

Parte I

Parte 1 - Fundamentos

Introdução

A Mecânica Quântica Antiga

A mecânica quântica nasce, no início do século XX, para explicar diversos resultados experimentais que não eram previstos pela mecânica clássica. Os mais relevantes são explicitados abaixo:

- A radiação de corpo negro [1899], que apresenta a catástrofe do ultravioleta (teoria clássica prevê que a taxa de energia deve ir ao infinito com o aumento da frequência da luz). A correta descrição da dependência desta taxa de energia com a frequência da luz é possibilitada postulando-se que a energia é entregue em pacotes (quanta), dados por $\Delta\epsilon = h\nu$, onde h é a constante de Planck e ν a frequência da luz.
- O efeito fotoelétrico [1886], que apresenta diversas características incompatíveis com a teoria clássica: a energia cinética de elétrons foto-emitidos depende da intensidade do campo; existe uma diferença de potencial limite para a emissão de elétrons que depende da frequência da luz; não existe retardo no acúmulo de energia pelos elétrons. Todas estas características foram explicadas por Einstein [1905] considerando que a luz é composta pelos pacotes de energia de Planck, sendo a energia cinética dos elétrons fotoemitidos dada por $K_e = h\nu - W$, onde W é a energia necessária para arrancar o elétron do material (função trabalho).
- Efeito Compton [1886], que é a observação do espalhamento inelástico de raios X por uma carga elétrica (geralmente um elétron), que pela teoria clássica deveria ser elástico (sem mudança na frequência do raio X). O efeito foi explicado considerando a propriedade corpuscular da luz, considerando que pacotes luminosos incidentes e espalhados e a carga espalhadora devem respeitar a conservação de energia e momento, similar a uma colisão de partículas. A dualidade onda-partícula [1924] foi introduzida por de Broglie, bem como o princípio de incerteza de

Heizemberg, importantes para entender os processos de espalhamento de campos e partítulas elementares.

- O calor específico de sólidos ($c_v = dE/dT$) vai a zero quando a temperatura é diminuída, e a tendência de decaimento segue uma lei universal quando normalizada pela temperatura de Debye. Este comportamento foi explicado por Einstein, com a introdução dos hoje chamados "phonons", quanta de energia de vibrações atômicas em matéria condensada.
- O espectro atômico do H, observado por Balmer [1885], que apresenta linhas de energia bem definidas. O espectro de emissão, e a próxima estrutura do átomo, sendo formado por elétrons, prótons e neutrons, não pode ser explicada por uma teoria clássica, tendo sido elucidada por Bohr [1913] com a introdução da quantização do momento angular orbital do elétron - $mvr = N\hbar$, com $\hbar = h/\pi$. O conceito foi generalizado pelas regras de Sommerfeld [1916], que estipulam que qualquer sistema físico no qual as coordenadas são funções periódicas do tempo, existe uma condição de quantização para cada coordenada.

O princípio da correspondência de Bohr [1923], propõe que as regras de seleção da física clássica são válidas na física quântica, e que a física quântica corresponde à física clássica quando o número N que estipula o orbital molecular tende a infinito. De fato, este limite faz sentido em sistemas específicos, onde a distância entre os níveis de energia vai tendendo a zero quando $N \rightarrow \infty$, mas não é uma regra de equivalência universal. De fato, a mecânica quântica estruturada até então, só descrevia bem sistemas periódicos, e limitava-se ao cálculo de energias, não sendo capaz de abordar as taxas de transição (probabilidades de ocorrência). Sem contar que descrevia bem o espectro do átomo de H (e outros com 1 elétron na camada de valência, como Li, Na, K...), mas já era falha na descrição da maioria dos átomos multieletrônicos, até mesmo do He.

Erwin Schroedinger [1925] propôs a formulação da mecânica quântica, que ao invés de determinar a posição de uma partícula em função do tempo, $\vec{r}(t)$, determina uma propriedade física do sistema, chamada função de onda, $\psi(\vec{r}, t)$, onde $|\psi(\vec{r}, t)|^2$ descreve, de forma estatística, a probabilidade de se encontrar o sistema (uma partícula, por exemplo), na posição \vec{r} no tempo t . Utilizando coordenadas esféricas, esta teoria descreve muito bem a física de moléculas e sólidos, postulando-se a existência do spin do elétron.

A mecânica quântica de Schroedinger passou ainda por diversos desenvolvimentos, incluindo a física relativística até a quantização de campos e os atuais modelos de unificação, mas isto está fora do escopo desta disciplina.

A Estrutura de uma Teoria

Como ponto de partida, vamos fazer um experimento imaginário, similar ao proposto por J. J. Sakurai em seu livro “Modern Quantum Mechanics”, e explorado em aula pelo Prof. Allan Adams, disponível no *MIT Open Course*¹. Tomemos um objeto O com uma propriedade que chamaremos LETRA, que pode ser A ou B , e outra que chamaremos NÚMERO, que pode ser 1 ou 2. Vamos definir que O tem 50% de chance de ser A e 50% de ser B ; tem também 50% de chance de ser 1 e 50% de ser 2; não há correlação estatística na distribuição entre LETRA e NÚMERO; por fim, o objeto está em estado estacionário, ou seja, ele não evolui no tempo. Suponha que tenhamos vários destes objetos O , todos respeitando as premissas estipuladas, para realizarmos uma sequência de testes:

$$1) \quad O \rightarrow \boxed{\text{LETRA}} \rightarrow A \rightarrow \boxed{\text{LETRA}} \rightarrow ? \quad (1)$$

$$2) \quad O \rightarrow \boxed{\text{LETRA}} \rightarrow A \rightarrow \boxed{\text{NÚMERO}} \rightarrow ? \quad (2)$$

$$3) \quad O \rightarrow \boxed{\text{LETRA}} \rightarrow A \rightarrow \boxed{\text{NÚMERO}} \rightarrow 1 \rightarrow \boxed{\text{LETRA}} \rightarrow ? \quad (3)$$

O que podemos dizer sobre o resultado destes testes? No primeiro teste, como A foi encontrado no primeiro procedimento e o sistema está em estado estacionário, então o segundo procedimento tem 100% de chance de encontrar A . No segundo teste, A foi encontrado no primeiro procedimento e o segundo procedimento questiona sobre a propriedade NÚMERO. Pelas premissas, temos 50% de chance de encontrar 1 e 50% de chance de encontrar 2. No terceiro teste, inicialmente A foi encontrado e, na sequência, este grupo foi questionado segundo NÚMERO, encontrando 1. Em seguida, o resultado novamente questiona os objetos segundo LETRA. Qual a LETRA ao final? Segundo a mecânica quântica, a resposta é: depende da natureza das propriedades LETRA e NÚMERO do objeto O e dos processos de questionamento, que se forem bem estipulados, fornecerão valores de probabilidade bem determinados. Você pode protestar: “Mas já havíamos encontrado A , então temos 100% de certeza de que teremos A ao final!” e você estaria errado. Os abstratos LETRAS e NÚMEROS aqui podem representar conceitos físicos que acreditamos ter clareza, como energia, posição, velocidade, momento angular, mas para dar fim aos protestos, é preciso reconstruir a teoria e os conceitos pela base.

Do início: toda teoria é fundamentada em proposições aceitas como verdadeiras, sem demonstração, e que servem de base para o seu desenvolvimento. Estas proposições são chamadas Postulados, ou Axiomas, o primeiro mais utilizado em teorias físicas e o segundo em matemática. Nem na matemática,

¹<https://www.youtube.com/watch?v=lZ3bPUKo5zc>

considerada hoje por alguns a rainha das ciências², existe a fé³ em uma completude e autoconsistência teórica. Esta fé foi demolida pelos teoremas da incompletude demonstrados por Kurt Gödel em 1931, que trouxeram a concepção de não ser possível encontrar um conjunto completo e consistente de axiomas para toda a aritmética. Seu segundo teorema da incompletude estipula que “Uma teoria, recursivamente enumerável e capaz de expressar verdades básicas da aritmética e alguns enunciados da teoria da prova, pode provar sua própria consistência se, e somente se, for inconsistente”.

A Mecânica Quântica, como toda teoria, é construída fundamentada em postulados. Diferentes livros texto trazem enunciados diferentes para os postulados, e mesmo um número de postulados diferentes pois, com a evolução do conhecimento, os postulados são revistos, demonstra-se que alguns estão contidos em outros, outros podem ser fundidos em um mais forte. Como exemplo, o livro Quantum Mechanics [1] (Cohen-Tannoudji, Diu e Laloë, 1977) enumera 6 postulados para a construção da mecânica quântica, a Equação de Schrödinger para evolução temporal dos sistemas físicos sendo um deles. Na sua concepção mais moderna, como encontrado no livro de Ballentine (2^a Ed., 2015) [2], são necessários apenas dois postulados bastante básicos para fundamentar toda a teoria:

1. Para cada variável dinâmica existe um operador hermitiano cujos autovalores representam os valores possíveis para a variável dinâmica.
2. Para cada estado existe um operador de estado único ρ , hermitiano, não-negativo e de traço unitário. O valor médio do observável R no estado ρ é dado por $\langle R \rangle = \text{Tr}(\rho R) / \text{Tr}(\rho)$.

Os conceitos citados nestes postulados pertencem a dois ramos da matemática, a álgebra linear e a probabilidade, que são pré-requisitos para o desenvolvimento da mecânica quântica. O primeiro postulado associa a cada atributo físico (aqui chamado de *variável dinâmica*) de um sistema, um objeto matemático (aqui chamado de operador hermitiano). E, ainda, associa os possíveis valores que podem ser medidos desse atributo físico com os autovalores desse operador. Já o segundo postulado associa ao *estado* de um sistema (note a sutil distinção entre *estado* e *variável dinâmica* aqui construída) um outro operador, único, de características peculiares: ser hermitiano, não-negativo e de traço unitário. Estes conceitos devem ser introduzidos e trabalhados para se construir a física quântica. Isto será feito ao longo deste texto.

Como exemplo, a equação de Schrödinger emerge destes dois postulados evocando-se condições de simetria do espaço-tempo, não precisa ser postulada.

²Título que um dia já foi da teologia, amanhã... quem sabe?

³Intrigante, mas a palavra correta é mesmo “fé”.

Outro exemplo, considere a posição de uma partícula. Classicamente posição é uma função do tempo (geralmente $x(t)$) que é solução da segunda lei de Newton. Os valores possíveis dessa variável dinâmica são todos os valores nos quais a função está definida. Quanticamente, entretanto, associado a esse ente físico posição existe um operador, chamado de operador posição, cujos valores que a posição da partícula poderão assumir serão apenas os autovalores desse operador. A relação deste operador com outros, como o operador momento, o operador energia, define o resultado dos experimentos.

Em Busca das Bases

No capítulo 1 lembraremos alguns aspectos e maquinário matemático importantes para se trabalhar com mecânica quântica, ligados aos pilares fundamentais da teoria moderna, que são a **álgebra linear** e a **teoria da probabilidade**. Mas primeiro façamos um exercício para uma compreensão mais profunda dos fundamentos da teoria. Sendo a álgebra linear um pilar, podemos dar um passo atrás e nos perguntar “o que é **álgebra**?” A palavra vem do árabe “al-jerb”, que significa “reunião de pedaços”, e denota o estudo de símbolos e das regras para manipulá-los. Como exemplo de um campo de aplicação da álgebra, vejamos a teoria de grupos [3], importante para lidar com operações de simetria.

Grupos

Uma coleção de elementos A, B, C, \dots forma um grupo quando as seguintes condições são satisfeitas:

- O produto de quaisquer dois elementos de um grupo é, ele próprio, um elemento do grupo. Por exemplo, relações do tipo $AB = C$ são válidas se A, B e C forem membros do mesmo grupo.
- A lei associativa é válida, isto é, $(AB)C = A(BC)$.
- Existe um elemento unidade E (também chamado identidade) tal que o produto de E com qualquer elemento do grupo deixa o elemento invariante: $AE = EA = A$.
- Para todo elemento A existe um elemento inverso A^{-1} tal que $A^{-1}A = AA^{-1} = E$.

Em geral os elementos do grupo não **comutam**, isto é, a ordem do produto dos elementos altera o resultado: $AB \neq BA$. Se todos os elementos do grupo comutarem, o grupo é chamado de grupo **Abeliano**.

Um exemplo simples de grupo é o conjunto das operações de permutação de três números, $P(3)$. Abaixo estão listadas as $3!=6$ possíveis formas de se permutar estes três elementos. A linha de cima denota o estado inicial, e a linha de baixo o estado dos números após a permutação. Cada permutação é um elemento do grupo $P(3)$.

$$\begin{aligned} E &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} & A &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} & B &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \\ C &= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} & D &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} & F &= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \end{aligned} \quad (4)$$

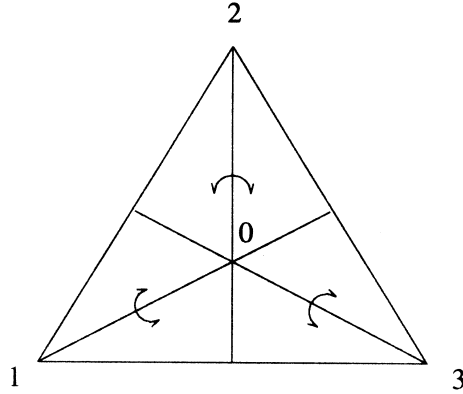


Figura 1: As operações de simetria em um triângulo equilátero são as rotações por $\pm 2\pi/3$ sobre a origem 0 e as rotações por π sobre os três eixos. Aqui os eixos são indicados por números em círculos.

Podemos também pensar nos elementos da Eq. 4 em termos dos 3 pontos de um triângulo equilátero (veja Fig. 1). Novamente, a linha superior indica o estado inicial e a linha inferior denota a posição final de cada número como o efeito das seis operações de simetria distintas que podem ser realizadas nesses três pontos (veja a legenda da Fig. 1). Nós podemos chamar cada operação de simetria de um **elemento** do grupo. Este grupo é, portanto, idêntico ao grupo para as operações de simetria em um triângulo equilátero, como mostrado na Fig. 1. O elemento D é uma rotação de $-2\pi/3$ (no sentido anti-horário) e F é uma rotação de $+2\pi/3$ (no sentido horário).

O que chamamos de produtos dos elementos do grupo pode ser definido e classificado:

$$\begin{aligned} (AB)C &= DC = B \\ A(BC) &= AD = B \end{aligned} \quad (5)$$

verificando a validade da **lei associativa** $(AB)C = A(BC)$ para estes elementos. Cada elemento do grupo de permutação $P(3)$ tem correspondência um

para um com as operações de simetria de um triângulo e, portanto, dizemos que estes dois grupos são **isomórficos**.

Para lidar com problemas físicos podemos associar grupos isomórficos e trabalhar com o grupo mais simples. Se podemos associar cada elemento de um grupo com uma matriz que obedece à mesma lógica de multiplicação dos elementos do grupo de interesse, isto é, se os elementos obedecem a $AB = D$, então as matrizes representando os elementos devem obedecer

$$M(A) M(B) = M(D). \quad (6)$$

Se esta relação é satisfeita, então podemos realizar todas as operações do grupo analiticamente em termos de operações aritméticas em matrizes, geralmente mais fáceis de executar. A identificação um-para-um de uma operação genérica de um grupo para uma operação com uma matriz é a ideia básica de uma **representação**.

O conjunto de matrizes abaixo satisfazem todas as relações de multiplicação para o grupo $P(3)$:

$$\begin{aligned} E &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & A &= \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} & B &= \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \\ C &= \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} & D &= \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} & F &= \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \end{aligned} \quad (7)$$

Observamos que a matriz correspondente à operação de identidade é sempre uma matriz unitária. As matrizes na Eq. 7 constituem uma representação de matriz do grupo que é isomorfo a $P(3)$ e às operações de simetria em um triângulo equilátero. A matriz A representa uma rotação de $\pm\pi$ sobre o eixo y enquanto as matrizes B e C , respectivamente, representam rotações de $\pm\pi$ sobre os eixos 2 e 3 na Fig. 1, e assim por diante.

Campos

Campo é outro conceito comum em física, e em matemática é definido como um conjunto de elementos que formam um grupo comutativo com respeito a duas operações compatíveis, a adição e a multiplicação, onde “compatível” é formalizado pelo conceito de “distributividade”, adicionando o detalhe de que o elemento identidade da adição (o zero), não tem inversa na multiplicação. A teoria para campos está fundamentada nos seguintes axiomas:

- fechamento para a adição e a multiplicação: $u + v = w$, $uv = w$, onde u, v, w pertencem ao campo;

- associatividade para a adição e a multiplicação: $(u+v)+w = u+(v+w)$, $(uv)w = u(vw)$;
- comutatividade para a adição e a multiplicação: $u+v = v+u$, $uv = vu$;
- contém identidade para a adição (0) e a multiplicação (1);
- contém inversa para a adição ($-a$) e a multiplicação (a^{-1}).
- distributividade da multiplicação sobre a adição: $a \cdot (a+b) = a \cdot b + a \cdot c$

Considerações finais

Tal procedimento de buscar as bases pode ser feito até chegar-se nos conceitos e definições básicas da teoria de números. Mas a mensagem que deve ficar aqui é que os conceitos que utilizaremos na física (como um grupo, operações de simetrias, campos... seguindo até energia, momento, posição) têm sua fundamentação em postulados e lógica matemática.

Capítulo 1

Pré-requisitos matemáticos

As estruturas que comentamos na seção anterior, onde falamos sobre operadores e estados, envolvem conceitos matemáticos que ainda precisamos construir (ou relembrar). Um dos pilares fundamentais da mecânica quântica moderna é a **álgebra linear**. O outro é a **teoria da probabilidade**, que será vista na sequência.

1.1 Álgebra Linear

A álgebra linear especifica as propriedades das equações lineares, espaços vetoriais e matrizes. Equações lineares são, por sua vez, equações algébricas que possuem constantes, o produto de uma constante por uma variável em primeira ordem de potência (sem expoente), e combinações de constantes e variáveis. Se $\psi_1(x)$ e $\psi_2(x)$ são soluções de uma equação linear, qualquer combinação linear das duas, $\psi_3(x) = c_1\psi_1(x) + c_2\psi_2(x)$, com c_1 e c_2 sendo constantes complexas, também é solução. Os conceitos de espaços vetoriais e matrizes serão vistos mais adiante.

1.1.1 Espaços vetoriais lineares

Consideremos um conjunto de elementos, que são objetos abstratos aqui representados pelo símbolo $|\rangle$ (já inserindo a notação de Dirac); para distinguir estes objetos, podemos dar a eles rótulos, como $|1\rangle$, $|a\rangle$, $|\alpha\beta\rangle$.

Um **espaço vetorial linear** \mathbb{V} é qualquer conjunto de elementos $|\rangle$ (que podem números, funções, matrizes, polinômios, etc.) munidos de duas operações, que chamaremos de adição de elementos e multiplicação de elemento por escalar, que obedecem às seguintes propriedades [4].

- Associatividade para a adição: $(|u\rangle + |v\rangle) + |w\rangle = |u\rangle + (|v\rangle + |w\rangle)$,

- Comutatividade para a adição: $|u\rangle + |v\rangle = |v\rangle + |u\rangle$,
- Contém elemento identidade para adição: $|u\rangle + \mathbf{0} = |u\rangle$,
- Contém elemento inverso para adição: $|u\rangle + |-u\rangle = \mathbf{0}$,
- Contém elemento identidade para multiplicação escalar: $\mathbf{1}|u\rangle = |u\rangle$,
- Compatibilidade entre multiplicação escalar e elementos: $(ab)|v\rangle = a(b|v\rangle)$,
- Distributividade da multiplicação escalar sobre a adição: $a(|u\rangle + |v\rangle) = a|u\rangle + a|v\rangle$,
- Distributividade da soma de escalares com respeito aos elementos: $(a+b)|v\rangle = a|v\rangle + b|v\rangle$,

em que $|u\rangle$, $|v\rangle$, $|w\rangle$ são os elementos de \mathbb{V} , enquanto a e b representam escalares, que podem ser números complexos. Os elementos de um espaço vetorial são chamados vetores de forma ampla. Exemplos comuns de espaços vetoriais são o de vetores discretos \mathbb{C}^n (cada vetor é composto por n números complexos), o espaço de matrizes, o espaço de polinômios e o espaço de funções diferenciáveis de uma variável.

Espaços vetoriais possuem conjuntos de vetores que podem ser usados para gerar o espaço inteiro. Esses conjuntos são chamados de **bases** do espaço vetorial e o número mínimo de vetores necessários para compor uma base é chamado de **dimensão** do espaço vetorial.

Exemplo 1 *Os seguintes vetores formam uma base no espaço \mathbb{R}^3 :*

$$|\hat{i}\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad |\hat{j}\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad |\hat{k}\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \quad (1.1)$$

Note que o espaço é infinito, mas sua dimensão é 3.

Além da hipótese de gerar o espaço inteiro, i.e. esse conjunto ser capaz de gerar todos os outros elementos, há ainda uma hipótese adicional. Suponha que $|\phi_n\rangle$ são n elementos da base de um espaço vetorial. A combinação linear:

$$\sum_n c_n |\phi_n\rangle = \mathbf{0} \Leftrightarrow c_n = 0 \forall n, \quad (1.2)$$

ou seja, se a soma só puder ser nula se, e somente se, todos os coeficientes forem zero, então dizemos que os vetores $|\phi_n\rangle$ são todos **linearmente independentes** (L.I., abreviadamente).

\Rightarrow *Resolva os exercícios 1.1 e 1.2.*

1.1.2 Produto interno (ou produto escalar)

O conceito de espaço vetorial é genérico e serve para descrever coisas diversas. Qualquer conjunto de elementos abstratos que satisfaça as propriedades de espaço vetorial, é considerado um espaço vetorial.

É possível, entretanto, incluir outras estruturas que restringem o conceito de espaço, levando a estruturas que são importantes em física, como o **produto interno** ou **produto escalar**, que associa um escalar à operação $\langle u|v\rangle$, tal que é um número complexo que segue os seguintes axiomas:

- Simetria Hermitiana: $\langle u|v\rangle = \langle v|u\rangle^*$, com o super-índice * indicando o complexo conjugado.
- Linearidade: se $|v\rangle = a|v_1\rangle + b|v_2\rangle$, $\langle u|v\rangle = a\langle u|v_1\rangle + b\langle u|v_2\rangle$.
- Positividade: $\langle u|u\rangle \geq 0$, com a igualdade apenas no caso de $|u\rangle = 0$

Considere dois vetores de entradas complexas num espaço vetorial finito de n dimensões:

$$|\psi\rangle = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \quad |\phi\rangle = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}. \quad (1.3)$$

O produto interno será definido como sendo:

$$\langle \psi|\phi\rangle = a_1^*b_1 + a_2^*b_2 + \cdots + a_n^*b_n, \quad (1.4)$$

Se $\langle \psi|\phi\rangle = \langle \phi|\psi\rangle = 0$ dizemos que os vetores são ortogonais entre si.

Note que o primeiro e o terceiro axiomas do produto interno, juntos, implicam que $\langle u|u\rangle$ seja real. Isso nos permite definir, a partir do produto interno, a **norma** de um vetor:

$$\text{norma } |\phi\rangle = \sqrt{\langle \phi|\phi\rangle}. \quad (1.5)$$

No espaço tridimensional essa norma está associada diretamente ao “comprimento” do vetor. Em espaços abstratos ela dá uma ideia desse mesmo comprimento ou da sua “magnitude”. É comum vermos o uso da operação de norma com dois traços verticais $|||\phi\rangle||$.

1.1.3 Base de um estado vetorial

Se um dado vetor possuir norma unitária, dizemos que ele é **normalizado**. Um conjunto de vetores que são todos ortogonais entre si e possuem norma unitária são chamados de um conjunto **ortonormal** de vetores.

Todo vetor $|v\rangle$ pode ser expandido em termos um conjunto ortonormal completo $\{\phi_i\}$, da seguinte forma:

$$|v\rangle = \sum_i v_i |\phi_i\rangle. \quad (1.6)$$

Damos a $\{\phi_i\}$ o nome de **base do estado vetorial**. Um resultado importante de vetores ortogonais está expresso pelo teorema a seguir:

Teorema 1 *Um conjunto de vetores não nulos ortogonais dois a dois, i.e. $\langle v_i | v_j \rangle = 0$ se $i \neq j$. Então o conjunto formado por todos os elementos $|v_i\rangle$ é L.I.*

Prova: Tomando o valor esperado da combinação linear da equação (1.2) de ambos lados e usando a condição de ortogonalidade o resultado sai diretamente.

Em outras palavras esse teorema estabelece que se um conjunto de vetores é ortogonal, eles são também L.I. Esse resultado é importante para se testar a independência linear de conjuntos que são difíceis de se aplicar diretamente a equação (1.2). Para mais detalhes, veja exercício 1.3.

A operação do produto interno também está definida em espaços de funções (podendo ser não finitas) usando o processo de integração. Sejam duas funções $\psi(x)$ e $\phi(x)$ mensuráveis e uma função peso $w(x)$ não negativa. O produto interno de ψ com ϕ é dado por:

$$(\psi, \phi) = \int_{\text{todo espaço}} \psi^*(x) \phi(x) w(x) dx, \quad (1.7)$$

\implies Resolva o exercício 1.3.

1.1.4 Espaço Euclidiano e Espaço de Hilbert

Espaços vetoriais que possuem a estrutura adicional do produto interno (ou produto escalar) são chamados genericamente de **espaços vetoriais com produto interno**. O espaço euclidiano é um espaço vetorial de dimensão finita munido de produto interno.

O espaço de Hilbert é uma generalização do espaço euclidiano, que não precisa estar restrito a um número finito de dimensões. Composto por todos os vetores $|h\rangle = \sum_n c_n |\phi_n\rangle$, escritos aqui na base ortonormal $|\phi_n\rangle$, com a limitação de que a norma de $|h\rangle$ $[\sqrt{\langle h|h\rangle} = \sqrt{\sum_n |c_n|^2}]$ é finita. Na grande maioria dos casos, a mecânica quântica é descrita no espaço de Hilbert, embora alguns vetores muito importantes, como o vetor posição, não estejam limitados a ele por poderem ter norma infinita.

1.1.5 Desigualdades de Schwarz e do triângulo

Um resultado importante sobre norma de vetores é conhecida como **desigualdade de Schwarz**:

$$|\langle \psi | \phi \rangle|^2 \leq \langle \psi | \psi \rangle \langle \phi | \phi \rangle. \quad (1.8)$$

A igualdade vale quando $|\psi\rangle$ e $|\phi\rangle$ são colineares.

\implies Resolva o exercício 1.4.

Outro resultado útil sobre normas é a desigualdade triangular, ou do triângulo. Seja $|\psi + \phi\rangle = |\psi\rangle + |\phi\rangle$, então:

$$\sqrt{\langle \psi + \phi | \psi + \phi \rangle} \leq \sqrt{\langle \psi | \psi \rangle} + \sqrt{\langle \phi | \phi \rangle}. \quad (1.9)$$

A desigualdade do triângulo é traduzida considerando que se $\rho(a, b)$ é a “distância” entre a e b , então $\rho(a, b) + \rho(b, c) \leq \rho(a, c)$, que no espaço euclidiano forma um triângulo, e a igualdade é válida se ab e bc são colineares (na Eq. 1.9, se $|\psi\rangle$ e $|\phi\rangle$ são colineares). Aplicando esta tradução para o espaço vetorial linear em geral, estende-se o conceito de distância, hoje comum em métodos computacionais de aprendizado de máquina, utilizados por exemplo por plataformas de oferta embasada em histórico de busca por músicas ou filmes. Isto significa que o espaço linear é um espaço métrico.

\implies Resolva o exercício 1.5.

Como veremos mais a frente, estas duas desigualdades estão relacionadas a alguns efeitos que distinguem, de forma quantitativa, as predições da mecânica quântica das da mecânica clássica.

1.1.6 Matrizes

Matrizes são arranjos retangulares de objetos matemáticos, para os quais operações como adição e multiplicação estão definidos. Exemplo: matrizes em um campo \mathbf{F} são arranjos retangulares de escalares de \mathbf{F} . Podem ser números reais ou complexos. A convenção para se montar uma matriz A na base $\{|e_i\rangle\}$, tal que

$$A |e_1\rangle = a_{11} |e_1\rangle + a_{12} |e_2\rangle + \dots + a_{1n} |e_n\rangle \quad (1.10)$$

$$A |e_2\rangle = a_{21} |e_1\rangle + a_{22} |e_2\rangle + \dots + a_{2n} |e_n\rangle \quad (1.11)$$

$$\vdots \quad (1.12)$$

$$A |e_m\rangle = a_{m1} |e_1\rangle + a_{m2} |e_2\rangle + \dots + a_{mn} |e_n\rangle \quad (1.13)$$

é:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix}. \quad (1.14)$$

É possível aplicar A em $|e_i\rangle$ mas não mapeá-la nesta base, pois não é necessário ficar no mesmo espaço. Entretanto, esta é a forma mais corriqueira.

Matrizes quadradas são obtidas quando o número de equações são iguais ao número de variáveis ($n = m$). Um exemplo interessante (e importante!) são as matrizes 2×2 . Para esse caso, faz-se a pergunta: *O conjunto de todas as matrizes dois por dois possui qual dimensão?* Como seria uma base nesse espaço?

Um exemplo trivial é a chamada *base canônica*

$$\mu_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \mu_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \mu_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \mu_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

já que qualquer matriz M 2×2 pode ser escrita como

$$M = m_{11}\mu_1 + m_{12}\mu_2 + m_{21}\mu_3 + m_{22}\mu_4.$$

Entretanto, a base mais importante no espaço das matrizes 2×2 é composta pelas chamadas **matrizes de Pauli**:

$$\begin{aligned} \sigma_0 = \mathbf{1} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mu_1 + \mu_4, \\ \sigma_1 = \sigma_x &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \mu_2 + \mu_3, \\ \sigma_2 = \sigma_y &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = -i(\mu_2 - \mu_3), \\ \sigma_3 = \sigma_z &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \mu_1 - \mu_4. \end{aligned} \quad (1.15)$$

É fácil verificar que $\text{Tr}(\sigma_i \sigma_j) = 2\delta_{ij}$. Assim, qualquer matriz M 2×2 pode ser escrita como

$$M = a_0 \mathbf{1} + a_1 \sigma_x + a_2 \sigma_y + a_3 \sigma_z,$$

com $a_i = \text{Tr}(M \sigma_i)/2$.

Veremos mais adiante que as matrizes de Pauli são de grande importância para tratar sistemas de dois níveis na mecânica quântica, como o momento angular de spin $\frac{1}{2}$ ou a polarização de um fóton. Interessante que nestes casos, teremos um espaço vetorial de matrizes, com a base de vetores $|\mathbf{1}\rangle = \mathbf{1}$, $|\sigma_x\rangle = \sigma_x$, $|\sigma_y\rangle = \sigma_y$ e $|\sigma_z\rangle = \sigma_z$.

1.1.7 Sobre a notação de Dirac

Damos o nome de **notação de Dirac** (ou **notação bra-ket**) à forma, inventada por Paul Dirac, de se representar os objetos matemáticos em mecânica quântica por $|\phi\rangle$, o que nos permite operá-los mais facilmente.

Aos elementos $|\phi\rangle$ damos o nome de **kets**. Existem elementos, chamados de **bras**, associados aos kets por isomorfismo (relação de um pra um, de acordo com o **teorema da representação de Riesz em espaços de Hilbert**). Esses elementos existem em um espaço diferente, chamado de **espaço dual**, que é o espaço de todos os funcionais lineares. Como há uma correspondência de um pra um, o bra associado ao ket $|\psi\rangle$ é denotado utilizando-se o mesmo rótulo, $\langle\psi|$. Existem casos onde esta correspondência não está presente, e espaços mais gerais do que o espaço de Hilbert devem ser considerados.

O produto interno consiste então do produto de um bra com um ket (nessa ordem):

$$\langle\psi|\phi\rangle = (\psi, \phi). \quad (1.16)$$

Q quando lidando com vetores discretos de dimensão finita n , a equação 1.16 resume-se à equação 1.4.

Note que a simetria Hermitiana do produto interno ($\langle u|v\rangle = \langle v|u\rangle^*$) faz com que bras e kets sejam anti-lineares. Explicando: se $|\psi\rangle = a_1|u_1\rangle + a_2|u_2\rangle$, então $\langle\phi|\psi\rangle = a_1\langle\phi|u_1\rangle + a_2\langle\phi|u_2\rangle$; entretanto, se $|\phi\rangle = a_1|u_1\rangle + a_2|u_2\rangle$, então

$$\begin{aligned} \langle\phi|\psi\rangle &= \langle\psi|\phi\rangle^* \\ &= [a_1\langle\psi|u_1\rangle + a_2\langle\psi|u_2\rangle]^* \\ &= a_1^*\langle\psi|u_1\rangle^* + a_2^*\langle\psi|u_2\rangle^* \\ &= a_1^*\langle u_1|\psi\rangle + a_2^*\langle u_2|\psi\rangle. \end{aligned} \quad (1.17)$$

ou seja, $\langle\phi|\psi\rangle$ é linear com respeito a $|\psi\rangle$ e anti-linear com respeito a $|\phi\rangle$. Logo, se o ket $|\psi\rangle = a_1|u_1\rangle + a_2|u_2\rangle$, o bra correspondente é dado por $\langle\psi| = a_1^*\langle u_1| + a_2^*\langle u_2|$.

Como já explicitado anteriormente, em muitos casos os kets podem ser representados por matrizes quadradas e seus bras equivalentes por operações de transposição de conjugação complexa, reduzindo o problema do cálculo de produto interno a um simples cálculo de produto de matrizes. Ou podem ser funções, e neste caso temos uma integral. Retornaremos a estes pontos ao longo do curso.

1.2 Operadores Lineares

Operadores são objetos matemáticos que mapeiam vetores (elementos de algum espaço vetorial \mathbb{V}) em outros vetores (que podem ser elementos tanto de V quanto de qualquer outro espaço vetorial \mathbb{W}). Os operadores são definidos de acordo com a forma que eles atuam em vetores de algum espaço vetorial:

$$|\phi\rangle = A|\psi\rangle, \quad (1.18)$$

sendo A o operador, $|\phi\rangle$ e $|\psi\rangle$ elementos de dois espaços vetoriais lineares arbitrários na notação de Dirac. Dois operadores A e B são iguais se $A|\phi\rangle = B|\phi\rangle$, para todo $|\phi\rangle$ no domínio comum de A e B .

Damos o nome de **operador linear** se ele satisfizer:

$$A(c_1|\psi_1\rangle + c_2|\psi_2\rangle) = c_1(A|\psi_1\rangle) + c_2(A|\psi_2\rangle).^1 \quad (1.19)$$

Além disso, operadores também satisfazem às propriedades:

$$(A + B)|\psi\rangle = A|\psi\rangle + B|\psi\rangle, \quad (1.20)$$

$$AB(|\psi\rangle) = A(B|\psi\rangle), \quad (1.21)$$

$$(AB)C = A(BC). \quad (1.22)$$

OBS: A propriedade de comutação não é exigida para operadores, sendo permitido que $AB \neq BA$. Em geral operadores que comutam possuem características inexistentes em operadores que não comutam. Falaremos deles mais adiante.

\implies *Resolva o exercício 1.6. Note que os vetores podem ser escritos como $|\vec{v}\rangle = 3|\hat{i}\rangle + 2|\hat{j}\rangle + |\hat{k}\rangle$ e $|\vec{u}\rangle = |\hat{i}\rangle + 2|\hat{j}\rangle + 3|\hat{k}\rangle$, sendo este um exemplo de um problema de espaço vetorial discreto \mathbb{V} de dimensão finita $n = 3$. Os elementos (vetores) desse espaço podem ser denotados por matrizes coluna com n entradas, e os operadores que levam elementos de \mathbb{V} em elementos do mesmo espaço podem ser descritos por matrizes quadradas de dimensão $n \times n$. Além disso qualquer equação de operadores em espaços de dimensão finita podem ser transformados em operações matriciais.*

Exemplo 2 (Operadores tipo rotação em 3D) *Sejam A uma matriz de rotação de 180° em torno do eixo y cartesiano e B a matriz de rotação de 120° no plano xy . Essas rotações podem ser representadas por matrizes 3×3 :*

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad B = \begin{pmatrix} \text{sen } 120^\circ & \text{cos } 120^\circ & 0 \\ \text{cos } 120^\circ & -\text{sen } 120^\circ & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (1.23)$$

¹Um operador anti-linear é tal que $A(c_1|\psi_1\rangle + c_2|\psi_2\rangle) = c_1^*(A|\psi_1\rangle) + c_2^*(A|\psi_2\rangle)$. Serão relevantes mais pra frente, mas por enquanto não trataremos deles.

Veja a conexão entre estas matrizes e aquelas da Eq. 7 da Introdução.

De forma mais genérica, pode-se definir a chamada **operação de similaridade** (ou equivalência, ou canônica), que equivale a uma transformação da matriz M do tipo UMU^{-1} , onde U é uma matriz unitária e U^{-1} a sua inversa ($UU^{-1} = U^{-1}U = I$, onde $I =$ matriz identidade). Como exemplo, as matrizes rotação do exemplo 3 podem ser utilizadas para rodar os eixos de referências de um operador.

1.2.1 Notação matricial para operadores

Podemos expandir os dois vetores em uma base comum ortonormal $|u_i\rangle$, tal que:

$$|\psi\rangle = \sum_j a_j |u_j\rangle, \quad |\phi\rangle = \sum_k b_k |u_k\rangle, \quad (1.24)$$

substituindo na equação de antes encontramos:

$$M \sum_j a_j |u_j\rangle = \sum_k b_k |u_k\rangle \Rightarrow \sum_j a_j M |u_j\rangle = \sum_k b_k |u_k\rangle, \quad (1.25)$$

multiplicando ambos lados pelo bra $\langle u_i|$ encontramos:

$$\sum_j a_j \langle u_i | M | u_j \rangle = \sum_k b_k \langle u_i | u_k \rangle, \quad (1.26)$$

mas, por hipótese:

$$\langle u_i | u_k \rangle = \delta_{ik}, \quad (1.27)$$

então encontramos:

$$\sum_j a_j \langle u_i | M | u_j \rangle = \sum_j a_j M_{ij} = b_i, \quad (1.28)$$

em que $M_{ij} = \langle u_i | M | u_j \rangle$ são chamados **elementos de matriz** que compõem a notação matricial do operador M . Esse procedimento nos permite construir representações matriciais de operadores em espaços de dimensão finita (a ideia pode ser estendida para operadores em dimensão infinita, porém contável).

Exemplo 3 (Operadores em espaços de funções) Considere a equação de operadores:

$$\frac{\partial}{\partial x} x = 1 + x \frac{\partial}{\partial x}, \quad (1.29)$$

consiste num operador definido em um espaço de funções. Ou seja, esse operador (que transforma a derivada em relação a x de uma função multiplicada

por x em uma soma da própria função com a sua derivada em relação a x multiplicada por x) existe se a equação a seguir valer para qualquer ψ :

$$O \equiv \frac{\partial}{\partial x} [x\psi(x)] = \psi(x) + x \frac{\partial \psi(x)}{\partial x}, \quad \forall \psi(x). \quad (1.30)$$

\implies Resolva o exercício 1.7.

Note que a Equação de Schroedinger

$$\frac{-\hbar}{2m} \frac{\partial^2}{\partial x^2} + V(x) = i\hbar \frac{\partial}{\partial t}, \quad (1.31)$$

é uma equação de operadores.

1.2.2 Traço de matrizes

Traço de uma matriz é definido como a soma dos elementos da diagonal principal. Em notação de Dirac, escrevemos:

$$\text{Tr } A = \sum_j \langle u_j | A | u_j \rangle. \quad (1.32)$$

O traço satisfaz a importante propriedade de ser independente da base, ou seja, ele é invariante sob qualquer mudança de base

\implies Resolva o exercício 1.8.

1.2.3 Operador agindo em um bra

Considere a atuação de um operador A em um ket:

$$A |\phi\rangle = |\psi\rangle, \quad (1.33)$$

tomando a operação “adjunta” (\dagger) de ambos os lados, que consiste na conjugação complexa e transposição, obtemos:

$$\langle \phi | A^\dagger = \langle \psi |, \quad (1.34)$$

essa é uma consequência do teorema da representação de Riesz que comentamos anteriormente. Ainda, da simetria Hermitiana do produto interno, $\langle \psi | \phi \rangle^* = \langle \phi | \psi \rangle$, temos que:

$$\langle \phi | A^\dagger |\psi\rangle^* = \langle \psi | A |\phi\rangle. \quad (1.35)$$

1.2.4 Produto externo

O produto interno, que comentamos anteriormente, gera um número complexo a partir de dois vetores. Porém existe um outro tipo de produto, conhecido como **produto externo**, cujo resultado da sua operação é um operador. Na notação de Dirac, pode ser denotado por:

$$|\phi\rangle\langle\psi| = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_n \end{pmatrix} (\psi_1^* \ \psi_2^* \ \cdots \ \psi_n^*) = \begin{pmatrix} \phi_1\psi_1^* & \phi_1\psi_2^* & \cdots & \phi_1\psi_n^* \\ \phi_2\psi_1^* & \phi_2\psi_2^* & \cdots & \phi_2\psi_n^* \\ \vdots & \vdots & \ddots & \vdots \\ \phi_n\psi_1^* & \phi_n\psi_2^* & \cdots & \phi_n\psi_n^* \end{pmatrix}, \quad (1.36)$$

que satisfaz à propriedade:

$$(|\psi\rangle\langle\phi|)^\dagger = |\phi\rangle\langle\psi|. \quad (1.37)$$

\implies Resolva o exercício 1.9.

1.2.5 Produto tensorial e de Kronecker

Sejam dois espaços vetoriais \mathbb{U} e \mathbb{V} , com os elementos $|U_k\rangle = \sum_{i=1}^m \alpha_k^i |u_i\rangle \in \mathbb{U}$ e $|V_k\rangle = \sum_{j=1}^n \beta_k^j |v_j\rangle \in \mathbb{V}$ descritos nas respectivas bases ortonormais $\{|u_i\rangle\}$ e $\{|v_j\rangle\}$. O produto tensorial $\mathbb{U} \otimes \mathbb{V}$ (que se lê “ \mathbb{U} tensor \mathbb{V} ”), gera um espaço bilinear \mathbb{W} com elementos

$$w(u, v) = \sum_k |U_k\rangle \otimes |V_k\rangle = \sum_{i,j,k} \alpha_k^i \beta_k^j w(u_i, v_j) \quad (1.38)$$

sendo $\{w(u_i, v_j) = |u_i\rangle \otimes |v_j\rangle\}$ (onde $1 \leq i \leq m$ e $1 \leq j \leq n$) a base ortonormal do espaço resultante \mathbb{W} de dimensão $d = nm$. Os elementos $w(u, v)$ do espaço vetorial \mathbb{W} são tensores de segunda ordem, que podem ser representados matricialmente por

$$w(u, v) = \sum_k \begin{bmatrix} \alpha_k^1 \beta_k^1 & \cdots & \alpha_k^1 \beta_k^n \\ \vdots & \ddots & \vdots \\ \alpha_k^m \beta_k^1 & \cdots & \alpha_k^m \beta_k^n \end{bmatrix}, \quad (1.39)$$

e são ditos decomponíveis.

Exemplo 4 (Produto tensorial de vetores) Qual o resultado do produto tensorial do vetor $|\vec{v}\rangle = 3|\hat{i}\rangle + 2|\hat{j}\rangle + |\hat{k}\rangle$ pelo vetor $|\vec{u}\rangle = |\hat{i}\rangle + 2|\hat{j}\rangle + 3|\hat{k}\rangle$? O resultado pode ser denotado por uma matriz 3×3 , dada por:

$$|\vec{v}\rangle \otimes |\vec{u}\rangle = \begin{pmatrix} 3 & 6 & 9 \\ 2 & 4 & 6 \\ 1 & 2 & 3 \end{pmatrix}. \quad (1.40)$$

Quando aplicado a matrizes, o produto tensorial leva o nome de produto de Kronecker, sendo uma operação em duas matrizes de tamanhos arbitrários resultando em uma matriz de blocos.² Se \mathbf{A} é uma matriz $m \times n$ e \mathbf{B} é uma matriz $p \times q$, então o produto de Kronecker $\mathbf{A} \otimes \mathbf{B}$ é a matriz de blocos $mp \times nq$:

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{bmatrix}. \quad (1.41)$$

Note que as componentes do resultado $\mathbf{C} = \mathbf{A} \otimes \mathbf{B}$ têm 4 índices, $c_{ij,kl}$, dois vindos da componente específica de \mathbf{A} , e dois de \mathbf{B} .

Sendo um caso especial do produto tensorial, o produto Kronecker é bilinear e associativo:

$$\mathbf{A} \otimes (\mathbf{B} + \mathbf{C}) = \mathbf{A} \otimes \mathbf{B} + \mathbf{A} \otimes \mathbf{C}, \quad (1.42)$$

$$(\mathbf{A} + \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes \mathbf{C} + \mathbf{B} \otimes \mathbf{C}, \quad (1.43)$$

$$(k\mathbf{A}) \otimes \mathbf{B} = \mathbf{A} \otimes (k\mathbf{B}) = k(\mathbf{A} \otimes \mathbf{B}), \quad (1.44)$$

$$(\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C}), \quad (1.45)$$

onde \mathbf{A} , \mathbf{B} e \mathbf{C} são matrizes e k é um escalar.

A construção baseada em produtos tensoriais é fundamental em mecânica quântica, na formação de estados compostos (capítulo 3), está sobretudo na análise do momento angular (vide capítulo 7), na análise das medições (vide capítulos 8 e 9) e especialmente em informação quântica [5].

1.2.6 Espaço fator

Dado um espaço \mathbb{W} gerado pelo produto tensorial de dois outros espaços, $\mathbb{W} = \mathbb{W}^{(1)} \otimes \mathbb{W}^{(2)}$, os espaços originários $\mathbb{W}^{(1)}$ e $\mathbb{W}^{(2)}$ são chamados de **espaços fatores**. Sob a limitação de que $\mathbb{W}^{(1)}$ e $\mathbb{W}^{(2)}$ contêm apenas elementos de traço unitário, os elementos $W^{(1)}$ do espaço fator $\mathbb{W}^{(1)}$ podem ser obtido por

$$W^{(1)} = \text{Tr}^{(2)} W \quad (1.46)$$

onde $\text{Tr}^{(2)}$ é o traço no espaço fator da componente 2, ficando os elementos de $W^{(1)}$ definidos por

$$\langle w_m | W^{(1)} | w_{m'} \rangle = \sum_n w_{mn, m'n}. \quad (1.47)$$

²O produto de Kronecker não deve ser confundido com a multiplicação usual de matrizes, que é uma operação totalmente diferente.

De forma similar, $W^{(2)} = \text{Tr}^{(1)} W$.
 \implies Resolva o exercício 1.10.

1.3 Operadores auto-adjuntos (Hermitianos)

1.3.1 Propriedades e definição

A classe de operadores mais importante de mecânica quântica são os chamados **operadores auto-adjuntos**, frequentemente chamados de **operadores Hermitianos**. Em espaços vetoriais finitos, não há distinção entre operadores auto-adjuntos e Hermitianos, mas em espaços vetoriais infinitos existe uma distinção formal, que não será abordada aqui. Operadores auto-adjuntos satisfazem a relação:

$$\langle \phi | A | \psi \rangle = \langle \psi | A | \phi \rangle^*, \quad (1.48)$$

ou, simplificadamente, $A^\dagger = A$, que é a definição de operador Hermitiano. Se um dado operador Hermitiano possuir representação matricial, a sua matriz conjugada transposta é igual a ela própria, ou seja, os elementos da matriz satisfazem: $M_{ij} = M_{ji}^*$. Para operadores Hermitianos temos a importante definição:

Definição: Se um operador Hermitiano A age em um vetor $|\phi\rangle$ e retorna um múltiplo desse mesmo vetor, isto é, $A|\phi\rangle = a|\phi\rangle$, dizemos que $|\phi\rangle$ é um *autovetor* de A e a é o *autovalor* de A associado a $|\phi\rangle$. Pode acontecer que dois ou mais autovetores linearmente independentes estejam associados ao mesmo autovalor. Neste caso, dizemos que o autovalor é *degenerado*. O número de autovetores linearmente independentes associados a um mesmo autovalor a é chamado de *índice de degenerescência* de a . Neste caso, os autovetores associados ao mesmo autovalor definem um subespaço cuja dimensão é o índice de degenerescência.

Para mais detalhes sobre o que são autovalores e autovetores, estruturas importantíssimas em mecânica quântica e que são oriundas da álgebra linear, recomendamos consultar o apêndice A.

\implies Resolva o exercício 1.11.

Operadores auto-adjuntos satisfazem três importantes teoremas que enunciaremos a seguir:

Teorema 2 Se $\langle \psi | A | \psi \rangle = \langle \psi | A | \psi \rangle^* \quad \forall |\psi\rangle$, então $A = A^\dagger$.

Teorema 3 Se A for Hermitiano todos os seus autovalores são reais.

Teorema 4 *Autovetores de um operador Hermitiano com autovalores distintos são ortogonais entre si.*

Quando há degenerescência, qualquer vetor pertencente ao subespaço associado a um mesmo autovalor é autovetor de A . Neste caso, é sempre conveniente escolher bases ortogonais nos subespaços degenerados, de modo que em conjunto com os demais autovetores não degenerados de A , formem uma base ortogonal em todo o espaço V .

\implies *Resolva o exercício 1.12.*

1.3.2 Projetor

Seja um ket $|v\rangle$ descrito em termos de uma base (conjunto ortonormal completo $\{\phi_i\}$) da seguinte forma:

$$|v\rangle = \sum_i v_i |\phi_i\rangle. \quad (1.49)$$

Multiplicando a esquerda pelo bra $\langle\phi_j|$ e considerando $\langle\phi_j|\phi_i\rangle = \delta_{ji}$, encontramos $v_i = \langle\phi_i|v\rangle$, ou seja, as componentes do ket $|v\rangle$ na base dos $|\phi_i\rangle$ são os produtos escalares $\langle\phi_i|v\rangle$.

Se $A|\phi_i\rangle = a_i|\phi_i\rangle$, então o operador A pode ser escrito em termos dos seus autovalores e autovetores:

$$A = \sum_i a_i |\phi_i\rangle \langle\phi_i|, \quad (1.50)$$

e seus autovalores e autovetores podem ser encontrados por:

$$\det|A - a_i I| = 0. \quad (1.51)$$

Diante disso podemos escrever o operador projeção (*projetor*) P , que projeta o vetor $|v\rangle$ (ou qualquer outro vetor) na base $\{\phi_i\}$ como:

$$P = \sum_i |\phi_i\rangle \langle\phi_i| = I, \quad (1.52)$$

de forma que

$$|v\rangle = P|v\rangle = \sum_i P_i |\phi_i\rangle. \quad (1.53)$$

O termo $P_i = |\phi_i\rangle \langle\phi_i|$ pode ser entendido como um projetor que projeta o ket $|v\rangle$ no vetor de base específico $|\phi_i\rangle$.

Como a soma de todos os projetores de uma base completa é igual à identidade (eq. (1.52)), ela pode ser usada em uma expressão algébrica para

realizar uma mudança de base (um procedimento bastante comum na resolução de problemas), como exemplificado abaixo:

$$|v\rangle = \sum_i v_i^\phi |\phi_i\rangle = \sum_i v_i^\phi I |\phi_i\rangle = \sum_i \sum_j v_i^\phi |\psi_j\rangle \langle\psi_j|\phi_i\rangle = \sum_j v_j^\psi |\psi_j\rangle, \quad (1.54)$$

ou, para elementos de matrizes:

$$M_{kl} = \langle\phi_k| M |\phi_l\rangle = \sum_i \sum_j \langle\phi_k|\psi_i\rangle \langle\psi_i| M |\psi_j\rangle \langle\psi_j|\phi_l\rangle = \langle\psi_k| M |\psi_l\rangle. \quad (1.55)$$

Operadores de projeção podem projetar também sobre subespaços. Por exemplo, seja o subespaço que contém mais de um $|\phi_{ij}\rangle$ associado a um mesmo autovalor degenerado a_i de um operador Hermitiano A . Podemos definir o projetor

$$P_i = \sum_{j=1}^d |\phi_{ij}\rangle \langle\phi_{ij}|. \quad (1.56)$$

Aqui d é o índice de degenerescência e $\{|\phi_{i1}\rangle, \dots, |\phi_{id}\rangle\}$ é uma base ortogonal nesse subespaço.

Um projetor P é um operador Hermitiano que satisfaz a condição de idempotência: $P^2 = P$.

\implies Resolva o exercício 1.13.

1.3.3 Teorema espectral

Todo esse arcabouço matemático nos permite enunciar um dos teoremas mais importantes para a mecânica quântica:

Teorema 5 (Teorema espectral - versão discreta) *Considere um operador R Hermitiano de espectro puramente discreto. Esse operador, então, pode ser escrito em termos de seus autovalores r_n e dos projetores sobre os subespaços associados da seguinte forma:*

$$R = \sum_n r_n P_n. \quad (1.57)$$

No caso não degenerado, a equação acima pode ser escrita em termos dos projetores sobre os autovetores $|r_n\rangle$ de R :

$$R = \sum_n r_n |r_n\rangle \langle r_n|. \quad (1.58)$$

Este teorema abarca sistemas onde os autovalores são discretos. Um exemplo físico seria o spin de um elétron. Entretanto, quando estamos falando de quantidades contínuas, como posição e momento, o sistema terá autovalores contínuos e será representado em uma base infinita. Para espaços de dimensão infinita (a versão contínua do teorema) teremos que ser cuidadosos na definição de operadores Hermitianos, pois as condições de contorno devem ser levadas em conta. Vejamos um exemplo:

Exemplo 5 *Seja o operador $D = -i\frac{d}{dx}$ definido no espaço de todas as funções diferenciáveis de uma variável no intervalo $a \leq x \leq b$. Usando a definição de adjunto:*

$$\begin{aligned} \langle \phi | D | \psi \rangle &= \langle \psi | D | \phi \rangle^* \\ \int_a^b \phi^*(x) D^\dagger \psi(x) dx &= \left[\int_a^b \psi^*(x) D \phi(x) dx \right]^*. \end{aligned} \quad (1.59)$$

Porém $\phi D \psi = D[\phi \psi] - (D\phi)\psi$. Então, integrando por partes o termo do lado direito encontramos:

$$\int_a^b \phi^*(x) D^\dagger \psi(x) dx = \int_a^b \phi^*(x) D \psi(x) dx + i \left[\psi(x) \phi(x)^* \right]_a^b. \quad (1.60)$$

O último termo do lado direito depende de condições de contorno (os valores das funções nos extremos do intervalo). Se escolhermos condições de contorno adequadas, por exemplo, $\phi(a) = \phi(b) = \psi(a) = \psi(b) = 0$, então o último termo se anula e o operador D é Hermitiano e, portanto, valerá o teorema espectral.

Assim sendo, para a equação diferencial

$$-i \frac{d}{dx} \phi(x) = \lambda \phi(x) \quad (1.61)$$

que tem como solução $\phi(x) = ce^{i\lambda x}$, onde c é uma constante, se considerada uma equação de autovalores para o operador D , as soluções válidas estarão restritas a um certo espaço vetorial, que pode ser definido dependendo das condições de contorno que são impostas. Um caso específico largamente utilizado na física são as condições de contorno periódicas, $\phi(a) = \phi(b)$. Considerando $a = -L/2$ e $b = +L/2$, os autovalores formarão um conjunto discreto $\lambda = \lambda_n = 2\pi n/L$, com n inteiro. Se nenhuma condição de contorno for imposta, D não será Hermitiana e não será de interesse para a mecânica quântica.

Antes de apresentar a forma mais geral do teorema espectral, é conveniente recordar ainda o conceito de *integral de Stieltjes* (de forma bem simplificada). Seja um intervalo $[a, b]$ sobre a reta real (eixo x) e duas funções reais $f(x)$ e $g(x)$ limitadas, definidas sobre $[a, b]$. Seja $p = \{a = x_0 < x_1 < \dots < x_n = b\}$ uma partição de $[a, b]$ e $\Delta x_i = x_{i+1} - x_i$. A integral de Stieltjes é definida como

$$\int_a^b f(x)dg(x) = \lim_{\substack{n \rightarrow \infty \\ \text{Max}(\Delta x_i) \rightarrow 0}} \sum_{i=0}^{n-1} f(x_i)[g(x_{i+1}) - g(x_i)]. \quad (1.62)$$

Se $g(x) = x$, a integral de Stieltjes se reduz à integral de Riemann. Se $g(x)$ é diferenciável,

$$\int_a^b f(x)dg(x) = \int_a^b f(x) \frac{dg}{dx} dx. \quad (1.63)$$

Exemplo 6 *Seja $g(x) = \sum_i \Theta(x - x_i)$, onde $\Theta(x)$ é a função degrau de Heaviside. Neste caso,*

$$\begin{aligned} \int_a^b f(x)dg(x) &= \int_a^b f(x) \frac{dg}{dx} dx \\ &= \int_a^b f(x) \sum_i \delta(x - x_i) dx \\ &= \sum_{x_i \in [a, b]} f(x_i). \end{aligned} \quad (1.64)$$

Vemos que dependendo de $g(x)$, a integral de Stieltjes pode se transformar em uma soma discreta.

Teorema 6 (Teorema espectral - versão contínua) *Para cada operador auto adjunto A , existe uma família única de projetores $E(\lambda)$ com as seguintes propriedades:*

- Se $\lambda_1 < \lambda_2$, então $E(\lambda_1)E(\lambda_2) = E(\lambda_2)E(\lambda_1) = E(\lambda_1)$.
- Se $\epsilon \rightarrow 0^+$, então $E(\lambda + \epsilon) |\psi\rangle \rightarrow E(\lambda) |\psi\rangle$.
- Se $\lambda \rightarrow -\infty$, então $E(\lambda) |\psi\rangle \rightarrow 0$.
- Se $\lambda \rightarrow +\infty$, então $E(\lambda) |\psi\rangle \rightarrow |\psi\rangle$.
- $\int_{-\infty}^{\infty} \lambda dE(\lambda) = A$. (1.65)

Exemplo 7 O teorema espectral em sua versão contínua aplica-se também ao caso discreto. Se a um operador Hermitiano A está associada a família de projetores $E(\lambda)$ dada por

$$E(\lambda) = \sum_i |\phi_i\rangle \langle \phi_i| \Theta(\lambda - a_i), \quad (1.66)$$

o teorema espectral diz que

$$\begin{aligned} A &= \int_{-\infty}^{\infty} \lambda dE(\lambda) \\ &= \int_{-\infty}^{\infty} \lambda \frac{dE}{d\lambda} d\lambda \\ &= \int_{-\infty}^{\infty} \lambda \sum_i |\phi_i\rangle \langle \phi_i| \delta(\lambda - a_i) d\lambda \\ &= \sum_i a_i |\phi_i\rangle \langle \phi_i|. \end{aligned} \quad (1.67)$$

Exemplo 8 Uma forma mais comum de se encontrar o teorema espectral nos textos de mecânica quântica é a seguinte. Considere o operador posição Q e seus autovetores $|q\rangle$. Embora os vetores $|q\rangle$ não sejam bem definidos no espaço de Hilbert, por terem norma infinita, a família de projetores

$$E(\lambda) = \int_{-\infty}^{\lambda} |q\rangle \langle q| dq \quad (1.68)$$

é bem definida. Levando $E(\lambda)$ no teorema espectral, chegamos a

$$Q = \int_{-\infty}^{\infty} q |q\rangle \langle q| dq. \quad (1.69)$$

Mencionamos as duas versões do teorema espectral, para espectro contínuo e discreto, comentamos sobre as condições para que ele seja válido (operadores Hermitianos e condições de contorno adicionais para operadores em dimensão infinita), mas omitimos a demonstração do teorema espectral devido à sua complexidade.

Felizmente, muitos dos resultados desenvolvidos para problemas de autovalores discretos podem ser generalizados para o caso contínuo. Vale a equação de autovalores ($Q|q\rangle = q|q\rangle$); pode-se substituir o símbolo de Kronecker pelo delta de Dirac

$$\langle r_n | r_m \rangle = \delta_{nm} \rightarrow \langle q' | q'' \rangle = \delta(q' - q''); \quad (1.70)$$

pode-se substituir a soma sobre todos os autovalores discretos pela integral na variável contíua

$$\sum_n |r_n\rangle \langle r_n| = 1 \rightarrow \int dq' |q'\rangle \langle q'| = 1, \quad (1.71)$$

e as relações de completeza 1.71 podem ser utilizadas para a substituição em trocas de base, como

$$|\alpha\rangle = \sum_n |r_n\rangle \langle r_n | \alpha \rangle \rightarrow |\alpha\rangle = \int dq' |q'\rangle \langle q' | \alpha \rangle, \quad (1.72)$$

$$\langle \beta | \alpha \rangle = \sum_n \langle \beta | r_n \rangle \langle r_n | \alpha \rangle \rightarrow \langle \beta | \alpha \rangle = \int dq' \langle \beta | q' \rangle \langle q' | \alpha \rangle. \quad (1.73)$$

Vale explicitar, apenas, que estando o espaço e Hilbert \mathcal{H} limitado a vetores de norma finita ($\langle \Psi | \Psi \rangle = \int_{-\infty}^{\infty} |\Psi(x)|^2 dx \equiv$ finito), algumas funções são trabalhadas no chamado espaço nuclear Ω , limitado a funções $\phi(x)$ que vão a zero mais rápido que qualquer potência de x para $|x| \rightarrow \infty$, outras no espaço estendido Ω^\times (conjugado de Ω), limitado a $\chi(x)$ que satisfaz $\langle \chi(x) | \phi(x) \rangle = \int_{-\infty}^{\infty} \chi(x)^* \phi(x) dx \equiv$ finito, $\forall \phi \in \Omega$. A condição em $\phi(x)$ limita Ω a funções que satisfazem $\int_{-\infty}^{\infty} |\phi(x)|^2 (1 + |x|)^m dx < \infty$ ($m = 0, 1, 2, \dots$). O tripleto $\Omega \subset \mathcal{H} \subset \Omega^\times$ contém, por exemplo, as funções $\delta(x - x')$, que é autofunção do operador Q , tal que $Q\delta(x - x') = x'\delta(x - x')$, e a função e^{ikx} , que é autofunção do operador $i\frac{d}{dx}$. Estas duas funções são de importância primordial para os conceitos de posição e momento em mecânica quântica, como veremos nos capítulos 4 e 5.

1.3.4 O teorema espectral e as soluções reais em física

Em muitos livros e comentários sobre mecânica quântica existe uma certa urgência em se usar operadores Hermitianos por causa do teorema 3 que garante os autovalores de um operador Hermitiano como sendo números reais.

Embora o tratamento de mensuráveis físicos como números reais seja algo intuitivo, ele é simplesmente uma convenção. Poderíamos, de igual forma e sem alterar nada de fundamentalmente importante em M.Q., usar números complexos para representar mensuráveis físicos. E, seguindo essa lógica, poderíamos deixar de lado a condição de hermiticidade e usar classes mais gerais de operadores, como por exemplo a dos operadores normais ($AA^\dagger = A^\dagger A$), que admitem a decomposição espectral com autovalores complexos. De fato, alguns autores já propuseram essa generalização [6, 7, 8]. Entretanto, a associação de resultados de medições a números reais e o uso de operadores Hermitianos para representar as respectivas quantidades físicas prevalece na grande maioria dos textos de mecânica quântica.

1.3.5 Teoremas importantes sobre operadores Hermitianos

Além da importância do teorema espectral que já mencionamos, existem outros dois teoremas importantes sobre operadores Hermitianos que gostaríamos de enunciar.

Teorema 7 *Sejam A e B dois operadores Hermitianos, cada um deles possuindo um conjunto completo de autovetores (uma base). Se os operadores comutam, ou seja, se $AB = BA$ então existe um conjunto completo de autovetores (uma base) para ambos A e B .*

Teorema 8 *Qualquer operador que comuta com todos os operadores de um conjunto completo de operadores que comutam entre si é uma função de operadores deste conjunto.*

O teorema 7 é de grande importância em mecânica quântica e está novamente associado a operadores Hermitianos. Porém, ele possui uma condição adicional: exige que dois operadores A e B , Hermitianos, também comutem entre si.

Caso esses dois operadores sejam os únicos que comutam entre si e que possuem bases completas, então dizemos que esses operadores formam um **conjunto completo de operadores que comutam (C.C.O.C.)**. O teorema 7 garante que sempre é possível encontrar uma base comum que diagonaliza todos os operadores de um C.C.O.C. Não necessariamente esse conjunto consistirá de apenas dois operadores.

A título de curiosidade, o C.C.O.C. do átomo de hidrogênio, um dos sistemas quânticos mais simples que existe, consiste nos operadores L^2 , L_z , H , S^2 e S_z , consistindo-se de cinco operadores. Veremos isso em detalhes mais tarde.

Exemplo 9 (Exemplo de função de operadores) *Seja o operador A Hermitiano $A|a_i\rangle = a_i|a_i\rangle$ com $\langle a_i|a_j\rangle = \delta_{ij}$. Pelo teorema espectral:*

$$A = \sum_i a_i |a_i\rangle \langle a_i|, \quad (1.74)$$

uma função $f(A)$ de operadores é tal que:

$$f(A) = \sum_i f(a_i) |a_i\rangle \langle a_i|, \quad (1.75)$$

um exemplo de função de operadores seria $f(A) = A^n$. Essa função satisfaz, assim como os números reais,

$$A^n A^m = A^{n+m}. \quad (1.76)$$

\implies Resolva o exercício 1.14.

1.4 Teoria de probabilidade

A teoria de probabilidade representa, junto com a álgebra linear, um pilar da mecânica quântica. Enquanto na mecânica clássica valores definidos são atribuídos aos mensuráveis físicos, na mecânica quântica calculam-se as probabilidades de determinados mensuráveis físicos terem determinados valores. Logo, similar ao que foi feito para a álgebra linear, começaremos com as definições importantes da teoria da probabilidade.

1.4.1 Definições iniciais

Seja $P(A|B)$ a probabilidade de encontrar o evento A sob condições determinadas pelo evento B.

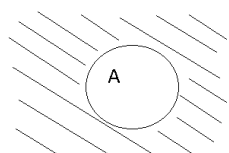
Exemplo 10 Qual é a probabilidade de se jogar um dado e tirar um número par?

Evento A = obter número par

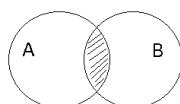
Evento B = jogar um dado

$$P(A|B) = \frac{3}{6} = \frac{1}{2} = 50 \%. \quad (1.77)$$

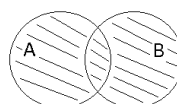
Definições:



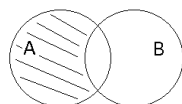
(a) Negação de A
($\sim A$)



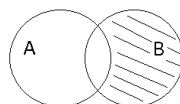
(b) Conjunção de A
com B ($A \& B$)



(c) Disjunção de A
e B ($A \vee B$)



(d) Conjunção de A
e não B ($A \& \sim B$)



(e) Conjunção de
não A e B ($\sim A \& B$)

1.4.2 Axiomas da teoria da probabilidade

A teoria da probabilidade está fundamentada nos seguintes axiomas:

Axioma 1 $0 \leq P(A|B) \leq 1$ (a probabilidade é não negativa e menor ou igual a 1)

Axioma 2 $P(A|A) = 1$ (a probabilidade de um evento condicionado a ele mesmo é igual a 1)

Axioma 3 $P(\sim A|B) = 1 - P(A|B)$ (probabilidade complementar)

Juntos, axiomas 2 e 3 implicam em $P(\sim A|A) = 0$

Axioma 4 $P(A \& B|C) = P(A|C) P(B|A \& C)$

Exemplo 11 Calcule a probabilidade de ao se jogar um dado obter um número par e esse número par ser 6.

Evento A = Sair um número par

Evento B = Sair o número 6

Evento C = Jogar um dado

Logo teremos $P(A|C) = \frac{3}{6}$ e também $P(B|A \& C) = \frac{1}{3}$ portanto:

$$P(A \& B|C) = P(A|C) P(B|A \& C) = \frac{3}{6} \cdot \frac{1}{3} = \frac{1}{6}. \quad (1.78)$$

Note que não há necessidade de um axioma para a disjunção:

$$P(A \vee B|C) = 1 - P(\sim A \& \sim B|C), \quad (1.79)$$

pois $A \vee B = \sim A \& \sim B$. No lado direito da equação anterior, podemos usar o axioma 4. Existe uma forma de calcular disjunção através de probabilidades de eventos:

$$P(A \vee B|C) = P(A|C) + P(B|C) - P(A \& B|C). \quad (1.80)$$

Note que num diagrama de conjuntos, o último termo corresponde à subtração do termo de conjunção para que este não seja contado duas vezes. Este resultado possui um caso especial para o qual os eventos A e B são ditos **mutuamente exclusivos** de C . Nesse caso $P(A \& B|C) = 0$ daí:

$$P(A \vee B|C) = P(A|C) + P(B|C). \quad (1.81)$$

Com essa condição de exclusividade mútua os eventos A e B são intercambiáveis:

$$P(A|C)P(B|A\&C) = P(B|C)P(A|B\&C), \quad (1.82)$$

disto podemos tirar o **teorema de Bayes**:

$$P(B|A\&C) = \frac{P(A|B\&C)P(B|C)}{P(A|C)}. \quad (1.83)$$

que é o princípio da probabilidade inversa, que conecta a probabilidade de B dado $A\&C$, com a probabilidade de A , dado $B\&C$.

Exemplo 12 Calcule a probabilidade de ao se jogar um dado sair um número 6 sendo que o resultado foi um número par.

Evento A = Resultado foi um número par

Evento B = Sair 6

Evento C = Jogar um dado

Logo teremos:

$$P(A|B\&C) = 1, \quad P(B|C) = \frac{1}{6}, \quad P(A|C) = \frac{3}{6} = \frac{1}{2}, \quad (1.84)$$

pelo teorema de Bayes teremos:

$$P(B|A\&C) = \frac{1 \times \frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}. \quad (1.85)$$

1.4.3 Eventos independentes

Dois eventos A e B são ditos estocasticamente (ou estatisticamente) independentes quando for válido:

$$P(B|A\&C) = P(B|C), \quad (1.86)$$

o que implica em:

$$P(A\&B|C) = P(A|C)P(B|C). \quad (1.87)$$

Um exemplo simples de eventos independentes são as múltiplas jogadas de uma moeda. Os resultados que podem ser obtidos em cada jogada (cara ou coroa) são independentes entre si. A independência de resultados também é importante na distribuição binomial, que veremos mais adiante. A equação (1.87) pode ser estendida para n eventos, resultado na multiplicação sucessiva das probabilidades independentes.

Exemplo 13 (Desigualdade de Bell) *Considere um evento ξ , e a ele estão ligados os eventos LETRA (com resultados A ou B), NÚMERO (resultados 1 ou 2) e SETA (resultados \uparrow ou \downarrow), sendo LETRA, NÚMERO e SETA mutuamente independentes entre si. Para sistemas desse tipo, vale a seguinte relação:*

$$P(A \& (\sim 1) | \xi) + P(1 \& (\sim \uparrow) | \xi) \geq P(A \& (\sim \uparrow) | \xi). \quad (1.88)$$

Para demonstrar esta desigualdade, note que existem 8 possíveis combinações de resultados $\{(A, 0, \uparrow), (A, 0, \downarrow), (A, 1, \uparrow), (A, 1, \downarrow), (B, 0, \uparrow), (B, 0, \downarrow), (B, 1, \uparrow), (B, 1, \downarrow)\}$, e que $P(A \& (\sim 1) | \xi) = P(A \& 0 \& \downarrow | \xi) + P(A \& 1 \& \uparrow | \xi)$; $P(1 \& (\sim \uparrow) | \xi) = P(A \& 1 \& \downarrow | \xi) + P(B \& 1 \& \uparrow | \xi)$; $P(A \& (\sim \uparrow) | \xi) = P(A \& 1 \& \uparrow | \xi) + P(A \& 1 \& \downarrow | \xi)$.

Exemplo 14 (Desigualdade de Bell-Wigner) *Sejam seis eventos $A_1, B_1, C_1, A_2, B_2, C_2$. Todos os eventos estão associados a valores $+$ e $-$, mutuamente exclusivos, sendo que α_1 e α_2 ($\alpha = A, B, C$) são anti correlacionados, isto é, se $\alpha_1 = +$, então $\alpha_2 = -$, e assim por diante. Chamando de $P(\alpha_j)$ a probabilidade de que $\alpha_j = +$, a desigualdade de Bell-Wigner [9] estabelece que*

$$P(A_1 \& C_2) \leq P(A_1 \& B_2) + P(B_1 \& C_2). \quad (1.89)$$

No seu artigo [9], Wigner usa esta desigualdade para provar a impossibilidade de se atribuir valores objetivos (ainda que aleatórios) simultaneamente às três componentes do spin de uma partícula de spin $\frac{1}{2}$.

Estas desigualdades são importantes em física (particularmente em informação quântica) e são adequações do que ficou conhecido como desigualdade de Bell, demonstrado pela primeira vez por John Bell em 1964 [10].

\implies *Resolva o exercício 1.15.*

1.5 Distribuições de probabilidade

Distribuição de probabilidade é uma forma de descrever as probabilidades de ocorrência de um evento através de uma função matemática. O uso de distribuições de probabilidade facilita a análise de diversos problemas probabilísticos. Podemos entender, de certa forma, que a função de onda (solução da equação de Schrödinger) é uma distribuição de probabilidade. Comentaremos sobre isso mais adiante.

Discutiremos aqui dois tipos de distribuições que são comumente observadas, embora muitas outras existam na natureza [11]. Mas inicialmente, definimos o conceito geral de momentos de uma distribuição.

1.5.1 Distribuição binomial

Considere um experimento aleatório E que possui apenas dois resultados A e $\sim A$ (note que essa é apenas uma notação, os resultados podem ser 0 ou 1, pra cima ou pra baixo, sim ou não, etc). O resultado A possui probabilidade p de ocorrer ($P(A|E) = p$) e o resultado $\sim A$ possui probabilidade $1 - p$ de ocorrer ($P(\sim A|E) = 1 - p$).

Cada experimento aleatório E é independente dos outros. Repete-se esse experimento um total de N vezes e obtêm-se n vezes o resultado A e $N - n$ vezes o resultado $\sim A$. A probabilidade de se encontrar o resultado A um número r de vezes ($r < N$) é dado pela **distribuição binomial**:

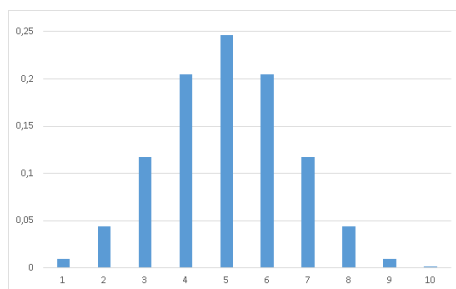
$$P(n = r|E^N) = \frac{N!}{r!(N-r)!} p^r (1-p)^{N-r}. \quad (1.90)$$

O primeiro termo é o número de permutações possíveis de N elementos (o total de vezes que o experimento é repetido) tomados r a r vezes (o número de vezes que obtemos A como resultado). O termo p^r é a probabilidade de se obter r vezes o resultado A . O último termo é a probabilidade de se medir $\sim A$ um número $N - r$ de vezes.

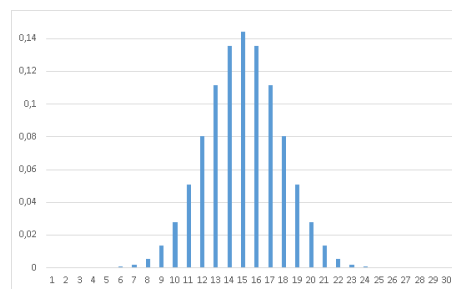
O valor médio da distribuição binomial e a variância são dados por:

$$\langle n \rangle = Np, \quad \langle (\Delta n)^2 \rangle = Np(1-p). \quad (1.91)$$

A título de ilustração, considere 10 jogadas sucessivas de uma moeda. A probabilidade de sair cara ou coroa é de $p = 1/2$. A seguir apresentamos dois gráficos com todos os resultados possíveis para a probabilidade de se obter de 1 até 10 caras e outro com todos os resultados possíveis para a probabilidade de se obter de 1 até 30 caras:



(f) Gráfico com $N = 10$



(g) Gráfico com $N = 30$

Figura 1.1: Gráficos de distribuições binomiais de jogadas de moedas.

O gráfico da direita possui um pico mais estreito e mais concentrado que o da esquerda. Na seção §1.5.2 comentaremos mais sobre isso.

⇒ Resolva o exercício 1.16.

1.5.2 Distribuição gaussiana

Outra distribuição de probabilidade de grande relevância em física é a **distribuição gaussiana** ou **distribuição normal**. Considere uma variável aleatória contínua X . A função densidade de probabilidade (a função matemática que descreve a distribuição de probabilidade) da distribuição gaussiana é dada por:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad (1.92)$$

com valor médio e variância dados, respectivamente, por:

$$\langle X \rangle = \mu, \quad \langle (\Delta X)^2 \rangle = \sigma^2. \quad (1.93)$$

A partir dessa distribuição de probabilidade é possível calcular a probabilidade de se medir X no intervalo de $-\infty$ até um dado valor a como sendo:

$$P(X < a) = \int_{-\infty}^a p(x) dx, \quad (1.94)$$

note que essa integral é numérica e, portanto, é muito comum o uso de softwares ou de tabelas para achar as probabilidades pedidas. Abaixo apresentamos o gráfico de uma distribuição gaussiana de média nula:

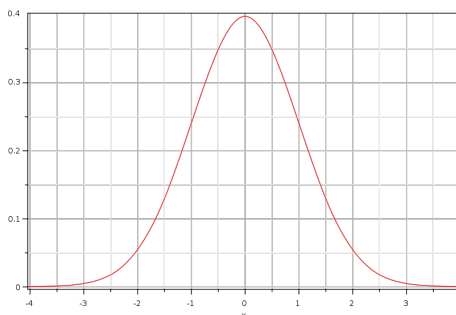


Figura 1.2: Gráfico de uma distribuição gaussiana com $\mu = 0$ e $\sigma = 1$.

Note que a forma dessa distribuição (a menos de sua “espessura”, que é regulada pelo desvio padrão σ) é muito similar à envoltória dos dados da distribuição binomial, em especial aquela para $N = 30$.

⇒ Resolva o exercício 1.17.

De fato, a probabilidade de se encontrar um resultado de uma distribuição binomial se aproxima de zero quando $N \rightarrow \infty$ (Lei dos grandes números).

Reparemos nos gráficos da figura 1.1: a probabilidade para $N = 30$ é muito pequena.

Além disso, quando N é grande a distribuição binomial se comporta como uma distribuição gaussiana. Se X é uma distribuição binomial com N experimentos e probabilidade p , então a variável aleatória $\frac{X-Np}{\sqrt{Np(1-p)}}$ é uma variável aleatória gaussiana com média nula e variância unitária (conhecida como distribuição normal padrão).

Em muitos problemas que envolvem distribuições binomiais com grandes números (valor de N muito grande) é mais adequado trabalhar com a distribuição gaussiana (a aproximação é ainda melhor quando N cresce).

Definindo frequência como $f = n/N$, onde n é o número de vezes que se obtém um resultado em um procedimento, e N o número total de vezes que o procedimento é repetido, para $N \rightarrow \infty$ a diferença entre a frequência e a probabilidade de se encontrar o resultado específico tende a zero.

\Rightarrow Resolva o exercício 1.18.

1.5.3 Momentos de uma distribuição

Os conceitos de *valor médio* e *variância* de uma função podem ser generalizados como *momentos de uma distribuição*. Considere uma função $p(x)$, diferenciável, que representa uma distribuição de probabilidade associada aos possíveis valores x de uma variável aleatória contínua X . Faremos uma discussão de momentos para variáveis contínuas, mas o mesmo segue válido para variáveis discretas.

Definimos o **primeiro momento** (ou o momento de primeira ordem) da distribuição de probabilidade de X como sendo a integral:

$$\langle X \rangle = \int_{\Omega} xp(x) dx, \quad (1.95)$$

sendo Ω o espaço onde a variável aleatória X apresenta valores. Para uma variável discreta a expressão anterior se reduz a um somatório. Sendo assim, o momento de n -ésima ordem é dado por:

$$\langle X^n \rangle = \int_{\Omega} x^n p(x) dx, \quad (1.96)$$

e para uma função de uma variável aleatória, $f(X)$, que também é uma variável aleatória, segue valendo:

$$\langle f(X) \rangle = \int_{\Omega} f(x) p(x) dx, \quad (1.97)$$

dizemos que o momento existe se a integral converge (em algumas distribuições mais complicadas pode acontecer dos momentos serem infinitos). Em alguns

problemas é costume calcular os chamados **momentos centrais** que são deslocados. O n -ésimo momento central é dado por:

$$\mu_n = \langle (X - \langle X \rangle)^n \rangle = \int_{\Omega} (x - \langle x \rangle)^n p(x) dx. \quad (1.98)$$

Em física geralmente chamamos os primeiros momentos de **valor esperado** ou de **valor médio**. O momento central de segunda ordem é normalmente conhecido como **variância** e é muito usado para caracterizar distribuições. Através da variância define-se o desvio padrão:

$$\sigma = \sqrt{\langle (\Delta x)^2 \rangle}. \quad (1.99)$$

1.5.4 Frequência e probabilidade de um evento

Vamos definir a frequência com que um certo procedimento gera um resultado específico r como $f_r = n_r/N$, onde n_r é o número de vezes que se obtém o resultado r , e N o número total de vezes que o procedimento é repetido. Para $N \rightarrow \infty$, a diferença entre a frequência e a probabilidade de se encontrar o resultado r tende a zero.

Como veremos durante o curso, esta sutileza, que desvincula os conceitos de frequência e probabilidade, está no cerne da interpretação da mecânica quântica enquanto teoria que descreve a realidade de um sistema individual ou de um conjunto de procedimentos de preparação e medida de estados quânticos.

1.6 Exercícios

Exercício 1.1 *Os números complexos poder ser considerados elementos (vetores) de um espaço vetorial linear de 2 dimensões? Explique (veja a seção §1.1.1).*

Exercício 1.2 *Prove que as funções $\sin x$ e $\sin(2x)$ são L.I. Utilize diretamente a equação (1.2).*

Exercício 1.3 *Prove que as funções do exercício 1.2 são ortogonais. Confirme o resultado do teorema 1.*

Exercício 1.4 *Prove a desigualdade de Schwarz, equação (1.8).*

Exercício 1.5 *Prove a desigualdade do triângulo, equação (1.9).*

Exercício 1.6 Qual é o operador que leva $\psi = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$ em $\phi = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ ambos elementos de \mathbb{R}^3 ? Esse operador é linear?

Exercício 1.7 Mostre que $\psi(x) = e^x$ é solução da equação de operadores (1.30). Nesse caso dizemos que $\psi(x)$ é autofunção do operador O .

Exercício 1.8 Prove a invariância do traço para mudanças de base. Use a notação de Dirac.

Exercício 1.9 Prove a propriedade de produto externo, equação (1.37).

Exercício 1.10 Seja $A^{(1)} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ e $A^{(2)} = \begin{pmatrix} 0 & 2 \\ 2 & 1 \end{pmatrix}$. Encontre $A = A^{(1)} \otimes A^{(2)}$. Partindo de A , encontre $A^{(1)}$ e $A^{(2)}$ através do traço parcial na outra componente.

Exercício 1.11 Encontre os autovalores e autovetores da matriz $A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$.

Exercício 1.12 Prove os teoremas (2)-(4) de operadores auto-adjuntos.

Exercício 1.13 Prove as propriedades de hermiticidade e idempotência³ do operador projeção, equação (1.56).

Exercício 1.14 Prove a equação (1.76).

Exercício 1.15 Prove a desigualdade de Bell-Wigner (1.89).

Exercício 1.16 Prove os momentos da distribuição binomial, equação (1.91).

Exercício 1.17 Prove os momentos da distribuição gaussiana, equação (1.93).

Exercício 1.18 Calcule a probabilidade de se encontrar cara 5 vezes jogando uma moeda 10 vezes. Calcule a probabilidade de encontrar cara 50 vezes jogando uma moeda 100 vezes. Compare os resultados considerando o que foi exposto na seção §1.5.2.

Exercício 1.19 Resolva os exercícios ainda não resolvidos do livro do Ballentine [2].

³Idempotência é a propriedade dos operadores que podem ser aplicados várias vezes sem que o valor do resultado se altere após a aplicação inicial.

