6 O Formalismo Matemático da Mecânica Quântica I

6.1 Espaços Vetoriais

Nesta seção expomos as noções básicas dos espaços vetoriais, pois o formalismo da mecânica quântica se baseia nestes conceitos.

Na formulação abstrata da mecânica quântica dizemos que $|\psi\rangle$ é um vetor num espaço de Hilbert. Para poder entender tal afirmação, precisamos saber o que é um espaço vetorial e, especialmente, o que é um espaço de Hilbert. E o que é um vetor? Comecemos com a noção do espaço vetorial (ou espaço linear).

Um espaço vetorial pode ser caracterizado por dez propriedades satisfeitas por seus elementos, chamados de **vetores**. Estes vetores podem ser de quaisquer natureza, contanto que cumpram os dez mandamentos. Especialmente as funções de onda da mecânica quântica têm o direito de serem chamados de vetores, pois, como veremos, elas pertencem ao grupo dos objetos escolhidos. (O nome "vetor" é devido à semelhança com quantidades que se comportam como flechas no plano ou no espaço físico. Alias, os espaços vetoriais da mecânica quântica são, geralmente, de dimensões infinitas, e são então chamados de espaços de Hilbert. David Hilbert foi um Matemático alemão, 1862-1943.)

Em mecânica quântica representamos geralmente os vetores pelos símbolos bra " $|\,\rangle$ " ou ket " $\langle\,|$ ", inventados por Dirac, conferir também "Mecânica",4.4.1 e 2.2.1 sobre vetores em geral. Aqui vem a lista dos 10 leis (também chamados de axiomas ou postulados) do espaço vetorial:

1. Vetores podem ser adicionados e sua soma é também um vetor

$$|x\rangle + |y\rangle = |z\rangle$$

2. A adição é comutativa: $|x\rangle + |y\rangle = |y\rangle + |x\rangle$

3. A adição é associativa: $(|x\rangle + |y\rangle) + |z\rangle = |x\rangle + (|y\rangle + |z\rangle)$

4. Há um vetor zero $|0\rangle$, tal que $|x\rangle + |0\rangle = |x\rangle$ para todos os vetores $|x\rangle$

5. Para cada vetor $|x\rangle$ há um vetor negativo $|y\rangle$ (ou $|-x\rangle$), tal que

$$|x\rangle + |y\rangle = |0\rangle$$

Observação: As propriedades 1,2,3,4,5 exprimem o fato de que um espaço vetorial é um **grupo** com a operação de adição de vetores. Um grupo com a propriedade 2 (comutatividade) é um grupo comutativo (ou Abeliano).

6. Vetores podem ser multiplicados por escalares a, b. O resultado sendo também um vetor. Se $|x\rangle$ for um vetor, então $a|x\rangle$ é também um vetor.

Observação: Os escalares podem ser números reais ou complexos. No espaço vetorial Hilbert da mecânica quântica, os escalares são números complexos.

7. A multiplicação por escalar é associativa: $a(b|x\rangle) = (ab)|x\rangle$

8. Vale a primeira lei distributiva: $(a+b)|x\rangle = a|x\rangle + b|x\rangle$

9. Vale a segunda lei distributiva: $a(|x\rangle + |y\rangle) = a|x\rangle + a|y\rangle$

10. Vale a invariância sob a multiplicação pela identidade: $1\,|x\rangle = |x\rangle$

Observação: A propriedade 10 não é óbvia, ela não pode ser deduzida das outras nove propriedades e deve, portanto, ser postulada.

Os espaços vetoriais usados na física possuem, quase sempre, um **produto interno** (ou produto escalar) que é um escalar, real ou complexo, formado de dois vetores, representado, geralmente, por $\langle x|y\rangle$ ou $(|x\rangle,|y\rangle)$.

As propriedades do produto interno são um pouco diferentes para espaços reais e espaços complexos:

1. Espaço real: $\langle x|y\rangle = \langle y|x\rangle$ o produto interno é simétrico

   Espaço complexo: $\langle x|y\rangle = \langle y|x\rangle^*$ diz-se que o produto interno possui

   simetria hermitiana. ($\langle y|x\rangle^*$ é o complexo conjugado do número $\langle y|x\rangle$)

2. Espaços reais e complexos: $\langle x|y+z\rangle = \langle x|y\rangle + \langle x|z\rangle$; $\langle x|ay\rangle = a\,\langle x|y\rangle$

   O produto interno é linear em relação ao segundo fator.

Observação: Por meio da primeira propriedade podemos comprovar que o produto interno é também linear em relação ao primeiro fator:

$\langle x+y|z\rangle = \langle x|z\rangle + \langle y|z\rangle$ ou também $\langle a+b|c+d\rangle = \langle a|c\rangle+\langle a|d\rangle + \langle b|c\rangle + \langle b|d\rangle$

mas para a complexo: $\langle ax|y\rangle = a^* \langle x|y\rangle$

Quando a é um número real, o asterisco é supérfluo. Diz-se que o produto interno em um espaço complexo é antilinear em relação ao primeiro fator.

A última propriedade do produto interno é

3. O produto interno possui uma **norma** positiva-definida:

$$\langle x|x\rangle \geq 0, \text{ onde } \langle x|x\rangle = 0 \text{ se e só se } |x\rangle = |0\rangle.$$

(A **norma** do vetor $|x\rangle$ definimos, também em C^n, como sendo a raiz quadrada de $\langle x|x\rangle$: $||x|| = \sqrt{(\langle x|x\rangle)}$ o que é um número real não-negativo. A norma escreve-se também como $||\,|x\rangle\,||$. Um vetor $|x\rangle$ é dito *normalizado* se $||x|| = 1$.)

Num espaço vetorial arbitrário (com produto interno) valem a Desigualdade de Schwarz e a Desigualdade Triangular.

A **Desigualdade de Schwarz**:

$$|\langle x|y\rangle| \leq ||x|| \cdot ||y|| \quad (1)$$

Note que | | à esquerda representa o valor absoluto de um número real, || || à direita representa o produto da norma de $|x\rangle$ com a norma de $|y\rangle$. Em vez de norma diz-se também comprimento.

A demonstração da desigualdade (1), embora não seja difícil, não é muito natural, mas exige um início criativo, ou seja um jeitinho. Podemos proceder da seguinte maneira:

Se $|x\rangle = |0\rangle$, então $||x|| = 0$ e $\langle x|y\rangle = 0$, logo (1) é válida. O mesmo argumento seria aplicável se $|y\rangle$ for $|0\rangle$. Suponha, agora, que nem $|x\rangle$ nem $|y\rangle$ é nulo e que c e d são duas constantes arbitrárias. Devido à propriedade 3 do produto interno podemos escrever

$$0 \leq \langle cx + dy|\, cx + dy\rangle$$

Por meio das outras propriedades introduzidas acima, podemos produzir os seguintes conclusões

$$0 \leq c^*\langle x|cx + dy\rangle + d^*\langle y|cx + dy\rangle$$

$$\leq c^*c\langle x|x\rangle + c^*d\langle x|y\rangle + d^*c\langle y|x\rangle + d^*d\langle y|y\rangle$$

Esta relação é válida para qualquer valor de c e d, então, deve ser válida para os valores especiais $c = -\langle x|y\rangle$ e $d = \langle x|x\rangle$. Usando estes valores, obtemos

$$0 \leq c^*cd - c^*dc - d^*cc^* + d^*d\langle y|y\rangle$$

$$\leq d[-c^*c + d^*\langle y|y\rangle]$$

$$\text{e} \quad 0 \leq - |\langle x|y\rangle|^2 + \langle x|x\rangle\langle y|y\rangle \quad \text{e daí}$$

$|\langle x|y\rangle|^2 \leq ||x||^2\ ||y||^2$ ou seja $|\langle x|y\rangle| \leq ||x||\ ||y||$.

A desigualdade de Schwarz podemos usar para definir um número θ tal que

$\cos \theta = \langle x|y\rangle / [\ ||x||\cdot||y||]$, $0 \leq \theta < \pi$ também no espaço R^n.

Outra desigualdade de grande importância, a Desigualdade Triangular, é só um corolário de (1). Vamos demonstrar, então, que

$$|| x + y || \leq ||x|| \cdot ||y|| \quad (2)$$

(**Desigualdade Triangular**)

Demonstração:

Sabemos da definição do comprimento de um vetor que

$$||x + y||^2 = \langle x+y|x+y\rangle = ||x||^2 + ||y||^2 + \langle x|y\rangle + \langle y|x\rangle$$

Com $\langle y|x\rangle = \langle x|y\rangle^*$ obtemos $\langle x|y\rangle + \langle y|x\rangle = 2\ \mathrm{Re}\langle x|y\rangle$, onde Re significa a parte real do número complexo que segue. (Em mecânica quântica o produto interno é, em geral, um número complexo.)

Temos, então, $||x + y||^2 = ||x||^2 + ||y||^2 + 2\mathrm{Re}\ \langle x|y\rangle$. Para todo número complexo $z = a + ib$ vale $a = \mathrm{Re}\ z \leq |z| = (a^2 + b^2)^{1/2}$, onde o signo = será válido quando z é real e positivo. Então

$$||x + y||^2 \leq ||x||^2 + ||y||^2 + 2|\langle x|y\rangle|$$

Do teorema anterior temos $|\langle x|y\rangle| \leq ||x|| \cdot ||y||$, onde o signo = só é válido quando $|x\rangle = \lambda\ |y\rangle$.

Consequentemente resulta

$$||x + y||^2 \leq (||x||^2 + ||y||^2 + 2\,|<x|y>|) = (||x|| + ||y||)^2$$

q.e.d;

6.1.1 Ortogonalidade

Uma **combinação linear** de um número finito de vetores é uma expressão

$$\sum_{i=1,n} a_i|x_i> = a_1|x_1> + a_2|x_2> + ... + a_n|x_n> \quad (3)$$

em que os a_i são escalares arbitrários. Um conjunto de vetores é chamado **linearmente independente,** se $\sum a_i\ |x_i> = |0>$ só se todos os a_i são zero. (Um conjunto de vetores é linearmente dependente, se pelo menos um dos a_i é diferente de zero.)

Num espaço vetorial V com produto interno, dois vetores $|x>$ e $|y>$ são ditos **ortogonais** se $<x|y> = 0$, e escrevemos $|x> \perp |y>$.

Um conjunto de vetores tal que quaisquer dois vetores são ortogonais é dito um **sistema ortogonal** de vetores. Um sistema ortogonal de vetores é dito um **sistema ortonormal** se cada vetor do sistema é normalizado ($||x_i|| = 1$ para cada $|x_i>$ do sistema. Num sistema ortonormal vale $<x_i|x_j> = \delta_{ij}$. (δ_{ij} é o chamado símbolo *delta* de Kronecker. $\delta_{ij} = 0$, se $i \neq j$ e $\delta_{ij} = 1$, se $i = j$)

Qualquer conjunto ortonormal de vetores é linearmente independente porque, se $\sum a_i\ |x_i> = |0>$, então temos para cada $|x_j>$

$$0 = <x_j|\sum_i a_i|x_i> > = \sum_i a_i <x_j|x_i> = \sum_i a_i\delta_{ij} = a_j$$

Assim, qualquer conjunto ortonormal de dimensão n é **base** de V.

Se o conjunto $\{|e_i>\}$ é uma base para o espaço de dimensão N, *qualquer* vetor $|x>$ pode ser representado por

$$|x\rangle = \sum_{i=1}^{N} a_i|e_i\rangle \quad (4)$$

(É costume designar os vetores da base de $\mathbf{e_i}$ ou $|e_i>$.) O conjunto dos N escalares $\{a_1, a_2, ..., a_n\}$ é chamado de **coordenadas** do vetor $|x>$ com relação à base $\{ |e_i>\}$. As coordenadas podemos calcular como

$$<e_j|x> = \sum_i <e_j|a_i|e_i> = \sum_i a_i <e_j|e_i> = \sum_i a_i\ \delta_{ij} = a_j \quad (5)$$

Portanto

$$|x\rangle = \sum_j \langle e_j|x\rangle\, |e_j\rangle \quad (6)$$

Os números (reais ou complexos) $\langle e_j|x\rangle$ são chamados de representantes de $|x\rangle$. O conjunto $\{a_i\}$ dos a_i é ordenado e único e é chamado de N-upla. A N-upla das coordenadas pode ser chamado de representação de $|x\rangle$ com relação à base dada.

É fácil demonstrar que o conjunto $\{a_i\}$ é único. Pois se $|x\rangle = \sum_i b_i\, |e_j\rangle$ for certo com outro conjunto $\{b_i\}$, então temos (subtraindo)

$$|0\rangle = \sum_{i=1,N} (a_i - b_i)\, |e_i\rangle$$

Já que os $|e_i\rangle$ são linearmente independentes, segue-se que $a_i - b_i = 0$, ou seja $a_i = b_i$ para todo i, como desejado.

A **dimensão** de um espaço vetorial é igual ao número máximo de vetores linearmente independentes.

Pode-se definir uma base para um espaço vetorial *infinito*, exigindo que cada vetor seja representado por uma combinação linear infinita dos vetores de base.

Na seguinte figura no plano R^2 temos uma ilustração da relação (6):

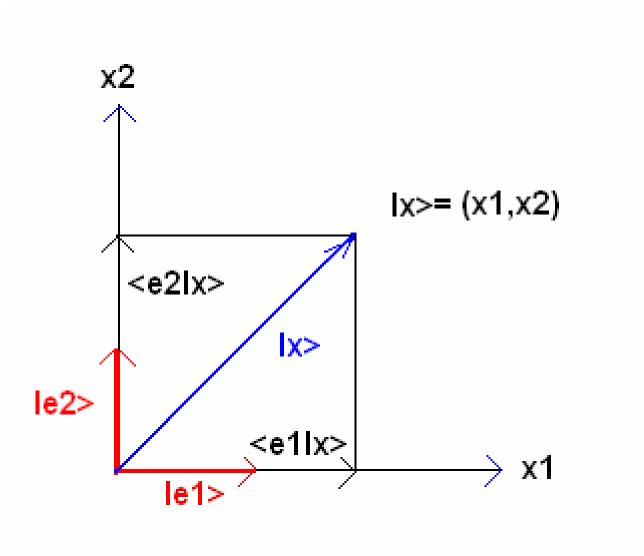

Fig. 1

O vetor $|x\rangle$ é projetado sobre os eixos x_1 e x_2. Os números $\langle e_1|x\rangle$ e $\langle e_2|x\rangle$ são as coordenadas do vetor $|x\rangle$ na direção dos eixos. Os vetores $\langle e_1|x\rangle\, |e_1\rangle$ e $\langle e_2|x\rangle\, |e_2\rangle$ são as componentes de $|x\rangle$ segundo os eixos x_1 e x_2, pois temos $|x\rangle = \sum_j \langle e_j|x\rangle\, |e_j\rangle = \langle e_1|x\rangle\, |e_1\rangle + \langle e_2|x\rangle\, |e_2\rangle$.

6.1.2 Funções ortogonais e expansões de uma função usando um sistema completo de autofunções.

Podemos generalizar o conceito de vetor para incluir também **funções** definidas sobre um intervalo [a,b] de números reais. Em outras palavras, uma função f(x) pode ser considerada como vetor num espaço de dimensões infinitas cujos eixos são designados pelo "índice" x. A "componente de f segundo a direção x" será o valor f(x) da função dada no ponto x do intervalo $a \le x \le b$.

A variável x que toma o lugar do índice i varia continuamente no intervalo [a,b]. Então, o vetor $|f\rangle$ tem um número infinito de componentes f(x).
O **produto interno** de duas funções (geralmente de valor complexo) f e g será definido como

$$\langle f|g\rangle = \int_{a,b} f(x)^* \, g(x)\, dx \qquad (7)$$

O somatório foi substituído por uma integral.

A ordem dos vetores no produto interno é importante, pois

$$\langle g|f\rangle = \int g(x)^* f(x)\, dx = \left(\int f(x)^* g(x)\, dx\right)^* = \langle f|g\rangle^* \qquad (8)$$

Dizemos que o produto interno possui simetria hermitiana.

O produto interno é linear em relação ao segundo fator (e antilinear em relação ao primeiro). Se a é um escalar, temos

$$\langle f|g + h\rangle = \langle f|g\rangle + \langle f|h\rangle \text{ e } \langle f|a\, g\rangle = a \langle f|g\rangle \qquad (9)$$

Podemos mostrar que o produto interno é positivo-definido

$$\langle f|f\rangle = \int_{a,b} |f(x)|^2\, dx \ge 0 \qquad (10)$$

Pois $f^*f = (\mathrm{Re}\, f - i\, \mathrm{Im}\, f)(\mathrm{Re}\, f + i\, \mathrm{Im}\, f) = (\mathrm{Re}\, f)^2 + (\mathrm{Im}\, f)^2$, o que é positivo definido.

(Uma função f é do *quadrado integrável* sobre um intervalo (a,b) se a integral

$\int_{a,b} |f(x)|^2\, dx$ existe.)

A *norma* de uma função (vetor) no intervalo [a,b] é, então, definido como

$$N(f) = \langle f | f \rangle = \int_{a,b} f^*f \, dx \quad (11)$$

Utiliza-se também aqui o símbolo $|| f(x) || = \langle f | f \rangle^{1/2}$ para designar a norma de f ou o comprimento de f.

A função f é *normalizada* quando a sua norma é 1.

Todos os teoremas anteriores são válidos para os novos vetores, também o conceito de ortogonalidade pode ser aplicado para funções:

Duas funções, f e g, são *ortogonais* no intervalo [a,b], se o seu produto interno é nulo em [a,b]:

$$\langle f | g \rangle = \int_{a,b} f(x)^* g(x) \, dx = 0 = \int_{a,b} g(x)^* f(x) \, dx = \langle g | f \rangle \quad (12)$$

Então, se o produto interno é nulo, a ordem das funções não tem importância.

Definimos um vetor do **espaço Hilbert**, **H**, sendo uma função complexa f de uma variável real x. Mas não todas essas funções são vetores em **H**, pois os vetores que constituem um espaço Hilbert (**H**) devem ter uma norma finita, ou seja $N(f) = \langle f | f \rangle = \int_{a,b} f^*f \, dx < \infty$.

Para generalizar a equação 4: $|x\rangle = \sum_{i=1}^{N} a_i |e_i\rangle$ para vetores em **H**, precisamos substituir os vetores ortonormais $\{|e_i\rangle\}$ por funções ortonormais $\{E_i(x)\}$. Mas não é suficiente o conjunto $\{E_i(x)\}$ ser um conjunto ortonormal, $\{E_i(x)\}$ também deve ser **completo** para que *qualquer* vetor de **H** pode ser escrito como combinação linear dos $E_i(x)$.

Os três vetores $\{|e_1\rangle, |e_2\rangle, |e_3\rangle\}$ constituem um conjunto completo e todo vetor no espaço geométrico R^3 pode ser escrito como $|x\rangle = c_1|e_1\rangle + c_2|e_2\rangle + c_3|e_3\rangle$. Os coeficientes c_i são as coordenadas (ou componentes) do vetor $|x\rangle$ e podem ser determinadas como *projeções* do vetor $|x\rangle$ sobre os vetores $|e_i\rangle$. Pois vimos acima que $c_1 = \langle e_1|x\rangle$, $c_2 = \langle e_2|x\rangle$ e temos agora também $c_3 = \langle e_3|x\rangle$.

(A demonstração de um conjunto ser completo não fazemos neste curso. Uma prova de integridade é sempre difícil.)

O conjunto $\{E_i(x)\}$ é um conjunto completo se, e só se, para qualquer vetor h(x) de **H** existe pelo menos um conjunto de escalares $\{c_i\}$ tal que

$$h(x) = \sum_{i=1,\infty} c_i E_i(x) \quad (13)$$

Os escalares c_i, normalmente complexos, são designados de componentes de h(x) ou coeficientes de expansão de h(x) na base ortonormal e completo $\{E_i(x)\}$.

Eles, os c_i, são os escalares $< E_i | h>$, assim como os $<e_j|x>$ na expansão $|x> = \sum_j <e_j|x> |e_j>$. Pois podemos escrever

$$<E_j | h> = < E_j | \sum c_i E_i > = \sum c_i <E_j | E_i> = \sum c_i \delta_{ij} = c_j$$

Por isso vale

$$h(x) = \sum_{i=1}^{\infty} \langle E_i | h \rangle E_i(x) \quad (14)$$

As **autofunções** de um **operador hermitiano**, veja mais adiante na seção 6.2, são ortogonais e completos (e os autovalores são reais) e toda função de onda pode ser expandida num conjunto completo de autofunções de um operador hermitiano, contanto que as autofunções satisfaçam as mesmas condições de contorno como a função de onda.

Se $\{\xi_i(x)\}$ é um conjunto completo de autofunções, então temos para a função de onda $\psi(x,t)$ de uma só partícula (em uma dimensão só) a seguinte expansão

$$\psi(x,t) = \sum_n c_n(t) \cdot \xi_n(x) \quad (15)$$

Os coeficientes $c_n(t) = <\xi_n|\psi>$ são as projeções de ψ sobre ξ_n.

(A notação $|\psi> = \sum_n c_n |\xi_n>$ com $c_n = <\xi_n|\psi>$ não é sempre a mais simples, por isso usa-se a miúdo uma simplificação, por exemplo $\psi(x,t) = \sum_n (\xi_n,\psi)\, \xi_n(x)$.)

Outro exemplo de um conjunto completo é o conjunto de funções trigonométricas $\{\cos(2n\pi x/L), \mathrm{sen}(2n\pi x/L)\}$, que podemos aplicar para a expansão de uma função f(x) (que deve cumprir com as exigências normais para uma função bem comportada, ou seja deve ser contínua e finita), que é periódica no intervalo [-L/2, L/2]. A expansão tem a forma

$$f(x) = a_0 + \sum_{n=1,\,\infty} [a_n \cos(2n\pi x/L) + b_n \,\mathrm{sen}\,(2n\pi x/L)] \quad (16)$$

Os coeficientes a_n, b_n nesta "Série de Fourier", confira "Mecânica", cap. 7, são deduzidos da função f(x):

$a_0 = 1/L \int f(x)\, dx$, a integral calcula-se aqui entre -L/2 e + L/2.

$b_n = 2/L \int f(x) \cos(2n\pi x/L)dx$, n = 1,2,3, ... ou: $b_n = 2/L <\cos(2n\pi x/L) | f >$

$c_n = 2/L \int f(x)\, \mathrm{sen}(2n\pi x/L)dx$, n = 1,2,3, ... (17)

A forma final de uma expansão contem as vezes somente poucos termos, já que a maioria dos coeficientes pode resultar nula.

Este caso se dá muitas vezes na mecânica quântica. Mas, em geral, é necessário utilizarmos todas as autofunções do operador hermitiano em questão.

As autofunções usadas para uma expansão na mecânica quântica não tem quer ser conectadas com o sistema cujo função de estado estamos expandindo. Por exemplo, poderíamos usar o sistema completo das autofunções do OHS para expandir uma função de onda do poço finito.

Se o espetro do operador contém autovalores discretos e também contínuos, como é o caso do poço finito o do átomo de hidrogênio, então, além da somatória sobre o espetro discreto, é necessário incluirmos na expansão de uma função também uma integral sobre o espetro contínuo, por exemplo

$$f = \sum_{1,\,N} a_i\xi_i + \int a(E)\,\xi(E)\,dE \quad (18)$$

Terminamos este parágrafo com um teorema que mostra uma expansão do produto interno.

Teorema de expansão

O produto interno $\langle f|g\rangle$ pode ser expandido em termos de um conjunto ortonormal $\{\varphi_i\}$ como

$$\langle f\,|\,g\rangle = \sum_i \langle f\,|\varphi_i\rangle\langle\varphi_i g\rangle \quad (19)$$

Prova:

Primeiro expandimos f e g: $f = \sum_i a_i\varphi_i$, $g = \sum_j b_j\varphi_j$. Logo temos

$$\langle f\,|g\rangle = \int f^*g = \sum_i\sum_j \int a_i^*\varphi_i^* b_j\varphi_j$$

$$= \sum_{ij} a_i^* b_j\langle\varphi_i|\varphi_j\rangle = \sum_{ij} a_i^* b_j\delta_{ij}$$

Agora, $a_i^* = \langle\varphi_i|f\rangle^* = \langle f\,|\varphi_i\rangle$ e $b_i = \langle\varphi_i|g\rangle$. Por isso temos

$$\langle f\,|g\rangle = \sum_i \langle f\,|\varphi_i\rangle\langle\varphi_i|g\rangle, \text{ q.e.d.}$$

A composição $\mathbf{|\varphi_i\rangle\langle\varphi_i|}$ aparecerá as vezes na mecânica quântica, de fato, apareceu já na seção 4.4 da Mecânica. Trata-se de um **operador**, e mais adiante vamos falar sobre estes objetos. O objeto $\mathbf{P}_i = |i\rangle\langle i|$ é denominado operador de projeção do ket $|i\rangle$. O conteúdo do teorema podemos, pelo momento, expressar como

$$\sum_i |\varphi_i\rangle\langle\varphi_i| = \mathbf{I} \quad (20)$$

ou seja, o objeto $\sum_i |\varphi_i><\varphi_i|$ funciona como um operador análogo ao número 1, ou seja, não faz nada, pois ele deixa inalterado todo ket $|v>$:

$$|v> = (\sum_i |i><i|)\, |v>$$

Chama-se $\sum_i |\varphi_i><\varphi_i|$ o operador identidade.

(Uma forma como $\mathbf{I} = \sum_{i=1,3} |e_i><e_i| = |e_1><e_1| + |e_2><e_2| + |e_3><e_3|$ chama-se em mecânica um diádico, especialmente diádico identidade. Cada $|e_i><e_i|$ é uma díada unitária, e um diádico é, por isso, uma soma de três díadas.

$P_1 := |e_1><e_1|$ projeta um vetor $|a>$ sobre o eixo $|e_1> := \mathbf{e}_1$, pois

$$P_1|a> = |e_1><e_1|a> = |e_1>\, a_1 = a_1\, \mathbf{e}_1.$$

O tensor $P_{12} := |e_1><e_1| + |e_2><e_2|$ projeta um vetor $|a>$ sobre o plano $(\mathbf{e_1},\mathbf{e_2})$, pois $P_{12}|a> = |e_1><e_1|a> + |e_2><e_2|a> = a_1\mathbf{e}_1 + a_2\, \mathbf{e}_2$. Os operadores P_1 e P_{12} são operadores de projeção. Na mecânica quântica utilizamos a mesma expressão.)

6.1.3 Funções ortonormais

Na mecânica quântica pedimos, geralmente, que as autofunções sejam normalizadas, ou seja pedimos que se cumpra a relação

$$\int_{-\infty}^{\infty} |\psi|^2 dx = 1 \qquad (21)$$

Até agora avaliamos a integral (21) diretamente e por meio do ajustamento de um fator de normalização que foi introduzido na função ψ. Existe, porem, um método "mecânico" com o qual se pode construir um conjunto ortonormal a partir de um conjunto completo de funções linearmente independentes. (Um conjunto completo de funções contém sempre um subconjunto linearmente independente.)

Trata-se do processo de ortogonalização de Gram-Schmidt. (J. P. Gram, 1850-1916, atuário dinamarquês, E. Schmidt, 1876-1959, matemático alemão, foi aluno de Hilbert.)

"Receita":

Suponha que queremos construir um conjunto ortonormal (= ortogonal e normalizado) $\{\varphi_i\}$ a partir do conjunto completo e linearmente independente $\{f_i\}$. (As φ_i são dois a dois ortonormais.)

Etapa 1. Seja $\varphi_0 = f_0/ N_0^{1/2}$, onde a norma de f_0 é $N_0 = <f_0|f_0>$. Nesta primeira etapa supomos que φ_0 seja proporcional a f_0.

Etapa 2. Seja φ_1 uma combinação linear de φ_0 e f_1, ou seja $\varphi_1 \approx c\varphi_0 + f_1$.
φ_1 seja ortogonal a φ_0 e normalizada. Então

$\varphi_1 = N_1^{-1/2} (c\varphi_0 + f_1)$, onde $<\varphi_0|\varphi_1> = 0$. Esta condição de ortogonalidade podemos escrever como

$$<\varphi_0|\varphi_1> = <\varphi_0|c\varphi_0+f_1> = c<\varphi_0|\varphi_0> + <\varphi_0|f_1> := 0$$

(Sabemos que $<f\,|cg> = c\,<f\,|g>$ e $<cf\,|g> = c^*<f\,|g>$.) Das φ_i pedimos ser ortogonais, ou seja $<\varphi_0|\varphi_0> = 0$. Isso dá para a constante c a expressão $c = -<\varphi_0|f_1>$. Agora temos $\varphi_1 = N_1^{-1/2} (f_1- <\varphi_0|f_1>\varphi_0)$.

A norma de φ_1 é
$N_1 = <\varphi_1|\varphi_1> = <f_1- <\varphi_0|f_1>\varphi_0|\, f_1- <\varphi_0|f_1>\varphi_0>$

$= <f_1|f_1> - <\varphi_0|f_1><f_1|\varphi_0> - <\varphi_0|f_1><\varphi_0|f_1> + <\varphi_0|f_1>^2<\varphi_0|\varphi_0>$

$= <f_1|f_1> - <\varphi_0|f_1><\varphi_0|f_1>^* = <f_1|f_1> - |<\varphi_0|f_1>|^2$

Continua-se este processo até o conjunto completo e ortonormal é gerado.

O termo geral pode ser escrito como

$$\varphi_k = \frac{1}{\sqrt{N_k}}(f_k - \sum_{j=0}^{k-1}\langle\varphi_j \mid f_k\rangle\varphi_j), \quad (22)$$

onde
$$N_k = \langle f_k \mid f_k\rangle - \sum_{j=0}^{k-1}|\langle\varphi_j \mid f_k\rangle|^2 \quad (23)$$

Como **exemplo** do processo de Gram-Schmidt vamos considerar as funções

$\{\exp(-x^2/2)x^n\}$ no intervalo $(-\infty, \infty)$.

Primeiro demonstraremos que essas funções têm uma norma finita:

$$\langle f_n|f_n\rangle = \int_{-\infty,\infty} \exp(-x^2/2)x^n \exp(-x^2/2)x^n\, dx = \int_{-\infty,\infty} \exp(-x^2)x^{2n}\, dx$$

Visto que o integrando é uma função par, podemos simplificar a integral

$\langle f_n|f_n\rangle = 2 \int_{0,\infty} \exp(-x^2)x^{2n}\, dx = \int_{0,\infty} e^{-y}\, y^{n-1/2}dx$ o que pode ser sucessivamente integrado por partes dando

$$\langle f_n|f_n\rangle=(n-1/2)(n-3/2)\ ...\ (1/2) \int_{0,\infty} e^{-y}y^{-1/2}dy = 2(n-1/2)(n-3/2)...(1/2) \int_{0,\infty} \exp(-x^2)dx$$

$$= (2n-1)(2n-3)\ ...(3)(1)\pi^{1/2}/2^n. \qquad (24)$$

A norma N_n é finita para qualquer n, o que significa que todas as funções no conjunto $\{\exp(-x^2/2)\ x^n\}$ são normalizáveis. (É a presença do termo decaindo, $\exp(-x^2/2)$, que torna a integral finita e as funções quadraticamente integráveis.) O conjunto $\{x^n\}$ não é normalizável sobre $(-\infty, \infty)$.

Etapa 1.

A norma de f_0 é $\langle f_0|f_0\rangle = \int_{-\infty,\infty} \exp(-x^2/2)x^0\, dx = \int_{-\infty,\infty}\exp(-x^2/2)dx = \pi^{1/2}$

$\rightarrow\ \varphi_0 = \pi^{-1/4}\exp(-x^2/2)$

Etapa 2.

Vamos calcular $\langle\varphi_0|f_1\rangle$ e $\langle f_1|f_1\rangle$:

$\langle\varphi_0|f_1\rangle = \int_{-\infty,\infty} \exp(-x^2/2)\pi^{-1/4} \exp(-x^2/2)x\, dx = 0$, pois o integrando é ímpar.

Da Eq. 24 obtemos $\langle f_1|f_1\rangle = \pi^{1/2}/2$ e da Eq. 23 segue

$N_1 = \langle f_1|f_1\rangle - |\langle\varphi_0|f_1\rangle|^2 = \pi^{1/2}/2$ e

$\varphi_1= N_1^{-1/2}(f_1 - \langle\varphi_0|f_1\rangle\varphi_0) = \sqrt{2} \cdot \pi^{-1/4} \exp(-x^2/2)\, x$

Etapa 3.

Note primeiro que todas as integrais $\langle\varphi_i|f_j\rangle$ com i = par e j = ímpar (ou vice versa) são zero.

Agora utilizaremos as fórmulas gerais (22) e (23)

$\varphi_2 = N_2^{-1/2} (f_2 - \langle\varphi_0|f_2\rangle\varphi_0 - \langle\varphi_{1(\text{ímpar})}|f_{2(\text{par})}\rangle\ \varphi_1 = N_2^{-1/2}(f_2 - \langle\varphi_0|f_2\rangle\varphi_0)$

$N_2 = <f_2|f_2> - |<\varphi_0|f_2>|^2 - |<\varphi_1|f_2>|^2 = <f_2|f_2> - |<\varphi_0|f_2>|^2$

Da Eq. 24 obtemos $<f_2|f_2> = 3\pi^{1/2}/4$.

$<\varphi_0|f_2> = \int_{-\infty,\infty} \pi^{-1/4}\exp(-x^2/2) \exp(-x^2/2)x^2\, dx = \pi^{-1/4} \int_{-\infty,\infty} x^2 \exp(-x^2)\, dx = \pi^{1/4}/2$

Finalmente:

$N_2 = 3/4 \cdot\pi^{1/2} - 1/4 \cdot\pi^{1/2} = \pi^{1/2}/2$

$\varphi_2 = 2^{1/2}\pi^{-1/4} [\exp(-x^2/2)x^2 - \pi^{1/4}/2\ \pi^{-1/4} \exp(-x^2/2)]$

$= \pi^{-1/4} \exp(-x^2/2) (2x^2 - 1) 2^{-1/2}$

O conjunto de funções ortonormais geradas dessa maneira será da forma

$\pi^{-1/4} \exp(-x^2/2) H_n(x)$, onde $H_0(x) = 1$; $H_1(x) = 2^{1/2} x$; $H_2 = 2x^2 - 1$, ...

Esses polinômios são proporcionais aos polinômios de Hermite, confira 4.4.1 no capítulo_5.

Problemas

1. Construa os polinômios de Legendre usando o processo de Gram-Schmidt para o conjunto linearmente independente e completo $\{x^n\}$ sobre [-1, +1]

2. Construir os polinômios de Laguerre usando o processo de Gram-Schmidt para o conjunto linearmente independente e completo $\{\exp(-x/2)x^n\}$ sobre $(0,+\infty)$

Solução:

Problema 1:

n = 0: $f_0= x^0 = 1$

$N_0 = <f_0|f_0> = \int_{-1,+1}1\cdot dx = 2$; $\varphi_0 = f_0/\sqrt{N_0} = 1/\sqrt{2}$

n = 1: $f_1 = x$

$N_1 = <f_1|f_1> - |<\varphi_0|f_1>|^2$; $<f_1|f_1> = \int_{-1,1}x^2dx = 2/3$

$<\varphi_0 I f_1> = \int_{-1,1} 2^{-1/2} x dx = 0$, então

$N_1 = <f_1 I f_1> = 2/3$; $\varphi_1 = N_1^{-1/2}(f_1 - <\varphi_0 I f_1>\varphi_0) = (3/2)^{1/2}\ x$

n = 2: $f_2 = x^2$

$N_2 = <f_2 I f_2> - |<\varphi_0 I f_2>|^2 - |<\varphi_1 I f_2>|^2$

$<f_2|f_2> = \int_{-1,1} x^4 dx = 2/5$; $<\varphi_0 I f_2> = \int_{-1,1} 2^{-1/2} x^2 dx = 2^{1/2}/3$

$<\varphi_1 I f_2> = \int_{-1,1} (3/2)^{1/2} x^3 dx = 0$, então

$N_2 = <f_2 I f_2> - |<\varphi_0 I f_2>|^2 = 2/5 - (2^{1/2}/3)^2 = 8/45$

$\varphi_2 = N_2^{-1/2}(f_2 - <\varphi_0 I f_2>\varphi_0) - <\varphi_1 I f_2>\varphi_1) = (5/2)^{1/2} \cdot 1/2 \cdot (3x^2-1)$

n = 3: $f_3 = x^3$

$N_3 = <f_3 I f_3> - |<\varphi_0 I f_3>|^2 - |<\varphi_1 I f_3>|^2 - |<\varphi_2 I f_3>|^2$

$<f_3 I f_3> = \int_{-1,1} x^6 dx = 2/7$; $<\varphi_1 I f_3> = \int_{-1,1} (3/2)^{1/2} x^4 dx = 6^{1/2}/5$

$<\varphi_0 I f_3> = \int_{-1,1} 2^{-1/2} x^3 dx = 0$; $<\varphi_2 I f_3> = \int_{-1,1} (5/8)^{1/2}(3x^2-1)dx = 0$

$N_3 = <f_3 I f_3> - |<\varphi_1 I f_3>|^2 = 8/175$

$\varphi_3 = N_3^{-1/2}(f_3 - <\varphi_0 I f_3>\varphi_0) - <\varphi_1 I f_3>\varphi_1 - <\varphi_2 I f_3>\varphi_2)$

$\varphi_3 = (7/8)^{1/2}\ (5x^3 - 3x) = (7/2)^{1/2} \cdot 1/2\ (5x^3 - 3x)$

Em geral temos $<f_n I f_n> = \int_{-1,1} x^{2n} dx = 2/(2n + 1)$

$<\varphi_i I f_j>$ 0 para i = par e j = ímpar (ou vice versa)

Resulta, então :

$\varphi_0 = (1/2)^{1/2} \cdot 1$; $\varphi_1 = (3/2)^{1/2} \cdot x$; $\varphi_2 = (5/2)^{1/2} \cdot 1/2 \cdot (3x^2 - 1)$

$\varphi_3 = (7/2)^{1/2} \cdot 1/2 \cdot (5x^3 - 3x)$ ... $\varphi_n = [(2n+1)/2]^{1/2}\ P_n(x)$

onde $P_n(x)$ são os polinômios de Legendre

Problema 2:

n = 0: $f_0 = e^{-x/2}$

$N_0 = \langle f_0 | f_0 \rangle = \int_{0,\infty} e^{-x}\, dx = 1$; $\varphi_0 = f_0/\sqrt{N_0} = e^{-x/2}$

n = 1: $f_1 = e^{-x/2} \cdot x$

$N_1 = \langle f_1 | f_1 \rangle - |\langle \varphi_0 | f_1 \rangle|^2$; $\langle f_1 | f_1 \rangle = \int_{0,\infty} e^{-x} x^2 dx = 2!$

$\langle \varphi_0 | f_1 \rangle = \int_{0,\infty} e^{-x} x dx = 1!$, então

$N_1 = 2! - 1! = 1$; $\varphi_1 = N_1^{-1/2}(f_1 - \langle \varphi_0 | f_1 \rangle \varphi_0) = (x-1)\, e^{-x/2}$

n = 2: $f_2 = e^{-x/2} x^2$

$N_2 = \langle f_2 | f_2 \rangle - |\langle \varphi_0 | f_2 \rangle|^2 - |\langle \varphi_1 | f_2 \rangle|^2$

$\langle f_2 | f_2 \rangle = \int_{0,\infty} e^{-x} x^4 dx = 4!$; $\langle \varphi_0 | f_2 \rangle = \int_{0,\infty} e^{-x} x^2 dx = 2!$

$\langle \varphi_1 | f_2 \rangle = \int_{0,\infty} (x-1) e^{-x} x^2 dx = 3! - 2! = 4$, então

$N_2 = 4! - (2!)^2 - 4^2 = 4$

$\varphi_2 = N_2^{-1/2}(f_2 - \langle \varphi_0 | f_2 \rangle \varphi_0) - \langle \varphi_1 | f_2 \rangle \varphi_1) = 1/2\, (x^2 - 4x + 2)\, e^{-x/2}$

n = 3: $f_3 = e^{-x/2} x^3$

$N_3 = \langle f_3 | f_3 \rangle - |\langle \varphi_0 | f_3 \rangle|^2 - |\langle \varphi_1 | f_3 \rangle|^2 - |\langle \varphi_2 | f_3 \rangle|^2$

$\langle f_3 | f_3 \rangle = \int_{0,\infty} e^{-x} x^6 dx = 720$; $\langle \varphi_0 | f_3 \rangle = \int_{0,\infty} e^{-x} x^3 dx = 3! = 6$

$\langle \varphi_1 | f_3 \rangle = \int_{0,\infty} (x-1) e^{-x} x^3 dx = 18$

$\langle \varphi_2 | f_3 \rangle = \int_{0,\infty} 1/2(x^2 - 4x + 2) e^{-x} x^3 dx = 18$

$N_3 = 36$

$\varphi_3 = N_3^{-1/2}(f_3 - \langle \varphi_0 | f_3 \rangle \varphi_0) - \langle \varphi_1 | f_3 \rangle \varphi_1 - \langle \varphi_2 | f_3 \rangle \varphi_2)$

$\varphi_3 = 1/6\, (x^3 - 9x^2 + 18x - 6)\, e^{-x/2}$

Resulta, então :

$\varphi_0 = e^{-x/2}$; $\varphi_1 = (x - 1)e^{-x/2}$; $\varphi_2 = 1/2\ (x^2 - 4x + 2)e^{-x/2}$

$\varphi_3 = 1/6 \cdot (x^3 - 9x^2 + 18x - 6)e^{-x/2}$ ... $\varphi_n = (-1)^n\ L_n(x)\ e^{-x/2}$

onde $L_n(x)$ são os polinômios de Laguerre.