $\{|e_1\rangle, |e_2\rangle\}$. Como as alternativas clássicas de um bit costumam ser denotadas 0 e 1, é comum utilizarmos a base $\{|0\rangle, |1\rangle\}$. O leitor deve ter muito cuidado para não confundir $|0\rangle$ com o vetor nulo do espaço $\mathbb{C}^2$. Claramente este não é o caso, pois $|0\rangle$ e $|1\rangle$ são linearmente independentes. O vetor nulo sempre será denotado pelo símbolo $\mathbf{0}$ para evitar que ele seja confundido com o vetor $|0\rangle$, que é um vetor não-nulo. Em geral, vamos utilizar a base formada pelos vetores $|0\rangle = (1, 0)$ e $|1\rangle = (0, 1)$. Assim, um estado para o bit quântico pode ser escrito na forma

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle \quad (4.1)$$

em que $|\alpha|^2 + |\beta|^2 = 1$ para que $|\psi\rangle$ seja um vetor unitário.

O segundo objeto matemático que precisamos em nossa descrição está relacionado às medições que podemos realizar sobre o sistema. O fato de que bits quânticos podem estar em superposições da forma (4.1) faz com que os testes que podemos fazer sobre eles tenham também uma estrutura muito mais rica. Em bits clássicos só podemos fazer um teste: o teste que discrimina entre as possibilidades 0 e 1. Em bits quânticos podemos fazer uma infinidade de testes.

Para um qbit, as medições que consideraremos nesse texto possuem sempre duas respostas possíveis. Tais respostas correspondem a alternativas distintas, que vamos associar a uma base *ortonormal*. Chegamos assim à importante noção de *teste* para qbits:

Postulado 2. *Um teste com alternativas a e b é associado a uma base ortonormal, denotada $\{|a\rangle, |b\rangle\}$. Aplicar um teste pode ser visto como decompor o vetor com relação a esta base, para em seguida selecionar apenas uma das alternativas.*

Definida uma base¹, todo vetor do espaço de estados pode ser escrito como combinação linear destes elementos. Para um qbit, então, seu estado será descrito por

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \quad (4.2)$$

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

Exercício 43. Lembrando que $\| |\psi\rangle \|^2 = \langle \psi | \psi \rangle$, obtenha a condição de normalização $|\alpha|^2 + |\beta|^2 = 1$ apresentada acima. Discuta a importância da base ser ortonormal nesse exercício.

De maneira geral, dada uma base ortonormal $\{|a\rangle, |b\rangle\}$, podemos escrever

$$|\psi\rangle = \delta |a\rangle + \gamma |b\rangle, \quad (4.3)$$

e para que $|\psi\rangle$ tenha norma 1, exigimos que $|\delta|^2 + |\gamma|^2 = 1$.

Uma das grandes novidades da mecânica quântica aparece na sua regra sobre como relacionar o estado $|\psi\rangle$ à medição das alternativas.

Postulado 3. Se um sistema quântico no estado $|\psi\rangle$ da eq. (4.2) é submetido a um teste com alternativas 0 e 1, a probabilidade de obter o resultado correspondente a 0 é dada por $|\alpha|^2$, enquanto a de obter 1 é dada por $|\beta|^2$.

Os coeficientes α e β da expansão do estado $|\psi\rangle$ com respeito à base $\{|0\rangle, |1\rangle\}$ são números complexos que permitem calcular probabilidades. Feynman batizou tais coeficientes *amplitudes de probabilidades*, ou simplesmente *amplitudes*.

Mesmo sem querer desviar para discussões sobre fundamentos de mecânica quântica, é necessário dizer que esta foi a primeira vez que uma teoria científica se assumiu probabilística *a priori*. Mesmo que conheçamos o estado $|\psi\rangle$ de uma partícula, o resultado de observações será, em geral, probabilístico. Em geral, o conceito de probabilidades é introduzido com a justificativa que, *na prática*, não podemos dar uma descrição precisa para um sistema macroscópico. De certa forma, é uma concessão que mentes determinísticas fizeram à dificuldade de trabalhar com 10^{23} coordenadas, ou mais. Mas mantinha-se a convicção que *em princípio* poderia se descrever microscopicamente um gás, por exemplo. Na mecânica quântica não; exceto se α ou β for zero, a mais completa descrição microscópica é incapaz de prever, senão probabilisticamente, o resultado do teste 0 ou 1.

Esta descrição probabilística da mecânica quântica tem uma consequência fundamental: embora gostemos muito de tratar de um sistema quântico específico, as previsões desta teoria só podem ser testadas quando preparamos igualmente um grande número de cópias do sistema, e agimos igualmente sobre todas elas e assim poderemos comparar as frequências obtidas com as probabilidades previstas.

De maneira geral, ao realizar um teste qualquer sobre um qbit, temos

¹Salvo menção contrária, daqui por diante todas as bases consideradas serão bases ortonormais.

Postulado 4. Se um sistema quântico no estado $|\psi\rangle$ da eq. (4.3) é submetido a um teste com alternativas a e b , a probabilidade de obter o resultado correspondente a a é dada por $|\delta|^2$, enquanto a de obter b é dada por $|\gamma|^2$.

Exercício 44. Considere os vetores

$$\begin{aligned} |+\rangle &= \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle); \\ |-\rangle &= \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle). \end{aligned} \quad (4.4)$$

1. Mostre que $\{|+\rangle, |-\rangle\}$ é uma base ortonormal para $\mathbb{C}^2$.
2. Considere o teste que discrimina entre as alternativas $+$ e $-$ associado à base $\{|+\rangle, |-\rangle\}$. Calcule a probabilidade de obter cada uma das alternativas ao aplicar esse teste a um sistema no estado

$$|\psi\rangle = \frac{\sqrt{2}}{\sqrt{3}} |+\rangle + \frac{i}{\sqrt{3}} |-\rangle.$$

3. Calcule a probabilidade de obter cada uma das alternativas ao aplicar esse teste a um sistema no estado $|0\rangle = (1, 0)$.
4. Calcule a probabilidade de obter cada uma das alternativas ao aplicar esse teste a um sistema no estado

$$|\phi\rangle = \frac{\sqrt{2}}{\sqrt{3}} |0\rangle + \frac{i}{\sqrt{3}} |1\rangle.$$

4.2 Depois das Medições

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

Postulado 5. Após a realização de um teste para discriminar entre as alternativas clássicas 0 e 1, se o resultado obtido foi 0, o sistema passa a ser descrito pelo estado $|0\rangle$; se o resultado obtido foi 1, o sistema passa a ser descrito pelo estado $|1\rangle$.

Este postulado está naturalmente associado à noção de *repetibilidade de testes*. Ou seja, se um teste é realizado e se obtém um resultado, repetições deste mesmo teste no mesmo sistema corroborarão o resultado obtido. É importante distinguir aqui entre “agir novamente no mesmo sistema” e “realizar o teste em outro elemento diferente, mas que esteja no mesmo estado”. Agir novamente no mesmo sistema é repetir o mesmo teste duas vezes, *no mesmo sistema físico*.

livremente outra base para E , gerando testes diferentes do teste que distingue entre as alternativas 0 e 1. A exigência de serem alternativas distinguíveis impõe ortonormalidade.

Como um exemplo, podemos considerar os vetores:

$$\begin{aligned} |+\rangle &= \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle); \\ |-\rangle &= \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle). \end{aligned} \quad (4.5)$$

Exercício 47. Mostre que $\{|+\rangle, |-\rangle\}$ é uma base ortonormal.

Podemos aplicar o teste $+$ ou $-$, que corresponde a esta base. Devemos aplicar a este teste as mesmas regras que antes usávamos para 0 e 1, com sua correspondente base. Chamemos o teste 0 ou 1 de Z e o teste $+$ ou $-$ de X , devido a uma convenção que pode ser justificada².

Exercício 48. Relação entre os testes X e Z .

1. Considere o estado inicial $|0\rangle$. Quais as probabilidades de cada alternativa para o teste Z ? E para o teste X ?
2. Suponha que foi realizado o teste X e obtido o resultado $+$. Qual a probabilidade de obter 0 em uma realização subsequente do teste Z ?

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

Vamos discutir essa situação em mais detalhe. Feito o primeiro teste Z , se foi obtido o resultado 0, sabemos que devemos passar a descrever o sistema pelo estado $|0\rangle$. Neste estado, o teste X terá o resultado $+$ ou $-$ de maneira equiprovável. Com isso, a melhor descrição do sistema será dada por $|+\rangle$ no primeiro caso e $|-\rangle$ no segundo. Em ambas as alternativas, o novo teste Z também terá os resultados 0 ou 1 de maneira equiprovável.

Exercício 49. Considere um sistema no estado inicial

$$|\psi\rangle = \frac{\sqrt{2}}{\sqrt{3}} |0\rangle + \frac{i}{\sqrt{3}} |1\rangle.$$

1. Calcule as probabilidades de obter cada resposta ao aplicar o teste Z a um sistema no estado acima.
2. Calcule as probabilidades de obter cada resposta ao aplicar o teste X a um sistema no estado acima.

²Veja, por exemplo, a sec. X.Y do livro [?].

Exercício 50. Bases mutuamente neutras.

1. Descreva um teste com resultados a e b , onde o estado $|0\rangle$ dá probabilidades p e $1 - p$.
2. Seja $|a\rangle$ o estado correspondente à alternativa a do teste anterior. Qual a probabilidade de obter 0 se um teste Z for aplicado a este estado?
3. Duas bases $\mathcal{B} = \{|b_0\rangle, |b_1\rangle\}$ e $\mathcal{C} = \{|c_0\rangle, |c_1\rangle\}$ são ditas mutuamente neutras se $|\langle b_i | c_j \rangle|$ é independente de i e j . Mostre que as bases $\mathcal{Z} = \{|0\rangle, |1\rangle\}$ e $\mathcal{X} = \{|+\rangle, |-\rangle\}$ são bases ortonormais mutuamente neutras.
4. É possível obter mais uma base, $\mathcal{Y}$, mutuamente neutra tanto com $\mathcal{Z}$ quanto com $\mathcal{X}$?

4.4 Quando perder é ganhar

Algumas tarefas muito simples do ponto de vista abstrato podem ser muito difíceis na prática³. Por exemplo, gerar números aleatórios. Um pensamento inocente diz que lançar uma moeda para cada bit (cara ou coroa) seria o suficiente. Mas não! Como garantir que a moeda é realmente honesta? Ou ainda, que seu lançamento é honesto?

Novamente atingimos o paradigma teórico onde aleatoriedade não surge a priori, mas da dificuldade de definir as condições iniciais com precisão, e de uma dinâmica muito sensível a tais condições. Os geradores de números “aleatórios” mais utilizados são sofisticações deste lançamento da moeda. Computadores calculam funções determinísticas mas extremamente sensíveis às condições iniciais, e estas condições iniciais envolvem dados razoavelmente aleatórios, como os últimos dígitos do relógio interno do computador, ou bits escolhidos dentro de um arquivo do qual nada se sabe... O que se obtém daí são números “suficientemente aleatórios” para a imensa maioria das aplicações: jogos de computador, simulações de Monte Carlo, geração de números primos muito grandes...

Mas a noção de “suficientemente aleatórios” é sutil. O que é suficientemente aleatório para quem só quer gerar números primos para criar uma chave RSA [?, ?] e usar na sua correspondência eletrônica privada pode não ser suficientemente aleatório para um banco que opera pela internet. O que é suficientemente aleatório para quem só quer se divertir com um jogo pode não ser suficientemente aleatório para uma empresa de jogos de azar *on line*! Pode parecer estranho, mas uma interessante aplicação da mecânica quântica⁴ é aproveitar a existência de testes incompatíveis para produzir números “suficientemente aleatórios”.

³E vice-versa.

⁴Já com algum sucesso comercial[10].

Exercício 51. Usando o que você já aprendeu até o presente momento, proponha uma máquina quântica de gerar bits aleatórios⁵.

De fato, já há trabalhos na linha de considerar aleatoriedade como um valioso *recurso*.

⁵Mais precisamente, gerar bits independentes, identicamente distribuídos, com distribuição equi-provável.

Capítulo 5

Um pouquinho de Física

5.1 Ondas eletromagnéticas

Em 1864, o físico escocês James Clerk Maxwell unificou todos os fenômenos elétricos e magnéticos observáveis, mostrando que eles podem ser descritos através de quatro equações, chamadas *equações de Maxwell*, as equações básicas do Eletromagnetismo.

A partir das equações de Maxwell é possível prever que um campo magnético $\vec{B}$ será produzido em um ponto do espaço se nas proximidades desse ponto houver um campo elétrico variável $\vec{E}$. Um campo magnético variável, por sua vez, pode produzir um campo elétrico variável.

Em algumas situações, esses campos variáveis interagem de tal forma a gerar uma onda de campos elétricos e magnéticos que se propaga até mesmo no vácuo (veja a figura 5.1). Essas ondas foram chamadas *ondas eletromagnéticas*. São tipos de ondas eletromagnéticas vários tipos de radiação, algumas das quais conhecemos muito bem: ondas de rádio, microondas, radiação infra-vermelha, a luz visível, radiação ultravioleta, raios X e raios gama. Vários fenômenos exibidos pela radiação eletromagnética só podem ser explicados através de seu caráter ondulatório: reflexão, refração, difração, interferência (figura 5.2), entre vários outros.

5.1.1 A dualidade onda-partícula

Apesar da formulação ondulatória da radiação eletromagnética descrever adequadamente vários fenômenos físicos como a interferência, ela não é capaz de explicar alguns fenômenos como o *efeito fotoelétrico*. O efeito fotoelétrico é a emissão de elétrons por um material, geralmente metálico, quando exposto à radiação eletromagnética de frequência suficientemente alta. Ele pode ser observado quando a luz incide numa placa de metal, literalmente arrancando elétrons da placa.

A explicação correta para esse fenômeno foi dada em 1905 por Einstein e lhe rendeu o prêmio Nobel de Física em 1921. Einstein admitiu que a luz e as demais radiações eletromagnéticas consistem de um feixe de partículas, denominadas *fótons*, cada uma carregando uma quantidade de energia $E = h \times f$, em que h é uma

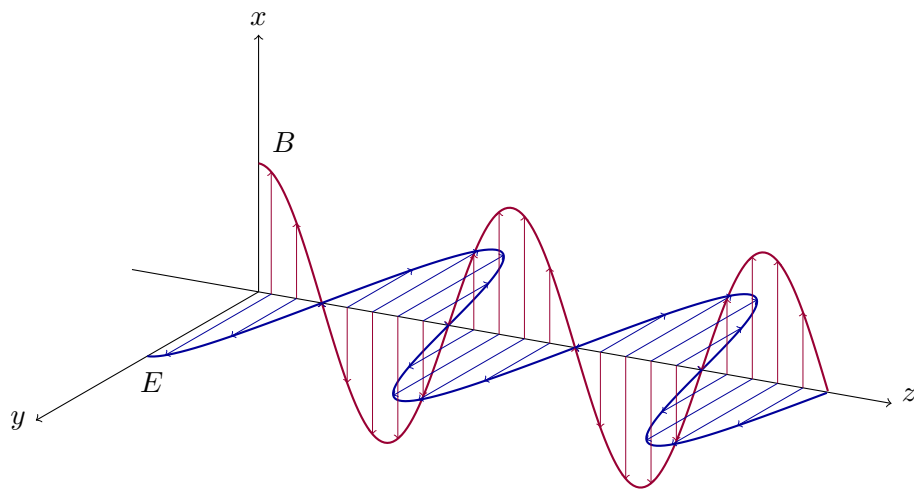


Figura 5.1: Uma onda eletromagnética composta por campos Elétrico e Magnético variáveis que se propagam pelo espaço. Os campos Elétrico E e Magnético B são perpendiculares entre si e ambos são perpendiculares à direção de propagação da onda, que nesse caso é o eixo z .

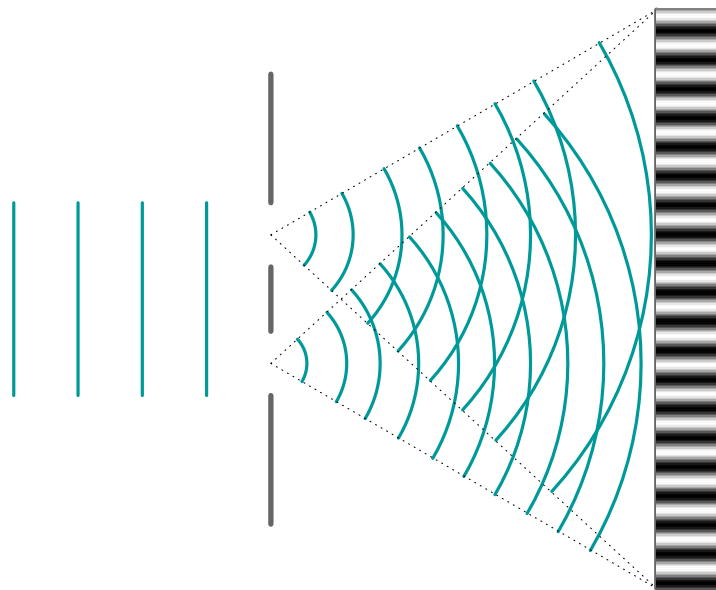


Figura 5.2: Fenômeno de interferência: superposição de duas ou mais ondas num mesmo ponto que pode ter um caráter de aniquilação (pontos escuros), quando as fases não são as mesmas (interferência destrutiva) ou pode ter um caráter de reforço (pontos claros) quando as fases combinam (interferência construtiva).

5.1. ONDAS ELETROMAGNÉTICAS

45

constante chamada *constante de Plank* e f é a frequência da radiação eletromagnética. Mas afinal de contas, o que é a luz? Uma onda ou uma partícula?

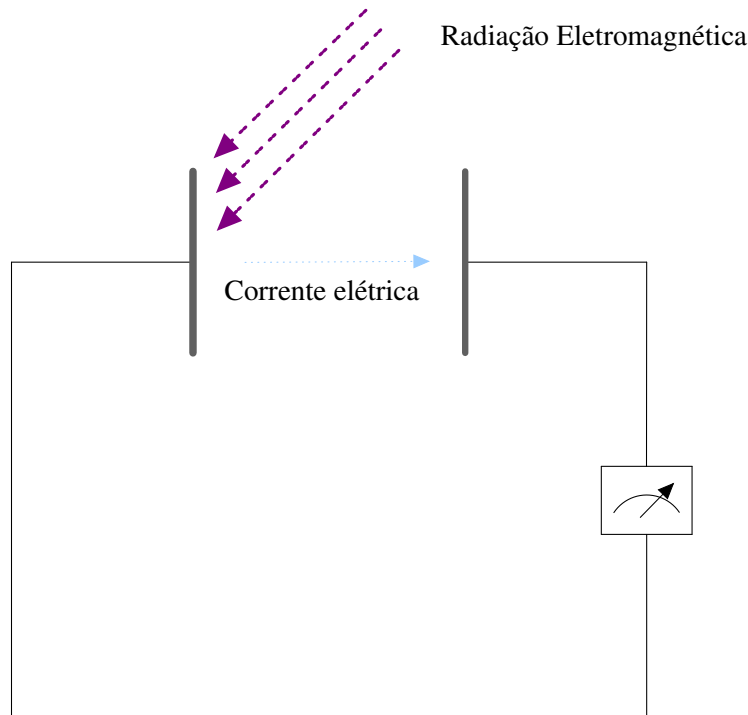


Figura 5.3: Efeito fotoelétrico: elétrons absorvem fótons da luz incidente e são ejetados do material, gerando uma corrente elétrica que pode ser experimentalmente registrada.

Se por um lado, somente o modelo de fótons explica adequadamente o efeito fotoelétrico, por outro, somente o modelo ondulatório explica fenômenos a difração e a interferência. A luz tem uma natureza dual, se comportando ora como uma onda, ora como um feixe de fótons, dependendo da maneira como interagimos com ela. Nenhum dos modelos deve ser utilizado isoladamente para descrever as propriedades da luz. Uma descrição completa do seu comportamento só é válida quando ambos os modelos são combinados de maneira complementar.¹

5.1.2 Polarização da Luz

Uma onda eletromagnética é formada por campos elétricos e magnéticos que variam no tempo e no espaço, perpendicularmente um ao outro, como vimos na

¹Na verdade, não só a luz se comporta ora como onda, ora como partícula. A matéria também exibe esse comportamento! Partículas muito pequenas, como átomos, moléculas, elétrons, prótons exibem comportamentos que só podem ser explicados através de um modelo ondulatório. Incrível, não é mesmo?

figura 5.1. A *direção de polarização* de uma onda eletromagnética é definida como a direção do campo elétrico dessa mesma onda. Assim, a onda mostrada na figura 5.1 é linearmente polarizada ao longo da direção y .

Como mencionamos anteriormente, a luz possui também características corpusculares que podem ser adequadamente descritas através dos fótons, as partículas elementares mediadoras da força eletromagnética. A polarização da luz pode ser vista como uma característica intrínseca dos fótons.

Existem vários dispositivos que permitem a passagem da luz polarizada em apenas uma direção. Chamaremos esses dispositivos de *polarizadores*. Um tipo comum de polarizador consiste em uma placa feita com um material que só deixa passar as componentes de campo elétrico da luz que estão em uma determinada direção. Um desses materiais, o polaróide, é constituído de longas cadeias de moléculas orientadas em uma direção. Essas cadeias são boas condutoras elétricas e absorvem luz incidente cujo campo elétrico é paralelo a elas e transmite luz cujo campo elétrico é perpendicular.

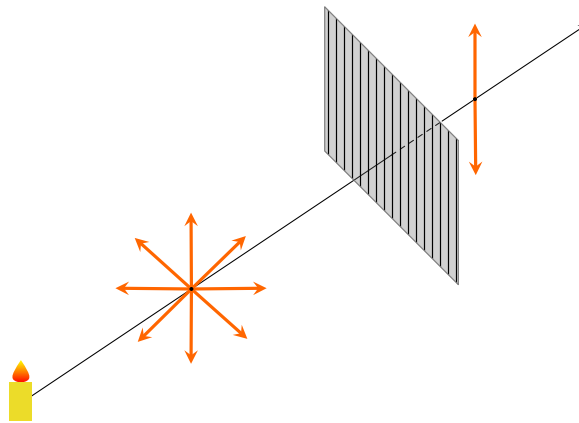


Figura 5.4: Um polarizador permite somente a passagem da luz polarizada em uma direção.

Utilizando fótons polarizados em duas direções perpendiculares, x e y , por exemplo, podemos criar uma representação física para um bit. Atribuíremos o valor 0 à situação em que o fóton está polarizado na direção x e o valor 1 à situação em que o fóton está polarizado na direção y . Utilizando um polarizador orientado na direção y podemos testar qual é o valor do bit: se o valor for 1 o fóton será transmitido e se o valor for 0 o fóton será absorvido. Quando o fóton estiver polarizado na direção x , seu estado será representado pelo símbolo $|H\rangle$. Quando ele estiver polarizado na direção y , utilizaremos o símbolo $|V\rangle$.

Apesar de podermos utilizar a polarização de um fóton para realizar um bit, a descrição completa da polarização é bem mais sofisticada e só pode ser feita adequadamente se utilizarmos os elementos de Física Quântica vistos no capítulo 4.

Acontece que os fótons podem estar polarizados em outras direções além das

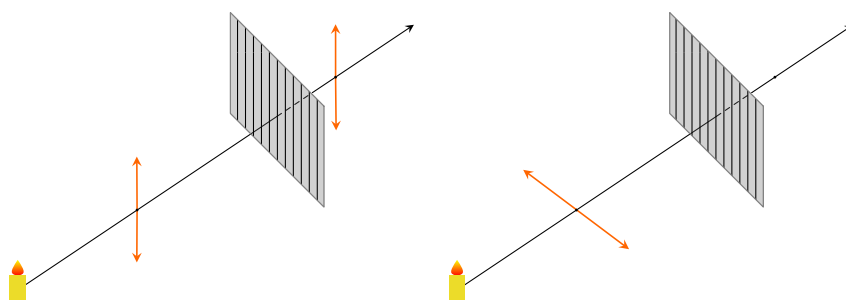


Figura 5.5: Um bit criado a partir da polarização de um fóton: se o fóton estiver polarizado na direção y , ele passará pelo polarizador. Se ele estiver polarizado na direção x ele não passará pelo polarizador.

direções x e y . Matematicamente, um fóton polarizado em outras direções pode ser representado por uma combinação dos símbolos $|H\rangle$ e $|V\rangle$:

$$|\phi\rangle = \alpha |H\rangle + \beta |V\rangle. \quad (5.1)$$

Assim, os estados de polarização podem ser adequadamente descritos se utilizarmos $\mathbb{C}^2$ como o espaço de estados associado. A polarização de um fóton é um exemplo de um bit quântico! Por convenção, associamos os estados de polarização $|H\rangle$ e $|V\rangle$ à base ortonormal

$$|H\rangle = (1, 0), \quad |V\rangle = (0, 1). \quad (5.2)$$

Os números α e β que aparecem na Equação (5.1) são números complexos que nos permitem descrever o comportamento da luz polarizada. Se $\alpha = 1$ e $\beta = 0$, o fóton está polarizado na direção x e não será transmitido ao passar por um polarizador orientado na direção y . Se $\alpha = 0$ e $\beta = 1$, o fóton está polarizado na direção y e será transmitido ao passar por esse mesmo polarizador. No entanto, se α e β forem não nulos, o fóton poderá ser transmitido ou não, sendo impossível prever de antemão qual das duas possibilidades irá ocorrer. Podemos apenas prever as probabilidades de cada alternativa: ao passar por um polarizador orientado na direção y , a probabilidade do fóton no estado de polarização $|\phi\rangle$ ser absorvido é igual a $|\alpha|^2$ e a probabilidade do fóton ser transmitido é $|\beta|^2$. A probabilidade dele ser absorvido mais a probabilidade dele ser transmitido devem somar 1 e portanto

$$|\alpha|^2 + |\beta|^2 = 1.$$

Se α e β forem números reais, dizemos que a luz está *linearmente polarizada*, como mostrado na figura 5.6. Nesse caso, a direção do campo elétrico é constante no tempo e é determinada pelos valores de α e β .

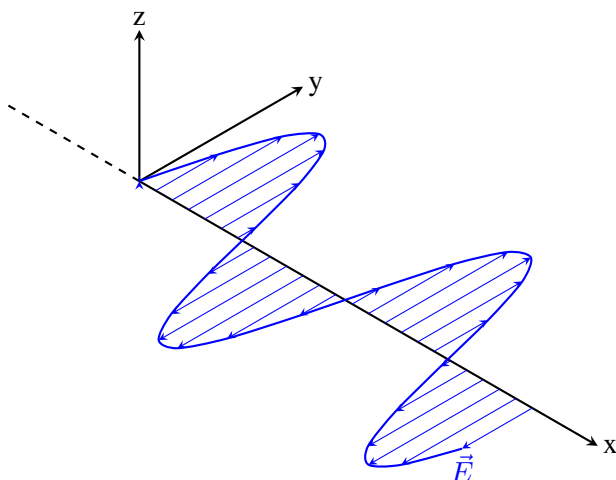


Figura 5.6: Luz linearmente polarizada.

A figura mostra a direção do campo elétrico para um fóton polarizado linearmente descrito pelo estado

$$|+\rangle = \frac{1}{\sqrt{2}} |H\rangle + \frac{1}{\sqrt{2}} |V\rangle.$$

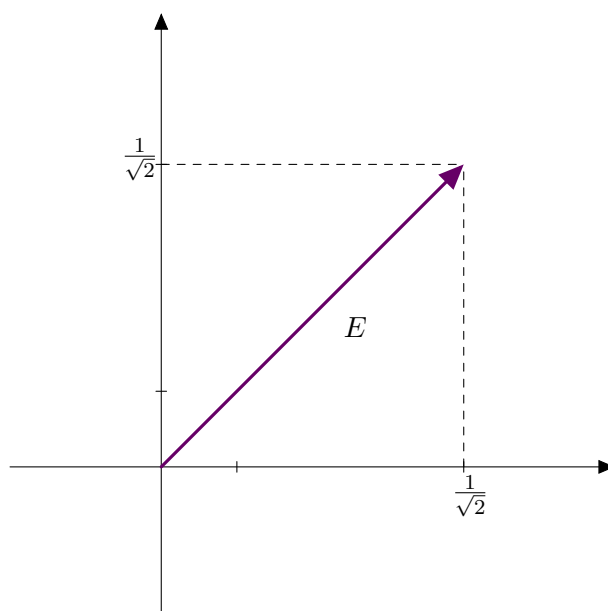


Figura 5.7: Direção do campo elétrico para um fóton cuja polarização é a superposição $|+\rangle$.

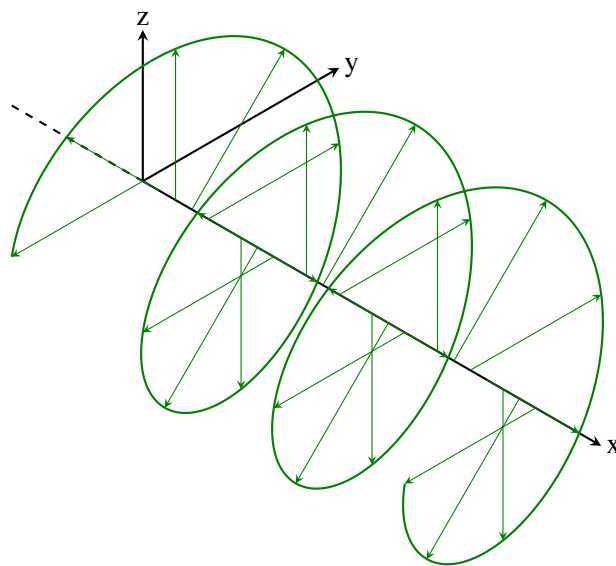


Figura 5.8: Luz polarizada circularmente.

Quando α ou β possuem parte imaginária diferente de zero, as componentes do campo elétrico nas direções x e y oscilam *fora de fase* e a direção do campo elétrico varia com o tempo. Nesse caso temos estados de polarização bem interessantes, chamados de polarização *circular* ou *elíptica*, como mostrado nas Figuras 5.8 e 5.9.

No Capítulo 4 vimos que o fato de os bits quânticos poderem estar em superposições da forma (4.1) faz com que os testes que podemos fazer sobre eles tenham também uma estrutura muito mais rica. Vejamos o que acontece no exemplo da polarização.

Se utilizarmos o polarizador orientado em outra direção, por exemplo a direção d mostrada na figura 5.7, criamos um novo teste para a polarização de um fóton que discrimina entre dois estados diferentes de $|H\rangle$ e $|V\rangle$. Nesse caso, o fóton será sempre transmitido se estiver no estado

$$|+\rangle = \frac{1}{\sqrt{2}}|H\rangle + \frac{1}{\sqrt{2}}|V\rangle,$$

que corresponde a polarização paralela à direção d , e será sempre absorvido se estiver no estado

$$|-\rangle = \frac{1}{\sqrt{2}}|H\rangle - \frac{1}{\sqrt{2}}|V\rangle,$$

que corresponde a polarização perpendicular à direção d . Chamaremos a alternativa em que o fóton é transmitido de resposta $+$ e a alternativa em que o fóton é absorvido de resposta $-$.

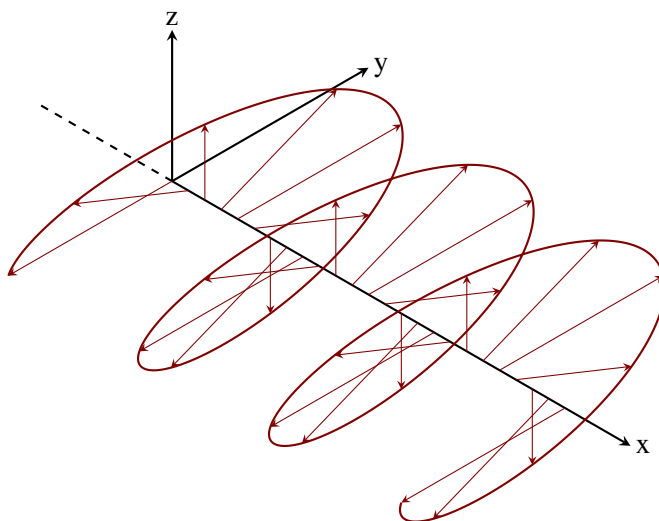


Figura 5.9: Luz elipticamente polarizada.

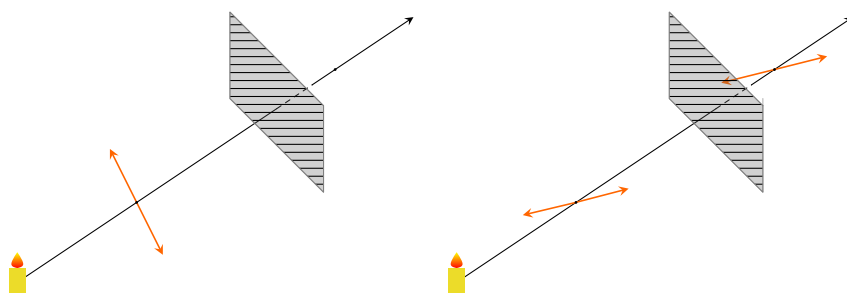


Figura 5.10: Ao posicionarmos o polarizador na direção d , criamos um teste que discrimina entre os estados de polarização $|+\rangle$ e $|-\rangle$. Se a polarização do fóton for paralela a d (estado $|+\rangle$), o fóton sempre será transmitido. Se a polarização do fóton for perpendicular a d (estado $|-\rangle$), o fóton sempre será absorvido. Para outros estados de polarização, as duas alternativas podem ocorrer, sendo impossível determinar de antemão qual será o resultado. Para determinar quais são as probabilidades de cada resposta, devemos escrever o estado de polarização do fóton como uma superposição dos estados $|+\rangle$ e $|-\rangle$.

Se o fóton estiver em qualquer outro estado de polarização $|\phi\rangle$, ele poderá tanto ser transmitido quanto absorvido, sendo impossível prever qual das duas alternativas irá ocorrer de fato. Para calcular a probabilidade de cada alternativa, escrevemos o estado do fóton como uma superposição dos estados $|+\rangle$ e $|-\rangle$:

$$|\phi\rangle = \gamma |+\rangle + \delta |-\rangle .$$

5.1. ONDAS ELETROMAGNÉTICAS

51

A probabilidade de obter a resposta $+$ é igual a $|\gamma|^2$ e a probabilidade de obter a resposta $-$ é $|\delta|^2$.

Por exemplo, o estado de polarização $|H\rangle$ pode ser escrito na forma

$$|H\rangle = \frac{1}{\sqrt{2}}|+\rangle + \frac{1}{\sqrt{2}}|-\rangle$$

e o estado de polarização $|V\rangle$ pode ser escrito na forma

$$|V\rangle = \frac{1}{\sqrt{2}}|+\rangle - \frac{1}{\sqrt{2}}|-\rangle.$$

Para ambos, a probabilidade de obter a resposta $+$ em um teste que discrimina os estados $|+\rangle$ e $|-\rangle$ é $\frac{1}{2}$, assim como a probabilidade de obter a resposta $-$.

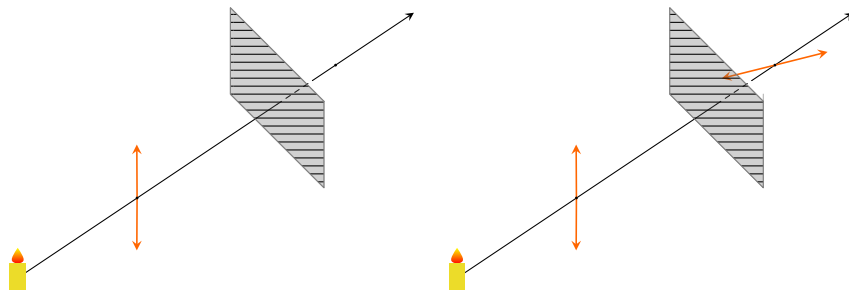


Figura 5.11: Um fóton no estado de polarização $|V\rangle$ passa por um polarizador orientado na direção d . O fóton pode tanto ser transmitido quanto ser absorvido, cada alternativa ocorrendo com probabilidade $\frac{1}{2}$.

Podemos também utilizar os polarizadores para realizar testes mais gerais. Quaisquer dois estados $|a\rangle$ e $|b\rangle$ que formem uma base ortonormal para $\mathbb{C}^2$ nos dão um teste que pode ser realizado sobre esse bit quântico, com duas respostas possíveis, que chamaremos a e b . De forma geral, podemos escrever

$$\begin{aligned} |a\rangle &= \mu |H\rangle + \nu |V\rangle \\ |b\rangle &= \nu |H\rangle - \mu |V\rangle. \end{aligned} \quad (5.3)$$

Para descobrir qual é a probabilidade de cada resposta para o teste que discrimina entre as alternativas a e b , precisamos escrever o estado do fóton como uma combinação linear de $|a\rangle$ e $|b\rangle$:

$$|\phi\rangle = \tau |a\rangle + \omega |b\rangle.$$

A probabilidade de obter a resposta a é igual a $|\tau|^2$ e a probabilidade de obter a resposta b é igual a $|\omega|^2$.

Se μ e ν são reais, para produzir experimentalmente o teste que distingue as alternativas a e b basta orientar um polarizador do tipo que mencionamos anteriormente, conhecido também como *polarizador linear*, na direção de $|a\rangle$ ou $|b\rangle$. Se

μ e ν são números complexos, também é possível realizar esse teste experimentalmente, utilizando elementos óticos um pouco mais sofisticados.

Por exemplo, se consideramos os estados de polarização circular à esquerda e à direita, representados matematicamente pelos vetores

$$\begin{aligned} |\odot\rangle &= \frac{1}{\sqrt{2}} |H\rangle + \frac{i}{\sqrt{2}} |V\rangle \\ |\oslash\rangle &= \frac{1}{\sqrt{2}} |H\rangle - \frac{i}{\sqrt{2}} |V\rangle, \end{aligned} \quad (5.4)$$

o teste que discrimina entre as alternativas $\odot$ e $\oslash$ pode ser implementado utilizando a chamada *placa de quarto de onda*, que atrasa a componente y do campo elétrico em relação à componente x , criando a diferença de fase desejada. Esse dispositivo óptico é também conhecido *polarizador circular* e assim como o polarizador linear pode ser utilizado em várias aplicações práticas. Por exemplo, eles são usados como filtros para reduzir reflexos indesejados na fotografia e compõe as lentes dos óculos usados para a visualização de alguns filmes 3D, onde a polarização da luz é usada para diferenciar que imagem deve ser vista pelo olho esquerdo e pelo olho direito.

5.2 Spin

Em 1922, os físicos alemães Otto Stern e Walther Gerlach realizaram um experimento em que átomos de prata eram enviados através de um íma até uma placa que indicava a localização em que os átomos colidiam com ela. Eles observaram que a posição final dos átomos era modificada pela presença do campo magnético produzido pelo ímã, e que ela podia ter apenas dois valores possíveis. Esse experimento provou que os átomos possuem uma característica física que responde à presença do campo magnético do ímã. Essa característica foi chamada pelos físicos de *spin*. As partículas que podem chegar a duas posições finais ao passar pelo ímã são chamadas de partículas de *spin* $\frac{1}{2}$. Outras partículas possuem valores diferentes de spin e nesse caso haverá mais do que duas possibilidades para a posição final quando elas são enviadas através do campo magnético. Nos concentraremos aqui nas partículas de spin $\frac{1}{2}$.

O spin de uma partícula de spin $\frac{1}{2}$ é mais um exemplo de um sistema físico que pode ser descrito através do formalismo de um bit quântico visto no Capítulo 4. Essas partículas podem estar em dois estados de spin que chamaremos de $|\uparrow\rangle$ e $|\downarrow\rangle$. Se o estado do sistema for $|\uparrow\rangle$, a partícula sempre colidirá com a placa na posição superior. Se o estado do sistema for $|\downarrow\rangle$, a partícula sempre colidirá com a placa na posição inferior. No entanto, a partícula pode estar em um estado de spin que é uma superposição dos estados $|\uparrow\rangle$ e $|\downarrow\rangle$:

$$|\phi\rangle = \alpha |\uparrow\rangle + \beta |\downarrow\rangle.$$

Para uma partícula no estado de spin $|\phi\rangle$, não podemos prever de antemão qual será a posição em que ela vai colidir com a placa. Podemos prever apenas as

probabilidades: com probabilidade $|\alpha|^2$ ela atinge a placa na parte superior e com probabilidade $|\beta|^2$ ela atinge a placa na parte inferior.

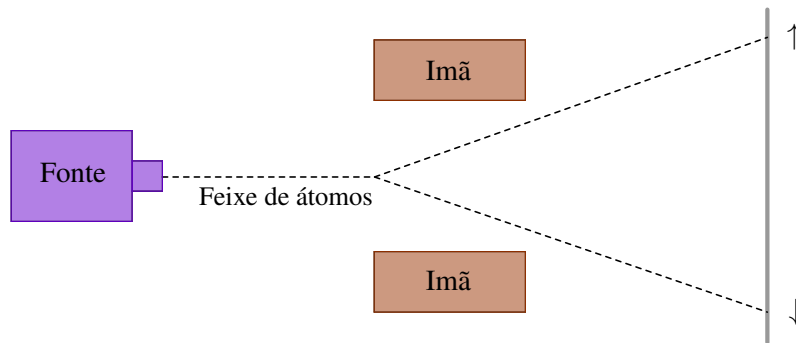


Figura 5.12: Um feixe de átomos de prata é enviado através de um campo magnético cuidadosamente elaborado. Existem duas possibilidades para cada átomo: ele pode ser desviado para cima, opção que indicaremos com o estado $|\uparrow\rangle$, ou pode ser desviado para baixo, opção que indicaremos com o estado $|\downarrow\rangle$. De maneira geral, um átomo está em um estado de superposição $\alpha|\uparrow\rangle + \beta|\downarrow\rangle$, e tem probabilidade $|\alpha|^2$ ela atinge a placa na parte superior e com probabilidade $|\beta|^2$ ela atinge a placa na parte inferior.

Para criar outros testes, podemos posicionar o ímã em outras direções. Por exemplo, ao posicionar os ímas um à direita, outro à esquerda, a trajetória dos átomos será deslocada ou para a direita, ou para a esquerda.

Através do estudo do spin podemos também entender porque a Física Quântica recebeu esse nome. A previsão clássica para o resultado desse experimento, sobre a qual não entraremos em detalhes aqui, diz que haveriam átomos colidindo com a placa em uma infinidade de posições entre a parte de cima e a de baixo. Isso não acontece. O que acontece é que os átomos colidem em apenas duas posições. Por isso, o resultado do experimento é *quantizado*, e esse termo levou ao nome da teoria que explica esse fenômeno.

Capítulo 6

Matrizes

Em matemática, uma *matriz* $m \times n$ é uma tabela de m linhas e n colunas de símbolos sobre um conjunto, normalmente um conjunto numérico, representada sob a forma de um quadro. Nesse capítulo veremos alguns aspectos da *álgebra matricial*.

Aplicações de matrizes são encontrados em inúmeras áreas da ciência. Em todos os ramos da física, incluindo a mecânica clássica, eletromagnetismo, óptica, mecânica quântica, e eletrodinâmica quântica, matrizes e suas generalizações são usadas para estudar fenômenos importantes, tais como o movimento de corpos rígidos. Matrizes também são utilizadas para representar estados, transformações e medições realizadas em determinados sistemas. Em computação gráfica, matrizes são usadas para projetar uma imagem tridimensional em uma tela bidimensional. Em teoria de probabilidade e estatística, matrizes estocásticas são usados para descrever conjuntos de probabilidades. Esse tipo de matriz aparece, por exemplo, no algoritmo *PageRank*, que ordena as páginas mostradas ao usuário em uma pesquisa no Google [5].

6.1 Definição

Definição 32. Uma *matriz* A é uma tabela de mn números dispostos em m linhas (horizontais) e n colunas (verticais):

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

Uma matriz com m linhas e n colunas é chamada de uma matriz m por n (escreve-se $m \times n$) e m e n são chamadas de suas dimensões, tipo ou ordem.

Aqui, consideraremos matrizes formadas por números reais ou complexos.

Exemplo 27. A matriz

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

é uma matriz do tipo 2×3 formada por números reais.

Exemplo 28. A matriz

$$B = \begin{bmatrix} 1+i & 2 \\ -i & 2-i \\ 3i & 2-4i \end{bmatrix}$$

é uma matriz do tipo 3×2 formada por números complexos.

Cada um dos símbolos que aparece em uma matriz é chamado de *elemento* ou *entrada* da matriz. Um elemento de uma matriz A que está na i -ésima linha e na j -ésima coluna é chamado de elemento ij ou (i, j) -ésimo elemento de A . Ele será denotado por a_{ij} ou $A[i, j]$. No Exemplo 28, o elemento a_{12} é 2, o número que aparece na primeira linha e segunda coluna do quadro.

O conjunto de todas as matrizes $m \times n$ com entradas reais será denotado por $M_{m \times n}(\mathbb{R})$. O conjunto de todas as matrizes $m \times n$ com entradas complexas será denotado por $M_{m \times n}(\mathbb{C})$.

Dizemos que duas matrizes A e B são iguais se elas têm o mesmo tamanho e os elementos correspondentes são iguais, ou seja, se A e B são ambas $m \times n$ e $a_{ij} = b_{ij}$ para todos os valores de i e j .

Três tipos de matrizes recebem nomes especiais:

- **Matriz linha ou vetor linha:** matriz do tipo $1 \times n$, ou seja, uma matriz que possui uma única linha.

Exemplo 29. A matriz 1×3

$$\begin{bmatrix} 3 & 7 & 2 \end{bmatrix} \in M_{1 \times 3}(\mathbb{R})$$

é uma matriz linha.

- **Matriz coluna ou vetor coluna:** matriz do tipo $n \times 1$, ou seja, uma matriz que possui uma única coluna.

Exemplo 30. A matriz 4×1

$$\begin{bmatrix} 1 \\ 4-i \\ i \\ 8-5i \end{bmatrix} \in M_{4 \times 1}(\mathbb{C})$$

é uma matriz coluna.

- **Matriz quadrada:** matriz do tipo $n \times n$, ou seja, uma matriz que tem o mesmo número de linhas e colunas.

6.2. OPERAÇÕES ENVOLVENDO MATRIZES

57

Exemplo 31. A matriz 3×3

$$\begin{bmatrix} 9 & 13 & 5 \\ 1 & 11 & 7 \\ 2 & 6 & 3 \end{bmatrix} \in M_{3 \times 3}(\mathbb{R})$$

é uma matriz quadrada.

Exemplo 32. Considere as seguintes matrizes com entradas reais:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, B = \begin{bmatrix} -1 & 2 \\ 1 & 3 \\ 0 & 1 \end{bmatrix}, C = \begin{bmatrix} 1 & 2 & 4 & 0 \end{bmatrix}, D = \begin{bmatrix} 1 \\ 3 \end{bmatrix},$$

$$E = \begin{bmatrix} -2 \end{bmatrix}, F = \begin{bmatrix} 1 & 2 & 4 & -1 & 0 \\ 0 & 3 & -1 & 4 & -5 \\ 7 & -3 & 1 & 4 & -4 \\ 2 & 1 & -1 & 0 & 5 \\ -5 & 1 & 2 & 0 & 1 \end{bmatrix}.$$

A é uma matriz quadrada 2×2 , B é uma matriz 3×2 , C é uma matriz linha 1×4 , D é uma matriz coluna 2×1 , E é uma matriz quadrada 1×1 e F é uma matriz quadrada 5×5 . Exemplos de elementos dessas matrizes: $a_{12} = 2$, $b_{32} = 1$, $c_{13} = 4$, $d_{21} = 3$, $e_{11} = -2$, $f_{54} = 0$.

Para matrizes quadradas, utilizaremos a notação mais compacta $M_{n \times n}(\mathbb{R}) = M_n(\mathbb{R})$ e $M_{n \times n}(\mathbb{C}) = M_n(\mathbb{C})$.

6.2 Operações envolvendo matrizes

6.2.1 Multiplicação de um escalar por uma matriz

A multiplicação de um número real ou complexo, chamado de *escalar*, por uma matriz é a operação matricial mais simples que podemos definir. Para multiplicar um número k por uma matriz A do tipo $n \times m$, basta multiplicar cada elemento a_{ij} de A por k . Assim, a matriz resultante B será também $n \times m$ e $b_{ij} = k \cdot a_{ij}$.

Exemplo 33. Dada a matriz

$$A = \begin{bmatrix} 1 & 8 & -3 \\ 4 & -2 & 5 \end{bmatrix},$$

ao multiplicá-la pelo escalar 2 obtemos

$$2A = 2 \begin{bmatrix} 1 & 8 & -3 \\ 4 & -2 & 5 \end{bmatrix} = \begin{bmatrix} 2 \times 1 & 2 \times 8 & 2 \times (-3) \\ 2 \times 4 & 2 \times (-2) & 2 \times 5 \end{bmatrix} = \begin{bmatrix} 2 & 16 & -6 \\ 8 & -4 & 10 \end{bmatrix}.$$

Exemplo 34. Dada a matriz

$$B = \begin{bmatrix} 1 & i \\ 0 & 2-i \end{bmatrix},$$

ao multiplicá-la pelo escalar i obtemos

$$iB = i \begin{bmatrix} 1 & i \\ 0 & 2-i \end{bmatrix} = \begin{bmatrix} i \times 1 & i \times i \\ 0 \times i & i \times (2-i) \end{bmatrix} = \begin{bmatrix} i & -1 \\ 0 & 1+2i \end{bmatrix}.$$

6.2.2 Adição e subtração de matrizes

Dadas as matrizes A e B do mesmo tipo $m \times n$, sua soma $A + B$ é a matriz $m \times n$ computada adicionando os elementos correspondentes:

$$(A + B)[i, j] = A[i, j] + B[i, j].$$

Exemplo 35. Considere as matrizes

$$A = \begin{bmatrix} 1 & 3 \\ 1 & 0 \\ 1 & 2 \end{bmatrix} \text{ e } B = \begin{bmatrix} 0 & 0 \\ 7 & 5 \\ 2 & 1 \end{bmatrix}.$$

Sua soma é obtida da seguinte maneira:

$$A + B = \begin{bmatrix} 1 & 3 \\ 1 & 0 \\ 1 & 2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 7 & 5 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 1+0 & 3+0 \\ 1+7 & 0+5 \\ 1+2 & 2+1 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 8 & 5 \\ 3 & 3 \end{bmatrix}.$$

Exemplo 36. Considere as matrizes

$$A = \begin{bmatrix} 3 & -1 & 2 \\ 1 & 1 & 2 \\ 3 & 2 & 2 \end{bmatrix} \text{ e } \begin{bmatrix} 1 & 0 & 4 \\ 2 & -5 & 0 \\ -2 & 1 & -1 \end{bmatrix}.$$

Sua soma é obtida da seguinte forma:

$$\begin{bmatrix} 3 & -1 & 2 \\ 1 & 1 & 2 \\ 3 & 2 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 4 \\ 2 & -5 & 0 \\ -2 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 3+1 & -1+0 & 2+4 \\ 1+2 & 1-5 & 2+0 \\ 3-2 & 2+1 & 2-1 \end{bmatrix} = \begin{bmatrix} 4 & -1 & 6 \\ 3 & -4 & 2 \\ 1 & 3 & 1 \end{bmatrix}.$$

Exemplo 37. Considere as matrizes

$$A = \begin{bmatrix} 1+i & 0 \\ -i & 2i \end{bmatrix} \text{ e } B = \begin{bmatrix} 1 & 2i \\ 0 & -1 \end{bmatrix}.$$

Sua soma é obtida da seguinte maneira:

$$A+B = \begin{bmatrix} 1+i & 0 \\ -i & 2i \end{bmatrix} + \begin{bmatrix} 1 & 2i \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} (1+i)+1 & 0+2i \\ -i+0 & 2i+(-1) \end{bmatrix} = \begin{bmatrix} 2+i & 2i \\ -i & -1+2i \end{bmatrix}.$$

6.2. OPERAÇÕES ENVOLVENDO MATRIZES

59

Dadas as matrizes A e B do mesmo tipo $m \times n$, sua subtração $A - B$ é a matriz $m \times n$ computada subtraindo os elementos correspondentes:

$$(A - B)[i, j] = A[i, j] - B[i, j].$$

Exemplo 38. Considere as matrizes

$$A = \begin{bmatrix} 1 & 3 \\ 1 & 0 \\ 1 & 2 \end{bmatrix} \text{ e } B = \begin{bmatrix} 0 & 0 \\ 7 & 5 \\ 2 & 1 \end{bmatrix}.$$

Sua subtração é obtida da seguinte maneira:

$$A - B = \begin{bmatrix} 1 & 3 \\ 1 & 0 \\ 1 & 2 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 7 & 5 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 1-0 & 3-0 \\ 1-7 & 0-5 \\ 1-2 & 2-1 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ -6 & -5 \\ -1 & 1 \end{bmatrix}.$$

Exemplo 39. Considere as matrizes

$$A = \begin{bmatrix} 3 & -1 & 2 \\ 1 & 1 & 2 \\ 3 & 2 & 2 \end{bmatrix} \text{ e } \begin{bmatrix} 1 & 0 & 4 \\ 2 & -5 & 0 \\ -2 & 1 & -1 \end{bmatrix}.$$

Sua subtração é obtida da seguinte forma:

$$\begin{bmatrix} 3 & -1 & 2 \\ 1 & 1 & 2 \\ 3 & 2 & 2 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 4 \\ 2 & -5 & 0 \\ -2 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 3-1 & -1-0 & 2-4 \\ 1-2 & 1-(-5) & 2-0 \\ 3-(-2) & 2-1 & 2-(-1) \end{bmatrix} = \begin{bmatrix} 2 & -1 & -2 \\ -1 & 6 & 2 \\ 5 & 1 & 3 \end{bmatrix}.$$

Observação. As operações $A - B$ e $A + (-1)B$ resultam na mesma matriz.

6.2.3 Multiplicação de matrizes

A multiplicação de duas matrizes é bem definida apenas se o número de colunas da matriz da esquerda é igual ao número de linhas da matriz da direita. Se A é uma matriz $m \times n$ e B é uma matriz $n \times p$, então seu produto AB é a matriz $m \times p$ (m linhas e p colunas) dada por:

$$(AB)[i, j] = A[i, 1]B[1, j] + A[i, 2]B[2, j] + \dots + A[i, n]B[n, j] = \sum_k A[i, k]B[k, j] \quad (6.1)$$

para cada par i e j .

A expressão acima pode parecer complicada, mas na prática o cálculo do elemento $(AB)[i, j]$ é bastante simples. Ele é obtido multiplicando-se os elementos da linha i de A pelos elementos correspondentes da coluna j de B e somando os resultados.

Exemplo 40. Dadas a matriz 2×3

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

e a matriz 3×2

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

seu produto AB é calculado da seguinte forma

$$\begin{bmatrix} 1 & 0 & 2 \\ -1 & 3 & 1 \end{bmatrix} \times \begin{bmatrix} 3 & 1 \\ 2 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} (1 \times 3 + 0 \times 2 + 2 \times 1) & (1 \times 1 + 0 \times 1 + 2 \times 0) \\ (-1 \times 3 + 3 \times 2 + 1 \times 1) & (-1 \times 1 + 3 \times 1 + 1 \times 0) \end{bmatrix} = \begin{bmatrix} 5 & 1 \\ 4 & 2 \end{bmatrix}.$$

Nesse caso também podemos calcular o produto BA :

$$\begin{aligned} \begin{bmatrix} 3 & 1 \\ 2 & 1 \\ 1 & 0 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 2 \\ -1 & 3 & 1 \end{bmatrix} &= \begin{bmatrix} (3 \times 1 + 1 \times (-1)) & 3 \times 0 + 1 \times 3 & 3 \times 2 + 1 \times 1 \\ 2 \times 1 + 1 \times 3 & 2 \times 0 + 1 \times 3 & 2 \times 2 + 1 \times 1 \\ 1 \times 1 + 0 \times (-1) & 1 \times 0 + 0 \times 3 & 1 \times 2 + 0 \times 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 3 & 7 \\ 5 & 3 & 5 \\ 1 & 0 & 2 \end{bmatrix}. \end{aligned}$$

Note que no exemplo acima $AB \neq BA$, já que a primeira é uma matriz 2×2 enquanto a segunda é uma matriz 3×3 . Isso mostra que a multiplicação de matrizes *não é comutativa*, ou seja, a ordem dos fatores pode alterar o valor do produto. Em alguns casos, pode ser possível calcular AB , mas o produto BA pode nem estar definido.

Exemplo 41. Dadas a matriz 1×3

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

e a matriz 3×2

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

seu produto AB é calculado da seguinte forma

$$\begin{bmatrix} 1 & 0 & 2 \end{bmatrix} \times \begin{bmatrix} 3 & 1 \\ 2 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} (1 \times 3 + 0 \times 2 + 2 \times 1) & (1 \times 1 + 0 \times 1 + 2 \times 0) \end{bmatrix} = \begin{bmatrix} 5 & 1 \end{bmatrix}.$$

Nesse caso NÃO podemos calcular o produto BA , já que uma matriz 3×2 não pode ser multiplicada por uma matriz 1×3 .

6.2. OPERAÇÕES ENVOLVENDO MATRIZES

61

Exemplo 42. Dadas a matriz 1×2

$$A = \begin{bmatrix} i & 1+i \end{bmatrix}$$

e a matriz 2×1

$$B = \begin{bmatrix} 1 \\ i \end{bmatrix}$$

seu produto AB é a matriz 1×1 dada por:

$$\begin{bmatrix} i & 1+i \end{bmatrix} \times \begin{bmatrix} 1 \\ i \end{bmatrix} = \begin{bmatrix} (i \times 1 + (1+i) \times i) \end{bmatrix} = \begin{bmatrix} -1 + 2i \end{bmatrix}.$$

Observe que nesse caso o produto BA está definido e será uma matriz 2×2 .**Teorema 14.** *Sejam α e β constantes reais ou complexas e A e B matrizes de tamanho apropriado. As operações matriciais satisfazem as seguintes propriedades:*

1. **Comutatividade da soma:** $A + B = B + A$;
2. **Associatividade da soma:** $A + (B + C) = (A + B) + C$;
3. **Associatividade do produto por escalar:** $\alpha(\beta A) = (\alpha\beta)A$;
4. **Distributividade do produto por escalar:** $(\alpha + \beta)A = \alpha A + \beta A$;
5. **Distributividade do produto por escalar:** $\alpha(A + B) = \alpha A + \alpha B$;
6. **Associatividade do produto:** $(AB)C = A(BC)$;
7. **Distributividade do produto de matrizes:** $A(B + C) = AB + AC$ e $(A + B)C = AC + BC$;
8. **Associatividade do produto por escalar:** $\alpha(AB) = (\alpha A)B = A(\alpha B)$.

Exercício 52. Prove a validade das propriedades acima.**Exercício 53.** Dadas as matrizes

$$A = \begin{bmatrix} 2 & -1 \\ 4 & -2 \end{bmatrix} \text{ e } B = \begin{bmatrix} 5 & 10 \\ 15 & 0 \end{bmatrix}$$

calcule $A + B$, $2A - 3B$, $\frac{1}{2}A + 5B$.**Exercício 54.** Utilizando as matrizes do exercício 53, calcule AB e BA . Conclua que o produto de matrizes não é comutativo.**Exercício 55.** Seja

$$C = \begin{bmatrix} 3 & 0 \\ -1 & 2 \end{bmatrix}.$$

1. Utilizando as matrizes A e B do exercício 53, calcule $(A + B) + C$ e $A + (B + C)$ (efetue primeiro a operação que está entre parênteses). As duas matrizes obtidas são iguais? Esse resultado ilustra a propriedade associativa da soma de matrizes.
2. Calcule $(AB)C$ e $A(BC)$. As duas matrizes obtidas são iguais? O resultado ilustra a propriedade associativa do produto de matrizes.

Exercício 56. Seja

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Utilizando as matrizes dos exercícios 53 e 55, calcule AI , IA , BI , IB , CI , IC (utilize um computador para realizar as contas). Você percebe alguma propriedade interessante ao calcular esses produtos? A matriz I é chamada *matriz identidade*. Veremos algumas propriedades dessa matriz na subseção 6.3.4.

Exercício 57. Entre as matrizes abaixo, quais podem ser somadas e quais podem ser multiplicadas? Justifique. Em caso afirmativo, calcule a soma ou o produto.

1. $X = \begin{bmatrix} 0 & 1 \\ 4 & -2 \end{bmatrix}, Y = \begin{bmatrix} 2 & -1 & 9 \\ 4 & 7 & -2 & 8 \end{bmatrix};$
2. $X = \begin{bmatrix} i & 0 \\ 1+i & -2 \end{bmatrix}, Y = \begin{bmatrix} 2 & -i \\ 4-2i & 0 \end{bmatrix};$
3. $X = \begin{bmatrix} 2 & 10 & 1 \\ 4 & 3 & -2 \\ 3 & 4 & 1 \end{bmatrix}, Y = \begin{bmatrix} 2 & -1 & 9 \\ 4 & -2 & 8 \end{bmatrix};$
4. $X = \begin{bmatrix} 0 & 1 & 3 \\ 4 & -2 & 2 \end{bmatrix}, Y = \begin{bmatrix} 2 & -1 & 9 \\ 4 & -2 & 8 \end{bmatrix};$
5. $X = \begin{bmatrix} 0 & 1 & 3 \\ 4 & -2 & 2 \end{bmatrix}, Y = \begin{bmatrix} 2 & -1 & 9 \\ 4 & -2 & 8 \\ -1 & 0 & 1 \end{bmatrix}.$

6.2.4 Transposição

Definição 33. A matriz transposta de uma matriz A do tipo $m \times n$ é a matriz A^T do tipo $n \times m$ que se obtém trocando as linhas pelas colunas de A , ou seja,

$$A^T[i, j] = A[j, i].$$

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \Leftrightarrow A^T = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix}.$$

Exemplo 43. 1. Se $A = \begin{bmatrix} 1 & 2 \end{bmatrix}$, então $A^T = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

2. Se $A = \begin{bmatrix} 1-i \\ 2+3i \end{bmatrix}$, então $A^T = [1-i \quad 2+3i]$.

3. Se $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ então $A^T = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$.

Teorema 1. *Seja c uma constante real ou complexa e A e B matrizes de tamanho adequado. As seguintes propriedades são válidas:*

1. $(A^T)^T = A$;
2. $(A + B)^T = A^T + B^T$;
3. $(cA)^T = cA^T$;
4. $(AB)^T = B^T A^T$.

A prova das propriedades acima pode ser encontrada nas referências [4, 5, 6].

Exercício 58. Seja

$$A = \begin{bmatrix} 2 & x^2 \\ 2x - 1 & 2 \end{bmatrix}.$$

Se $A = A^T$, encontre o valor de x .

6.2.5 Matriz Hermitiana Conjugada

Definição 34. A *matriz hermitiana conjugada* de uma matriz A do tipo $m \times n$ é a matriz $A^\dagger$ do tipo $n \times m$ que se obtém trocando as linhas pelas colunas de A e em seguida trocando cada entrada pelo seu complexo conjugado, ou seja,

$$A^\dagger[i, j] = \overline{A[j, i]}.$$

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{m,n} \end{bmatrix} \Leftrightarrow A^\dagger = \begin{bmatrix} \overline{a_{11}} & \overline{a_{21}} & \dots & \overline{a_{m1}} \\ \overline{a_{12}} & \overline{a_{22}} & \dots & \overline{a_{m2}} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{a_{1n}} & \overline{a_{2n}} & \dots & \overline{a_{mn}} \end{bmatrix}. \quad (6.2)$$

Observe que se A é real, $A^\dagger = A^T$.

Exemplo 44. 1. Se $A = \begin{bmatrix} 1 & 2 \end{bmatrix}$, então $A^\dagger = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

2. Se $A = \begin{bmatrix} 1-i \\ 2+3i \end{bmatrix}$, então $A^\dagger = [1+i \quad 2-3i]$.

$$3. \text{ Se } A = \begin{bmatrix} 1-i & 2i \\ 1+3i & 4 \end{bmatrix} \text{ então } A^\dagger = \begin{bmatrix} 1+i & 1-3i \\ -2i & 4 \end{bmatrix}.$$

Teorema 15. *Seja c uma constante real ou complexa e A e B matrizes de tamanho adequado. As seguintes propriedades são válidas:*

1. $(A^\dagger)^\dagger = A$;
2. $(A + B)^\dagger = A^\dagger + B^\dagger$;
3. $(cA)^\dagger = \bar{c}A^\dagger$;
4. $(AB)^\dagger = B^\dagger A^\dagger$.

Exercício 59. Seja

$$A = \begin{bmatrix} 2 & x \\ 1+i & 2 \end{bmatrix}.$$

Se $A = A^\dagger$, encontre o valor de x .

6.3 Algumas matrizes especiais

6.3.1 Matrizes Diagonais

Definição 35. A *diagonal principal* de uma matriz quadrada corresponde aos elementos a_{ij} com $i = j$.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}. \quad (6.3)$$

Definição 36. Uma *matriz diagonal* é uma matriz quadrada cujos elementos exteriores à diagonal principal são nulos.

Exemplo 45. As matrizes

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2-i & 0 \\ 0 & 0 & 3i \end{pmatrix}, \begin{pmatrix} 3-i & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

são matrizes diagonais. Observe que a definição de uma matriz diagonal permite que o elementos que pertencem à diagonal principal de uma matriz diagonal sejam nulos.

Várias operações matriciais preservam a forma de matrizes diagonais:

- O produto de um escalar por uma matriz diagonal é uma matriz diagonal;
- A soma de duas matrizes diagonais é uma matriz diagonal;
- O produto de duas matrizes diagonais é uma matriz diagonal.

6.3.2 Matrizes triangulares

Definição 37. Uma matriz quadrada é chamada *triangular* quando os elementos acima ou abaixo da diagonal principal são zero. Mais especificamente, uma *matriz triangular superior* é aquela em que os elementos abaixo da diagonal principal são nulos, ou seja, $a_{ij} = 0$ sempre que $i > j$; uma *matriz triangular inferior* é aquela em que os elementos acima da diagonal principal são nulos, ou seja, $a_{ij} = 0$ sempre que $i < j$.

Várias operações matriciais preservam a forma de matrizes triangulares:

- O produto de uma matriz triangular superior por uma constante é uma matriz triangular superior;
- A soma de duas matrizes triangulares superiores é uma matriz triangular superior;
- O produto de duas matrizes triangulares superiores é uma matriz triangular superior.

Analogamente, temos que:

- O produto de uma matriz triangular inferior por uma constante é uma matriz triangular inferior;
- A soma de duas matrizes triangulares inferiores é uma matriz triangular inferior;
- O produto de duas matrizes triangulares inferiores é uma matriz triangular inferior.

Exemplo 46. A matriz

$$A = \begin{bmatrix} 1 & 4 - i & 2i \\ 0 & 3 - 2i & 4 \\ 0 & 0 & i \end{bmatrix}$$

é uma matriz triangular superior e a matriz

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 8 & 0 \\ 4 & 9 & 7 \end{bmatrix}$$

é uma matriz triangular inferior.

Exercício 60. Mostre que se uma matriz é triangular superior e inferior simultaneamente, então ela é uma matriz diagonal.

6.3.3 Matrizes Nulas

A *matriz nula* $0_{m \times n}$ é a matriz $m \times n$ com todos os elementos iguais a zero.

Exemplo 47. As matrizes

$$0_{1 \times 1} = [0], \quad 0_{2 \times 2} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad 0_{2 \times 3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (6.4)$$

são matrizes nulas.

Em geral, a matriz nula $m \times n$ tem a forma

$$0_{m \times n} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}. \quad (6.5)$$

A matriz nula $m \times n$ é o elemento neutro para a soma das matrizes de tamanho $m \times n$, ou seja, para toda matriz A do tipo $m \times n$ valem as igualdades

$$A + 0_{m \times n} = 0_{m \times n} + A = A. \quad (6.6)$$

Em geral, o tamanho da matriz fica claro do contexto e escrevemos apenas 0 para denotar a matriz nula. Fiquem sempre atentos para não confundir com o número 0 ou as diferentes matrizes nulas entre si. Sempre que aparecer o símbolo 0, ele representará a matriz nula de tamanho adequado.

6.3.4 Matrizes Identidade

A *matriz identidade* $n \times n$ é uma matriz diagonal, cujos elementos da diagonal são todos iguais a 1. É denotada por I_n ou simplesmente I , quando o tamanho da matriz for claro do contexto. A matriz identidade I_n tem a seguinte forma:

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}. \quad (6.7)$$

A matriz I_n é o *elemento neutro* da multiplicação de matrizes $n \times n$. Mais precisamente, para qualquer matriz A do tipo $n \times n$, as seguintes igualdades são válidas:

$$AI_n = I_n A = A. \quad (6.8)$$

6.3.5 Matrizes Simétricas

Definição 38. Uma matriz é chamada *simétrica* se $A = A^T$.

Teorema 16. *Toda matriz simétrica é uma matriz quadrada.*

Demonstração. Seja A uma matriz $m \times n$. Sabemos que a transposição inverte as linhas e as colunas da matriz, A^T é uma matriz $n \times m$. Se A é uma matriz simétrica, $A = A^T$ e como duas matrizes são iguais somente se as dimensões são iguais, temos que $m = n$. $\square$

As entradas de uma matriz simétrica são *simétricas* em relação à diagonal principal.

Teorema 17. Se A é uma matriz simétrica, então $a_{ij} = a_{ji}$.

Demonstração. Primeiramente notamos que se $B = A^T$ então $b_{ij} = a_{ji}$, uma vez que a transposição troca as linhas e colunas de A . Da igualdade $A = A^T = B$ segue que $a_{ij} = b_{ij} = a_{ji}$ e o resultado está provado. $\square$

Exemplo 48. A matriz

$$\begin{bmatrix} 1 & 7 & 3 \\ 7 & 4 & -5 \\ 3 & -5 & 6 \end{bmatrix}$$

é uma matriz simétrica.

6.3.6 Matrizes Anti-simétricas

Definição 39. Uma matriz é chamada *anti-simétrica* se $A = -A^T$. Como duas matrizes são iguais somente se as dimensões são iguais, uma matriz só é anti-simétrica se ela é quadrada.

Exercício 61. Utilize um argumento semelhante ao utilizado no Teorema 16 para mostrar que toda matriz anti-simétrica é uma matriz quadrada.

Exercício 62. Utilize um argumento semelhante ao utilizado no Teorema 17 para mostrar que se A é uma matriz anti-simétrica, então $a_{ij} = -a_{ji}$.

Exemplo 49. A matriz

$$\begin{bmatrix} 0 & 2 & -1 \\ -2 & 0 & -4 \\ 1 & 4 & 0 \end{bmatrix}$$

é uma matriz anti-simétrica.

6.3.7 Matrices Hermitianas

Definição 40. Uma matriz é chamada *hermitiana* se $A = A^\dagger$.

Teorema 18. *Toda matriz hermitiana é uma matriz quadrada.*

Teorema 19. *Se A é uma matriz simétrica, então $a_{ij} = \overline{a_{ji}}$.*

Exercício 63. Utilize argumentos semelhantes aos utilizados nos Teoremas 16 e 17 para provar os Teoremas 18 e 19.

Exemplo 50. A matriz

$$\begin{bmatrix} 1 & 1-i \\ 1+i & -2 \end{bmatrix}$$

é uma matriz hermitiana.

6.3.8 Matrizes Anti-hermitianas

Definição 41. Uma matriz é chamada *anti-hermitiana* se $A = -A^\dagger$.

Teorema 20. *Toda matriz hermitiana é uma matriz quadrada.*

Teorema 21. *Se A é uma matriz anti-hermitiana, então $a_{ij} = -\overline{a_{ji}}$.*

Exercício 64. Utilize argumentos semelhantes aos utilizados nos Teoremas 16 e 17 para provar os Teoremas 20 e 21.

Exemplo 51. A matriz

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

é uma matriz anti-hermitiana.

6.3.9 Matrizes Ortogonais

Definição 42. Uma matriz quadrada $A \in M_n(\mathbb{R})$ é dita *ortogonal* se sua transposta coincide com a sua inversa. Isto é, A é ortogonal se

$$A^T = A^{-1} \quad (6.9)$$

ou, alternativamente:

$$A^T A = A A^T = I_n. \quad (6.10)$$

Exemplo 52. A matriz identidade

$$I_n = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

é uma matriz ortogonal.

Exemplo 53. A matriz

$$R_\theta = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

é uma matriz ortogonal.

Exemplo 54. A matriz

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

é uma matriz ortogonal.

Exercício 65. Mostre que a matriz

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

é uma matriz ortogonal.

6.3.10 Matrizes Unitárias

Definição 43. Uma matriz quadrada $U \in M_n(\mathbb{C})$ é dita *unitária* se sua transposta coincide com a sua inversa. Isto é, U é unitária se

$$U^{-1} = U^\dagger \quad (6.11)$$

ou, équivalentement, se

$$U^\dagger U = U U^\dagger = I_n. \quad (6.12)$$

Exemplo 55. A matriz identidade

$$I_n = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

é uma matriz unitária.

Exemplo 56. A matriz

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

é uma matriz unitária.

Exercício 66. Mostre que a matriz

$$A = \begin{bmatrix} \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \end{bmatrix}$$

é uma matriz unitária.

6.4 Representação matricial de $\mathbb{R}^2$ e $\mathbb{C}^2$

Utilizando matrizes coluna, podemos representar o conjunto $\mathbb{R}^2$ de maneira alternativa identificando-o com o conjunto de matrizes 2×1 com entradas reais:

$$\mathbb{R}^2 = \left\{ |u\rangle = \begin{bmatrix} x \\ y \end{bmatrix} \mid x, y \in \mathbb{R} \right\}. \quad (6.13)$$

De forma análoga, podemos também representar o conjunto $\mathbb{C}^2$ de maneira alternativa identificando-o com o conjunto de matrizes 2×1 com entradas complexas:

$$\mathbb{C}^2 = \left\{ |u\rangle = \begin{bmatrix} z \\ w \end{bmatrix} \mid z, w \in \mathbb{C} \right\}. \quad (6.14)$$

Observe que as operações de soma e produto por escalar em $\mathbb{R}^2$ e $\mathbb{C}^2$, com essa identificação, são equivalentes à soma e ao produto por escalar definidos na Seção 6.2 para as matrizes 2×1 correspondentes.

Também é possível representar o produto interno nessa nova notação matricial. No caso real, se

$$|u_1\rangle = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \text{ e } |u_2\rangle = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$$

então

$$\langle u_1 | u_2 \rangle = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}^T \times \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 & y_1 \end{bmatrix} \times \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = x_1 x_2 + y_1 y_2. \quad (6.15)$$

No caso complexo, se

$$|u_1\rangle = \begin{bmatrix} z_1 \\ w_1 \end{bmatrix} \text{ e } |u_2\rangle = \begin{bmatrix} z_2 \\ w_2 \end{bmatrix}$$

então

$$\langle u_1 | u_2 \rangle = \begin{bmatrix} z_1 \\ w_1 \end{bmatrix}^\dagger \times \begin{bmatrix} z_2 \\ w_2 \end{bmatrix} = \begin{bmatrix} \bar{z}_1 & \bar{w}_1 \end{bmatrix} \times \begin{bmatrix} z_2 \\ w_2 \end{bmatrix} = \bar{z}_1 z_2 + \bar{w}_1 w_2. \quad (6.16)$$

6.5 Transformações Lineares

Utilizando a representação matricial para $\mathbb{R}^2$, as matrizes em $M_2(\mathbb{R})$ definem uma aplicação que leva vetores de $\mathbb{R}^2$ em vetores de $\mathbb{R}^2$.

Definição 44. Dada uma matriz qualquer $A \in M_n(\mathbb{R})$, a função

$$\begin{aligned} T : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ |u\rangle &\longmapsto A |u\rangle \end{aligned} \quad (6.17)$$

é chamada uma *transformação linear* de $\mathbb{R}^2$ em $\mathbb{R}^2$.

6.5. TRANSFORMAÇÕES LINEARES

71

A nomenclatura para esse tipo de aplicação vem do fato de que elas satisfazem uma propriedade chamada *linearidade*: dados $|u_1\rangle, |u_2\rangle \in \mathbb{R}^2$ e $\lambda \in \mathbb{R}$ temos

$$\begin{aligned} T(u_1 + u_2) &= A(|u_1\rangle + |u_2\rangle) = A|u_1\rangle + A|u_2\rangle = T(|u_1\rangle) + T(|u_2\rangle), \\ T(\lambda u_1) &= A(\lambda |u_1\rangle) = \lambda A|u_1\rangle = \lambda T(|u_1\rangle). \end{aligned} \quad (6.18)$$

De maneira análoga, utilizando a representação matricial para $\mathbb{C}^2$, as matrizes em $M_2(\mathbb{C})$ definem uma aplicação que leva vetores de $\mathbb{C}^2$ em vetores de $\mathbb{C}^2$.

Definição 45. Dada uma matriz qualquer $A \in M_2(\mathbb{C})$, a função

$$\begin{aligned} T: \mathbb{C}^2 &\longrightarrow \mathbb{C}^2 \\ |u\rangle &\longmapsto A|u\rangle \end{aligned} \quad (6.19)$$

é chamada uma *transformação linear* de $\mathbb{C}^2$ em $\mathbb{C}^2$.

As transformações lineares em $\mathbb{C}^2$ também satisfazem a propriedade de *linearidade*: dados $|u_1\rangle, |u_2\rangle \in \mathbb{C}^2$ e $\lambda \in \mathbb{C}$ temos

$$\begin{aligned} T(u_1 + u_2) &= A(|u_1\rangle + |u_2\rangle) = A|u_1\rangle + A|u_2\rangle = T(|u_1\rangle) + T(|u_2\rangle), \\ T(\lambda u_1) &= A(\lambda |u_1\rangle) = \lambda A|u_1\rangle = \lambda T(|u_1\rangle). \end{aligned} \quad (6.20)$$

Exemplo 57. Seja T a transformação linear em $\mathbb{R}^2$ definida pela matriz

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

Se

$$|v_1\rangle = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

então

$$T(|v_1\rangle) = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \times \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

Exercício 67. Seja T a transformação linear em $\mathbb{R}^2$ definida pela matriz

$$A = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}.$$

Calcule $T(|v_1\rangle)$ e $T(|v_2\rangle)$ em que

$$|v_1\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ e } |v_2\rangle = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Exercício 68. Seja T a transformação linear em $\mathbb{C}^2$ definida pela matriz

$$A = \begin{bmatrix} 1 & 2i \\ -2 & i \end{bmatrix}.$$

Calcule $T(|v_1\rangle)$ e $T(|v_2\rangle)$ em que

$$|v_1\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ e } |v_2\rangle = \begin{bmatrix} 1 \\ -i \end{bmatrix}.$$

6.6 Algumas definições especiais para matrizes quadradas

6.6.1 Traço

Definição 46. O traço de uma matriz quadrada A de elementos a_{ij} é definido como sendo a soma dos elementos da diagonal principal, ou seja,

$$\text{Tr} A := \sum_{i=1}^n a_{ii} = a_{11} + a_{22} + \dots + a_{nn}. \quad (6.21)$$

Isso então define uma função que leva cada matriz em um número real ou complexo. Algumas de suas propriedades básicas estão condensadas no Teorema abaixo:

Teorema 22. Dadas matrizes A, B de tamanho adequado e λ uma constante real ou complexa, temos:

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

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

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

□

Exercício 69. Justifique cada passagem utilizada na demonstração do Teorema anterior.

Exemplo 58. Dada a matriz

$$A = \begin{bmatrix} 9 & 13 & 5 \\ 1 & 11 & 7 \\ 2 & 6 & 3 \end{bmatrix} \in M_{3 \times 3}(\mathbb{R}),$$

temos que

$$\text{Tr}(A) = 9 + 11 + 3 = 23.$$

Exercício 70. Calcule o traço das matrizes

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, B = \begin{bmatrix} -i & 2+i \\ 1 & 3+5i \end{bmatrix}, C = \begin{bmatrix} 1 & 2 & 4 & -1 & 0 \\ 0 & 3 & -1 & 4 & -5 \\ 7 & -3 & 1 & 4 & -4 \\ 2 & 1 & -1 & 0 & 5 \\ -5 & 1 & 2 & 0 & 1 \end{bmatrix}.$$

6.7 Matriz Inversa

Sabemos que todo número real não nulo a possui um inverso multiplicativo $a^{-1} = \frac{1}{a}$ tal que $a \times a^{-1} = a^{-1} \times a = 1$. Quando temos uma equação do tipo $ax = b$ em que a e b são números reais conhecidos e x é uma incógnita que deve ser encontrada, basta dividirmos ambos os lados da equação por a para encontrar sua solução

$$x = \frac{b}{a}.$$

Esse procedimento é possível sempre que $a \neq 0$.

Sabemos que para matrizes a divisão não está definida e portanto o mesmo procedimento não pode ser aplicado à equação matricial $AX = B$ para encontrar a solução de um sistema linear. No entanto, em alguns casos é possível definir o que chamamos de *matriz inversa* da matriz A , denotada por A^{-1} , de modo que $A^{-1}A = AA^{-1} = I$. Assim, uma equação matricial envolvendo a matriz A do tipo $AX = B$ pode ser resolvida multiplicando-se ambos os lados da equação por A^{-1} .

$$\begin{aligned} A^{-1}(AX) &= A^{-1}B \\ (A^{-1}A)X &= A^{-1}B \\ IX &= A^{-1}B \\ X &= A^{-1}B. \end{aligned}$$

6.7.1 Definição e propriedades

Definição 47. Uma matriz quadrada A $n \times n$ é dita invertível quando existe outra matriz A^{-1} $n \times n$ tal que

$$A^{-1}A = AA^{-1} = I \quad (6.22)$$

onde I é a matriz identidade $n \times n$.

Caso a inversa de A não exista, dizemos que A é uma matriz *singular*.

Se A uma matriz invertível, valem as seguintes propriedades:

1. A matriz inversa é única. De fato, supondo que A^{-1} e B sejam duas inversas para a matriz A temos

$$B = IB = (A^{-1}A)B = A^{-1}(AB) = A^{-1}I = A^{-1}.$$

2. A matriz inversa de uma matriz invertível é também invertível, sendo que a inversa da inversa de uma matriz é igual à própria matriz, ou seja, $A = (A^{-1})^{-1}$.

3. A matriz transposta de uma matriz invertível é também invertível, e a inversa da transposta é a transposta da inversa, ou seja, $(A^T)^{-1} = (A^{-1})^T$. De fato

$$A^T(A^{-1})^T = (A^{-1}A)^T = I^T = I \text{ e } (A^{-1})^T A^T = (AA^{-1})^T = I^T = I.$$

4. A inversa de uma matriz multiplicada por um número (diferente de zero) é igual à matriz inversa multiplicada pelo inverso desse número, ou seja

$$(\alpha A)^{-1} = \alpha^{-1} A^{-1}$$

5. O inverso do produto de matrizes invertíveis é igual aos produtos das inversas dessas matrizes com a ordem trocada, ou seja,

$$(AB)^{-1} = B^{-1}A^{-1}.$$

6. A matriz inversa de uma matriz identidade é sempre igual à própria matriz identidade, ou seja, $I^{-1} = I$. Essa propriedade decorre da igualdade $I \times I = I$.

Exercício 71. Dada uma matriz A , suponha que B é a matriz inversa de A^2 . Mostre que a matriz $C = AB$ é a inversa de A .

Exercício 72. Utilizando o exercício anterior, mostre que A é invertível se, e somente se, A^2 é invertível.

6.7.2 Determinação da inversa

O método de cálculo da inversa que utilizaremos aqui consiste em partir de uma matriz quadrada genérica, com incógnitas em vez de valores e aplicar a condição

$$AA^{-1} = I.$$

Exercício 73. Vamos calcular a inversa da matriz

$$A = \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}.$$

Solução. Sabemos que a matriz inversa tem que ser também uma matriz 2×2 e portanto é da forma

$$A^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

6.7. MATRIZ INVERSA

75

O objetivo é determinar os valores de a, b, c e d . Para isso aplicaremos a definição de inversa:

$$\begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix} \times \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Resolvendo essa multiplicação de matrizes obtemos:

$$\begin{bmatrix} 2a + c & 2b + d \\ 4a + 3c & 4b + 3d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

o que nos leva ao sistema de equações:

$$\begin{cases} 2a + c = 1, \\ 2b + d = 0, \\ 4a + 3c = 0, \\ 4b + 3d = 1. \end{cases}$$

Esse sistema pode ser resolvido facilmente (por substituição, por exemplo), e a solução nos leva a $a = \frac{3}{2}, b = \frac{-1}{2}, c = -2$ e $d = 1$. Assim, temos

$$A^{-1} = \begin{bmatrix} \frac{3}{2} & \frac{-1}{2} \\ -2 & 1 \end{bmatrix}.$$

□

Caso a matriz que queremos inverter não fosse invertível, chegaríamos a um sistema de equações que não possui solução.

Esse método se torna bastante trabalhoso para matrizes 3×3 e impraticável para matrizes maiores. Não abordaremos aqui outros métodos mais avançados para o cálculo da inversa. O leitor interessado pode encontrar uma explicação detalhada de outros métodos nas referências [4, 5, 6].

Exercício 74. Encontre, se possível, as inversas das matrizes abaixo utilizando o método mostrado no exemplo acima. Utilize um computador para escalonar as matrizes envolvidas.

$$1. A = \begin{bmatrix} 0 & 2 \\ -3 & 0 \end{bmatrix};$$

$$2. A = \begin{bmatrix} 2 & 0 \\ 4 & 2 \end{bmatrix};$$

$$3. A = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix};$$

$$4. A = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}.$$

Exercício 75. Encontre cinco matrizes 2×2 distintas de modo que cada uma delas seja sua própria inversa, ou seja, $A^2 = I$.

6.8 Determinantes

O determinante é uma função matricial que associa a cada matriz quadrada um escalar. Esta função fornece várias informações sobre a matriz A , algumas das quais veremos mais adiante.

Existem várias maneiras diferentes de definir o determinante de uma matriz. Adotaremos aqui uma abordagem operacional, definido o determinante a partir do cálculo recursivo através de matrizes menores e cofatores. O leitor interessado em definições alternativas pode procurar as referências [4, 5, 6].

6.8.1 Cálculo do determinante

Calcular o determinante de matrizes grandes é um processo extremamente trabalhoso. Para encontrar uma maneira de fazer esse cálculo procederemos recursivamente: calcularemos o determinante de matrizes 1×1 , em seguida calcularemos o determinante de matrizes 2×2 . Utilizando esses determinantes calcularemos o determinante de matrizes 3×3 e assim sucessivamente. Para calcular o determinante de matrizes $n \times n$ utilizaremos o determinante de matrizes $(n - 1) \times (n - 1)$.

Matrizes 1×1

Definição 48. Se $A = [a_{11}]$ então

$$\det(A) = a_{11}.$$

Exemplo 59. Se $A = [3]$, então $\det(A) = 3$.

Matrizes 2×2

Definição 49. O determinante de uma matriz 2×2 é a diferença entre o produto dos termos da diagonal principal e o produto dos termos da diagonal secundária.

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

Exemplo 60. O determinante da matriz

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

é dado por $\det(A) = 0 \times (-1) - 2 \times 1 = 0 - 2 = -2$.

Exemplo 61. O determinante da matriz

$$A = \begin{bmatrix} 1 & i \\ i & -1 \end{bmatrix}$$

é dado por $\det(A) = i \times (-1) - i \times i = -1 + 1 = 0$.

Menores e Cofatores

Para calcular o determinante recursivamente, vamos precisar da definição de *menor* e *cofator*.

Definição 50. Dada uma matriz $n \times n$, a *matriz menor* $\tilde{A}_{ij}$ relativa ao elemento A_{ij} é a matriz $(n - 1) \times (n - 1)$ obtida de A eliminando-se a linha i e a coluna j , ou seja, a linha e a coluna do elemento A_{ij} .

Exemplo 62. Dada a matriz

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

a matriz menor $\tilde{A}_{11}$ é obtida de A eliminando-se a primeira linha e a primeira coluna:

$$\begin{bmatrix} 1 & -1 & 0 \\ -1 & 3 & 1 \\ -2 & 2 & -1 \end{bmatrix} \Rightarrow \tilde{A}_{11} = \begin{bmatrix} 3 & 1 \\ 2 & -1 \end{bmatrix}.$$

A matriz menor $\tilde{A}_{12}$ é obtida de A eliminando-se a primeira linha e a segunda coluna:

$$\begin{bmatrix} 1 & -1 & 0 \\ -1 & 3 & 1 \\ -2 & 2 & -1 \end{bmatrix} \Rightarrow \tilde{A}_{12} = \begin{bmatrix} -1 & 1 \\ -2 & -1 \end{bmatrix}.$$

A matriz menor $\tilde{A}_{31}$ é obtida de A eliminando-se a terceira linha e a primeira coluna:

$$\begin{bmatrix} 1 & -1 & 0 \\ -1 & 3 & 1 \\ -2 & 2 & -1 \end{bmatrix} \Rightarrow \tilde{A}_{31} = \begin{bmatrix} -1 & 0 \\ 3 & 1 \end{bmatrix}.$$

A matriz menor $\tilde{A}_{33}$ é obtida de A eliminando-se a terceira linha e a terceira coluna:

$$\begin{bmatrix} 1 & -1 & 0 \\ -1 & 3 & 1 \\ -2 & 2 & -1 \end{bmatrix} \Rightarrow \tilde{A}_{33} = \begin{bmatrix} -1 & -1 \\ -1 & 3 \end{bmatrix}.$$

Definição 51. Dada uma matriz A , o *cofator* associado ao elemento A_{ij} é

$$C_{ij} = (-1)^{i+j} \det \left(\tilde{A}_{ij} \right).$$

Exercício 76. Dada a matriz

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

temos que

$$C_{11} = (-1)^{1+1} \det(\tilde{A}_{11}) = (-1)^{1+1} \det \begin{bmatrix} 3 & 1 \\ 2 & -1 \end{bmatrix} = -5;$$

$$C_{12} = (-1)^{1+2} \det(\tilde{A}_{12}) = (-1)^{1+2} \det \begin{bmatrix} -1 & 1 \\ -2 & -1 \end{bmatrix} = -3;$$

$$C_{31} = (-1)^{3+1} \det(\tilde{A}_{31}) = (-1)^{3+1} \det \begin{bmatrix} -1 & 0 \\ 3 & 1 \end{bmatrix} = -1;$$

$$C_{33} = (-1)^{3+3} \det(\tilde{A}_{33}) = (-1)^{3+3} \det \begin{bmatrix} -1 & -1 \\ -1 & 3 \end{bmatrix} = 3.$$

Observe que na definição dos cofatores de uma matriz $n \times n$, são utilizados determinantes de matrizes $(n-1) \times (n-1)$. Com o que temos até aqui só conseguimos calcular cofatores de matrizes 3×3 . Veremos agora que os cofatores de uma matriz 3×3 nos permitem calcular o determinante dessa matriz. Com isso poderemos calcular os cofatores de matrizes 4×4 . Os cofatores de uma matriz 4×4 , por sua vez, nos permitem calcular o determinante dessa matriz. Podemos então calcular os cofatores de matrizes 5×5 e assim sucessivamente.

Cálculo do determinante através de cofatores

Para calcular o determinante de uma matriz A $n \times n$, escolhemos uma linha qualquer i e somamos os elementos dessa linha multiplicados pelos seus cofatores:

$$\det(A) = a_{i1}C_{i1} + a_{i2}C_{i2} + a_{i3}C_{i3} + \dots + a_{in}C_{in} = \sum_{j=1}^n a_{ij}C_{ij}.$$

Observe que na soma acima o coeficiente i permanece fixo, enquanto o coeficiente j percorre todos os valores possíveis: a linha i fica fixa enquanto a soma percorre todas as colunas da matriz A .

Podemos calcular o determinante escolhendo uma coluna qualquer j e somando os elementos dessa coluna multiplicados pelos seus cofatores

$$\det(A) = a_{1j}C_{1j} + a_{2j}C_{2j} + a_{3j}C_{3j} + \dots + a_{nj}C_{nj} = \sum_{i=1}^n a_{ij}C_{ij}.$$

Observe que na soma acima o coeficiente j permanece fixo, enquanto o coeficiente i percorre todos os valores possíveis: a coluna j fica fixa enquanto a soma percorre todas as linhas da matriz A .

Observação. 1. A escolha da linha ou da coluna que utilizaremos para calcular o determinante é arbitrária. Qualquer escolha deverá produzir o mesmo valor para $\det(A)$. Escolha sempre a linha ou coluna que simplifique as contas o máximo possível. Quanto mais zeros, melhor.

2. A equação para o cálculo do determinante através de cofatores parece simples, mas na verdade ela não é tão simples assim. Ao utilizar essa equação trocamos o problema de calcular o determinante de uma matriz $n \times n$ pelo cálculo de n determinantes $(n-1) \times (n-1)$. Cada um desses determinantes, por sua vez, se transforma em $(n-1)$ determinantes $(n-2) \times (n-2)$ e esse processo deve ser repetido até que os cofatores possam ser calculados através de determinantes de matrizes 2×2 . De forma geral, o número de operações necessárias para calcular o determinante por esse método é maior que $n!$, o que é extremamente ineficiente, mesmo em computadores muito rápidos. O determinante pode ser calculado por outro método, o chamado método de Gauss-Jordan, que utiliza aproximadamente n^3 operações. O leitor interessado pode encontrar uma apresentação completa sobre esse método nas referências [4, 5, 6, 7].

O determinante possui diversas propriedades interessantes:

Teorema 23. 1. Se A é uma matriz quadrada triangular inferior, então $\det(A)$ é o produto dos elementos da diagonal principal.

2. Se uma matriz A possui duas linhas iguais, então $\det(A) = 0$.
3. $\det(A^T) = \det(A)$.
4. $\det(AB) = \det(A) \det(B)$.

Exercício 77. Utilizando o teorema 23, mostre que se A é uma matriz invertível então

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

Exercício 78. Se $\det(A) = 3$ encontre

1. $\det(A^2)$;
2. $\det(A^4)$;
3. $\det(A^{-1})$;
4. $\det(A^T)$.

É possível calcular $\det(2A)$?

Exercício 79. Suponha que B seja uma matriz invertível. Prove que para qualquer matriz A vale

$$\det(B^{-1}AB) = \det(A).$$

O determinante de uma matriz A está intimamente relacionado a existência da inversa A^{-1} .

Teorema 24. *Uma matriz A $n \times n$ é invertível se, e somente se, $\det(A) \neq 0$.*

A prova dos resultados 23 e 24 está fora do escopo desse texto. O leitor interessado pode encontrar mais detalhes nas referências [4, 5, 6, 7].

Exercício 80. Mostre que se $\det(AB) = 0$, ou A é singular ou B é singular.

Exercício 81. Seja A uma matriz invertível tal que $A^2 = A$. Calcule $\det(A)$.

Exercício 82. Seja A uma matriz tal que $A^k = 0$ para algum $k \in \mathbb{N}$. Mostre que A é uma matriz singular.

Exercício 83. Suponhamos que A seja uma matriz tal que $A^T = A^{-1}$. Verifique quais são os possíveis valores de $\det(A)$.

Capítulo 7

Generalização para dimensões maiores

Os resultados vistos no Capítulo 2 podem ser generalizados para vetores com mais coordenadas, o que nos leva naturalmente à definição do espaço vetorial $\mathbb{R}^n$. Apesar da generalidade com que pretendemos tratar o assunto, devemos sempre nos lembrar dos espaços vetoriais $\mathbb{R}^2$ e $\mathbb{R}^3$ onde toda a teoria pode ser traduzida de maneira geométrica, fazendo com que todos os resultados enunciados possam ser compreendidos de maneira intuitiva. Podemos sempre trabalhar com as duas abordagens: a abordagem algébrica e abstrata, que pode ser aplicada igualmente a todos os espaços vetoriais, e a abordagem geométrica que possibilita a visualização dos conceitos da álgebra linear nos espaços $\mathbb{R}^2$ e $\mathbb{R}^3$.

7.1 O Espaço $\mathbb{R}^n$

Definição 52. Representamos por $\mathbb{R}^n$ o conjunto de todas as n -úplas de números reais, que denotaremos por

$$|u\rangle = (x_1, x_2, \dots, x_n). \quad (7.1)$$

Definição 53. Dados $|u\rangle = (x_1, x_2, \dots, x_n)$ e $|v\rangle = (y_1, \dots, y_n)$ em $\mathbb{R}^n$ definimos a soma $|u\rangle + |v\rangle$

$$|u\rangle + |v\rangle = (x_1 + y_1, \dots, x_n + y_n) \quad (7.2)$$

e o produto de $|u\rangle$ por um escalar $\lambda \in \mathbb{R}$

$$\lambda |u\rangle = (\lambda x_1, \dots, \lambda x_n). \quad (7.3)$$

Assim como fizemos para $\mathbb{R}^2$, será útil posteriormente identificar vetores em $\mathbb{R}^n$ e matrizes coluna $n \times 1$:

$$(x_1, \dots, x_n) \leftrightarrow \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}. \quad (7.4)$$

Teorema 25. Para todos $|u\rangle, |v\rangle, |w\rangle \in \mathbb{R}^n$ e $\lambda, \nu \in \mathbb{R}$ temos

1. (Associatividade) $|u\rangle + (|v\rangle + |w\rangle) = (|u\rangle + |v\rangle) + |w\rangle$;
2. (Comutatividade) $|u\rangle + |v\rangle = |v\rangle + |u\rangle$;
3. (Existência de zero) O vetor $\mathbf{0} = (0, 0, \dots, 0)$ é tal que $|u\rangle + \mathbf{0} = |u\rangle$;
4. (Existência de inverso aditivo) Dado $|u\rangle \in V$ o vetor $-|u\rangle$ é tal que $|u\rangle + (-|u\rangle) = \mathbf{0}$;
5. (Associatividade) $\lambda(\nu |u\rangle) = (\lambda\nu) |u\rangle$;
6. (Distributividade) $\lambda(|u\rangle + |v\rangle) = \lambda |u\rangle + \lambda |v\rangle$;
7. (Distributividade) $(\lambda + \nu) |u\rangle = \lambda |u\rangle + \nu |u\rangle$;
8. $1 |u\rangle = |u\rangle$.

As operações de soma e multiplicação por escalar nos espaços vetoriais $\mathbb{R}^2$ e $\mathbb{R}^3$ podem ser representados geometricamente, o que faz com que eles sejam importantes também do ponto de vista pedagógico.

Para $\mathbb{R}^2$, a representação geométrica de vetores, soma e produto por escalar foram mostradas nas Figuras 2.3, 2.4 e 2.5. Em $\mathbb{R}^3$, podemos proceder de maneira análoga: cada vetor $|u\rangle = (x, y, z)$ é representado por um segmento orientado que liga a origem ao ponto com coordenadas (x, y, z) . A figura 7.1 mostra a representação geométrica do vetor $|u\rangle = (1, 1, 2) \in \mathbb{R}^3$.

A soma de dois vetores também pode ser representada geometricamente. Dados dois vetores $|u\rangle$ e $|v\rangle$, sua soma é igual ao vetor obtido da seguinte forma: tomamos o segmento orientado que representa $|u\rangle$; em seguida, tomamos o segmento orientado que representa $|v\rangle$ com origem na extremidade de $|u\rangle$; o vetor $|u\rangle + |v\rangle$ é representado pelo segmento orientado que vai da origem até a extremidade de $|v\rangle$. A soma dos vetores $|u\rangle = (2, 2, 1)$ e $|v\rangle = (1, -1, 3)$ é ilustrada na Figura 7.2.

A multiplicação de um vetor por um escalar também pode ser representada geometricamente. Dado um vetor $|v\rangle \in \mathbb{R}^3$ e $\lambda \in \mathbb{R}$, o vetor $\lambda|v\rangle$ é encontrado da seguinte forma: se $\lambda = 0$ então $\lambda|v\rangle = \mathbf{0}$; caso contrário, $\lambda|v\rangle$ tem comprimento $|\lambda|$ vezes o comprimento de $|v\rangle$ e mesma direção de $|v\rangle$ (dizemos que eles são *paralelos*); $\lambda|v\rangle$ tem o mesmo sentido de $|v\rangle$ se $\lambda > 0$ e sentido oposto se $\lambda < 0$. A Figura 7.3 ilustra os vetores $|v\rangle = (1, 1, 1)$ e $2|v\rangle = (2, 2, 2)$.

7.2 Subespaços Vetoriais

Definição 54. Um *subespaço vetorial* S do espaço vetorial $\mathbb{R}^n$ é um subconjunto de $\mathbb{R}^n$ que é preservado pelas operações de soma e multiplicação por escalar definidas em $\mathbb{R}^n$. Para isso, precisamos que as seguintes propriedades sejam satisfeitas:

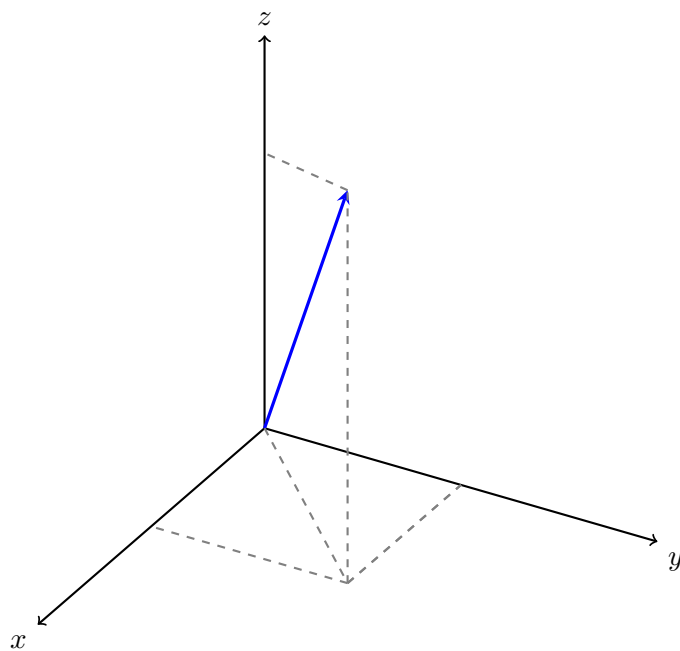


Figura 7.1: Representação geométrica do vetor $|u\rangle = (1, 1, 2)$ em $\mathbb{R}^3$.

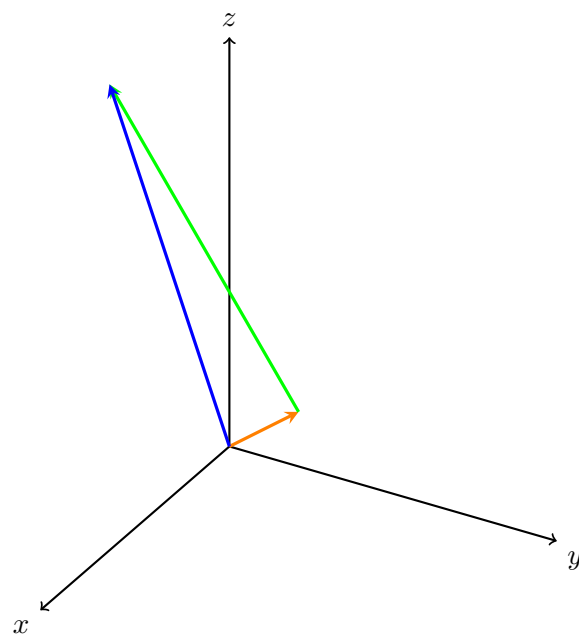


Figura 7.2: Representação geométrica da soma $(2, 2, 1) + (1, -1, 3) = (3, 1, 4)$.

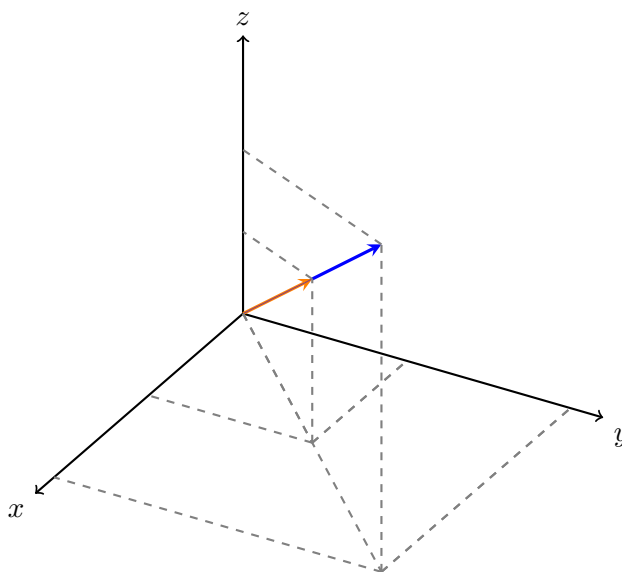


Figura 7.3: Representação geométrica de $(1, 1, 1)$ e $2 \times (1, 1, 1) = (2, 2, 2)$.

1. $|u\rangle + |v\rangle \in S$ para todo par $|u\rangle$ e $|v\rangle \in S$;
2. $\lambda |u\rangle \in S$ para todo $\lambda \in \mathbb{R}$ e todo $|v\rangle \in S$.

Para mostrar que S é um subespaço vetorial de $\mathbb{R}^n$ devemos verificar se as operações de soma e multiplicação por escalar *preservam* S : ao somar dois elementos de S devemos permanecer em S e ao multiplicar um elemento de S por um escalar devemos permanecer em S .

Exemplo 63. Em $\mathbb{R}^2$, os subespaços correspondem a todo o espaço $\mathbb{R}^2$, retas passando pela origem ou o conjunto $\{\mathbf{0}\}$. Em $\mathbb{R}^3$, os subespaços correspondem a todo o espaço $\mathbb{R}^3$, planos passando pela origem, retas passando pela origem ou o conjunto $\{\mathbf{0}\}$.

Exemplo 64. O subconjunto

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

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

Solução. A solução de um problema como esse passa sempre por três passos. O primeiro deles é verificar qual é a propriedade que define o subconjunto S . Nesse exemplo, S é o conjunto dos vetores de $\mathbb{R}^n$ que possuem todas as coordenadas iguais a zero a partir da segunda. A primeira coordenada pode assumir qualquer valor.

O segundo passo é verificar se a soma de quaisquer elementos de S permanece em S . Suponhamos então que $|u\rangle$ e $|v\rangle \in S$. Então $|u\rangle = (t_1, 0, \dots, 0)$ e $|v\rangle =$

7.3. COMBINAÇÕES LINEARES

85

$(t_2, 0, \dots, 0)$ com t_1 e $t_2 \in \mathbb{R}$. Assim

$$|u\rangle + |v\rangle = (t_1 + t_2, 0, \dots, 0)$$

também pertence a S .

O terceiro passo é verificar se a multiplicação de elementos de S por constante permanece em S . Tomando $|u\rangle = (t_1, 0, \dots, 0) \in S$ qualquer e $\lambda \in \mathbb{R}$ temos

$$\lambda |u\rangle = (\lambda t_1, 0, \dots, 0)$$

que também pertence a S . Como as propriedades 1 e 2 são satisfeitas, S é um subespaço de $\mathbb{R}^n$.

Observe que em $\mathbb{R}^2$ e $\mathbb{R}^3$ esse subespaço corresponde aos vetores sobre o eixo x e portanto S é uma reta passando pela origem, o que coincide com o que já havíamos visto anteriormente.

□

Exercício 84. Verifique quais dos subconjuntos de $\mathbb{R}^3$ abaixo são subespaços vetoriais.

1. O subconjunto de vetores que possuem primeira coordenada igual a zero;
2. O subconjunto de vetores que possuem primeira coordenada igual a 1;
3. O subconjunto de vetores (x, y, z) tais que $xy = 0$;
4. O subconjunto de vetores (x, y, z) tais que $z - y + 3x = 0$.

7.3 Combinações Lineares

Em $\mathbb{R}^n$ os elementos podem ser combinados livremente através da soma e da multiplicação por escalar.

Definição 55. Dizemos que $|v\rangle$ é uma *combinação linear* de $|u_1\rangle, \dots, |u_k\rangle$ se existem constantes $a_1, \dots, a_k \in \mathbb{R}$ tais que

$$|v\rangle = a_1 |u_1\rangle + \dots + a_k |u_k\rangle. \quad (7.5)$$

Exemplo 65. O vetor $(7, 2, 9) \in \mathbb{R}^3$ é uma combinação linear dos vetores $(2, 1, 3)$ e $(1, 0, 1)$ porque

$$(7, 2, 9) = 2(2, 1, 3) + 3(1, 0, 1).$$

Exemplo 66. O vetor nulo $\mathbf{0}$ é sempre combinação linear de qualquer conjunto de vetores $|v_1\rangle, \dots, |v_k\rangle$. De fato,

$$\mathbf{0} = 0 |v_1\rangle + \dots + 0 |v_k\rangle.$$

Exemplo 67. Considere os vetores $|e_1\rangle = (1, 0, 0)$, $|e_2\rangle = (0, 1, 0)$ e $|e_3\rangle = (0, 0, 1)$. Qualquer vetor em $\mathbb{R}^3$ é combinação linear desses três vetores. De fato, dado $|v\rangle = (a_1, a_2, a_3)$ podemos escrever

$$\begin{aligned}(a_1, a_2, a_3) &= (a_1, 0, 0) + (0, a_2, 0) + (0, 0, a_3) \\ &= a_1(1, 0, 0) + a_2(0, 1, 0) + a_3(0, 0, 1) \\ &= a_1|e_1\rangle + a_2|e_2\rangle + a_3|e_3\rangle.\end{aligned}$$

De maneira geral, qualquer vetor de $\mathbb{R}^n$ pode ser escrito como uma combinação linear dos vetores $|e_1\rangle = (1, 0, \dots, 0)$, $|e_2\rangle = (0, 1, \dots, 0), \dots, |e_n\rangle = (0, 0, \dots, 1)$.

Exemplo 68. Se um vetor $|v\rangle$ é combinação linear de um vetor $|u\rangle$ então

$$|v\rangle = \lambda |u\rangle$$

ou seja, $|v\rangle$ é múltiplo de $|u\rangle$.

Exemplo 69. Considere os vetores $|u_1\rangle = (1, 0, 0)$ e $|u_2\rangle = (1, 1, 0)$. O vetor $|v\rangle = (4, 2, 0)$ é combinação linear de $|u_1\rangle$ e $|u_2\rangle$. Para provar esse fato, suponhamos inicialmente que $|v\rangle$ possa ser escrito como combinação de $|u_1\rangle$ e $|u_2\rangle$. Isso quer dizer que existem $a_1, a_2 \in \mathbb{R}$ tais que

$$\begin{aligned} |v\rangle &= a_1 |u_1\rangle + a_2 |u_2\rangle \\ (4, 2, 0) &= a_1 (1, 0, 0) + a_2 (1, 1, 0). \end{aligned} \quad (7.6)$$

Utilizando a identificação entre vetores de $\mathbb{R}^n$ e matrizes coluna, podemos escrever

$$\begin{bmatrix} 4 \\ 2 \\ 0 \end{bmatrix} = a_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a_1 + a_2 \\ a_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}.$$

Para encontrar as constantes a_1 e a_2 devemos então resolver o sistema de três equações e duas incógnitas

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 0 \end{bmatrix}.$$

Esse sistema possui solução única $a_1 = 2$ e $a_2 = 2$. Logo $|v\rangle$ é combinação de $|u_1\rangle$ e $|u_2\rangle$. De fato,

$$(4, 2, 0) = 2(1, 0, 0) + 2(1, 1, 0).$$

Exemplo 70. Considere os vetores $|u_1\rangle = (1, 0, 0)$ e $|u_2\rangle = (1, 1, 0)$. O vetor $|v\rangle = (1, 2, -1)$ não é combinação linear de $|u_1\rangle$ e $|u_2\rangle$. Para provar esse fato,

7.3. COMBINAÇÕES LINEARES

87

suponhamos inicialmente que $|v\rangle$ possa ser escrito como combinação de $|u_1\rangle$ e $|u_2\rangle$. Isso quer dizer que existem $a_1, a_2 \in \mathbb{R}$ tais que

$$\begin{aligned} |v\rangle &= a_1 |u_1\rangle + a_2 |u_2\rangle \\ (1, 2, -1) &= a_1(1, 0, 0) + a_2(1, 1, 0). \end{aligned} \quad (7.7)$$

Utilizando a identificação entre vetores de $\mathbb{R}^n$ e matrizes coluna, podemos escrever

$$\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = a_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a_1 + a_2 \\ a_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}.$$

Para encontrar as constantes a_1 e a_2 devemos então resolver o sistema de três equações e duas incógnitas

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}.$$

Esse sistema não possui solução, uma vez que a última equação equivale a $0 = -1$.

De fato, qualquer combinação linear desses vetores é da forma

$$a_1(1, 0, 0) + a_2(1, 1, 0) = (a_1 + a_2, a_2, 0),$$

que nunca pode ser igual ao vetor $(1, 2, -1)$.

De maneira geral, dados vetores $|u_1\rangle, \dots, |u_m\rangle \in \mathbb{R}^n$, um vetor $|v\rangle$ será combinação linear de $|u_1\rangle, \dots, |u_m\rangle$ se existirem constantes $a_1, \dots, a_m$ tais que

$$|v\rangle = a_1 |u_1\rangle + \dots + a_m |u_m\rangle.$$

Essa equação nos levará a um sistema de equações lineares para as constantes $a_1, \dots, a_m$. Se esse sistema possuir solução, tal solução nos dará valores de $a_1, \dots, a_m$ que podem ser utilizados para escrever $|v\rangle$ como combinação linear de $|u_1\rangle, \dots, |u_m\rangle$. Se esse sistema não possui solução, $|v\rangle$ não pode ser escrito como combinação linear de $|u_1\rangle, \dots, |u_m\rangle$.

Exercício 85. Verifique se o vetor $|v\rangle$ é combinação linear dos vetores $|u_1\rangle, \dots, |u_n\rangle$ dados abaixo. Em caso afirmativo, encontre uma combinação linear desses vetores que gera $|v\rangle$.

1. $|v\rangle = (1, 1); |u_1\rangle = (1, 0);$
2. $|v\rangle = (2, 2); |u_1\rangle = (1, 1);$
3. $|v\rangle = (2, 3); |u_1\rangle = (1, 0), |u_2\rangle = (0, 1);$
4. $|v\rangle = (2, 3); |u_1\rangle = (1, 0), |u_2\rangle = (1, 1);$

5. $|v\rangle = (1, 1, 1); |u_1\rangle = (1, 0, 0), |u_2\rangle = (0, 1, 0);$
6. $|v\rangle = (2, 2, 2); |u_1\rangle = (1, 0, 1), |u_2\rangle = (0, 1, 0);$
7. $|v\rangle = (2, 2, 2); |u_1\rangle = (1, 0, 1), |u_2\rangle = (0, 1, 0), |u_3\rangle = (1, 1, 1);$
8. $|v\rangle = (2, 3, 1); |u_1\rangle = (1, 0, 0), |u_2\rangle = (0, 1, 0), |u_3\rangle = (0, 0, 1);$
9. $|v\rangle = (2, 3, 1); |u_1\rangle = (1, 1, 0), |u_2\rangle = (0, 1, 0), |u_3\rangle = (0, 1, 1);$
10. $|v\rangle = (2, 3, 1, -1); |u_1\rangle = (1, 0, 0, 0), |u_2\rangle = (0, 1, 0, 0), |u_3\rangle = (0, 0, 1, 0);$
11. $|v\rangle = (2, 3, 1, -1); |u_1\rangle = (1, 0, 0, 0), |u_2\rangle = (0, 1, 0, 0), |u_3\rangle = (0, 0, 1, 0), |u_4\rangle = (0, 0, 0, 1);$
12. $|v\rangle = (2, 3, 1, -1); |u_1\rangle = (1, 1, 0, 0), |u_2\rangle = (0, 1, 0, 0), |u_3\rangle = (0, 0, 1, 0), |u_4\rangle = (0, 0, 1, 1);$
13. $|v\rangle = (2, 3, 1, -1); |u_1\rangle = (1, 1, 0, 0), |u_2\rangle = (0, 1, 0, 0), |u_3\rangle = (0, 0, 1, 0), |u_4\rangle = (0, 0, 1, 1), |u_5\rangle = (1, 1, 1, 1).$

7.4 Dependência e Independência linear

Queremos encontrar condições que informem se dentre um conjunto de vetores, algum deles pode ser escrito como combinação linear dos outros. Para isso definimos a noção de dependência linear.

Definição 56. Dizemos que um conjunto de vetores $\{|u_1\rangle, \dots, |u_k\rangle\} \subset \mathbb{R}^n$ é *linearmente independente* (LI) se a equação

$$a_1 |u_1\rangle + \cdots + a_k |u_k\rangle = \mathbf{0} \quad (7.8)$$

só admite a solução trivial $a_1 = \dots = a_k = 0$, ou seja, se a única combinação linear dos vetores que gera o vetor nulo é aquela em que todas as constantes são nulas.

Caso contrário, dizemos que os vetores são *linearmente dependentes* (LD). Um conjunto de vetores é LD se existe uma combinação linear deles que gera o vetor nulo em que pelo menos uma das constantes é não-nula.

Exemplo 71. Um conjunto de vetores que contém o vetor nulo é sempre LD. De fato, tomando o conjunto $\{|u_1\rangle, \dots, |u_k\rangle, \mathbf{0}\}$ podemos escrever a combinação linear

$$0 \times |u_1\rangle + \cdots + 0 \times |u_k\rangle + 1 \times \mathbf{0} = \mathbf{0}. \quad (7.9)$$

Temos então uma combinação linear dos vetores que gera o vetor nulo em que a última constante utilizada é diferente de zero. Isso implica que esses vetores são LD.

Exemplo 72. Um conjunto formado por um único vetor não nulo $|u\rangle$ é sempre LI. Nesse caso, as únicas combinações lineares possíveis são múltiplos de $|u\rangle$ e para que um desses múltiplos seja o vetor nulo temos que

$$\lambda |u\rangle = \mathbf{0},$$

o que implica que $\lambda = 0$ uma vez que $|u\rangle \neq 0$.

Três ou mais vetores em $\mathbb{R}^2$, assim como quatro ou mais vetores em $\mathbb{R}^3$, assim como mais de n vetores em $\mathbb{R}^n$ são sempre LD. Isso acontece porque o problema de identificar se esses vetores são ou não LI leva a um sistema de equações com mais incógnitas que equações que sempre possui solução não trivial.

Teorema 26. *O número máximo de vetores LI em $\mathbb{R}^n$ é igual a n .*

A definição de dependência linear é importante porque ela permite identificar se dentro de um conjunto de vetores, um deles pode ser escrito como combinação linear dos outros ou não. De fato, as duas propriedades são equivalentes.

Teorema 27. *Um conjunto de vetores é LD se, e somente se, podemos expressar ao menos um dos vetores como combinação linear dos outros.*

Demonstração. Suponhamos que $\{|v_1\rangle, \dots, |v_k\rangle\}$ seja um conjunto LD. Então existe uma combinação linear desses vetores que resulta no vetor nulo

$$a_1 |v_1\rangle + \cdots + a_k |v_k\rangle = \mathbf{0} \quad (7.10)$$

com pelo menos uma das constantes $a_i \neq 0$. Suponhamos que $a_1 \neq 0$. Podemos então escrever

$$|v_1\rangle = -\frac{a_2}{a_1}|v_2\rangle - \dots - \frac{a_k}{a_1}|v_k\rangle \quad (7.11)$$

e $|v_1\rangle$ é combinação linear dos outros vetores.

Por outro lado, se um dos vetores, digamos $|v_1\rangle$, é combinação linear dos outros,

$$|v_1\rangle = b_2 |v_2\rangle + \cdots + b_k |v_k\rangle \quad (7.12)$$

então

$$|v_1\rangle - b_2|v_2\rangle - \cdots - b_k|v_k\rangle = \mathbf{0} \quad (7.13)$$

é uma combinação linear dos vetores que resulta no vetor nulo sem que todas as constantes sejam nulas, o que implica que eles são LD.

☐

Exemplo 73. Um conjunto com dois vetores é LD se, e somente se, um deles é múltiplo do outro. De fato, se dois vetores são LD, um deles pode ser escrito como combinação linear do outro.

Exemplo 74. Em $\mathbb{R}^3$, três vetores são LD apenas em três casos: ou os três são paralelos, ou dois deles são paralelos, ou os três vetores são coplanares.

Exercício 86. Verifique se os conjuntos de vetores abaixo são LI ou LD.

1. $(1, 1)$;
2. $(1, 2), (3, 6)$;
3. $(1, 2), (3, 1)$;
4. $(1, 2), (3, 10), (1, 1)$;
5. $(1, 1, 2), (1, 0, 0), (4, 6, 12)$;
6. $(1, 1, 1), (2, 3, 1), (3, 1, 2)$;
7. $(1, -2, 3), (-2, 4, -6)$;
8. $(4, 2, -1), (6, 5, -5), (2, -1, 3)$.
9. $(1, 1, 2, 0), (1, -1, 0, 1)$;
10. $(1, 1, 0, 2), (1, -1, 0, 0), (1, 0, 1, 0)$;
11. $(1, 1, 0, 0), (1, -2, 0, 0), (1, 0, 1, -1), (0, 0, 1, 1)$;
12. $(1, 1, 0, 0, 2), (1, -1, 0, 0, 0), (1, 0, 1, 0, 0), (0, 0, 1, 2, 0)$;
13. $(1, 1, 0, 0, 3), (1, -1, 0, 0, 0), (2, 0, 1, 0, 0), (0, 0, -1, 1, 0), (0, 0, 0, 0, 1)$.

Exercício 87. Em cada item do exercício anterior em que os vetores são LD, escreva uma combinação linear dos vetores que resulta o vetor nulo em que pelo menos uma das constantes usadas seja diferente de zero. Escreva um dos vetores como combinação linear dos outros.

7.4.1 Subespaços gerados

Definição 57. Dado um conjunto de vetores $|u_1\rangle, \dots, |u_k\rangle$, o *subespaço gerado* por esses vetores é o conjunto de todas as suas combinações lineares:

$$\langle |u_1\rangle, \dots, |u_k\rangle \rangle = \{a_1 |u_1\rangle + \dots + a_k |u_k\rangle; a_i \in \mathbb{R}\}. \quad (7.14)$$

Exercício 88. Prove que $\langle |u_1\rangle, \dots, |u_k\rangle \rangle$ é de fato um subespaço de $\mathbb{R}^n$.

Exemplo 75. O vetor $(7, 2, 9) \in \mathbb{R}^3$ é uma combinação linear dos vetores $(2, 1, 3)$ e $(1, 0, 1)$ porque

$$(7, 2, 9) = 2(2, 1, 3) + 3(1, 0, 1).$$

Portanto $(7, 2, 9) \in \langle (2, 1, 3), (1, 0, 1) \rangle$.

Exemplo 76. O subespaço gerado por um vetor não nulo $|v\rangle \in \mathbb{R}^2$ ou $\mathbb{R}^3$ é uma reta passando pela origem. De fato

$$\langle |v\rangle \rangle = \{t |v\rangle ; t \in \mathbb{R}\}$$

que corresponde justamente à reta passando pela origem na direção do vetor $|v\rangle$.

Exemplo 77. O subespaço gerado pelos vetores $(1, 0, 0)$ e $(0, 1, 0)$ em $\mathbb{R}^3$ corresponde ao plano xy . De fato, as combinações lineares de $(1, 0, 0)$ e $(0, 1, 0)$ são vetores da forma

$$a_1(1, 0, 0) + a_2(0, 1, 0) = (a_1, a_2, 0)$$

que correspondem justamente aos vetores pertencentes ao plano xy

De forma análoga, o subespaço gerado pelos vetores $(1, 0, 0)$ e $(0, 0, 1)$ corresponde ao plano xz e o subespaço gerado pelos vetores $(0, 1, 0)$ e $(0, 0, 1)$ corresponde ao plano yz .

Exemplo 78. O subespaço gerado por dois vetores $|u\rangle$ e $|v\rangle$ não paralelos em $\mathbb{R}^3$ é um plano passando pela origem que contém os vetores $|u\rangle$ e $|v\rangle$.

7.4.2 Base e Dimensão

Podemos nos perguntar se existe um conjunto de vetores LI de forma que todo elemento do espaço $\mathbb{R}^n$ possa ser escrito como combinação linear dos elementos desse conjunto. É possível mostrar que o espaço $\mathbb{R}^n$ possui um conjunto LI com essa propriedade para todo $n \in \mathbb{N}$.

Definição 58. Uma base para $\mathbb{R}^n$ é um conjunto LI

$$\mathcal{B} = \{|u_1\rangle, \dots, |u_k\rangle\} \quad (7.15)$$

tal que todo vetor de $\mathbb{R}^n$ é combinação linear de $|u_1\rangle, \dots, |u_k\rangle$, ou seja, dado qualquer vetor $|v\rangle \in \mathbb{R}^n$, existem constantes $a_1, \dots, a_k \in \mathbb{R}$ tais que

$$|v\rangle = a_1 |u_1\rangle + a_2 |u_2\rangle + \dots + a_k |u_k\rangle. \quad (7.16)$$

De forma equivalente, dizemos que $\mathcal{B} = \{|u_1\rangle, \dots, |u_k\rangle\}$ é uma base para $\mathbb{R}^n$ se $\mathcal{B}$ é um conjunto LI e

$$\langle |u_1\rangle, \dots, |u_k\rangle \rangle = \mathbb{R}^n. \quad (7.17)$$

Exercício 89. Os vetores $|u_1\rangle = (1, 0)$ e $|u_2\rangle = (0, 1)$ formam uma base para $\mathbb{R}^2$, como vimos no Exemplo 19. De maneira análoga, como mostrado no exemplo 67, os vetores $|u_1\rangle = (1, 0, 0)$, $|u_2\rangle = (0, 1, 0)$ e $|u_3\rangle = (0, 0, 1)$ formam uma base para $\mathbb{R}^3$. De maneira geral, os vetores

$$|u_1\rangle = (1, 0, 0, \dots, 0, 0), |u_2\rangle = (0, 1, 0, \dots, 0, 0), \dots, |u_n\rangle = (0, 0, 0, \dots, 0, 1) \quad (7.18)$$

formam uma base para $\mathbb{R}^n$.

Em geral, não é difícil provar que um conjunto de vetores é LI, pois isso é equivalente a resolver um sistema linear. Com o que temos até agora, o difícil é provar se um conjunto de vetores gera ou não todo o espaço $\mathbb{R}^n$. Antes de vermos mais exemplos, vamos provar alguns resultados que facilitam essa tarefa.

Teorema 28. *Se $\{|u_1\rangle, \dots, |u_k\rangle\}$ é uma base para o espaço $\mathbb{R}^n$, então qualquer conjunto com mais de k vetores é LD.*

Demonstração. Seja $\{|v_1\rangle, \dots, |v_m\rangle\}$ um conjunto qualquer com $m > k$. Seja

$$b_1 |v_1\rangle + b_2 |v_2\rangle + \dots + b_m |v_m\rangle = \mathbf{0} \quad (7.19)$$

uma combinação linear desses vetores que gera o vetor nulo. Nosso objetivo é provar que é possível escolher os valores das constantes b_i na equação acima sem que todas elas sejam nulas.

Utilizando o fato de que o conjunto $\{|u_1\rangle, \dots, |u_k\rangle\}$ é uma base, vamos reescrever a equação acima em termos dos vetores $|u_i\rangle$. Isso é possível porque todo vetor $|v_j\rangle$ pode ser escrito como combinação linear de $\{|u_1\rangle, \dots, |u_k\rangle\}$:

$$\begin{aligned} |v_1\rangle &= x_{11} |u_1\rangle + x_{12} |u_2\rangle + \dots + x_{1k} |u_k\rangle \\ |v_2\rangle &= x_{21} |u_1\rangle + x_{22} |u_2\rangle + \dots + x_{2k} |u_k\rangle \\ &\vdots \\ |v_m\rangle &= x_{m1} |u_1\rangle + x_{m2} |u_2\rangle + \dots + x_{mk} |u_k\rangle \end{aligned} \quad (7.20)$$

Substituindo as igualdades acima na equação (7.19), temos

$$\begin{aligned} &b_1 (x_{11} |u_1\rangle + x_{12} |u_2\rangle + \dots + x_{1k} |u_k\rangle) \\ &+ b_2 (x_{21} |u_1\rangle + x_{22} |u_2\rangle + \dots + x_{2k} |u_k\rangle) \\ &+ \dots + b_m (x_{m1} |u_1\rangle + x_{m2} |u_2\rangle + \dots + x_{mk} |u_k\rangle) = \mathbf{0} \end{aligned} \quad (7.21)$$

Colocando em evidência os vetores $|u_i\rangle$ na equação acima chegamos à

$$\begin{aligned} &(b_1 x_{11} + b_2 x_{21} + \dots + b_m x_{m1}) |u_1\rangle \\ &+ (b_1 x_{12} + b_2 x_{22} + \dots + b_m x_{m2}) |u_2\rangle \\ &+ \dots + (b_1 x_{1k} + b_2 x_{2k} + \dots + b_m x_{mk}) |u_k\rangle = \mathbf{0}. \end{aligned} \quad (7.22)$$

A equação acima fornece uma combinação linear dos vetores $|u_1\rangle, \dots, |u_k\rangle$ que resulta no vetor nulo. Como esses vetores são LI, isso só pode acontecer se todas as constantes forem zero, o que implica que

$$\begin{aligned} b_1 x_{11} + b_2 x_{21} + \dots + b_m x_{m1} &= 0 \\ b_1 x_{12} + b_2 x_{22} + \dots + b_m x_{m2} &= 0 \\ &\vdots \\ b_1 x_{1k} + b_2 x_{2k} + \dots + b_m x_{mk} &= 0. \end{aligned}$$

Esse é um sistema de equações nas variáveis b_i que possui k equações e m incógnitas, ou seja, mais incógnitas que equações, e portanto possui solução não-trivial (veja as referências [4, 5, 6, 7]). Isso quer dizer que é possível encontrar constantes $b_1, \dots, b_m$, não todas nulas, tais que a equação (7.19) seja verdadeira. Isso implica que os vetores $\{|v_1\rangle, \dots, |v_m\rangle\}$ são LD. $\square$

O teorema acima permite provar que duas bases de $\mathbb{R}^n$ possuem o mesmo número de elementos. De fato, suponhamos que $\{|u_1\rangle, \dots, |u_k\rangle\}$ e $\{|v_1\rangle, \dots, |v_m\rangle\}$ sejam bases para $\mathbb{R}^n$. Se $m > k$, o teorema anterior implica que o conjunto $\{|v_1\rangle, \dots, |v_m\rangle\}$ é LD, e portanto não poderia ser uma base, o que é uma contradição. Se $m < k$, o teorema anterior implica que $\{|u_1\rangle, \dots, |u_k\rangle\}$ é LD, o que também é uma contradição. Logo $k = m$.

Definição 59. A *dimensão* de $\mathbb{R}^n$ é o número de vetores em uma base. O espaço vetorial $\mathbb{R}^2$ tem dimensão 2; O espaço vetorial $\mathbb{R}^3$ tem dimensão 3; Em geral, o espaço vetorial $\mathbb{R}^n$ tem dimensão n .

Teorema 29. Em $\mathbb{R}^n$ qualquer conjunto LI $\{|u_1\rangle, \dots, |u_n\rangle\}$ com n vetores é uma base para $\mathbb{R}^n$.

Demonstração. Para provar que um conjunto é uma base para $\mathbb{R}^n$, devemos provar que, além de ser LI, esse conjunto gera $\mathbb{R}^n$, ou seja, qualquer vetor $|v\rangle \in \mathbb{R}^n$ pode ser escrito como combinação linear de $\{|u_1\rangle, \dots, |u_n\rangle\}$.

Sabemos pelo Teorema 28 que o conjunto $\{|u_1\rangle, \dots, |u_n\rangle, |v\rangle\}$ é LD. Isso quer dizer que existe uma combinação linear

$$a_1 |u_1\rangle + a_2 |u_2\rangle + \dots + a_n |u_n\rangle + b |v\rangle = \mathbf{0} \quad (7.23)$$

que gera o vetor nulo em que pelo menos uma das constantes é não nula. Afirmamos que $b \neq 0$. Se b fosse igual a 0, poderíamos eliminar o vetor $|v\rangle$ da equação acima e teríamos

$$a_1 |u_1\rangle + a_2 |u_2\rangle + \dots + a_n |u_n\rangle = \mathbf{0} \quad (7.24)$$

em que pelo menos uma das constantes é não nula. Como $\{|u_1\rangle, \dots, |u_n\rangle\}$ é uma base, é, em particular LI, e portanto isso não pode acontecer. Logo concluímos que $b \neq 0$.

Podemos então isolar $|v\rangle$ na equação (7.23), o que nos leva a

$$|v\rangle = -\frac{a_1}{b} |u_1\rangle - \frac{a_2}{b} |u_2\rangle - \dots - \frac{a_n}{b} |u_n\rangle. \quad (7.25)$$

Isso mostra que, independente de quem seja o vetor $|v\rangle$, ele pode ser escrito como combinação linear de $\{|u_1\rangle, \dots, |u_n\rangle\}$. Logo esse conjunto é uma base para $\mathbb{R}^n$. $\square$

Assim, para verificar se um conjunto de vetores é ou não uma base para $\mathbb{R}^n$, basta verificarmos se ele possui o número correto de vetores e se esses vetores são LI. O Teorema 29 garante que nessas condições, o conjunto gera $\mathbb{R}^n$ e portanto é uma base. Podemos resumir esse resultado da seguinte forma:

Uma base para $\mathbb{R}^n$ é um conjunto LI com exatamente n vetores.

Exercício 90. Verifique se os vetores abaixo formam uma base para o espaço ao qual pertencem.

1. $(1, 1)$;
2. $(1, 2), (1, 0)$;
3. $(1, 2), (1, 0), (0, 1)$;
4. $(1, 0, 0), (0, 1, 1)$;
5. $(1, 0, 0), (0, 1, -1), (0, 1, 1)$;
6. $(0, 1, 0), (0, 1, -1), (0, 1, 1)$;
7. $(2, 1, -1), (3, 1, -2), (1, 0, -1), (1, 0, 1)$;
8. $(3, 1, -2), (1, 0, -1), (1, 0, 1)$;
9. $(2, 1, -1), (3, 1, -2), (1, 0, -1)$;
10. $(2, 1, -1), (3, 1, -2), (1, 0, 2)$;
11. $(1, 1, 0, 0), (1, -1, 0, 0)$;
12. $(1, 1, 0, 0), (1, -1, 0, 0), (1, 0, 1, 0)$;
13. $(1, 1, 0, 0), (1, -1, 0, 0), (1, 0, 1, 0), (0, 0, 1, 1)$;
14. $(1, 1, 0, 0, 1), (1, -1, 0, 0, 0), (1, 0, 1, 0, 0), (0, 0, 1, 1, 0)$;
15. $(1, 1, 0, 0, 1), (1, -1, 0, 0, 0), (1, 0, 1, 0, 0), (0, 0, 1, 1, 0), (0, 0, 0, 0, 1)$.

Exercício 91. Dados $|u_1\rangle = (2, 1, 3)$ e $|u_2\rangle = (2, 6, 4)$, responda:

1. O conjunto $\{|v_1\rangle, |v_2\rangle\}$ é uma base para $\mathbb{R}^3$? Descreva geometricamente o subespaço gerado por esses vetores.
2. Quais são as condições sobre $|v_3\rangle$ para que $\{|v_1\rangle, |v_2\rangle, |v_3\rangle\}$ seja uma base de $\mathbb{R}^3$?
3. Encontre um vetor $|v_3\rangle$ de modo que $\{|v_1\rangle, |v_2\rangle, |v_3\rangle\}$ seja uma base de $\mathbb{R}^3$.

7.5 Produto Interno

Definição 60. Dados dois vetores $|u\rangle = (x_1, \dots, x_n)$ e $|v\rangle = (y_1, \dots, y_n)$ em $\mathbb{R}^n$ definimos o *produto interno* entre $|u\rangle$ e $|v\rangle$ por

$$\langle u | v \rangle = (x_1, \dots, x_n) \cdot (y_1, \dots, y_n) = \sum_{i=1}^n x_i y_i. \quad (7.26)$$

Exemplo 79. O produto interno entre os vetores $(1, 3, 5)$ e $(4, 2, 1)$ em $\mathbb{R}^3$ é igual a

$$(1, 3, -5) \cdot (4, -2, -1) = (1)(4) + (3)(-2) + (-5)(-1) = 4 - 6 + 5 = 3.$$

Teorema 30. O produto interno em $\mathbb{R}^n$ satisfaz as seguintes propriedades:

1. $(\lambda |u\rangle + \mu |v\rangle) \cdot |w\rangle = \lambda (\langle u | w \rangle) + \mu (\langle v | w \rangle);$
2. $\langle u | v \rangle = \langle v | u \rangle;$
3. $\langle u | u \rangle \geq 0;$
4. Se $\langle u | u \rangle = 0$ então $|u\rangle = 0$.

Definição 61. A partir do produto interno podemos definir a *norma* de um vetor $|v\rangle$ escrevendo

$$\|v\| = \sqrt{\langle v | v \rangle}.$$

O produto interno nos permite introduzir uma noção que generaliza a um espaço $\mathbb{R}^n$ a ideia de perpendicularidade nos espaços $\mathbb{R}^2$ e $\mathbb{R}^3$, com a qual já estamos familiarizados.

Definição 62. Dizemos que dois vetores $|u\rangle$ e $|v\rangle$ são *ortogonais* se $\langle u | v \rangle = 0$. Dizemos que um conjunto $E = \{|v_1\rangle, \dots, |v_k\rangle\}$ é *ortogonal* se seus elementos são dois a dois ortogonais. Dizemos que um conjunto $E = \{|v_1\rangle, \dots, |v_k\rangle\}$ é *ortonormal* se é ortogonal e $\langle v_i | v_i \rangle = 1$ para todo i .

Exercício 92. Mostre que se u e v são ortogonais, então

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

Teorema 31. Se $\{|v_1\rangle, \dots, |v_k\rangle\}$ é um conjunto ortogonal, então valem as seguintes propriedades:

1. O conjunto $\{|v_1\rangle, \dots, |v_k\rangle\}$ é LI;
2. Se $|v\rangle = a_1 |v_1\rangle + \dots + a_k |v_k\rangle$, então

$$a_i = \frac{\langle v | v_i \rangle}{\| |v_i\rangle \|^2}.$$

Demonstração.

1. Suponhamos que

$$\lambda_1 |v_1\rangle + \lambda_2 |v_2\rangle + \dots + \lambda_k |v_k\rangle = \mathbf{0}. \quad (7.27)$$

Fazendo o produto interno por $|v_i\rangle$ em ambos os lados da equação acima, temos que

$$\begin{aligned} (\lambda_1 |v_1\rangle + \lambda_2 |v_2\rangle + \dots + \lambda_k |v_k\rangle) \cdot |v_i\rangle &= \mathbf{0} \cdot |v_i\rangle \\ \lambda_1 (|v_1\rangle \cdot |v_i\rangle) + \lambda_2 (|v_2\rangle \cdot |v_i\rangle) + \dots + \lambda_k (|v_k\rangle \cdot |v_i\rangle) &= 0 \\ \lambda_i (|v_i\rangle \cdot |v_i\rangle) &= 0 \\ \lambda_i \| |v_i\rangle \|^2 &= 0. \end{aligned} \quad (7.28)$$

Como $\| |v_i\rangle \| \neq 0$, a equação acima implica que $\lambda_i = 0$. Logo a equação (7.29) só pode ser verdadeira se todas as constantes λ_i são nulas e portanto o conjunto $\{|v_1\rangle, \dots, |v_k\rangle\}$ é LI.

2. Suponhamos que

$$|v\rangle = a_1 |v_1\rangle + a_2 |v_2\rangle + \dots + a_k |v_k\rangle. \quad (7.29)$$

Para provar que o conjunto $\{|v_1\rangle, \dots, |v_k\rangle\}$ é LI, devemos mostrar que a única possibilidade para as constantes a_i na equação acima é que todas elas sejam nulas.

Fazendo o produto interno por $|v_i\rangle$ em ambos os lados da equação acima, temos que

$$\begin{aligned} (a_1 |v_1\rangle + a_2 |v_2\rangle + \dots + a_k |v_k\rangle) \cdot |v_i\rangle &= |v\rangle \cdot |v_i\rangle \\ a_1 (|v_1\rangle \cdot |v_i\rangle) + a_2 (|v_2\rangle \cdot |v_i\rangle) + \dots + a_k (|v_k\rangle \cdot |v_i\rangle) &= |v\rangle \cdot |v_i\rangle \\ a_i (|v_i\rangle \cdot |v_i\rangle) &= |v\rangle \cdot |v_i\rangle \\ a_i \| |v_i\rangle \|^2 &= |v\rangle \cdot |v_i\rangle \end{aligned} \quad (7.30)$$

A equação acima implica que $a_i = \frac{|v\rangle \cdot |v_i\rangle}{\| |v_i\rangle \|^2}$, como queríamos provar.

□

Definição 63. Dizemos que uma base $\mathcal{B} = \{|v_1\rangle, \dots, |v_k\rangle\}$ para $\mathbb{R}^n$ é uma *base ortogonal* se seus elementos são dois a dois ortogonais. Dizemos uma base $\mathcal{B} = \{|v_1\rangle, \dots, |v_k\rangle\}$ é uma *base ortonormal* se é ortogonal e $\langle v_i | v_i \rangle = 1$ para todo i .

Exemplo 80. A base

$$\mathcal{B} = \{|e_1\rangle = (1, 0, \dots, 0), |e_2\rangle = (0, 1, \dots, 0), \dots, |e_n\rangle = (0, 0, \dots, 1)\} \quad (7.31)$$

de $\mathbb{R}^n$ é uma base ortonormal. De fato, é fácil mostrar que o produto interno $\langle e_i | e_j \rangle$ é igual a zero se $i \neq j$ e igual a 1 se $i = j$.

Exemplo 81. Mostre que o conjunto

$$\mathcal{B} = \left\{ |v_1\rangle = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right), |v_2\rangle = \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}} \right), |v_3\rangle = (0, 1, 0) \right\}$$

é uma base ortonormal de $\mathbb{R}^3$.

Solução. Calculando os produtos internos temos

$$\langle v_1 | v_2 \rangle = \langle v_1 | v_3 \rangle = \langle v_2 | v_3 \rangle = 0, \quad \langle v_1 | v_1 \rangle = \langle v_2 | v_2 \rangle = \langle v_3 | v_3 \rangle = 1$$

o que mostra que $\mathcal{B}$ é um conjunto ortonormal. Pelo teorema 31, $\mathcal{B}$ é um conjunto LI e portanto é uma base, já que possui três vetores. $\square$

Outra consequência do Teorema 31 é que dado um vetor qualquer $|v\rangle$ e uma base ortogonal $\mathcal{B} = \{|v_1\rangle, \dots, |v_k\rangle\}$ de $\mathbb{R}^n$, é fácil encontrar as constantes que devemos usar para escrever $|v\rangle$ como combinação linear de $|v_1\rangle, \dots, |v_k\rangle$.

Exercício 93. Escreva $|v\rangle = (2, 1, -1)$ como combinação linear dos vetores

$$|v_1\rangle = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right), |v_2\rangle = \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right), |v_3\rangle = (0, 1, 0).$$

Solução. Pelo Teorema 31, sabemos que se

$$|v\rangle = a_1 |v_1\rangle + a_2 |v_2\rangle + a_3 |v_3\rangle$$

, então cada constante a_i é dada por

$$a_i = \frac{\langle v \mid v_i \rangle}{\| |v_i\rangle \|^2} = \langle v \mid v_i \rangle.$$

Observe que a expressão se simplifica ainda mais, já que os vetores tem norma igual a 1. Calculando os produtos internos, temos

$$\begin{aligned} a_1 &= \langle v \mid v_1 \rangle = \frac{2}{\sqrt{2}} + 0 - \frac{1}{\sqrt{2}} = \frac{1}{\sqrt{2}} \\ a_2 &= \langle v \mid v_2 \rangle = \frac{2}{\sqrt{2}} + 0 + \frac{1}{\sqrt{2}} = \frac{3}{\sqrt{2}} \\ a_3 &= \langle v \mid v_3 \rangle = 0 + 1 + 0. \end{aligned} \quad (7.32)$$

Temos então que

$$|v\rangle = \frac{1}{\sqrt{2}}|v_1\rangle + \frac{3}{\sqrt{2}}|v_2\rangle + |v_3\rangle.$$

☐

Exercício 94. Verifique se os conjuntos de vetores abaixo são ortogonais. Verifique quais deles são ortonormais. Verifique quais deles constituem bases de $\mathbb{R}^3$.

1. $\{(1, 2, -1), (1, 0, 1)\}$;
2. $\left\{\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{-1}{\sqrt{3}}\right), \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)\right\}$;
3. $\{(1, 2, 2), (1, 2, -1), (0, 0, 1)\}$;
4. $\left\{\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(\frac{-1}{\sqrt{6}}, \frac{-1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\right), \left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, 0\right)\right\}$;

Exercício 95. Considere os conjuntos acima que são bases de $\mathbb{R}^3$. Escreva o vetor $(1, 1, 1)$ como combinação linear dos elementos dessas bases.

Exercício 96. Sejam $|v\rangle = (1, 1, -2)$ e $|u\rangle = (a, -1, 2)$. Para quais valores da constante a os vetores $|v\rangle$ e $|u\rangle$ são ortogonais?

Exercício 97. Sejam $|v\rangle = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)$ e $|u\rangle = \left(a, \frac{1}{\sqrt{3}}, b\right)$. Para quais valores de a e b o conjunto $\{|u\rangle, |v\rangle\}$ é ortonormal?

Exercício 98. Mostre que se $|u\rangle$ é ortogonal a $|v\rangle$ então $\lambda |u\rangle$ é ortogonal a $|v\rangle$ para todo valor de $\lambda \in \mathbb{R}$.

Exercício 99. Mostre que se $|v\rangle$ é ortogonal a $|v_1\rangle, \dots, |v_k\rangle$ então $|v\rangle$ é ortogonal a qualquer combinação linear de $|v_1\rangle, \dots, |v_k\rangle$.

7.6 O Espaço $\mathbb{C}^n$

Vamos agora modificar o que vimos para espaços vetoriais reais substituindo o conjunto numérico $\mathbb{R}$ por $\mathbb{C}$. Veremos que todas as definições funcionam de maneira idêntica, com exceção do produto interno que deve ser ligeiramente modificado.

7.7 Definição e Propriedades

Definição 64. Representamos por $\mathbb{C}^n$ o conjunto de todas as n -úplas de números complexos, que denotaremos por

$$|u\rangle = (z_1, z_2, \dots, z_n), z_i \in \mathbb{C} \quad (7.33)$$

Definição 65. Dados $|u\rangle = (z_1, z_2, \dots, z_n)$ e $|v\rangle = (w_1, \dots, w_n)$ em $\mathbb{C}^n$ definimos a soma

$$|u\rangle + |v\rangle = (z_1 + w_1, \dots, z_n + w_n) \quad (7.34)$$

e o produto de $|u\rangle$ por um escalar $\lambda \in \mathbb{C}$

$$\lambda |u\rangle = (\lambda z_1, \dots, \lambda z_n). \quad (7.35)$$

7.8. COMBINAÇÕES LINEARES

99

Assim como fizemos para $\mathbb{C}^2$, será útil posteriormente identificar vetores em $\mathbb{C}^n$ e matrizes coluna $n \times 1$:

$$(z_1, \dots, z_n) \leftrightarrow \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}. \quad (7.36)$$

Teorema 32. Para todos $|u\rangle, |v\rangle, |w\rangle \in \mathbb{C}^n$ e $\lambda, \nu \in \mathbb{C}$ temos

1. (Associatividade) $|u\rangle + (|v\rangle + |w\rangle) = (|u\rangle + |v\rangle) + |w\rangle$;
2. (Comutatividade) $|u\rangle + |v\rangle = |v\rangle + |u\rangle$;
3. (Existência de zero) O vetor $\mathbf{0} = (0, 0, \dots, 0)$ é tal que $|u\rangle + \mathbf{0} = |u\rangle$;
4. (Existência de inverso aditivo) Dado $|u\rangle$, o vetor $-|u\rangle$ é tal que $|u\rangle + (-|u\rangle) = \mathbf{0}$;
5. (Associatividade) $\lambda(\nu |u\rangle) = (\lambda\nu) |u\rangle$;
6. (Distributividade) $\lambda(|u\rangle + |v\rangle) = \lambda |u\rangle + \lambda |v\rangle$;
7. (Distributividade) $(\lambda + \nu) |u\rangle = \lambda |u\rangle + \nu |u\rangle$;
8. $1 |u\rangle = |u\rangle$.

A partir de $n = 2$ os conjuntos $\mathbb{C}^n$ precisam de pelo menos quatro eixos reais para serem representados e portanto nesse caso perdemos a interpretação geométrica que temos para os casos $\mathbb{R}^2$ e $\mathbb{R}^3$.

7.8 Combinações Lineares

Em $\mathbb{C}^n$ os elementos podem ser combinados livremente através da soma e da multiplicação por escalar.

Definição 66. Dizemos que $|v\rangle$ é uma *combinação linear* de $|u_1\rangle, \dots, |u_k\rangle$ se existem constantes $a_1, \dots, a_k \in \mathbb{C}$ tais que

$$|v\rangle = a_1 |u_1\rangle + \dots + a_k |u_k\rangle.$$

Exemplo 82. O vetor $(3i, 0, 2 + i) \in \mathbb{C}^3$ é uma combinação linear dos vetores $(i, 0, 1)$ e $(1, 0, 1)$ porque

$$(3i, 0, 2 + i) = 2(i, 0, 1) + i(1, 0, 1).$$

Exemplo 83. O vetor nulo $\mathbf{0}$ é sempre combinação linear de qualquer conjunto de vetores $|v_1\rangle, \dots, |v_k\rangle$. De fato,

$$\mathbf{0} = 0 |v_1\rangle + \dots + 0 |v_k\rangle.$$

Exemplo 84. Considere os vetores $|e_1\rangle = (1, 0, 0)$, $|e_2\rangle = (0, 1, 0)$ e $|e_3\rangle = (0, 0, 1)$. Qualquer vetor em $\mathbb{C}^3$ é combinação linear desses três vetores. De fato, dado $|v\rangle = (a_1, a_2, a_3)$ podemos escrever

$$\begin{aligned}(a_1, a_2, a_3) &= (a_1, 0, 0) + (0, a_2, 0) + (0, 0, a_3) \\ &= a_1(1, 0, 0) + a_2(0, 1, 0) + a_3(0, 0, 1) \\ &= a_1|e_1\rangle + a_2|e_2\rangle + a_3|e_3\rangle.\end{aligned}\tag{7.37}$$

De maneira geral, qualquer vetor de $\mathbb{C}^n$ pode ser escrito como uma combinação linear dos vetores

$$|e_1\rangle = (1, 0, \dots, 0), |e_2\rangle = (0, 1, \dots, 0), \dots, |e_n\rangle = (0, 0, \dots, 1).\tag{7.38}$$

Exercício 100. Se um vetor $|v\rangle$ é combinação linear de um vetor $|u\rangle$ então

$$|v\rangle = \lambda |u\rangle$$

ou seja, $|v\rangle$ é múltiplo de $|u\rangle$.

Assim como acontece em $\mathbb{R}^n$, dados vetores $|u_1\rangle, \dots, |u_m\rangle \in \mathbb{C}^n$, um vetor $|v\rangle$ será combinação linear de $|u_1\rangle, \dots, |u_m\rangle$ se existirem constantes $a_1, \dots, a_m$ tais que

$$|v\rangle = a_1 |u_1\rangle + \dots + a_m |u_m\rangle.$$

Essa equação nos levará a um sistema de equações lineares para $a_1, \dots, a_m$.

Exercício 101. Verifique se o vetor $|v\rangle$ é combinação linear dos vetores $|u_1\rangle, \dots, |u_n\rangle$ dados abaixo. Em caso afirmativo, encontre a combinação linear desses vetores que gera $|v\rangle$.

1. $|v\rangle = (1, i); |u_1\rangle = (i, 0);$
2. $|v\rangle = (2, 2); |u_1\rangle = (i, i);$
3. $|v\rangle = (2i, 3); |u_1\rangle = (i, 0), |u_2\rangle = (0, 1 + i);$
4. $|v\rangle = (2, 3i); |u_1\rangle = (i, 0), |u_2\rangle = (1, 1);$
5. $|v\rangle = (i, i, i); |u_1\rangle = (1, 0, 0), |u_2\rangle = (0, -i, 0);$
6. $|v\rangle = (2, 2, 2); |u_1\rangle = (i, 0, i), |u_2\rangle = (0, 1, 0);$
7. $|v\rangle = (2, 3 - i, 1); |u_1\rangle = (1, 0, 0), |u_2\rangle = (0, 1 - i, 0), |u_3\rangle = (0, 0, 1);$
8. $|v\rangle = (2 - i, 3, i); |u_1\rangle = (1, i, 0), |u_2\rangle = (0, 1, 0), |u_3\rangle = (0, 1, i);$
9. $|v\rangle = (2, 3, 1, -1); |u_1\rangle = (1, 1, 0, 0), |u_2\rangle = (0, i, 0, 0), |u_3\rangle = (0, 0, -i, 0), |u_4\rangle = (0, 0, 1, 1), |u_5\rangle = (1, 1, i, i).$

7.9 Subespaços Vetoriais

Definição 67. Um *subespaço vetorial* S do espaço vetorial $\mathbb{C}^n$ é um subconjunto de $\mathbb{C}^n$ que é preservado pelas operações de soma e multiplicação por escalar definidas em $\mathbb{C}^n$. Para isso, precisamos que as seguintes propriedades sejam satisfeitas:

1. $|u\rangle + |v\rangle \in S$ para todo par $|u\rangle$ e $|v\rangle \in S$;
2. $\lambda |u\rangle \in S$ para todo $\lambda \in \mathbb{C}$ e todo $|v\rangle \in S$.

Para mostrar que S é um subespaço vetorial de $\mathbb{C}^n$ devemos verificar se as operações de soma e multiplicação por escalar *preservam* S : ao somar dois elementos de S devemos permanecer em S e ao multiplicar um elemento de S por um escalar devemos permanecer em S .

Exemplo 85. O subconjunto

$$S = \{(t, 0, \dots, 0) \in \mathbb{C}^n; t \in \mathbb{C}\}$$

é um subespaço vetorial de $\mathbb{C}^n$.

Solução. Novamente, a solução do problema passa por três passos. O primeiro deles é verificar qual é a propriedade que define o subconjunto S . Nesse exemplo, o subconjunto S é o conjunto de todos os vetores de $\mathbb{C}^n$ que possuem todas as coordenadas iguais a zero a partir da segunda. A primeira coordenada pode assumir qualquer valor.

O segundo passo é verificar se a soma de quaisquer elementos de S permanece em S . Suponhamos que $|u\rangle$ e $|v\rangle \in S$. Então $|u\rangle = (t_1, 0, \dots, 0)$ e $|v\rangle = (t_2, 0, \dots, 0)$ com t_1 e $t_2 \in \mathbb{C}$. Assim

$$|u\rangle + |v\rangle = (t_1 + t_2, 0, \dots, 0)$$

também pertence a S .

O terceiro passo é verificar se a multiplicação de elementos de S por constante permanece em S . Tomando $|u\rangle = (t_1, 0, \dots, 0) \in S$ qualquer e $\lambda \in \mathbb{C}$ temos

$$\lambda |u\rangle = (\lambda t_1, 0, \dots, 0)$$

que também pertence a S . Como as propriedades 1 e 2 são satisfeitas, S é um subespaço de $\mathbb{C}^n$. □

Exercício 102. Verifique quais dentre os subconjuntos de $\mathbb{C}^3$ abaixo são subespaços vetoriais.

1. O subconjunto de vetores que possuem primeira coordenada igual a zero;
2. O subconjunto de vetores que possuem primeira coordenada igual a i ;
3. O subconjunto de vetores (x, y, z) tais que $xy = 0$;
4. O subconjunto de vetores (x, y, z) tais que $z - iy + 3x = 0$.

7.10 Dependência e Independência linear

Queremos encontrar condições que informem se dentre um conjunto de vetores, algum deles pode ser escrito como combinação linear dos outros. Para isso definimos a noção de dependência linear.

Definição 68. Dizemos que um conjunto de vetores $\{|u_1\rangle, \dots, |u_k\rangle\} \subset \mathbb{C}^n$ é *linearmente independente* (LI) se a equação

$$a_1 |u_1\rangle + \cdots + a_k |u_k\rangle = \mathbf{0} \quad (7.39)$$

só admite a solução trivial $a_1 = \dots = a_k = 0$, ou seja, se a única combinação linear dos vetores que gera o vetor nulo é aquela em que todas as constantes são nulas.

Caso contrário, dizemos que os vetores são *linearmente dependentes* (LD). Um conjunto de vetores é LD se existe uma combinação linear deles que gera o vetor nulo em que pelo menos uma das constantes é não-nula.

Exemplo 86. Um conjunto de vetores que contém o vetor nulo é sempre LD. De fato, tomando o conjunto $\{|u_1\rangle, \dots, |u_k\rangle, \mathbf{0}\}$ podemos escrever a combinação linear

$$0 \times |u_1\rangle + \cdots + 0 \times |u_k\rangle + 1 \times 0 = \mathbf{0}. \quad (7.40)$$

Temos então uma combinação linear dos vetores que gera o vetor nulo em que a última constante utilizada é diferente de zero. Isso implica que esses vetores são LD.

Exemplo 87. Um conjunto formado por um único vetor não nulo $|u\rangle$ é sempre LI. Nesse caso, as únicas combinações lineares possíveis são múltiplos de $|u\rangle$ e para que um desses múltiplos seja o vetor nulo temos que

$$\lambda|u\rangle = \mathbf{0} \quad (7.41)$$

o que implica que $\lambda = 0$ uma vez que $|u\rangle \neq 0$.

Três ou mais vetores em $\mathbb{C}^2$, assim como quatro ou mais vetores em $\mathbb{C}^3$, assim como mais de n vetores em $\mathbb{C}^n$ são sempre LD. Isso acontece porque o problema de identificar se esses vetores são ou não LI leva a um sistema com mais incógnitas que equações que sempre possui solução não trivial.

Teorema 33. *O número máximo de vetores LI em $\mathbb{C}^n$ é igual a n .*

A definição de dependência linear é importante porque ela permite identificar se dentro de um conjunto de vetores, um deles pode ser escrito como combinação linear dos outros ou não. De fato, as duas propriedades são equivalentes.

Teorema 34. *Um conjunto de vetores é LD se e somente se podemos expressar ao menos um dos vetores como combinação linear dos outros.*

Exemplo 88. Um conjunto com dois vetores é LD se, e somente se, um deles é múltiplo do outro. De fato, se dois vetores são LD, um deles pode ser escrito como combinação linear do outro.

Exercício 103. Verifique se os conjuntos de vetores abaixo são LI ou LD.

1. $(1, i)$;
2. $(1, 2 + i), (3, 6 + 3i)$;
3. $(1, 2), (3 - i, 1)$;
4. $(1, 2), (3, 1), (i, i)$;
5. $(1, 1, 2), (1, 0, 0), (4i, 6i, 12i)$;
6. $(i, 1, i), (2, 3, 1), (3, 1 - i, 2)$;
7. $(1, -2, 3), (-2i, 4i, -6i)$;
8. $(1, 1, 2, 0), (1, -i, 0, i)$;
9. $(1, 1, 0, 2), (i, -i, 0, 0), (1, 0, 1, 0)$;

Exercício 104. Em cada item do exercício anterior em que os vetores são LD, escreva uma combinação linear dos vetores que resulta o vetor nulo em que pelo menos uma das constantes usadas seja diferente de zero. Escreva um dos vetores como combinação linear dos outros.

7.10.1 Subespaços gerados

Definição 69. Dado um conjunto de vetores $|u_1\rangle, \dots, |u_k\rangle$ em $\mathbb{C}^n$, o *subespaço gerado* por esses vetores é o conjunto de todas as suas combinações lineares:

$$\langle |u_1\rangle, \dots, |u_k\rangle \rangle = \{a_1 |u_1\rangle + \dots + a_k |u_k\rangle; a_i \in \mathbb{C}\}.$$

Exercício 105. Prove que $\langle |u_1\rangle, \dots, |u_k\rangle \rangle$ é de fato um subespaço de $\mathbb{C}^n$.

Exemplo 89. O vetor $(i, 2, 2 + i) \in \mathbb{C}^3$ é uma combinação linear dos vetores $(i, 0, 2)$ e $(0, 2, i)$ porque

$$(i, 2, 2 + i) = (i, 0, 2) + (0, 2, i). \quad (7.42)$$

Portanto $(i, 2, 2 + i) \in \langle (i, 0, 2), (0, 2, i) \rangle$.

7.10.2 Base e Dimensão

Podemos nos perguntar se existe um conjunto de vetores LI de forma que todo elemento do espaço $\mathbb{C}^n$ possa ser escrito como combinação linear dos elementos desse conjunto.

Definição 70. Uma base para $\mathbb{C}^n$ é um conjunto LI

$$\mathcal{B} = \{|u_1\rangle, \dots, |u_k\rangle\}$$

tal que todo vetor de $\mathbb{C}^n$ é combinação linear de $|u_1\rangle, \dots, |u_k\rangle$, ou seja, dado qualquer vetor $|v\rangle \in \mathbb{C}^n$, existem constantes $a_1, \dots, a_k \in \mathbb{C}$ tais que

$$|v\rangle = a_1 |u_1\rangle + a_2 |u_2\rangle + \dots + a_k |u_k\rangle. \quad (7.43)$$

Exemplo 90. Os vetores $|u_1\rangle = (1, 0)$ e $|u_2\rangle = (0, 1)$ formam uma base para $\mathbb{C}^2$, como vimos no Exemplo 23. De maneira análoga, como mostrado no exemplo 84, os vetores

$$|u_1\rangle = (1, 0, 0), |u_2\rangle = (0, 1, 0), |u_3\rangle = (0, 0, 1) \quad (7.44)$$

formam uma base para $\mathbb{C}^3$. De maneira geral, os vetores

$$|u_1\rangle = (1, 0, 0, \dots, 0, 0), |u_2\rangle = (0, 1, 0, \dots, 0, 0), \dots, |u_n\rangle = (0, 0, 0, \dots, 0, 1) \quad (7.45)$$

formam uma base para $\mathbb{C}^n$, como pode ser verificado de maneira análoga aos casos de $\mathbb{C}^2$ e $\mathbb{C}^3$.

Teorema 35. Se $\{|u_1\rangle, \dots, |u_k\rangle\}$ é uma base para $\mathbb{C}^n$, então qualquer conjunto com mais de k vetores é LD.

O teorema acima permite provar que duas bases de $\mathbb{C}^n$ possuem o mesmo número de elementos. De fato, suponhamos que $\{|u_1\rangle, \dots, |u_k\rangle\}$ e $\{|v_1\rangle, \dots, |v_m\rangle\}$ sejam bases para $\mathbb{C}^n$. Se $m > k$, o teorema anterior implica que o conjunto $\{|v_1\rangle, \dots, |v_m\rangle\}$ é LD, e portanto não poderia ser uma base, o que é uma contradição. Se $m < k$, o teorema anterior implica que $\{|u_1\rangle, \dots, |u_k\rangle\}$ é LD, o que também é uma contradição. Logo $k = m$.

Definição 71. A *dimensão* de $\mathbb{C}^n$ é o número de vetores em uma base. O espaço vetorial $\mathbb{C}^2$ tem dimensão 2; O espaço vetorial $\mathbb{C}^3$ tem dimensão 3; Em geral, o espaço vetorial $\mathbb{C}^n$ tem dimensão n .

Teorema 36. Em $\mathbb{C}^n$ qualquer conjunto LI $\{|u_1\rangle, \dots, |u_n\rangle\}$ com n vetores é uma base.

Assim, para verificar se um conjunto de vetores é ou não uma base para o espaço $|CC^n$, basta verificarmos se ele possui o número correto de vetores e se esses vetores são LI. O Teorema 36 garante que nessas condições, o conjunto gera $\mathbb{C}^n$ e portanto é uma base. Podemos resumir esse resultado da seguinte forma:

Uma base para $\mathbb{C}^n$ é um conjunto LI com exatamente n vetores.

As demonstrações dos resultados acima são idênticas às mostradas na seção anterior para $\mathbb{R}^n$ e serão omitidas aqui.

Exercício 106. Verifique se os vetores abaixo formam uma base para o espaço vetorial ao qual pertencem.

1. $(1, 1 - i);$
2. $(1, 2 + i), (i, 0);$
3. $(1, 2), (i, 0), (0, 1);$
4. $(1, 0, 0), (0, 1, -i);$
5. $(1, 0, 0), (0, i, -i), (0, 1, 1);$
6. $(0, 1, 0), (0, i, -i), (0, 1, 1);$
7. $(2 + 3i, 1, -1), (3, 1, -2), (1, 0, -i), (i, 0, 1);$
8. $(1, 1, 0, 0), (5, -5 - 7i, 0, 0);$
9. $(1, 1, 0, 0), (i, -i, 0, 0), (1, 0, 1, 0);$

7.10.3 Produto Interno

Definição 72. Dados dois vetores $|u\rangle = (z_1, \dots, z_n)$ e $|v\rangle = (w_1, \dots, w_n)$ em $\mathbb{C}^n$ definimos o *produto interno* entre $|u\rangle$ e $|v\rangle$ por

$$\langle u | v \rangle = (z_1, \dots, z_n) \cdot (w_1, \dots, w_n) = \sum_{i=1}^n \bar{z}_i w_i. \quad (7.46)$$

Exemplo 91. O produto interno entre os vetores $(i, 3i, -5)$ e $(4, -2, -i)$ em $\mathbb{C}^3$ é igual a

$$(i, 3i, -5) \cdot (4, -2, -i) = \bar{i} \times 4 + \overline{3i} \times (-2) + (-5) \times (-i) = -4i + 6i + 5i = 7i.$$

Teorema 37. O produto interno em $\mathbb{C}^n$ satisfaz as seguintes propriedades:

1. $(\lambda |u\rangle + \mu |v\rangle) \cdot |w\rangle = \bar{\lambda} (\langle u | w \rangle) + \bar{\mu} (\langle v | w \rangle);$
2. $\langle u | v \rangle = \overline{\langle v | u \rangle};$
3. $\langle u | u \rangle \geq 0;$
4. Se $\langle u | u \rangle = 0$ então $|u\rangle = 0$.

Definição 73. A partir do produto interno podemos definir a *norma* de um vetor $|v\rangle$ escrevendo

$$\|v\| = \sqrt{\langle v | v \rangle}.$$

O produto interno nos permite introduzir uma noção que generaliza a um espaço vetorial qualquer a ideia de perpendicularidade nos espaços $\mathbb{R}^2$ e $\mathbb{R}^3$, com a qual já estamos familiarizados.

Definição 74. Dizemos que dois vetores $|u\rangle$ e $|v\rangle$ são *ortogonais* se $\langle u | v \rangle = 0$. Dizemos que um conjunto $E = \{|v_1\rangle, \dots, |v_k\rangle\}$ é *ortogonal* se seus elementos são dois a dois ortogonais. Dizemos que um conjunto $E = \{|v_1\rangle, \dots, |v_k\rangle\}$ é *ortonormal* se é ortogonal e $\langle v_i | v_i \rangle = 1$ para todo i .

Teorema 38. Se $\{|v_1\rangle, \dots, |v_k\rangle\}$ é um conjunto ortogonal, então valem as seguintes propriedades:

1. O conjunto $\{|v_1\rangle, \dots, |v_k\rangle\}$ é LI;
2. Se $|v\rangle = a_1|v_1\rangle + \dots + a_k|v_k\rangle$, então

$$a_i = \frac{\langle v | v_i \rangle}{|| |v_i\rangle ||^2}. \quad (7.47)$$

Definição 75. Dizemos que uma base $\mathcal{B} = \{|v_1\rangle, \dots, |v_k\rangle\}$ para $\mathbb{C}^n$ é uma *base ortogonal* se seus elementos são dois a dois ortogonais. Dizemos uma base $\mathcal{B} = \{|v_1\rangle, \dots, |v_k\rangle\}$ é uma *base ortonormal* se é ortogonal e $\langle v_i | v_i \rangle = 1$ para todo i .

Exemplo 92. A base

$$\mathcal{B} = \{|e_1\rangle = (1, 0, \dots, 0), |e_2\rangle = (0, 1, \dots, 0), \dots, |e_n\rangle = (0, 0, \dots, 1)\} \quad (7.48)$$

de $\mathbb{C}^n$ é uma base ortonormal. De fato, é fácil mostrar que o produto interno $\langle e_i | e_j \rangle$ é igual a zero se $i \neq j$ e igual a 1 se $i = j$.

Exemplo 93. Mostre que o conjunto

$$\mathcal{B} = \{|v_1\rangle = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right), |v_2\rangle = \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right), |v_3\rangle = (0, i, 0)\}$$

é uma base ortonormal de $\mathbb{C}^3$.

Solução. Calculando os produtos internos temos

$$\langle v_1 | v_2 \rangle = \langle v_1 | v_3 \rangle = \langle v_2 | v_3 \rangle = 0, \quad \langle v_1 | v_1 \rangle = \langle v_2 | v_2 \rangle = \langle v_3 | v_3 \rangle = 1$$

o que mostra que B é um conjunto ortonormal. Pelo teorema 38, B é um conjunto LI e portanto é uma base, já que possui três vetores. $\square$

Exercício 107. Verifique se os conjuntos de vetores abaixo são ortogonais. Verifique quais deles são ortonormais. Verifique quais deles constituem bases de $\mathbb{C}^3$.

7.10. *DEPENDÊNCIA E INDEPENDÊNCIA LINEAR*

107

1. $\{(1, 2i, -1), (i, 0, i)\}$;
2. $\left\{\left(\frac{i}{\sqrt{3}}, \frac{i}{\sqrt{3}}, \frac{-i}{\sqrt{3}}\right), \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)\right\}$;
3. $\{(-i, -2i, -2i), (1, 2, -1), (0, 0, i)\}$;
4. $\left\{\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(\frac{-i}{\sqrt{6}}, \frac{-i}{\sqrt{6}}, \frac{2i}{\sqrt{6}}\right) \left(\frac{i}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, 0\right)\right\}$.

Capítulo 8

Evoluções Unitárias

No capítulo 4 discutimos os processos de preparação e medição de um qbit. O processo de medição foi associado a uma escolha de base ortonormal do espaço $\mathbb{C}^2$. Existe um conjunto especial de matrizes, que chamamos de *matrizes unitárias*, que tem em suas colunas bases ortonormais. Assim, também podemos padronizar o processo de preparação e medição de um qbit como a seqüência: preparar o estado $|0\rangle$, atuar nele com uma matriz unitária, U , e fazer o teste 0 ou 1, também conhecido como *medir na base computacional*.

O procedimento descrito acima pode ser entendido como uma *computação quântica* que utiliza um único qbit. Melhor ainda, ela é a computação quântica mais geral possível, para um qbit. Assim, vemos que há uma bijeção entre as possíveis computações quânticas de um qbit e as matrizes unitárias 2×2 .

8.1 O grupo $U(2)$

Começamos por definir mais uma estrutura algébrica muito interessante:

Definição 76. Um *grupo* é um conjunto G em que podemos definir uma operação

$$\cdot : G \times G \longrightarrow G$$

$$(a, b) \longmapsto a \cdot b = ab$$

tal que, para todos $a, b, c \in G$, valem

1. Associatividade: $a \cdot (b \cdot c) = (a \cdot b) \cdot c$;
2. Existência do elemento neutro: existe em G o elemento neutro e , tal que $e \cdot a = a \cdot e = a$;
3. Existência de inverso: existe $a^{-1} \in G$ tal que $a \cdot a^{-1} = e$.

Note que não é exigida comutatividade. De fato, se também valer a propriedade comutativa, $a \cdot b = b \cdot a$, o grupo é dito *abeliano*. Para o nosso caso motivador, veremos que a não-comutatividade tem consequências muito interessantes.

Uma matriz complexa, A , tem o que chamamos de hermitiana conjugada, $A^\dagger$, que pode ser obtida tomando o conjugado complexo da matriz transposta de A . Ou seja

$$A_{ij}^\dagger = \overline{A_{ji}}.$$

E agora podemos definir o grupo $U(n)$, em especial, o $U(2)$. Uma matriz quadrada U é dita unitária se $U^\dagger U = I$, onde I denota a matriz identidade. $U(n)$ é o conjunto de todas as matrizes $n \times n$, unitárias.

Exercício 108. Mostre que $U(n)$, com a multiplicação usual de matrizes complexas, é um grupo.

Queremos agora entender melhor o grupo $U(n)$, e em especial, o $U(2)$.

Exercício 109. Mostre que, se considerarmos cada coluna de uma matriz $n \times n$ como um vetor de $\mathbb{C}^n$, a matriz M é unitária se, e somente se, suas colunas formam uma base ortonormal.

8.2 Subgrupos de $U(2)$

Um conceito muito importante, em teoria de grupos, é o conceito de subgrupo. Se G é um grupo, qualquer subconjunto de G que, com a mesma operação, forma um grupo é um subgrupo de G . Dois exemplos triviais de subgrupos de G são $\{e\}$ e G . Qualquer outro subgrupo é um subgrupo próprio.

Para discutir alguns exemplos interessantes, consideremos alguns elementos especiais do grupo $U(2)$, conhecidas como *matrizes de Pauli*.

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \quad (8.1)$$

Exercício 110. Mostre que para toda matriz de Pauli, σ_k , tem-se $\{I, \sigma_k\}$ um grupo.

Outro conceito interessante é o de subgrupo gerado por um subconjunto, S , de um grupo. Isso pode ser entendido como o menor subgrupo que contém S .

Exercício 111. Obtenha o subgrupo de $U(2)$ gerado por $\{\sigma_x, \sigma_z\}$. (Dica: esse grupo tem 8 elementos).

Um isomorfismo entre grupos G e H é dado por uma função inversível $f : G \rightarrow H$ tal que $f(a \cdot b) = f(a) \cdot f(b)$. Note que nessa definição usamos a mesma notação para a operação do grupo G e do grupo H . Mais comum ainda é omitir tal notação e entender que ab denota $a \cdot b$. Caso exista um isomorfismo $f : G \rightarrow H$, dizemos que G e H são isomorfos.

Exercício 112. Mostre que os subgrupos gerados por $\{\sigma_x, \sigma_z\}$, por $\{\sigma_x, \sigma_y\}$ e por $\{\sigma_y, \sigma_z\}$ são todos isomorfos. Da mesma forma, mostre que todo subgrupo gerado por uma única matriz de Pauli é isomorfo ao $\mathbb{Z}_2$, o grupo $\{0, 1\}$ com a adição módulo 2.

Se todos os subgrupos gerados por uma ou duas matrizes de Pauli já foram caracterizados, o próximo passo é usar as três.

Exercício 113. Obtenha o subgrupo de $U(2)$ gerado pelas três matrizes de Pauli. Quanto elementos ele tem?

Este grupo é chamado de *grupo de Pauli*.

Um outro elemento interessante do grupo $U(2)$, para quem trabalha com computação quântica, é a chamada matriz de Hadamard:

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

Observe que as colunas da matriz de Hadamard formam uma base bastante especial.

Exercício 114. Obtenha o subgrupo gerado pelas três matrizes de Pauli e pela matriz de Hadamard.

8.2.1 Subgrupos a um parâmetro

Até agora vimos exemplos de subgrupos finitos de $U(2)$. Você já deve ter percebido que o $U(2)$ possui uma infinidade de elementos. Assim, é razoável perguntar se há subgrupos próprios infinitos.

Alguns exemplos interessantes aparecem de forma bem simples:

Exemplo 94. Considere as matrizes

$$P(\phi) = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{bmatrix},$$

chamadas matrizes de fase. Note que $P(\phi)P(\theta) = P(\phi + \theta)$, com o que é fácil perceber que $\{P(\phi)\}$ é um subgrupo, isomorfo ao grupo $U(1) = \{e^{i\phi}, \phi \in \mathbb{R}\}$.

Este subgrupo é um primeiro exemplo de subgrupo a um parâmetro e todo subgrupo a um parâmetro¹ de $U(n)$ é isomorfo ao $U(1)$.

Exercício 115. Mostre que a notação $U(1)$ é consistente com a definição geral de $U(n)$. Ou seja, verifique que ela corresponde ao caso $n = 1$ para $U(n)$.

8.2.2 A exponencial

Outro exemplo, com aplicações em várias áreas de matemática, vem da função exponencial. Existem várias maneiras de definir a exponencial de uma matriz.

¹Nesse texto, e em vários outros lugares, subgrupo a um parâmetro significa subgrupo a um parâmetro *real*. Outros exemplos interessante, e não isomorfos a $U(1)$, podem aparecer com um parâmetro racional, em outros casos, complexo...

Duas formas particularmente interessantes são pelo uso de séries de potências ou pelo uso de autovalores e autovetores. Mas ambos são caminhos mais avançados do que queremos aqui.

Quando introduzimos a exponencial de números imaginários puros,

$$e^{i\phi} = \cos \phi + i \operatorname{sen} \phi.$$

era essencial o fato que $i^2 = -1$.

Exercício 116. Verifique a validade da fórmula $(i\sigma_k) = -I$ para $k = x, y, z$.

Com isso, podemos ver que I e $i\sigma_k$, para cada k fixado, jogam o mesmo papel que 1 e i para os números complexos. Podemos então tomar como uma definição que

$$\exp(i\phi\sigma_k) = \cos(\phi)I + i\sin(\phi)\sigma_k. \quad (8.2)$$

Exercício 117. Verifique que

$$\exp(i\phi\sigma_z) = \begin{bmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{bmatrix}. \quad (8.3)$$

Exercício 118. Obtenha explicitamente $\exp(i\phi\sigma_x)$ e $\exp(i\phi\sigma_y)$.

De fato, as fórmulas aqui obtidas são exemplos especiais de algo bem mais geral. As matrizes de Pauli nos permitem usar uma notação de produto escalar para escrever

$$\vec{v} \cdot \vec{\sigma} = \begin{bmatrix} v_z & v_x - iv_y \\ v_x + iv_y & -v_z \end{bmatrix}, \quad (8.4)$$

onde $\vec{v} = (v_x, v_y, v_z) \in \mathbb{R}^3$.

Exercício 119. Mostre que $\text{Tr}(\vec{v} \cdot \vec{\sigma}) = 0$ e que $\det \vec{v} \cdot \vec{\sigma} = \|\vec{v}\|$.

Exercício 120. Mostre que $(i\vec{v} \cdot \vec{\sigma})^2 = -\|\vec{v}\|^2 I$.

E assim, generalizamos a definição anterior para

$$\exp(i\phi\vec{v}\cdot\vec{\sigma}) = \cos(\phi)I + i\sin(\phi)\vec{v}\cdot\vec{\sigma}, \quad (8.5)$$

onde $\|\vec{v}\| = 1$.

Exercício 121. Fixado $\vec{v}$, mostre que $\{\exp(i\phi\vec{v} \cdot \vec{\sigma})\}$ é um subgrupo a um parâmetro de $U(2)$.

Poderíamos ir muito mais longe, associando esses subgrupos a rotações, generalizando para exponenciais de outras matrizes, comentando a relação dessa construção com grupos e álgebras de Lie... mas tudo isso fica como convite ao leitor/estudante para aprofundar seus estudos.

8.2.3 O grupo $SU(2)$

Outro subgrupo interessante de $U(2)$ é dado por $SU(2) = \{U \in U(2) : \det U = 1\}$. Nessa notação, a letra S vem de *Special*, devido à condição especial do determinante ser um.

Exercício 122. Mostre que $SU(2)$ é subgrupo.

Exercício 123. Sejam a e b números complexos tais que $|a|^2 + |b|^2 = 1$. Mostre que

$$\begin{bmatrix} a & -\bar{b} \\ b & a \end{bmatrix}$$

pertence a $SU(2)$. Mais ainda, mostre que todo elemento de $SU(2)$ é dessa forma.

No exercício anterior você mostrou que há uma bijeção² entre $SU(2)$ e $S^3 = \{x \in \mathbb{R}^4 : \|x\| = 1\}$.

Do ponto de vista da teoria quântica, como fases globais são irrelevantes, podemos dizer que as computações quânticas de um qbit podem ser descritas pelo $SU(2)$, ou ainda, pela esfera S^3 .

Exercício 124. Mostre ainda que todo elemento de $SU(2)$ pode ser escrito como $\exp(i\vec{v} \cdot \vec{\sigma})$, para uma escolha adequada de $\vec{v} \in \mathbb{R}^3$.

8.3 Alguns outros grupos interessantes

Já falamos do $U(n)$, mas esses grupos também possuem irmãos para matrizes com entradas reais. Uma matriz $n \times n$ com entradas reais, A , é dita *ortogonal* se $A^t A = I$, onde A^t denota a transposta de A .

Exercício 125. Prove que $O(n) = \{A \in M_{n \times n} : A^t A = I\}$ é grupo.

Exercício 126. Mostre que se A é uma matriz ortogonal, suas colunas formam uma base ortonormal para $\mathbb{R}^n$.

Um caso particularmente interessante é $O(2)$. Para entendê-lo, é interessante notar que ele se divide em “dois pedaços”.

Exercício 127. Mostre que se $A \in O(2)$, $\det A = \pm 1$.

Como $\det : M_{n \times n} \rightarrow \mathbb{C}$, é uma função contínua, os conjuntos $\det^{-1}(1)$ e $\det^{-1}(-1)$ em $O(2)$ serão isolados. É interessante notar que em $U(2)$ as coisas são diferentes. As matrizes $P(\phi)$, do exemplo 94, com $\phi \in [0, \pi]$ formam um caminho em $U(2)$ ligando uma matriz com determinante 1 a uma matriz com determinante -1 .

²Há um pouco mais, mas não vamos nos aprofundar...

Um interessante subgrupo de $O(n)$ é dado por $SO(n)$, onde a letra S tem o mesmo significado de antes. Ou seja, $SO(n) = \{A \in O(n) : \det A = 1\}$.

Exercício 128. Para não perder o hábito, mostre que $SO(n)$ é grupo. Por que $SO^-(n) = \{A \in O(n) : \det A = -1\}$ não é um grupo?

Seguindo a mesma estratégia do exercício 123, podemos parametrizar o grupo $SO(2)$. Mas com uma vantagem, como agora $a, b \in \mathbb{R}$, podemos recorrer à trigonometria para usar apenas um ângulo na parametrização.

Exercício 129. Mostre que

$$A = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix},$$

estabelece uma bijeção entre $S^1 = \{x \in \mathbb{R}^2 : \|x\| = 1\}$ e $SO(2)$.

O exercício anterior nos mostra que os elementos de $SO(2)$ podem ser vistos como rotações. Em especial, rotações por um ângulo θ . Lido de outra forma, ele diz que podemos escolher qualquer vetor unitário para ser a primeira coluna da matriz, e que a segunda coluna está definida pela primeira, girada de $\pi/2$.

O que acontece com os demais elementos de $O(2)$?

Exercício 130. Mostre que todo elemento de $SO^-(2)$ é da forma

$$B(\theta) = \begin{bmatrix} \sin \theta & \cos \theta \\ \cos \theta & -\sin \theta \end{bmatrix}.$$

Interprete geometricamente $B(0)$. Mostre que $B(\theta) = B(0)A$ e interprete esse resultado em termos de composição de transformações.

Por fim, você já deve ter percebido um interessante isomorfismo:

Exercício 131. Mostre que $SO(2)$ é isomorfo ao $U(1)$.

Há também uma interessante relação entre $SU(2)$ e $SO(3)$, mas não vamos investir nela aqui.

Capítulo 9

Aproximação de Unitárias

Nesse curto capítulo vamos apresentar alguns resultados simples, mas cuja generalização ganhou notoriedade por sua aplicabilidade. Se queremos construir um computador quântico, em princípio precisamos ser capazes de construir “máquinas” capazes de implementar todas as operações possíveis em um conjunto de vários qbits. Nesse texto só estamos interessados em qbits individuais¹, mas vamos pensar um pouquinho em vários qbits. Se para cada computação quântica que quiséssemos realizar, fosse necessário construir um computador especialmente dedicado e este computador precisasse incluir um grande número de interações entre grandes números de qbits, o desafio seria ainda maior do que já é. O resultado muito celebrado garante que podemos trabalhar com uma noção de “computador quântico universal”, parecido com os computadores clássicos a que estamos acostumados: um único *hardware* pode rodar diferentes programas: essa noção de computador programável amplia imensamente a utilidade dos dispositivos!

Melhor que isso, também assim como nos computadores clássicos, é interessante ter o que, em engenharia, se chama *modularização*; ou seja, que algumas peças básicas possam ser “encaixadas” de diferentes maneiras, gerando diferentes equipamentos. Vamos argumentar aqui pela existência dessas peças básicas em computação quântica.

No problema prático, é importante trabalhar com n qbits, onde n varia com o tipo de problema ou aplicação em mente. Vamos nos manter no caso $n = 1$ e mostrar que basta a capacidade de fazer poucas operações para ser capaz de *aproximar* qualquer computação quântica.

9.1 Operações Quânticas Universais

Em computação, um conjunto de operações é chamado *universal* quando é possível gerar as demais operações compondo estas. É o mesmo conceito que um

¹Você pode aprender sobre mais qbits com alguns textos da bibliografia.

conjunto de geradores para um grupo². Em palavras mais relaxadas, são os “blocos fundamentais” a partir dos quais construímos nosso quebra-cabeça, “encaixando-os” conforme as regras do jogo.

Em computação quântica, esse conceito é um pouco diferente, ganhando um espírito mais próximo à análise³. Como o conjunto das possíveis computações é um conjunto infinito e com uma noção natural de proximidade⁴, não é razoável querer gerar todos os elementos a partir de alguns poucos. Mas para toda implementação prática, é suficiente estar razoavelmente próximo da operação desejada. Assim, o ideia de universalidade muda um pouco, ainda que ganhando algum aspecto técnico que também não pretendemos detalhar aqui.

Um conjunto de operações quânticas é dito universal se, compondo seus elementos, é possível aproximar qualquer computação quântica de maneira eficiente⁵. Aqui, podemos traduzir isso como $S = \{u_i\}$ é um conjunto universal se o subgrupo gerado por S , $\langle S \rangle$, for denso em $U(2)$. Esse é um bom momento para esclarecer o que queremos dizer com *denso*.

9.2 Subconjuntos densos

Este é um conceito que só faz sentido para conjuntos infinitos e com uma noção natural de vizinhança para cada ponto. Um subconjunto $D \subset A$ é *denso em A* se todo elemento de A pode ser aproximado por elementos de D . Isso quer dizer que, dado $a \in A$, existe sequência em D , $(d_i)_i$ que *converge* para a .

É inadequado querermos detalhar o conceito de convergência de sequências. Mas podemos dar uma ideia...

O exemplo mais famoso de subconjunto denso é $\mathbb{Q} \in \mathbb{R}$. O conjunto dos racionais é denso nos reais. Uma maneira de entender isso é pensar na expansão decimal de um número real a . Cada vez que truncamos uma aproximação decimal na n -ésima casa, temos um número racional que está à distância do número desejado menor que 10^{-n} . Assim, podemos pensar em cada d_i como a expansão até o i -ésimo algarismo depois da vírgula, e teremos uma sequência de números racionais onde, para qualquer distância que fixarmos, “quase todos” os temos da sequência estão a uma distância de a menor que tal distância. Como uma primeira ideia, isso significa que a sequência $(d_i)_i$ converge para a . Como podemos fazer isso para todo número real, temos que $\mathbb{Q}$ é denso em $\mathbb{R}$.

O segundo exemplo bastante famoso em matemática já é mais próximo do problema que queremos resolver. Se lembrarmos que o grupo $U(1)$ corresponde à circunferência S^1 , podemos agora entender os subgrupos de $U(1)$ gerados por um

²Ainda que sem qualquer relação com outros significados da palavra universal, por exemplo, em álgebra.

³Análise é uma das grandes áreas da Matemática. As outras são Álgebra e Geometria/Topologia, ainda que vários dos avanços mais significativos da matemática ignorem essa separação em áreas.

⁴Em termos mais precisos, com uma topologia métrica.

⁵Não vamos querer detalhar o conceito de eficiência, típico da computação, e que só tem importância para número arbitrário de qbits.

único elemento. Seja $u = \exp(i\phi) \in U(1)$. Temos duas alternativas: ou existe $n \in \mathbb{N}$ tal que $u^n = 1$, ou não existe tal n . No primeiro caso (existe n), é verdade, também, que existe um menor tal n (todo conjunto não vazio de naturais possui um menor elemento). Chamamos esse menor n de ordem de u em $U(1)$. Se estamos trabalhando com o subgrupo gerado por u , segue um resultado muito interessante:

Exercício 132. Se u tem ordem n , mostre que o subgrupo gerado por u tem n elementos.

Como nosso assunto agora são conjuntos densos, este é um caso desinteressante.

Exercício 133. Mostre que o subgrupo de $U(1)$ gerado por u , onde u tem ordem n , não é denso.

Já quando não existe n natural tal que $u^n = 1$ (e dizemos que u tem ordem infinita), a situação é mais interessante. Se tomamos $\phi \in (0, 2\pi)$, notamos que $u^n = \exp(in\phi)$. Assim, não importa se tomarmos ϕ um ângulo grande ou pequeno, para o primeiro valor de $n > \frac{2\pi}{\phi}$ teremos completado uma volta e sobrá um ângulo menor que ϕ . Se pensarmos na circunferência, esse elemento do grupo fica entre $1 = u^0$ e u .

Exercício 134. Chame de v esse elemento do subgrupo gerado por u que se encontra entre 1 e u . Prove que há outro elemento do subgrupo gerado por u entre 1 e v .

Claro que esse processo não pára por aqui...

Exercício 135. Lembrando que $U(1)$ pode ser visto como S^1 , mostre que se u tem ordem infinita, escolhidos $m \neq n \in \mathbb{Z}$, existem elementos do subgrupo gerado por u nos dois arcos definidos por u^m e u^n .

Exercício 136. Mostre que, se u tem ordem infinita, então o subgrupo gerado por u é denso em $U(1)$.

O produto cartesiano de duas circunferências dá um *toro*, a figura de um pneu ou de uma rosquinha: $T^2 = S^1 \times S^1$. Como $U(1)$ corresponde à circunferência S^1 , só precisamos agora entender que o produto cartesiano de dois grupos ganha naturalmente uma estrutura de grupo, onde $(a_1, a_2) \cdot (b_1, b_2) = (a_1 \cdot b_1, a_2 \cdot b_2)$. Com isso, podemos generalizar os exemplos anteriores.

Exercício 137. Sejam $u, v \in U(1)$, mostre que

1. Se u e v têm ordem finita, o subgrupo gerado por $(u, 1)$ e por $(1, v)$ é finito e não é denso em $U(1)$;
2. Se u tem ordem finita e v tem ordem infinita, o subgrupo gerado por $(u, 1)$ e por $(1, v)$ é infinito e não é denso em $U(1)$;

3. Se u e v têm ordem infinita, o subgrupo gerado por $(u, 1)$ e por $(1, v)$ é infinito e é denso em $U(1)$.

Como visão geométrica sempre ajuda...

Exercício 138. Tente esboçar figuras em T^2 descrevendo os subgrupos gerados por dois elementos da forma $(u, 1)$ e $(1, v)$. Se quiser se divertir ainda mais, pense no que pode acontecer primeiro com subgrupos a um gerador em $U(1) \times U(1)$, depois com subgrupos a dois geradores, quaisquer, em $U(1) \times U(1)$.

9.3 Operações Universais de um qbit

De certa forma, o que queremos agora é trabalhar com subgrupos finitamente gerados de $U(2)$ e $SU(2)$, generalizando os exemplos anteriores em $U(1)$ e $U(1) \times U(1)$. Mais que isso, trabalharemos com poucos geradores.

Exercício 139. Seja $u \in U(2)$, com $u \neq e$. Mostre que se a ordem de u for finita, o subgrupo gerado por u é finito; se a ordem de u for infinita, o subgrupo gerado por u é infinito, mas não é denso em $U(2)$. (Sugestão: pode ser interessante mostrar que, nesse caso, ele é denso em um subgrupo a um parâmetro, e que esse subgrupo a um parâmetro, isomorfo a $U(1)$, não é denso em $U(2)$.)

Até aí está tudo razoavelmente parecido com o caso de $U(1) \times U(1)$. Mas enquanto este é um grupo abeliano, $U(2)$ e $SU(2)$ não são. E isso traz muita riqueza...

Considere dois elementos de $SU(2)$ da forma $u_1 = \exp(\vec{v}_1 \cdot \vec{\sigma})$ e $u_2 = \exp(\vec{v}_2 \cdot \vec{\sigma})$.

Exercício 140. Obtenha condições em v_1 e v_2 para que:

1. O subgrupo gerado por u_i seja finito;
2. O subgrupo gerado por u_1 seja um subgrupo do subgrupo gerado por u_2 .

Agora suponha que v_1 e v_2 sejam linearmente independentes. O ponto essencial é que, nesse caso, teremos $u_1 u_2 \neq u_2 u_1$.

Exercício 141. Faça o caso $v_1 = (1, 0, 0)$ e $v_2 = (0, 0, 1)$ para se convencer. Se tiver disposição, mostre o caso geral.

Moral da história, o subgrupo de $SU(2)$ gerado por u_1 e u_2 não é abeliano.

O ponto central vem quando colocamos álgebra e geometria para conversar. Já sabemos que o grupo $SU(2)$ corresponde à esfera S^3 ; ou seja, é um objeto tridimensional. Seria natural, então, pensarmos que ao “juntar” dois subgrupos a um parâmetro como $\{\exp(s\vec{v}_1 \cdot \vec{\sigma})\}$ e $\{\exp(s\vec{v}_2 \cdot \vec{\sigma})\}$, cada um deles isomorfo a $U(1)$ e, geometricamente, circunferências S^1 , obteríamos algo como o toro $T^2 = S^1 \times S^1$ discutido anteriormente. Mas isso não é verdade! O fato já apresentado de

$u_1 u_2 \neq u_2 u_1$ pode ser generalizado. Usamos a notação $u_1(s) = \exp(s\vec{v}_1 \cdot \vec{\sigma})$ e $u_2(t) = \exp(t\vec{v}_2 \cdot \vec{\sigma})$. É genericamente⁶ verdade que $u_1(s) u_2(t) \neq u_2(t) u_1(s)$. Podemos pensar os dois lados dessa desigualdade como, de um lado primeiro andamos t unidades na direção de $\vec{v}_2$ e depois s unidades na direção de $\vec{v}_1$, e do outro fazemos na ordem contrária, primeiro s unidades na direção de $\vec{v}_1$ e depois t unidades na direção de $\vec{v}_2$. Se fizéssemos isso em $\mathbb{R}^3$, chegaríamos ao mesmo lugar, fechando um paralelogramo, e não teríamos uma desigualdade. O ponto é que, em S^3 , não temos essa mesma noção de paralelismo. O fato de não fechar um paralelogramo, ao invés de ser um problema, é uma riqueza.

A sutileza adicional está no fato dos elementos da forma $u_1(s)u_2(t)$ não formarem um grupo. O inverso de $u_1(s)u_2(t)$ é $u_2(-t)u_1(-s)$ e, genericamente, ele não tem a forma em questão (por conta da não comutatividade! Fosse em um grupo abeliano, sim, teria). Da mesma forma, o produto de dois elementos dessa forma, não dá (genericamente) um elemento dessa forma.

Isso diz que, apenas com v_1 e v_2 , mas com a liberdade de “andar quanto quiser” em cada uma dessas duas direções, quantas vezes quiser, podemos gerar elementos como

$$u_1(s_1)u_2(t_1)u_1(s_2)u_2(t_2)\dots u_1(s_n)u_2(t_n). \quad (9.1)$$

Sem irmos demasiadamente aos detalhes, podemos recordar do exercício 124 para argumentar que bastariam três direções independentes para gerar todos os elementos de $SU(2)$. O ponto central é que, genericamente, a direção correspondente ao elemento $u_1 u_2$ forma um conjunto independente com as direções dadas por u_1 e u_2 , respectivamente $\vec{v}_1$ e $\vec{v}_2$.

O argumento final passa por voltar à expressão (9.1) e trocar os parâmetros $s_i, t_j \in \mathbb{R}$ por inteiros. Ou seja, estaremos voltando aos elementos u_1 e u_2 e construindo o subgrupo gerado por eles. Genericamente, esse grupo é infinito. Melhor que isso, é denso.

Ou seja, concluímos que essencialmente qualquer par de elementos de $SU(2)$ é um conjunto universal para computação quântica de um qbit!

Exercício 142. Justifique por que σ_x e σ_z não formam um conjunto universal para computação quântica de um qbit.

Exercício 143. Exiba um conjunto universal para computação quântica de um qbit.

9.4 E para mais qbits?

Não é nosso foco, aqui, mas podemos adiantar que para mais qbits segue-se a mesma noção de densidade. É comum, por questões práticas, assumir que tem-se disponível todas as operações unitárias de um qbit, mas, como vimos, do ponto

⁶Esse é um termo matemático muito interessante... algo é genericamente válido se, para quase toda escolha dos parâmetros envolvidos, é verdade. No caso em questão, se fizermos $t = 0$, por exemplo, não é verdade, bem como para $t = k\pi / \left\| \vec{v}_2 \right\|$, com $k \in \mathbb{Z}$.

matemático isso é equivalente a ter um par de operações universais para um qubit e concatená-las adequadamente. A questão séria, depois, é: será que precisamos atuar conjuntamente em muitos qbits? A interessante resposta é que não. Basta a liberdade de atuar com uma operação adequada em qualquer par de qbits para se chegar a um conjunto universal.

Referências Bibliográficas

- [1] Portal da Matemática - OBMEP, *Números Complexos – Forma Algébrica*.
- [2] Portal da Matemática - OBMEP, *Números Complexos – Forma Geométrica*.
- [3] M.G. Soares, “Cálculo de uma Variável Complexa,” Coleção Matemática Universitária, IMPA (2009).
- [4] I. Vainsencher, “Notas de Aula de Álgebra Linear II”.
- [5] B. Amaral, “Introdução à Álgebra Linear”.
- [6] R. Santos, “Um curso de Geometria Analítica e Álgebra Linear,” Imprensa UFMG, 2010.
- [7] E.L. Lima, “Álgebra Linear,” Coleção Matemática Universitária, IMPA (2008).
- [8] Bárbara Amaral, *Uma Breve Introdução aos Conceitos Matemáticos da Física Quântica*. Notas e vídeo-aulas desenvolvidas para o curso Mentores - OBMEP.
- [9] B. Amaral, M. O. Terra Cunha, A. T. Baraviera, “*Mecânica Quântica para Matemáticos em Formação*” Monografias de Matemática, IMPA-SBM (2011).
- [10] <http://www.idquantique.com> e <http://www.magiqtech.com> são bons exemplos.
- [11] A. T. Baraviera, “*Introdução à Mecânica Quântica*,” 1º Colóquio de Matemática da Região Sul (2010).
- [12] R.P. Feynman, R.B. Leighton e M. Sands, “*The Feynman Lectures on Physics*,” vol. 3, Addison-Wesley publishing company (1965).
- [13] M.A. Nielsen e I.L. Chuang, “*Quantum Computation and Quantum Information*,” Cambridge University Press (2000).

- [14] A. Peres, “*Quantum Theory: Concepts and Methods*,” Kluwer Academic Publishers (1995).
- [15] P. C.M. Penteado, C. M. A. Torres, “*Física, Ciência e Tecnologia*,” Moderna (2005).
- [16] D. Halliday, R. Resnick, K. S. Krane, “*Física 4*,” LTC (2004).
- [17] G. J. Biscuola, N. Villas Bôas, R. H. Doca, “*Física*” Saraiva (2010).
- [18] M. Terra Cunha, “*Noções de Informação Quântica*,” Monografias de Matemática, IMPA-SBM (2007).