

Anexo à Teoria de Perturbações  
Degenerada, até pag. 36

## § TEORIA DE PERTURBAÇÕES DEGENERADA a 2ª ORDEM

Usamos as mesmas equações das páginas 30-31.  
O ket pode ser expandido em série perturbativa

$$|l\rangle = |l^{(0)}\rangle + g|l^{(1)}\rangle + g^2|l^{(2)}\rangle + \dots$$

Sejam  $\{|n_\alpha^{(0)}\rangle\}$  o conjunto de estados degenerados em ordem zero. Para o ket  $|l\rangle$  temos:

$$\lim_{g \rightarrow 0} |l\rangle = |l^{(0)}\rangle = \sum_{\alpha} \langle n_\alpha^{(0)} | l^{(0)} \rangle |n_\alpha^{(0)}\rangle$$

O índice  $\alpha$  percorre o grau de degenerescência.

Para a energia é feita uma expansão da mesma forma:

$$\begin{aligned} \bar{E}_n &= E_n^{(0)} + g E_n^{(1)} + g^2 E_n^{(2)} + \dots \\ &\equiv E_n^{(0)} + \Delta_n \end{aligned}$$

Analisamos o caso que tem  $E_n^{(1)} = 0$  em 1ª ordem de teoria de perturbações degenerada. Especificamente, tratamos o caso de

$$\langle n_\alpha^{(0)} | V | n_\beta^{(0)} \rangle = 0,$$

para a perturbação. Neste caso a degenerescência não é removida em 1ª ordem. Agora temos:

$$\Delta_n = g^2 E_n^{(2)} + g^3 E_n^{(3)} + \dots$$

A eq. de Schrödinger é:

$$(\mathcal{H}_0 + gV)|e\rangle = (E_m^{(0)} + g^2 E_m^{(2)} + \dots)|e\rangle$$

ou

$$(\mathcal{H}_0 - E_m^{(0)})|e\rangle = (\Delta_m - gV)|e\rangle,$$

e em ordens sucessivas obtemos:

$$o(g^0): (\mathcal{H}_0 - E_m^{(0)})|e^{(0)}\rangle = 0 \quad (\text{I})$$

$$o(g): (\mathcal{H}_0 - E_m^{(0)})|e^{(1)}\rangle = -V|e^{(0)}\rangle \quad (\text{II})$$

$$o(g^2): (\mathcal{H}_0 - E_m^{(0)})|e^{(2)}\rangle = E_m^{(2)}|e^{(0)}\rangle - V|e^{(1)}\rangle \quad (\text{III})$$

⋮

Como o potencial perturbativo não tem elementos de matriz entre estados degenerados, projetamos a eq. (II)

como:

( $m \neq n$ )

$$\langle m^{(0)} | (\mathcal{H}_0 - E_m^{(0)}) | e^{(1)} \rangle = - \langle m^{(0)} | V | e^{(0)} \rangle$$

$$(E_m^{(0)} - E_m^{(0)}) \langle m^{(0)} | e^{(1)} \rangle = - \sum_{\alpha} \langle m^{(0)} | V | n_{\alpha}^{(0)} \rangle \langle n_{\alpha}^{(0)} | e^{(0)} \rangle$$

com  $\langle e^{(1)} | e^{(0)} \rangle = 0$ ,  $|e^{(1)}\rangle$  só tem componentes fora do espaço de degenerescência. Os coeficientes lineares de  $|e^{(1)}\rangle$  são:

$$\langle m^{(0)} | \ell^{(1)} \rangle = \sum_{(\alpha)} \frac{\langle m^{(0)} | V | n_{\alpha}^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} \langle n_{\alpha}^{(0)} | \ell^{(0)} \rangle \quad (\text{IV})$$

$(m \neq n)$

Para calcular a energia perturbada  $E_n^{(2)}$ , projetamos a eq. (III) sobre o espaço degenerado:

$$0 = \langle n_{\beta}^{(0)} | (H_0 - E_n^{(0)}) | \ell^{(2)} \rangle = E_n^{(2)} \langle n_{\beta}^{(0)} | \ell^{(0)} \rangle - \langle n_{\beta}^{(0)} | V | \ell^{(1)} \rangle$$

e como  $V$  não tem elementos de matriz no espaço degenerado:

$$0 = E_n^{(2)} \langle n_{\beta}^{(0)} | \ell^{(0)} \rangle - \sum_{m \neq n} \langle n_{\beta}^{(0)} | V | m^{(0)} \rangle \langle m^{(0)} | \ell^{(1)} \rangle \quad (\text{V})$$

e substituímos (IV) em (V), obtendo umas equações lineares homogêneas para os coeficientes  $\langle n_{\beta}^{(0)} | \ell^{(0)} \rangle$  da expansão de  $|\ell^{(0)}\rangle$ :

$$0 = E_n^{(2)} \langle n_{\beta}^{(0)} | \ell^{(0)} \rangle - \sum_{\alpha} \sum_{m \neq n} \frac{\langle n_{\beta}^{(0)} | V | m^{(0)} \rangle \langle m^{(0)} | V | n_{\alpha}^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} \times \langle n_{\alpha}^{(0)} | \ell^{(0)} \rangle,$$

que pode ser escrita na forma evidentemente homogênea:

36<sup>14</sup>

$$0 = \sum_{\alpha} \left\{ \sum_{m \neq n} \frac{\langle n_{\beta}^{(0)} | V | m^{(0)} \rangle \langle m^{(0)} | V | n_{\alpha}^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} - \delta_{\beta\alpha} E_m^{(2)} \right\} \langle n_{\alpha}^{(0)} | l^{(0)} \rangle$$

Solução não trivial no caso:

$$\det \left[ \sum_{m \neq n} \frac{\langle n_{\beta}^{(0)} | V | m^{(0)} \rangle \langle m^{(0)} | V | n_{\alpha}^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} - \delta_{\beta\alpha} E_m^{(2)} \right] = 0,$$

que é uma equação secular para  $E_m^{(2)}$  de dimensão igual à dimensão do espaço de degenerescência (grau  $\nu$ ).

Problema 5,12 (Sakurai, "MQM")

1

Matriz Hamiltoniana

$$H = \begin{pmatrix} E_1 & 0 & a \\ 0 & E_1 & b \\ a^* & b^* & E_2 \end{pmatrix}, \quad E_2 > E_1$$

A eq. secular exata é dada por:

$$0 = \det \begin{pmatrix} E_1 - E & 0 & a \\ 0 & E_1 - E & b \\ a^* & b^* & E_2 - E \end{pmatrix} =$$

$$= (E_1 - E)^2 (E_2 - E) - (E_1 - E) |a|^2 - (E_1 - E) |b|^2$$

Temos:

$$(E_1 - E) \left[ (E_1 - E)(E_2 - E) - |a|^2 - |b|^2 \right] = 0$$

Soluções:  $E = E_1$

$$E^2 - E(E_1 + E_2) + E_1 E_2 - |a|^2 - |b|^2 = 0$$

$$E_{\pm} = \frac{(E_1 + E_2)}{2} \pm \frac{1}{2} \sqrt{(E_1 - E_2)^2 + 4(|a|^2 + |b|^2)}$$

Def:  $0 < \Delta \equiv E_2 - E_1 = |E_2 - E_1|$

$$\begin{cases} E_{\pm} = \frac{E_1 + E_2}{2} \pm \frac{\Delta}{2} \sqrt{1 + 4 \frac{|a|^2 + |b|^2}{\Delta^2}} \\ E = E_1 \end{cases}$$

São as três soluções exatas do problema. No caso que  $(|a|^2 + |b|^2) \ll \Delta^2$ , podemos expandir a raiz:

$$\sqrt{1 + 4 \frac{(|a|^2 + |b|^2)}{\Delta^2}} \approx 1 + \frac{2(|a|^2 + |b|^2)}{\Delta^2}$$

e as soluções ficam:

$$E_{\pm} = \frac{E_1 + E_2}{2} \pm \frac{E_2 - E_1}{2} \pm \frac{|a|^2 + |b|^2}{\Delta}$$

$$E_{\pm} \approx \begin{cases} E_2 + \frac{|a|^2 + |b|^2}{E_2 - E_1} \\ E_1 - \frac{|a|^2 + |b|^2}{E_2 - E_1} \end{cases}, E = E_1$$

Para usar (e comparar com) a Teoria de Perturbações, separamos o Hamiltoniano em uma parte diagonal  $H_0$ , e uma parte não diagonal que será o potencial perturbativo. A validade da TP significa que os parâmetros  $|a|, |b|$  são pequenos em comparação com  $(E_2 - E_1) = \Delta$ .

$$H_0 = \begin{pmatrix} E_1 & 0 & 0 \\ 0 & E_1 & 0 \\ 0 & 0 & E_2 \end{pmatrix}, \quad V = \begin{pmatrix} 0 & 0 & a \\ 0 & 0 & b \\ \alpha^* & b^* & 0 \end{pmatrix}$$

$$V = \begin{pmatrix} V_{00} & V_{01} & V_{02} \\ V_{10} & V_{11} & V_{12} \\ V_{20} & V_{21} & V_{22} \end{pmatrix}$$

Todos os elementos de matriz de  $V$  são nulos no sub-  
espaço de degenerescência. Portanto a degenerescência  
não é removida em 1ª ordem. Usamos TP a 2ª  
ordem (cegamente) no caso não degenerado:

$$E_0 = E_1 + \frac{|\langle 0 | V | 2 \rangle|^2}{E_1 - E_2} = E_1 + \frac{|a|^2}{E_1 - E_2}$$

$$E_1 = E_1 + \frac{|\langle 1 | V | 2 \rangle|^2}{E_1 - E_2} = E_1 + \frac{|b|^2}{E_1 - E_2}$$

$$E_2 = E_2 + \frac{|\langle 0 | V | 2 \rangle|^2}{E_2 - E_1} + \frac{|\langle 1 | V | 2 \rangle|^2}{E_2 - E_1}$$

$$= E_2 + \frac{|a|^2 + |b|^2}{E_2 - E_1}$$

As correções, no sub-espaço degenerado, estão erradas,

mas nada justifica o uso de TP não degenerada a 2ª ordem quando a degenerescência não foi removida a 1ª ordem.

Usamos então TP degenerada a 2ª ordem (ver notas de aula). A equação secular que temos que resolver para o sub-espaço degenerado é

$$\det(M_{\beta\alpha} - \delta_{\beta\alpha} E_m^{(2)}) = 0$$

Com

$$M_{\beta\alpha} = \sum_{m \neq n} \frac{\langle m_{\beta}^{(0)} | V | m^{(0)} \rangle \langle m^{(0)} | V | n_{\alpha}^{(0)} \rangle}{E_m^{(0)} - E_n^{(0)}}$$

$$M_{00} = \frac{\langle 0 | V | 2 \rangle \langle 2 | V | 0 \rangle}{E_1 - E_2} = \frac{|a|^2}{E_1 - E_2}$$

$$M_{11} = \frac{|\langle 1 | V | 2 \rangle|^2}{E_1 - E_2} = \frac{|b|^2}{E_1 - E_2}$$

$$\begin{aligned} M_{01} &= \frac{\langle 0 | V | 2 \rangle \langle 2 | V | 1 \rangle}{E_1 - E_2} = \frac{V_{02} V_{21}}{E_1 - E_2} \\ &= \frac{a b^*}{E_1 - E_2} = M_{10}^* \end{aligned}$$

A equação secular é:

5

$$\begin{vmatrix} -\frac{|a|^2}{(\epsilon_1 - \epsilon_2)} - E & -\frac{ab^*}{\Delta} \\ -\frac{a^*b}{\Delta} & -\frac{|b|^2}{\Delta} - E \end{vmatrix} = 0$$

$$0 = \begin{vmatrix} +|a|^2 + E\Delta & +ab^* \\ +a^*b & +|b|^2 + E\Delta \end{vmatrix} =$$

$$= (E\Delta + |a|^2)(E\Delta + |b|^2) - |a|^2|b|^2$$

$$= E^2\Delta^2 + E\Delta(|a|^2 + |b|^2) + \cancel{|a|^2|b|^2} - \cancel{|a|^2|b|^2} = 0$$

$$E\Delta(E\Delta + |a|^2 + |b|^2) = 0$$

Soluções:  $E = 0$

$$E = -\frac{|a|^2 + |b|^2}{\Delta} = -\frac{|a|^2 + |b|^2}{\epsilon_2 - \epsilon_1}$$

As energias corrigidas em 2ª ordem são:

$$\begin{cases} \mathcal{E}_0 = E_1 & ) \\ \mathcal{E}_1 = E_1 - \frac{|a|^2 + |b|^2}{E_2 - E_1} & , \\ \mathcal{E}_2 = E_2 + \frac{|a|^2 + |b|^2}{E_2 - E_1} & , \end{cases}$$

Resultado que coincide com a expansão da solução exata para

$$\frac{|a|^2 + |b|^2}{E_2 - E_1} \ll 1$$

Como fato curioso, note <sup>que</sup> Brillouin-Wigner aplicado ao nível não degenerado, fornece a solução exata (a toda ordem):

$$E = E_2 + \frac{|a|^2 + |b|^2}{E - E_1}$$

Solução de:

$$E^2 - (E_1 + E_2)E + E_1 E_2 - (|a|^2 + |b|^2) = 0$$

que são

$$E_{\pm} = \frac{(E_1 + E_2)}{2} \pm \frac{1}{2} \sqrt{(E_1 - E_2)^2 + 4(|a|^2 + |b|^2)}$$

Páginas 41-43

'Em Branco'

## Métodos variacionais

Os métodos perturbativos se baseiam no fato de que existe um problema relativamente semelhante que tem solução exata. O método variacional não tem esta limitação, e justamente é amplamente usado nesta situação.

Considere nos um sistema físico arbitrário com Hamiltoniano  $H$ , independente do tempo. Assumimos que o espectro de  $H$  é discreto e não degenerado (mesmo sem conhecer a solução):

$$H |n\rangle = E_n |n\rangle, \quad n=0, 1, 2, \dots (\infty)$$

Seja agora  $|\psi\rangle$  um ket arbitrário no espaço dos estados do sistema. Temos o:

Teorema. O valor médio de  $H$  com  $|\psi\rangle$  é tal que satisfaz a desigualdade

$$\langle H \rangle = \frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle} \geq E_0,$$

onde  $E_0$  é o valor da energia para o estado fundamental "verdadeiro" de  $H$ .

Dem.

Mesmo não conhecendo a base  $\{|n\rangle\}$ , mas assumindo que existe, desenvolvemos o ket dado  $|\psi\rangle$  nesta base:

$$|\psi\rangle = \sum_n \langle n | \psi \rangle |n\rangle = \sum_n C_n |n\rangle$$

Assim temos:

$$\langle \psi | \mathcal{H} | \psi \rangle = \sum_n E_n |c_n|^2 \geq E_0 \sum_n |c_n|^2$$

e  $\langle \psi | \psi \rangle = \sum_n |c_n|^2$ , normalização. Obtemos

$$\frac{\langle \psi | \mathcal{H} | \psi \rangle}{\langle \psi | \psi \rangle} \geq E_0 \quad \text{c.q.d.}$$

Caso particular:  $c_n \equiv 0$ , exceto  $c_0 \neq 0 \Rightarrow |\psi\rangle$  coincide com o estado fundamental e vale a igualdade.

Este teorema é a base de um método aproximado para determinar  $E_0$ . Escolhemos (usando critérios físicos), uma família de kets  $|\psi(\alpha)\rangle$  dependentes de um conjunto de parâmetros que simbolizamos por  $\{\alpha\}$ . Calculamos o valor médio  $\langle \mathcal{H} \rangle(\alpha)$  do Hamiltoniano para estes estados; depois minimizamos  $\langle \mathcal{H} \rangle(\alpha)$  em relação aos parâmetros  $\{\alpha\}$ ; o valor mínimo obtido  $\langle \mathcal{H} \rangle_{\min}$  é a melhor aproximação do nível fundamental  $E_0$ .

$\{|\psi(\alpha)\rangle\}$  : estados de prova

Este resultado é consequência de um teorema geral:

### Teorema de Ritz

O valor médio do Hamiltoniano  $H$  é estacionário na vizinhança de seus autovalores discretos (estacionário aqui significa de variação nula em 1ª ordem)

Dem.

$$\langle H \rangle = \frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle},$$

ou

$$\langle H \rangle \cdot \langle \psi | \psi \rangle = \langle \psi | H | \psi \rangle.$$

Neste caso,  $\langle H \rangle$  é um funcional do ket  $|\psi\rangle$ . Por exemplo, usando a representação de coordenadas, temos

$$\langle H \rangle = \langle \psi | H | \psi \rangle = \int d^3x \psi^*(\vec{x}) H \psi(\vec{x}).$$

Queremos calcular a variação  $\delta \langle H \rangle$  do valor médio quando variamos o ket  $|\psi\rangle \rightarrow |\psi\rangle + |\delta\psi\rangle$ , onde assumimos que  $|\delta\psi\rangle$  é infinitesimal. Tomamos variações na expressão acima:

$$\begin{aligned} \langle \psi | \psi \rangle \delta \langle H \rangle + \langle H \rangle (\langle \psi | \delta\psi \rangle + \langle \delta\psi | \psi \rangle) &= \\ &= \langle \psi | H | \delta\psi \rangle + \langle \delta\psi | H | \psi \rangle. \end{aligned}$$

Por exemplo, se  $|\psi\rangle$  depende de um parâmetro,  $|\psi(\alpha)\rangle$ , temos

$$|\delta\psi\rangle = \delta\alpha \left| \frac{d\psi}{d\alpha}(\alpha) \right\rangle$$

O valor médio  $\langle \mathcal{H} \rangle$  é um número. Assim:

$$\langle \psi | \psi \rangle \delta \langle \mathcal{H} \rangle = \langle \psi | (\mathcal{H} - \langle \mathcal{H} \rangle) | \delta \psi \rangle + \langle \delta \psi | (\mathcal{H} - \langle \mathcal{H} \rangle) | \psi \rangle$$

O valor médio é estacionário  $\Leftrightarrow \delta \langle \mathcal{H} \rangle = 0$ , implicando

$$\langle \psi | (\mathcal{H} - \langle \mathcal{H} \rangle) | \delta \psi \rangle + \langle \delta \psi | (\mathcal{H} - \langle \mathcal{H} \rangle) | \psi \rangle = 0,$$

onde a variação  $|\delta \psi\rangle$  é arbitrária. Definindo

$$|\varphi\rangle \equiv (\mathcal{H} - \langle \mathcal{H} \rangle) |\psi\rangle,$$

a condição acima fica:

$$\langle \varphi | \delta \psi \rangle + \langle \delta \psi | \varphi \rangle = 0,$$

porque  $\mathcal{H}^\dagger = \mathcal{H}$  (hermiteano).  $|\delta \psi\rangle$  é arbitrário e em particular pode ser escolhido como sendo

$$|\delta \psi\rangle = \delta \alpha |\varphi\rangle, \text{ com } \delta \alpha \text{ real}$$

$$0 = \delta \alpha \langle \varphi | \varphi \rangle + \delta \alpha \langle \varphi | \varphi \rangle = 2 \delta \alpha \langle \varphi | \varphi \rangle$$

$$\Rightarrow |\varphi\rangle = (\mathcal{H} - \langle \mathcal{H} \rangle) |\psi\rangle = 0, \text{ ou seja}$$

$$\hat{H}|\psi\rangle = \langle \hat{H} \rangle |\psi\rangle$$

Resultado:  $\delta\langle \hat{H} \rangle = 0 \iff |\psi\rangle$  é autoket de  $\hat{H}$ , e o valor estacionário  $\langle \hat{H} \rangle_0$  corresponde ao autovvalor do Hamiltoniano.

Consequência: Se o funcional  $\langle \hat{H} \rangle(\alpha)$  obtido a partir de "estados de prova"  $|\psi(\alpha)\rangle$  apresenta várias extremos, estes fornecem valores aproximados de algumas das energias  $E_n$  de  $\hat{H}$ .

O método variacional permite estabelecer um limite superior para o estado fundamental. Não estabelece porém um correspondente limite inferior.

Existe um método para estimar limites superior e inferior dos autovvalores. Queremos resolver aproximadamente a equação:

$$\hat{H}|m\rangle = E_m|m\rangle,$$

e seja  $|\psi\rangle$  um estado de prova que assumimos normalizado  $\langle \psi|\psi\rangle = 1$ . Nossa aproximação para o autovvalor é

$$\Lambda \equiv \langle \psi|\hat{H}|\psi\rangle$$

Para estimar quão acurado é este valor, definimos um bet de erro:

$$|R\rangle \equiv (\mathcal{H} - \Lambda)|\psi\rangle.$$

Temos:

$$\begin{aligned} \langle R|R\rangle &= \sum_m \langle \psi | (\mathcal{H} - \Lambda) | m \rangle \langle m | (\mathcal{H} - \Lambda) | \psi \rangle \\ &= \sum_m (E_m - \Lambda)^2 |\langle \psi | m \rangle|^2 \\ &= \sum_m |c_m|^2 (E_m - \Lambda)^2. \end{aligned}$$

Seja  $E_R$  o autovalor mais perto de  $\Lambda$ . Temos:

$$\begin{aligned} \langle R|R\rangle &= \sum_m |c_m|^2 (E_m - \Lambda)^2 \geq \sum_m |c_m|^2 (E_R - \Lambda)^2 \\ &= (E_R - \Lambda)^2 \sum |c_m|^2 = (E_R - \Lambda)^2 \langle \psi | \psi \rangle \\ &= (E_R - \Lambda)^2. \end{aligned}$$

Seja então  $\Delta \equiv \sqrt{\langle R|R\rangle}$ , ou  $\Delta^2 = \langle R|R\rangle > 0$ ,

$$(E_R - \Lambda)^2 \leq \Delta^2,$$

$\Rightarrow$

$$\Lambda - \Delta \leq E_R \leq \Lambda + \Delta$$

Este método só funciona se sabemos que  $E_R$  é o autovalor mais próximo a  $\Lambda$ .

Assumamos um conjunto l.i. de kets  $|\psi_i\rangle$ ,  $i=1,2,\dots,N$   
 Formamos um ket variacional por uma combinação linear:

$$|\varphi\rangle = \sum_{i=1}^N c_i |\psi_i\rangle,$$

onde consideramos os  $\{c_i\}$  como um conjunto de parâmetros variacionais, que assumimos como sendo reais.

$$\text{Seja: } \Delta_{ij} \equiv \langle \psi_i | \psi_j \rangle = \langle \psi_j | \psi_i \rangle = \Delta_{ji}^*$$

$$H_{ij} \equiv \langle \psi_i | H | \psi_j \rangle = H_{ji}^*$$

$$\langle H \rangle = \frac{\langle \varphi | H | \varphi \rangle}{\langle \varphi | \varphi \rangle} = \frac{\sum_{n,m}^N c_n c_m H_{nm}}{\sum_{n,m}^N \Delta_{nm} c_n c_m} \equiv E$$

ou

$$E \sum_{ij} \Delta_{ij} c_i c_j = \sum_{ij} c_i c_j H_{ij}$$

Condição estacionária:

$$\frac{\delta E}{\delta c_k} = 0, \quad k=1,2,\dots,N$$

$$\begin{aligned} \partial_k E \sum_{(i,j)} c_i c_j \Delta_{ij} + E \sum_{ij} (\delta_{ik} c_j + \delta_{jk} c_i) \Delta_{ij} \\ = \sum_{ij} (\delta_{ik} c_j + \delta_{jk} c_i) \mathcal{H}_{ij} \end{aligned}$$

com  $\partial_k E = 0$  temos

$$\begin{aligned} E \left( \sum_j c_j \Delta_{kj} + \sum_i c_i \Delta_{ik} \right) \\ = \sum_j c_j \mathcal{H}_{kj} + \sum_i c_i \mathcal{H}_{ik} \end{aligned}$$

$$E \sum_j c_j (\Delta_{kj} + \Delta_{kj}^*) = \sum_j c_j (\mathcal{H}_{kj} + \mathcal{H}_{kj}^*)$$

com termos todos reais temos:

$$\sum_j c_j (\mathcal{H}_{jk} - E \Delta_{jk}) = 0$$

que é um conjunto de eqs. lineares, homogêneas para os coeficientes  $\{c_j\}$ . Para termos soluções não triviais, o determinante do sistema deve ser nulo

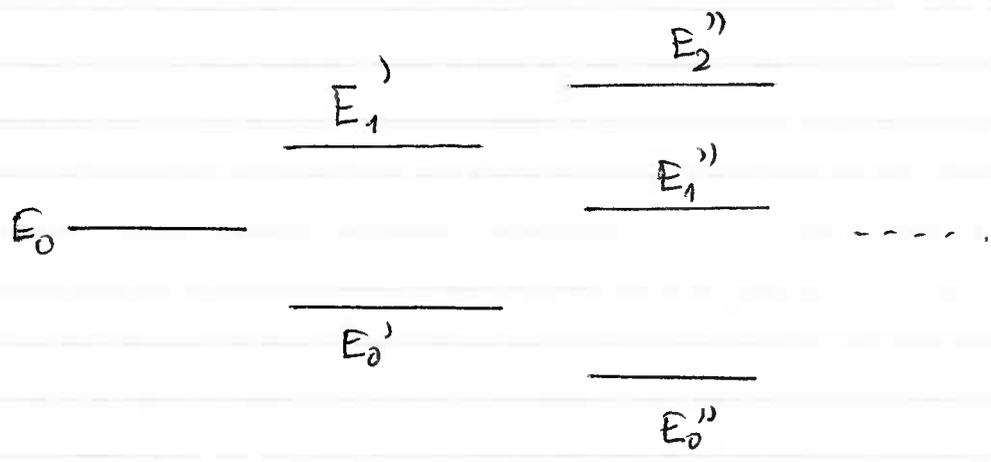
$\Rightarrow$  obtemos equações seculares para a energia  $E$

$$\begin{vmatrix}
 H_{11} - E \Delta_{11} & H_{12} - \Delta_{12} E & \dots & H_{1N} - \Delta_{1N} E \\
 \vdots & H_{22} - E \Delta_{22} & & \\
 \vdots & & \ddots & \\
 \vdots & & & H_{NN} - E \Delta_{NN}
 \end{vmatrix} = 0$$

A menor raiz desta equação  $E_{min}$  é um limite superior para a energia  $E_0$ . As outras raízes, pode ser mostrado, são limites superiores para outras energias do espectro.

Propriedade importante:

para cada etapa sucessiva, em que acrescentamos um novo ket, as raízes se separam



com  $E_0' \leq E_1'' \leq E_1'$  ,

ou em geral  $E_n' \leq E_n'' \leq E_{n+1}'$

► Exemplo: a) Outra vez o Oscilador Harmônico 1-dim

$$\hat{H} = -\frac{\hbar^2}{2m} \partial_x^2 + \frac{1}{2} m \omega^2 x^2 = T + V(x)$$

Como o potencial  $V(x) = \frac{1}{2} m \omega^2 x^2$  é par, procuramos uma função de prova par para aproximar o estado fundamental.

Escolhemos uma família de 1-parâmetros:

$$\psi_\alpha(x) = e^{-\alpha x^2} \quad ; \quad \alpha > 0, \alpha \in \mathbb{R} \\ = \langle x | \psi_\alpha \rangle$$

Temos:

$$\langle \psi_\alpha | \psi_\alpha \rangle = \int_{-\infty}^{\infty} dx e^{-2\alpha x^2} = C^2 > 0. \quad \text{Também}$$

temos:

$$\begin{aligned} \langle T \rangle_\alpha &= \int_{-\infty}^{\infty} dx e^{-\alpha x^2} \left( -\frac{\hbar^2}{2m} \right) \left( \partial_x^2 e^{-\alpha x^2} \right) \\ &= \int_{-\infty}^{\infty} dx \left( -\frac{\hbar^2}{2m} \right) \partial_x \left( -2\alpha x e^{-\alpha x^2} \right) e^{-\alpha x^2} \\ &= \int_{-\infty}^{\infty} dx e^{-\alpha x^2} \left( -\frac{\hbar^2}{2m} \right) \left( -2\alpha e^{-\alpha x^2} + 4\alpha^2 x^2 e^{-\alpha x^2} \right) \\ &= \frac{\hbar^2}{2m} (2\alpha) C^2 - \frac{\hbar^2}{2m} 4\alpha^2 \int_{-\infty}^{\infty} dx e^{-2\alpha x^2} x^2 \end{aligned}$$

$$\langle V \rangle_\alpha = \frac{1}{2} m \omega^2 \int_{-\infty}^{\infty} dx e^{-2\alpha x^2} x^2 \equiv \frac{1}{2} m \omega^2 D^2$$

Temos os resultados:

$$C^2 = \frac{\sqrt{\pi}}{\sqrt{2\alpha}}, \quad D^2 = \frac{1}{2} \sqrt{\frac{\pi}{(2\alpha)^3}} = \frac{1}{2} \frac{C^2}{2\alpha}$$

Temos:

$$\langle H \rangle(\alpha) = \frac{\left( \frac{\hbar^2 \alpha}{m} - \frac{\hbar^2}{2m} 4\alpha^2 \frac{1}{4\alpha} + \frac{1}{2} m \omega^2 \frac{1}{4\alpha} \right) C^2}{C^2}$$

$$= \frac{\hbar^2 \alpha}{2m} + \frac{m \omega^2}{8\alpha}$$

Variando o parâmetro para um extremo:

$$0 = \frac{d}{d\alpha} \langle H \rangle(\alpha) = \frac{\hbar^2}{2m} - \frac{m \omega^2}{8\alpha^2} \Rightarrow \frac{4\alpha^2}{m \omega^2} = \frac{m}{\hbar^2}$$

$$\alpha = \frac{m \omega}{2\hbar}$$

$$\sqrt{\frac{\hbar}{m \omega}} \equiv x_0, \quad \alpha = \frac{1}{2x_0^2}$$

Obtemos a solução exata para o estado fundamental

$$\psi_0(x) = e^{-\frac{x^2}{2x_0^2}}$$

com valor estacionário:

$$\begin{aligned} \langle H \rangle_0 &= \frac{\hbar^2}{2m} \frac{m\omega}{2\hbar} + \frac{m\omega^2}{8} \frac{2\hbar}{m\omega} \\ &= \frac{\hbar\omega}{4} + \frac{\hbar\omega}{4} = \frac{\hbar\omega}{2} = E_0 \end{aligned}$$

b) Pôço unidimensional infinito

Estimar o estado fundamental do potencial

$$V(x) = \begin{cases} 0, & \text{para } |x| < a \\ \infty, & \text{para } |x| > a \end{cases}$$

A solução é conhecida  $\langle x|0 \rangle = \psi_0(x) = \frac{1}{\sqrt{a}} \cos\left(\frac{\pi x}{2a}\right)$ ,

$$\text{com } E_0 = \frac{\hbar^2}{2m} \left(\frac{\pi^2}{4a^2}\right).$$

Suponhamos que queremos estimar esta energia variacionalmente. Sabemos que a função de onda deve ser nula em  $x = \pm a$  (não pode penetrar a barreira infinita). Esperamos que seja par e não tenha nó entre  $x = \pm a$ . Usamos a função de prova:

$$\psi_0(x) = a^2 - x^2$$

Calculamos a energia:

$$\langle \overline{H} \rangle = \frac{\left(-\frac{\hbar^2}{2m}\right) \int_{-a}^a dx (a^2 - x^2) \partial_x^2 (a^2 - x^2)}{\int_{-a}^a dx (a^2 - x^2)^2}$$

$$\partial_x^2 (a^2 - x^2) = \partial_x (-2x) = -2$$

$$\int_{-a}^a dx (a^2 - x^2) = 2 \int_0^a dx (a^2 - x^2) = 2 \left( a^2 x - \frac{1}{3} x^3 \right) \Big|_0^a = \frac{4}{3} a^3$$

$$\int_{-a}^a dx (a^2 - x^2)^2 = 2 \int_0^a dx (a^4 - 2a^2 x^2 + x^4) = 2 \left( a^5 - \frac{2}{3} a^2 x^3 + \frac{1}{5} x^5 \right) \Big|_0^a =$$

$$= 2a^5 \left( 1 - \frac{2}{3} + \frac{1}{5} \right) = \frac{16}{15} a^5,$$

logo:

$$\langle \overline{H} \rangle = \frac{10}{4} \left( \frac{\hbar^2}{2ma^2} \right) = \frac{10}{\pi^2} \left( \frac{\hbar^2 \pi^2}{8ma^2} \right) = \frac{10}{\pi^2} E_0$$

$$\approx 1.0132 E_0$$

Erro relativo desta estimativa: 1.32%

Calcular agora quão bem a função de prova simula o estado fundamental. Escrevamos:

$$\psi_0(x) = a^2 - x^2 = \langle x | 0 \rangle$$

para a função de prova. Queremos estimar  $\langle 0 | 0 \rangle$ ,

isto é a projeção de  $|\bar{0}\rangle$  sobre o verdadeiro estado fundamental  $|0\rangle$ . É conhecido que o espectro deste problema é dado por

$$E_n = \frac{\hbar^2 \pi^2}{8ma^2} n^2, \quad n=1,2,3,\dots$$

Assumindo que o estado de prova está normalizado,  $\langle \bar{0} | \bar{0} \rangle = 1$ , temos:

$$\begin{aligned} \langle \bar{H} \rangle &= \sum_k |\langle k | \bar{0} \rangle|^2 E_k \\ &\geq |\langle 0 | \bar{0} \rangle|^2 E_0 + E_2 (1 - |\langle 0 | \bar{0} \rangle|^2), \end{aligned}$$

porque sendo  $|0\rangle$  uma função par,  $\langle 1 | \bar{0} \rangle = 0$ .

Assim usamos o nível seguinte com  $E_2 = 9E_0$ .

Escrevemos

$$|\langle 0 | \bar{0} \rangle|^2 = c_0^2, \quad \text{supondo } c_0 \text{ real.}$$

$$1.0132 E_0 \geq c_0^2 E_0 + 9 E_0 (1 - c_0^2)$$

$$\geq -8 c_0^2 E_0 + 9 E_0$$

$$-1.0132 E_0 \leq 8 c_0^2 E_0 - 9 E_0$$

$$8 c_0^2 \geq 9 - 1.0132 \Rightarrow c_0^2 \geq \frac{9 - 1.0132}{8} \approx 0.99835$$

## Resumo

Faz-se uma revisão completa das propriedades e usos do bem conhecido método de Lanczos, dando-se uma visão mais física do processo. Compara-se o método com a teoria de perturbações em mecânica quântica. Para isso a função de prova que inicializa o processo de Lanczos é escolhida como sendo uma autofunção do Hamiltoniano não perturbado, com a idéia de compará-lo com a teoria de perturbações ordinária "a la Rayleigh-Schrödinger". Esta escolha reduz a eficiência potencial do método, porém este converge mais rapidamente que a série de perturbações e funciona inclusive em situações onde esta série diverge.

# Capítulo 1

## O Método de Lanczos

Em 1950 C. Lanczos [1] propôs um método iterativo, através do qual reduzimos matrizes simétricas reais à forma tridiagonal. Em seu trabalho, Lanczos aplica este método a operadores diferenciais e integrais, mostrando uma grande eficiência para calcular alguns autovalores e autovetores do espectro através da diagonalização de pequenas matrizes que são obtidas a cada passo do método.

Este fato fez com que tradicionalmente o método fosse visto como um dispositivo para diagonalizar matrizes de grande porte (o que até certo ponto é acidental; como veremos), apresentando grande confiabilidade e estabilidade no cálculo numérico.

De uma maneira resumida o método funciona a partir de uma matriz simétrica real  $\mathbf{H}$  de ordem  $(n \times n)$  e um vetor inicial  $|v_1\rangle$ .

$$\mathbf{H} |v_1\rangle = \alpha_1 |v_1\rangle + \beta_1 |v_2\rangle \quad (1.1)$$

$$\mathbf{H} |v_2\rangle = \beta_1 |v_1\rangle + \alpha_2 |v_2\rangle + \beta_2 |v_3\rangle \quad (1.2)$$

$$\mathbf{H} |v_3\rangle = \beta_2 |v_2\rangle + \alpha_3 |v_3\rangle + \beta_3 |v_4\rangle \quad (1.3)$$

$$\vdots$$

$$\mathbf{H} |v_n\rangle = \beta_{n-1} |v_{n-1}\rangle + \alpha_n |v_n\rangle + \beta_n |v_{n+1}\rangle \quad (1.4)$$

Cada vetor  $|v_{n+1}\rangle$  é obtido ortogonalizando-se  $\mathbf{H} |v_n\rangle$  com respeito a  $|v_n\rangle$  e  $|v_{n-1}\rangle$ . Na base dos  $|v_n\rangle$ 's, então vemos que

$$\mathbf{H} \rightarrow \begin{bmatrix} \alpha_1 & \beta_1 & 0 & 0 & \cdots \\ \beta_1 & \alpha_2 & \beta_2 & 0 & \cdots \\ 0 & \beta_2 & \alpha_3 & \beta_3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad (1.5)$$

já é escrita diretamente numa forma tridiagonal. Este outro fato levou muitos pesquisadores, durante muito tempo, a condenar o método, pois do ponto de vista numérico este mecanismo era superado por outros dois métodos matematicamente equivalentes, de Givens e Householder [2].

Só recentemente, estudiosos de vários campos da física (e de outras áreas) redescobriram o método e mostraram ser ele de grande utilidade, com resultados que têm surpreendido a todos.

Apesar de nem sempre a equivalência matemática ser a mesma, este método tem sido aplicado em física nuclear [3], física estatística [4], em física do estado sólido [5] como meio de calcular densidade de estados; além de sua extensão de aplicabilidade à química.

Nas próximas seções revisitaremos o método de Lanczos, de uma maneira um tanto quanto diferente de como ele vem sendo apresentado na literatura, procurando

de uma forma simplificada e geral ilustrar uma série de ligações interessantes e muito importantes em matemática e principalmente em física. A intenção desta parte do trabalho é de servir como uma espécie de manual de uso e extensões do método para as várias áreas onde ele pode (ou poderá) ser utilizado.

Deixamos de lado, até certo ponto, a análise numérica do método, considerando que isto já está feito de uma forma completa na literatura [6], e por fugir dos objetivos deste trabalho. Começamos pela descrição matemática do modelo, seus atrativos computacionais; bem como sua ligação com o método dos momentos e portanto à teoria dos polinômios ortogonais e frações contínuas. Do ponto de vista de soluções recursivas da equação de Schrödinger, veremos as várias faces do método; desde variacional até perturbativa.

Vale a pena frisar, por fim, que este capítulo, apesar de ser uma revisão do método de Lanczos, contém não só uma análise matemática completa do método, bem como uma aproximação maior com a sua significação física, sempre de uma maneira simples e concisa o que, como já foi dito, não pode ser encontrado na literatura existente.

## 1.1 Descrição do Método

Para construirmos um algoritmo para o processo de Lanczos, primeiramente trabalharemos com matrizes hermiteanas ao invés de matrizes simétricas, especificamente consideradas por Lanczos em seu trabalho original. Futuramente estenderemos os resultados aqui obtidos à matrizes não-hermiteanas com autovalores reais e

complexos. Além dessa pequena diferença em relação ao trabalho de Lanczos, nós o adaptamos ao simbolismo da Mecânica Quântica para indicar vetores e multiplicação de matrizes por vetores.

Portanto, seguindo Lanczos, consideramos uma matriz hermiteana  $\mathbf{H}$  e operamos com ela sobre um vetor inicial  $|v_1\rangle$ . O que nós queremos estabelecer é uma combinação linear do vetor iterado  $\mathbf{H}|v_1\rangle$  com  $|v_1\rangle$ . Nós não podemos esperar, entretanto, que esta combinação linear esteja inicialmente correta, porém podemos aproximar tal combinação de maneira a fazer com que o módulo do novo vetor seja tão pequeno quanto possível. Logo, nós queremos escolher como novo vetor  $|v_2\rangle$  a combinação:

$$|v_2\rangle = \mathbf{H}|v_1\rangle - \alpha_1 |v_1\rangle \quad (1.6)$$

onde, tomamos inicialmente  $\alpha_1$  como um número complexo (por causa da generalidade de  $|v_1\rangle$ ).

Explorando a hermiticidade de  $\mathbf{H}$ , nós calculamos:

$$\langle v_2|v_2\rangle = \langle v_1|\mathbf{H}^2|v_1\rangle - 2\alpha_1 \langle v_1|\mathbf{H}|v_1\rangle + \alpha_1^2 \langle v_1|v_1\rangle. \quad (1.7)$$

A minimização de  $\langle v_2|v_2\rangle$  com respeito ao parâmetro variacional  $\alpha_1$  nos leva a manipular a eq.1.7 fazendo-se

$$0 = \frac{\partial \langle v_2|v_2\rangle}{\partial \alpha_1} = -2 \langle v_1|\mathbf{H}|v_1\rangle + 2\alpha_1 \langle v_1|v_1\rangle \quad (1.8)$$

e portanto,

$$\alpha_1 = \frac{\langle v_1|\mathbf{H}|v_1\rangle}{\langle v_1|v_1\rangle}, \quad (1.9)$$

que é real ( $\mathbf{H}$  é hermiteana) independentemente do campo em que  $|v_1\rangle$  é inicialmente definido.

Além disso, é fácil ver que o novo vetor  $|v_2\rangle$  definido na eq. 1.6 é ortogonal ao vetor original  $|v_1\rangle$ , por causa da eq. 1.9.

Nós agora continuamos o nosso processo. De  $|v_2\rangle$  vamos obter uma combinação linear do vetor iterado  $\mathbf{H}|v_2\rangle$  com  $|v_2\rangle$  e  $|v_1\rangle$ ,

$$|v_3\rangle = \mathbf{H}|v_2\rangle - \alpha_2|v_2\rangle - \beta_1|v_1\rangle, \quad (1.10)$$

com a condição de que a amplitude do novo vetor torne-se tão pequena quanto possível. Agora  $\alpha_2$  e  $\beta_1$  são os novos parâmetros variacionais. Desde que  $\mathbf{H}$  é hermiteana, nós podemos provar que  $\alpha_2$  e  $\beta_1$  são números reais, como verificado anteriormente para  $\alpha_1$ .

A amplitude do vetor  $|v_3\rangle$  é

$$\langle v_3|v_3\rangle = \langle v_2|\mathbf{H}^2|v_2\rangle - 2\alpha_2\langle v_2|\mathbf{H}|v_2\rangle - 2\beta_1\langle v_2|\mathbf{H}|v_1\rangle + \alpha_2^2\langle v_2|v_2\rangle + \beta_1^2\langle v_1|v_1\rangle. \quad (1.11)$$

Da condição de mínimo, estabelecemos que

$$\frac{\partial\langle v_3|v_3\rangle}{\partial\alpha_2} = 0 \quad \text{e} \quad \frac{\partial\langle v_3|v_3\rangle}{\partial\beta_1} = 0,$$

obtendo-se respectivamente,

$$\alpha_2 = \frac{\langle v_2|\mathbf{H}|v_2\rangle}{\langle v_2|v_2\rangle} \quad (1.12)$$

$$\beta_1 = \frac{\langle v_1|\mathbf{H}|v_2\rangle}{\langle v_1|v_1\rangle} = \frac{\langle v_2|v_2\rangle}{\langle v_1|v_1\rangle}. \quad (1.13)$$

O processo pode ser continuado e nós devemos pensar, por analogia, que para o próximo passo a melhor combinação é do tipo

$$|v_4\rangle = \mathbf{H}|v_3\rangle - \alpha_3|v_3\rangle - \beta_2|v_2\rangle - \gamma|v_1\rangle. \quad (1.14)$$

Mas nós podemos mostrar que, na eq.1.14,  $\gamma = 0$ ; realmente, da condição de mínimo para o comprimento de  $|v_4\rangle$ , fazemos

$$0 = \frac{\partial \langle v_4|v_4\rangle}{\partial \gamma} = -\langle v_1|\mathbf{H}|v_3\rangle - \langle v_3|v_1\rangle + 2\gamma \langle v_1|v_1\rangle \quad (1.15)$$

ou,

$$\gamma = \frac{\langle v_1|\mathbf{H}|v_3\rangle + \langle v_3|v_1\rangle}{2\langle v_1|v_1\rangle}, \quad (1.16)$$

o que do fato de  $|v_3\rangle$  ser ortogonal aos vetores prévios, nós facilmente mostramos na eq.1.16 que  $\gamma$  vai a zero. Do discutido fica claro ainda, que a construção da base para Lanczos é mais sutil comparada à ortogonalização de Schimdt, a qual usa todos os vetores em cada estágio.

Portanto, como já havia sido notado por Lanczos, “a característica mais marcante deste processo de minimizações sucessivas é que a melhor combinação linear nunca inclui mais que três termos...” [1]. Ou seja, todo novo passo do processo de minimização requer somente dois termos de correção.

Nós podemos sumarizar o método de Lanczos escrevendo a hierarquia da melhor combinação linear, começando de  $|v_1\rangle$ .

$$|v_{n+1}\rangle = \mathbf{H}|v_n\rangle - \alpha_n|v_n\rangle - \beta_{n-1}|v_{n-1}\rangle, \quad (1.17)$$

para  $n \geq 1; |v_0\rangle = 0$  e,

$$\alpha_n = \frac{\langle v_n | \mathbf{H} | v_n \rangle}{\langle v_n | v_n \rangle}, \quad n = 1, 2, 3, \dots \quad (1.18)$$

$$\beta_n = \frac{\langle v_n | \mathbf{H} | v_{n+1} \rangle}{\langle v_n | v_n \rangle} = \frac{\langle v_{n+1} | v_{n+1} \rangle}{\langle v_n | v_n \rangle} n = 1, 2, 3, \dots \quad (1.19)$$

Se o conjunto de vetores ortogonais  $\{|v_1\rangle, |v_2\rangle, \dots\}$  é então normalizado e usado como uma base para a matriz  $\mathbf{H}$ , nós obtemos sua forma tridiagonal (eq.1.5) com elementos diagonais  $\{\alpha_1, \alpha_2, \dots\}$  e os elementos fora da diagonal  $\{\beta_1, \beta_2, \dots\}$ .

O processo terminará automaticamente quando o conjunto de vetores  $\{|v_1\rangle, \dots\}$  expandirem o espaço inteiro de  $\mathbf{H}$  (se ela for finita). Isto ocorrerá quando o novo vetor tiver comprimento zero. O término do processo de Lanczos, significa que o menor subespaço invariante contendo o vetor inicial é dimensionalmente finito e expandido pelos vetores obtidos no processo. Vale frisar ainda, que cada um destes vetores tem a dimensão da matriz  $\mathbf{H}$ . Deve ser notado aqui também que se  $|v_1\rangle$  é ortogonal a qualquer autovetor de  $\mathbf{H}$ , então a série de operações descritas nas equações 1.17 a 1.19 termina depois de  $m$  passos, isto é, para algum  $m < n$ ,  $\beta_n = 0$ . Um caso importante para o qual isto pode acontecer é quando  $\mathbf{H}$  tem algum autovalor degenerado. Neste evento um novo vetor pode ser escolhido ortogonal aos  $|v_i\rangle$ 's já obtidos, e as operações indicadas pelas equações 1.17 a 1.19 são continuadas e/ou repetidas (cf. seção 1.5).

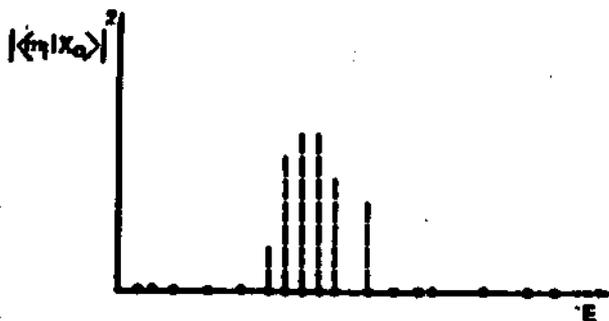
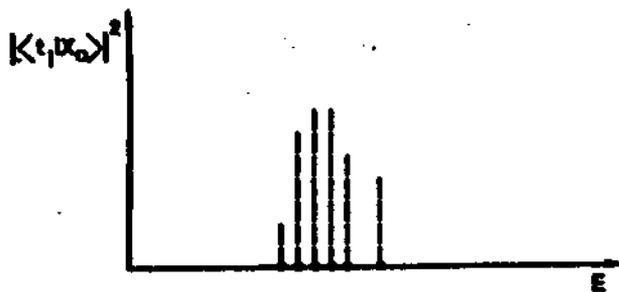
(a) diagonalização de  $M$ (b) diagonalização de  $T$ 

Figura 1: Os autovetores de  $M$  (aqui suposta hermiteana) são indicados por  $|m_i\rangle$ , e os autovalores são reportados no eixo real  $E$ . Os autovetores da matriz tridiagonal gerados por  $|x_0\rangle$  são denotados por  $|t_i\rangle$ ; os autovalores selecionados pelo estado teste são também indicados por conveniência.

A característica mais atraente do método de Lanczos é não precisarmos completar o processo para termos precisão nas estimativas de alguns dos autovalores extremos de  $\mathbf{H}$ , ou seja, podemos “truncar” a cada passo o processo e então diagonalizar as pequenas matrizes tridiagonais obtidas; os autovalores destas serão excelentes aproximações de alguns autovalores de  $\mathbf{H}$ , mesmo para poucos passos do processo (cf. apêndice I e seção 1.2).

Na figura 1 da página anterior fazemos um quadro esquemático para o método de Lanczos, onde partimos de um vetor de teste inicial  $|x_0\rangle$  e uma matriz  $\mathbf{M}$  grande. Depois da tridiagonalização de  $\mathbf{M}$  obtemos pequenas matrizes tridiagonais  $\mathbf{T}$ .

Uma outra característica do método, um tanto óbvia a esta altura, é que  $\mathbf{H}$  não necessita ser uma matriz, ela pode ser um operador abstrato (o hamiltoniano) e os  $|v\rangle$ 's podem ser combinações lineares de vetores abstratos (os estados nos quais o hamiltoniano atua). Para formar  $\mathbf{H}|v_i\rangle$  nós usamos as regras para operar com  $|v_i\rangle$  ao invés da multiplicação ordinária de matrizes (cf. seção 1.3 e 1.4).

Em geral os autovalores que aparecem e convergem melhor, são aqueles com o maior “overlap” com o estado inicial. Se por exemplo, sabe-se que um vetor  $|v_1\rangle$  envolve os mais baixos autovalores de  $\mathbf{H}$ , nós podemos escolhê-lo como vetor inicial e determinar com certa confiança o mais baixo autovalor de  $\mathbf{H}$ .

Para finalizar esta seção, mostramos o procedimento a ser seguido em Lanczos para quando  $\mathbf{H}$  é não hermiteana. A diferença essencial é que nós devemos agora operar com  $\mathbf{H}$  e  $\mathbf{H}^\dagger$ . Nós temos então que considerar o conjunto de vetores biortog-

nais,

$$\{|v_1\rangle, |v_2\rangle, \dots\} \text{ e } \{|\tilde{v}_1\rangle, |\tilde{v}_2\rangle, \dots\}, \quad (1.20)$$

cuja definição é a generalização de 1.17, ou seja,

$$\begin{aligned} |v_1\rangle &= |\tilde{v}_1\rangle \\ |v_2\rangle &= \mathbf{H}|v_1\rangle - \alpha_1|v_1\rangle \\ |\tilde{v}_2\rangle &= \mathbf{H}^\dagger|\tilde{v}_1\rangle - \alpha_1^*|\tilde{v}_1\rangle \\ |v_{n+1}\rangle &= \mathbf{H}|v_n\rangle - \alpha_n|v_n\rangle - \beta_{n-1}|v_{n-1}\rangle \\ |v_{n+1}^\sim\rangle &= \mathbf{H}^\dagger|\tilde{v}_n\rangle - \alpha_n^*|v_n\rangle - \beta_{n-1}^*|v_{n-1}^\sim\rangle, \end{aligned} \quad (1.21)$$

com a prescrição de que ao se fazer produtos escalares entre vetores opostos, não consideramos a amplitude  $\langle \tilde{v}_n | v_n \rangle$  e procuramos por extremos ( no apêndice I fazemos a transformação numérica do método para matrizes reais e simétricas, mas a partir do discutido anteriormente fica fácil a extensão do método.).

Outros refinamentos de técnica original de Lanczos vêm sendo amplamente usados em vários ramos da física, tais como cálculos em modelos de camadas na física nuclear [8], determinação da densidade de estados em matéria condensada [5], solução de modelos em teoria de campos [11], além de outras aplicações importantes em diversas áreas da computação científica [12].

## 1.2 Convergência, Rapidez e Estocagem

Nesta seção discutimos algumas características importantes do Método de Lanczos do ponto de vista computacional.

Como já foi discutido na seção anterior, a característica mais interessante do método é a de conseguirmos boa precisão nas estimativas de alguns autovalores do espectro de  $\mathbf{H}$ . Para entendermos um pouco melhor esta propriedade do método (além de podermos otimizá-la em alguns problemas), vamos definir a truncagem da base de  $N$  vetores da matriz tridiagonal  $N \times N$  (eq.1.5), que representa  $\mathbf{H}$  na base finita  $\{|v_1\rangle, |v_2\rangle, \dots, |v_n\rangle\}$ . Para isto tomamos esta base dos  $n$  vetores obtidos por meio de Lanczos. A matriz  $\mathbf{H}$  em nossa base truncada é obviamente representada por  $\mathbf{H}^{(n)}$ , a matriz  $n \times n$  no canto superior de  $\mathbf{H}$ .

O determinante  $\det |\lambda \mathbf{I} - \mathbf{H}^{(n)}|$ , o qual produz o polinômio característico pertencente a  $\mathbf{H}^{(n)}$ , possui propriedades [4] que nos garantem que os autovalores de  $\mathbf{H}$  em nosso espaço truncado têm propriedades variacionais. Isto é, nós podemos cortar arbitrariamente a sequência de operações definidas pelas eqs.1.17 e 1.18, depois de obter  $n$  vetores  $|v_i\rangle$ , gerando  $\mathbf{H}^{(n)}$ . Além disso, ficamos seguros de que os autovalores de  $\mathbf{H}^{(n)}$ , tomados em ordem ascendente, são maiores que o correspondente autovalor exato de  $\mathbf{H}$  no espaço completo. Ou seja, se denotarmos os autovalores de  $\mathbf{H}^{(m)}$ , em ordem ascendente, por  $\eta_1^{(m)}, \eta_2^{(m)}, \dots, \eta_m^{(m)}$ ,

$$\eta_i^{(m)} > \eta_i^{(m+1)} > \dots > \lambda_i, \quad (1.22)$$

com  $\lambda_i$  sendo o  $i$ -ésimo autovalor de  $\mathbf{H}$  tomado em ordem ascendente. A possibilidade, portanto, definitivamente existe de que muitos dos menores autovalores  $\eta_i$  convergirão com grande precisão muito antes de que o espaço de Hilbert seja completamente percorrido.

Como um exemplo simples deste fenômeno, consideremos um Hamiltoniano

$\mathbf{H}$  com o menor autovalor sendo  $\lambda_0$ , distinto dos outros. Escolhendo um vetor inicial  $|v_1\rangle$ , o qual não é deficiente do estado fundamental deste Hamiltoniano, nós podemos expandir  $|v_1\rangle$  em termos dos autovalores exatos  $|\psi_i^{Ex}\rangle$ , e então operarmos sobre  $|v_1\rangle$  com potências de  $\mathbf{H}$  obtendo-se

$$\mathbf{H}^r |v_1\rangle = a_0 \lambda_0^r |\psi_0^{Ex}\rangle + \sum_{i=1}^N a_i \lambda_i^r |\psi_i^{Ex}\rangle. \quad (1.23)$$

Eventualmente o termo  $a_0 \lambda_0^r |\psi_0^{Ex}\rangle$  irá sobrepujar todos os outros; o valor de  $r$  para o qual o lado direito normalizado da eq.1.23 é considerado uma boa aproximação dependerá apenas fracamente do tamanho do espaço de Hilbert e dependerá principalmente da separação de  $\lambda_0$  dos outros autovalores. Este valor de  $r$  deve, de fato, estar simplesmente relacionado à dimensão do espaço truncado necessário para descrever o estado fundamental de  $\mathbf{H}$ .

De uma maneira geral, podemos observar que autovalores bem isolados perto dos extremos do espectro podem ser “encontrados” depois de um pequeno número de passos (20 ou 30 iterações) independentemente da ordem da matriz  $\mathbf{H}$ . Uma convergência rápida também ocorre para autovalores bem separados do centro do espectro ou que estejam bem representados no vetor inicial  $|v_1\rangle$  (cf. seção 1.4); no entanto, se o autovalor desejado estiver nos extremos do espectro, porém pouco separado do resto do espectro, sua convergência será lenta. Aproximações acuradas aos correspondentes autovetores também podem ser calculadas. Neste caso é necessário reter na memória os vetores  $\{|v_1\rangle, |v_2\rangle, \dots, |v_n\rangle\}$  obtidos no processo. A convergência pode ser monitorada numericamente diagonalizando a matriz tridiagonal depois de cada passo e observando a variação dos autovalores e autovetores

Da seção precedente observamos (no mínimo teoricamente) que para a recursão básica de Lanczos, a ortogonalização com respeito aos dois vetores gerados mais recentemente é suficiente para garantir que cada vetor sucessivo é ortogonal com respeito aos gerados previamente. O procedimento de Lanczos pode ser visto portanto como uma ortogonalização de Gram-Schmidt do conjunto de vetores  $\{|v_1\rangle, |Hv_1\rangle, \dots, |H^{(n-1)}v_n\rangle\}$  (vetores de Krylov [6]). No entanto, em cálculos computacionais, erros de arredondamento nas operações aritméticas básicas levam à perda de ortogonalidade dos vetores  $\{|v_1\rangle, |v_2\rangle, \dots, |v_n\rangle\}$ , fazendo com que as quantidades computadas pelo algoritmo diverjam fortemente dos seus valores na aritmética exata. O próprio Lanczos, diante deste comportamento, sugeriu uma reortogonalização completa de cada novo vetor gerado no processo, a todos os previamente obtidos. O custo deste novo arranjo do procedimento original, é uma maior demanda em termos de memória e tempo de processamento, limitando em alguns casos o número de autovalores que podem ser computados, tanto quanto o tamanho da matriz a ser guardada.

Através de uma análise numérica detalhada do método de Lanczos, Paige [8] elaborou uma variante computacional do algoritmo de Lanczos que não usa reortogonalização. Mesmo com a perda total de ortogonalidade, é possível obter autovalores acurados usando o método de forma iterativa. Na verdade, a ortogonalidade dos vetores não é realmente necessária, mas somente uma conveniência para alguns outros usos. O requerimento real é que os vetores sejam linearmente independentes. Paige mostrou que os autovalores obtidos no processo necessariamente convergem quando a dimensão da matriz truncada aumenta, e que diferem dos autovalores de  $\mathbf{H}$  por

quantidades que dependem da precisão da máquina. Erros de arredondamento são minimizados e boa convergência é obtida mesmo para autovalores muito próximos. Além disto, necessitamos menos memória, desde que somente dois vetores ao longo do processo de iteração necessitam ser armazenados.

O método de Lanczos aparece então como uma poderosa ferramenta para calcular o estado fundamental de Hamiltonianos (matrizes) com grandes dimensões, com a imensa vantagem de não ser preciso uma diagonalização direta da matriz completa. A convergência é acelerada se o vetor  $|v_1\rangle$ , que inicializa o processo, contém todas as simetrias da solução (cf. figura 1 e seção 1.5). Ainda, com a escolha do vetor inicial, podemos explorar diferentes regiões do espectro, e conseqüentemente o método não se limita apenas à determinação do estado fundamental (cf. seção 2.4).

### 1.3 O modelo da cadeia

Uma das características destacadas do método de Lanczos era o fato de podermos trabalhar com as matrizes e vetores envolvidos, interpretando-os como, por exemplo, hamiltonianos e estados nos quais estes atuam. Com isso abrimos a perspectiva de ir além da importância matemática do método e olhá-lo como um modo de melhor entender um sistema físico particular.

Para trabalharmos os aspectos físicos do método vamos partir de outro método recursivo, formulado por Haydoch [10], no qual se propõe a transformação de um modelo quântico geral em um dito modelo da cadeia, partindo-se da premissa de que tal transformação sempre pode ser feita para qualquer modelo da mecânica quântica

e então resolve-lo.

O ponto inicial da transformação é algum modelo quântico geral que nós investigaremos convertendo-o a um modelo da cadeia. O modelo geral consiste de um conjunto de estados, sua interpretação física, e um hamiltoniano que determina o movimento do sistema entre seus estados. O modelo pode ser especificado por funções de onda e um operador diferencial. A menos que o sistema físico seja unidimensional de algum modo fundamental, o hamiltoniano conectará cada estado com muitos (ou um número infinito) de outros estados. O método aqui apresentado é um procedimento estável de gerar uma representação tridiagonal de um dado operador.

Consideremos então, o operador hamiltoniano  $H$  (supomos aqui ele sendo hermitiano). Escolhemos um estado inicial  $|u_1\rangle$  normalizado. Esta escolha irá nos direcionar ao sistema que a cadeia representa. Por exemplo, se nós queremos explorar a estrutura eletrônica numa superfície, então  $|u_1\rangle$  deve ser um estado localizado na superfície metálica, ou mais simplesmente, se quisermos estudar o espectro e os estados de um oscilador harmônico, podemos escolher  $|u_1\rangle$  como sendo o estado fundamental ou o primeiro estado excitado do oscilador. Cada escolha de  $|u_1\rangle$  produz uma cadeia diferente, e portanto nós escolhemos  $|u_1\rangle$  para obtermos uma cadeia contendo a informação que queremos.

A partir disso, nós construímos uma hierarquia de estados ortonormais de acordo com o seguinte esquema de recorrência:

$$\begin{aligned}
\beta_1 |u_2\rangle &= H |u_1\rangle - \alpha_1 |u_1\rangle \\
\beta_2 |u_3\rangle &= H |u_2\rangle - \alpha_2 |u_2\rangle - \beta_1 |u_1\rangle \\
&\vdots \\
\beta_n |u_{n+1}\rangle &= H |u_n\rangle - \alpha_n |u_n\rangle - \beta_{n-1} |u_{n-1}\rangle
\end{aligned} \tag{1.24}$$

onde,

$$\begin{cases} \alpha_n = \langle u_n | H | u_n \rangle & , n = 1, 2, 3, \dots \\ e \\ \beta_n = \langle u_n | H | u_{n+1} \rangle & , n = 1, 2, 3, \dots \end{cases} \tag{1.25}$$

Nós geramos então uma cadeia de variáveis com elementos da diagonal da matriz  $\alpha_n$  e a interação de vizinhos mais próximos  $\beta_n$ , os quais descrevem a ação do hamiltoniano  $\mathbf{H}$  sobre os estados por meio de uma relação de recorrência com tres termos e simétrica. A recorrência é chamada simétrica por causa da componente  $|u_{n+1}\rangle$  em  $\mathbf{H}|u_n\rangle$  ser a mesma daquela de  $|u_n\rangle$  em  $\mathbf{H}|u_{n+1}\rangle$ . Isto é particularmente um resultado da hermiticidade do hamiltoniano.

Na fig.2.a fazemos um quadro esquemático do modelo em uma simples rede quadrada bi-dimensional.

O esquema mostrado na fig. 2.b equivale a escrever o modelo Hamiltoniano  $\mathbf{H}$  como

$$\mathbf{H} = \alpha_1 |u_1\rangle \langle u_1| + \beta_1 |u_1\rangle \langle u_2| + \beta_1 |u_2\rangle \langle u_1| + \dots + \alpha_n |u_n\rangle \langle u_n| \dots$$

A seqüência de estados  $|u_1\rangle, |u_2\rangle, \dots$  definida pelas equações 1.24 e 1.25 é formalmente equivalente a seqüência de vetores definida nas equações 1.17 a 1.19 do método de Lanczos; a única (não essencial) diferença é que nós preferimos construir, começando do requerimento de normalização, de modo que  $|u_i\rangle$  constitui um conjunto ortonormal, com interação de vizinhos próximos.

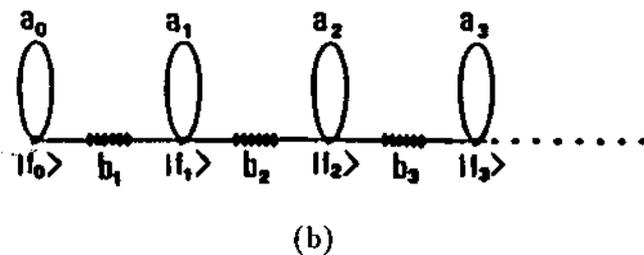
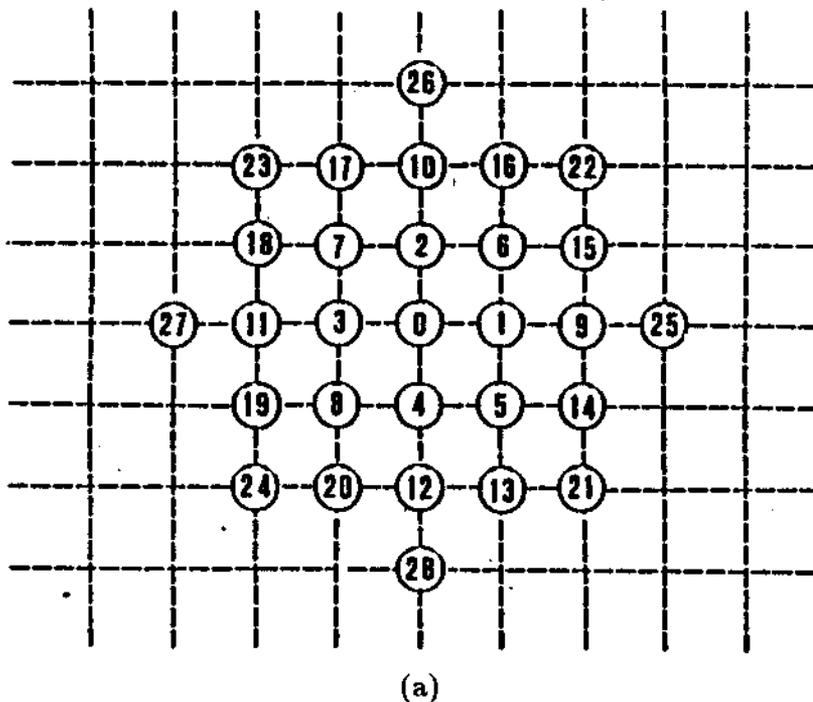


Figura 2: Rede quadrada bidimensional (a) no esquema de recursão (b). Definindo-se

$|\psi_0\rangle = |f_0\rangle = |\phi_0\rangle$ ;  $|\psi_1\rangle = \sum_1^4 |\phi_i\rangle$ ;  $|\psi_2\rangle = \sum_5^8 |\phi_i\rangle$ ;  $|\psi_3\rangle = \sum_9^{12} |\phi_i\rangle$ ;  $|\psi_4\rangle = \sum_{13}^{20} |\phi_i\rangle$ ;  $|\psi_5\rangle = \sum_{21}^{24} |\phi_i\rangle$ ;  $|\psi_6\rangle = \sum_{25}^{28} |\phi_i\rangle$ ; os primeiros estados  $f_i$  são dados por  $|f_0\rangle = |\psi_0\rangle$ ;  $|f_1\rangle = |\psi_1\rangle$ ;  $|f_2\rangle = 2|\psi_2\rangle + |\psi_3\rangle$ ;  $|f_3\rangle = 2|\psi_4\rangle + |\psi_6\rangle$ ; e os primeiros parâmetros  $b_i^2$  são  $b_1^2 = 4$ ,  $b_2^2 = 5$ ,  $b_3^2 = 19/2, \dots$

Este modelo é matematicamente equivalente a expressar o hamiltoniano  $H$  em uma forma tridiagonal como em 1.5. Como no procedimento de Lanczos, podemos diagonalizar a cada passo da cadeia, dada a sua característica recursiva, e todas as propriedades observadas em Lanczos permanecem válidas na construção deste modelo. Originalmente porém, este método recursivo, em oposição ao de Lanczos não focalizava a diagonalização de 1.5, mas sim a construção dos elementos da matriz da função de Green associada a um operador particularmente estudado. Voltaremos a este ponto na próxima secção com ênfase no uso do processo de Lanczos.

Há também que se ressaltar aqui uma interpretação física do modelo da cadeia (até onde isso é possível de se fazer). Da figura 2.a precedente, fica claro que a partir de um estado inicial do sistema, a cadeia estabelece uma ordem de estados da base em termos de como muitos estados intermediários devem ser passados para alcançar um dado estado. Outro modo de expressar isto é dizer que o comportamento do sistema inicialmente em  $|u_1\rangle$  é mais fortemente influenciado por  $|u_2\rangle$  e menos portanto, por cada um dos estados sucessivos da cadeia. A maior influência no movimento do sistema vem de estados próximos no sentido de serem acessíveis por meio de poucos estados intermediários. O modelo da cadeia é a expressão matemática do conceito de condições locais. Cada estado na cadeia representa uma parte mais distante das condições do estado inicial e os parâmetros de recorrência especificam o efeito daquelas condições no movimento do sistema.

A cadeia descreve a evolução do sistema a partir de um estado inicial. Os estados que não estão acoplados ao estado inicial não aparecem na cadeia. A ortonormal-

idade dos estados na cadeia significa simplesmente que a recorrência será simétrica para um hamiltoniano hermiteano.

Podemos dizer que uma cadeia é um autoestado generalizado. A cadeia simples contém um estado  $|v_1\rangle$ , e neste caso a equação 1.24 torna-se

$$\mathbf{H} |v_1\rangle = \alpha_1 |v_1\rangle , \quad (1.26)$$

que quer dizer obviamente que  $|v_1\rangle$  é um autoestado de  $\mathbf{H}$  (este é o caso em que o primeiro vetor gerado em Lanczos tem comprimento nulo, ou seja, o nosso estado teste é exatamente a solução do problema). Uma cadeia maior, no sentido de serem necessárias mais iterações para expandir o espaço inteiro de  $\mathbf{H}$ , representa um subespaço invariante contendo o espaço inicial, tal que  $\mathbf{H}$  atuando sobre qualquer estado, dá outro estado no subespaço. O modelo da cadeia corresponde ao menor subespaço invariante contendo o estado inicial, ou ainda, qualquer outro estado da cadeia. Isto nos leva (como já observado) a utilizar um número de estados reduzido em todo o processo; nós estudamos somente estados translacionalmente invariantes (usamos condições de contorno periódicas do começo ao fim).

Do colocado aqui, vamos à próxima secção discutir alguns resultados interessantes para ilustrar as propriedades observadas anteriormente.

## 1.4 O método de Lanczos e o método dos momentos

Nossa preocupação agora, é de estabelecer alguma relação entre o método de Lanczos e de momentos e com isso ligar Lanczos com a teoria de polinômios ortogonais e frações contínuas, bem como a possibilidade de obtenção de uma função de Green ligada a um operador particular. Porém, vale frisar, será feito o tratamento mais elementar possível, ressaltando apenas alguns resultados, deixando de lado muitas das ramificações possíveis. Meramente nos preocupamos em mostrar algumas conexões interessantes, colocando uma vasta literatura [2] [9] [11] [12] à disposição para um estudo mais apurado da questão.

Um grande número de problemas físicos são descritos por operadores os quais, em geral são muito complicados para uma diagonalização direta mas são ainda suficientemente manipuláveis para permitir, de um modelo razoavelmente simples, o cálculo dos momentos. Neste caso nós podemos obter informações preciosas sobre o sistema físico de interesse através de uma expansão em frações contínuas da função de Green de um dado operador. Este procedimento matemático é conhecido como o problema dos momentos, que tem sido amplamente estudado, em geral, tendo-se em mente operadores hermiteanos (problema dos momentos clássico). Consideremos então uma matriz  $\mathbf{H}$  hermiteana e um vetor  $|v\rangle$  arbitrário. Nós definimos uma matriz de momentos geral como tendo os elementos

$$\mu_n = \langle v | H^n | v \rangle , \quad (1.27)$$

com a ajuda da identidade

$$1 = \sum_{\alpha} |\Psi_{\alpha}\rangle \langle \Psi_{\alpha}| , \quad (1.28)$$

onde  $|\Psi_{\alpha}\rangle$  denota os autoestados de  $\mathbf{H}$  com energia (real)  $E_{\alpha}$ . Podemos escrever a eq. 1.28 como

$$\mu_n = \sum_{\alpha} E_{\alpha}^n |\langle \Psi_{\alpha}|v\rangle|^2 . \quad (1.29)$$

Se a dimensão de  $\mathbf{H}$  é densamente ocupada sobre a região de energia a eq. 1.29 torna-se

$$\mu_n = \int E^n n(E) dE , \quad (1.30)$$

e aqui  $\mu_n$  fica sendo o  $n$ -ésimo momento da distribuição

$$n(E) = a^2(E)P(E) , \quad (1.31)$$

onde  $a^2(\lambda)$  é a média quadrada da projeção de  $|v\rangle$  com um autovetor de  $\mathbf{H}$ , e  $P(E)$  o número de autovalores entre  $E$  e  $E + dE$  (densidade de autovalores). Se  $|v\rangle$  é tal que  $a^2(E)$  é constante (os  $|\langle \Psi_{\alpha}|v\rangle|^2$  são iguais na distribuição discreta)  $n(E)$  fica sendo justamente a densidade de estados.

Considerando-se os determinantes:

$$L_n = \begin{vmatrix} 1 & \mu_1 & \mu_2 & \dots & \mu_n \\ \mu_1 & \mu_2 & \dots & & \mu_{n+1} \\ \dots & & & & \vdots \\ \mu_n & \dots & \dots & \dots & \mu_{2n} \end{vmatrix} , \quad (1.32)$$

$$M_n = \begin{vmatrix} 1 & \mu_1 & \dots & \mu_{n-1} & \mu_{n+1} \\ \mu_1 & \mu_2 & \dots & \mu_n & \mu_{n+2} \\ \dots & & & & \vdots \\ \mu_n & \dots & \dots & \mu_{2n-1} & \mu_{2n+1} \end{vmatrix} . \quad (1.33)$$

Podemos construir uma matriz simétrica tridiagonal associada com os momentos dados, com elementos diagonais

$$a_i = M_{i-1}/L_{i-1} - M_{i-2}/L_{i-2} \quad (1.34)$$

e elementos fora da diagonal

$$b_i = \sqrt{L_i L_{i-2}}/L_{i-1} , \quad (1.35)$$

com  $M_{-1} = 0$ ,  $M_0 = \mu$ ,  $L_0 = 1$ .

Os  $a$ 's e os  $b$ 's podem ser olhados como "momentos reduzidos" no sentido de que eles carregam informação sobre a distribuição de momentos. A solução do problema é agora como segue: os autovalores e autovetores dessa matriz tridiagonal construída a partir dos  $a$ 's e  $b$ 's são os mesmos da matriz dos momentos definidos na equação 1.27 . Mais do que essa característica o procedimento acima nos dá uma sequência de soluções no sentido de se aproximar mais e mais da distribuição original, o que para algum propósito específico pode ser o suficiente. Portanto, truncando-se o procedimento de se obter a matriz tridiagonal associada aos momentos a qualquer ordem  $i$ , produz-se uma solução tal que seus  $2i - 1$  momentos independentes concordam exatamente com os primeiros  $2i - 1$  momentos da solução exata. Apesar desse importante resultado, este procedimento requer que calculemos determinantes como os das equações 1.32 e 1.33, o que vem a ser um procedimento extremamente instável, resultando na necessidade de se conhecer os momentos com quão pequena que seja a precisão e carregar esta precisão através dos cálculos. Então para uma precisão

aritmética finita, em geral, as delicadas relações que garantem que os momentos correspondem a uma distribuição são violadas.

Para o problema da matriz dos momentos portanto, nós podemos ver o método de Lanczos como um modo alternativo de se obter a matriz tridiagonal, sem realmente calcular os momentos o que, como observado anteriormente, garante mais estabilidade e maior precisão. Então, se nós quisermos formar a matriz  $n \times n$  da matriz tridiagonal associada aos momentos nós necessitamos somente fazer  $n$  passos do procedimento de Lanczos.

Esta estreita relação entre Lanczos e os Momentos nos permite especular um pouco mais sobre a convergência dos autovalores de  $\mathbf{H}$  no processo de Lanczos. Note-mos em primeiro lugar que passando-se  $\mathbf{H} \rightarrow \mathbf{H} - \mu_1 \mathbf{I}$ , com  $\mathbf{I}$  sendo a identidade, meramente fazemos com que  $a_i \rightarrow a_i - \mu_1$ , com os  $b$ 's permanecendo os mesmos. Isto corresponde a representar os momentos dados pela equação 1.28 como momentos centrais ou momentos sobre a média. Agora os altos momentos centrais são muito sensíveis aos autovalores extremos, mas depois de  $k$  passos nós calculamos, com efeito, o  $(2k - 1)$ -ésimo momento exatamente (ou tão precisamente quanto for necessário para calcular os  $a$ 's e os  $b$ 's). Então a convergência aos autovalores extremos é uma consequência do poder do método para calcular altos momentos. Não é correto que o fato de os autovalores extremos convergirem primeiro, eles dependem dos baixos momentos.

Continuando nossa "busca" de algumas conexões interessantes com o método de Lanczos, vamos estudar os autoestados e autovalores do modelo original obtidos

pelo método (ou se quisermos, na transformação para a cadeia). Isto nos levará a teoria de polinômios ortogonais (o que já era esperado da relação com os momentos derivada anteriormente).

Suponhamos portanto, que se queira achar os autoestados de uma dada matriz tridiagonal correspondendo aos autovalores desta. Para isso precisamos resolver a equação matricial

$$(1.36) \quad \begin{bmatrix} E - a_1 & -b_1 & 0 & 0 & \cdots \\ -b_1 & E - a_2 & -b_2 & 0 & \cdots \\ 0 & -b_2 & E - a_3 & -b_3 & \cdots \\ 0 & 0 & -b_3 & E - a_4 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ \vdots \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \end{bmatrix}$$

a qual é justamente a equação de Schrödinger independente do tempo

$$\mathbf{H} \sum_{n=1}^{\infty} \mathbf{P}_n |u_n\rangle = \mathbf{E} \sum_{n=1}^{\infty} \mathbf{P}_n |u_n\rangle \quad (1.37)$$

A equação 1.36 ainda define um conjunto de equações lineares as quais podem ser consideradas na relação de recorrência

$$b_n P_{n+1}(E) = (E - a_n) P_n(E) - b_{n-1} P_{n-1}(E), \quad (1.38)$$

que é a relação de recorrência fundamental para uma família de polinômios ortogonais  $\{P_n(E)\}$ . Esta sequência infinita (ou finita) termina se nós considerarmos os valores  $E = E_i$  nos quais  $P_n(E)$  vai a zero, ou seja, os autovalores da matriz truncada a  $n \times n$  são os zeros do polinômio de grau  $n$ ,  $P_n(E)$ . A recorrência apresentada em

1.38 é a contrapartida, em teoria de polinômios ortogonais, da relação 1.24 obtida na construção do modelo da cadeia (ou ainda a equação 1.17 do processo de Lanczos).

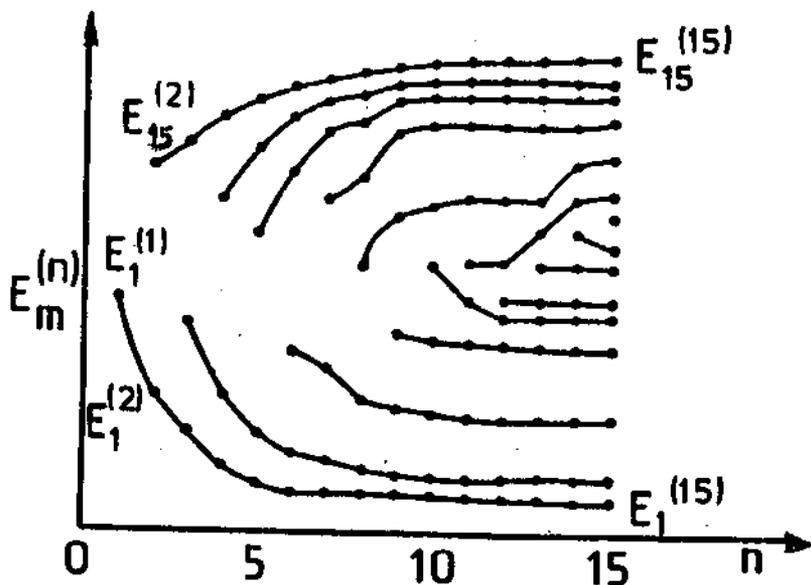


Figura 3: Comportamento esquemático das raízes de  $P_n(E)$  como uma função de  $n$

Voltando-se a questão da convergência dos autovalores, desde que são uma função de  $n$ , seus extremos convergem rapidamente e os centrais mais vagorosamente. Se nós calcularmos as raízes  $\{E_m^{(n)}\}$  de  $\{P_n(E)\}$  nós achamos que as maiores e menores raízes variam pouco com  $n$ .

Esta convergência move-se do extremo do espectro para o centro. Uma explicação qualitativa para este fenômeno é que  $(H - a_n) |u_n\rangle$  contém proporcionalmente

autovalores mais distantes de  $a_n$ ; isto é contrabalançado pela subtração de  $b_{n-1} |u_{n-1}\rangle$ . Para cada autovalor há um valor de  $n$  para o qual as subtrações sobrepujam as multiplicações e as raízes correspondentes cessam de variar (um esquema desse fenômeno é plotado na figura 3).

Para finalizar esta secção vamos agora focar nossa atenção na construção da função de Green (ou resolvente) recursivamente, tendo os momentos como ingrediente. Esta aproximação é muito mais conveniente para manusearmos operadores com um contínuo no espectro de energia, bem como, de um ponto de vista mais particular, calcular a densidade de estados para alguma distribuição. Portanto, tomando-se os momentos definidos na equação 1.29 podemos escrever a função de Green  $G(E)$ , relativa a estes momentos como :

$$G(E) = \sum_{n=1}^{\infty} \mu_n / E^{n+1} \quad (1.39)$$

para propósitos computacionais a função  $G(E)$  pode ser expandida na forma de frações continuadas ficando:

$$G(E) = \frac{1}{E - a_1 - \frac{b_1}{E - a_2 - \frac{b_2}{E - a_3 - \dots}}} \quad (1.40)$$

com os  $a_n$ 's e os  $b_n$ 's dados respectivamente por 1.34 e 1.35. Aqui os valores assintóticos (desde que a distribuição  $n(E)$  nos quais os momentos estão relacionados é tomada num contínuo de energia) para  $a_n$  e  $b_n$  produzem importantes informações físicas; de fato  $a_{\infty}$  é a posição do centro da banda de energia, enquanto a largura da banda é obtida de  $b_{\infty}$ .

Mais uma vez, obtemos parâmetros de recursão que são os mesmos dos gerados

no processo de Lanczos, nos permitindo agora escrever o resolvente iterativamente, de uma forma mais precisa e estável que quando trabalhado a partir do método dos momentos.

Existe uma vasta literatura a ser consultada [11] para uma análise mais profunda dos resultados derivados nesta seção, e por fugir ao escopo deste trabalho, não foi levada a cabo.

## 1.5 Simetrias

Como já observado anteriormente (cf.seção 1.2), tomando-se vantagem da simetria em um hamiltoniano, podemos economizar tempo e estocagem na computação do processo de Lanczos, ou simplificar consideravelmente o processo analítico de se obter as iterações de Lanczos, (transformação a uma cadeia). Ainda, uma simplificação do problema pode ser obtida com a escolha do estado inicial, além de com isso ser possível explorar todo o espectro de soluções. A única dificuldade é que uma análise mais apurada das simetrias pode consumir muito tempo ou ainda ser complicada por si só. Nesta secção, nós estudamos brevemente a aplicação de simetria à transformação do modelo da cadeia.

A simetria do hamiltoniano pode levar a degenerescência em seu espectro. Seja  $\mathbf{H}$  um hamiltoniano com autovalores  $E_n$ , possivelmente degenerados, correspondendo a uma escolha de autovetores ortonormais  $\{|\omega_\alpha\rangle\}$ . Se em termos dos  $\{|\omega_\alpha\rangle\}$ ,  $\mathbf{T}$  é a representação unitária de alguma simetria de  $\mathbf{H}$ , então  $\mathbf{H}$  e  $\mathbf{T}$  comutam

$$\mathbf{TH} = \mathbf{HT}, \quad (1.41)$$

portanto

$$\mathbf{H}\{|\omega_\alpha\rangle\} = E_\alpha\{|\omega_\alpha\rangle\}, \quad (1.42)$$

implica em

$$\mathbf{HT}\{|\omega_\alpha\rangle\} = E_\alpha\mathbf{T}\{|\omega_\alpha\rangle\} \quad (1.43)$$

Então  $\mathbf{T}\{|\omega_\alpha\rangle\}$  é também um autoestado de  $\mathbf{H}$  com a mesma energia que  $\{|\omega_\alpha\rangle\}$ . A menos que  $\exp(i\Phi_\alpha)|\omega_\alpha\rangle = \mathbf{T}|\omega_\alpha\rangle$ , para todo  $\alpha$ ,  $\mathbf{T}|\omega_\alpha\rangle$  é linearmente independente de  $|\omega_\lambda\rangle$  para algum  $\lambda$  e  $E_\lambda$  é então degenerado. Inversamente, supondo  $E_\lambda = E_\mu$ , para  $\lambda \neq \mu$ , e  $\mathbf{I}$  sendo a identidade; nós podemos construir uma transformação unitária

$$\mathbf{T} = \mathbf{I} - \{|\omega_\lambda\rangle - |\omega_\mu\rangle\}\{\langle\omega_\lambda| - \langle\omega_\mu|\} \quad (1.44)$$

O que neste caso é fácil verificar que  $\mathbf{T}$  comuta com  $\mathbf{H}$  e portanto representa uma simetria de  $\mathbf{H}$ . Não é sempre possível dar uma interpretação física para transformações que comutam com  $\mathbf{H}$ , mas nós as tratamos todas do mesmo modo.

Em contrapartida a possibilidade de simetria do hamiltoniano (e uma possível degenerescência), o modelo da cadeia não tem simetrias. Isto segue da natureza não degenerada do espectro de autovalores da cadeia. Para os propósitos desta seção, nós podemos ampliar o entendimento do modelo contendo estados de simetria simples somente, todas as outras simetrias tendo sido “jogadas fora”.

Quanto a escolha do estado inicial, o uso mais eficiente da transformação ao modelo da cadeia, resulta quando este estado pertence a uma representação irreduzível das simetrias de  $\mathbf{H}$ . Quando este é o caso, os únicos estados acoplados a  $|u_1\rangle$  (e portanto aparecendo no modelo) são aqueles pertencentes a mesma linha da

mesma representação irredutível das simetrias de  $\mathbf{H}$ . Isto quer dizer que somente tais estados necessitam ser retidos, e portanto na computação a mesma resolução pode ser guardada com muito menos estocagem e gasto de tempo porque a dimensionalidade da matriz  $\mathbf{H}$  é reduzida.

As cadeias produzidas por estados iniciais correspondentes a diferentes linhas da mesma representação irredutível são similares. Supondo  $\mathbf{T}$  como sendo uma representação unitária que leva  $|u_1\rangle$  em  $\mathbf{T}|u_1\rangle$ ; então porque  $\mathbf{T}$  comuta com  $\mathbf{H}$ , o  $n$ -ésimo estado  $|u_n\rangle$  na cadeia, gerado por  $|u_1\rangle$  vai para  $\mathbf{T}|u_n\rangle$ , o  $n$ -ésimo estado na cadeia gerado por  $\mathbf{T}|u_1\rangle$  e os parâmetros da cadeia são idênticos. Isto significa portanto que, um modelo da cadeia necessita ser calculado somente uma vez para cada representação irredutível.

Do colocado acima, enfatizamos que uma “boa” escolha do nosso vetor inicial simplifica o problema de transformação ao modelo da cadeia. Este fato nos sugere a possibilidade de algum sub-modelo para se obter um bom vetor inicial. Essa possibilidade, a nosso ver, seria um outro trabalho elaborado em um sentido um tanto diferente do que aqui nos propomos, e portanto deixamos essa possibilidade como uma sugestão de trabalho futuro. Entretanto, de modo a ampliar o sugerido, partimos da afirmação de que qualquer modelo quântico (aqui representado por uma hamiltoniana e os estados nos quais ela atua) complicado no sentido de suas simetrias, pode ser reduzido a uma “soma” (ou acoplamento) de modelos simples os quais sabemos sobre suas simetrias com isso podemos sempre escolher, pelo menos em princípio, como um “bom” estado inicial, o pertencente ao espaço de estados de algum desses modelos

simples componentes do modelo estudado efetivamente.

Voltaremos a discutir um pouco mais essa questão de uma maneira mais pragmática no próximo capítulo para esclarecer melhor a proposição.

# Bibliografia

- [1] C. Lanczos, J. Res. Natl. Bur. Stad., **45**, 255 (1950).
- [2] J. H. Wilkinson, *The Algebraic Eigenvalue Problem* (Clarendon Press, Oxford, 1965).
- [3] R. R. Whitehead *et al.*, Adv. Nucl. Phys., **9**, 123 (1977).
- [4] H. Mori, Jpn. Prog. Theor. Phys., **34**, 399 (1965).
- [5] R. Haydock *et al.*, J. Phys. **C8**, 2591 (1975).
- [6] J. K. Cullum and R. A. Willoughby, *Lanczos Algorithms for Large Symmetric Eigenvalue Computations*, vol.I - Theory (Progress in Scientific Computing - vol.3).
- [7] T. Sebe and J. Nachamkim, Ann. Phys., **51**, 100 (1969).
- [8] C. C. Paige, J. Inst. Math. Its. Appl., **10**, 373 (1972); **18**, 341 (1976).
- [9] R. Haydock, Sol. Stat. Phys., **35**, 215 (1980); (revisão).
- [10] R. Haydock, J.Phys. A: Math. Gen., **10**(4), 461 (1977).

- [11] R. R. Whitehead, in *Theory and Applications of Moment Methods in Many-Fermion Systems*, eds. B. J. Dalton *et al.* (Plenum Press, New York, 1980), p.235; (Revisão).
- [12] A. Tria *et al.*, *Phys. Rev.*, **B28**(4), 1859 (1983).
- [13] H. H. Romany, H. W. Wyld and L. E. Holloway, *Phys. Rev.*, **D21**, 1557 (1980).
- [14] M. Znojil, *Phys. Rev.*, **A135**(23), 87 (1991).
- [15] M. Znojil, *Phys. Rev.*, **A35**(6), 2448 (1987).
- [16] A. S. Davidov, *Quantum Mechanics* (Pergamon Press, New York).
- [17] N. H. March, W. H. Young and S Sampanthar, *The Many-Body Problem in Quantum Mechanics*, (Cambridge University Press, Cambridge, 1967).