

CAPITOLO 10

Legge di trasformazione di osservabili non simmetriche

Abbiamo già avuto occasione di osservare, sull'esempio dell'effetto Zeeman, che la riduzione di una r.i. rispetto a un sottogruppo può dare informazioni sulla rottura della degenerazione degli autovalori di un'osservabile simmetrica, ma non sulla separazione che ne risulta per gli autovalori. Vogliamo ora vedere che cosa si può dire in proposito, se si aggiunge la conoscenza della “legge di trasformazione” dell'osservabile.

Per chiarire che cosa s'intende per legge di trasformazione, pensiamo appunto all'effetto Zeeman. In quel caso la perturbazione che distrugge l'invarianza della hamiltoniana per il gruppo $SO(3)$ è l'interazione del momento magnetico dell'atomo col campo esterno: supponendo per ora che il momento magnetico sia proporzionale al momento angolare orbitale, e prendendo l'asse z nella direzione del campo, la perturbazione è proporzionale a L_z .

Si dice di solito che \vec{L} “per rotazioni si trasforma come un vettore”: con ciò s'intende che se R è una matrice di rotazione, e $U(R)$ il corrispondente operatore unitario sullo spazio di Hilbert \mathcal{H} degli stati, la trasformazione (passiva) di \vec{L} è data da

$$L'_i = U(R) L_i U(R)^+ = L_k R^k{}_i. \quad (10-1)$$

Si noti che la (10-1) può essere letta come segue: le tre osservabili L_1, L_2, L_3 generano, mediante combinazioni lineari, un s.s. di dimensione 3 nell'insieme degli operatori su \mathcal{H} . Su questo s.s. è definita una r.i. di $SO(3)$, e precisamente quella 3-dimensionale che chiamiamo “dei vettori.”

La stessa situazione si può presentare in molti altri casi: dovremo comunque considerare un s.s. vettoriale nell'insieme degli operatori, e sarà per noi interessante il caso in cui tale s.s. sia invariante rispetto a un certo gruppo di simmetria.

Se abbiamo a che fare con un'unica osservabile *simmetrica*, siamo ovviamente in un caso particolare, per due ragioni:

- il s.s. in questione è unidimensionale (quindi irriducibile)
- la r.i. in questione è quella banale (identità).

Trasformazione di elementi di matrice

Consideriamo un generico gruppo di simmetria \mathcal{G} e due suoi s.s.i. minimi in \mathcal{H} , \mathcal{V}_a e \mathcal{V}_b ; indicheremo con D_a e D_b le corrispondenti matrici di r. del gruppo. Sia poi dato un s.s.i. \mathcal{S} di operatori, che supporremo anch'esso minimo (se non lo fosse, procederemo a decomporlo in s.s.i. minimi) e siano ancora D_c le matrici

di r . Ciò vuol dire che se $|a, i\rangle, |b, j\rangle$ sono basi risp. di $\mathcal{V}_a, \mathcal{V}_b$, avremo

$$\begin{aligned}\varrho_a(g) &: |a, i\rangle \mapsto U(g)|a, i\rangle = |a, i'\rangle [D_a(g)]^{i'_i} \\ \varrho_b(g) &: |b, j\rangle \mapsto U(g)|b, j\rangle = |b, j'\rangle [D_b(g)]^{j'_j}.\end{aligned}$$

Analogamente, se $\{A_k\}$ è una base di \mathcal{S} :

$$\varrho_c(g) : A_k \mapsto U(g) A_k U(g)^+ = A_{k'} [D_c(g)]^{k'_k}.$$

Vogliamo ora studiare l'effetto di \mathcal{G} sugli elementi di matrice $\langle a, i|A_k|b, j\rangle$. Si ha:

$$\begin{aligned}\langle a, i|A_k|b, j\rangle &= \langle a, i|U(g)^+ U(g) A_k U(g)^+ U(g)|b, j\rangle \\ &= \sum_{i'j'k'} \langle a, i'|A_{k'}|b, j'\rangle ([D_a(g)]^{i'_i})^* [D_b(g)]^{j'_j} [D_c(g)]^{k'_k}.\end{aligned}\quad (10-2)$$

Si vede che a destra abbiamo il prodotto di tre r.i.: se fossero solo due, potremmo sfruttare le relazioni di ortogonalità viste al cap. prec. Ma consideriamo il prodotto $[D_b(g)]^{j'_j} [D_c(g)]^{k'_k}$: questo è esattamente ciò che in precedenza abbiamo indicato con $D_b \otimes D_c$, ossia il prodotto diretto delle due r . Tale prodotto sarà in generale riducibile, e potremo ridurlo decomponendo il prodotto tensoriale $\mathcal{V}_b \otimes \mathcal{S}$ in una somma di s.s.i. minimi.

Convieni a questo punto cambiare notazione: invece di \mathcal{S} e della sua base $\{A_k\}$, scriviamo \mathcal{V}_c e $|c, k\rangle$. Allora la riduzione del prodotto tensoriale $\mathcal{V}_b \otimes \mathcal{V}_c$ darà questo risultato:

$$|b, j\rangle|c, k\rangle = \sum_{r,l} |r, l\rangle \langle r, l|bc, jk\rangle \quad (10-3)$$

dove r è l'indice di una r.i. e l quello del generico vettore del s.s.i. \mathcal{V}_r . I coefficienti $\langle r, l|bc, jk\rangle$ che compiono la trasformazione di base non sono altro che i già citati *coeff. di Clebsch-Gordan*.

Se applichiamo al primo membro della (10-3) la simmetria g , otteniamo

$$\sum_{j'k'} |b, j'\rangle|c, k'\rangle [D_b(g)]^{j'_j} [D_c(g)]^{k'_k}$$

che grazie alla stessa (10-3) è anche

$$\sum_{j',k'} \sum_{r,l'} |r, l'\rangle \langle r, l'|bc, j'k'\rangle [D_b(g)]^{j'_j} [D_c(g)]^{k'_k}.$$

Se applichiamo la stessa simmetria al secondo membro abbiamo

$$\sum_{r,l} \sum_{l'} |r, l'\rangle [D_r(g)]^{l'_l} \langle r, l|bc, jk\rangle$$

e le due espressioni debbono coincidere: quindi

$$\sum_{j'k'} \langle r, l' | bc, j'k' \rangle [D_b(g)]^{j'}_j [D_c(g)]^{k'}_k = \sum_l [D_r(g)]^{l'}_l \langle r, l | bc, jk \rangle. \quad (10-4)$$

Ma la matrice formata dai coeff. di CG è unitaria, essendo un cambiamento di base:

$$\sum_{r,l'} \langle bc, jk | r, l' \rangle \langle r, l' | bc, j'k' \rangle = \delta_{jj'} \delta_{kk'}$$

e perciò dalla (10-4) si arriva a

$$[D_b(g)]^{j'}_j [D_c(g)]^{k'}_k = \sum_{r,l'} \sum_l \langle bc, j'k' | r, l' \rangle [D_r(g)]^{l'}_l \langle r, l | bc, jk \rangle.$$

Sostituendo nella (10-2):

$$\begin{aligned} \langle a, i | A_k | b, j \rangle = \\ \sum_{i'j'k'} \langle a, i' | A_{k'} | b, j' \rangle ([D_a(g)]^{i'}_i)^* \sum_{r,l'} \sum_l \langle bc, j'k' | r, l' \rangle [D_r(g)]^{l'}_l \langle r, l | bc, jk \rangle. \end{aligned}$$

Il teorema di Wigner–Eckart

Possiamo finalmente sfruttare le relazioni di ortogonalità (9-6), integrando sul gruppo, e arriviamo a

$$\langle a, i | A_k | b, j \rangle = \sum_{i'j'k'} \langle a, i' | A_{k'} | b, j' \rangle \langle bc, j'k' | a, i' \rangle \langle a, i | bc, jk \rangle. \quad (10-5)$$

Si vede che il primo membro contiene tre indici liberi (i, j, k) che a secondo membro figurano solo nell'ultimo fattore, che è un coeff. di CG. Se poniamo

$$\langle a || A || b \rangle \stackrel{\text{def}}{=} \sum_{i'j'k'} \langle a, i' | A_{k'} | b, j' \rangle \langle bc, j'k' | a, i' \rangle$$

(si chiama “elemento di matrice ridotto”) arriviamo al *teorema di Wigner–Eckart*:

$$\langle a, i | A_k | b, j \rangle = \langle a || A || b \rangle \langle a, i | bc, jk \rangle. \quad (10-6)$$

In parole: *gli elementi di matrice dell'insieme di osservabili $\{A_k\}$, che si trasformano secondo la r.i. D_c , tra i due sottospazi \mathcal{V}_a e \mathcal{V}_b , risp. supporti delle r.i. D_a e D_b , sono dati dal prodotto di un elemento di matrice ridotto — che dipende solo dalle particolari osservabili e dalle r . — e del coeff. di CG di D_a nella riduzione del prodotto diretto $D_b \otimes D_c$.*

È però necessaria una precisazione: nella dimostrazione del teorema abbiamo dato per scontato che nella riduzione del prodotto $D_b \otimes D_c$ ciascuna r.i. figuri *una volta sola*. Questo è vero per $SO(3)$, ma non in generale: fra i gruppi importanti per la fisica, già $SU(3)$ porta a una situazione più complicata. In altre parole, può accadere che nel prodotto tensoriale $\mathcal{V}_b \otimes \mathcal{V}_c$ esistano più s.s.i. minimi equivalenti.

In questo caso nella somma a destra della (10-3) compariranno più termini con lo stesso indice r . Procedendo nel calcolo, nel secondo membro della (10-5) resterà ancora una somma su r . equivalenti:

$$\langle a, i | A_k | b, j \rangle = \sum_{i' j' k'} \sum_p \langle a, i' | A_{k'} | b, j' \rangle \langle bc, j' k' | a, i' \rangle_p \langle a, i | bc, jk \rangle_p$$

dove è stato aggiunto un indice p che serve a distinguere le diverse r . equivalenti a D_a . Per ciascuna di queste il cambiamento di base potrà essere diverso, e avranno quindi diversi coeff. di CG. Dovremo perciò definire più e.m.r.:

$$\langle a || A || b \rangle_p \stackrel{\text{def}}{=} \sum_{i' j' k'} \langle a, i' | A_{k'} | b, j' \rangle \langle bc, j' k' | a, i' \rangle_p$$

e la (10-6) si modificherà di conseguenza:

$$\langle a, i | A_k | b, j \rangle = \sum_p \langle a || A || b \rangle_p \langle a, i | bc, jk \rangle_p. \quad (10-7)$$

Il grande valore del teorema di WE sta nell'economia che consente nel calcolo degli elementi di matrice: per un intero insieme di osservabili, e per tutti gli stati, basta conoscere uno (o pochi) e.m.r. (e i coeff. di CG) per ricavare tutti gli elementi di matrice.

Nell'applicazione pratica, spesso si procede così: supponiamo, per cominciare, di essere nel caso semplice in cui c'è un solo e.m.r., e di avere due insiemi di osservabili, A_k e B_k , che si trasformano secondo la stessa r.i. di \mathcal{G} . Allora $\langle a, i | A_k | b, j \rangle$ e $\langle a, i | B_k | b, j \rangle$ sono proporzionali, nel senso che il loro rapporto non dipende da i, j, k . Se accade che gli elementi di matrice di una delle A siano facilmente calcolabili, se ne ricavano quelli della corrispondente B *a meno di un fattore comune*.

Se gli e.m.r. sono più di uno (diciamo m), si può applicare il ragionamento a $m+1$ insiemi di osservabili: $A_{1k} \dots, A_{mk}$ e B_k . Gli elementi di matrice delle B saranno combinazioni lineari di quelli delle A :

$$\langle a, i | B_k | b, j \rangle = \sum_{p=1}^m c^p \langle a, i | A_{pk} | b, j \rangle$$

ma i coefficienti c^p non possono essere dedotti per questa via. Restano parametri liberi, che a seconda dei casi potremo ottenere da esperimenti o da altre parti della teoria.

Esempio 1: l'effetto Zeeman anomalo

L'esempio più ovvio di applicazione del teorema di WE è l'effetto Zeeman anomalo. Se teniamo conto anche del momento magnetico di spin, l'interazione dell'atomo col campo magnetico si scrive

$$H' = -\mu_0 B (L_z + 2S_z)$$

dove, come di solito, μ_0 indica il magnetone di Bohr, L è il momento angolare orbitale dell'atomo e S quello di spin.

Faremo il calcolo nell'ipotesi che si possano classificare i livelli imperturbati coi numeri quantici L, S, J (accoppiamento di Russell–Saunders) e che l'effetto del campo magnetico sia piccolo rispetto alla separazione dovuta all'interazione spin-orbita.

Dovremo allora calcolare soltanto gli elementi di matrice relativi a uno stesso s.s.i. minimo di $\text{SO}(3)$: $\langle LSJ, m | H' | LSJ, m' \rangle$. Qui J indica la r.i. (corrisponde quindi agli a e b della teoria generale) mentre m, m' (autovalori di J_z) corrispondono a i, j . Le osservabili A_k sono ora le componenti di $\vec{L} + 2\vec{S}$, che ovviamente si trasformano secondo la r.i. $J = 1$ di $\text{SO}(3)$.

Nota: A rigore qui non dovremmo parlare di r. di $\text{SO}(3)$, bensì di $\text{SU}(2)$. Riprenderemo la questione nel prossimo capitolo; per ora accontentiamoci dell'identificazione “classica” del gruppo delle rotazioni con $\text{SO}(3)$, anche se in questo modo non si spiegano i valori semiinteri di J .

Per applicare la tecnica descritta sopra, basta osservare che J_z si trasforma come $L_z + 2S_z$ ed è diagonale, con autovalori m : pertanto avremo

$$\langle LSJ, m | H' | LSJ, m' \rangle = -g\mu_0 B m \delta_{mm'}$$

il che già mostra che i livelli sono equidistanti. Resta solo da determinare la costante g (fattore di Landé).

Osserviamo anche che

$$\begin{aligned} \langle LSJ, m | H' | LSJ, m' \rangle &= -\mu_0 B \langle LSJ, m | (L_z + 2S_z) | LSJ, m' \rangle \\ &= -\mu_0 B \langle LSJ, m | (J_z + S_z) | LSJ, m' \rangle \\ &= -\mu_0 B (m \delta_{mm'} + \langle LSJ, m | S_z | LSJ, m' \rangle) \end{aligned}$$

per cui il solo problema è il calcolo degli elementi di matrice di S_z . Posto

$$\langle LSJ, m | \vec{S} | LSJ, m' \rangle = c \langle LSJ, m | \vec{J} | LSJ, m' \rangle$$

(sempre per il teorema di WE) avremo in particolare

$$\langle LSJ, m | S_z | LSJ, m' \rangle = c m \delta_{mm'}$$

da cui $g = 1 + c$.

Per arrivare agli elementi di matrice di S_z possiamo procedere così:

$$\begin{aligned} \langle LSJ, m | (\vec{J} \cdot \vec{S}) | LSJ, m' \rangle &= \sum_{m''} \langle LSJ, m | \vec{J} | LSJ, m'' \rangle \cdot \langle LSJ, m'' | \vec{S} | LSJ, m' \rangle \\ &= c \sum_{m''} \langle LSJ, m | \vec{J} | LSJ, m'' \rangle \cdot \langle LSJ, m'' | \vec{J} | LSJ, m' \rangle \\ &= c \langle LSJ, m | (\vec{J} \cdot \vec{J}) | LSJ, m' \rangle = c J(J+1) \delta_{mm'} \end{aligned}$$

(il primo e il terzo passaggio sono giustificati dal fatto che \vec{J} non ha elementi di matrice fra stati con diverso J nello stesso multipletto di struttura fina, né fra stati appartenenti a diversi multipletti, perché le altre osservabili assunte diagonali nella base sono scalari, quindi commutano con \vec{J}). D'altra parte vale l'identità

$$\vec{J} \cdot \vec{S} = \frac{1}{2} (|\vec{J}|^2 + |\vec{S}|^2 - |\vec{L}|^2)$$

che fornisce

$$\langle LSJ, m | (\vec{J} \cdot \vec{S}) | LSJ, m' \rangle = \frac{1}{2} [J(J+1) + S(S+1) - L(L+1)] \delta_{mm'}.$$

Confrontando

$$c = \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)}$$

e infine

$$g = \frac{3}{2} + \frac{S(S+1) - L(L+1)}{2J(J+1)}.$$

Esempio 2: la struttura iperfina

La struttura iperfina dei livelli atomici è dovuta a due effetti: l'interazione fra il momento magnetico degli elettroni e quello del nucleo, e la modifica al campo elettrico del nucleo che si presenta quando questo non ha simmetria sferica. Qui vogliamo occuparci del secondo, limitandoci al caso di un solo elettrone ottico.

In termini generali, l'interazione elettrostatica fra l'elettrone e il nucleo si può scrivere come somma delle interazioni con i singoli protoni:

$$V = -e^2 \sum |\vec{r} - \vec{r}_p|^{-1}$$

dove \vec{r} e \vec{r}_p sono posizione dell'elettrone e del p -mo protone, rispettivamente. Sebbene \vec{r} e \vec{r}_p siano osservabili del sistema, ossia operatori e non numeri, è lecito uno sviluppo in serie di multipoli, della forma

$$V = -\frac{Ze^2}{r} - \frac{e^2}{r^3} \sum_p \vec{r} \cdot \vec{r}_p - \frac{e^2}{2r^5} \sum_p [3(\vec{r} \cdot \vec{r}_p)^2 - r^2 r_p^2] - \dots \quad (10-8)$$

Questo perché su tutti gli stati dell'elettrone esclusi quelli con $l = 0$ (che però, come vedremo, non contano) la funzione d'onda dell'elettrone internamente al nucleo è trascurabile.

Il primo termine della (10-8) (monopolo) è l'energia potenziale dell'elettrone nel campo a simmetria sferica del nucleo supposto puntiforme (o sferico). Quanto al secondo (dipolo) se non ci fossero i neutroni potrebbe essere reso nullo semplicemente prendendo come origine dei vettori \vec{r}_p il centro di massa del nucleo. La presenza dei neutroni complica le cose, ma vedremo poi che il termine di dipolo non ha effetto per il nostro problema. Resta, come primo termine importante, il terzo, che è l'interazione di quadrupolo.

Dobbiamo ora descrivere gli stati sui quali c'interessa studiare l'effetto della perturbazione. In presenza del solo monopolo, possiamo classificare gli autovettori della hamiltoniana coi soliti numeri quantici n, l, j, m per l'elettrone; per il nucleo servirà il solo livello fondamentale, che potrà essere degenerare e descriveremo coi due numeri quantici I e M , che indicano rispettivamente il modulo e la componente z del momento angolare (spin) del nucleo.

Complessivamente gli stati che c'interessano stanno nel prodotto tensoriale, e come base possiamo prendere $|nljm, IM\rangle$. Dobbiamo aspettarci degenerazione su m e su M , ossia in totale $(2j+1)(2I+1)$ stati con la stessa energia. Il nostro problema è come il resto dell'interazione risolve questa degenerazione.

Osserviamo ora che gli stati dell'elettrone (e anche quelli del nucleo, ma non è necessario pensarci) hanno parità definita, la stessa per tutti quelli con lo stesso l , mentre \vec{r} è dispari per inversione spaziale: ne segue che elementi di matrice del tipo $\langle nljm | \vec{r} | nljm \rangle$ sono nulli. È questa la ragione per cui il termine di dipolo non conta.

Nota 1: A prima vista potrebbe sembrare che l'argomento non sia concludente, perché l'inversione spaziale cambia segno tanto a \vec{r} quanto a \vec{r}_p , per cui il loro prodotto scalare è pari, non dispari. Ma siamo proprio in un caso in cui serve considerare non già l'inversione spaziale di *tutte* le osservabili, bensì solo di quelle relative all'elettrone. L'argomento si appoggia sul fatto che gli autostati della hamiltoniana dell'elettrone nel campo a simmetria sferica del nucleo hanno parità definita rispetto alla sola inversione spaziale dell'elettrone.

Nota 2: Poiché c'interessano solo e.m. fra stati con gli stessi nl , d'ora in poi ometteremo d'indicare questi numeri quantici.

Riassumendo, dobbiamo studiare gli e.m. del solo termine di quadrupolo V_Q :

$$\begin{aligned}
& \langle jm, IM | V_Q | jm', IM' \rangle \\
&= -\frac{1}{2} e^2 \langle jm, IM | r^{-5} \sum_p [3 (\vec{r} \cdot \vec{r}_p)^2 - r^2 r_p^2] | jm', IM' \rangle \\
&= -\frac{1}{2} e^2 \langle jm, IM | r^{-5} \sum_{ik} (3 x^i x^k - r^2 \delta^{ik}) \sum_p (3 x_p^i x_p^k - r_p^2 \delta^{ik}) | jm', IM' \rangle \\
&= -\frac{1}{2} e^2 \sum_{ik} \langle jm | r^{-5} (3 x^i x^k - r^2 \delta^{ik}) | jm' \rangle \langle IM | \sum_p (3 x_p^i x_p^k - r_p^2 \delta^{ik}) | IM' \rangle.
\end{aligned} \tag{10-9}$$

Nell'ultima riga della (10-9) compaiono degli e.m. ai quali possiamo finalmente applicare il teorema di WE. In entrambi i casi gli operatori di cui vogliamo gli e.m. sono tensori simmetrici a traccia nulla, il che ci permette intanto di asserire che *l'effetto dell'interazione di quadrupolo è nullo a meno che non sia $j \geq 1$ e $I \geq 1$* . In particolare *non c'è struttura iperfina se $l = 0$* : è questo il motivo per cui nello sviluppo in multipoli non ci dobbiamo preoccupare della penetrazione dell'elettrone nel nucleo.

Possiamo anche dire, più in generale, che *nuclei con spin minore di 1 non hanno momento di quadrupolo*.

Più in dettaglio:

$$\begin{aligned}
& \langle jm | r^{-5} (3 x^i x^k - r^2 \delta^{ik}) | jm' \rangle = A \langle jm | [3 (J^i J^k + J^k J^i) - 2j(j+1) \delta^{ik}] | jm' \rangle \\
& \langle IM | \sum_p (3 x_p^i x_p^k - r_p^2 \delta^{ik}) | IM' \rangle = B \langle IM | [3 (I^i I^k + I^k I^i) - 2I(I+1) \delta^{ik}] | IM' \rangle
\end{aligned}$$

dove A e B sono coefficienti da calcolare a parte. (È stato necessario scrivere $J^i J^k + J^k J^i$ per avere un tensore simmetrico, dato che J^i e J^k non commutano se $i \neq k$; analogamente per I^i e I^k .)

Quando si va a sostituire nella (10-9) c'è da calcolare

$$\sum_{ik} [3 (J^i J^k + J^k J^i) - 2j(j+1) \delta^{ik}] [3 (I^i I^k + I^k I^i) - 2I(I+1) \delta^{ik}].$$

Sfruttando le relazioni di commutazione si arriva a

$$36 (\vec{J} \cdot \vec{I})^2 + 18 \vec{J} \cdot \vec{I} - 12 j(j+1) I(I+1).$$

Al prodotto scalare $\vec{J} \cdot \vec{I}$ si applicano considerazioni analoghe a quelle viste, per l'effetto Zeeman, a proposito di $\vec{J} \cdot \vec{S}$: s'introduce il vettore $\vec{F} = \vec{J} + \vec{I}$ e si riduce il prodotto tensoriale con base $|jm, IM\rangle$ in una somma diretta, descritta dalla base $|jIFM_F\rangle$. Si avrà ancora degenerazione su M_F , mentre

$$2 \vec{J} \cdot \vec{I} = F(F+1) - j(j+1) - I(I+1).$$

Ne segue che l'interazione di quadrupolo separa gli autovalori dell'energia in funzione del numero quantico F , che come al solito va da $|j - I|$ a $j + I$.