

# Fundamentos da Mecânica Quântica

**Vitor Oguri**

Departamento de Física Nuclear e Altas Energias (DFNAE)  
Instituto de Física Armando Dias Tavares (IFADT)  
Universidade do Estado do Rio de Janeiro (UERJ)

Essas notas constituem um complemento às abordagens dos tópicos contidos no livro texto<sup>1</sup> da disciplina Mecânica Quântica I, do curso de Física da UERJ.

13 de agosto de 2010

---

<sup>1</sup> D. J. Griffiths, *Introduction to Quantum Mechanics*, Prentice Hall, 1995.

# Índice

<b>1</b>	<b>Introdução</b>	<b>3</b>
1.1	Processos aleatórios . . . . .	3
1.2	Grandezas e variáveis aleatórias . . . . .	4
1.3	Probabilidades a priori . . . . .	4
1.3.1	Eventos equiprováveis . . . . .	4
1.3.2	Regras básicas . . . . .	5
1.3.3	Eventos compostos . . . . .	6
1.4	Probabilidades a posteriori . . . . .	9
1.4.1	Distribuição de probabilidades contínua . . . . .	10
1.5	Probabilidade axiomática . . . . .	13
<b>2</b>	<b>Física e probabilidades</b>	<b>14</b>
2.1	Teorias probabilísticas . . . . .	14
2.2	Analogias e dualidade . . . . .	15
<b>3</b>	<b>Origens da Mecânica Quântica</b>	<b>17</b>
3.1	A quantização de Planck da energia . . . . .	17
3.2	A quantização de Einstein da luz . . . . .	17
3.3	O átomo de Bohr . . . . .	18
3.4	A equação de Schrödinger . . . . .	19
3.5	A interpretação probabilística de Born . . . . .	20
3.6	O princípio da incerteza de Heisenberg . . . . .	22
3.7	Comentários sobre alguns livros de Mecânica Quântica . . . . .	23
<b>4</b>	<b>O movimento da partícula na descrição de Schrödinger</b>	<b>24</b>
4.1	Estados estacionários . . . . .	25
4.2	A partícula livre . . . . .	26
4.3	Operador momentum linear . . . . .	27
4.4	Generalizações . . . . .	28
4.5	As equações de Ehrenfest . . . . .	31
4.6	Problema de estado inicial . . . . .	32
4.7	Transições e regras de seleção . . . . .	33
4.8	Modelos e problemas típicos . . . . .	34
<b>5</b>	<b>Leis de conservação</b>	<b>35</b>
5.1	Momentum linear e translações . . . . .	35
5.2	Grupos de simetria e leis de conservação . . . . .	37
5.3	Geradores e regras de comutação . . . . .	39
5.4	Momentum angular e rotações . . . . .	41
<b>A</b>	<b>Operadores lineares</b>	<b>42</b>
A.1	Produto escalar e ortogonalidade . . . . .	42
A.2	Operadores lineares, adjuntos e hermitianos . . . . .	43
A.3	Autovalores e autofunções . . . . .	47
A.4	Representação matricial . . . . .	49

# 1 Introdução

A Mecânica Quântica foi criada e desenvolvida, no início do século XX, para resolver os problemas associados aos sistemas (microscópicos) atômicos, moleculares e nucleares. Utilizando-se de métodos probabilísticos, a teoria estabelece os fundamentos para a abordagem de fenômenos que ocorrem em dimensões da ordem de  $10^{-6}$  a  $10^{-13}$  cm, e intervalos de tempo da ordem de  $10^{-8}$  s.

## 1.1 Processos aleatórios

Os fenômenos físicos podem ser classificados como processos determinísticos ou aleatórios. Se os efeitos associados a um fenômeno devido a determinadas influências ou causas são inequivocamente previsíveis, diz-se que os processos envolvidos são determinísticos. Por outro lado, se os efeitos associados a um fenômeno não são exatamente previsíveis, mas podem ser associados a certas expectativas relativas de ocorrência ou probabilidades, os processos envolvidos são ditos aleatórios.

Em geral, a não previsibilidade dos efeitos de um fenômeno está associada a processos complexos que envolvem a interação de um grande número de sistemas simples. Nesse sentido, até o surgimento da Mecânica Quântica, em 1926, as teorias físicas probabilísticas que descreviam fenômenos ou sistemas físicos, por envolverem um grande número de partículas, eram também teorias estatísticas. Assim, o conceito de probabilidade estava vinculado ao comportamento estatístico das moléculas de um gás, ou à enorme quantidade de núcleos que participam do fenômeno da radioatividade.

Não existe nenhum método para se prever exatamente quantos átomos de rádio se desintegrarão em um dado intervalo de tempo, mas com base na observação repetida de um grande número de desintegrações de uma amostra, o neozelandês Ernest Rutherford (1871-1937), em 1916, estabeleceu que a média dos intervalos de tempo entre sucessivas desintegrações tende a um valor que caracteriza a taxa de desintegração de um núcleo.

Uma vez que a teoria fundamental da Física Clássica – a Mecânica de Newton – descreve o comportamento de um sistema com um pequeno número de partículas por equações diferenciais, pressupunha-se que esse comportamento seria completamente determinado por sua condição inicial.<sup>2</sup> A aleatoriedade e o acaso em um fenômeno, ou em um experimento, eram atribuídos a incapacidade do observador determinar as condições iniciais, ou à complexidade dos arranjos experimentais. Em princípio, a partir de uma dada condição inicial, uma teoria fundamental causal como a Mecânica seria determinística, e permitiria a predição de somente um único resultado para a evolução de um sistema físico. Portanto, teorias probabilísticas não seriam fundamentais, uma vez que poderiam admitir vários possíveis resultados para a evolução de um sistema, ao associar probabilidades a cada um desses possíveis resultados.<sup>3</sup>

---

<sup>2</sup> Caracterizado pelas posições e velocidades iniciais de suas partículas constituintes.

<sup>3</sup> Sabe-se hoje que, mesmo para sistemas com poucos graus de liberdade descritos por teorias causais, em princípio, determinísticas, pequenas perturbações iniciais podem dar origem a fenômenos caóticos não previsíveis.

## 1.2 Grandezas e variáveis aleatórias

Se as medidas associadas a uma grandeza não são inteiramente previsíveis quando efetuadas sob as mesmas condições experimentais pré-determinadas, diz-se que a grandeza é uma **variável aleatória**. No entanto, variáveis aleatórias não se manifestam de modo totalmente imprevisível, uma vez que, em geral, seus valores ou medidas, além de limitados a um intervalo definido, estão associados a certas expectativas de ocorrência. A aleatoriedade significa uma impredecabilidade parcial, tal que a ocorrência de qualquer valor, em vez deterministicamente previsível, é de certo modo probabilisticamente previsível, ou seja, a cada valor dentro desse intervalo atribui-se uma medida da expectativa de sua ocorrência.

## 1.3 Probabilidades a priori

A teoria de probabilidades teve sua origem na análise dos jogos de azar, ao se quantificar a expectativa de ocorrência do resultado associado a um fenômeno não aleatório, como um jogo de cartas ou o lançamento de dados. Essa quantificação iniciou-se com a correspondência entre os matemáticos franceses Blaise Pascal (1623-1662) e Pierre Fermat (1601-1665), em 1654, e foi sintetizada pelo também francês Pierre Simon Laplace (1749-1827), em sua clássica obra *Théorie Analytique des Probabilités*, em 1812.

Devido a essa gênese, os conceitos e métodos probabilísticos elementares são, usualmente, apresentados a partir de problemas que envolvem processos combinatoriais.

### 1.3.1 Eventos equiprováveis

Nesse sentido, o lançamento de dados proporciona um exemplo simples e acessível para se estudar e estabelecer as regras básicas dos fenômenos aleatórios ou do acaso.

Os únicos resultados possíveis para os lançamentos de um dado são os números naturais  $\{1, 2, 3, 4, 5, 6\}$ . Para um dado não viciado, devido à simetria do problema, a probabilidade *a priori* atribuída a ocorrência de um determinado número possível é igual a  $1/6$ . Assim, a definição de probabilidade, segundo Laplace, é dada pela razão entre o número de casos favoráveis possíveis para a ocorrência de um evento e o número total de alternativas igualmente possíveis, ou equiprováveis.

$$P(i) = \frac{1}{6} \quad (i = 1, 2, 3, 4, 5, 6)$$

Assim, o cálculo de probabilidades consiste na contagem das possibilidades de ocorrência de um evento. Mas, esse tipo de cálculo ou estimativa é possível somente em situações simples, nas quais os possíveis resultados, igualmente prováveis, em número finito, são conhecidos *a priori*.

$$P(\text{evento}) = \frac{\text{número de casos favoráveis}}{\text{número de casos possíveis}} \quad (\text{probabilidade clássica})$$

Um exemplo ainda mais simples de eventos equiprováveis são os lançamentos de uma moeda. Nesse caso, a probabilidade associada a cada um dos dois únicos resultados possíveis e equiprováveis,  $\{\text{cara}, \text{coroa}\}$ , é igual a  $1/2$ , ou seja,

$$P(\text{cara}) = P(\text{coroa}) = \frac{1}{2}$$

### 1.3.2 Regras básicas

A partir da análise de ocorrência de eventos simples, no entanto, pode-se estabelecer as regras básicas para o cálculo de probabilidades.

Por exemplo, no lançamento de dados, a probabilidade de ocorrência de um resultado ( $i$ ) maior que 4 é igual a probabilidade associada a a ocorrência dos dois resultados 5 ou 6, entre os seis igualmente possíveis e mutuamente excludentes,

$$P(i > 4) = P(5 \text{ ou } 6) = \frac{2}{6} = \frac{1}{3}$$

ou seja, igual à soma das respectivas probabilidades,

$$P(5 \text{ ou } 6) = P(5) + P(6) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

Por outro lado, a probabilidade de um resultado par,  $\{2, 4, 6\}$ , ou maior que 4,  $\{5, 6\}$ , é dada por

$$P(i = \text{par ou } i > 4) = \frac{4}{6} = \frac{2}{3}$$

pois, entre os seis igualmente possíveis, mas não mutuamente excludentes, os resultados distintos são  $\{2, 4, 5, 6\}$ .

Nesse caso,

$$P(i = \text{par ou } i > 4) = P(i = \text{par}) + P(i > 4) - P(i = \text{par e } i > 4) = \frac{3}{6} + \frac{2}{6} - \frac{1}{6} = \frac{2}{3}$$

Assim, quando a ocorrência de um evento implica necessariamente a não ocorrência do outro, como a ocorrência dos números  $A = 5$  ou  $B = 6$  no lançamento de um dado, ou seja, quando os eventos são ditos **mutuamente excludentes**, a probabilidade associada a ocorrência de um ou de outro é dada pela soma das probabilidades,

$$\boxed{P(A \text{ ou } B) = P(A) + P(B)} \quad (A \text{ e } B \text{ mutuamente excludentes})$$

Por outro lado, se a ocorrência de um evento  $A = \{i = \text{par}\}$  não exclui a ocorrência de outro evento  $B = \{i > 4\}$ , a probabilidade associada a ocorrência de  $A$  ou  $B$  é dada por

$$\boxed{P(A \text{ ou } B) = P(A) + P(B) - P(A \text{ e } B)} \quad (A \text{ e } B \text{ não excludentes})$$

onde  $P(A \text{ e } B)$  é a probabilidade de ocorrência simultânea dos dois eventos  $A$  e  $B$ .

Se os eventos são independentes, no sentido que a ocorrência de um não afeta a ocorrência do outro, ou seja, não está condicionado ao outro,

$$\boxed{P(A \text{ e } B) = P(A) \times P(B)} \quad (A \text{ e } B \text{ independentes})$$

### 1.3.3 Eventos compostos

Para eventos compostos, mesmo que em número finito, a situação pode ser mais complicada. Galileu Galileu (1564-1642), em um artigo escrito por volta de 1613, identificou corretamente os 216 resultados equiprováveis no lançamento de três dados, sendo capaz de prever a ligeira diferença entre as probabilidades dos três dados somarem 9 e 10. Apesar de cada resultado estar associado a seis partições distintas, cada partição corresponde a diferentes multiplicidades.

soma ( $S$ ) de três dados	partições	partições equivalentes	multiplicidade
9	(126)	126 – 162 – 216 – 261 – 612 – 621	6
	(135)	135 – 153 – 315 – 351 – 513 – 531	6
	(144)	144 – 414 – 441	3
	(225)	225 – 252 – 522	3
	(234)	234 – 243 – 324 – 342 – 423 – 432	6
	(333)	333	1
10	(136)	136 – 163 – 316 – 361 – 613 – 631	6
	(145)	145 – 154 – 415 – 451 – 514 – 541	6
	(226)	226 – 262 – 622	3
	(235)	235 – 253 – 325 – 352 – 523 – 532	6
	(244)	244 – 424 – 442	3
	(334)	334 – 343 – 433	3

Uma vez que o número total de casos possíveis é 216, com um total de 25 resultados possíveis para a soma 9, e de 27 para a soma 10, as respectivas probabilidades são dadas por

$$P(S = 9) = 25/216 \quad \text{e} \quad P(S = 10) = 27/216$$

Um exemplo um pouco mais complicado, consiste em um dos problemas propostos pelo jogador francês Chevalier de Méré a Pascal, que deu origem à correspondência entre Pascal e Fermat.

Méré queria saber o porquê de que, em quatro lançamentos de um mesmo dado de pôquer, as chances de ocorrência de uma face marcada com um *ás* ( $i = 1$ ) eram maiores do que, em vinte e quatro lançamentos de dois dados, a ocorrência de dois *azes* ( $i = 2$ ).

Uma vez que, em um único lançamento ( $N = 1$ ) de um dado, a probabilidade,  $P(i = 1)$ , de ocorrência de um *ás* é igual a  $1/6$ , a probabilidade de não ocorrência,  $P(i \neq 1)$ , é dada por

$$P(i \neq 1) = 1 - P(i = 1) = 1 - \frac{1}{6} = \frac{5}{6}$$

Assim, pela regra da intersecção de eventos independentes, a probabilidade de não ocorrer pelo menos um *ás* em quatro lançamentos ( $N = 4$ ) do mesmo dado é igual a

$$P(i \neq 1, N = 4) = P(i \neq 1, N = 1)^4 = \left(\frac{5}{6}\right)^4$$

e, a probabilidade de ocorrer pelo menos um *ás* em quatro lançamentos ( $N = 4$ ) do mesmo dado, é igual a

$$P(i = 1, N = 4) = 1 - P(i \neq 1, N = 4)^4 = 1 - \left(\frac{5}{6}\right)^4 \simeq 0,52$$

ou seja, ligeiramente superior a 50%.

Esse resultado poderia ser obtido, a partir da contagem de vezes que pelo menos um *ás* ( $i = 1$ ) ocorre na listagem das 1296 ( $6^{N=4}$ ) possíveis sequências dos resultados de eventos equiprováveis nos quatro lançamentos de um dado.

(1 1 1 1), (1 1 1 2), ... (1 1 1 6), (1 1 2 1), ... (1 1 2 6), ... (1 1 6 1), ... (1 1 6 6) ... (1 6 6 6)

(2 1 1 1), (2 1 1 2), ... (2 1 1 6), (2 1 2 1), ... (2 1 2 6), ... (2 1 6 1), ... (2 1 6 6) ... (2 6 6 6)

(3 1 1 1), (3 1 1 2), ... (3 1 1 6), (3 1 2 1), ... (3 1 2 6), ... (3 1 6 1), ... (3 1 6 6) ... (3 6 6 6)

(4 1 1 1), (4 1 1 2), ... (4 1 1 6), (4 1 2 1), ... (4 1 2 6), ... (4 1 6 1), ... (4 1 6 6) ... (4 6 6 6)

(5 1 1 1), (5 1 1 2), ... (5 1 1 6), (5 1 2 1), ... (5 1 2 6), ... (5 1 6 1), ... (5 1 6 6) ... (5 6 6 6)

(6 1 1 1), (6 1 1 2), ... (6 1 1 6), (6 1 2 1), ... (6 1 2 6), ... (6 1 6 1), ... (6 1 6 6) ... (6 6 6 6)

Esse processo trabalhoso, seria impraticável no caso de 24 lançamentos ( $N = 24$ ) de dois dados. No entanto, sabendo-se que a probabilidade de ocorrência de dois *azes*,  $P(i = 2, N = 1)$ , em um único lançamento de dois dados é igual a

$$P(i = 2, N = 1) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$$

a probabilidade de não ocorrência de duplo *azes* é dada por

$$P(i \neq 2, N = 1) = 1 - \frac{1}{36} = \frac{35}{36}$$

Assim, em 24 lançamentos ( $N = 24$ ) de dois dados, a probabilidade de ocorrência de pelo menos dois *azes* é dada por

$$P(i = 2, N = 24) = 1 - \left(\frac{35}{36}\right)^{24} \simeq 0,49$$

ou seja, ligeiramente inferior a 50%.

O fato de que muitos resultados, apesar de conhecidos pelos jogadores mais experientes da época, não serem por eles explicados, mostra que eventos compostos podem ser complexos demais para que suas alternativas ou resultados equiprováveis sejam diretamente determinados.

No entanto, utilizando-se as propriedades estabelecidas para a união e a intersecção de conjuntos de eventos, pode-se fazer prever probabilidades associadas a eventos compostos. Por

exemplo, a partir dos 36 pares ordenados de resultados possíveis e equiprováveis no lançamento de dois dados,

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

pode-se definir os seguintes conjuntos de pares de eventos

- $A = \{(i, j) \mid i + j = 7\}$  (pares com soma igual a 7)
- $B = \{(i, j) \mid i = 1\}$  (o primeiro valor é igual a 1)
- $C = \{(i, j) \mid i = 1 \text{ e } j \leq 4\}$  (o primeiro é 1 e o segundo menor ou igual a 4)

Desse modo, de acordo com as regras de cálculo de probabilidades, pode-se escrever para cada conjunto de eventos

- $P(A) = \frac{6}{36} = \frac{1}{6}$

- $P(B) = \frac{6}{36} = \frac{1}{6}$

- $P(C) = \frac{4}{36} = \frac{1}{9}$

- $A$  e  $B$  independentes  $\implies \begin{cases} P(A \text{ e } B) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36} \\ P(A \text{ ou } B) = \frac{1}{6} + \frac{1}{6} - \frac{1}{36} = \frac{11}{36} \end{cases}$

- $A$  e  $C$  excludentes  $\implies \begin{cases} P(A \text{ e } C) = 0 \\ P(A \text{ ou } C) = \frac{1}{6} + \frac{1}{9} = \frac{5}{18} \end{cases}$

- $C \subset B \implies \begin{cases} P(C \text{ e } B) = \frac{4}{36} = \frac{1}{9} \\ P(C \text{ ou } B) = \frac{1}{9} + \frac{1}{6} - \frac{1}{9} = \frac{1}{6} \end{cases}$

## 1.4 Probabilidades a posteriori

A probabilidade de  $1/6$ , que pode ser *a priori* atribuída à ocorrência de um determinado resultado no lançamento de um dado não viciado, pode ser obtida, também, *a posteriori*.

Na distribuição de resultados de um grande número de lançamentos, é possível verificar que a proporção de ocorrências de cada face é da ordem de  $1/6$ . Diz-se que a **distribuição de frequência** de ocorrência nesse tipo de fenômeno é uniforme.

face ( $i$ )	freq. ( $n_i$ )
1	22
2	16
3	22
4	22
5	17
6	21

Tabela 1: Distribuição de frequência das faces de um dado em 120 lançamentos.

A Fig. 1 mostra o **histograma** correspondente à Tab. 1, de distribuição de frequência ( $n_i$ ) de cada uma das faces ( $i$ ) de um dado, resultante da simulação de  $N = 120$  lançamentos.

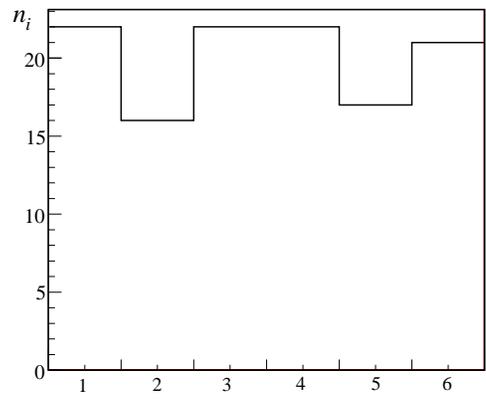


Figura 1: Histograma da distribuição de frequências das faces do dado correspondente à Tab. 1.

Desse modo, a medida que o número de lançamentos aumenta, a **frequência relativa** ( $f_i$ ) de ocorrência de uma face  $i$ ,

$$f_i = \frac{n_i}{N}$$

aproxima-se cada vez mais de um número definido entre 0 e 1, nesse caso, igual a  $1/6$ .

Considerando que o número total  $N$  de medidas  $x_i$  de uma grandeza  $x$  em um experimento, ou que o número  $n_i$  de ocorrências de um evento seja “suficientemente grande”, a probabilidade de ocorrência  $P(x_i)$  da medida ou do evento em questão é definida como o limite da razão entre a frequência de observação do evento e o número total de eventos observados, ou seja, como o limite da frequência relativa,

$$P(x_i) = \lim_{N \gg 1} f_i = \lim_{N \gg 1} \frac{n_i}{N}$$

A concepção frequentista ou estatística de probabilidade *a posteriori*, introduzida em 1718 pelo francês Abraham De Moivre (1667-1754), na obra *The Doctrine of Chances*, foi adotada pelo inglês John Venn (1834-1923), em 1822, e sistematizada pelo austríaco Richard von Mises (1883-1953), em 1931[13], e pode ser aplicada mesmo quando o número total de eventos possíveis seja infinito, bastando que a frequência relativa aproxime-se de um limite experimental.

Do mesmo modo que o valor médio ( $\bar{x}$ ) de uma coleção  $\{x_i\}$  de  $N$  valores distribuídos segundo um conjunto de frequências  $\{n_i\}$  é caracterizado pela média dos valores ponderados pelas respectivas frequências relativas ( $f_i = n_i/N$ ),

$$\bar{x} = \sum_{i=1}^N x_i f_i$$

o **valor médio**,  $\langle x \rangle$ , associado a uma coleção  $\{x_i\}$  de  $N$  valores associados a uma distribuição de probabilidades  $\{P(x_i)\}$  é dado por

$$\langle x \rangle = \sum_{i=1}^N x_i P(x_i)$$

Analogamente, a dispersão desses valores pode ser caracterizada pela **variância** ( $\sigma_x^2$ ), definida por

$$\sigma_x^2 = \sum_{i=1}^N (x_i - \langle x \rangle)^2 P(x_i)$$

A raiz quadrada positiva da variância, denotada por  $\sigma_x$ , é o chamado **desvio-padrão**.

Do ponto de vista prático, a variância pode ser calculada por

$$\sigma_x^2 = \langle x^2 \rangle - \langle x \rangle^2$$

onde

$$\langle x^2 \rangle = \sum_{i=1}^N x_i^2 P(x_i)$$

é a média dos quadrados dos valores da coleção  $\{x_i\}$

### 1.4.1 Distribuição de probabilidades contínua

Para o caso de grandezas cujas variações são hipoteticamente contínuas, como as coordenadas ( $x$ ) de uma pedra atirada aleatoriamente sobre um canaleta, os dados observados, que são as medidas das coordenadas, podem ser agrupados em intervalos discretos  $[x_k, x_{k+1})$ , de amplitudes  $\Delta x_k = x_{k+1} - x_k$ .<sup>4</sup> Desse modo, se  $n_k$  é a frequência das medidas associadas ao intervalo  $(x_k, x_{k+1})$ , obtém-se o histograma das coordenadas mostrado na Fig. 2.

Assim, a probabilidade de ocorrência de uma medida da coordenada  $x$  em um dado intervalo  $[x_k, x_{k+1})$ , denotada por  $P(x_k \leq x < x_{k+1})$ , é dada pela razão entre a área hachureada ( $n_k \cdot \Delta x_k$ )

<sup>4</sup> Em geral, mas não necessariamente, esses intervalos possuem a mesma amplitude.

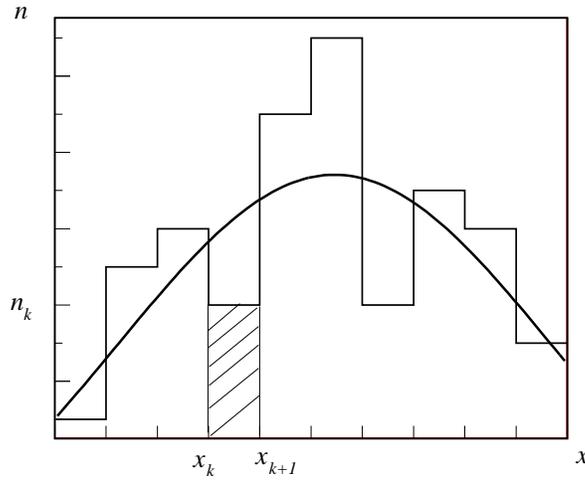


Figura 2: Histograma das coordenadas ( $x$ ) de uma pedra e a distribuição de frequência limite.

da parte do histograma correspondente ao dado intervalo e a área total  $\left(\sum_{i=1}^N n_i \cdot \Delta x_i = A\right)$  do histograma, ou seja,

$$P(x_k \leq x < x_{k+1}) = \frac{n_k \cdot \Delta x_k}{\sum_{i=1}^N n_i \cdot \Delta x_i} = \frac{n_k}{A} \cdot \Delta x_k \quad \Longrightarrow \quad \sum_{k=1}^N P(x_k \leq x < x_{k+1}) = 1$$

onde  $N$  é o número total de medidas.

A representação gráfica dos termos  $\rho(x_k) = n_k/A$  associados aos intervalos  $[x_k, x_{k+1})$  é denominada **distribuição normalizada de frequência**.

No limite de um grande número de observações ( $N \gg 1$ ), os intervalos  $[x_k, x_{k+1})$  ou **classes de frequência** podem ser tais que a amplitude  $\Delta x_k$  de cada classe seja tão pequena quanto se queira. Nessas condições, os termos limites

$$\rho(x)|_{x_k} = \lim_{\Delta x_k \rightarrow 0} \frac{P(x_k < x < x_{k+1})}{\Delta x_k} = \lim_{N \gg 1} \frac{n_k}{A} \quad k = 1, 2, 3, \dots, N \gg 1$$

definem uma distribuição contínua normalizada de frequências relativas por unidade de medida da coordenada  $x$ , denominada **função densidade de probabilidade**,  $\rho(x)$ , para os possíveis resultados contínuos de  $x$ , tal que a probabilidade associada a medida de  $x$  entre dois valores  $a$  e  $b$  é dada por

$$P(a < x < b) = \int_a^b \rho(x) dx$$

e, a chamada **condição de normalização** da distribuição, por

$$\int_{x_{\min}}^{x_{\max}} \rho(x) dx = 1$$

onde  $x_{\min}$  e  $x_{\max}$  são, respectivamente, o mínimo e o máximo valores possíveis para a medida de  $x$ , que delimitam o domínio ( $D$ ) de valores possíveis para as ocorrências de  $x$ .

Quaisquer outras questões referentes à variável aleatória  $x$  pode se obtida a partir dessa distribuição  $\rho(x)$ . Por exemplo, o valor médio de  $x$  é dado por

$$\langle x \rangle = \int_D x \rho(x) dx$$

e o desvio-padrão, por

$$\sigma_x = \sqrt{\int_D (x - \langle x \rangle)^2 \rho(x) dx} = \sqrt{\langle x^2 \rangle - \langle x \rangle^2}$$

Por exemplo, seja  $x$  uma variável contínua, como as possíveis posições de uma partícula confinada em uma região de dimensão  $a$ , cuja densidade de probabilidade  $\rho(x)$  é proporcional a função  $\text{sen}^2 \frac{\pi}{a} x$ .

Determinando-se a constante de normalização,

$$\int_0^a \rho(x) dx = A \int_0^a \text{sen}^2 \frac{\pi}{a} x dx = \frac{A}{2} \int_0^a \left(1 - \cos \frac{2\pi}{a} x\right) dx = 1 \implies A = \frac{2}{a}$$

o valor médio da posição é dado por

$$\langle x \rangle = \frac{2}{a} \int_0^a x \text{sen}^2 \frac{\pi}{a} x dx = \frac{1}{a} \left\{ \underbrace{\int_0^a x dx}_{a^2/2} - \underbrace{\int_0^a x \cos \frac{2\pi}{a} x dx}_0 \right\} = \frac{a}{2}$$

a média dos quadrados por

$$\left\{ \begin{aligned} \langle x^2 \rangle &= \frac{1}{a} \left\{ \int_0^a x^2 dx - \int_0^a x^2 \cos \frac{2\pi}{a} x dx \right\} \\ &= \frac{1}{a} \left\{ \frac{a^3}{3} - \frac{a}{\pi} \int_0^a x \text{sen} \frac{2\pi}{a} x dx \right\} \\ &= \left\{ \frac{a^2}{3} + \frac{a}{2\pi^2} x \cos \frac{2\pi}{a} x \Big|_0^a \right\} = \frac{a^2}{3} + \frac{a^2}{2\pi^2} \end{aligned} \right.$$

e, o desvio-padrão, por

$$\left\{ \begin{aligned} \sigma_x^2 &= \langle x^2 \rangle - \langle x \rangle^2 = \frac{a^2}{3} + \frac{a^2}{2\pi^2} - \frac{a^2}{4} = a^2 \left( \frac{1}{12} + \frac{1}{2\pi^2} \right) \\ \sigma_x &= \frac{a}{2\sqrt{3}} \sqrt{1 + \frac{6}{\pi^2}} \end{aligned} \right.$$

Além de estar associada à distribuição de valores da variável  $x$ , a densidade de probabilidade  $\rho(x)$  determina também a expectativa de ocorrência dos valores de qualquer outra função  $f(x)$ . Por exemplo, se  $x$  representa as possíveis medidas para o raio de um círculo, os possíveis valores para a área  $f(x) = \pi x^2$  estarão associadas à mesma distribuição de probabilidade  $\rho(x)$ , de modo que o valor médio  $\langle f \rangle$  será dado por

$$\langle f \rangle = \int_D f(x) \rho(x) dx$$

e, a respectiva variância ( $\sigma_f^2$ ), por

$$\sigma_f^2 = \langle f^2 \rangle - \langle f \rangle^2$$

## 1.5 Probabilidade axiomática

O conceito de probabilidade foi apropriadamente axiomatizado no início do século XX[9], em 1933, pelo matemático russo Andrei Nikolaevich Kolmogorov (1903-1987), ao estabelecer que qualquer conjunto discreto de números não-negativos  $\{P_i\}$ , que esteja associado a um outro conjunto discreto de elementos  $S = \{x_1, x_2, x_3, \dots\}$ , como os possíveis valores para as medidas de uma grandeza, tal que

$$\sum_i P(x_i) = 1 \quad (\text{condição de normalização})$$

e, para elementos distintos,<sup>5</sup>

$$P(x_i \text{ ou } x_j) = P_i + P_j \quad (x_i \neq x_j)$$

é denominado um conjunto de probabilidades  $\{P_i\}$ , associado ao conjunto  $S = \{x_k\}$ .

Ou seja, do ponto de vista matemático, não é necessário atribuir nenhum significado ao conceito de probabilidade, como a expectativa ou a crença na ocorrência de um evento; basta que se associe aos possíveis resultados independentes de um experimento, como as medidas de uma grandeza, um número positivo menor que a unidade, cuja adição sobre todos os possíveis eventos (igual à unidade), expressa a certeza em se obter um dos possíveis resultados (condição de normalização).

Nesse sentido, tanto a concepção clássica *a priori* de Laplace (expectativa de ocorrência), como a concepção frequentista *a posteriori* de von Mises satisfazem aos chamados axiomas de Kolmogorov para probabilidades discretas.

---

<sup>5</sup> Na linguagem estatística, eventos mutuamente excludentes.

## 2 Física e probabilidades

### 2.1 Teorias probabilísticas

Ao contrário das atividades experimentais, a introdução de métodos estatísticos e probabilidades no domínio da Física Teórica foi, e ainda é, questão de grandes controvérsias.

Do ponto de vista experimental, a determinação do valor da componente  $A$  de uma grandeza associada a um sistema em uma dada condição (**estado**), em geral, resulta de um conjunto  $\{a_1, a_2, \dots, a_N\}$  de  $N$  medidas diretas de  $A$ . Se  $n_j$  é o n.º de ocorrências (frequência) de uma medida particular  $a_j$ , diz-se que o **valor esperado** de  $A$  é dado pelo valor médio  $\langle A \rangle$ ,

$$\langle A \rangle = \sum_{j=1}^m a_j f_j$$

onde  $f_j = n_j/N$  é a frequência relativa associada a uma medida  $a_j$ , e  $m$  é o n.º de medidas distintas.

A **dispersão** ou **incerteza**  $\Delta A$  associada ao conjunto de medidas é caracterizada pela raiz quadrada da variância, ou seja, pelo desvio-padrão,

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

onde  $\langle A^2 \rangle = \sum_{j=1}^m a_j^2 f_j$ .

Uma vez que o processo de medição é aleatório, no limite de um grande número de medidas, as frequências relativas representam as probabilidades associadas às respectivas medidas.

De modo equivalente, uma descrição ou teoria probabilística para o estudo dos fenômenos ou sistemas físicos deve prover regras, ou basear-se em hipóteses, que permitam calcular:

- os valores (medidas) possíveis para as componentes das grandezas associadas ao fenômeno ou ao sistema físico;
- as respectivas probabilidades de ocorrência, ou a distribuição dessas probabilidades.

As distribuições de probabilidades associadas às medidas de uma grandeza física deveriam ser determinadas *a posteriori*, uma vez que a aleatoriedade de um processo de medição não pode ser descrita ou prevista por qualquer teoria. No entanto, devido ao próprio grau de desconhecimento em relação ao processo, podem ser estabelecidas distribuições de probabilidades *a priori* (**binomial, Gauss e Poisson**)[18] tão gerais que servem de fundamentos para a construção dos métodos estatísticos para a análise de erros.

Por outro lado, o conceito *a priori* de probabilidade é amplamente usado na fundamentação de estudos teóricos, nos quais são utilizados conceitos probabilísticos, como no desenvolvimento da Mecânica Estatística[11, 14]. Nesses casos, distribuições de probabilidades especiais (**Maxwell-Boltzmann, Planck, Fermi-Dirac e Bose-Einstein**), adequadas à descrição de sistemas de muitas partículas quase-independentes, são deduzidas e mostram-se compatíveis com os comportamentos experimentais observados em diversos sistemas como os gases moleculares a baixa pressão, os elétrons em metais e semicondutores e, ainda, em diversos fenômenos, como a variação do calor específico dos sólidos a baixas temperaturas e a radiação de corpo negro.

Uma abordagem mais geral da Mecânica Estatística, iniciada pelo físico austríaco Ludwig Boltzmann (1844-1906) e sintetizada pelo físico americano Josiah Willard Gibbs (1839-1903), em 1901[3, 8, 11, 14], utilizando-se também do conceito *a priori* de probabilidade, permite o estabelecimento de distribuições de probabilidades gerais (**microcanônica, canônica e gran-canônica**) não associadas apenas a gases. Desse modo, para cada classe de problema, a partir das formas gerais dessas distribuições, é possível a obtenção de uma distribuição de probabilidade específica.

Até o primeiro quarto do século XX, a utilização na Física Teórica do conceito *a priori* de probabilidade, estava associado a problemas estatísticos, devido à impossibilidade prática da caracterização simultânea do estado de todas as partículas de um sistema com um grande número de graus de liberdade. Ou seja, o problema resultava da complexidade dos sistemas observados. Com o surgimento da teoria da Mecânica Quântica, que utiliza também um conceito *a priori* de probabilidade, o não determinismo passa a ser uma característica intrínseca da evolução ou descrição dos fenômenos e sistemas, mesmos aqueles com poucos graus de liberdade. A Mecânica Quântica estabelece que, para cada problema, há de se calcular uma distribuição de probabilidade que conterá as informações necessárias à descrição do fenômeno estudado.

## 2.2 Analogias e dualidade

De acordo com o filósofo de ciência austríaco Karl Popper (1902-1994)[16], uma das principais fontes de mal entendidos em uma abordagem probabilística da Física Moderna é a não discriminação entre categorias distintas, ou seja, a associação dos comportamentos análogos de algumas propriedades<sup>6</sup> dos elementos de um sistema com a natureza desses elementos.

Por exemplo, a distribuição (gaussiana) dos pesos das pessoas residentes em uma certa rua não revela ou está associada a nenhuma característica intrínseca de cada indivíduo. Os salários dos leitores de um jornal ou as notas de prova dos alunos de uma escola, em geral, apresentam essa mesma distribuição. Apesar dos elementos de cada conjunto terem todos um “caráter gaussiano”, ou seja, seus atributos apresentarem comportamentos análogos, suas naturezas são completamente distintas.

Assim, os elementos de sistemas que apresentam comportamentos análogos, no sentido de que alguns de seus atributos (grandezas) obedecem à mesma distribuição (equação), não possuem, necessariamente, a mesma natureza.

Um outro exemplo desse tipo de analogia é o das partículas sub-atômicas, como os elétrons ( $e$ ), com a luz. O fato do campo  $\Psi$ , denominado **função de onda**, associado a uma partícula como o elétron obedecer a uma equação diferencial linear<sup>7</sup> reflete ou implica propriedades que não devem ser atribuídas à sua natureza. Entretanto, devido a não visualização dos fenômenos microscópicos, ao se analisar um experimento imaginário de dupla fenda de Young com partículas, usualmente, considera-se a distribuição resultante das partículas como consequência ou evidência de suas naturezas, em certas situações, ondulatórias. Ou seja, de um ponto de vista não-quântico, as partículas teriam um comportamento ambíguo, uma natureza dual. Esse argumento foi utilizado no século XIX para “estabelecer” o caráter ondulatório e a natureza contínua dos fenômenos óticos e da luz.

Apesar da descrição ondulatória, ou seja, da equação que descreve a propagação de ondas

---

<sup>6</sup> Por exemplo, uma função densidade de probabilidade de ocorrência das medidas de uma grandeza.

<sup>7</sup> E, portanto, obedecer ao Princípio da Superposição. Para Dirac, esse é o princípio fundamental da Mecânica Quântica, a partir do qual se estabelece a estrutura linear da teoria.

acústicas, a equação de d'Alembert, poder ser deduzida a partir de um enfoque microscópico, baseado nas leis da Mecânica Clássica, como um limite para as pequenas vibrações de um grande número de partículas constituintes de um meio, esse fato não suscitou grandes discussões e controvérsias sobre a questão da natureza contínua ou discreta do fenômeno ondulatório acústico.

Na tentativa para se construir uma imagem da propagação de uma onda eletromagnética no vácuo, em analogia com os fenômenos acústicos, associou-se ao espaço uma estranha subestrutura, o *éter*. Esse meio teria propriedades mecânicas muito especiais para suportar as oscilações de altíssimas frequências ( $10^{14}$  Hz) dos campos eletromagnéticos associados à propagação da luz. O éter foi abolido dos modelos físicos a partir dos experimentos de Michelson e Morley, em 1887, e das análises de Albert Einstein (1879-1955),<sup>8</sup> em 1905, para estabelecer a Teoria da Relatividade Restrita.

Por não ser necessário a existência de um meio material para a propagação da luz, ou de qualquer perturbação eletromagnética, a descrição dos fenômenos eletromagnéticos baseia-se apenas nas equações de Maxwell, e não em modelos materiais mecânicos. A propagação da luz passou a ser encarada, então, como um processo que obedece a uma equação de onda, e que, portanto, reflete a natureza contínua do processo ondulatório de transporte de energia entre dois pontos do espaço, sem, necessariamente, a presença de matéria.

Apesar do caráter não comprobatório das analogias, elas são frutíferas pois, foi a partir de tentativas para se estabelecer analogias mais do que formais entre os fenômenos ondulatórios acústicos e eletromagnéticos,<sup>9</sup> que se configurou a grande crise na Física, ao final do século XIX<sup>10</sup>. Dessa crise, surgiram as duas teorias fundamentais da Física Moderna: a Teoria da Relatividade Restrita e a Mecânica Quântica, que exigiram a revisão e a modificação de vários conceitos geométricos e dinâmicos, acerca da natureza, criados e aceitos até então pelo homem.

Com a criação da Mecânica Quântica, ao final do primeiro quarto do século XX, surgiu uma nova teoria fundamental da Física, a qual, a partir de então, qualquer modelo mecânico, material ou não, teria que obedecer. No entanto, a utilização de analogias baseadas na Mecânica Clássica se constituiu em um procedimento tão forte, durante a construção da teoria quântica, que mascarou várias questões de cunho interpretativo, que constituem motivos de controvérsias até os dias de hoje, como em nenhuma outra teoria da Física.

O resultado dessas analogias é que tanto a Teoria da Relatividade Restrita quanto a Mecânica Quântica possuem estruturas formais condicionadas pela Mecânica Clássica, através de correspondências estabelecidas para a forma e expressão de seus conceitos e grandezas.

---

<sup>8</sup> Sobre os conceitos de espaço, tempo, massa e energia.

<sup>9</sup> O que no fundo seria a explicação dos fenômenos elétricos a partir de modelos materiais mecânicos que obedecessem às leis de movimento de Newton da Mecânica Clássica, ou seja, uma redução do Eletromagnetismo à Mecânica.

<sup>10</sup> Essa foi, essencialmente, uma crise da Mecânica Clássica como teoria fundamental, e se estendeu até que se percebesse que a Mecânica Clássica é que teria que ser modificada ou substituída por outras teorias.

## 3 Origens da Mecânica Quântica

### 3.1 A quantização de Planck da energia

A questão da natureza contínua ou discreta dos fenômenos remonta à própria gênese da Mecânica Quântica, quando o alemão Max Planck (1858-1947), em 1900, em seus primeiros passos para elucidar o problema da **radiação do corpo negro** em equilíbrio térmico, considera que os valores das energias ( $\varepsilon$ ) dos átomos osciladores que o constituíam só podiam ser múltiplos de uma certa quantidade mínima, proporcional à frequência  $\nu$  da radiação[4, 15, 20]. Ou seja, só poderiam assumir valores múltiplos de um *quantum* de energia  $h\nu$ ,

$$\varepsilon_n = nh\nu \quad \{n = 0, 1, 2, \dots\}$$

onde  $h \simeq 6.63 \times 10^{-34}$  J.s é a constante de Planck.

A **quantização da energia** de um sistema mecânico oscilante marca a introdução de uma idéia que contrariava um resultado básico da Física Clássica, o fato da energia de um sistema poder assumir qualquer valor real, ou seja, variar continuamente. Essa hipótese é tida por muitos pesquisadores, historiadores e filósofos da ciência a idéia revolucionária que se constituiu num divisor marcante na evolução das idéias e conceitos da Física.

Assim como a velocidade da luz no vácuo ( $c \simeq 3.0 \times 10^8$  m/s) é a constante fundamental da Física que estabelece o limite para uma descrição não-relativística dos fenômenos, a constante de Planck é a constante fundamental que estabelece um limite clássico não quântico.

Em primeira aproximação, diz-se que um fenômeno é quântico se apresentar uma combinação de grandezas relevantes que tenha a mesma dimensão e valor comparável à constante de Planck.

### 3.2 A quantização de Einstein da luz

Abordando o problema da radiação com um enfoque distinto de Planck, Einstein[4, 7], em 1905, a partir da Lei de Wien, e concentrando-se na própria radiação como Lord Rayleigh (1842-1919)[17], estende a idéia de quantização à própria radiação.

Mostrando que a energia de uma onda eletromagnética de frequência  $\nu$  estaria distribuída discretamente por corpúsculos (fótons) de energia  $\varepsilon = h\nu$ , onde a constante de proporcionalidade era a constante de Planck, ele explica as propriedades básicas do **efeito fotoelétrico**, descoberto pelo alemão Heinrich Rudolf Hertz (1857-1894), em 1887, e, extensivamente, estudado pelo húngaro Philipp Lenard (1862-1947), em 1900.

Desse modo, a interação da radiação com a matéria pôde ser parcialmente descrita como resultado das interações individuais de fótons e elétrons e, assim, foi reforçada a idéia de que a luz, em alguns processos, possuía um caráter corpuscular.

A **quantização da radiação** por Einstein foi realizada a partir de argumentos termodinâmicos e probabilísticos. No entanto, esse caráter corpuscular pode ser estabelecido por um procedimento utilizado no estudo da propagação de ondas acústicas em sólidos. O fenômeno decorre das pequenas vibrações dos átomos em torno de suas posições de equilíbrio em uma rede cristalina. Nesse caso, pode-se associar ao conjunto de átomos da rede um conjunto de osciladores acoplados e, posteriormente, expressar a energia total do conjunto por uma soma da energia de osciladores virtuais independentes (desacoplados). Utilizando-se a quantização

de Planck para os osciladores virtuais, os fenômenos de natureza acústica são associados ao movimento de quase-partículas (fónons) cujas energias são proporcionais às frequências próprias de oscilação da rede ou de propagação da onda acústica na rede cristalina.

De maneira análoga, a quantização da energia de uma onda eletromagnética equivale a representá-la também por um conjunto de osciladores virtuais desacoplados e, posteriormente, introduzir os *quanta* de energia do campo, os fótons.

Apesar do americano Robert Millikan (1868-1953), em 1914, ter determinado a constante de Planck a partir da teoria de Einstein para o efeito fotoelétrico[12], o argumento de quantização da radiação só se tornou convincente mais tarde, quando o também americano Arthur Compton (1892-1962), em 1923, descobre um outro mecanismo pelo qual os raios X são espalhados pela matéria, perdendo parte da energia, com mudanças no comprimento de onda e frequência [4, 5]. Classicamente, isso não ocorreria pois, a frequência de uma onda não sofre variação após a interação com a matéria, ou seja, a frequência não pode ser alterada pela mudança na direção de propagação de uma onda.

O fenômeno de mudança da frequência da radiação ao ser espalhada pela matéria, doravante chamado **efeito Compton**, foi interpretado pelo próprio Compton como devido às colisões de fótons com os elétrons livres dos átomos do meio, reforçando novamente a idéia de que a luz teria um caráter corpuscular.

### 3.3 O átomo de Bohr

Ao final do século XIX, um outro problema que, posteriormente, reforçou a idéia de quantização da energia de um sistema foi a separação da luz emitida por gases aquecidos, em suas componentes espectrais, ou seja, o estudo do **espectros dos gases**. Nesse contexto, foram estabelecidas regras que culminaram com a fórmula de de Balmer (1885) e Rydberg (1890) para o espectro de emissão do hidrogênio,

$$\frac{1}{\lambda} = R \left( \frac{1}{2^2} - \frac{1}{n^2} \right)$$

onde  $\lambda$  é o comprimento de onda da luz emitida,  $n$  é um inteiro maior que 3, e  $R \simeq 1.097 \times 10^7 \text{ m}^{-1}$  é a constante de Rydberg.

Apresentando um modelo atômico no qual cargas negativas (elétrons) oscilavam em um substrato positivamente carregado, o inglês J. J. Thomson (1856-1940), em 1903, tenta explicar o fenômeno identificando as frequências de oscilações dos elétrons com as frequências da luz emitida.

Uma explicação mais conveniente e consistente para o espectro do hidrogênio só foi obtida a partir de uma série de experimentos de espalhamentos, nos quais Rutherford, Geiger e Marsden, em 1909, observaram que, muitas vezes, as trajetórias das chamadas partículas alfas ( $\alpha$ )<sup>11</sup> sofriam fortes desvios quando bombardeavam uma folha fina de ouro. Com base nesses experimentos, Rutherford, em 1911, propõe um modelo atômico nuclear, análogo ao modelo planetário, e o dinamarquês Niels Bohr (1885-1962), em 1913, utilizando-se da quantização da energia do oscilador de Planck, deriva os resultados do suíço Johan Balmer (1825-1898)[1, 4].

---

<sup>11</sup> Núcleos de hélio emitidos por materiais radioativos.

### 3.4 A equação de Schrödinger

Uma vez “estabelecido” o caráter corpuscular de uma onda eletromagnética, um outro passo decisivo na evolução da Mecânica Quântica foi a hipótese do francês Louis de Broglie (1892-1987), em 1924.

Uma onda eletromagnética plana monocromática  $\Psi$ , de frequência  $\nu$  e comprimento de onda  $\lambda$ , que se propaga no vácuo, com velocidade  $c$ , na direção do vetor  $\vec{k}$ , onde  $k = \omega/c = 2\pi\nu/c = 2\pi/\lambda$ , pode ser descrita por

$$\Psi \sim e^{i(\vec{k}\cdot\vec{r} - \omega t)}$$

Associando-se um momentum  $\vec{p} = \hbar\vec{k} = (h/\lambda) \hat{k}$ , e uma energia  $E = \hbar\omega = h\nu$ , que obedecem a relação de Einstein para uma partícula sem massa,  $E = pc$ , pode-se representá-la por

$$\Psi \sim e^{i(\vec{p}\cdot\vec{r} - Et)/\hbar}$$

De maneira recíproca, uma vez que a função de onda  $\Psi$  associada a uma partícula livre não-relativística de momentum  $\vec{p}$  e energia  $E$  possui uma dependência funcional análoga, mas com  $E = p^2/2m$ , é costume associar-se à partícula um comprimento de onda  $\lambda = h/p$  e uma frequência  $\nu = E/h$ , que obedecem à relação de dispersão  $\nu = h/(2m\lambda^2)$ .

Com base nessas idéias, alguns anos depois, os americanos Clinton Davisson (1881-1958) e Lester Germer (1896-1971)[4, 6], em 1927, realizaram a experiência de **difração de elétrons**.

As hipóteses de Louis de Broglie e Einstein originaram a idéia de uma “dualidade onda-partícula” que, segundo o austríaco Erving Schrödinger (1887-1961)[19], foi fundamental para que ele, em 1926, estabelecesse a equação diferencial linear que descreve o comportamento de uma partícula não-relativística de massa  $m$ , em um campo conservativo  $V(\vec{r})$ , em uma dada região do espaço,<sup>12</sup>

$$\left[ -\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r}) \right] \psi(\vec{r}) = E\psi(\vec{r})$$

onde  $\hbar = h/2\pi$ , e  $\psi(\vec{r})$  é a função de onda associada à partícula.

A partir dessa equação, Schrödinger resolveu os problemas da quantização da energia do elétron no átomo de hidrogênio e do oscilador harmônico.

No processo de construção e aprendizagem da Mecânica Quântica, pode-se lançar mão de vários argumentos que tornam a equação de Schrödinger mais ou menos plausível. Entretanto, como as leis de Newton ou as equações de Maxwell, por ser o alicerce de uma teoria fundamental da Física, a partir dela é que devem e são deduzidas as principais características da Mecânica Quântica não-relativística da partícula ou de um sistema de partículas.

Apesar da grande ruptura com conceitos alguns dos da Mecânica Clássica (como os de trajetória ou velocidade de uma partícula), um procedimento que permaneceu na formulação da Mecânica Quântica é o método analítico newtoniano de focalizar ou isolar um objeto de estudo e representar suas interações com o restante do Universo (ou sua vizinhança) por um campo de forças que atua sobre o objeto.

---

<sup>12</sup> Apesar da importância histórica dessa analogia, cabe realçar que ela é simplesmente um procedimento mnemônico que não justifica qualquer teoria.

Essa característica só foi modificada a partir dos trabalhos dos americanos Richard Feynman (1918-1988) e Julian Schwinger (1918-1994), e do japonês Shin-ichiro Tomonaga (1906-1979), de elaboração da Eletrodinâmica Quântica (QED),<sup>13</sup> em 1947, quando foram estabelecidos métodos aproximados (perturbativos) que permitiram o cálculo de processos de interações entre partículas que incorporam de modo eficaz suas ações recíprocas.

### 3.5 A interpretação probabilística de Born

Com a evolução da Mecânica Quântica, a função de onda  $\Psi$  associada às partículas já não seria mais identificada como uma onda no sentido acústico ou eletromagnético pois, na generalização da equação de Schrödinger para sistemas de  $N$  partículas, a função de onda  $\Psi$  é um campo definido no espaço de configuração, *i. e.*, uma função de  $N$  variáveis.

De fato, o alemão Max Born (1882-1970), em 1926, guiado pela identificação da intensidade de uma onda<sup>14</sup> eletromagnética como uma densidade de fótons numa dada região, interpreta a função de onda  $\Psi$  de uma partícula como uma **amplitude de probabilidade** de presença associada à partícula, de modo que o quadrado de seu módulo seria igual a uma **densidade de probabilidade**  $\rho$ , ou seja,

$$\rho = |\Psi|^2 = \Psi^* \Psi$$

Segundo Born[2], seu argumento baseou-se em uma analogia devido a Einstein. Se em um experimento de espalhamento da luz por um dielétrico admite-se que a luz espalhada é composta de fótons, a intensidade do campo que a descreve (igual ao quadrado do módulo da amplitude do campo) determina a distribuição de fótons espalhados. De maneira análoga, se em um experimento de espalhamento de Rutherford interpreta-se o quadrado do módulo da função de onda associada aos elétrons espalhados como uma medida da distribuição desses elétrons, obtém-se a fórmula de Rutherford para a seção de choque.

A interpretação probabilística e a linearidade da equação de Schrödinger explica o porque do chamado comportamento ondulatório do elétron, como no clássico exemplo de Feynman do experimento de dupla fenda com elétrons (Fig. 3).

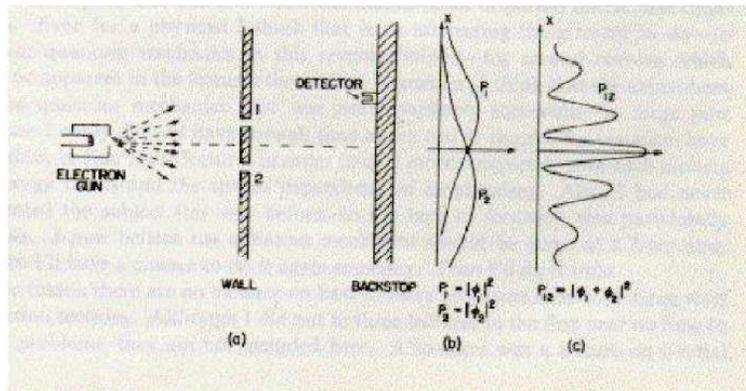


Figura 3: Experimento de dupla fenda com elétrons.

Seja  $x$  a coordenada sobre um eixo perpendicular ao anteparo e ao arranjo de fendas. Se  $\psi_1(x)$  é a função de onda associada ao comportamento de uma partícula quando passa pela fenda 1

<sup>13</sup> R. P. Feynman, *Quantum Electrodynamics - 1961*, reprinted by The Benjamin/Cummings Pub., 1983.

<sup>14</sup> Proporcional ao quadrado de sua amplitude.

(enquanto a fenda 2 está fechada), e  $\psi_2(x)$  é a função de onda associada ao comportamento da partícula quando passa pela fenda 2 (enquanto a fenda 1 está fechada), pelo princípio da superposição (linearidade da equação de Schrödinger), a função de onda associada à partícula quando ambas as fendas estão abertas será dada por

$$\psi(x) = \psi_1(x) + \psi_2(x)$$

e, de acordo com a interpretação de Born, a distribuição de probabilidade de presença da partícula ser detectada em qualquer ponto do anteparo é dada por

$$P = |\psi_1 + \psi_2|^2$$

Assim, a distribuição de probabilidade resultante é formalmente idêntica à distribuição da intensidade resultante no experimento de dupla fenda de Young, com a luz (Fig. 4).

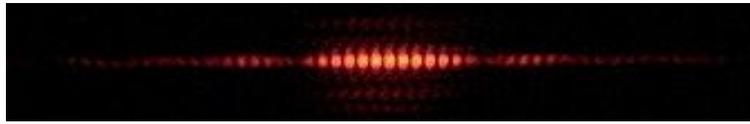


Figura 4: Interferência de Young.

Experimentos de interferência de partículas em arranjos de dupla fenda só foram realizados com nêutrons, em 1988, pelo austríaco Anton Zeilinger (n. 1945), e com elétrons, em 1989, pelo japonês Akira Tonomura, nos laboratórios da Hitachi (Fig. 5).

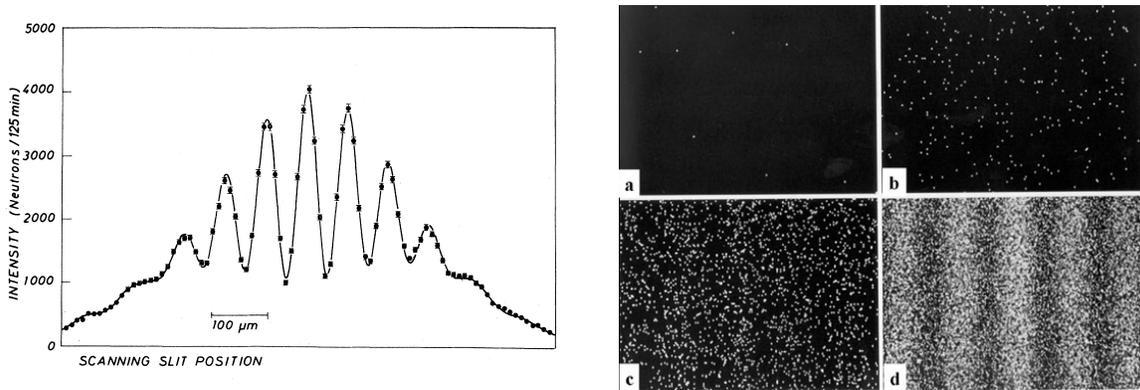


Figura 5: Resultados dos experimentos de Zeilinger, com nêutrons (à esquerda), e de Tonomura com elétrons (à direita).

No sentido estrito, o campo de probabilidades associado a uma partícula, ou a um conjunto de partículas, não deve ser identificado como uma propriedade ou característica intrínseca de uma partícula, mas sim como uma medida da distribuição de probabilidades de ocorrências de eventos associados a ela, que depende da interação com a sua vizinhança.

### 3.6 O princípio da incerteza de Heisenberg

Uma outra consequência da interpretação probabilística de Born para a Mecânica Quântica, implica que os resultados da medição de uma grandeza física associada a uma partícula, ou a um sistema de partículas, são aleatórios e, portanto, distribuídos em um intervalo ou espectro de valores.

Assim, não se pode atribuir ao movimento da partícula uma trajetória definida como na Mecânica Clássica pois, nem a posição e nem o *momentum* da partícula são univocamente definidos. O que se pode afirmar é que a localização de uma partícula está associada a uma certa região do espaço, onde ela se desloca com *momentum* indefinido.

Essa idéia foi expressa pelo alemão Werner Heisenberg (1901-1976), em 1927, como um princípio fundamental da Mecânica Quântica, que estabelece vínculos entre as incertezas associadas às correspondentes componentes da posição e do *momentum* de uma partícula, em torno de seus valores médios, quantificados pelas chamadas relações de incertezas de Heisenberg,

$$\begin{cases} \Delta x \Delta p_x \geq \hbar/2 \\ \Delta y \Delta p_y \geq \hbar/2 \\ \Delta z \Delta p_z \geq \hbar/2 \end{cases}$$

Para uma partícula, como o elétron em um átomo, confinada em uma região de dimensão linear da ordem de  $0.5 \times 10^{-8}$  cm (esse valor corresponderia à incerteza máxima na posição em relação ao núcleo), a incerteza mínima associada ao *momentum* seria da ordem de

$$(\Delta p)_{\min} \simeq 10^{10} \hbar \simeq 10^{-24} \text{ kg.m/s}$$

Esse valor é aproximadamente um valor mínimo para o próprio *momentum* da partícula. A partir desse valor, pode-se estimar, por exemplo, o valor mínimo para a energia do elétron no átomo de hidrogênio.

$$E_{\min} = \frac{(p_{\min})^2}{2m} \simeq 10^{-18} \text{ J} \simeq 10 \text{ eV}$$

### 3.7 Comentários sobre alguns livros de Mecânica Quântica

A abordagem inicial do livro texto, de Griffiths, é a de trabalhar em alguns problemas típicos que podem ser resolvidos pela Mecânica Quântica, antes de se tentar apresentar uma interpretação dos princípios básicos da teoria. A teoria não é introduzida a partir de uma pequena motivação histórica, como apresentada nessas notas. Entretanto, é apresentada uma boa discussão inicial, a ser retomada ao final do livro, sobre algumas interpretações da teoria. O livro é dividido em duas partes: *teoria* e *aplicações*. Pressupõe o que normalmente é exigido em seu nível de apresentação: Cálculo Avançado, noções de Álgebra Linear, Mecânica e Eletricidade elementares.

Três outros livros que contêm e desenvolvem de forma mais avançada os tópicos abordados no livro texto são:

- W. Greiner, *Quantum Mechanics – An Introduction*, Springer, 1994.
- C. Cohen-Tannoudji, B. Diu & F. Laloë, *Quantum Mechanics Vol I & II*, Wiley, 1977.
- P. A. M. Dirac, *The Principles of Quantum Mechanics*, Oxford University Press, 1976.

A parte histórica de motivação, como a quantização de grandezas físicas e da radiação, é apresentada de modo completo nos capítulos 1, 2 e 3 do livro de Greiner, e no capítulo I de Cohen *et al.* .

Enquanto o livro de Greiner possui uma estrutura similar ao do Griffiths, o de Cohen *et al.* é estruturado em capítulos, que contêm o núcleo básico da teoria, e em complementos, que contêm os tópicos e aplicações especiais escolhidos pelo autor.

O livro do inglês Paul Dirac (1902-1984) foi escrito por um dos principais personagens no desenvolvimento da Mecânica Quântica e criador da equação utilizada em todas as teorias quânticas relativísticas, a **equação de Dirac**. A partir de princípios gerais como o de Superposição Linear e de Incerteza, Dirac desenvolve a teoria da Mecânica Quântica como uma teoria linear em termos de vetores e operadores abstratos, independentemente de uma base para representá-los. O livro apresenta apenas duas aplicações elementares: o oscilador harmônico e as características não relativísticas do átomo de hidrogênio.

De certa forma, todo o material apresentado no livro de Dirac resultou de trabalhos realizados por ele mesmo; comutadores e parênteses de Poisson quânticos, teoria da representação e vetores *bra* e *ket*, função delta de Dirac, teoria da perturbação dependente do tempo e transições, operadores de permutação e sistemas de partículas, distribuição de Fermi-Dirac, interação da radiação com a matéria, espinores de Dirac, equação relativística para o elétron, estrutura fina do átomo de hidrogênio, teoria do pósitron, ... Uma das obras primas da Física.

## 4 O movimento da partícula na descrição de Schrödinger

Na Mecânica Quântica, o movimento de uma partícula não é descrito por uma trajetória<sup>15</sup> no espaço. Suas propriedades são descritas por um campo  $\Psi(x, t)$  que determina a probabilidade de sua presença, em qualquer instante  $t$  e posição  $x$  no espaço. Sob ação de um campo conservativo  $V(x)$ , uma partícula não-relativística de massa  $m$ , obedece à equação de Schrödinger

$$H\Psi(x, t) = i\hbar \frac{\partial}{\partial t} \Psi(x, t)$$

onde  $H = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x)$  é o operador linear denominado **hamiltoniano**<sup>16</sup> da partícula, o qual descreve a sua evolução e interação com o potencial.

O campo  $\Psi(x, t)$ , denominado **função de onda** ou **amplitude de probabilidade**, caracteriza o estado da partícula, e

$$\rho(x, t) = \Psi^*(x, t)\Psi(x, t)$$

é a **densidade de probabilidade**<sup>17</sup> de presença da partícula, isto é, a distribuição de probabilidades para as posições da partícula ou de qualquer grandeza que dependa da posição.

Desse modo, o **valor esperado**  $\langle x \rangle$  para a posição da partícula, cujo estado é caracterizado por  $\Psi(x, t)$ , é dado por<sup>18</sup>

$$\langle x \rangle = \int_{-\infty}^{\infty} x\Psi^*(x, t)\Psi(x, t) dx = (\Psi, x\Psi)$$

e, a chamada **incerteza**  $\Delta x$  na posição, por

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

Decorre também da equação de Schrödinger, que a densidade de probabilidade obedece a uma equação análoga à equação de continuidade da massa ou da carga,

<sup>15</sup>Uma curva  $x(t)$  parametrizada pelo tempo.

<sup>16</sup> Para campos conservativos é o operador associado à energia da partícula.

<sup>17</sup> A interpretação de  $\Psi^*\Psi$  como uma densidade de probabilidade impõe restrições à função de onda  $\Psi$ :

- A integração de  $\Psi^*\Psi$  sobre todo o espaço ( $-\infty < x < \infty$ ) deve ser realizável e igual à unidade,

$$\int_{-\infty}^{\infty} \Psi^*\Psi dx = 1$$

ou seja, a função de onda  $\Psi$  deve ser normalizável.

- Uma vez normalizada ela deve permanecer como tal, ou seja, a normalização deve ser independente do tempo.

$$\frac{d}{dt} \int_{-\infty}^{\infty} \Psi^*\Psi dx = 0$$

<sup>18</sup> Em geral, para qualquer grandeza  $f(x)$  que depende apenas da posição,

$$\langle f(x) \rangle = \int_{-\infty}^{\infty} f(x)\Psi^*(x, t)\Psi(x, t) dx$$

$$\begin{aligned}
\frac{\partial \rho}{\partial t} &= \underbrace{\frac{\partial \Psi^*}{\partial t}}_{-iH\Psi^*/\hbar} \Psi + \Psi^* \underbrace{\frac{\partial \Psi}{\partial t}}_{-iH\Psi/\hbar} \\
&= \frac{i}{\hbar} \Psi \left[ -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right] \Psi^* - \frac{i}{\hbar} \Psi^* \left[ -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right] \Psi \\
&= -\frac{i\hbar}{2m} \frac{\partial}{\partial x} \left( \Psi \frac{\partial \Psi^*}{\partial x} - \Psi^* \frac{\partial \Psi}{\partial x} \right) = -\frac{\partial J_x}{\partial x} \\
&\quad \Downarrow
\end{aligned}$$

$$\boxed{\frac{\partial \rho}{\partial t} + \frac{\partial J_x}{\partial x} = 0}$$

onde  $J_x = \frac{i\hbar}{2m} \left( \Psi \frac{\partial \Psi^*}{\partial x} - \Psi^* \frac{\partial \Psi}{\partial x} \right)$  é a chamada **densidade de corrente de probabilidade**<sup>19</sup> (seção 4.6) associada ao estado  $\Psi$ .

Assim, de acordo com a equação de continuidade, pode-se estabelecer a conservação da normalização,

$$\begin{aligned}
\frac{d}{dt} \int_{-\infty}^{\infty} \Psi^* \Psi \, dx &= \frac{d}{dt} \int_{-\infty}^{\infty} \rho \, dx \\
&= \int_{-\infty}^{\infty} \frac{\partial \rho}{\partial t} \, dx = - \int_{-\infty}^{\infty} \frac{\partial J_x}{\partial x} \, dx \\
&= - \left[ J_x(\infty, t) - J_x(-\infty, t) \right] = 0
\end{aligned}$$

uma vez que qualquer campo físico se anula no infinito.

#### 4.1 Estados estacionários

Escrevendo-se a equação de Schrödinger como

$$\left[ -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right] \Psi(x, t) = i\hbar \frac{\partial}{\partial t} \Psi(x, t)$$

e utilizando-se o método de separação das variáveis,

$$\Psi(x, t) = \psi(x)\phi(t)$$

obtém-se

$$\frac{1}{\psi(x)} \left[ -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right] \psi(x) = i\hbar \frac{1}{\phi(t)} \frac{d\phi}{dt} = E$$

<sup>19</sup>Para uma partícula no espaço, a densidade de corrente de probabilidade é generalizada e definida por

$$\vec{J} = \frac{i\hbar}{2m} (\Psi \nabla \Psi^* - \Psi^* \nabla \Psi)$$

onde a constante de separação  $E$  é a energia da partícula.

Assim, resultam duas equações diferenciais,

$$\begin{cases} i\hbar \frac{d\phi}{dt} = E\phi(t) \\ H\psi(x) = E\psi(x) \end{cases}$$

A primeira equação diferencial, no domínio do tempo, pode ser imediatamente integrada e resulta

$$\phi(t) = e^{-i(E/\hbar)t}$$

A segunda equação, chamada **equação de Schrödinger independente do tempo**, está associada a um *problema de autovalor* (Apêndice A.3), tal que, dependendo das condições impostas à função de onda  $\psi(x)$ , o autovalor  $E$ , isto é, a energia da partícula, pode assumir valores discretos  $\{E_n\}$  ou contínuos ( $E$ ).

Desse modo, a quantização da energia para os estados ligados de uma partícula é naturalmente obtida a partir de condições de contorno impostas à função de onda. Ou seja, na abordagem de Schrödinger, o problema de autovalor para estados ligados são problemas de contorno que envolvem uma equação diferencial.

Uma vez que a equação de Schrödinger independente do tempo é de ordem 2, a solução  $\Psi$  e sua derivada devem ser contínuas em todo o espaço<sup>20</sup>.

A cada valor possível de energia  $E$  pode corresponder um ou mais estados da partícula.<sup>21</sup> Por serem representados por autofunções  $\psi_E$  do operador hamiltoniano, são chamados de **autoestados de energia** e o conjunto dos autovalores de **espectro de energia**.

Os autoestados de energia encontrados para qualquer problema espacialmente unidimensional em um campo conservativo são também denominados **estados estacionários**, pois as densidades de probabilidades de presença associadas às soluções que evoluem desses estados independem do tempo

$$\rho = \Psi^*(x, t)\Psi(x, t) = \psi_E^*(x)\psi_E(x)$$

## 4.2 A partícula livre

Para uma partícula livre ( $V = 0$ ), com energia  $E > 0$ , as autofunções  $\psi_E(x)$  de seu hamiltoniano  $H$  são dadas por

$$\psi_E^+(x) \sim e^{i\frac{\sqrt{2mE}}{\hbar}x} \quad \text{e} \quad \psi_E^-(x) \sim e^{-i\frac{\sqrt{2mE}}{\hbar}x}$$

onde  $\psi_E^+$  corresponde à autofunção na qual a partícula se desloca no sentido  $+x$ , com momentum  $p = \sqrt{2mE}$  e  $\psi_E^-$  corresponde à que se desloca no sentido  $-x$ , com momentum  $p = -\sqrt{2mE}$ .

<sup>20</sup>Exceto para potenciais impulsivos do tipo delta de Dirac.

<sup>21</sup> Se a um dado valor de energia estão associados dois ou mais estados independentes esse autovalor é dito degenerado.

$$\begin{aligned}
H \left( e^{\pm i \frac{\sqrt{2mE}}{\hbar} x} \right) &= -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \left( e^{\pm i \frac{\sqrt{2mE}}{\hbar} x} \right) \\
&= E \left( e^{\pm i \frac{\sqrt{2mE}}{\hbar} x} \right)
\end{aligned}$$

Assim, a evolução da função de onda associada a uma partícula livre, a partir de um dos autoestados de energia  $E$ , será dada por

$$\Psi(x, t) = \psi_E^\pm(x) e^{-i(E/\hbar)t}$$

### 4.3 Operador momentum linear

Escrevendo-se o hamiltoniano numa forma análoga à relação clássica,

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

pode-se identificar o **operador momentum linear** na direção  $x$ , por

$$p_x = -i\hbar \frac{\partial}{\partial x}$$

Uma vez que

$$\begin{aligned}
p_x \left( e^{i \frac{\sqrt{2mE}}{\hbar} x} \right) &= -i\hbar \frac{\partial}{\partial x} \left( e^{i \frac{\sqrt{2mE}}{\hbar} x} \right) \\
&= \underbrace{\sqrt{2mE}}_p \left( e^{i \frac{\sqrt{2mE}}{\hbar} x} \right)
\end{aligned}$$

$\psi_E^+(x) = \psi_p(x) = e^{i \frac{\sqrt{2mE}}{\hbar} x}$  também é autofunção do operador  $p_x$  associada ao autovalor  $p = \sqrt{2mE}$ . Analogamente,  $\psi_E^-(x) = \psi_{-p}(x) = e^{-i \frac{\sqrt{2mE}}{\hbar} x}$  é autofunção do operador  $p_x$  associada ao autovalor  $p = -\sqrt{2mE}$ .

Desse modo, usualmente, representa-se o estado  $\Psi(x, t)$  de uma partícula livre de massa  $m$ , que se desloca no sentido  $+x$  com energia  $E$  e momentum  $p = \sqrt{2mE}$ , como

$$\Psi(x, t) = C e^{i(px - Et)/\hbar}$$

onde  $C$  é uma constante de normalização.

## 4.4 Generalizações

Diferentemente de uma partícula livre, os autoestados de energia de uma partícula em um poço de potencial constituem um conjunto discreto de autofunções, ou seja,

$$H\phi_n = E_n\phi_n \quad (n = 1, 2, 3, \dots)$$

Assim, se o estado inicial  $\Psi(x, 0) = \Psi_0(x)$  da partícula não for um autoestado de energia, a sua evolução pode ser calculada a partir da hipótese de que os autoestados de energia constituem um conjunto completo, no sentido que qualquer estado inicial  $\Psi_0(x)$  pode ser expresso por uma combinação linear dos autoestados  $\phi_n(x)$  de  $H$ ,

$$\Psi_0(x) = \sum_n c_n(0)\phi_n(x)$$

onde  $c_n(0)$  é a projeção de  $\Psi_0$  sobre o autoestado  $\phi_n$ ,

$$c_n(0) = (\phi_n, \Psi_0) = \int_{-\infty}^{\infty} \phi_n^*(x)\Psi_0(x)dx$$

de tal modo que o estado da partícula, em qualquer instante posterior  $t$ , será caracterizado por

$$\Psi(x, t) = \sum_n c_n(0)\phi_n(x)e^{-i(E_n/\hbar)t} = \sum_n c_n(t)\phi_n(x)$$

Calculando-se a energia média,

$$\begin{aligned} \langle H \rangle &= (\Psi, H\Psi) \\ &= \sum_{m,n} c_m^*(t)c_n(t) \underbrace{(\phi_m, H\phi_n)}_{E_n\delta_{mn}} \\ &= \sum_n |c_n(t)|^2 E_n = \sum_n |c_n(0)|^2 E_n \end{aligned}$$

a expressão indica que o quadrado dos módulos dos coeficientes da expansão do estado  $\Psi$  representam as probabilidades de ocorrência  $P(E_n)$  dos valores  $E_n$  ao se realizar uma medida da energia, ou seja,

$$P(E_n) = |c_n(0)|^2$$

De modo geral,<sup>22</sup>

*Toda grandeza é representada por um operador linear hermitiano cujos autovetores constituem um conjunto completo, e suas possíveis medidas são dadas por seus autovalores.*

Se uma partícula encontra-se num estado  $\Psi_t$  e  $\phi_n$  e  $a_n$  são os respectivos autovetores e autovalores associados a uma grandeza  $A$ , define-se a **probabilidade de ocorrência**  $P(a_n, t)$  de um autovalor  $a_n$ , ao se realizar uma medida da grandeza  $A$ , como

$$P(a_n, t) = |(\phi_n, \Psi_t)|^2$$

Em linguagem matemática,<sup>23</sup>

<sup>22</sup> O fato de grandezas serem representadas por operadores resulta em uma característica sempre presente nos cálculos e deduções da teoria, a não-comutatividade das operações.

<sup>23</sup> Apêndice A.4

Para um sistema num estado  $\Psi_t$ , a probabilidade de ocorrência de um autovalor  $a_n$  de um operador linear hermitiano  $A$ , associado a uma dada grandeza, é dada pela projeção do estado  $\Psi_t$  sobre o subespaço próprio gerado pelos autoestados  $\phi_n$  de  $A$ , associados ao autovalor  $a_n$ .

É claro que, enquanto as possíveis medidas para uma grandeza não dependem do estado atual de um sistema, suas probabilidades de ocorrência ou as distribuições dessas probabilidades dependem. Além da determinação do espectro de energia e dos autoestados de energia associados a um sistema em um campo conservativo, a importância fundamental da equação de Schrödinger decorre do fato de que ela descreve a evolução temporal do estado de um sistema. Desse modo, se for possível associar um estado inicial a um sistema, a probabilidade de ocorrência, em qualquer instante posterior, de quaisquer das medidas de uma grandeza, pode ser calculada a partir da solução da equação de Schrödinger, que representa o estado atual do sistema.

Decorre da definição de probabilidade de ocorrência que

**O valor esperado** de uma grandeza representada por um operador  $A$ , cujos autovalores  $a_n$  e respectivos autoestados  $\phi_n$  satisfazem

$$A\phi_n = a_n\phi_n \quad e \quad (\phi_n, \phi_m) = \delta_{nm}$$

para um sistema num estado  $\Psi_t = \sum_n c_n \phi_n$ , onde  $c_n(t) = (\phi_n, \Psi_t)$ ,

é dado por

$$\langle A \rangle(t) = \sum_n a_n |c_n(t)|^2$$

que pode ser expresso também por <sup>a</sup>

$$\langle A \rangle(t) = (\Psi_t, A\Psi_t)$$

---


$$^a \langle A \rangle = \sum_n a_n \underbrace{(\phi_n, \Psi_t)^*}_{(\Psi_t, \phi_n)} \underbrace{(\phi_n, \Psi_t)}_{c_n} = \sum_n c_n (\Psi_t, \underbrace{a_n \phi_n}_{A\phi_n}) = (\Psi_t, A \underbrace{\sum_n c_n \phi_n}_{\Psi_t})$$

Em Mecânica Quântica uma grandeza  $A$  é dita conservada se a probabilidade de ocorrência de qualquer valor  $a_n$  para sua medida não depende do tempo,

$$P(a_n, t) = P(a_n, 0)$$

ou seja, quando<sup>24</sup>

$$\frac{d\langle A \rangle}{dt} = 0 \quad \iff \quad \langle A \rangle = \text{cte.}$$

De acordo com a equação de Schrödinger,

$$i\hbar \frac{\partial \Psi}{\partial t} = H\Psi$$

---

<sup>24</sup>Esse é sempre o caso quando o sistema estiver num autoestado (estacionário) de energia.

a evolução do valor esperado de uma grandeza  $A$ , para um sistema em um estado  $\Psi$ , é dado por

$$\begin{aligned} \frac{d\langle A \rangle}{dt} &= \left( \underbrace{\frac{\partial \Psi}{\partial t}}_{-iH\Psi/\hbar}, A\Psi \right) + \left( \Psi, A \underbrace{\frac{\partial \Psi}{\partial t}}_{-iH\Psi/\hbar} \right) \\ &= \frac{i}{\hbar} \underbrace{(H\Psi, A\Psi)}_{(\Psi, HA\Psi)} - \frac{i}{\hbar} (\Psi, AH\Psi) \\ &= \frac{i}{\hbar} \langle [H, A]_- \rangle \end{aligned}$$

Desse modo, independentemente do estado em que se encontra um sistema, qualquer grandeza  $A$  que comute com seu hamiltoniano  $H$  é conservada.

$$\boxed{[H, A]_- = 0 \iff \langle A \rangle = \text{cte.}}$$

De maneira geral, os operadores associados a grandezas cinemáticas<sup>25</sup> não exibem dependência temporal. Apenas operadores que representam grandezas que surgem ou se manifestam na presença de interações<sup>26</sup> podem exibir dependência temporal explícita.

Apesar de não ter sido, historicamente, a primeira formulação da Mecânica Quântica, a construção ou elaboração da teoria quântica a partir desse formalismo, denominada “*descrição de Schrödinger*”, passou a ser o método tradicional, quase canônico, de apresentação da Mecânica Quântica.

Por outro lado, numa forma mais análoga à formulação da Mecânica Clássica, chamada “*descrição de Heisenberg*”, que precede ligeiramente à de Schrödinger, os operadores que representam grandezas cinemáticas ou dinâmicas, em geral, passam a exibir uma dependência temporal explícita, análoga aos casos clássicos similares, quando esses existirem.

A conexão e distinção entre as duas abordagens é que, enquanto a equação de Schrödinger descreve a evolução temporal do estado de um sistema, na descrição de Heisenberg o estado do sistema não varia com o tempo, apenas as grandezas associadas o fazem.

Em outra abordagem, denominada “*descrição de Dirac*”<sup>27</sup>, tanto os operadores quanto os estados passam a depender explicitamente do tempo, mas de tal maneira que as ações internas do sistema não aparecem mais explicitamente, e a equação de evolução é determinada explicitamente apenas por ações externas.

A teoria da Mecânica Quântica, em sua versão criada, formulada e interpretada, principalmente, por Schrödinger, Heisenberg, Dirac, Born, Jordan, Kramers, Pauli e Bohr, culmina com um esquema duplamente linear. Ao contrário da Mecânica Clássica e do Eletromagnetismo, onde as grandezas cinemáticas e dinâmicas (posições velocidades, acelerações, forças e campos) podem ser vistas como realizações de uma estrutura algébrica linear (conjuntos de vetores) cujas equações não são necessariamente lineares, na Mecânica Quântica, tanto os estados, as grandezas e as equações de evolução dos estados (equações de Schrödinger, Heisenberg, Dirac, Klein-Gordon) são representadas por elementos lineares (vetores e operadores). Essa simplicidade, entretanto, é apenas aparente, pois suas aplicações mais simples, em geral, envolvem espaços de dimensão infinita ou problemas de contorno não triviais.

<sup>25</sup> Grandezas que são definidas mesmo na ausência de interações internas ou de ações externas, como momentum, posição, spin,...

<sup>26</sup> Como energia potencial de interação, polarização, magnetização,...

<sup>27</sup> Ou *Interaction Picture*.

## 4.5 As equações de Ehrenfest

A equação de evolução do valor esperado de uma grandeza  $A$ , que pode ser escrita numa forma similar à equação de Schrödinger,

$$i\hbar \frac{d\langle A \rangle}{dt} = \langle [A, H]_- \rangle$$

permite derivar as chamadas **equações de Ehrenfest** e, a partir delas, estabelecer a validade dos procedimentos que utilizam soluções clássicas como soluções de problemas quânticos.

Essas equações expressam que as variações temporais dos valores médios da posição  $x$  e do momentum linear  $p_x$  de uma partícula em campo conservativo  $V(x)$  obedecem a expressões análogas às equações de movimento clássicas.

$$\left\{ \begin{array}{l} \frac{d\langle x \rangle}{dt} = \frac{\langle p_x \rangle}{m} \\ \frac{d\langle p_x \rangle}{dt} = \left\langle -\frac{\partial V}{\partial x} \right\rangle \end{array} \right. \implies m \frac{d^2\langle x \rangle}{dt^2} = \left\langle -\frac{\partial V}{\partial x} \right\rangle$$

Uma vez que para  $H = p_x^2/2m + V(x) \Rightarrow [x, H(x, p_x)]_- = [x, p_x^2/2m]_- = \frac{1}{2m}[x, p_x^2]_- = i\hbar \frac{p_x}{m}$

obtem-se  $\frac{d\langle x \rangle}{dt} = \frac{\langle p_x \rangle}{m}$  e  $[p_x, H(x, p_x)]_- = [p_x, V(x)]_- = -i\hbar \frac{\partial V}{\partial x}$

Assim,

$$\frac{d\langle p_x \rangle}{dt} = \left\langle -\frac{\partial V}{\partial x} \right\rangle = \langle -F(x) \rangle$$

Esse resultado é tido algumas vezes como uma expressão da 2ª lei de Newton, ou seja, que os valores médios das grandezas cinemáticas obedecem às equações de movimento da Mecânica Clássica, o que não é correto. Pois, expandindo-se o termo de força  $F(x)$  em série de Taylor, em torno do valor médio  $\langle x \rangle$  da posição,

$$F(x) = F(\langle x \rangle) + (x - \langle x \rangle) \left( \frac{\partial F}{\partial x} \right)_{\langle x \rangle} + \frac{1}{2}(x - \langle x \rangle)^2 \left( \frac{\partial^2 F}{\partial x^2} \right)_{\langle x \rangle} + \dots$$

e calculando-se o valor médio, resulta

$$m \frac{d^2\langle x \rangle}{dt^2} = F(\langle x \rangle) + \frac{(\Delta x)^2}{2} \left( \frac{\partial^2 F}{\partial x^2} \right)_{\langle x \rangle} + \dots$$

Desse modo, para todo sistema que utiliza o oscilador harmônico ou a partícula livre<sup>28</sup> como modelo, as soluções clássicas expressam qualitativamente o comportamento do sistema, e alguns resultados quantitativos podem ser obtidos por analogia.<sup>29</sup>

Por esse mesmo motivo Bohr obteve resultados corretos para o átomo de hidrogênio. No modelo de Bohr, a quantização do momentum angular equivale a considerar os aspectos periódicos do movimento e, portanto, utiliza o modelo do oscilador.

<sup>28</sup> Para os quais  $-\frac{\partial^3 V}{\partial x^3} = \frac{\partial^2 F}{\partial x^2} = 0$ .

<sup>29</sup> Foi desse modo que, basicamente, ocorreu o processo de construção da teoria quântica (mecânica das matrizes) por Heisenberg, Born, Jordan e Pauli.

Uma outra situação em que alguns conceitos clássicos, como os de trajetória e velocidade, podem ser utilizados é quando potencial varia lentamente num intervalo comparável à incerteza na posição da partícula. Esse é o caso das trajetórias de feixes de partículas em aceleradores ou câmaras de ionização, sujeitos a campos que variam sobre distâncias macroscópicas.

## 4.6 Problema de estado inicial

A abordagem tradicional da Mecânica Quântica associa a cada instante  $t$  e região do espaço  $\vec{r}$  ocupado por uma partícula, sujeita ou não a ações externas, um campo  $\Psi(\vec{r}, t)$  cuja evolução temporal obedece à equação de Schrödinger,

$$i\hbar \frac{\partial \Psi}{\partial t} = H\Psi$$

de tal modo que o valor médio de qualquer grandeza  $A$  associada à partícula é dado por

$$\langle A \rangle(t) = (\Psi, A\Psi)$$

Como a equação de Schrödinger é uma equação diferencial linear homogênea<sup>30</sup>, qualquer autoestado ou combinação linear de autoestados será solução. E, por ser de ordem 1 com relação ao tempo, para se determinar uma única solução para o problema, torna-se necessário a especificação de uma condição inicial (estado inicial) para a função de onda. Ou seja, a equação de Schrödinger determina a evolução de um sistema, a partir de um estado inicial  $\Psi(x, 0)$ .

Por exemplo, se uma partícula em  $t = 0$  possui energia  $E$  e momentum  $p = \sqrt{2mE}$  definidos, a sua amplitude de probabilidade ou estado inicial  $\Psi(x, 0)$  é proporcional ao autoestado correspondente,

$$\Psi(x, 0) = \Psi_0(x) \sim e^{i\frac{p}{\hbar}x}$$

e evoluirá segundo

$$\Psi(x, t) \sim e^{i(p x - Et) / \hbar}$$

Entretanto, em alguns problemas como, por exemplo, o do comportamento de uma partícula em um poço de potencial, nos interessa, não a evolução de um particular estado da partícula e sim, o seu espectro de energia ou as possíveis transições entre os respectivos autoestados. Nesses casos, mesmo que não saibamos associar ou desconhecamos um estado inicial para a partícula é possível a previsão das principais características do sistema.

A não ser que observemos uma única partícula ou um feixe homogêneo e mono-energético de partículas idênticas, em geral, não podemos associar um estado inicial a um sistema, mesmo do ponto de vista clássico. Para sistemas com muitos graus de liberdade devemos utilizar os métodos estatísticos (ensembles) desenvolvidos por Boltzmann e Gibbs [8].

Do ponto de vista experimental, podemos associar a um feixe de partículas idênticas, cada qual com massa  $m$  e carga  $e$ , um estado inicial, se as acelerarmos, durante um certo trecho, por um potencial elétrico fixo  $V$ . Ao final desse período, obtem-se um feixe onde a energia de cada partícula é igual à  $E = |eV|$  e cujo estado inicial ( $\Psi_0$ ) do será proporcional ao autoestado de uma partícula livre de energia  $E$ , momentum  $p = \sqrt{2mE}$  e massa  $m$ ,

$$\Psi(x, 0) = \Psi_0(x) = C e^{i\sqrt{2mE}x/\hbar}$$

---

<sup>30</sup> Isto é, vale o Princípio da Superposição.

Uma vez que a densidade de corrente associada a esse estado é dada por

$$J_x = \frac{i\hbar}{2m} \left( \Psi \underbrace{\frac{\partial \Psi^*}{\partial x}}_{-ip\Psi^*/\hbar} - \Psi^* \underbrace{\frac{\partial \Psi}{\partial x}}_{ip\Psi/\hbar} \right) = \frac{p}{m} \underbrace{(\Psi_0^* \Psi_0)}_{\rho}$$

pode-se interpretar  $\rho$  como a densidade de partículas do feixe, e normalizar o estado inicial de um feixe, de tal modo que a sua integral numa dada região, de dimensão (comprimento)  $L$ , seja igual ao número ( $N$ ) de partículas nessa região.

$$\underbrace{\int_L \Psi_0^* \Psi_0 dx}_{|C|^2 L} = N \implies C = \sqrt{\frac{N}{L}}$$

Essa é a normalização utilizada na análise de espalhamento de feixes de partículas não-relativísticas por sistemas atômicos.

#### 4.7 Transições e regras de seleção

Uma propriedade dos sistemas conservativos, confinados numa dada região do espaço e caracterizados por um hamiltoniano  $H$ , é que seus espectros constituem conjuntos discretos, ou seja,

$$H\phi_n(x) = E_n\phi_n(x) \quad (n = 0, 1, \dots)$$

onde os  $\phi_n$  são os autoestados de  $H$ , correspondentes às energias  $E_n$ .

Se o estado inicial  $\Psi(x, 0) = \Psi_0(x)$  do sistema for um desses autoestados,  $\Psi_0(x) = \phi_n(x)$ , os estados subsequentes,  $\Psi(x, t)$ , serão fisicamente indistinguíveis, ou seja, suas propriedades, como os valores médios de quaisquer grandezas, não variarão com o tempo,

$$\Psi(x, t) = e^{-i\frac{E_n}{\hbar}t} \phi_n(x)$$

Mesmo que o estado inicial  $\Psi_0(x)$  não seja um dos autoestados de  $H$ , ou mesmo algum outro estado conhecido, em princípio, pode-se expandi-lo também em relação aos autoestados de  $H$ ,

$$\Psi_0(x) = \sum_n \underbrace{(\phi_n, \Psi_0)}_{c_n(0)}$$

de tal modo que sua evolução será dada por

$$\Psi(x, t) = \sum_n c_n(0) e^{-i\frac{E_n}{\hbar}t} \phi_n(x)$$

e o valor médio de qualquer grandeza  $A$  será dado por

$$\langle A \rangle_t = \sum_{m,n} c_m^*(0) c_n(0) \underbrace{(\phi_m, A\phi_n)}_{A_{mn}} e^{i\frac{(E_m - E_n)}{\hbar}t}$$

Assim, se  $A$  não depende explicitamente do tempo, os elementos de matriz  $A_{mn}$  serão constantes e a evolução temporal do valor médio será descrita por uma série de termos oscilantes, cujas frequências, denominadas **frequências de Bohr**,

$$\omega_{mn} = \left| \frac{(E_m - E_n)}{\hbar} \right|$$

são características do sistema, independentes da grandeza  $A$  e do estado inicial.

Desse modo, é razoável supor que as frequências de Bohr, as quais os valores médios de grandezas atômicas como o momento dipolar oscilam, são as únicas frequências permitidas para a luz<sup>31</sup> absorvida ou emitida por um sistema.

Como o peso de cada termo oscilante é dado pelos elementos de matriz  $A_{mn}$ , os elementos nulos correspondem às frequências ausentes na absorção ou emissão da luz por um átomo. Essa é a origem das chamadas **regras de seleção**.

## 4.8 Modelos e problemas típicos

Do mesmo modo que os processos e fenômenos macroscópicos abordados pela Mecânica Clássica são representados por idealizações, que permitem que encontremos soluções analíticas ou aproximadas, o mesmo ocorre no domínio microscópico. Ou seja, a partícula livre, o oscilador harmônico e o sistema de dois corpos constituem os modelos básicos para a abordagem de qualquer problema em nível quântico.

Uma característica dos fenômenos microscópicos é que as interações entre elétrons, prótons, átomos e moléculas, no domínio não-relativístico, são de origem eletromagnética.

Entretanto, apesar da estrutura linear da teoria, mesmo esses modelos básicos envolvem dificuldades matemáticas de tal ordem que, inicialmente, utilizamos modelos ainda mais simplificados e afastados dos fenômenos reais, como partículas em poços e barreiras de potencial retangulares ou infinitos. Assim, o problema do oscilador unidimensional, de massa  $m$  e frequência própria  $\omega$ , que envolve um potencial do tipo  $V = m\omega^2 x^2/2$ , e o problema de dois corpos que envolve um potencial central  $V = -k/r$ , são deixados para serem analisados numa fase mais madura do estudante.

Uma outra dificuldade deve-se à existência de novos graus de liberdade, mesmo para uma partícula livre, como o *spin*, que só é devidamente descrito ao introduzirmos a representação espinorial.<sup>32</sup>

Uma outra classe de problemas, que se tornaram os protótipos para os métodos de investigações das propriedades dinâmicas dos sistemas microscópicos, é a de espalhamentos de feixes de partículas por átomos, núcleos, hádrons, mésons e outros feixes de partículas. Nesses casos, tanto o feixe incidente quanto o espalhado são caracterizados por superposições de estados da partícula livre.

A partir da análise de experimentos desses tipos de espalhamentos, iniciados com Rutherford, é que a atual visão de uma natureza discreta da matéria foi consolidada. Além da determinação da estrutura atômica da matéria (sólidos cristalinos e amorfos, líquidos e gases), se estabeleceu

<sup>31</sup> Radiação eletromagnética

<sup>32</sup> Ou seja, o conceito de espinor.

também a estrutura quarkônica das partículas não-elementares hadrônicas como os prótons, os pions e os nêutrons<sup>33</sup>.

## 5 Leis de conservação

A conexão entre grandezas que obedecem a leis de conservação e grupos de transformações associados às simetrias de um sistema, de uma grandeza ou equação física é de caráter geral [21] e permite a ampliação de vários conceitos e grandezas originários da Física Clássica, como a energia e o momentum.

Geometricamente, se a forma de um objeto não se modifica sob a ação de uma certa operação, esta é dita uma operação de simetria. Diz-se também que esse objeto é simétrico com relação à operação. Do mesmo modo, uma lei e uma equação física podem também não se alterarem ao serem descritas segundo referenciais distintos. Nesse contexto, diz-se que a lei e a equação são invariantes com relação a mudanças de referenciais.

### 5.1 Momentum linear e translações

O operador momentum linear pode ser associado a uma ação espacial (translação) que permite uma ampla generalização do próprio conceito de momentum.

Se  $\psi$  é a função de onda associada a uma partícula livre, de massa  $m$ , que se desloca na direção  $x$ , com energia e momentum relacionados por  $E = p^2/(2m)$ , suas imagens em dois pontos,  $x$  e  $x + d$ , cuja distância é  $d$ ,

$$\psi(x) = e^{ipx/\hbar} \quad \text{e} \quad \psi(x + d) = e^{ip(x + d)/\hbar}$$

estão relacionadas por

$$\psi(x + d) = e^{ipd/\hbar}\psi(x)$$

De maneira geral, as imagens de qualquer campo, em quaisquer dois pontos  $x$  e  $x + d$ , podem ser relacionadas por um operador  $U_d$ , que representa para essas funções o efeito da translação de  $x$  a  $x + d$ ,

$$\psi(x + d) = U_d\psi(x)$$

Expandindo-se  $\psi(x + d)$  em uma série de Taylor,

$$\begin{aligned} \psi(x + d) &= \psi(x) + \underbrace{\psi'(x)}_{\partial\psi/\partial x} d + \underbrace{\psi''(x)}_{\partial^2\psi/\partial x^2} \frac{d^2}{2} + \dots \\ &= \underbrace{\left(1 + d\frac{\partial}{\partial x} + \frac{d^2}{2}\frac{\partial^2}{\partial x^2} + \dots\right)}_{e^{d\partial/\partial x}} \psi(x) = e^{i\frac{d}{\hbar}\left(-i\hbar\frac{\partial}{\partial x}\right)} \psi(x) \end{aligned}$$

<sup>33</sup> Isso aconteceu em meados do século XX, nos anos 60, quando os experimentos de Friedman *et al.* [?], interpretados por Feynman e Bjorken, estabeleceram que o próton teria uma estrutura interna. Sendo composto pelos *quarks* já previstos por Gell-Man e Zweig, de maneira análoga à tabela periódica de Mendeleiev, ou seja, por argumentos de simetrias.

pode-se identificar e escrever o operador de translação  $U_d$ , para o campo  $\psi$ , como um operador unitário<sup>34</sup>,

$$U_d = e^{i\frac{d}{\hbar}p_x}$$

para qualquer que seja o valor do parâmetro  $d$ . Nesse sentido, o operador momentum linear  $p_x = -i\hbar\partial/\partial x$  na direção  $x$  é identificado como um **gerador de translações** nessa direção.

Do ponto de vista matemático, diz-se que o operador  $U_d$  é uma **representação unitária induzida** do grupo das translações no eixo  $x$ , cujo gerador é o operador hermitiano<sup>35</sup>

$$p_x = -i\hbar\frac{\partial}{\partial x}$$

Do ponto de vista físico, o caráter unitário do operador  $U_d$  pode ser derivado do **Princípio da Homogeneidade Espacial**, no sentido que o estado de uma partícula livre não depende da origem do referencial, ou seja,  $\psi(x+d)$  e  $\psi(x)$  representam estados que possuem a mesma

<sup>34</sup> Apêndice A.2

<sup>35</sup> Para uma translação infinitesimal  $\epsilon$ , o operador de translação na direção  $x$  pode ser escrito

$$U_\epsilon = 1 + i\frac{\epsilon}{\hbar}p_x$$

e, uma translação finita  $d = N\epsilon$  pode ser obtida a partir de uma sucessão de translações infinitesimais,

$$U_d = \lim_{N \rightarrow \infty} \left(1 + i\frac{d}{N}\frac{p_x}{\hbar}\right)^N = e^{i\frac{d}{\hbar}p_x}$$

Nesse sentido diz-se que o operador de translação é uma transformação contínua ou que o grupo das translações é contínuo.

Para uma translação  $\vec{d}$  numa direção qualquer do espaço, o operador de translação unitário induzido é expresso por

$$U(\vec{d}_j) = e^{i\frac{\sum_{j=1}^3 d_j p_{x_j}}{\hbar}} = e^{i\frac{\vec{d} \cdot \vec{p}}{\hbar}}$$

onde  $p_{x_j} = -i\hbar\partial/\partial x_j$  é o gerador de translações na direção  $x_j$  e os  $d_j$  são os chamados parâmetros da translação  $T_{\vec{d}}$ .

Para uma translação infinitesimal,

$$1 + i\frac{\sum_{j=1}^3 \epsilon_j p_{x_j}}{\hbar}$$

distribuição de probabilidades de presença,<sup>36</sup>

$$\psi^*(x+d)\psi(x+d) = \psi^*(x)\psi(x)$$

Desse modo, as translações representam operações de simetria para uma partícula livre, no sentido que sua densidade de probabilidade de presença permanece invariante em seu movimento uniforme. De maneira geral, a probabilidade de ocorrência de qualquer evento associado a uma partícula livre permanece invariante sob ação de uma translação.

Por sua vez, o Princípio da Homogeneidade Espacial ou **Princípio de Invariância à Translação** reflete o fato de que o valor médio do momentum de uma partícula livre permanece inalterado em seu movimento de translação uniforme, ou seja, deriva da **Lei de Conservação do Momentum** para uma partícula livre.

## 5.2 Grupos de simetria e leis de conservação

Do ponto de vista formal, o conceito de simetria pode ser estabelecido a partir de um estudo mais detalhado das transformações lineares definidas num dado espaço.

As ações de **translações ativas**  $T_a$  de uma partícula em um ponto  $x$  a um ponto  $x+d$ ,

$$T_a x = x + d$$

estão ligadas também às ações de **translações passivas**  $T_p$  (ou mudança de sistemas de coordenadas), que relacionam as representações de um mesmo ponto  $P$  segundo dois sistemas de coordenadas  $S$  e  $S'$ , cujas origens estão separadas por uma distância  $d$ . Se  $x$  é a coordenada de  $P$  em  $S$  e  $x'$  em  $S'$ ,

$$x' = T_p x = x - d$$

então,

$$T_a = T_p^{-1}$$

---

<sup>36</sup> De modo geral, esse caráter unitário das transformações induzidas por um grupo de transformações cinemáticas decorre do fato de que essas transformações não modificam as amplitudes de probabilidades associadas a um sistema.

Por exemplo, se  $c_n = (\phi_n, \Psi)$  é a probabilidade de ocorrência do autovalor  $a_n$  de  $A$ , associado ao autoestado  $\phi_n$ , para um sistema num estado  $\Psi$ , e  $U$  é uma representação induzida por um grupo de transformações sobre as coordenadas do sistema ou sobre seus momenta,

$$U\phi_n = \phi'_n \quad \text{e} \quad U\Psi = \Psi'$$

para que a amplitude de probabilidade seja invariante,

$$\begin{aligned} (\phi'_n, \Psi') &= (\phi_n, \Psi) \\ &\implies U^\dagger U = 1 \\ (U\phi_n, U\Psi) &= (U^\dagger U\phi_n, \Psi) \end{aligned}$$

ou seja, a representação induzida deve ser unitária.

O fato dos geradores de um grupo contínuo, associado a uma transformação unitária, serem hermitianos decorre de

$$\begin{aligned} U(\epsilon) = 1 + i\epsilon G &\implies U^\dagger = 1 - i\epsilon G^\dagger \\ U^\dagger U = 1 + i\epsilon(G - G^\dagger) + \dots &\implies G^\dagger = G \end{aligned}$$

As representações de um **campo**<sup>37</sup>,  $\psi$  segundo  $S$  e  $\psi'$  segundo  $S'$ , devem ter dependências funcionais tais que suas imagens obedeçam a

$$\psi'(x') = \psi(x) = \psi(T_p^{-1}x')$$

Intercambiando-se as variáveis  $x$  e  $x'$ , resulta

$$\boxed{\psi'(x) = \psi(T_p^{-1}x) = \psi(T_a x)}$$

ou seja, numa relação entre as representações de um campo, segundo dois sistemas de coordenadas de tal modo que a imagem, num ponto  $x$ , da representação em  $S'$  é a mesma que a imagem de sua representação em  $S$ , num outro ponto trasladado de  $d$ .

Se  $U_d$  é uma representação induzida qualquer do grupo das translações de  $d$ , na direção  $x$ ,

$$\psi(T_a x) = \psi(x + d) = U_d \psi(x) = \psi'(x)$$

Transformações de outras entidades tensoriais, tais como o hamiltoniano ou qualquer operador linear  $H$ ,

$$H(x)\psi(x) = \phi(x)$$

podem ser diretamente deduzidas, pois se  $T$  é uma representação de um grupo de transformações ativas de  $x$ , e  $U$  é uma representação induzida,

$$UH(x)U^{-1} \underbrace{U\psi(x)}_{\psi'(x)} = U\phi(x) = \phi(Tx) = \underbrace{H(Tx)}_{H'(x)} \underbrace{\psi(Tx)}_{\psi'(x)}$$

as representações  $H'(x)$  e  $H(x)$  de um operador linear  $H$  estarão relacionadas por

$$\boxed{H'(x) = UH(x)U^{-1}}$$

Desse modo, a condição para que um operador linear  $H$  seja invariante sob ação de um grupo,

$$H'(x) = H(x) \iff HU = UH$$

é expressa pela anulação do chamado comutador de  $H$  com alguma representação  $U$  do grupo,

$$\boxed{HU - UH = [H, U]_- = 0}$$

ou pela anulação do comutador de  $H$  com algum gerador  $G$  do grupo,

$$\boxed{[H, G]_- = 0}$$

Uma vez que o momentum linear associado a uma partícula livre é o gerador das translações na direção do movimento da partícula, o hamiltoniano de uma partícula livre é invariante a qualquer operação de translação. Diz-se que para uma partícula livre, a operação de translação é uma **operação de simetria**.

Nesse sentido, os grupos que comutam com o hamiltoniano de um sistema são chamados de grupos de operações de simetria do sistema.

<sup>37</sup> Um invariante a transformações passivas.

*As operações de simetria de um sistema são as transformações sob as quais o hamiltoniano de um sistema é invariante.*

Desse modo, se o operador momentum linear, que é o gerador das translações, comuta com o hamiltoniano de um sistema, isto é, o sistema exibe simetria a translações, o valor esperado do momentum linear não depende do tempo. Diz-se, então, que o momentum linear é conservado.

Assim como o momentum linear é o gerador de translações, o momentum angular é o gerador de rotações e o hamiltoniano é o próprio gerador de translações temporais. Portanto, para sistemas que são simétricos com relação a essas operações<sup>38</sup> cada um desses geradores obedecem a uma lei de conservação.

*As simetrias de um sistema ou os grupos de operações sob as quais o seu hamiltoniano é invariante implicam leis de conservação para os geradores desses grupos.*

Do mesmo modo que as leis e equações fundamentais da Física,<sup>39</sup> as leis de conservação e os princípios de invariância são postulados *a posteriori* e não verdades *a priori*. Entretanto, a utilização desses princípios foi tão frutífera nas mãos de Einstein e Dirac que, a partir de meados do século XX, passa a vigorar na Física Teórica a tendência de se construir teorias e tentar derivar as outras leis e equações dinâmicas da Física a partir desses princípios de invariância, adotando-os como princípios fundamentais. Tal ênfase é oposta à que levou ao estabelecimento das teorias clássicas da Mecânica e do Eletromagnetismo, e mesmo à Mecânica Quântica Não-Relativística, construídas e elaboradas a partir de generalizações de algumas leis particulares e não de princípios tão gerais.

Sobre a utilização indiscriminada desse procedimento, vale lembrar as palavras do matemático Henri Lebesgue,[10]

*“Uma generalização realizada, não pelo vão prazer de uma generalização mas, para resolver problemas previamente existentes é sempre uma generalização frutífera.”*

### 5.3 Geradores e regras de comutação

Em vez de definir alguns outros operadores a partir de generalizações de relações clássicas, as representações de grandezas dinâmicas associadas a leis de conservação podem ser construídas a partir das representações unitárias dos chamados grupos de simetria, ou seja, determinando-se seus geradores.

As simetrias associadas a leis de conservação clássicas, do ponto de vista matemático, estão associadas aos grupos contínuos, ou seja, operações que podem ser obtidas a partir de transformações infinitesimais.

Se  $T(\epsilon_i)$  é uma representação de um grupo, onde os  $\epsilon_i$  são parâmetros do grupo, as transformações unitárias infinitesimais induzidas  $S(T)$  podem ser expressas por

<sup>38</sup> Por exemplo, sistemas livres de ações externas.

<sup>39</sup> Como as *Leis de Newton* e as *Equações de Maxwell, Schrödinger e Dirac*.

$$S(T) = 1 + i \sum_j \epsilon_j G_j$$

onde os  $G_j$  são os geradores (hermitianos) do grupo.

Desse modo, resultam dois procedimentos:

- um para determinar os geradores do grupo,
- outro para determinar seus comutadores com algum operador  $L(x)$ .

$$\begin{cases} S(T)\psi(x) = \psi(Tx) \\ SL(x)S^{-1} = L(Tx) \end{cases}$$

$$\Downarrow$$

$$\begin{cases} i \sum_j \epsilon_j G_j \psi(x) = \psi(Tx) - \psi(x) \\ -i \sum_j \epsilon_j [L(x), G_j] = L(Tx) - L(x) \end{cases}$$

Por exemplo, no caso de translações tri-dimensionais  $T_{\vec{a}}$ ,

$$x'_k = x_k + a_k = (T_{\vec{a}}x)_k$$

escrevendo-se o operador de translação infinitesimal como

$$S = 1 + i \sum_k a_k p_k / \hbar$$

onde  $p_k$  é a componente  $k$  do gerador de translação, resulta

$$i a_k \frac{p_k}{\hbar} \psi(\vec{r}) = \psi(\vec{x} + \vec{a}) - \psi(\vec{x}) = a_l \frac{\partial \psi}{\partial x_l} = a_k \frac{\partial \psi}{\partial x_k}$$

ou seja,

$$p_k = -i\hbar \frac{\partial}{\partial x_k}$$

Das regras de comutação,

$$-i \frac{a_k}{\hbar} [x_l, p_k]_- = (x_l + a_l) - x_l = a_l = \delta_{lk} a_k$$

resulta

$$[x_l, p_k]_- = i\hbar \delta_{lk}$$

## 5.4 Momentum angular e rotações

No caso de rotações em torno de um eixo  $z$ ,

$$T \begin{pmatrix} x \\ y \end{pmatrix} \rightarrow \begin{pmatrix} x - \theta y \\ \theta x + y \end{pmatrix}$$

Escrevendo-se o operador de rotação infinitesimal como

$$S = 1 + i\theta L_z / \hbar$$

resulta que

$$\begin{aligned} i\theta \frac{L_z}{\hbar} \psi(x, y, z) &= \psi(x - \theta y, y + \theta x, z) - \psi(x, y, z) \\ &= \theta \left( x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right) \psi \end{aligned}$$

ou seja, o gerador de rotações em torno do eixo  $z$  é a componente do **operador momento angular**  $\vec{L}$  na direção  $z$ ,

$$\begin{aligned} L_z &= x \underbrace{\left( -i\hbar \frac{\partial}{\partial y} \right)}_{p_y} - y \underbrace{\left( -i\hbar \frac{\partial}{\partial x} \right)}_{p_x} \\ &= xp_y - yp_x \end{aligned}$$

De maneira geral, o gerador de rotações  $L_j$  em torno de um eixo genérico  $x_j$  é expresso por

$$L_j = \sum_{k,l} \epsilon_{jkl} x_k p_l$$

Uma vez que  $\vec{L}$  é um operador vetorial,

$$T \begin{pmatrix} L_x \\ L_y \end{pmatrix} \rightarrow \begin{pmatrix} L_x - \theta L_y \\ \theta L_x + L_y \end{pmatrix}$$

implica

$$\begin{cases} -i\frac{\theta}{\hbar} [L_x, L_z]_- = (L_x - \theta L_y) - L_x \\ -i\frac{\theta}{\hbar} [L_y, L_z]_- = (L_y + \theta L_x) - L_y \end{cases}$$

$$\Downarrow$$

$$\begin{cases} [L_x, L_z]_- = -i\hbar L_y \\ [L_y, L_z]_- = i\hbar L_x \end{cases}$$

que pode ser resumido nas seguintes relações de comutação,

$$[L_j, L_k]_- = i\hbar \epsilon_{ljk} L_l$$

## A Operadores lineares

De maneira geral, um **operador** é qualquer relação definida entre os membros de um conjunto.

Por exemplo, a derivada de uma função  $f$  em uma variável  $x$ , real ou complexa,

$$f'(x) = \frac{d}{dx}f(x) = D_x f(x) = g(x)$$

resulta em outra função  $g$  de  $x$ . Nesse sentido, a derivação é um operador representado por

$$\frac{d}{dx} \quad \text{ou} \quad D_x$$

A multiplicação de uma função  $f(x)$  por um número real ou complexo ( $c$ ),  $cf(x)$ , é um outro operador simplesmente representado por  $c$ .

Do mesmo modo, a integração (indefinida) de uma função  $f(x)$ ,

$$\int f(x) dx = h(x)$$

é um operador que pode ser representado por

$$\mathcal{L}_x f = h$$

### A.1 Produto escalar e ortogonalidade

A classificação dos operadores definidos num conjunto de funções<sup>40</sup> está associada, principalmente, à definição do **produto escalar**.

O produto escalar definido num conjunto  $L^2 = \{f, g, h, \dots\}$  de funções em uma variável  $x$ , que se anulam no infinito, é uma operação binária entre duas funções  $f$  e  $g$  que resulta em um número complexo ou real, definido e representado por<sup>41</sup>

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

Se  $c$  é um número real ou complexo, as propriedades básicas do produto escalar são:

$$\begin{aligned}(f, g) &= (g, f)^* \\(f, g + h) &= (f, g) + (f, h) \\(f, cg) &= c(f, g) \\(cf, g) &= c^*(f, g) \\(f, f) &\geq 0 \quad \text{e} \quad (f, f) = 0 \quad \text{se e somente se} \quad f = 0\end{aligned}$$

<sup>40</sup>Ou qualquer conjunto de vetores.

<sup>41</sup>O produto escalar de duas funções  $f$  e  $g$  de uma variável  $x$ , em um intervalo  $a < x < b$ , com relação a uma função peso real positiva  $w(x)$  é dado por

$$\int_a^b w(x) f^*(x) g(x) dx$$

Por suas propriedades básicas,<sup>42</sup> o produto escalar em  $L^2$ , ou em qualquer conjunto de funções onde ele seja definido, é uma generalização do produto escalar ( $\vec{a} \cdot \vec{b}$ ) entre dois vetores geométricos  $\vec{a}$  e  $\vec{b}$ , que fazem um ângulo  $\theta$  entre si.

$$\vec{a} \cdot \vec{b} = |\vec{a}| |\vec{b}| \cos \theta$$

Se  $\vec{a}$ ,  $\vec{b}$  e  $\vec{c}$  são vetores e  $\lambda$  é um número real,

$$\begin{aligned} \vec{a} \cdot \vec{b} &= \vec{b} \cdot \vec{a} \\ \vec{a} \cdot (\vec{b} + \vec{c}) &= \vec{a} \cdot \vec{b} + \vec{a} \cdot \vec{c} \\ \vec{a} \cdot (\lambda \vec{b}) &= \lambda (\vec{a} \cdot \vec{b}) \\ \vec{a} \cdot \vec{a} &\geq 0 \quad \text{e} \quad \vec{a} \cdot \vec{a} = 0 \quad \text{se e somente se} \quad \vec{a} = 0 \end{aligned}$$

Assim, o número positivo  $\|f\| = (f, f)^{1/2}$  é denominado **norma** de  $f$ .

Se um vetor  $\vec{a}$  possui norma unitária, o produto escalar  $\vec{a} \cdot \vec{b}$ , em linguagem geométrica, é uma medida da projeção de  $\vec{b}$  sobre a direção de  $\vec{a}$ .

O produto escalar permite a definição de uma das relações mais utilizadas num conjunto de funções, a **ortogonalidade**: duas funções  $f$  e  $g$  são ditas ortogonais<sup>43</sup> se:

$$(f, g) = 0$$

## A.2 Operadores lineares, adjuntos e hermitianos

Um operador  $A$  que atua sobre as funções de  $L^2$  é dito linear se:

$$\begin{cases} A(f + g) = Af + Ag \\ A(cf) = c(Af) \end{cases}$$

Por exemplo, o operador derivada,  $d/dx$  é linear,

$$\begin{cases} \frac{d}{dx}(f + g) = \frac{df}{dx} + \frac{dg}{dx} \\ \frac{d}{dx}(cf) = c \left( \frac{df}{dx} \right) \end{cases}$$

São lineares também a multiplicação por um número e a integração de funções.

Um outro exemplo é a projeção de uma função  $f$  na “direção” de uma função de norma unitária  $\phi$ , definida por

$$P_\phi f = (\phi, f)\phi$$

Tal operador é linear,

$$P_\phi(cf) = (\phi, cf)\phi = c(\phi, f)\phi = c(P_\phi f)$$

<sup>42</sup>É importante notar que para satisfazer as propriedades básicas, é necessário que a integral

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

de qualquer função de  $L^2$  exista. Nesse sentido,  $L^2$  é dito o conjunto das funções de quadrado somável.

<sup>43</sup>Cabe notar que a relação de ortogonalidade em um conjunto de funções depende do intervalo de definição do produto escalar.

e, usualmente, chamado de **projedor**.

A **adição** de operadores lineares é naturalmente definida. Por exemplo, multiplicar uma função  $f(x)$  por um número  $c$  e adicionar o resultado à sua derivada,

$$c f(x) + f'(x)$$

pode ser atribuído ao operador

$$c + \frac{d}{dx} = c + D_x$$

denominado **soma** do operador multiplicação por um número com a derivada.

A ação sucessiva de uma mesma operação, ou de operações distintas, leva ao conceito de **multiplicação** de operadores. Por exemplo, a derivada segunda, que resulta da aplicação sucessiva da derivada primeira,

$$f''(x) = \frac{d}{dx} \left( \frac{d}{dx} f(x) \right)$$

é um outro operador linear, denominado **produto** de  $d/dx$  por  $d/dx$ , e representado por

$$\frac{d^2}{dx^2} \quad \text{ou} \quad D_x^2$$

A definição do produto permite representar de modo mais simples a característica fundamental de um projetor. Uma vez aplicado a uma função o resultado é invariante a aplicações sucessivas do mesmo,

$$P_\phi(P_\phi f) = P_\phi^2 f = P_\phi f \implies \boxed{P_\phi^2 = P_\phi}$$

ou seja, os projetores têm a propriedade de idempotência.

A partir do produto, surge também o conceito de operador **inverso**  $A^{-1}$  de um operador linear  $A$ ,

$$A^{-1}(Af) = f \Leftrightarrow A^{-1}A = 1$$

Nesse sentido, a integral indefinida e a derivada de uma função são operadores inversos um do outro

$$\int \frac{d}{dx} f(x) dx = \frac{d}{dx} \int f(x) dx = f(x)$$

$$L_x D_x = D_x L_x = 1 \implies D_x^{-1} = L_x \text{ ou } L_x^{-1} = D_x$$

Apesar de  $L_x$  e  $D_x$  serem operações comutativas, em geral a multiplicação entre operadores  $A$  e  $B$ , ao contrário da adição, não é comutativa,

$$AB \neq BA \quad (\text{em geral})$$

Desse modo, diz-se que as operações de multiplicar por  $x$  e a derivação  $d/dx$  não comutam,

$$\left\{ \begin{array}{l} \left( x \frac{d}{dx} \right) f = x \left( \frac{df}{dx} \right) = x f'(x) \\ \left( \frac{d}{dx} x \right) f = \frac{d}{dx}(xf) = x f'(x) + f(x) \end{array} \right.$$

$$x \frac{d}{dx} \neq \frac{d}{dx} x$$

ou

$$x \frac{d}{dx} - \frac{d}{dx} x = -1$$

Nesse contexto, a diferença entre os produtos  $x(d/dx)$  e  $(d/dx)x$ , que é igual ao operador  $-1$ , é denominado **comutador** de  $x$  e  $d/dx$  e, genericamente, é representado por

$$\left[ x, \frac{d}{dx} \right]_- = x \frac{d}{dx} - \frac{d}{dx} x = -1$$

Uma vez definidas a adição e a multiplicação de operadores, pode-se definir também as **funções de operadores lineares**<sup>44</sup>. A exponenciação  $e^A$ , ou em geral, qualquer forma funcional de um operador linear  $A$  é definida a partir da expansão em série de Taylor

$$e^A = 1 + A + \frac{A^2}{2} + \frac{A^3}{3!} + \dots$$

O produto escalar em  $L^2$ ,  $\left( f, \frac{d}{dx} g \right)$

$$\begin{aligned} \left( f, \frac{d}{dx} g \right) &= \int_{-\infty}^{\infty} f^* \frac{dg}{dx} dx \\ &= \underbrace{\left[ f^* g \right]_{-\infty}^{\infty}}_0 - \int_{-\infty}^{\infty} g \frac{df^*}{dx} dx \\ &= \int_{-\infty}^{\infty} \left( -\frac{df}{dx} \right)^* g dx \end{aligned}$$

pode ser escrito também como  $\left( -\frac{d}{dx} f, g \right)$

Assim, com relação ao produto escalar, pode-se associar ao operador linear  $A = \frac{d}{dx}$  um outro operador linear  $A^\dagger = -\frac{d}{dx}$ , denominado operador **adjunto**, tal que

$$\boxed{(f, Ag) = (A^\dagger f, g)}$$

Para a operação de multiplicação por um número complexo  $c$ ,

$$(f, cg) = c(f, g) = (c^* f, g)$$

de acordo com as propriedades do produto escalar,

$$c^\dagger = c^*$$

o adjunto de uma constante multiplicativa é o seu conjugado.

Um operador linear  $H$  é dito **hermitiano** ou **auto-adjunto** se  $H^\dagger = H$

---

<sup>44</sup>Que são também operadores lineares.

$$\boxed{(f, Hf) = (Hf, f) \text{ (real)}}^{45}$$

Assim como a ortogonalidade, o caráter hermitiano e a determinação do adjunto de um operador linear dependem do intervalo de definição do produto escalar.

Por exemplo, o operador  $i\frac{d}{dx}$  é hermitiano em  $L^2$ , pois

$$\begin{aligned} \left(f, i\frac{d}{dx}g\right) &= \int_{-\infty}^{\infty} f^* i\frac{dg}{dx} dx = i \underbrace{\left[f^*g\right]_{-\infty}^{\infty}}_0 - i \int_{-\infty}^{\infty} g \frac{df^*}{dx} dx \\ &= \int_{-\infty}^{\infty} \left(i\frac{df}{dx}\right)^* g dx = \left(i\frac{d}{dx}f, g\right) \end{aligned}$$

ou seja,  $\left(i\frac{d}{dx}\right)^\dagger = i\frac{d}{dx}$

Algumas propriedades da associação de um conjunto de operadores lineares  $\{A, B, C, \dots\}$  e do respectivo conjunto de operadores adjuntos  $\{A^\dagger, B^\dagger, C^\dagger, \dots\}$  são:

- $(f, [A + B]g) = (f, Ag + Bg) = (f, Ag) + (f, Bg)$   
 $= (A^\dagger f, g) + (B^\dagger f, g) = (A^\dagger f + B^\dagger f, g)$   
 $= ([A + B]^\dagger f, g) = ([A^\dagger + B^\dagger]f, g)$

↓

$$\boxed{[A + B]^\dagger = A^\dagger + B^\dagger}$$

Por exemplo,  $\left(\frac{d}{dx} + c\right)^\dagger = -\frac{d}{dx} + c^*$  em  $L^2$ .

- $(f, A^\dagger g) = (A^\dagger f, g)^* = (g, Af)^* = ([A^\dagger]^\dagger f, g) = (Af, g)$

↓

$$\boxed{[A^\dagger]^\dagger = A}$$

Um operador linear  $U$  que não modifica a norma de uma função de  $L^2$ ,

$$Uf = g \implies \|f\| = \|g\| \implies (f, f) = (g, g)$$

é dito **unitário**.

<sup>45</sup>Uma vez que,  $(f, Hf) = (H^\dagger f, f) = (f, H^\dagger f)^*$   
se  $H^\dagger = H \implies (f, Hf) = (f, Hf)^*$  (real)

Assim:

- o inverso de um operador unitário é o seu adjunto,

$$(Uf, Uf) = (U^\dagger Uf, f) = (f, f)$$

↓

$$U^\dagger U = 1 \implies U^{-1} = U^\dagger$$

- se  $H$  é um operador hermitiano,  $U = e^{iH}$  é unitário,

$$\left\{ \begin{array}{l} e^{iH} = 1 + iH - \frac{H^2}{2} - i\frac{H^3}{3!} + \dots \\ (e^{iH})^{-1} = e^{-iH} = 1 - iH - \frac{H^2}{2} + i\frac{H^3}{3!} + \dots \end{array} \right.$$

pois o operador adjunto de  $e^{iH}$  é  $e^{-iH}$ ,

$$(e^{iH})^\dagger = 1 - iH^\dagger - \frac{H^{\dagger 2}}{2} + i\frac{H^{\dagger 3}}{3!} + \dots$$

se  $H^\dagger = H$ ,

$$(e^{iH})^\dagger = 1 - iH - \frac{H^2}{2} + i\frac{H^3}{3!} + \dots = e^{-iH}$$

↓

$$(e^{iH})^\dagger e^{iH} = 1 \implies U = e^{iH} \text{ é unitário,}$$

ou seja,  $U^\dagger = U^{-1} = e^{-iH}$

### A.3 Autovalores e autofunções

Se  $A$  é um operador linear e  $\phi_n$  é tal que

$$A\phi_n = a_n\phi_n \quad (a_n \text{ real ou complexo})$$

ou seja, a ação do operador sobre  $\phi_n$  é simplesmente multiplicativa,  $\phi_n$  é dita uma **autofunção** de  $A$ , associada ao **autovalor**  $a_n$ .

Desse modo, um operador hermitiano  $H$  possui autovalores  $\{\epsilon_n\}$  reais e autofunções  $\{\phi_n\}$ , associadas a distintos autovalores, ortogonais.

$$\bullet \underbrace{(\phi_n, H\phi_n)}_{\epsilon_n} = \epsilon_n(\phi_n, \phi_n) \implies \epsilon_n = \frac{(\phi_n, H\phi_n)}{(\phi_n, \phi_n)} \quad (\text{real})$$

$$\bullet \underbrace{(\phi_n, H\phi_m)}_{\epsilon_m} = \underbrace{(H\phi_n, \phi_m)}_{\epsilon_n} \implies (\epsilon_n - \epsilon_m)(\phi_n, \phi_m) = 0$$

$$\text{se } \epsilon_n \neq \epsilon_m \implies (\phi_n, \phi_m) = 0$$

Por exemplo,  $e^{ax}$  e  $e^{bx}$  são autofunções do operador  $d/dx$ , associadas, respectivamente, aos autovalores positivos  $a$  e  $b$ ,

$$\left\{ \begin{array}{l} \frac{d}{dx} e^{ax} = a e^{ax} \\ \frac{d}{dx} e^{bx} = b e^{bx} \end{array} \right.$$

Por  $d/dx$  não ser hermitiano, no intervalo  $(-\infty, 0)$ <sup>46</sup>,  $e^{ax}$  e  $e^{bx}$  não são, necessariamente, ortogonais,

$$\begin{aligned} (e^{ax}, e^{bx}) &= \int_{-\infty}^0 e^{ax} e^{bx} dx \\ &= \frac{1}{a+b} [e^{(a+b)x}]_{-\infty}^0 = \frac{1}{a+b} \neq 0 \end{aligned}$$

No caso do operador  $id/dx$ , as funções  $e^{-iax}$  e  $e^{-ibx}$  são autofunções ortogonais, no intervalo  $(-\infty, \infty)$ , associadas aos autovalores reais e positivos  $a$  e  $b$ ,

$$\left\{ \begin{array}{l} i \frac{d}{dx} e^{-iax} = a e^{-iax} \\ i \frac{d}{dx} e^{-ibx} = b e^{-ibx} \end{array} \right.$$

e,

$$\begin{aligned} (e^{-iax}, e^{-ibx}) &= \int_{-\infty}^{\infty} e^{iax} e^{-ibx} dx = \int_{-\infty}^{\infty} e^{i(a-b)x} dx \\ &= \lim_{\epsilon \rightarrow 0} \left\{ \int_{-\infty}^0 e^{i(a-b)x} e^{\epsilon x} dx + \int_0^{\infty} e^{i(a-b)x} e^{-\epsilon x} dx \right\} \\ &= \lim_{\epsilon \rightarrow 0} \left[ \frac{1}{i(a-b) + \epsilon} - \frac{1}{i(a-b) - \epsilon} \right] = 0 \quad (a \neq b) \end{aligned}$$

---

<sup>46</sup>

$$\begin{aligned} \left( f, \frac{d}{dx} g \right) &= \int_{-\infty}^0 f \frac{dg}{dx} dx = [fg]_{-\infty}^0 - \int_{-\infty}^0 g \frac{df}{dx} dx \\ &= \underbrace{f(0)g(0)}_{\int_{-\infty}^{0+} \delta(x)f(x)g(x) dx} + \int_{-\infty}^0 \left( -\frac{df}{dx} \right) g dx \\ &= \left( \left( \frac{d}{dx} \right)^\dagger f, g \right) = \int_{-\infty}^{0+} \left[ \left( \delta(x) - \frac{d}{dx} \right) f \right] g dx \\ &\quad \downarrow \\ &\left( \frac{d}{dx} \right)^\dagger = \delta(x) - \frac{d}{dx} \end{aligned}$$

## A.4 Representação matricial

Normalizando-se<sup>47</sup> as autofunções  $\{\phi_n\}$  de um operador hermitiano  $H$ , esse conjunto pode ser expresso por

$$\{\phi_n \mid H\phi_n = \epsilon_n\phi_n \text{ e } (\phi_n, \phi_m) = \delta_{nm}\}$$

Em linguagem geométrica, se  $f$  é uma função de  $L^2$  e  $\phi_n(x)$  é uma autofunção normalizada de um operador hermitiano  $H$ , uma vez que o produto escalar  $(\phi_n, f)$  é a medida da projeção de  $f$  na direção de  $\phi_n$ <sup>48</sup>, pode-se representar, ao menos formalmente, qualquer função de  $L^2$  pela soma de suas projeções sobre todas as direções das autofunções  $\{\phi_n\}$ <sup>49</sup> de um operador linear hermitiano  $H$ , ou seja, por uma combinação linear do tipo,

$$f = \sum_n f_n \phi_n$$

onde os  $f_n$  são as medidas das respectivas projeções de  $f$  nas direções de  $\phi_n$

$$f_n = (\phi_n, f)$$

Assim,

$$f = \sum_n (\phi_n, f) \phi_n$$

Nem sempre a cada autovalor corresponde uma única autofunção,

$$A\phi_n^i = a_n\phi_n^i \quad \{i = 1, \dots, g_n\}$$

Nesse caso, diz-se que o autovalor  $a_n$  é degenerado e que o grau de degenerescência é  $g_n$ .

O conjunto de todas as funções obtidas pela combinação linear das  $g_n$  autofunções  $\phi_n$  de um operador hermitiano  $A$ , correspondentes a um autovalor  $a_n$ , é denominado **subespaço próprio** de  $a_n$ .

A utilização de autofunções ortogonais e normalizadas<sup>50</sup> de um operador hermitiano permite que equações ou expressões envolvendo operadores sejam representadas por matrizes.

Por exemplo, se duas funções  $f$  e  $g$  estão relacionadas por um operador linear  $L$ ,

$$f = Lg$$

e  $\{\phi_n\}$  é um conjunto de autofunções ortonormais, associadas aos autovalores  $a_n$  do operador hermitiano  $A$ , ou seja,

$$A\phi_n = a_n\phi_n$$

multiplicando-se a equação que envolve  $f$  e  $g$  escalarmente por  $\phi_n$  e utilizando-se a expansão

$$g = \sum_k g_k \phi_k$$

<sup>47</sup>O processo de normalização de uma função  $f$  consiste em dividi-la por sua norma,

$$\frac{f}{\|f\|} = h \implies \|h\| = \sqrt{(h, h)} = 1$$

<sup>48</sup>De modo mais geral, no subespaço próprio (ver no texto, adiante) associado ao autovalor  $\epsilon_n$ .

<sup>49</sup>Como essas direções são ortogonais entre si, o conjunto é denominado também de **base própria ortogonal** de  $H$ .

<sup>50</sup>Denominadas **bases ortonormais**.

$$\underbrace{(\phi_n, f)}_{f_n} = (\phi_n, Lg) = \sum_k \underbrace{(\phi_n, L\phi_k)}_{L_{nk}} g_k$$

ou seja,

$$f_k = \sum_k L_{nk} g_k$$

Uma vez que cada  $L_{nk}$ , denominado elemento de matriz, é univocamente determinado pelo operador  $L$  e pelas autofunções  $\phi_n$  de  $A$ , a matriz  $\mathcal{L}$ , constituída por esses elementos, é denominada **representação matricial** de  $L$  na base própria de  $A$ , e a equação pode ser expressa por

$$\boxed{\mathcal{F} = \mathcal{L} \mathcal{G}}$$

onde  $\mathcal{F}$  e  $\mathcal{G}$  são as representações matriciais (matrizes colunas) das funções  $f$  e  $g$ .

Desse modo, qualquer teoria física linear, como a Mecânica Newtoniana para sistemas lineares, o Eletromagnetismo em meios lineares ou a Mecânica Quântica de Schrödinger, pode ser expressa numa forma matricial. Historicamente, essa foi a primeira forma de apresentação da Mecânica Quântica por Heisenberg (1925).

## Bibliografia

- [1] N. Bohr, *Sobre a constituição de átomos e moléculas*, tradução por Fundação Calouste Gulbenkian, 1969.
- [2] M. Born, *Physics in my Generation*, Springer, 1969.
- [3] L. Boltzmann, *Lectures on Gas Theory*, University of California Press, Berkeley & Los Angeles, 1964.
- [4] F. Caruso, V. Oguri, *FÍSICA MODERNA - Origens Clássicas e Fundamentos Quânticos*, Elsevier Ed. Ltda., 2006.
- [5] A. H. Compton, *A Quantum Theory of the Scattering of X-Rays by Light Elements*, Phys. Rev. **21**, 484, 1923.
- [6] C. Davisson, L. H. Germer, *Diffraction of Electrons by a Crystal of Nickel*, Phys. Rev. **30**, 706, 1927.
- [7] A. Einstein, *Sobre um ponto de vista heurístico a respeito da produção e transformação da luz - 1905*, traduzido em O Ano Miraculoso de Einstein, Ed. UFRJ, 2001.
- [8] J. W. Gibbs, *Elementary Principles in Statistical Mechanics - 1904*, Ox Bow Press, 1981.
- [9] A. N. Kolmogorov, *Foundations of the Theory of Probability - 1933*, Chelsea Pub. Company, 1950.
- [10] H. Lebesgue, *Sur le développement de la notion d'intégrale*, Revue de Métaphysique et de Morale, 1927.
- [11] L. D. Landau, E. M. Lifshitz, *Statistical Physics*, Pergamon Press, 1968.
- [12] R. A. Millikan, *A Direct Determination of "h"*, Phys. Rev. **4**, n. 1, p. 73-75.
- [13] R. von Mises, *Probability, Statistics and Truth*, 2nd. edition, Dover Publications, Inc., 1981.
- [14] P. M. Morse, *Thermal Physics*, Benjamin, 1965.
- [15] M. Planck, *Sobre a Lei da Distribuição de Energia no Espectro Normal*, Annal. Phys. **4**, 1901, traduzido na Revista Brasileira de Ensino de Física, Vol 22, 4, 2000.
- [16] K. R. Popper, *Quantum Theory and the Schism in Physics - 1982*, Routledge, 1995.
- [17] J. W. S. Rayleigh, *Remarks upon the Law of Complete Radiation*, Philos. Mag., XLIX, 539, 1900.
- [18] C. Ruhla, *The Physics of Chance - from Blaise Pascal to Niels Bohr*, Oxford University Press, 1992.
- [19] E. Schrödinger, *Quantisation as a Problem of Proper Values*, Part I, II, III & IV, Annal. Phys. V. 79, 80 & 81, 1926, translated in Collected Papers on Quantum Mechanics, Chelsea Pub., 1982.
- [20] N. Studart, *A Invenção do Conceito de Quantum de Energia Segundo Planck*, Revista Brasileira de Ensino de Física, Vol 22, 4, 2000.
- [21] E. P. Wigner, *Symmetries and Reflections - Scientific Essays - 1967*, Ox Bow Press, 1979.