

Metodi elementari di calcolo chimico

Appunti di lezione ed esercizi

`antonino.polimeno@unipd.it`

Indice

1	Alcune nozioni di base	7
1.1	Derivate parziali	7
1.2	Differenziali lineari	9
2	Algebra lineare: principi	11
2.1	Spazi vettoriali	11
2.2	Prodotti scalari	13
2.3	Vettori e matrici	14
2.4	Matrici	16
2.5	Sistemi lineari	22
2.6	Trasformazioni	23
2.7	Problemi agli autovalori	24
2.8	Funzioni di matrici	26
2.9	Spazi di Hilbert	27
2.10	Operatori lineari	28
2.11	Postulati della quantomeccanica in una terminologia algebrica	29
2.12	Esercizi	30
3	Algebra lineare: applicazioni	31
3.1	Sistemi lineari: esempi e metodi di soluzione	31
3.1.1	Sistema di equazioni lineari	31
3.1.2	Bilanciamento di una reazione chimica	33
3.1.3	Eliminazione di Gauss-Jordan	35
3.1.4	Decomposizioni	37
3.1.5	Matrici sparse	39
3.1.6	Metodo del gradiente coniugato	40
3.2	Problemi agli autovalori: esempi e metodi di soluzione	42
3.2.1	Problemi semplici agli autovalori in meccanica quantistica	42
3.2.2	Un sistema a due livelli	42
3.2.3	Il metodo di Hückel per i polieni coniugati	48
3.2.4	Problemi agli autovalori di grandi dimensioni	50
3.2.5	Metodi numerici	54
4	Integrazione di funzioni	57
4.1	Definizioni	57
4.1.1	Integrali indefiniti, definiti, generalizzati, multipli	57
4.1.2	Metodi di integrazione analitici	59
4.2	Metodi numerici di integrazione	61
5	Equazioni non lineari	65
5.1	Definizioni	65
5.2	Soluzioni esatte	66
5.3	Soluzioni numeriche	67
5.3.1	Una sola variabile	67

5.3.2	Molte variabili	69
6	Equazioni differenziali	71
6.1	Introduzione	71
6.1.1	Equazioni differenziali ordinarie	72
6.1.2	Equazioni lineari	73
6.1.3	Sistemi lineari del primo ordine a coefficienti costanti	75
6.2	Equazioni alle derivate parziali	77
6.2.1	Condizioni al contorno	79
6.3	Esercizi	80
7	Equazioni differenziali: applicazioni	83
7.1	Esempi: termodinamica e cinetica chimica	83
7.1.1	Termodinamica	83
7.1.2	Cinetica	85
7.2	Esempi: sistemi quantistici risolvibili analiticamente	91
7.2.1	Fuga dal pozzo	94
7.2.2	Una scatola con una barriera interna	95
7.2.3	Scatola tridimensionale	97
7.2.4	Particella <i>in</i> e <i>su</i> un anello	98
7.2.5	Particella <i>in</i> e <i>su</i> una sfera	100
7.2.6	Oscillatore armonico	102
7.2.7	Potenziale di Morse	103
7.2.8	Atomo idrogenoide	104
7.3	Metodi numerici per le equazioni differenziali	105
7.3.1	Differenze finite	105
7.3.2	Schemi di discretizzazione	106
7.3.3	Applicazioni	107
7.4	Dinamica molecolare: metodologie ed applicazioni	109
7.4.1	Potenziali di interazione	110
7.4.2	Liquidi molecolari	111
7.4.3	Metodi computazionali di MD	114
7.4.4	Metodi alle differenze finite	115
7.4.5	Molecole lineari	116
7.5	Cenni di teoria delle perturbazioni	117
7.5.1	Perturbazioni dipendenti dal tempo	120
A	Funzioni speciali	121
A.1	Integrali ellittici	121
A.2	Funzione gamma	122
A.3	Funzione dell'errore	122
A.4	Funzioni di Bessel	122
A.5	Polinomi ortogonali	123
A.6	Polinomi di Legendre	123
A.7	Polinomi di Hermite	123
A.8	Polinomi di Laguerre	124
B	L'atomo idrogenoide	125
C	Problema generalizzato agli autovalori	127
D	Studio MD di un fluido atomico	129
D.1	Il sistema	129
D.2	Metodo di calcolo	130
D.3	Codice	131

Caveat

In queste note, gli studenti interessati troveranno spunti di discussione e proposte di esercizi, utili per la comprensione dei modelli analitici e numerici applicati correntemente alla descrizione di problemi comuni della chimica inorganica, fisica, organica, analitica. L'intento è quello di (ri)svegliare la curiosità verso una serie di *tecniche*, utili per applicare meglio i vari strumenti computazionali che dovrebbero oggi fare parte del bagaglio culturale di uno studente di chimica o chimica industriale, per il quale saper risolvere un sistema lineare di equazioni è utile, tanto quanto il saper maneggiare una bilancia di precisione o uno spettrometro di massa.

Nel testo sono segnalati come *esempi* alcune applicazioni semplici delle metodologie illustrate; in blu sono evidenziate alcune *applicazioni* più elaborate. Infine sono citati alcuni spunti di discussione relativi gli argomenti principali del testo indicati, in rosso, come *approfondimenti* .

Capitolo 1

Alcune nozioni di base

1.1 Derivate parziali

È data una funzione generica reale $f(x_1, x_2, \dots, x_m)$ di m variabili indipendenti reali, che per ora assumiamo definita in \mathbb{R}^m . Nel seguito, assumiamo che la funzione f abbia un buon comportamento (derivabile ovunque, invertibile rispetto a tutte le coordinate etc., salvo al più che in un numero finito di punti). La **derivata parziale** di f rispetto a x_i è il limite del rapporto tra la variazione della funzione e la variazione di x_i mantenendo tutte le altre $m - 1$ coordinate costanti

$$\left(\frac{\partial f}{\partial x_i}\right)_{x_j \neq i} = \lim_{\delta x_i \rightarrow 0} \frac{f(x_1, \dots, x_i + \delta x_i, \dots, x_m) - f(x_1, \dots, x_i, \dots, x_m)}{\delta x_i} \quad (1.1)$$

L'ordine di differenziazione è ininfluente sul risultato

$$\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{x_k \neq i, j} = \left(\frac{\partial^2 f}{\partial x_j \partial x_i}\right)_{x_k \neq i, j} \quad (1.2)$$

Il **differenziale totale** df di una funzione f è

$$df = \sum_{i=1}^m \left(\frac{\partial f}{\partial x_i}\right)_{x_j \neq i} dx_i \quad (1.3)$$

dove dx_i è una variazione infinitesimale e arbitraria della i -esima variabile indipendente x_i .

Esempio 1. Data la relazione $p(n, V, T) = nRT/V$ calcoliamo le derivate prime di p . Le derivate sono

$$\begin{aligned} \left(\frac{\partial p}{\partial n}\right)_{V, T} &= \frac{RT}{V} \\ \left(\frac{\partial p}{\partial V}\right)_{n, T} &= -\frac{nRT}{V^2} \\ \left(\frac{\partial p}{\partial T}\right)_{n, V} &= \frac{nR}{V} \end{aligned} \quad (1.4)$$

Esempio 2. Calcoliamo il differenziale totale di p , definita dalla relazione precedente. Si trova subito che

$$dp = \left(\frac{\partial p}{\partial n}\right)_{V, T} dn + \left(\frac{\partial p}{\partial V}\right)_{n, T} dV + \left(\frac{\partial p}{\partial T}\right)_{n, V} dT = \frac{R}{V} \left(Tdn - \frac{nT}{V}dV + ndT\right)$$

Data la funzione $Z = Z(X, Y)$ dimostriamo che

$$\left(\frac{\partial Y}{\partial Z}\right)_X = \frac{1}{\left(\frac{\partial Z}{\partial Y}\right)_X} \quad (1.5)$$

e la cosiddetta "regola ciclica"

$$\left(\frac{\partial X}{\partial Y}\right)_Z \left(\frac{\partial Y}{\partial Z}\right)_X \left(\frac{\partial Z}{\partial X}\right)_Y = -1 \quad (1.6)$$

Partiamo dal differenziale totale di Z

$$dZ = \left(\frac{\partial Z}{\partial X}\right)_Y dX + \left(\frac{\partial Z}{\partial Y}\right)_X dY \quad (1.7)$$

Dato che abbiamo assunto un buon comportamento, possiamo ricavare in principio $Y = Y(X, Z)$ e calcolare il differenziale totale

$$dY = \left(\frac{\partial Y}{\partial X}\right)_Z dX + \left(\frac{\partial Y}{\partial Z}\right)_X dZ \quad (1.8)$$

Sostituendo la prima nella seconda per dZ

$$dY = \left(\frac{\partial Y}{\partial X}\right)_Z dX + \left(\frac{\partial Y}{\partial Z}\right)_X \left[\left(\frac{\partial Z}{\partial X}\right)_Y dX + \left(\frac{\partial Z}{\partial Y}\right)_X dY \right] \quad (1.9)$$

vale a dire

$$\left[\left(\frac{\partial Y}{\partial Z}\right)_X \left(\frac{\partial Z}{\partial X}\right)_Y + \left(\frac{\partial Y}{\partial X}\right)_Z \right] dX + \left[\left(\frac{\partial Y}{\partial Z}\right)_X \left(\frac{\partial Z}{\partial Y}\right)_X - 1 \right] dY = 0 \quad (1.10)$$

poiché le variazioni di X e Y sono arbitrarie, affinché la relazione precedente sia vera, i contenuti delle parentesi quadre devono essere nulli. Quindi, ponendo a zero il coefficiente di dY

$$\left(\frac{\partial Y}{\partial Z}\right)_X \left(\frac{\partial Z}{\partial Y}\right)_X = 1 \Rightarrow \left(\frac{\partial Y}{\partial Z}\right)_X = \frac{1}{\left(\frac{\partial Z}{\partial Y}\right)_X} \quad (1.11)$$

Naturalmente si potevano scegliere come variabili indipendenti una coppia qualunque tra X , Y e Z quindi vale anche che

$$\left(\frac{\partial X}{\partial Z}\right)_Y = \frac{1}{\left(\frac{\partial Z}{\partial X}\right)_Y} \quad \left(\frac{\partial X}{\partial Y}\right)_Z = \frac{1}{\left(\frac{\partial Y}{\partial X}\right)_Z} \quad (1.12)$$

Se ora nell'equazione 1.10 mandiamo a zero il coefficiente di dX otteniamo

$$\left(\frac{\partial Y}{\partial Z}\right)_X \left(\frac{\partial Z}{\partial X}\right)_Y = - \left(\frac{\partial X}{\partial Y}\right)_Z \Rightarrow \left(\frac{\partial X}{\partial Y}\right)_Z \left(\frac{\partial Y}{\partial Z}\right)_X \left(\frac{\partial Z}{\partial X}\right)_Y = -1 \quad (1.13)$$

Esempio 3. Date le grandezze $\alpha = \frac{1}{V} \left(\frac{\partial V}{\partial T}\right)_p$ (coefficiente di espansione termica) e $\kappa = -\frac{1}{V} \left(\frac{\partial V}{\partial p}\right)_T$ (coefficiente di compressibilità), dimostriamo che $\left(\frac{\partial p}{\partial T}\right)_V = \frac{\alpha}{\kappa}$.

Dalla regola ciclica abbiamo che $\left(\frac{\partial p}{\partial T}\right)_V \left(\frac{\partial V}{\partial p}\right)_T \left(\frac{\partial T}{\partial V}\right)_p = -1$, quindi

$$\left(\frac{\partial p}{\partial T}\right)_V = -\frac{1}{\left(\frac{\partial V}{\partial p}\right)_T \left(\frac{\partial T}{\partial V}\right)_p} = -\frac{\left(\frac{\partial V}{\partial T}\right)_p}{\left(\frac{\partial V}{\partial p}\right)_T} = \frac{V\alpha}{V\kappa} = \frac{\alpha}{\kappa}$$

Date due funzioni $F = F(X, Y)$ e $G = G(X, Y)$, si può dimostrare la regola della "catena" (*chain rule*)

$$\left(\frac{\partial F}{\partial G}\right)_X = \left(\frac{\partial F}{\partial Y}\right)_X \left(\frac{\partial Y}{\partial G}\right)_X \quad (1.14)$$

e che

$$\left(\frac{\partial F}{\partial X}\right)_G = \left(\frac{\partial F}{\partial X}\right)_Y + \left(\frac{\partial F}{\partial Y}\right)_X \left(\frac{\partial Y}{\partial X}\right)_G \quad (1.15)$$

Consideriamo i differenziali totali di $F = F(X, Y)$ e $Y = Y(X, G)$

$$\begin{aligned} dF &= \left(\frac{\partial F}{\partial X}\right)_Y dX + \left(\frac{\partial F}{\partial Y}\right)_X dY \\ dY &= \left(\frac{\partial Y}{\partial X}\right)_G dX + \left(\frac{\partial Y}{\partial G}\right)_X dG \end{aligned} \quad (1.16)$$

Sostituendo la seconda nella prima otteniamo

$$dF = \left[\left(\frac{\partial F}{\partial X}\right)_Y + \left(\frac{\partial F}{\partial Y}\right)_X \left(\frac{\partial Y}{\partial X}\right)_G \right] dX + \left(\frac{\partial F}{\partial Y}\right)_X \left(\frac{\partial Y}{\partial G}\right)_X dG \quad (1.17)$$

ma F può anche essere considerata funzione di X e G , $F = F(X, G)$

$$dF = \left(\frac{\partial F}{\partial X}\right)_G dX + \left(\frac{\partial F}{\partial G}\right)_X dG \quad (1.18)$$

Le ultime due espressioni per dF devono essere identiche, per dX , dG arbitrarie. Quindi devono essere vere le due regole enunciate inizialmente

1.2 Differenziali lineari

Una quantità infinitesimale

$$dz = \sum_{i=1}^m M_i(x_1, \dots, x_m) dx_i \quad (1.19)$$

è detta **differenziale lineare**. In due variabili:

$$dz = M(x, y)dx + N(x, y)dy \quad (1.20)$$

Un differenziale si dice esatto se esiste una funzione F tale che $dF = dz$. In due variabili, la condizione necessaria e sufficiente perchè il differenziale sia esatto è che

$$\left(\frac{\partial M}{\partial y}\right)_x = \left(\frac{\partial N}{\partial x}\right)_y \quad (1.21)$$

mentre in generale dovrà essere $\left(\frac{\partial M_i}{\partial x_j}\right)_{x_k \neq j} = \left(\frac{\partial M_j}{\partial x_i}\right)_{x_k \neq i}$

Considerate ora una curva \mathcal{C} nel piano xy , che connetta due punti $P_{\text{in}}(x_{\text{in}}, y_{\text{in}})$ e $P_{\text{fin}}(x_{\text{fin}}, y_{\text{fin}})$. Dato il differenziale 1.20, definiamo l'integrale di linea I sulla curva come

$$I = \lim_{n \rightarrow \infty} \sum_{i=1}^n [M(x_i, y_i)\Delta x_i + N(x_i, y_i)\Delta y_i] \equiv \int_{\mathcal{C}} dz$$

dove i punti (x_i, y_i) sono i vertici di una spezzata che parte da P_{in} e arriva a P_{fin} .

Si dicono **integrali ciclici** $I = \oint_{\mathcal{C}} dz$ gli integrali di linea di un differenziale dz con un cammino di integrazione rappresentato da una curva chiusa ($P_{\text{in}} = P_{\text{fin}}$). Valgono le seguenti importanti proprietà

1. se dz è esatto $I = \oint_{\mathcal{C}} dz = 0$ per ogni curva chiusa \mathcal{C}
2. se dz è esatto $I = \int_{\mathcal{C}} dz = I' = \int_{\mathcal{C}'} dz$ se \mathcal{C} e \mathcal{C}' hanno lo stesso punto iniziale e finale (l'integrale è indipendente dal cammino)

Se il differenziale è esatto, infatti, esiste una funzione F tale che $dz = dF$, e quindi $\int_{\mathcal{C}} dF = F(x_{\text{fin}}, y_{\text{fin}}) - F(x_{\text{in}}, y_{\text{in}})$, indipendentemente dalla forma del cammino \mathcal{C} . Le proprietà dei differenziali lineari sono di fondamentale interesse per esempio, in termodinamica. Per ogni trasformazione infinitesimale (cioè cambiamento di stato di un sistema termodinamico), le variazioni infinitesimale delle varie funzioni di stato dp , dV , dU , dH , dS , dG , dA e cosívia sono differenziali esatti; il lavoro dw ed il calore dq scambiati dal sistema invece non lo sono. Per una trasformazione finita dallo stato 1 allo stato 2, la variazione finita ΔF di una funzione di stato F è ottenuta come $\Delta F = F_2 - F_1$; il lavoro ed il calore totali scambiati si devono invece calcolare dagli integrali di linea relativi al cammino specifico della trasformazione (se esiste un cammino, cioè una successione continua di punti di equilibrio del sistema!).

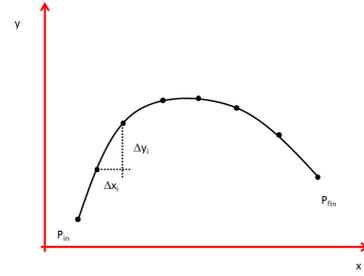


Figura 1.1: Integrale di linea

Esempio 4. La variazione infinitesima dell'energia libera di Gibbs di un sistema di n componenti è data dal differenziale lineare

$$dG = -SdT + Vdp + \sum_{i=1}^n \mu_i dn_i$$

Capitolo 2

Algebra lineare: principi

2.1 Spazi vettoriali

Iniziamo con una breve presentazione degli *spazi vettoriali*. Le definizioni che seguono sono del tutto generali: vale a dire le proprietà definite per uno spazio vettoriale non dipendono dalla natura dei suoi componenti, o *vettori*, che possono essere i vettori di uno spazio tridimensionale noti dai corsi di analisi matematica, o un insieme di funzioni e così via.

La definizione di uno spazio vettoriale V è fondata sulla precedente definizione di un insieme \mathcal{K} di quantità *scalari*, cioè un insieme di numeri. Il *corpo* \mathcal{K} gode delle proprietà:

- Se a e b appartengono a \mathcal{K} , allora $a + b$ e ab appartengono a \mathcal{K}
- 0 e 1 appartengono a \mathcal{K}
- Se a appartiene a \mathcal{K} , $-a$ appartiene a \mathcal{K} ; se a non è 0 allora a^{-1} appartiene a \mathcal{K}

Esempio 5. Sia \mathcal{N} l'insieme dei numeri naturali; si vede subito che *non* è un corpo, poiché se n appartiene a \mathcal{N} , $-n$ non è un numero naturale; analogamente per l'insieme dei numeri interi \mathcal{Z} , per un cui elemento generico z si ha che $1/z$ non appartiene a \mathcal{Z} .

In effetti si parla di spazio vettoriale V sul corpo \mathcal{K} di scalari. Il corpo \mathcal{K} è in genere dato dall'insieme dei numeri razionali \mathcal{Q} , o reali \mathcal{R} o complessi \mathcal{C} , o loro sottoinsiemi. Uno spazio vettoriale V sul corpo \mathcal{K} è dunque un insieme di oggetti che possono essere addizionati fra loro e moltiplicati per elementi di \mathcal{K} (nel seguito semplicemente *numeri*), in modo che il risultato appartenga ancora a V , con le seguenti proprietà aggiuntive:

Associativa. Se \mathbf{u} , \mathbf{v} , \mathbf{w} sono elementi di V , allora:

$$(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w}) \tag{2.1}$$

Neutro. Esiste un elemento $\mathbf{0}$ di V , tale che

$$\mathbf{0} + \mathbf{u} = \mathbf{u} + \mathbf{0} = \mathbf{u} \tag{2.2}$$

per ogni \mathbf{u} di V

Opposto. Per ogni \mathbf{u} di V esiste $-\mathbf{u}$ di V tale che

$$\mathbf{u} + (-\mathbf{u}) = (-\mathbf{u}) + \mathbf{u} = \mathbf{0} \tag{2.3}$$

Commutativa. Per ogni coppia \mathbf{u} , \mathbf{v} vale che

$$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u} \tag{2.4}$$

Proprietà in \mathcal{K} . Se a, b, c sono numeri

- $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$
- $(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$
- $(ab)\mathbf{u} = a(b\mathbf{u})$
- $1\mathbf{u} = \mathbf{u}$

Esempio 6. Alcuni esempi di spazi vettoriali:

1. lo spazio $P_N(x)$ di tutti i polinomi a coefficienti reali o complessi di grado N nella coordinata x ; lo spazio F di tutte le funzioni di variabile complessa definite in un intervallo.
2. Dato l'insieme \mathcal{R}^n di tutte le n -uple di numeri reali, dove la somma di due elementi è definita come la somma ordinata delle componenti, $(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$ e il prodotto per uno scalare come $a(x_1, \dots, x_n) = (ax_1, \dots, ax_n)$, allora \mathcal{R}^n è uno spazio vettoriale. Lo studente provi a dimostrarlo verificando i vari punti della definizione di spazio vettoriale.
3. Dato l'insieme di numeri reali positivi \mathcal{R}^+ , definiamo le seguenti operazioni: 'addizione', $x \otimes y = xy$ (cioè l'ordinaria moltiplicazione di due reali), 'moltiplicazione per un numero reale a ', $a \cdot x = x^a$ (potenza); allora si dimostra che \mathcal{R}^+ è uno spazio vettoriale rispetto alle due operazioni appena definite. Lo studente provi a dimostrarlo verificando i vari punti della definizione di spazio vettoriale.

Una somma generica di elementi di uno spazio vettoriale V ,

$$\sum_{i=1}^n a_i \mathbf{x}_i = a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2 + \dots + a_n \mathbf{x}_n \quad (2.5)$$

è detta *combinazione lineare* del set di vettori $\{\mathbf{x}_i\}$; gli n vettori sono *linearmente dipendenti* se esistono n scalari a_i non tutti nulli tali che la combinazione lineare sia zero, o *linearmente indipendenti* se non esistono. Se è possibile esprimere tutti gli elementi di V come combinazioni lineari del set $\{\mathbf{x}_i\}$, allora si dice che il set *genera* (in inglese, *span*) lo spazio vettoriale. Una *base* dello spazio vettoriale è definita come un insieme di vettori $\{\mathbf{e}_i\}$ che sia linearmente indipendente e generi lo spazio. Si può dimostrare che due basi qualunque di uno spazio, se esistono, hanno lo stesso numero di elementi. La *dimensione* n dello spazio V è il numero di elementi di una base, e può anche essere infinita. Per un vettore generico possiamo scrivere perciò:

$$\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i \quad (2.6)$$

il coefficiente x_i è la proiezione o *componente* di \mathbf{x} sull'elemento i -esimo della base. L'insieme dei coefficienti è anche detto *coordinate*.

Esempio 7. Quali sono le coordinate della funzione $f(\theta) = 3 \sin \theta + 5 \cos \theta$ in un opportuno spazio di dimensione 2? Evidentemente, se si considerano come elementi di base $\sin \theta$ e $\cos \theta$, le componenti sono 3 e 5. E quali sono le coordinate della funzione $df/d\theta$? deriando si ottiene $df/d\theta = 3 \cos \theta - 5 \sin \theta$, di componenti -5 e 3.

In generale se consideriamo lo spazio vettoriale delle funzioni $f(\theta)$ a valori reali e periodiche in $[-\pi, \pi]$, l'insieme degli elementi $\sin \theta, \sin 2\theta, \sin 3\theta$ etc. e $\cos \theta, \cos 2\theta, \cos 3\theta$ etc. costituisce una base

2.2 Prodotti scalari

Se V è uno spazio vettoriale definito su un corpo \mathcal{K} , definiamo con un *prodotto scalare* un modo per associare ad ogni coppia di elementi \mathbf{u} e \mathbf{v} di V uno scalare di \mathcal{K} , che indichiamo con $\mathbf{u} \cdot \mathbf{v}$. Valgono le seguenti proprietà

Commutativa. Per ogni coppia \mathbf{u}, \mathbf{v} vale che $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$

Distributiva. Per ogni terna $\mathbf{u}, \mathbf{v}, \mathbf{w}$ vale che $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$

Scalare. Se a appartiene a \mathcal{K} allora $(a\mathbf{u}) \cdot \mathbf{v} = a(\mathbf{u} \cdot \mathbf{v})$ e $\mathbf{u} \cdot (a\mathbf{v}) = a(\mathbf{u} \cdot \mathbf{v})$.

Due vettori \mathbf{u}, \mathbf{v} si dicono *ortogonali* o perpendicolari se $\mathbf{u} \cdot \mathbf{v} = 0$. Data una base $\mathbf{e}_1, \dots, \mathbf{e}_n$, questa si dice *ortogonale*, se per ogni coppia di indici $\mathbf{e}_i \cdot \mathbf{e}_j = 0$.

Esempio 8. Consideriamo lo spazio vettoriale delle coppie di numeri reali \mathcal{R}^2 . Si noti che tale spazio è in corrispondenza biunivoca con lo spazio dei vettori (in senso tradizionale, quantità con una intensità, direzione e verso) nel piano. Definiamo il prodotto scalare di due vettori $\mathbf{u} = (u_1, u_2)$ e $\mathbf{v} = (v_1, v_2)$ come:

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2$$

Torniamo all'insieme di tutte le funzioni periodiche a valori reali in $[-\pi, \pi]$. Definiamo il prodotto

$$f(\theta) \cdot g(\theta) = \int_{-\pi}^{\pi} d\theta f(\theta)g(\theta)$$

Lo studente può verificare facilmente che si tratta di un prodotto scalare.

Consideriamo ora nel seguito di questo paragrafo uno spazio vettoriale sul corpo \mathcal{R} dei numeri reali. Definiamo un prodotto scalare. Si dice che il prodotto è *definito positivo* se per ogni \mathbf{u} di V si ha che $\mathbf{u} \cdot \mathbf{u} \geq 0$, e per ogni \mathbf{u} non nullo, $\mathbf{u} \cdot \mathbf{u} > 0$. Per una coppia di vettori \mathbf{u} e \mathbf{v} , se \mathbf{v} non è nullo definiamo la *proiezione* di \mathbf{u} lungo \mathbf{v} come:

$$a = \frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{v} \cdot \mathbf{v}} \quad (2.7)$$

detta anche *coefficiente di Fourier*. Data un base ortogonale, si può verificare che le coordinate di un elemento generico di V rispetto alla base stessa sono i coefficienti di Fourier rispetto agli elementi della base

$$\mathbf{u} = \sum_i u_i \mathbf{e}_i, \quad u_i = \frac{\mathbf{u} \cdot \mathbf{e}_i}{\mathbf{e}_i \cdot \mathbf{e}_i} \quad (2.8)$$

Dato un vettore \mathbf{u} , Si dice *norma* la grandezza scalare $u = \sqrt{\mathbf{u} \cdot \mathbf{u}}$. Se la norma vale 1, il vettore \mathbf{u} si dice *normalizzato*, o *vettore unità*, o *versore*¹.

Esempio 9. Consideriamo l'insieme dei polinomi $P_2(x)$ di grado 2 in $[0, 1]$ a valori reali. È uno spazio vettoriale (definendo l'addizione come somma dei coefficienti)? Di che dimensione? Possiamo definire una base? Ed un prodotto scalare? Ed una base ortonormale? *Suggerimenti:* la dimensione coincide con il numero massimo di coefficienti; una base è data semplicemente da $1, x$ e \dots ; il prodotto scalare può essere definito come l'integrale da 0 ad 1 del prodotto di \dots

Consideriamo nel seguito la generalizzazione del concetto di prodotto scalare che è valido per spazi vettoriali sul corpo dei numeri reali, al caso di spazi vettoriali sul corpo dei complessi. Parleremo in questo caso di *prodotto hermitiano* tra due vettori \mathbf{u} e \mathbf{v} , con le seguenti proprietà

Commutativa. Per ogni coppia \mathbf{u}, \mathbf{v} vale che $\mathbf{u} \cdot \mathbf{v} = (\mathbf{v} \cdot \mathbf{u})^*$

¹Supponiamo che V abbia dimensione finita n . Una sua base $\mathbf{e}_1, \dots, \mathbf{e}_n$ si dice *ortonormale* se essa è ortogonale e se ogni suo elemento è un vettore unità; da una base ortogonale si può sempre ricavare una base ortonormale, dividendo ogni elemento della base per la sua norma.

Distributiva. Per ogni terna \mathbf{u} , \mathbf{v} , \mathbf{w} vale che $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$

Scalare. Se a è un numero complesso allora $(a\mathbf{u}) \cdot \mathbf{v} = a^*(\mathbf{u} \cdot \mathbf{v})$ e $\mathbf{u} \cdot (a\mathbf{v}) = a(\mathbf{u} \cdot \mathbf{v})$.

dove con a^* indichiamo il complesso coniugato di a . Tutte le definizioni precedenti relative al prodotto scalare (definito positivo, ortogonalità, basi ortogonali, etc.) sono analoghe.

Esempio 10. Ecco due esempi di prodotto hermitiano

1. Consideriamo lo spazio delle n -uple complesse, \mathcal{C}^n ; un elemento generico è $\mathbf{u} = (u_1, \dots, u_n)$ dove u_i appartiene a \mathcal{C} . La dimensione dello spazio è evidentemente n ; definiamo il prodotto hermitiano

$$\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^n u_i^* v_i$$

Il prodotto è definito positivo poiché se $\mathbf{u} \neq \mathbf{0}$, allora almeno uno dei numeri u_i è diverso da zero (complesso) e $u_i^* u_i$ è sempre positivo.

2. Consideriamo lo spazio delle funzioni $f(\theta)$ a valori complessi definite in $[-\pi, \pi]$. È uno spazio vettoriale, di dimensione infinita, e possiamo introdurre il prodotto hermitiano definito positivo:

$$f \cdot g = \int_{-\pi}^{\pi} d\theta f(\theta)^* g(\theta)$$

Non è difficile verificare che l'insieme di funzioni $\{1, e^{\pm i\theta}, e^{\pm 2i\theta}, \dots\}$ costituisce una base ortogonale del problema; se indichiamo con $f_n = e^{in\theta}$ dove n è un numero intero, la norma di ciascuna funzione di base è

$$f_n \cdot f_n = \int_{-\pi}^{\pi} d\theta e^{-in\theta} e^{in\theta} = 2\pi$$

Ogni elemento generico dello spazio può essere espresso come combinazione lineare delle funzioni di base; il coefficiente di Fourier n -esimo è il coefficiente che misura il contributo di ciascuna funzione di base:

$$f(\theta) = \sum_n c_n f_n = \sum_n c_n e^{in\theta}$$

$$c_n = \frac{f \cdot f_n}{f_n \cdot f_n} = \frac{1}{2\pi} \int_{-\pi}^{\pi} d\theta f(\theta)^* e^{in\theta}$$

2.3 Vettori e matrici

Nel corso di questa Sezione riassumeremo alcune proprietà già note allo studente dei vettori tridimensionali reali. Quindi procederemo alla definizione e allo studio delle grandezze matriciali.

Un vettore nello spazio tridimensionale ordinario è definito come una coppia ordinata di punti dello spazio. In Figura 2.1 è riportata la rappresentazione di due vettori \vec{u} e \vec{v} , definiti rispetto ad un punto comune di origine. Si deve notare che lo spazio vettoriale delle triple di numeri reali \mathcal{R}^3 è in corrispondenza biunivoca con lo spazio dei vettori ordinari, che possono perciò essere identificati da una terna di numeri reali (le *componenti* rispetto ad un sistema di riferimento, vedi Figura 2.1). Comunque la definizione dello spazio vettoriale dei vettori ordinari può essere data prescindendo dalla rappresentazione: il vettore si può identificare come una freccia orientata, con una direzione, un verso ed un'intensità. Nel seguito useremo sempre caratteri in grassetto per indicare n -ple di numeri o matrici (es. la tripla di numeri v_1, v_2, v_3 sarà \mathbf{v}); e caratteri tipografici per indicare il vettore o la matrice indipendentemente dalla sua rappresentazione (es. v). Dove non sia necessario specificare la differenza useremo caratteri in grassetto.

La somma di due vettori è descritta come un vettore ottenibile con la *regola del parallelogramma* (vedi Figura 2.1). Il prodotto scalare è ottenuto in accordo con la definizione nota di prodotto scalare, $\vec{u} \cdot \vec{v} = uv \cos \theta$, dove u e v sono il modulo dei due vettori e θ è l'angolo compreso. Alternativamente, se si considera l'equivalente spazio \mathcal{R}^3 , introducendo una base di tre versori $\vec{e}_1, \vec{e}_2, \vec{e}_3$, ed ogni vettore viene rappresentato dalle sue componenti sulla base, la somma e il prodotto scalare sono definiti

$$\mathbf{u} + \mathbf{v} = \sum_{i=1}^3 (u_i + v_i) \mathbf{e}_i \quad (2.9)$$

$$\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^3 u_i v_i \quad (2.10)$$

Altri prodotti possono essere definiti con un chiaro significato geometrico. Il *prodotto vettoriale* di \vec{u} e \vec{v} è un vettore che ha per modulo la quantità $uv \sin \theta$, per direzione la perpendicolare al piano definito da \vec{u} e \vec{v} e per verso quello ottenuto avvitando una vite destrorsa lungo la direzione definita, in modo da portare \vec{u} su \vec{v} (regola della vite o della mano destra). Il prodotto vettoriale è un esempio di *applicazione* definita in uno spazio vettoriale: ovvero un'operazione che associa elementi dello spazio ad altri elementi dello spazio. Ecco alcune utili identità:

$$\vec{a} \times \vec{b} = -\vec{b} \times \vec{a} \quad (2.11)$$

$$\vec{a} \cdot (\vec{b} \times \vec{c}) = \vec{b} \cdot (\vec{c} \times \vec{a}) = \vec{c} \cdot (\vec{a} \times \vec{b}) \quad (2.12)$$

$$\vec{a} \times (\vec{b} \times \vec{c}) = \vec{b}(\vec{a} \cdot \vec{c}) - \vec{c}(\vec{a} \cdot \vec{b}) \quad (2.13)$$

$$(\vec{a} \times \vec{b}) \cdot (\vec{c} \times \vec{d}) = (\vec{a} \cdot \vec{c})(\vec{b} \cdot \vec{d}) - (\vec{a} \cdot \vec{d})(\vec{b} \cdot \vec{c}) \quad (2.14)$$

La notazione vettoriale è di grande utilità perché permette di riassumere in modo compatto equivalenze ed equazioni in tre dimensioni, o in n dimensioni, la generalizzazione a \mathcal{R}^n essendo molto semplice. In \mathcal{R}^3 così l'equazione che definisce un piano nello spazio passante per un punto \mathbf{x}_0 è

$$(\mathbf{x} - \mathbf{x}_0) \cdot \mathbf{n} = 0 \quad (2.15)$$

dove \mathbf{n} è un versore normale al piano. L'angolo diedro tra due piani è invece dato come

$$\cos \Theta = |\mathbf{n}_1 \cdot \mathbf{n}_2| \quad (2.16)$$

Molte altre identità e definizioni della geometria analitica nello spazio tridimensionale sono note allo studente dai corsi di matematica. L'uso della notazione vettoriale è utile anche per definire *funzioni vettoriali*. Possiamo distinguere funzioni scalari, o campi scalari, che associano elementi di \mathcal{R}^3 a \mathcal{R} o funzioni vettoriali che associano elementi di \mathcal{R} a \mathcal{R}^3 . Un campo vettoriale è infine una funzione che associa elementi di \mathcal{R}^3 a \mathcal{R}^3 .

Esempio 11. Ecco alcuni esempi di campi scalari, funzioni vettoriali e campi vettoriali:

Campi scalari: la temperatura di un laboratorio, intesa come funzione delle coordinate spaziali; la funzione $f(\mathbf{x}) = |\mathbf{x}|$; la funzione $f(\mathbf{x}) = x_1^2 - x_2 x_3$; la concentrazione di cloruro di terbutile in un reattore contenente alcol terbutilico, cloro e cloruro di terbutile.

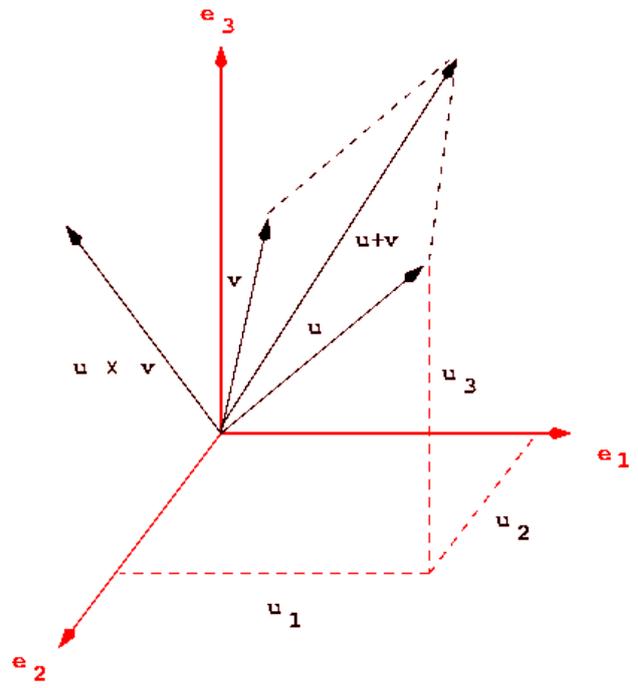


Figura 2.1: Vettori nello spazio 3D

Funzioni vettoriali: la retta $\mathbf{x}(t) = \mathbf{x}_0 + \mathbf{n}t$; la funzione $\mathbf{x}(t) = r \cos \omega t \mathbf{e}_1 + r \sin \omega t \mathbf{e}_2 + v t \mathbf{e}_3$, che rappresenta la traiettoria di una particella che si muova lungo un'elica destrorsa, con asse parallelo a \mathbf{e}_3 (vedi Figura 2.2).

Campi di vettori: la forza di gravità tra due corpi; la funzione $\mathbf{f}(\mathbf{x}) = \mathbf{x} \times \mathbf{n}$; l'orientazione delle molecole di cristallo liquido che formano il display di un orologio digitale.

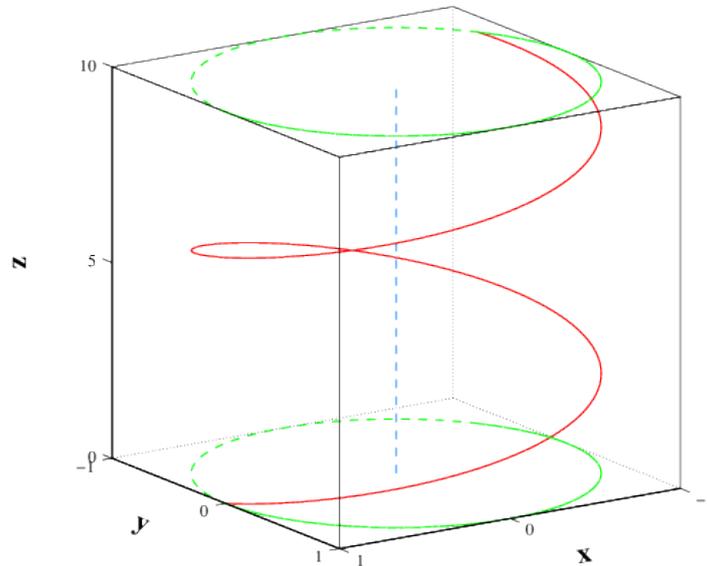


Figura 2.2: Una spirale destrorsa.

2.4 Matrici

Una matrice di mn elementi dove m ed n sono rispettivamente il numero di *righe* ed il numero di *colonne*, è definita come una tabella rettangolare di numeri, il cui elemento generico è a_{ij} (dove $1 \leq i \leq m$ e $1 \leq j \leq n$), e si indica nel modo seguente

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix} \quad (2.17)$$

o con notazioni equivalenti: p.es. doppie linee invece di parentesi.

Consideriamo l'insieme delle matrici $m \times n$. Possiamo definire la moltiplicazione per uno scalare α di una matrice \mathbf{A} come la matrice ottenuta moltiplicando ciascun elemento di \mathbf{A} per α

$$(\alpha \mathbf{A})_{ij} = \alpha a_{ij} \quad (2.18)$$

La somma di due matrici si ottiene sommando gli elementi delle matrici addende

$$(\mathbf{A} + \mathbf{B})_{ij} = a_{ij} + b_{ij} \quad (2.19)$$

La matrice nulla è formata solo da zeri

$$\mathbf{0}_{ij} = 0 \quad (2.20)$$

ed infine la matrice opposta di una matrice generica è

$$(-\mathbf{A})_{ij} = -a_{ij} \quad (2.21)$$

Si verifica facilmente che l'insieme così definito è uno spazio vettoriale di dimensione mn . Una base è costituita dalle mn matrici con elementi ovunque nulli, tranne che in una posizione

$$\mathbf{e}_{11} = \begin{pmatrix} 1 & \dots & 0 \\ 0 & \dots & 0 \end{pmatrix}, \dots, \mathbf{e}_{mn} = \begin{pmatrix} 0 & \dots & 0 \\ 0 & \dots & 1 \end{pmatrix} \quad (2.22)$$

Esempio 12. Definiamo un prodotto tra due matrici $m \times n$ come la matrice ottenuta moltiplicando gli elementi nella stessa posizione $(\mathbf{A} * \mathbf{B})_{ij} = a_{ij}b_{ij}$. L'operazione $*$ è un prodotto scalare?

Il prodotto tra matrici è un'applicazione che associa a due matrici generiche \mathbf{A} $m \times n$ e \mathbf{B} $n \times p$ una matrice \mathbf{AB} $m \times p$ (il numero di colonne della prima matrice deve essere uguale al numero di righe della seconda):

$$(\mathbf{AB})_{ij} = \sum_{k=1}^n a_{ik}b_{kj} \quad (2.23)$$

L'insieme delle n -uple di numeri, cioè gli elementi di \mathcal{R}^n o \mathcal{C}^n si possono mettere in relazione biunivoca con le matrici $1 \times n$ (matrici riga) oppure $n \times 1$ (matrici colonna). Così la moltiplicazione di una matrice $m \times n$ per un vettore n dimensionale significa in pratica la moltiplicazione per una matrice colonna $n \times 1$ ed il risultato è una matrice colonna $m \times 1$:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \dots \\ u_n \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ \dots \\ v_m \end{pmatrix} \equiv \mathbf{A}\mathbf{u} = \mathbf{v} \quad (2.24)$$

dove $v_i = \sum_{j=1}^n a_{ij}u_j$.

Esempio 13. Il prodotto tra matrici non è commutativo, ovvero $\mathbf{AB} \neq \mathbf{BA}$. Si definisce il *commutatore* come la matrice: $[\mathbf{A}, \mathbf{B}] = \mathbf{AB} - \mathbf{BA}$. Evidentemente perché il commutatore esista deve essere $m = p$; si dice che due matrici commutano se $[\mathbf{A}, \mathbf{B}] = \mathbf{0}$.

Consideriamo nel seguito solo le *matrici quadrate* $n \times n$ (nel seguito, di dimensione n), e vettori (matrici riga o colonna). Alcune matrici quadrate speciali di dimensione n sono le seguenti:

- la matrice identità

$$\mathbf{1} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \quad (2.25)$$

- le matrici scalari, $\mathbf{A} = a \mathbf{1}$
- le matrici diagonali

$$\mathbf{D} = \begin{pmatrix} d_1 & 0 & 0 & \dots & 0 \\ 0 & d_2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & d_n \end{pmatrix} \quad (2.26)$$

- le matrici tridiagonali

$$\mathbf{T}_d = \begin{pmatrix} d_1 & e_1 & 0 & \dots & 0 \\ e_1 & d_2 & 0 & \dots & 0 \\ & & \dots & & \\ 0 & 0 & 0 & \dots & d_n \end{pmatrix} \quad (2.27)$$

- le matrici triangolari (superiori od inferiori)

$$\mathbf{T} = \begin{pmatrix} t_{11} & t_{12} & t_{13} & \dots & t_{1n} \\ 0 & t_{22} & t_{23} & \dots & t_{2n} \\ & & \dots & & \\ 0 & 0 & 0 & \dots & t_{nn} \end{pmatrix} \quad (2.28)$$

- la potenza di una matrice: $\mathbf{A}^p = \mathbf{A} \cdot \dots \cdot \mathbf{A}$ (p volte)
- le matrici idempotenti: $\mathbf{P}^2 = \mathbf{P}$

In associazione con una generica matrice \mathbf{A} con elementi a_{ij} si possono definire alcune matrici strettamente collegate

- la matrice trasposta

$$\mathbf{A}^T = \begin{pmatrix} a_{11} & a_{21} & a_{31} & \dots & a_{n1} \\ a_{12} & a_{22} & a_{32} & \dots & a_{n2} \\ & & \dots & & \\ a_{1n} & a_{2n} & a_{3n} & \dots & a_{nn} \end{pmatrix} \quad (2.29)$$

- la matrice complessa coniugata

$$\mathbf{A}^* = \begin{pmatrix} a_{11}^* & a_{12}^* & a_{13}^* & \dots & a_{1n}^* \\ a_{21}^* & a_{22}^* & a_{23}^* & \dots & a_{2n}^* \\ & & \dots & & \\ a_{1n}^* & a_{2n}^* & a_{3n}^* & \dots & a_{nn}^* \end{pmatrix} \quad (2.30)$$

- la matrice aggiunta

$$\mathbf{A}^\dagger = \begin{pmatrix} a_{11}^* & a_{21}^* & a_{31}^* & \dots & a_{n1}^* \\ a_{12}^* & a_{22}^* & a_{32}^* & \dots & a_{n2}^* \\ & & \dots & & \\ a_{1n}^* & a_{2n}^* & a_{3n}^* & \dots & a_{nn}^* \end{pmatrix} \quad (2.31)$$

evidentemente $(\mathbf{A}^T)^* = (\mathbf{A}^*)^T = \mathbf{A}^\dagger$. Prima di procedere oltre consideriamo il concetto di *funzione scalare di una matrice*, ovvero una corrispondenza tra una matrice e l'insieme dei reali o complessi. Le funzioni scalari più importanti sono la *traccia*

$$\text{Tr}(\mathbf{A}) = \sum_{i=1}^n a_{ii} \quad (2.32)$$

che è la somma di tutti gli elementi diagonali di una matrice; ed il *determinante*

$$\det(\mathbf{A}) = \sum_{\mathcal{P}} (-1)^P \prod_{j=1}^n a_{j\mathcal{P}(j)} \quad (2.33)$$

di definizione un poco più complicata: si tratta della somma di $n!$ addendi, ciascun dei quali è ottenuto moltiplicando n quantità che coincidono con un elemento per ciascuna riga; vi sono appunto $n!$ possibili scelte, che coincidono con le *permutazioni* tra gli n indici $1, 2, \dots, n$; ogni permutazione è caratterizzata dal

simbolo \mathcal{P} ed è pari, $(-1)^{\mathcal{P}} = 1$ o dispari $(-1)^{\mathcal{P}} = -1$ a seconda che sia ottenuto con un numero pari o dispari di trasposizioni degli indici. In pratica il calcolo dei determinanti è effettuato applicando il seguente teorema

$$\det(\mathbf{A}) = (-1)^{1+j} a_{1j} \det(\mathbf{A}_{1j}) + \dots + (-1)^{n+j} a_{nj} \det(\mathbf{A}_{nj}) \quad (2.34)$$

cioè scegliendo una colonna della matrice (indice j) e sommando i determinanti delle sottomatrici di dimensione $n - 1$ ottenute cancellando la colonna e la riga corrispondente a ciascun elemento (e.g. \mathbf{A}_{12} è la matrice ottenuta cancellando la prima riga e la seconda colonna). Un'analoga relazione sussiste per uno sviluppo lungo una riga. Varie proprietà elementari dei determinanti sono facilmente dimostrabili:

- $\det(\mathbf{1}) = 1$
- $\det(\mathbf{D}) = \prod_{i=1}^n d_i$
- $\det(\mathbf{T}) = \prod_{i=1}^n t_{ii}$
- $\det(\mathbf{A}^T) = \det(\mathbf{A})$
- $\det(\mathbf{A}^*) = \det(\mathbf{A})^*$
- $\det(\mathbf{A}^\dagger) = \det(\mathbf{A})^*$
- $\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B})$
- $\det(\alpha \mathbf{B}) = \alpha^n \det(\mathbf{A})$

Possiamo ora definire esplicitamente la *matrice inversa* di una matrice generica come la matrice \mathbf{A}^{-1} tale che

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{1} \quad (2.35)$$

Gli elementi della matrice inversa sono calcolabili infatti come:

$$(\mathbf{A}^{-1})_{ij} = (-1)^{i+j} \frac{\det(\mathbf{A}_{ji})}{\det(\mathbf{A})} \quad (2.36)$$

quindi una matrice ammette inversa solo se il determinante non è nullo (altrimenti si dice *singolare*). Le matrici associate ad applicazioni fisiche od utilizzate nei calcoli quantomeccanici godono spesso di alcune proprietà speciali, tra le quali

- matrice simmetrica: $\mathbf{A}^T = \mathbf{A}$
- matrice hermitiana: $\mathbf{A}^\dagger = \mathbf{A}$
- matrice ortogonale: $\mathbf{A}^T = \mathbf{A}^{-1}$
- matrice unitaria: $\mathbf{A}^\dagger = \mathbf{A}^{-1}$

Una matrice reale simmetrica o ortogonale è anche hermitiana o unitaria.

Esempio 14. Nel trattamento di Hückel dei polieni lineari è necessario calcolare il determinante della matrice tridiagonale

$$\mathbf{A}^T = \begin{pmatrix} x & 1 & 0 & \dots & 0 \\ 1 & x & 1 & \dots & 0 \\ & & \dots & & \\ 0 & 0 & 0 & \dots & x \end{pmatrix}$$

Come si può procedere per ridurlo ad un polinomio?

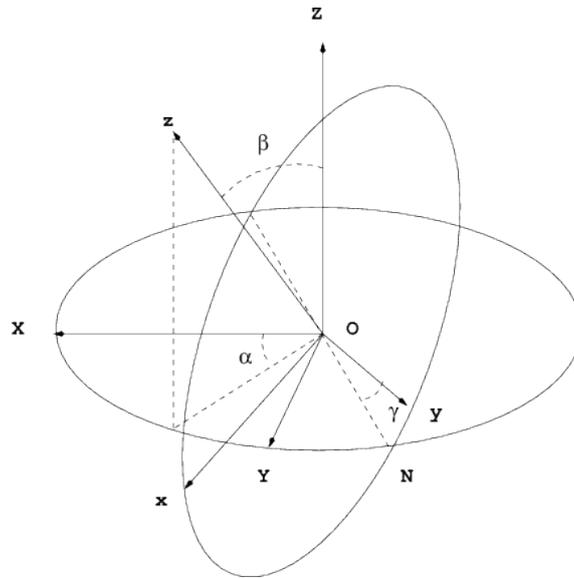


Figura 2.3: Angoli di Eulero α, β, γ per specificare la rotazione di un sistema di riferimento molecolare rispetto ad un sistema di riferimento di laboratorio.

In generale la moltiplicazione di un vettore per una matrice denota una *trasformazione* del vettore. Nello spazio tridimensionale possiamo usare analoghi geometrici per indicare una trasformazione (rotazione, traslazione etc.) del vettore stesso. In uno spazio n -dimensionale questo non è ovviamente possibile. Consideriamo per esempio il problema di descrivere la rotazione di un vettore nello spazio 3D: in generale una rotazione si può identificare mediante tre angoli, detti angoli di Eulero. Consideriamo un vettore generico \vec{u} che abbia le coordinate (X, Y, Z) in un dato sistema di riferimento fisso (per esempio solidale con il laboratorio, e definito dai versori $(\vec{X}, \vec{Y}, \vec{Z})$); vogliamo sapere quali sono le sue coordinate in un sistema di riferimento ruotato (per esempio solidale con una molecola, e definito dai versori $\vec{x}, \vec{y}, \vec{z}$). Il sistema di riferimento ruotato è definito in modo complicato ma rigoroso, tramite tre rotazioni elementari successive (vedi Figura 2.3):

- una rotazione antioraria di un angolo α intorno a \vec{Z} ; questa rotazione porta \vec{Y} nella *linea dei nodi* \vec{N} ;
- una rotazione antioraria di un angolo β intorno a \vec{N} ; questa rotazione porta \vec{Z} in \vec{z} ;
- una rotazione antioraria di un angolo γ intorno a \vec{z} ; questa rotazione porta \vec{N} in \vec{y} ;

Dopo la prima rotazione, le coordinate del vettore siano (x', y', z') ; dopo la seconda rotazione (x'', y'', z'') ; dopo la terza, nel sistema di riferimento molecolare (x, y, z) . Applicando la trigonometria, si vede che:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos \alpha & \sin \alpha & 0 \\ -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \quad (2.37)$$

$$\begin{pmatrix} x'' \\ y'' \\ z'' \end{pmatrix} = \begin{pmatrix} \cos \beta & 0 & -\sin \beta \\ 0 & 1 & 0 \\ \sin \beta & \cos \beta & 0 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \quad (2.38)$$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cos \gamma & \sin \gamma & 0 \\ -\sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x'' \\ y'' \\ z'' \end{pmatrix} \quad (2.39)$$

Moltiplicando le tre matrici di rotazione elementari si ottiene:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cos \alpha \cos \beta \cos \gamma - \sin \alpha \sin \gamma & \sin \alpha \cos \beta \cos \gamma + \cos \alpha \sin \gamma & -\sin \beta \cos \gamma \\ -\cos \alpha \cos \beta \sin \gamma - \sin \alpha \cos \gamma & -\sin \alpha \cos \beta \sin \gamma + \cos \alpha \cos \gamma & \sin \beta \sin \gamma \\ \cos \alpha \sin \beta & \sin \alpha \sin \beta & \cos \beta \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$$

(2.40)

La matrice di trasformazione è detta *matrice di Eulero*, \mathbf{E} .

Esempio 15. Qual è la trasformazione lineare che moltiplica le componenti di un vettore n -dimensionale per uno scalare a ? Scrivere la matrice che rappresenta la trasformazione. E qual è la trasformazione che scambia la componente i -esima con la componente j -esima?

Matrici e vettori sono molto utili per la rappresentazione matematica di proprietà fisiche che dipendono dalla direzione di misura. Consideriamo per esempio il momento di dipolo indotto $\vec{\mu}$ in una molecola da un campo elettrico \mathbf{E} : in generale si dovrà scrivere

$$\vec{m} = \mathbf{a}\vec{E} \quad (2.41)$$

dove \vec{m} e \vec{E} sono vettori tridimensionali, mentre \mathbf{a} è un *tensore cartesiano del secondo ordine* , cioè una proprietà rappresentabile da una matrice 3×3 (Figura 2.4). Se la nostra molecola ha una simmetria cilindrica, e si esprimono tutte le grandezze vettoriali in un sistema di riferimento solidale con la molecola ed avente l'asse z diretto lungo l'asse principale della molecola, allora la matrice di polarizzabilità assume una forma particolarmente semplice:

$$\mathbf{a} = \begin{pmatrix} a_{\perp} & 0 & 0 \\ 0 & a_{\perp} & 0 \\ 0 & 0 & a_{\parallel} \end{pmatrix} \quad (2.42)$$

In generale, le proprietà fisiche rappresentabili da matrici 3×3 sono dette proprietà tensoriali, e si può

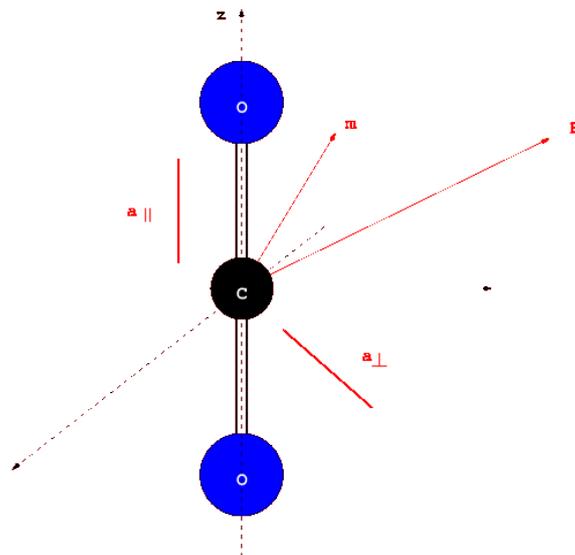


Figura 2.4: Schema del momento dipolare *indotto* in una molecola di CO_2 da un campo elettrico esterno.

sempre identificare un sistema di riferimento (di solito solidale con la molecola), tale che siano rappresentate da matrici diagonali (i valori diagonali si chiamano allora *valori principali*). Le grandezze tensoriali in Tabella (2.1) sono sempre rappresentabili come matrici reali simmetriche. Se i tre valori principali sono uguali, la rappresentazione della proprietà tensoriale è la stessa in qualunque sistema di riferimento, e si parla di proprietà *scalare*.

Esempio 16. Consideriamo il seguente problema; è data una molecola del complesso $\text{Cu}(\text{NH}_3)_6^{2+}$. Calcoliamo il tensore di inerzia nel sistema di riferimento indicato in Figura 2.5 dai versori $\vec{x}, \vec{y}, \vec{z}$. Dal corso di Fisica I è noto che per un sistema di punti materiali con masse m_i e coordinate $\mathbf{r}_i = (x_i, y_i, z_i)$, il momento

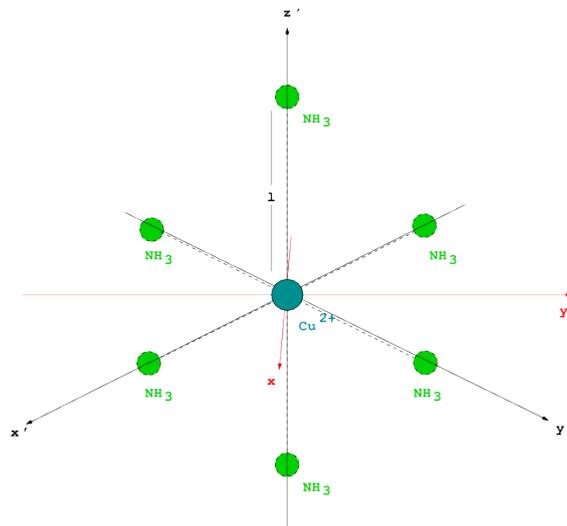
Relazione tensoriale	\vec{y}	\mathbf{A}	\vec{x}
Momento angolare = tensore di inerzia \times velocità angolare	\vec{L}	\mathbf{I}	$\vec{\omega}$
Momento di dipolo = polarizzabilità \times campo elettrico	\vec{m}	\mathbf{a}	\vec{E}
Induzione elettrica = tensore dielettrico \times campo elettrico	\vec{D}	ϵ	\vec{E}
Magnetizzazione = suscettibilità magnetica \times campo magnetico	\vec{M}	χ	\vec{H}
Velocità della luce = indice di rifrazione \times velocità nel vuoto	\vec{v}	\mathbf{n}	\vec{v}_0

Tabella 2.1: Relazioni tensoriali nella forma generica $\vec{y} = \mathbf{A}\vec{x}$

di inerzia ha componenti $I_{11} = \sum_i m_i(r_i^2 - x_i^2)$, $I_{12} = -\sum_i m_i x_i y_i$ e così via. Nel sistema di riferimento indicato in Figura 2.5 si ha perciò:

$$\mathbf{I} = 4m_{\text{NH}_3}l^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Quindi il sistema di riferimento $\vec{x}, \vec{y}, \vec{z}$ è tale che il tensore di inerzia è già in forma diagonale. Lo studente può verificare facilmente che anche il sistema ruotato $\vec{x}', \vec{y}', \vec{z}'$ gode di questa proprietà).

Figura 2.5: Calcolo del momento di inerzia del complesso $\text{Cu}(\text{NH}_3)_6^{2+}$.

2.5 Sistemi lineari

Dopo aver discusso le proprietà principali delle matrici e dei vettori, riassumiamo brevemente alcune applicazioni a problemi matematici che ricorrono sistematicamente nello studio della chimica (e non solo della chimica quantistica).

La soluzione dei sistemi lineari di equazioni si presta molto bene ad una implementazione in termini di matrici e vettori. Consideriamo solo il caso di n equazioni lineari in n incognite. Possiamo scrivere un sistema lineare in forma estesa

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n &= b_n \end{aligned} \tag{2.43}$$

o in forma compatta $\mathbf{Ax} = \mathbf{b}$; la soluzione si ottiene semplicemente nella forma $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$. Naturalmente la soluzione esiste solo se la matrice \mathbf{A} è invertibile, cioè se il suo determinante non è nullo.

Esempio 17. 100 ml di una soluzione di iodio I^- e Cl^- titola 100 ml 0.1 N di $AgNO_3$. 100 ml della stessa soluzione, trattati con un eccesso di $AgNO_3$, danno 1.5 g di precipitato. Calcolare la concentrazione di iodio e cloro nella soluzione.

Di interesse sono anche i *sistemi omogenei*:

$$\mathbf{Ax} = \mathbf{0} \quad (2.44)$$

in cui $\mathbf{b} = \mathbf{0}$; se \mathbf{A} non è singolare, l'unica soluzione possibile è $\mathbf{A}^{-1}\mathbf{0} = \mathbf{0}$, cioè l' n -upla nulla. Se \mathbf{A} è singolare, $\det(\mathbf{A}) = 0$, l'equazione ammette infinite soluzioni, multiple della soluzione generica:

$$x_i = (-1)^{i+j} \det(\mathbf{A}_{ji}) \quad (2.45)$$

dove j è un indice di riga qualunque.

2.6 Trasformazioni

Come un vettore tridimensionale ha diverse componenti in un diverso sistema di riferimento (o, nel linguaggio degli spazi vettoriali, in una diversa base), così una matrice ha diverse componenti in un diverso sistema di riferimento.

Nella Sezione precedente era stato discusso brevemente il concetto di rotazione di un sistema di riferimento; inoltre nell'esempio riferito alla polarizzabilità, il momento di dipolo, il campo elettrico e la polarizzabilità erano stati espressi implicitamente in un sistema di riferimento 'conveniente', che permetteva di esprimere la polarizzabilità in una forma semplice. In generale, qual è la relazione tra la trasformazione di un sistema di riferimento e le componenti di un vettore o di una matrice? Ancora una volta si tratta di un problema facilmente descrivibile nel linguaggio degli spazi vettoriali. di n -uple di numeri. Consideriamo in generale una base e_i (per esempio, nel caso della rotazione tridimensionale, la base di versori che identifica gli assi del laboratorio). Un vettore generico \mathbf{v} ha componenti v_1, \dots, v_n nella base, cioè l' n -upla \mathbf{v} . Introduciamo un secondo sistema di riferimento, cioè una seconda base e'_i , in cui il vettore \mathbf{v} ha componenti v'_1, \dots, v'_n , o \mathbf{v}' . La relazione tra le componenti del vettore nei due sistemi di riferimento è in generale esprimibile come una trasformazione lineare

$$\mathbf{v} = \mathbf{T}\mathbf{v}' \quad (2.46)$$

Nell'esempio della rotazione del sistema di riferimento \mathbf{T} era l'inversa della matrice di Eulero, che coincide con la trasposta della matrice di Eulero (che è ortogonale), $\mathbf{T} = \mathbf{E}^T$. Consideriamo ora una relazione tra due vettori generici, \mathbf{u} e \mathbf{v} , nella forma

$$\mathbf{u} = \mathbf{A}\mathbf{v} \quad (2.47)$$

Nelle due basi e_i e e'_i (o nei due sistemi di riferimento, se si preferisce), si possono scrivere due distinte relazioni matriciali:

$$\mathbf{u} = \mathbf{A}\mathbf{v} \quad (2.48)$$

$$\mathbf{u}' = \mathbf{A}'\mathbf{v}' \quad (2.49)$$

Sostituendo nella prima equazione le espressioni $\mathbf{u} = \mathbf{T}\mathbf{u}'$ e $\mathbf{v} = \mathbf{T}\mathbf{v}'$ si ottiene

$$\mathbf{T}\mathbf{u}' = \mathbf{A}\mathbf{T}\mathbf{v}' \rightarrow \mathbf{u}' = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}\mathbf{v}' \quad (2.50)$$

da cui si vede che la matrice trasformata è

$$\mathbf{A}' = \mathbf{T}^{-1}\mathbf{A}\mathbf{T} \quad (2.51)$$

La precedente equazione è detta anche *trasformazione di similitudine*; se \mathbf{T} è unitaria o ortogonale, parleremo di trasformazione unitaria o ortogonale.

2.7 Problemi agli autovalori

Un problema molto comune che ricorre nelle applicazioni fisiche è il calcolo degli *autovettori* e degli *autovet-tori* di una matrice, o della risoluzione del problema agli autovalori relativo ad una matrice. Un'autovettore \mathbf{x} di una matrice \mathbf{A} è un vettore tale che:

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x} \quad (2.52)$$

cioè tale che la moltiplicazione per \mathbf{A} non ne modifica altro che la lunghezza, di un coefficiente scalare λ , detto autovalore. L'insieme degli autovalori è detto *spettro* della matrice. La n -upla nulla, $\mathbf{0}$ è sempre una soluzione, triviale; dato un autovettore \mathbf{x} , ogni vettore $a\mathbf{x}$ è pure un autovettore. E' possibile avere più di un autovettore linearmente indipendente corrispondente ad uno stesso autovalore. Il numero di autovettori linearmente indipendenti che condividono lo stesso autovalore è detto *molteplicità* dell'autovalore; se la molteplicità è diversa da uno, l'autovalore è degenere. La soluzione del problema agli autovalori è impostabile come il calcolo degli zeri di un polinomio di grado n . Infatti l'equazione agli autovalori si può scrivere come un sistema omogeneo

$$(\mathbf{A} - \lambda\mathbf{1})\mathbf{x} = \mathbf{0} \quad (2.53)$$

che ammette soluzioni non triviali solo se la matrice è singolare:

$$\det(\mathbf{A} - \lambda\mathbf{1}) = 0 \quad (2.54)$$

L'equazione precedente è detta *equazione secolare*. Se si è in grado di trovare gli autovalori e i corrispondenti autovettori, i quali possono essere eventualmente normalizzati per evitare l'arbitrarietà di un fattore scalare, possiamo costruire una matrice diagonale $\mathbf{\Lambda}$ con gli autovalori in diagonale ed una matrice \mathbf{X} la cui colonna i -esima è l'autovettore i -esimo. L'insieme delle n equazioni lineari diviene dunque

$$\mathbf{A}\mathbf{X} = \mathbf{X}\mathbf{\Lambda} \quad (2.55)$$

od anche

$$\mathbf{\Lambda} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X} \quad (2.56)$$

La matrice degli autovettori è dunque una matrice che trasforma \mathbf{A} secondo una trasformazione di similitudine che la rende diagonale: se due matrici generiche \mathbf{A} , \mathbf{B} *commutano*, $[\mathbf{A}, \mathbf{B}] = 0$, allora si possono scegliere combinazioni lineari di autovettori di \mathbf{A} che sono anche autovettori di \mathbf{B} o viceversa. In altre parole, due matrici con commutatore nullo si possono diagonalizzare simultaneamente con la stessa trasformazione; questa affermazione si generalizza anche ad un numero generico di matrici: se $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_n$ commutano fra loro, possono essere diagonalizzate simultaneamente.

Approfondimento 1 . Le definizioni precedenti sono generali, ma si può essere un po' più specifici. Per una matrice generica \mathbf{A} , la scrittura $\mathbf{A}\mathbf{x}_R = \lambda_R\mathbf{x}_R$ definisce gli *autovettori destri* della matrice; la scrittura

$$\mathbf{x}_L^T\mathbf{A} = \lambda_L\mathbf{x}_L^T \quad (2.57)$$

definisce gli autovalori sinistri della matrice. Si trova subito che gli autovettori sinistri della matrice \mathbf{A} sono i trasposti degli autovettori destri della trasposta di \mathbf{A} . Poichè gli autovalori destri di una matrice si trovano dall'equazione secolare, eq. 2.54, ed il determinante di una matrice è uguale al determinante della sua trasposta, si può dire che gli autovalori destri e sinistri sono uguali, e li indicheremo nel seguito solo con λ .

Formiamo ora le matrici \mathbf{X}_R (le cui colonne sono gli autovettori destri; è la matrice \mathbf{X} definita poco sopra) e \mathbf{X}_L (le cui righe sono gli autovettori sinistri). Abbiamo che $\mathbf{A}\mathbf{X}_R = \mathbf{X}_R\mathbf{\Lambda}$ e $\mathbf{X}_L\mathbf{A} = \mathbf{\Lambda}\mathbf{X}_L$; pre-moltiplicando la prima di queste uguaglianze per \mathbf{X}_L , e post-moltiplicando la seconda per \mathbf{X}_R si ottiene $\mathbf{X}_L\mathbf{A}\mathbf{X}_R = \mathbf{X}_L\mathbf{X}_R\mathbf{\Lambda} = \mathbf{\Lambda}\mathbf{X}_L\mathbf{X}_R$. Quindi possiamo dire che la matrice $\mathbf{X}_L\mathbf{X}_R$ commuta con la matrice diagonale $\mathbf{\Lambda}$. Si possono avere tre casi:

1. non sono presenti autovalori degeneri; ad ogni autovalore è associata una coppia di autovalori destro e sinistro, **ortogonali** e quindi, scegliendo una opportuna costante di normalizzazione, ortonormali
2. sono presenti autovalori degeneri, ma per ogni autovalore degenero è possibile costruire un sottospazio di autovettori destri e sinistri ortonormali, come succede per tutti i casi di interesse in ambito chimico, a cui d'ora in poi ci limiteremo (matrici reali simmetriche, complesse hermitiane o complesse simmetriche)
3. sono presenti autovalori degeneri, ma non è possibile la costruzione di autovettori ortonormali

Tralasciamo il caso 3., in generale avremo quindi $\mathbf{X}_L \mathbf{X}_R = \mathbf{1}$, quindi *la matrice degli autovettori sinistri è l'inversa della matrice degli autovettori destri*. D'ora in poi quindi ignoreremo gli autovettori sinistri e tralascieremo di indicare l'indice R al piede della matrice \mathbf{X} .

Esempio 18. Nella teoria MO, basata sull'approssimazione di Huckel, per calcolare gli orbitali molecolari π dell'alchene 1,4-butadiene è necessario diagonalizzare la seguente matrice

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Il determinante secolare è dato da

$$\lambda^4 - 3\lambda^2 + 1 = 0$$

Lo studente verifichi che è un'equazione biquadratica con le seguenti soluzioni

$$\lambda_1 = -\sqrt{\frac{3+\sqrt{5}}{2}} \quad \lambda_2 = -\sqrt{\frac{3-\sqrt{5}}{2}} \quad \lambda_3 = \sqrt{\frac{3-\sqrt{5}}{2}} \quad \lambda_4 = \sqrt{\frac{3+\sqrt{5}}{2}}$$

Nelle applicazioni quantomeccaniche, i problemi agli autovalori coinvolgono sempre matrici hermitiane. Le matrici hermitiane \mathbf{H} godono delle seguenti importanti proprietà

- gli autovalori di \mathbf{H} sono reali
- gli autovettori corrispondenti ad autovalori distinti sono ortogonali

Poiché gli autovettori corrispondenti ad un autovalore degenero costituiscono comunque un set linearmente indipendente, si può trovare una trasformazione lineare che li trasforma in un set ortogonale. Di conseguenza, il set complessivo di autovettori normalizzati di una matrice hermitiana costituisce una base ortonormale, tale che per una coppia qualunque

$$\mathbf{x}_i^\dagger \mathbf{x}_j = \delta_{ij} \tag{2.58}$$

dove δ_{ij} è il simbolo di Kronecker (0 se $i \neq j$, 1 se $i = j$). In definitiva, se \mathbf{X} è la matrice degli autovettori di \mathbf{H} , e $\mathbf{\Lambda}$ è la matrice diagonale degli autovalori, si ha che:

$$\mathbf{X}^\dagger \mathbf{H} \mathbf{X} = \mathbf{\Lambda} \tag{2.59}$$

dove \mathbf{X} è unitaria e $\mathbf{\Lambda}$ è reale.

2.8 Funzioni di matrici

Concludiamo questo capitolo introducendo il concetto di funzione di matrice. Data una funzione continua $f(x)$ nella variabile x definita dalla serie convergente

$$f(x) = \sum_{n=0}^{\infty} c_n x^n \quad (2.60)$$

si definisce la funzione della matrice \mathbf{A} come la matrice ottenuta dalla serie

$$f(\mathbf{A}) = \sum_{n=0}^{\infty} c_n \mathbf{A}^n \quad (2.61)$$

Esempio 19. Alcune funzioni di matrici:

- Polinomio in \mathbf{A} : $f(\mathbf{A}) = \sum_{n=0}^N c_n \mathbf{A}^n$
- Esponenziale in \mathbf{A} : $f(\mathbf{A}) = \exp(\mathbf{A}) = \sum_{n=0}^{\infty} \mathbf{A}^n / n!$
- Coseno in \mathbf{A} : $f(\mathbf{A}) = \cos(\mathbf{A}) = \sum_{n=0}^{\infty} (-1)^n \mathbf{A}^{2n} / (2n)!$
- Seno in \mathbf{A} : $f(\mathbf{A}) = \sin(\mathbf{A}) = \sum_{n=0}^{\infty} (-1)^n \mathbf{A}^{2n+1} / (2n+1)!$
- Inversa di $\mathbf{1} + \mathbf{A}$: $f(\mathbf{A}) = (\mathbf{1} + \mathbf{A})^{-1} = \sum_{n=0}^{\infty} (-1)^n \mathbf{A}^n$; con la condizione $\lim_{n \rightarrow \infty} \mathbf{A}^n = \mathbf{0}$
- Inversa della radice di $\mathbf{1} + \mathbf{A}$: $f(\mathbf{A}) = (\mathbf{1} + \mathbf{A})^{-1/2} = \sum_{n=0}^{\infty} (-1)^n (2n)! \mathbf{A}^n / 2^{2n} n!^2$; con la condizione $\lim_{n \rightarrow \infty} (2n)! \mathbf{A}^n / 2^{2n} n!^2 = \mathbf{0}$

Dato lo sviluppo in serie di una funzione, la valutazione della funzione di matrice si persegue utilizzando il concetto di diagonalizzazione. Sia data la funzione generica (2.61) in \mathbf{A} . Siano note la matrice degli autovalori $\mathbf{\Lambda}$ e la matrice degli autovettori \mathbf{X} di \mathbf{A} , allora

$$f(\mathbf{A}) = \sum_{n=0}^{\infty} c_n \mathbf{A}^n = \sum_{n=0}^{\infty} c_n (\mathbf{X} \mathbf{\Lambda} \mathbf{X}^{-1})^n \quad (2.62)$$

e lo studente può verificare per ispezione che $(\mathbf{X} \mathbf{\Lambda} \mathbf{X}^{-1})^n = \mathbf{X} \mathbf{\Lambda}^n \mathbf{X}^{-1}$ da cui segue

$$f(\mathbf{A}) = \mathbf{X} \left(\sum_{n=0}^{\infty} c_n \mathbf{\Lambda}^n \right) \mathbf{X}^{-1} = \mathbf{X} f(\mathbf{\Lambda}) \mathbf{X}^{-1} \quad (2.63)$$

e $f(\mathbf{\Lambda})$ è immediatamente valutabile come una matrice diagonale il cui i -esimo elemento diagonale è $f(\lambda_i)$.

Esempio 20. È data la matrice 2×2 :

$$\mathbf{A} = \begin{pmatrix} \sqrt{2} & 1 \\ 1 & \sqrt{2} \end{pmatrix}$$

Valutare $\exp(\mathbf{A})$. Le matrici $\mathbf{\Lambda}$ e \mathbf{X} (gli autovettori non sono normalizzati) sono:

$$\mathbf{\Lambda} = \begin{pmatrix} \sqrt{2} + 1 & 0 \\ 0 & \sqrt{2} - 1 \end{pmatrix} \quad \mathbf{X} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

Quindi $\exp(\mathbf{A})$ è data da:

$$\exp(\mathbf{A}) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \exp(\sqrt{2} + 1) & 0 \\ 0 & \exp(\sqrt{2} - 1) \end{pmatrix} \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix}$$

ed infine

$$\mathbf{A} = \begin{pmatrix} \exp(\sqrt{2} + 1) + \exp(\sqrt{2} - 1) & \exp(\sqrt{2} + 1) - \exp(\sqrt{2} - 1) \\ \exp(\sqrt{2} + 1) - \exp(\sqrt{2} - 1) & \exp(\sqrt{2} + 1) + \exp(\sqrt{2} - 1) \end{pmatrix} / 2$$

La soluzione di un sistema di equazioni lineari differenziali ordinarie può essere espressa formalmente usando la funzione esponenziale della matrice dei coefficienti.

Esempio 21. La soluzione di un sistema di equazioni lineari differenziali ordinarie per n specie chimiche aventi concentrazioni al tempo t date da $c_1(t), c_2(t), \dots, c_n(t)$ si presta molto bene ad evidenziare l'utilità delle funzioni di matrici. Il sistema si può scrivere in forma compatta come:

$$\frac{d\mathbf{c}(t)}{dt} = -\mathbf{K}\mathbf{c}(t)$$

dove \mathbf{K} è una matrice di coefficienti costanti. Le condizioni iniziali, che specificano la concentrazione al tempo iniziale, sono $\mathbf{c}(0) = \mathbf{c}_0$. La soluzione del sistema è semplicemente

$$\mathbf{c}(t) = \exp(-\mathbf{K}t)\mathbf{c}_0$$

che si può scrivere anche come: $\mathbf{c}(t) = \mathbf{X} \exp(-\mathbf{\Lambda}t)\mathbf{X}^{-1}\mathbf{c}_0$. Quindi il problema è equivalente alla diagonalizzazione della matrice \mathbf{K} .

2.9 Spazi di Hilbert

L'estensione dei concetti introdotti per spazi vettoriali di dimensioni finite è abbastanza ovvia. Un concetto completamente nuovo che deve però essere introdotto è quello di *completezza*, essenziale per il calcolo di autovalori ed autovettori e per il trattamento computazionale di matrici di dimensioni formalmente infinite. Gli spazi vettoriali di dimensioni infinite per i quali si possa definire la proprietà addizionale di completezza sono anche detti *spazi di Hilbert* e sono di importanza fondamentale per lo studio della meccanica quantistica. Utilizziamo nel seguito la *notazione di Dirac* o notazione *bracket* tipica delle applicazioni quanto-meccaniche. Un vettore di uno spazio Hilbertiano è rappresentato da un *ket* $|\phi\rangle$.

Consideriamo uno spazio vettoriale di dimensione infinita, con un prodotto hermitiano. Per due vettori generici, $|\phi\rangle$ e $|\psi\rangle$, il prodotto è $\langle\phi|\psi\rangle$ (dove il primo vettore appare in forma *bra*). L'opposto di $|\phi\rangle$ è $-|\phi\rangle$, l'elemento nullo è $|0\rangle$ e così via. Il modulo di un vettore è $\langle\phi|\phi\rangle$. Sia ora data una successione di vettori $|\phi_1\rangle, \dots, |\phi_n\rangle, \dots$ ed un *vettore limite* $|\phi\rangle$. La successione converge² al vettore limite se *per ogni numero ϵ positivo piccolo a piacere* esiste un intero $N(\epsilon)$ tale che per $n > N(\epsilon)$ vale che

$$\langle\delta\phi_n|\delta\phi_n\rangle < \epsilon \tag{2.64}$$

dove $|\delta\phi_n\rangle = |\phi_n\rangle - |\phi\rangle$. Si dice che lo spazio vettoriale è uno spazio di Hilbert, di cui la successione ϕ_n è una base, se la successione converge a $|\phi\rangle$ per ogni $|\phi\rangle$. Nel seguito considereremo solo spazi hilbertiani.

Esempio 22. Lo spazio delle funzioni periodiche a valori complessi è uno spazio Hilbertiano. Qual è una sua base completa?

Una volta definita una base, possiamo rappresentare ogni vettore nella base definendone le componenti. Se pensiamo agli elementi dello spazio hilbertiano come a delle funzioni generiche, queste possono essere rappresentate da vettori colonna di dimensioni infinite. Un'applicazione ovvero una corrispondenza che mette in relazione un vettore con un altro vettore può essere espressa tramite il concetto, fondamentale, di *operatore*

$$\mathcal{O}|\phi\rangle = |\psi\rangle \tag{2.65}$$

²In senso *forte*: esiste anche una definizione di convergenza *debole*, più generale.

Nell'equazione precedente l'operatore \mathcal{O} agisce sul vettore $|\phi\rangle$ trasformandolo nel vettore $|\psi\rangle$. Sia ora data una base ortonormale $\{|i\rangle\}$; si ha che $\langle i|i\rangle = \delta_{ij}$. Un vettore generico è

$$|\phi\rangle = \sum_j \phi_j |j\rangle \quad (2.66)$$

sostituendo nell'equazione (2.65) e premoltiplicando per un generico $|i\rangle$:

$$\sum_j \langle i|\mathcal{O}|j\rangle \phi_j = \psi_j \quad (2.67)$$

che si può scrivere in termini matriciali come

$$\mathbf{O}\phi = \psi \quad (2.68)$$

La rappresentazione di un operatore è una matrice di dimensioni infinite.

2.10 Operatori lineari

Nella meccanica quantistica sono impiegati solo gli *operatori lineari* che godono della proprietà

$$\mathcal{O}(a|\phi\rangle + b|\psi\rangle) = a\mathcal{O}|\phi\rangle + b\mathcal{O}|\psi\rangle \quad (2.69)$$

dove \mathcal{O} è un operatore lineare, a e b due scalari generici. Tutti i concetti definiti in precedenza per le matrici possono essere generalizzati per gli operatori

- operatore unità: $\hat{1} = \sum_i |i\rangle\langle i|$
- operatore aggiunto di un operatore \mathcal{O} : $\langle \psi|\mathcal{O}^\dagger|\phi\rangle^* = \langle \phi|\mathcal{O}|\psi\rangle$
- commutatore di due operatori: $[\mathcal{O}_1, \mathcal{O}_2] = \mathcal{O}_1\mathcal{O}_2 - \mathcal{O}_2\mathcal{O}_1$

e così via: per esempio, un'operatore \mathcal{O} si dice hermitiano se $\mathcal{O}^\dagger = \mathcal{O}$.

Esempio 23. Si dimostri, usando le proprietà degli operatori lineari, dei prodotti hermitiani e le definizioni precedenti, che

$$(a_1\mathcal{O}_1 + a_2\mathcal{O}_2)^\dagger = a_1^*\mathcal{O}_1^\dagger + a_2^*\mathcal{O}_2^\dagger$$

$$(\mathcal{O}_1\mathcal{O}_2)^\dagger = \mathcal{O}_2^\dagger\mathcal{O}_1^\dagger$$

Il problema agli autovalori per un operatore si pone analogamente a quello di una matrice di dimensioni finite:

$$\mathcal{O}|\phi\rangle = \lambda|\phi\rangle \quad (2.70)$$

anzi, *coincide* con un problema matriciale a dimensioni infinite una volta che si sia introdotto una base:

$$\mathbf{O}\phi = \lambda\phi \quad (2.71)$$

Valgono per gli operatori hermitiani le medesime proprietà introdotte per le matrici hermitiane: gli autovalori sono reali, autovettori (o *autofunzioni*) corrispondenti ad autovalori distinti sono ortogonali; l'operatore di trasformazione è unitario. Nella notazione di Dirac un autovettore con autovalore λ si indica con $|\lambda\rangle$, ovvero $\mathcal{O}|\lambda\rangle = \lambda|\lambda\rangle$.

2.11 Postulati della quantomeccanica in una terminologia algebrica

Formuliamo i principi generali della meccanica quantistica ³ utilizzando la simbologia appena introdotta.

Principio I: lo stato fisico di un sistema isolato è descritto da un vettore $|\phi\rangle$ di un appropriato spazio hilbertiano.

Principio II: una grandezza fisica *osservabile* è rappresentato da un operatore lineare hermitiano \hat{F} .

Principio III: i risultati possibili della misura di un'osservabile \mathcal{F} sono gli autovalori f_i di \mathcal{F} :

$$\mathcal{F}|f_i\rangle = f_i|f_i\rangle \quad (2.72)$$

se $|\phi\rangle = |f_i\rangle$, il sistema è in un autostato di \mathcal{F} e la misura dà con certezza f_i .

Principio IV: il *valore d'attesa* di un'osservabile \mathcal{F} è definito come

$$\bar{F} = \frac{\langle\phi|\mathcal{F}|\phi\rangle}{\langle\phi|\phi\rangle} \quad (2.73)$$

dove \bar{F} è il risultato medio di una serie di misure di \mathcal{F}

Principio V: l'operatore relativo ad un'osservabile si può esprimere in funzione di operatori spostamento \hat{x} e momento \hat{p} (in una dimensione), che obbediscono al principio

$$[\hat{x}, \hat{p}] = i\hbar \quad (2.74)$$

dove $\hbar = h/2\pi = 1.054 \times 10^{-34}$ Js.

Principio VI: lo stato al tempo t di un sistema è ottenibile come la soluzione dell'equazione di Schrödinger

$$i\hbar \frac{\partial}{\partial t} |\phi(t)\rangle = \mathcal{H}|\phi(t)\rangle \quad (2.75)$$

dove \mathcal{H} è l'operatore *hamiltoniano* del sistema isolato. Gli autovalori dell'hamiltoniano sono le autoenergie del sistema:

$$\mathcal{H}|E\rangle = E|E\rangle \quad (2.76)$$

$|E\rangle$ è un autostato del sistema con energia E .

Cosa significa 'spazio *appropriato*'? Dipende dalla natura fisica del sistema. Consideriamo il caso di una particella che si muova lungo una retta: la coordinata (nel senso della meccanica classica) è x , il momento associato è p . Gli operatori quantomeccanici corrispondenti sono \hat{x} e \hat{p} . Definiamo la *base continua* $|x\rangle$ come gli autostati di \hat{x} con autovalori x :

$$\hat{x}|x\rangle = x|x\rangle \quad (2.77)$$

Nella *base continua* $|x\rangle$, un vettore $|\phi\rangle$ è rappresentato come una funzione $\phi(x)$, l'operatore spostamento \hat{x} è rappresentato dalla moltiplicazione per la funzione x , cioè $\hat{x}|\phi\rangle = x\phi(x)$ l'operatore \hat{p} è rappresentato da $\hbar i \frac{\partial}{\partial x}$. L'hamiltoniano è invece dato dall'espressione

$$\mathcal{H} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + U(x) \quad (2.78)$$

dove il primo termine rappresenta l'energia cinetica (m è la massa), il secondo l'energia potenziale.

³Non si tratta di assiomi matematici da cui dedurre una struttura matematica coerente; piuttosto sono principi fisici, esposti in modo semplificato per comprendere le applicazioni della meccanica quantistica nella sua formulazione *non-relativistica*.

2.12 Esercizi

1. Siano $\mathbf{x} \otimes_1 \mathbf{y}$ e $\mathbf{x} \otimes_2 \mathbf{y}$ due prodotti scalari definiti nello stesso spazio vettoriale. Dimostrare che anche le seguenti operazioni: $\mathbf{x} \odot_1 \mathbf{y} = \mathbf{x} \otimes_1 \mathbf{y} + \mathbf{x} \otimes_2 \mathbf{y}$ e $\mathbf{x} \odot_2 \mathbf{y} = \lambda \mathbf{x} \otimes_1 \mathbf{y} + \mu \mathbf{x} \otimes_2 \mathbf{y}$, sono prodotti scalari dove λ e μ sono numeri arbitrari non negativi e non simultaneamente nulli.
2. Definire il prodotto scalare dello spazio dei polinomi di grado $\leq n$ in modo che la base $1, t, t^2/2!, \dots, t^n/n!$ sia ortonormale.
3. Valutare il determinante di una matrice triangolare superiore di dimensione n .
4. Valutare il determinante di una matrice triangolare inferiore di dimensione n .
5. Porre in forma di polinomio di t il determinante della matrice

$$\mathbf{A} = \begin{pmatrix} -t & 0 & 0 & \dots & 0 & a_1 \\ a_2 & -t & 0 & \dots & 0 & 0 \\ 0 & a_3 & -t & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & -t & 0 \\ 0 & 0 & 0 & \dots & a_n & -t \end{pmatrix}$$

6. Per quali valori di λ il seguente sistema ammette soluzione?

$$x_1 + x_2 + \lambda x_3 = 1$$

$$x_1 + \lambda x_2 + x_3 = 1$$

$$\lambda x_1 + x_2 + x_3 = 1$$

7. Trovare l'inversa della matrice

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & \dots & 0 & a_1 \\ 0 & 0 & \dots & a_2 & 0 \\ & & \dots & & \\ a_n & 0 & \dots & 0 & 0 \end{pmatrix}$$

8. Trovare l'inversa della matrice

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ a & 1 & \dots & 0 & 0 \\ 0 & a & \dots & 0 & 0 \\ & & \dots & & \\ 0 & 0 & \dots & a & 1 \end{pmatrix}$$

9. Esprimere l'equazione secolare della seguente matrice in forma polinomiale:

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & \dots & 0 & b_1 \\ 0 & 0 & \dots & 0 & b_2 \\ & & \dots & & \\ 0 & 0 & \dots & 0 & b_n \\ c_1 & c_2 & \dots & c_n & a \end{pmatrix}$$

10. Trovare gli autovalori di una matrice triangolare superiore di dimensione n .

Capitolo 3

Algebra lineare: applicazioni

3.1 Sistemi lineari: esempi e metodi di soluzione

I metodi numerici basati sui concetti generali riassunti nel Capitolo 2 trovano moltissime applicazioni nello studio delle scienze chimiche. Gli esempi che potremo discutere in questo testo sono necessariamente limitate, ma lo studente dovrebbe comunque essere consapevole che il calcolo numerico basato sui principi dell'algebra lineare è alla base di quasi tutte le descrizioni quantitative e modellistiche della chimica fisica, analitica, inorganica e così via.

Distinguiamo due classi di applicazioni/metodi/esempi: soluzione di sistemi lineari e/o calcolo di proprietà di matrici e soluzione di problemi agli autovalori. Nella prima categoria ricordiamo per esempio i seguenti problemi

1. il bilanciamento di una reazione chimica;
2. il calcolo del numero di processi stechiometrici indipendenti in un sistema termodinamico multicomponente all'equilibrio;
3. la determinazione delle concentrazioni di equilibrio di un insieme di composti chimici in presenza di vari bilanci di massa;
4. l'ottimizzazione di un modello dipendente da un insieme di parametri, sulla base di una serie di dati sperimentali (metodo dei minimi quadrati)

Nella seconda categoria troviamo per esempio

1. il calcolo dei modi normali di una molecola poliatomiche, tipica dello studio delle proprietà vibrazionali e l'interpretazione degli spettri IR o Raman dei sistemi molecolari;
2. il calcolo dei livelli energetici e delle funzioni d'onda (orbitali molecolari) elettroniche di una molecola a partire da una base di orbitali atomici, a vari livelli di approssimazione (metodi semimpirici, metodi HF, DFT, post-HF);
3. lo studio della cinetica un sistema di reazioni chimiche del primo ordine (che sarà oggetto di discussione nel capitolo dedicato alla soluzione delle equazioni differenziali)

3.1.1 Sistema di equazioni lineari

Discutiamo ora alcune specifiche metodologie per il trattamento di un sistema di equazioni lineari. Consideriamo n variabili x_j ($j = 1, \dots, n$) definite in un campo \mathcal{K} , per esempio \mathcal{R} o \mathcal{C} ¹ e supponiamo che siano

¹Nel seguito, se non altrimenti specificato, assumeremo sempre di lavorare con numeri reali o complessi

note m relazioni lineari

$$\begin{aligned}
 a_{11}x_1 + a_{12}x_2 + \dots + a_{1j}x_j + \dots + a_{1n}x_n &= b_1 \\
 a_{21}x_1 + a_{22}x_2 + \dots + a_{2j}x_j + \dots + a_{2n}x_n &= b_2 \\
 &\dots \\
 a_{i1}x_1 + a_{i2}x_2 + \dots + a_{ij}x_j + \dots + a_{in}x_n &= b_j \\
 &\dots \\
 a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mj}x_j + \dots + a_{mn}x_n &= b_m
 \end{aligned} \tag{3.1}$$

con $i = 1, \dots, m$. Il sistema si può scrivere in modo compatto come $\mathbf{Ax} = \mathbf{b}$ dove \mathbf{A} è la matrice $m \times n$ con elemento a_{ij} , \mathbf{x} è il vettore (matrice colonna $n \times 1$ avente elemento generico x_j e \mathbf{b} è il vettore (matrice colonna $m \times 1$ avente elemento generico b_i).

Alcune definizioni utili: un sistema si dice *sottodeterminato* se $n > m$, *sovradeterminato* se $n < m$. In generale, si danno tre possibilità: il sistema lineare non ammette soluzioni, cioè non esiste alcun vettore \mathbf{x} che soddisfi la (3.2); ammette una sola soluzione, cioè esiste un unico vettore \mathbf{x} che soddisfi la (3.2); oppure ammette infinite soluzioni, ed esistono quindi un numero infinito di vettori \mathbf{x} che soddisfino la (3.2). Il teorema di Rouchè-Capelli ci permette di calcolare, dati \mathbf{A} e \mathbf{b} , in quale caso siamo.

Teorema di Rouchè-Capelli - Consideriamo la matrice (detta *associata*)

$$(\mathbf{A}|\mathbf{b}) = \begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ \dots & \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} & b_m \end{pmatrix} \tag{3.2}$$

Il sistema ammette soluzioni se e solo se il rango di $(\mathbf{A}|\mathbf{b})$ è uguale al rango di \mathbf{A} , vale a dire $\text{rk}(\mathbf{A}|\mathbf{b}) = \text{rk}(\mathbf{A})$. In questo caso, se $n = \text{rk}$ (numero di incognite uguale al rango di \mathbf{A}), esiste una sola soluzione; altrimenti le soluzioni sono infinite.

Approfondimento 2 . Il *rango* $\text{rk}(\mathbf{A})$ di una matrice generica \mathbf{A} è il numero massimo di righe (o colonne) linearmente indipendenti della matrice; il rango è evidentemente minore o uguale al minimo tra il numero di righe e il numero di colonne (per esempio, se una matrice è 3×5 , il rango sarà al massimo 3).

Si può notare che un *sistema lineare omogeneo*, cioè tale che $\mathbf{b} = \mathbf{0}$, il rango di $(\mathbf{A}|\mathbf{0})$ è sempre uguale al rango di \mathbf{A} , quindi ammette sempre almeno una soluzione (infatti $\mathbf{x} = \mathbf{0}$ è la soluzione triviale di ogni sistema omogeneo).

Se esistono soluzioni, queste formeranno uno *spazio affine* a \mathcal{K}^n di dimensione $n - \text{rk}(\mathbf{A})$: per esempio se $n - \text{rk}(\mathbf{A}) = 1$, avremo infinite soluzioni formate da n -uple di elementi di \mathcal{K} ottenute come multipli di una n -upla specifica etc. Alcuni esempi:

- per $n < m$, abbiamo un sistema sovradeterminato (numero di incognite n minore del numero di equazioni m); se esistono soluzioni (cioè $\text{rk}(\mathbf{A}|\mathbf{b}) = \text{rk}(\mathbf{A})$) possiamo avere una soluzione (nel qual caso possiamo costruire n equazioni linearmente indipendenti a partire dalle m date), oppure possiamo avere infinite soluzioni;
- per $n = m$ (numero di equazioni uguale al numero di incognite), se $\text{rk}(\mathbf{A}|\mathbf{b}) = \text{rk}(\mathbf{A}) = n$, abbiamo una sola soluzione oppure infinite soluzioni;
- per $n > m$, abbiamo un sistema sottodeterminato (numero di incognite n maggiore del numero di equazioni m); se esistono soluzioni, sono infinite.

Il caso più comune, che si incontra nella maggioranza delle applicazioni è quello in cui il numero di incognite n sia uguale al numero di equazioni m , quindi \mathbf{A} è una matrice quadrata $n \times n$. Se esistono soluzioni, cioè $\text{rk}(\mathbf{A}|\mathbf{b}) = \text{rk}(\mathbf{A})$, troviamo che

1. se $\text{rk}(\mathbf{A}) = n$, allora esiste una sola soluzione; si noti che in questo caso abbiamo una matrice quadrata il cui rango coincide con il numero di righe e colonne. In questo caso si può dimostrare che la matrice ammette un'inversa, \mathbf{A}^{-1} . La soluzione del sistema lineare si può perciò scrivere come

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} \quad (3.3)$$

questo significa che se $\mathbf{b} = \mathbf{0}$ abbiamo un sistema lineare omogeneo che ammette la sola soluzione triviale $\mathbf{0}$; se invece $\mathbf{b} \neq \mathbf{0}$, abbiamo il classico caso di un sistema lineare non omogeneo con un'unica soluzione diversa da $\mathbf{0}$, data dalla (3.3);

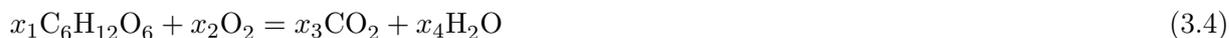
2. se $\text{rk}(\mathbf{A}) < n$, la matrice \mathbf{A} non è invertibile, ed esistono infinite soluzioni; il caso più comune è un sistema lineare omogeneo che ammette infinite soluzioni non triviali (troveremo questo concetto molto utile quando discuteremo i problemi agli autovalori).

Si noti che una matrice quadrata $n \times n$ \mathbf{A} è invertibile (esiste l'inversa \mathbf{A}^{-1}) se $\text{rk}(\mathbf{A}) = n$, il rango della matrice è uguale ad n , in questo caso $\det(\mathbf{A}) \neq 0$ il determinante della matrice è diverso da zero. In pratica, nella maggioranza delle applicazioni di interesse per un chimico, dato un sistema lineare in n incognite ed n equazioni $\mathbf{Ax} = \mathbf{b}$, che ammetta soluzioni ($\text{rk}(\mathbf{A}|\mathbf{b}) = \text{rk}(\mathbf{A})$), verificheremo prima di tutto se il sistema è omogeneo, $\mathbf{b} = \mathbf{0}$, oppure no. Nel caso in cui sia omogeneo, se il determinante della matrice dei coefficienti è diverso da zero, il sistema ammetterà solo la soluzione triviale $\mathbf{b} = \mathbf{0}$, altrimenti ammetterà infinite soluzioni non triviali. Nel caso sia non omogeneo, se il determinante della matrice dei coefficienti è diverso da zero, il sistema ammetterà la soluzione $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.

3.1.2 Bilanciamento di una reazione chimica

Cerchiamo di applicare alcuni dei concetti esposti sopra ad un caso di interesse: il bilanciamento di una o più *reazioni chimiche*, intese come *processi stechiometrici che definiscano leggi di conversione quantitative tra diverse specie chimiche*.

Esempio 24. Un caso semplice, per cominciare, la reazione del glucosio con ossigeno per dare anidride carbonica e acqua



Una reazione chimica, nella sua interpretazione più semplice (stechiometrica) è un insieme di bilanci di massa, relativi ai vari elementi che compaiono nelle specie coinvolte nella reazione. In questo caso, abbiamo quattro specie chimiche e tre elementi (carbonio, idrogeno, ossigeno), quindi tre bilanci di massa

$$\begin{aligned} 6x_1 + 0x_2 &= x_3 + 0x_4 \\ 12x_1 + 0x_2 &= 0x_3 + 2x_4 \\ 6x_1 + 2x_2 &= 2x_3 + x_4 \end{aligned} \quad (3.5)$$

che riscriviamo come $\mathbf{Ax} = \mathbf{b}$, dove

$$\mathbf{A} = \begin{pmatrix} -6 & 0 & 1 & 0 \\ -12 & 0 & 0 & 2 \\ -6 & -2 & 2 & 1 \end{pmatrix} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad (3.6)$$

Abbiamo quindi un sistema sottodeterminato (3 equazioni e 4 incognite) omogeneo. Il rango di \mathbf{A} , uguale al rango di $(\mathbf{A}|\mathbf{0})$ naturalmente, è 3 (verificate ...) quindi sono ammesse infinite soluzioni, tutte multiple di una quaterna di valori che possiamo facilmente trovare scegliendo di fissare una delle x_j e ricavando le altre; per esempio ponendo $x_1 = 1$, ricaviamo facilmente $x_3 = 6$, $x_4 = 6$ e $x_2 = 6$, quindi le infinite soluzioni sono

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = c \begin{pmatrix} 1 \\ 6 \\ 6 \\ 6 \end{pmatrix} \quad (3.7)$$

dove c è una costante arbitraria (che per motivi ovvi scegliamo reale positiva). La reazione bilanciata dunque è



o un multiplo qualunque.

Come secondo esempio consideriamo la reazione del cloruro di fosforo (V) e acqua per dare acido fosforico e acido cloridrico:



in questo caso abbiamo 4 incognite (4 specie chimiche) e 4 relazioni (4 elementi: P, Cl, O, H). Il sistema lineare è

$$\begin{aligned} x_1 + 0x_2 &= x_3 + 0x_4 \\ 5x_1 + 0x_2 &= 0x_3 + x_4 \\ 0x_1 + x_2 &= 4x_3 + 0x_4 \\ 0x_1 + 2x_2 &= 3x_3 + x_4 \end{aligned} \quad (3.10)$$

che riscriviamo come $\mathbf{Ax} = \mathbf{b}$, dove

$$\mathbf{A} = \begin{pmatrix} -1 & 0 & 1 & 0 \\ -5 & 0 & 0 & 1 \\ 0 & -1 & 4 & 0 \\ 0 & -2 & 3 & 1 \end{pmatrix} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (3.11)$$

il rango di \mathbf{A} è 3, quindi ancora una volta sono ammesse come soluzioni i multipli di una quaterna di coefficienti, che troviamo semplicemente ponendo $x_1 = 1$, da cui $x_3 = 1$, $x_4 = 5$, $x_2 = 4$. La reazione bilanciata è



Approfondimento 3 . Bilanciamento di una reazione chimica - È data una reazione che coinvolga n specie chimiche, C_j con $j = 1, \dots, n$, formate da m elementi $\text{E}_1, \dots, \text{E}_m$, $i = 1, \dots, m$.

1. Scriviamo la reazione come $\sum_{j=1}^n x_j \text{C}_j = 0$ dove $k = 1, \dots, r$; x_j è il coefficiente stechiometrico della specie j -esima.
2. Ogni specie chimica si scriverà come $\text{C}_j = \text{E}_{1a_{1j}}\text{E}_{2a_{2j}} \cdots \text{E}_{ma_{mj}}$, dove a_{ij} è il numero di atomi di E_i presenti nella specie C_j .
3. Possiamo scrivere le m relazioni lineari nella forma $\mathbf{Ax} = \mathbf{0}$
4. Dire che la reazione è bilanciabile equivale ad affermare che $n - \text{rk}(\mathbf{A}) = 1$.
5. in questo caso, fissando uno dei coefficienti stechiometrici, per esempio $x_1 = -1$, si possono ricavare i valori degli altri coefficienti per ispezione diretta (il che equivale formalmente a risolvere il sistema non omogeneo ottenuto da $n - 1$ relazioni lineari nei restanti $n - 1$ coefficienti).

La determinazione delle proprietà delle matrici come il rango e per le matrici quadrate il determinante o l'inversa (se esiste), insieme al calcolo della o delle soluzioni di una sistema lineare sono problemi intimamente collegati fra loro, la loro soluzione è solitamente affrontata mediante l'impiego di algoritmi che trasformano la matrice di interesse in una forma relativamente semplice da trattare, ma che mantenga le stesse proprietà della matrice iniziale. Se in linea di principio i vari algoritmi di trasformazione sono equivalenti, in pratica la loro implementazione numerica varia molto in termini di robustezza, stabilità, precisione numerica, risorse computazionali necessarie. Nel seguito, presentiamo una breve discussione, senza pretesa di completezza (si vedano allo scopo alcuni dei testi suggeriti nella bibliografia) dei metodi più noti, analizzando anche nel contempo alcune loro applicazioni dirette a problemi di interesse chimico.

3.1.3 Eliminazione di Gauss-Jordan

Il primo metodo che analizziamo, di interesse storico² e pedagogico, è il metodo di *eliminazione di Gauss*, strettamente collegato al metodo di *eliminazione di Gauss-Jordan* e alla *decomposizione LU*. Consideriamo il sistema lineare generico (3.2), $\mathbf{Ax} = \mathbf{b}$; il metodo si basa sostanzialmente sull'applicazione di una serie di operazioni elementari effettuate sulle righe e/o colonne della matrice $(\mathbf{A}|\mathbf{b})$: lo scambio di due righe (1), la moltiplicazione di una riga per uno scalare (2), l'aggiunta ad una riga di un'altra, moltiplicata per uno scalare (3). Tutte queste operazioni lasciano invariata la soluzione \mathbf{x} .

Consideriamo una matrice $m \times n$; si dice che la matrice è a *scalini* (in inglese *echelon matrix*) se il primo elemento diverso da zero di una riga è più a destra del primo elemento diverso da zero della riga precedente. Il primo elemento diverso da zero su ogni riga, se esiste, è il pivot della riga. Per esempio, la matrice che segue è scalini

$$\begin{array}{ccccc} 0 & 3 & 0 & -2 & 1 \\ 0 & 0 & 7 & 0 & 3 \\ 0 & 0 & 0 & 0 & -9 \end{array}$$

con pivot 3, 7 e -9. In pratica, il metodo di eliminazione di Gauss (con pivot) per la soluzione di un sistema di equazioni lineari procede come segue

1. se la prima riga ha il primo elemento nullo, si scambia con un'altra riga avente il primo elemento non nullo; per assicurare una maggiore stabilità all'algoritmo, si sceglie quella riga con il primo elemento non nullo avente il massimo valore assoluto;³ se tutte le righe hanno il primo elemento nullo, si passa al punto 3;
2. per ogni riga i con primo elemento non nullo, eccetto la prima ($i > 1$): si moltiplica la prima riga per un coefficiente scelto in maniera tale che la somma tra la prima riga e la riga i -esima abbia il primo elemento nullo, usando quindi il coefficiente $-a_{i1}/a_{11}$; si sostituisce la riga i -esima con la somma;
3. si ritorna al punto 1 considerando la sottomatrice ottenuta cancellando la prima riga e la prima colonna.

In questo modo, il sistema è ridotto ad una forma facilmente risolvibile; nel caso più comune, $m = n$, $\mathbf{b} \neq \mathbf{0}$, si ha semplicemente una matrice triangolare superiore dei coefficienti, risolvibile dal basso verso l'altro in modo triviale.

Consideriamo il seguente esempio generico:

$$\begin{array}{rcl} -x_1 + 4x_2 + x_3 & = & 1 \\ x_1 + x_2 + x_3 & = & 0 \\ 2x_1 - x_2 & = & -1 \end{array}$$

Scriviamo la matrice completa:

$$\begin{pmatrix} -1 & 4 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 2 & -1 & 0 & -1 \end{pmatrix}$$

scambiamo la prima e la terza riga

$$\begin{pmatrix} 2 & -1 & 0 & -1 \\ -1 & 4 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

²*Jiuzhang Suanshu*, I nove capitoli dell'arte matematica, II secolo; attribuito a Johann Carl Friedrich Gauss, XIX secolo

³Il metodo di eliminazione senza pivot non è stabile se implementato in un calcolatore a precisione finita; si noti inoltre che la scelta del pivot come l'elemento massimo in valore assoluto è empirica, cioè non necessariamente la migliore in assoluto

moltiplichiamo la prima riga per $1/2$ e aggiungiamola alla seconda; moltiplichiamo la prima riga per $-1/2$ e aggiungiamola alla terza

$$\begin{pmatrix} 2 & -1 & 0 & -1 \\ 0 & 7/2 & 1 & -1/2 \\ 0 & 3/2 & 1 & 3/2 \end{pmatrix}$$

moltiplichiamo la seconda riga per $-3/7$ e aggiungiamola alla terza

$$\begin{pmatrix} 2 & -1 & 0 & -1 \\ 0 & 7/2 & 1 & 1/2 \\ 0 & 0 & 4/7 & 12/7 \end{pmatrix}$$

Se la procedura viene ripetuta partendo dal basso verso l'alto, la matrice viene trasformata nella forma a scalini ridotta (con tutti i pivot unitari), e si parla di algoritmo di Gauss-Jordan. Continuando con l'esempio precedente, dividiamo la terza riga per $4/7$

$$\begin{pmatrix} 2 & -1 & 0 & -1 \\ 0 & 7/2 & 1 & 1/2 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

moltiplichiamo la terza riga per -1 e sommiamola alla seconda; moltiplichiamo la seconda riga per $4/7$

$$\begin{pmatrix} 2 & -1 & 0 & -1 \\ 0 & 1 & 0 & -5/7 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

aggiungiamo la seconda riga alla prima; moltiplichiamo la prima riga per $1/2$

$$\begin{pmatrix} 1 & 0 & 0 & -6/7 \\ 0 & 1 & 0 & -5/7 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

quindi $x_1 = -6/7$, $x_2 = -5/7$, $x_3 = 3$.

La procedura di Gauss-Jordan o Gauss può essere usata per calcolare il rango di una matrice generica $m \times n$, il determinante e (se esiste) l'inversa di una matrice $n \times n$. Per il rango, si applica la procedura di eliminazione di Gauss fino alla riduzione della matrice nella sua forma a scalini: il rango è dato dal numero di righe della matrice trasformata che non siano ad elementi tutti nulli. Il determinante di una matrice quadrata \mathbf{A} si ottiene applicando la procedura di Gauss ed ottenendo una matrice triangolare superiore, dove il determinante è dato dal prodotto degli elementi diagonali (verificate ...) (tenendo conto però che il segno del prodotto deve essere cambiato se il numero di scambi di righe effettuato è dispari). Infine l'inversa si calcola costruendo la matrice completa $(\mathbf{A}|\mathbf{1})$ di dimensioni $n \times 2n$ ed applicando l'algoritmo di Gauss-Jordan fino ad ottenere la matrice trasformata $(\mathbf{1}|\mathbf{A}^{-1})$. Quest'ultima procedura equivale a risolvere n sistemi lineari contemporaneamente.

Approfondimento 4 . L'algoritmo di Gauss-Jordan applicato ad una matrice quadrata di dimensioni $n \times n$ richiede $O(n^3)$ operazioni matematiche.

Molti problemi pratici nella chimica fisica, analitica etc. si prestano a essere risolti mediante sistemi lineari. Consideriamo l'esempio seguente.

Applicazione 1 . Abbiamo a disposizione $v^A = 10000$ litri di alcol (una soluzione di acqua e etanolo) contenenti il 96 % in peso di etanolo. Vogliamo trasformare l'alcol in vodka, cioè in alcol al 56 % in peso di etanolo, aggiungendo acqua. Alla temperatura di lavoro la densità dell'acqua è 0.9991 g ml^{-1} . I volumi parziali molari dell'acqua \bar{V}_{aq}^A e dell'etanolo \bar{V}_{et}^A nell'alcol al 96 % sono 14.61 e $58.01 \text{ ml mol}^{-1}$

rispettivamente, e i volumi parziali molari dell'acqua \bar{V}_{aq}^V e dell'etanolo \bar{V}_{et}^V sono 17.11 e 56.58 ml mol⁻¹ nella vodka. Quanti litri di vodka v^V si ottengono? Definiamo le incognite: $x_1 = v^V$, $x_2 = p_{et}^A$, $x_3 = p_{aq}^A$, $x_4 = p_{et}^V$, $x_5 = p_{aq}^V$, cioè rispettivamente il volume finale di vodka (V), il peso dell'etanolo (et) ed il peso dell'acqua (aq) nella soluzione di alcol iniziale (A), il peso dell'etanolo (et) ed il peso dell'acqua (aq) nella soluzione di vodka. Le relazioni lineari che legano queste cinque variabili sono:

$$\begin{aligned} p_{et}^A &= p_{et}^V \\ \frac{p_{et}^A}{M_{et}} \bar{V}_{et}^A + \frac{p_{aq}^A}{M_{aq}} \bar{V}_{aq}^A &= v^A \\ \frac{p_{et}^V}{M_{et}} \bar{V}_{et}^V + \frac{p_{aq}^V}{M_{aq}} \bar{V}_{aq}^V &= v^V \\ \frac{p_{et}^A}{p_{et}^A + p_{aq}^A} &= r^A \\ \frac{p_{et}^V}{p_{et}^V + p_{aq}^V} &= r^V \end{aligned}$$

dove r^A e r^V sono le percentuali in peso dell'etanolo nelle due soluzioni, M_{aq} e M_{et} sono i pesi molecolari di acqua ed etanolo. Ponendo: $a_{11} = 0$, $a_{12} = 1$, $a_{13} = 0$, $a_{14} = -1$, $a_{15} = 0$, $b_1 = 0$, $a_{21} = 0$, $a_{22} = \frac{\bar{V}_{et}^A}{M_{et}}$, $a_{23} = \frac{\bar{V}_{aq}^A}{M_{aq}}$, $a_{24} = 0$, $a_{25} = 0$, $b_2 = v^A$, $a_{31} = -1$, $a_{32} = 0$, $a_{33} = 0$, $a_{34} = \frac{\bar{V}_{et}^V}{M_{et}}$, $a_{35} = \frac{\bar{V}_{aq}^V}{M_{aq}}$, $b_3 = 0$, $a_{41} = 0$, $a_{42} = 1 - r^A$, $a_{43} = -r^A$, $a_{44} = 0$, $a_{45} = 0$, $b_4 = 0$, $a_{51} = 0$, $a_{52} = 0$, $a_{53} = 0$, $a_{54} = 1 - r^V$, $a_{55} = -r^V$, $b_5 = 0$, otteniamo un sistema lineare di cinque equazioni in cinque incognite $\mathbf{Ax} = \mathbf{b}$



Figura 3.1: Chemistry at work.

3.1.4 Decomposizioni

Quando si devono trattare matrici di dimensioni piccole, dell'ordine dell'unità, con carta e penna o mediante l'uso di un software di manipolazione algebrica, il metodo di Gauss è l'approccio naturale. Nel caso si vogliano affrontare problemi numerici di dimensioni fino al migliaio, mediante l'impiego di programmi di calcolo, il metodo di Gauss può essere comunque preferito, per la sua semplicità di programmazione. Ma per matrici di dimensioni superiori, esistono tecniche ed algoritmi molto più avanzati, che riducono drasticamente il tempo di calcolo mediante la riduzione del numero di operazioni complessive. Oggi giorno esistono librerie ottimizzate di routine di calcolo per l'algebra lineare, che possono essere direttamente incluse in codici di calcolo. Gli algoritmi su cui sono basate sono diversi, e di solito sono collegati a trasformazioni delle

matrici in forme convenienti, e particolarmente adatti ad essere implementati efficacemente con i linguaggi di programmazione moderni (C, C++, Python e così via), anche per architetture in parallelo. Ricordiamo tra tutte **LINPACK**, e soprattutto **LAPACK**, basate sulle **BLAS (Basic Linear Algebra Subprograms)**, che sono disponibili presso www.netlib.org.

Discutiamo brevemente alcuni di questi approcci⁴.

Decomposizione LU

Iniziamo dalla decomposizione LU, che si pone il problema di scomporre una matrice generica \mathbf{A} nel prodotto di una matrice triangolare inferiore a diagonale unitaria e di una matrice triangolare superiore. Da un punto di vista strettamente computazionale, è comparabile all'eliminazione di Gauss-Jordan (cioè richiede un numero $O(n^3)$ di operazioni per un problema di dimensioni n , ed anzi in pratica è equivalente all'algoritmo di Gauss. Si noti che in generale una matrice quadrata invertibile è sempre decomponibile nella forma **LU** previa permutazione di alcune righe, cioè $\mathbf{PA} = \mathbf{LU}$, dove \mathbf{P} è una matrice di permutazione, tale cioè che \mathbf{PA} dà \mathbf{A} con una o più righe scambiate. Per esempio, data la matrice $n \times n$ \mathbf{A} , cerchiamo le matrici \mathbf{L} (triangolare inferiore a diagonale unitaria) e \mathbf{U} (triangolare superiore) tali che $\mathbf{PA} = \mathbf{LU}$. Perché?

- supponiamo di volere risolvere il sistema $\mathbf{Ax} = \mathbf{b}$, che diventa equivalente a $\mathbf{Ly} = \mathbf{P}^{-1}\mathbf{b}$ e $\mathbf{Ux} = \mathbf{y}$ cioè possiamo risolvere i due sistemi di seguito in \mathbf{L} e \mathbf{U} per trovare \mathbf{x} (la soluzione dei due sistemi è molto facile per la struttura triangolare delle due matrici);
- oppure possiamo calcolare il determinante di \mathbf{A} : $\det(\mathbf{A}) = \det(\mathbf{P}^{-1})\det(\mathbf{L})\det(\mathbf{U})$; il determinante dell'inversa di \mathbf{P} vale 1 (se \mathbf{P} descrive un numero pari di permutazioni) o -1 (numero dispari di permutazioni), ovvero in generale $(-1)^P$, dove P è la parità di \mathbf{P} (numero di permutazioni); quindi $\det(\mathbf{A}) = (-1)^P \prod_{i=1}^n l_{ii} \prod_{i=1}^n u_{ii} = (-1)^P \prod_{i=1}^n u_{ii}$
- infine il calcolo dell'inversa è semplificato, infatti $\mathbf{A}^{-1} = \mathbf{U}^{-1}\mathbf{L}^{-1}\mathbf{P}^{-1} = \mathbf{U}^{-1}\mathbf{L}^{-1}\mathbf{P}$ (l'inversa di \mathbf{P} è uguale a \mathbf{P} , cioè \mathbf{P} è idempotente); il calcolo dell'inversa di una matrice triangolare è triviale.

Il processo di decomposizione LU si effettua mediante forme modificate dell'eliminazione di Gauss, come per esempio l'*algoritmo di Crout*. Senza entrare in dettagli eccessivi, che esulano dallo scopo di questo manuale, si tratta comunque di una successione di operazioni di eliminazioni di elementi, permutazioni e pivoting, che consentono, con un numero di operazioni $O(n^3)$ di decomporre una matrice invertibile generica per poi risolvere un sistema lineare o calcolare il determinante e l'inversa.

Altri metodi di decomposizione

Una matrice generica (reale) \mathbf{A} di dimensione $m \times n$ può sempre essere scritta come il prodotto di una matrice $m \times n$, \mathbf{U} , di una matrice diagonale $n \times n$, \mathbf{w} , con elementi non negativi, e della trasposta di una matrice \mathbf{V} ,

$$\mathbf{A} = \mathbf{U}\mathbf{w}\mathbf{V}^T \quad (3.13)$$

dove la matrici \mathbf{U} e \mathbf{V} sono ortonormali a colonne, cioè $\mathbf{U}^T\mathbf{U} = \mathbf{V}^T\mathbf{V} = \mathbf{1}$, dove $\mathbf{1}$ è intesa qui come la matrice identità $n \times n$. Questa decomposizione si chiama *decomposizione ai valori singolari* (SVD, dall'inglese *Singular Value Decomposition*). Il metodo SVD è molto utile per il calcolo dell'inversa di una matrice quadrata, $m = n$; se la matrice invertibile (cioè non è *singolare*)

$$\mathbf{A}^{-1} = \mathbf{V}\mathbf{w}^{-1}\mathbf{U}^T \quad (3.14)$$

il calcolo dell'inversa di \mathbf{w} è ovvio (basta costruire la matrice diagonale con gli inversi degli elementi diagonali della matrice \mathbf{w} originaria). Per un sistema lineare $\mathbf{Ax} = \mathbf{b}$, la soluzione è perciò

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} = \mathbf{V}\mathbf{w}^{-1}\mathbf{U}^T\mathbf{b} \quad (3.15)$$

⁴Ci concentriamo d'ora in poi solo sul caso delle matrici quadrate $m = n$, con elementi reali

L'*indice di condizionamento* della matrice $n \times n$ \mathbf{A} è il rapporto fra il più grande ed il più piccolo degli elementi (diagonali) di w_i ; se l'indice è infinito, la matrice è singolare.

Perchè il metodo SVD è importante? Prima di tutto trova un'applicazione naturale nella soluzione dei sistemi lineari che nascono dalla risoluzione di problemi di ottimizzazione basati sul *metodo dei minimi quadrati* (discussi in un capitolo successivo). Ma soprattutto, da un punto di vista pratico, permette di trattare meglio problemi numerici che coinvolgono matrici *malcondizionate*, cioè con un indice di condizionamento molto alto. In questo caso, una routine di calcolo basata sul metodo LU dà una soluzione formalmente corretta, ma affetta da un errore molto grande (a causa degli errori accumulati nel moltiplicare e dividere numeri molto grandi, causati da pivot vicini a zero). Il metodo SVD può invece fornire soluzioni molto più accurate basate sull'equazione (3.15), purchè si trascurino gli elementi di \mathbf{w} molto vicini a zero (il che equivale a trascurare il contributo di alcune delle equazioni lineari che sono così affette da errori di arrotondamento da corrompere la soluzione).

3.1.5 Matrici sparse

Come ultimo argomento in questa sezione, consideriamo ora l'importante - per i chimici - argomento delle *matrici sparse*. Una matrice sparsa è semplicemente una matrice con un grande numero di elementi nulli⁵. le matrici sparse si presentano spesso con schemi particolari, (diagonali a bande, diagonali a blocchi etc.) ma anche in forme molto complesse non riconducibili a classificazioni semplici. Le matrici sparse sono di

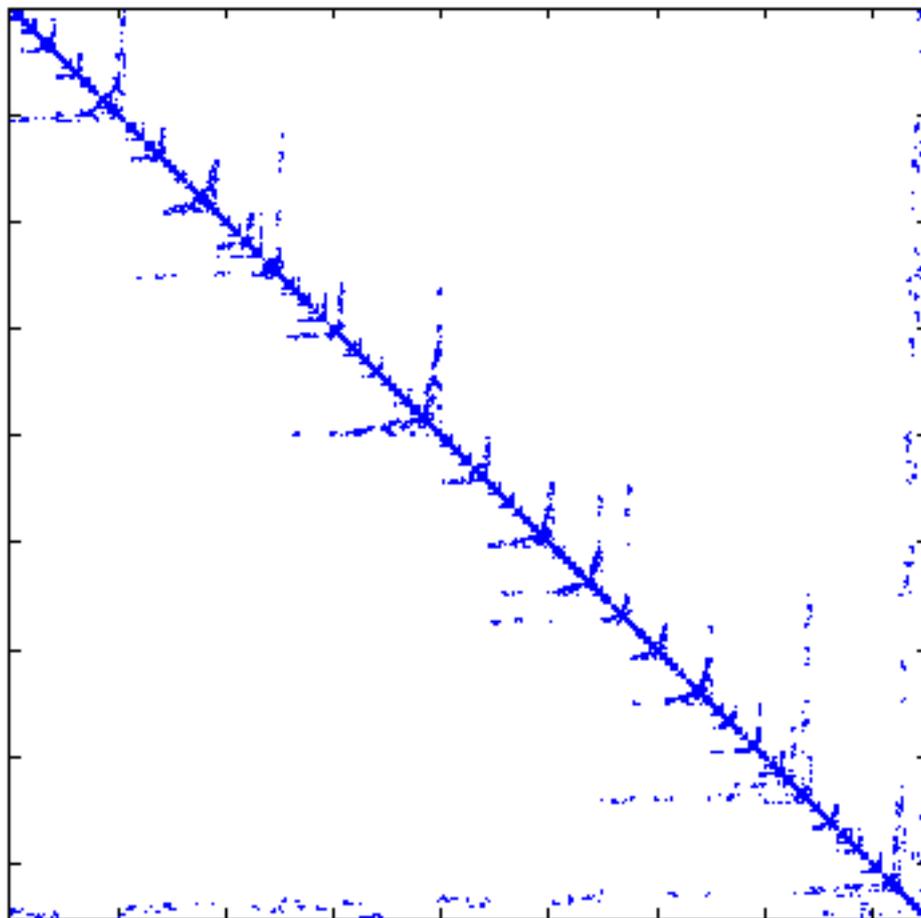


Figura 3.2: Esempio di matrice sparsa.

fondamentale interesse in chimica quantistica. Come vedremo un po' più avanti, e come sa ogni studente che abbia già affrontato il problema del calcolo dei livelli energetici di un sistema molecolare, con uno qualunque dei metodi più comuni usati nell'ambito della chimica computazionale, il calcolo delle matrici rappresentative di operatori specifici (e.g. la matrice di Fock nel metodo HF) su una base di orbitali atomici, porta di solito

⁵una matrice con la maggior parte degli elementi non nulli si dice *densa*

a matrici grandi (per molecole di qualche decina di atomi del primo e secondo periodo, le dimensioni sono di ordine $O(10^3) - O(10^4)$) e sparse; altre applicazioni che coinvolgono matrici sparse si trovano in quasitutte le moderne implementazioni di metodi per la soluzioni di equazioni differenziali, con i vari metodi di discretizzazione (elementi finiti, differenze finite, volumi finiti etc.): per esempio lo studio della reattività combinata con la convezione e diffusione di specie chimiche in un reattore, tipicamente affrontata nel contesto della chimica industriale e dell'ingegneria chimica, mediante equazioni idrodinamiche accoppiate a equazioni di diffusione-convezione-reazione porta spesso alla soluzione di matrici sparse di grande dimensioni, anche di ordine $O(10^6) - O(10^7)$ e più.

Il grande vantaggio delle matrici sparse è però che il numero di elementi non nulli cresce in modo, nella maggior parte dei casi, lineare con le loro dimensioni. Questo fatto consente di identificare algoritmi speciali per il trattamento di queste matrici, che eliminano automaticamente la maggior parte delle operazioni matematiche che coinvolgano elementi nulli. In questo modo, per una matrice sparsa di dimensione $n \times n$ si possono ottenere codici di calcolo che in termini di memoria e tempo scalano come $O(n^r)$ con r significativamente più piccolo di 3 (il tipico valore dei metodi di decomposizione discussi sopra).

Spesso, la natura di questi algoritmi è di tipo iterativo (sequenza di operazioni descrivibili come una successione di prodotti matrici/vettori e loro combinazioni lineari, in cui il risultato della k -esima operazione viene utilizzato nell'operazione $k + 1$ -esima). Tra i più comuni e noti ricordiamo il metodo del *gradiente coniugato* (CG, *conjugate gradient*) ed il *metodo generalizzato dei residui minimi* (GMRES, *generalized minimal residuals*) per la soluzione di una sistema lineare sparso. Naturalmente si deve ricordare che i metodi iterativi non sono necessariamente limitati alle matrici(sistemi) sparse; sono però sicuramente gli unici disponibili per matrici di grandi dimensioni e come abbiamo già detto, si prestano particolarmente ad essere implementati in programmi di calcolo con una relativa semplicità e soprattutto 'economia' di risorse computazionali.

3.1.6 Metodo del gradiente coniugato

Discutiamo a titolo di esempio, ed anche perchè tra gli algoritmi iterativi è tra i più diffusi, la famiglia di metodi del gradiente coniugato. Consideriamo il solito sistema di equazioni lineari a coefficienti e soluzione reali $\mathbf{Ax} = \mathbf{b}$, con $m = n$. Gli algoritmi di tipo CG sono sostanzialmente legati - almeno nel caso la matrice \mathbf{A} sia reale simmetrica e definita positiva (cioè con autovalori tutti maggiori di zero) - all'idea di cercare la soluzione (che supponiamo esistere, cioè \mathbf{A} si suppone non singolare e $\mathbf{b} \neq \mathbf{0}$), come il vettore nello spazio \mathcal{R}^n che minimizza la funzione $f(\mathbf{x}) = \mathbf{x}^T \mathbf{Ax} / 2 - \mathbf{x}^T \mathbf{b}$, cioè per la quale il gradiente della funzione valga zero $\hat{\nabla} f = \mathbf{Ax} - \mathbf{b} = \mathbf{0}$ ⁶

In pratica si genera una successione di vettori \mathbf{x}_k via via (si suppone) sempre più vicine alla soluzione esatta; ogni vettore si ottiene dal precedente come $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k$, avendo definito un opportuno vettore direzione \mathbf{p}_k ed uno scalare α_k , che minimizza $f(\mathbf{x}_{k+1})$. Per una matrice generica l'interpretazione geometrica è più complessa, ed è meglio presentare il metodo nella sua forma algoritmica, per la costruzione di una successione di vettori $\mathbf{x}_k, \mathbf{r}_k, \bar{\mathbf{r}}_k, \mathbf{p}_k, \bar{\mathbf{p}}_k$, con $k = 1, 2, \dots$. Consideriamo una approssimazione iniziale \mathbf{x}_1 come soluzione del sistema lineare, e definiamo $\mathbf{r}_1 = \bar{\mathbf{r}}_1 = \mathbf{b} - \mathbf{Ax}_1$; \mathbf{r}_1 e $\bar{\mathbf{r}}_1$ sono i *residui biconiugati*

$${}^6\hat{\nabla} = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \dots \\ \frac{\partial}{\partial x_n} \end{pmatrix}$$

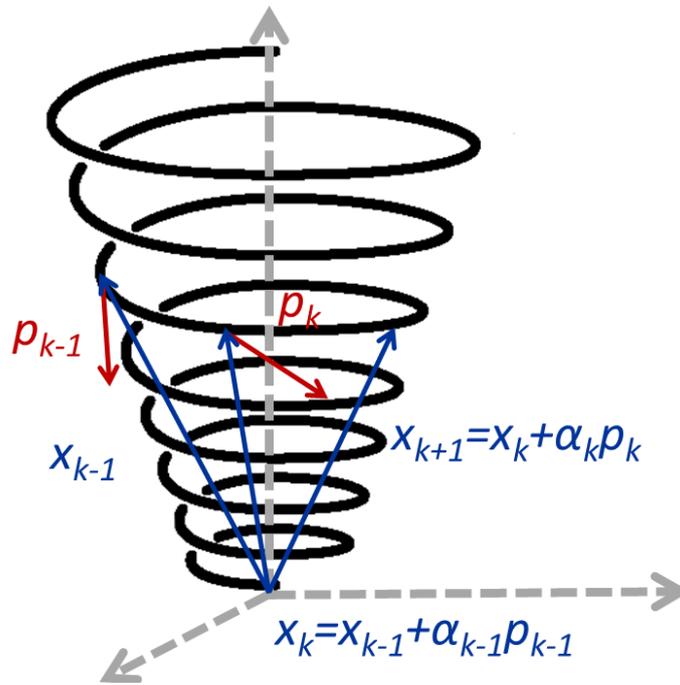


Figura 3.3: Schema CG in tre dimensioni.

alla prima iterazione. Dati \mathbf{r}_1 e $\bar{\mathbf{r}}_1$, $\mathbf{p}_1 = \mathbf{r}_1$, $\bar{\mathbf{p}}_1 = \bar{\mathbf{r}}_1$ si genera la successione

$$\begin{aligned}
 \alpha_k &= \frac{\bar{\mathbf{r}}_k^T \mathbf{r}_k}{\bar{\mathbf{p}}_k^T \mathbf{A} \mathbf{p}_k} \\
 \mathbf{r}_{k+1} &= \mathbf{r}_k - \alpha_k \mathbf{A} \mathbf{p}_k \\
 \bar{\mathbf{r}}_{k+1} &= \bar{\mathbf{r}}_k - \alpha_k \mathbf{A}^T \bar{\mathbf{p}}_k \\
 \mathbf{x}_{k+1} &= \mathbf{x}_k + \alpha_k \mathbf{p}_k \\
 \beta_k &= \frac{\bar{\mathbf{r}}_{k+1}^T \mathbf{r}_{k+1}}{\bar{\mathbf{r}}_k^T \mathbf{r}_k} \\
 \mathbf{p}_{k+1} &= \mathbf{p}_k + \beta_k \mathbf{p}_k \\
 \bar{\mathbf{p}}_{k+1} &= \bar{\mathbf{p}}_k + \beta_k \bar{\mathbf{p}}_k
 \end{aligned} \tag{3.16}$$

\mathbf{r}_k e $\bar{\mathbf{r}}_k$ sono i *residui biconiugati* alla k -iterazione. Il residuo \mathbf{r}_k è in effetti $\mathbf{r}_k = \mathbf{b} - \mathbf{A} \mathbf{x}_k$, e quindi indica quanto il vettore \mathbf{x}_k sia vicino alla soluzione esatta del sistema lineare. Valgono le proprietà: $\bar{\mathbf{r}}_i^T \mathbf{r}_j = \mathbf{r}_i^T \bar{\mathbf{r}}_j = 0$, $\bar{\mathbf{p}}_i^T \mathbf{A} \mathbf{p}_j = \mathbf{p}_i^T \mathbf{A}^T \bar{\mathbf{p}}_j = 0$, $\bar{\mathbf{r}}_i^T \mathbf{p}_j = \mathbf{r}_i^T \bar{\mathbf{p}}_j = 0$ per $j < i$.

La soluzione, se la precisione del calcolo è infinita, si dovrebbe trovare per $k = n$; in pratica, gli errori di arrotondamento rendono la procedura non esatta, ed il calcolo termina quando per esempio $|\mathbf{r}_k|$ è minore di qualche parametro di accuratezza prestabilito dal ricercatore. L'algoritmo è stato descritto nella sua forma *biconiugata*, adatta per matrici reali generiche; esistono versioni per matrici complesse, complesse simmetriche, hermitiane etc. Si deve notare che la convergenza dell'algoritmo non è dimostrabile a priori. In effetti l'applicazione numerica degli algoritmi iterativi può essere affetta dall'accumulo di errori di arrotondamento che possono in alcuni casi rendere la soluzione trovata alquanto inaccurata.

Prima di concludere, razionalizziamo il motivo per cui algoritmi iterativi come il metodo CG sono particolarmente efficienti nel caso di matrici sparse: di per sé, un algoritmo iterativo nel contesto dell'algebra lineare è una sequenza di operazioni prodotto tra una matrice ed un vettore, $\mathbf{A} \mathbf{v}$, e di operazioni di combinazione lineare di vettori; le seconde sono relativamente poco costose computazionalmente se rapportate alle prime, che tipicamente per una matrice generica di dimensioni $n \times n$, richiedono $O(n^3)$ operazioni matematiche elementari. Nel caso però di matrici sparse, è possibile però riscrivere in modo intelligente l'algoritmo per il calcolo del prodotto di una matrice con un vettore, tenendo conto dei soli elementi non nulli. Se n' è il numero di elementi non nulli, e questo numero è molto più piccolo di n (come è il caso delle matrici sparse), il guadagno in termini di tempi di calcolo e di memoria per il salvataggio dei risultati è molto grande: in

pratica, si guadagna molto tempo perchè il prodotto $\mathbf{A}\mathbf{v}$ è calcolato in modo molto più veloce per una matrice sparsa. Inoltre, se non sono presenti patologie particolari nella forma di \mathbf{A} , la convergenza rispetto al numero di iterazioni è spesso molto rapida, con una precisione accettabile della soluzione ottenuta entro un numero di iterazioni pari ad una frazione delle dimensioni di \mathbf{A} .

3.2 Problemi agli autovalori: esempi e metodi di soluzione

Consideriamo il problema del calcolo degli autovalori (destri) di una matrice \mathbf{A} , come è descritto nel Capitolo precedente, sezione (2.7)

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x} \quad (3.17)$$

dove \mathbf{A} è una matrice quadrata di dimensioni $n \times n$. Questa è l'applicazione più comune dei principi di base dell'algebra lineare, per un chimico. Il calcolo numerico od analitico degli autovalori di una matrice è, più o meno consapevolmente, uno degli strumenti principali che vengono utilizzati da un chimico per discutere innumerevoli proprietà dei sistemi molecolari. Prima di tutto, ma non solo, la struttura elettronica e nucleare di una molecola generica, descritta con i metodi della meccanica quantistica, si traduce naturalmente nel linguaggio del calcolo degli autovettori (le "funzioni d'onda" o gli "stati") ed autovalori (i "livelli energetici") di una matrice, che è a sua volta direttamente collegata ad un operatore agente sul sistema di interesse (vedi il Capitolo precedente).

In realtà la soluzione di un problema agli autovalori è richiesta anche in settori della chimica diversi dalla chimica quantistica. Ricordiamo qui per esempio lo studio delle cinetiche chimiche risultanti da meccanismi di reazione contenenti solo reazioni unimolecolari. Inoltre la soluzione di problemi agli autovalori ottenuti a partire da problemi elementari della meccanica quantistica si possono in alcuni casi trattare evitando un linguaggio algebrico e rimanendo nell'ambito dell'analisi matematica di base (vedi il Capitolo dedicato alle equazioni differenziali).

Varrebbe la pena, prima di procedere oltre, e per mettere nel giusto contesto il problema, ri-discutere l'approccio generale allo studio generale delle proprietà elettroniche di una molecola. Supporremo comunque nel seguito, per ragioni di spazio, che lo studente abbia già chiare le basi della meccanica quantistica necessaria per comprendere le principali applicazioni elementari in ambito chimico. Approfitteremo comunque di questa discussione per richiamare prima di tutto esempi di *calcolo analitico* dell'insieme degli autovalori e degli autovettori, per matrici con dimensioni n dell'ordine dell'unità, o per matrici anche grandi ma con strutture particolari; il metodo analitico è semplicemente effettuato risolvendo l'equazione secolare e poi risolvendo i sistemi lineari omogenei associati a ciascun autovalore. Al più la difficoltà del problema è limitata alla soluzione di un'equazione secolare polinomiale, le cui radici possono essere ottenibili solo numericamente (vedi il Capitolo dedicato alle equazioni non-lineari).

3.2.1 Problemi semplici agli autovalori in meccanica quantistica

Come è noto, l'hamiltoniano elettronico⁷ \mathcal{H} di una molecola un operatore complesso, che agisce su tutte le coordinate elettroniche (per esempio, per l'etilene C_2H_4 da $6 \times 2 + 1 \times 6 = 18$ elettroni otteniamo $18 \times 3 = 54$ coordinate). La funzione d'onda multielettronica viene costruita in modo approssimato (secondo l'approssimazione *monodeterminante di Slater*) a partire un certo numero di orbitali molecolari, vale a dire funzioni monoelettroniche, ottenute combinando linearmente un insieme di funzioni di base, solitamente forme approssimate degli orbitali atomici degli atomi costituenti la molecola. La maggior parte degli approcci di base della chimica computazionale risolve (per esempio il metodo di Hartree-Fock-Roothaan) risolve il problema del calcolo dei coefficienti che esprimono gli orbitali molecolari in funzione degli orbitali atomici **risolvendo una serie di problemi agli autovalori**.

3.2.2 Un sistema a due livelli

Consideriamo un primo caso semplice, analizzando il sistema molecolare H_2^+ . Prima di procedere, però, un'avvertenza: per quanto semplice un problema possa essere dal punto di vista computazionale, nella sua

⁷Consideriamo i nuclei a riposo e a geometria fissa, ovvero in altri termini cerchiamo le funzioni d'onda elettroniche nell'ambito dell'*approssimazione di Born-Oppenheimer*

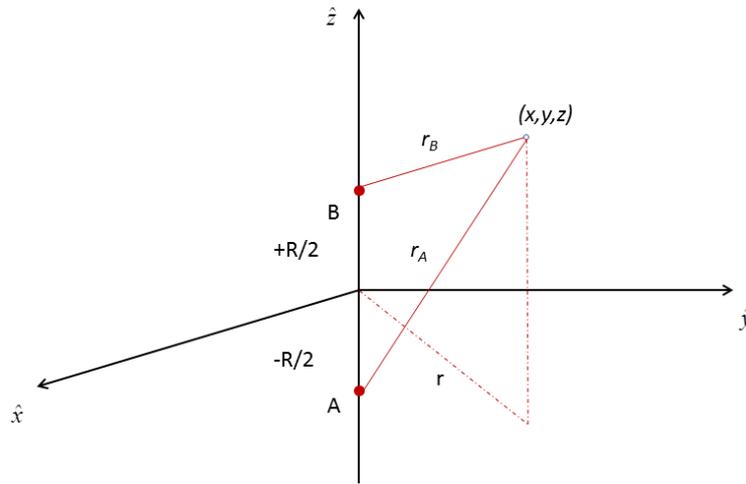


Figura 3.4: Sistema di riferimento e coordinate H_2^+

forma conclusiva, è necessario - soprattutto in un scienza come la chimica, che si occupa della materia in modo dettagliato, specifico, *molecolare* - definire bene le ipotesi e le approssimazioni necessarie. Quindi, quello che in realtà alla fine è un semplice studio di un sistema a due livelli, deve essere accuratamente compreso a partire dalle ipotesi di base della meccanica quantistica (non relativistica). Stabilendo che i nuclei siano in posizioni prefissate, possiamo definire appropriate coordinate interne nucleari - che possiamo scegliere in modo conveniente come R , θ and ϕ dove R è la distanza internucleare $|\vec{R}_A - \vec{R}_B|$, θ e ϕ le coordinate polari del vettore $\vec{R}_A - \vec{R}_B$. L'hamiltoniano elettronico è perciò

$$\mathcal{H} = -\frac{1}{2}\hat{\nabla}^2 - \frac{1}{r_A} - \frac{1}{r_B} + \frac{1}{R} \tag{3.18}$$

dove $r_A = \sqrt{r^2 + (z + R/2)^2}$ e $r_B = \sqrt{r^2 + (z - R/2)^2}$, $r^2 = x^2 + y^2$, e la posizione generica dell'elettrone è individuata dalle coordinate x , y , z . Si noti che nella discussione si applicazioni quantistiche adotteremo sempre, salvo indicazioni specifiche, *unità di misura atomiche*, che permettono di definire grandezze fisiche atomico/molecolari riscalate in modo semplice. Nella Tabella sono definite le principali unità atomiche: di particolare importanza sono l'unità di lunghezza, detta *Bohr*, e l'unità di energia, chiamata *Hartree*. Gli

massa	m_u	$m_e = 9.1091 \times 10^{-31}$ Kg
lunghezza	l_u	$a_0 = \frac{4\pi\epsilon_0\hbar^2}{m_e e^2} = 0.52917 \times 10^{-10}$ m
momento angolare	L_u	$\hbar = 4.16336 \times 10^{-33}$ J s ⁻¹
energia	E_u	$\frac{e^2}{4\pi\epsilon_0 a_0} = 4.359 \times 10^{-18}$ J
carica	e	$e = 1.602 \times 10^{-19}$ C

Tabella 3.1: Unità atomiche

autostati e i livelli energetici della molecola di idrogeno ionizzata saranno in generale funzioni parametriche di R . Per $R \rightarrow 0$ la molecola diviene l'atomo di elio ionizzato, ed i suoi orbitali e stati energetici sono quelli di un atomo idrogenoide con $Z = 2$. Per $R \rightarrow \infty$ la molecola diviene equivalente alla coppia $H^+ + H$. Scriviamo l'equazione di Schrödinger nella forma

$$\left(-\frac{1}{2}\hat{\nabla}^2 - \frac{1}{r_A} - \frac{1}{r_B} + \frac{1}{R} \right) \Psi = E\Psi \tag{3.19}$$

dove E è l'energia dell'autostato elettronico descritto dalla autofunzione Ψ .

Approfondimento 5 . La soluzione del problema (3.19) è possibile in modo analitico. Per quanto questo argomento esuli dal tema trattato in questo Capitolo, e sia piuttosto legato ai metodi di soluzione delle equazioni differenziali, riassumiamo qui per convenienza la procedura di soluzione analitica. Introduciamo

le coordinate ellissoidali μ , ν e l'angolo azimutale ϕ relativo alla rotazione intorno all'asse internucleare; valgono le relazioni dirette ed inverse:

$$\begin{aligned}\mu &= r_A + r_B/R \\ \nu &= r_A - r_B/R \\ x &= \sqrt{(\mu^2 - 1)(1 - \nu^2)}/R \cos \phi \\ y &= \sqrt{(\mu^2 - 1)(1 - \nu^2)}/R \sin \phi\end{aligned}\quad (3.20)$$

$$z = \mu\nu R/2 \quad (3.21)$$

con $\mu \geq 1$, $-1 \leq \nu \leq 1$; μ descrive gli ellissoidi di rivoluzione aventi come fuochi i nuclei, mentre ν descrive gli iperboloidi confocali di rivoluzione [cfr. Fig. (3.5)]. Il laplaciano può essere espresso in queste coordinate

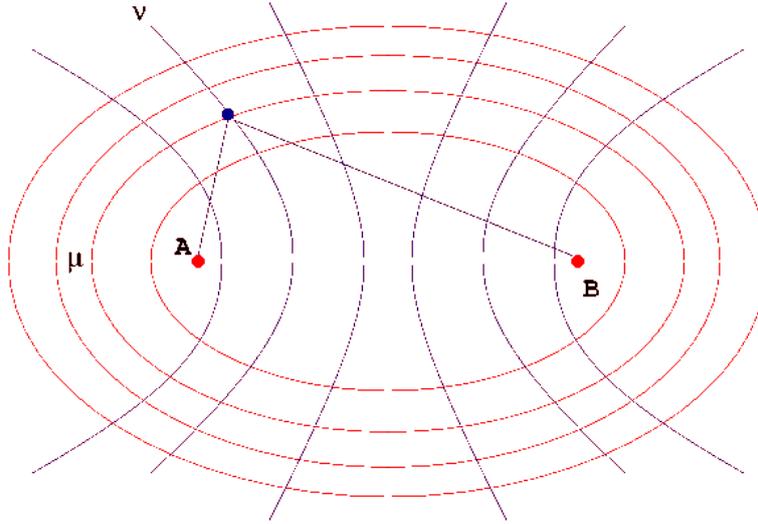


Figura 3.5: Coordinate ellissoidali H_2^+

nella forma:

$$\hat{\nabla}^2 = \frac{4}{R^2(\mu^2 - \nu^2)} \left[\frac{\partial}{\partial \mu} (\mu^2 - 1) \frac{\partial}{\partial \mu} + \frac{\partial}{\partial \nu} (1 - \nu^2) \frac{\partial}{\partial \nu} + \frac{\partial}{\partial \phi} \frac{(\mu^2 - \nu^2)}{(\mu^2 - 1)(1 - \nu^2)} \frac{\partial}{\partial \phi} \right] \quad (3.22)$$

Sostituendo nell'equazione di Schrödinger, si ottiene *un'equazione differenziale separabile*, cioè ponendo

$$\Psi(\mu, \nu, \phi) = M(\mu)N(\nu)F(\phi) \quad (3.23)$$

si trovano le tre equazioni separate:

$$\left[\frac{d}{d\mu} (\mu^2 - 1) \frac{d}{d\mu} - \epsilon \mu^2 + 2R\mu - \frac{\lambda^2}{\nu^2 - 1} + \kappa \right] M(\mu) = 0 \quad (3.24)$$

$$\left[\frac{d}{d\nu} (1 - \nu^2) \frac{d}{d\nu} + \epsilon \nu^2 + \frac{\lambda^2}{1 - \nu^2} - \kappa \right] N(\nu) = 0 \quad (3.25)$$

$$\left(\frac{d^2}{d\phi^2} + \lambda^2 \right) F(\phi) = 0 \quad (3.26)$$

λ e κ sono costanti di separazione, mentre $\epsilon = -R^2 E/2$. In particolare, λ è il numero quantico associato con la componente del momento angolare elettronico lungo l'asse internucleare, ed assume i valori $0, \pm 1, \pm 2, \dots$. L'autovalore E dipende da λ^2 , ovvero è presente una doppia degenerazione dovuta all'equivalenza delle due

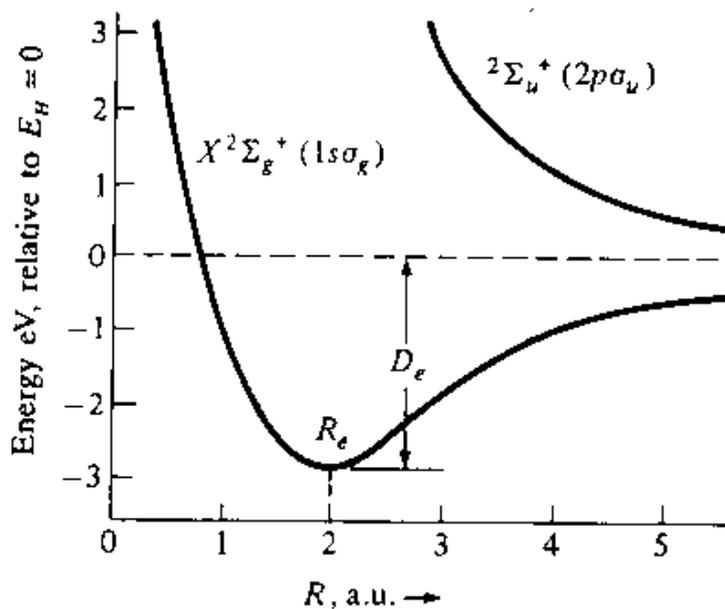


Figura 3.6: I primi due livelli energetici della molecola di H_2^+ [E. Teller, *Z. Physik*, **61**, 474 (1930)]

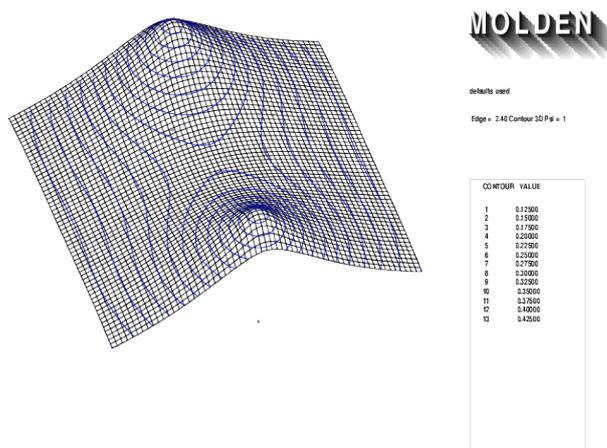


Figura 3.7: Superficie che rappresenta l'orbitale $1s\sigma_g$ in un piano molecolare

direzioni di rotazione attorno all'asse. I valori permessi di κ e ϵ sono determinabili numericamente. In generale i vari stati, doppiamente degeneri, di H_2^+ sono simboleggiati da una notazione del tipo $nl\alpha_{u,g}$ dove: n ed l sono riferiti ai numeri quantici dell'orbitale atomico idrogenoide, cui l'orbitale molecolare tende per $R \rightarrow 0$; α è una lettera greca $\sigma, \pi, \delta, \phi, \dots$ corrispondente a $|\lambda| = 0, 1, 2, 3, \dots$; u o g corrispondono alla simmetria dell'orbitale molecolare rispetto all'operazione di inversione rispetto al punto centrale tra i nuclei - si può dimostrare che l'orbitale molecolare cambia di segno (u) o resta invariato (g) rispetto questa operazione di simmetria.

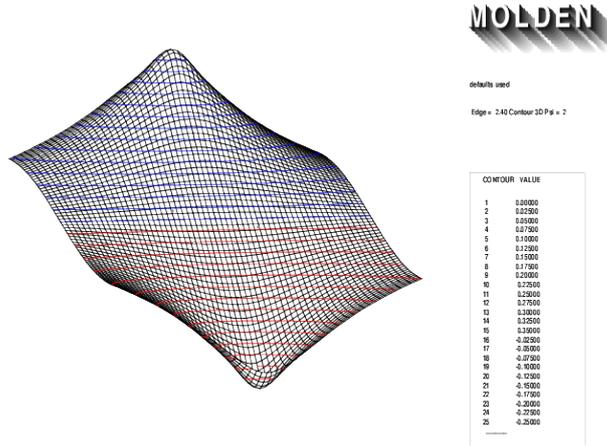


Figura 3.8: Superficie che rappresenta l'orbitale $2p\sigma_u$ in un piano molecolare

La funzione d'onda soluzione del problema agli autovalori (3.19) è esprimibile come combinazione lineare di funzioni, detta *base*, note (vedi la sezione successiva). Nel caso della molecola di idrogeno ionizzata introduciamo una base di sole due funzioni, $1s_A(\vec{r})$ e $1s_B(\vec{r})$, rispettivamente il primo orbitale idrogenoide centrato sul nucleo A ed il primo orbitale idrogenoide centrato sul nucleo B. Due funzioni sono poche? Il risultato però è sorprendentemente accurato se confrontato con la soluzione esatta, ed è in realtà una conseguenza della semplicità del sistema. Utilizziamo la notazione di Dirac: $|\Psi\rangle$ è la funzione d'onda, $|1\rangle$ è $1s_A(\vec{r})$ e $|2\rangle$ è $1s_B(\vec{r})$; tutte le funzioni sono reali, quindi possiamo lavorare solo in \mathcal{R} . Le due funzioni di base $|1\rangle$ e $|2\rangle$ hanno la forma esplicita

$$|1\rangle = \frac{1}{\sqrt{\pi}} e^{-r_A} \quad |2\rangle = \frac{1}{\sqrt{\pi}} e^{-r_B} \quad (3.27)$$

e sono normalizzate $\langle 1|1\rangle = \langle 2|2\rangle = 1$, mentre NON sono ortogonali, $\langle 1|2\rangle = \langle 2|1\rangle = s$, dove

$$s = \frac{1}{\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-r_A - r_B} dx dy dz \quad (3.28)$$

per $R \rightarrow \infty$ le due funzioni hanno overlap nullo, cioè $\lim_{R \rightarrow \infty} s = 0$. Il calcolo della matrice rappresentativa dell'hamiltoniano (3.18) porge invece $\langle 1|\mathcal{H}|1\rangle = \langle 2|\mathcal{H}|2\rangle = a + \frac{1}{R}$ e $\langle 1|\mathcal{H}|2\rangle = \langle 2|\mathcal{H}|1\rangle = b + \frac{s}{R}$ dove

$$a = \frac{1}{\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-r_A} \left(-\frac{1}{2} \nabla^2 - \frac{1}{r_A} - \frac{1}{r_B} + \frac{1}{R} \right) e^{-r_A} dx dy dz \quad (3.29)$$

$$b = \frac{1}{\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-r_A} \left(-\frac{1}{2} \nabla^2 - \frac{1}{r_A} - \frac{1}{r_B} + \frac{1}{R} \right) e^{-r_B} dx dy dz \quad (3.30)$$

quindi a è l'elemento di matrice di $1s_A(\vec{r})$ dell'hamiltoniano elettronico a meno del termine $1/R$ rispetto a $1s_A(\vec{r})$, b è l'elemento di matrice di $1s_A(\vec{r})$ dell'hamiltoniano elettronico a meno del termine $1/R$ rispetto

a $1s_B(\vec{r})$, s è l'elemento di matrice di overlap tra le due funzioni, che sono normalizzate ma non ortogonali. La funzione d'onda è $|\Psi\rangle = c_1|1\rangle + c_2|2\rangle$. Il problema agli autovalori in forma matriciale assume perciò la forma

$$\begin{pmatrix} a + \frac{1}{R} & b + \frac{s}{R} \\ b + \frac{s}{R} & a + \frac{1}{R} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = E \begin{pmatrix} 1 & s \\ s & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \xrightarrow{\text{in forma algebrica}} \mathbf{H}\mathbf{c} = E\mathbf{S}\mathbf{c} \quad (3.31)$$

Questo è un problema generalizzato agli autovalori, dovuto alla presenza di una matrice di overlap diversa da zero; se trascuriamo s (per esempio assumendo che i due nuclei siano abbastanza lontani da rendere piccola la sovrapposizione tra i due orbitali atomici), ritroviamo il problema classico

$$\begin{pmatrix} a + \frac{1}{R} & b \\ b & a + \frac{1}{R} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = E \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = E \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \xrightarrow{\text{in forma algebrica}} \mathbf{H}\mathbf{c} = E\mathbf{c} \quad (3.32)$$

Le forme esplicite di a , b e s si possono calcolare analiticamente con metodi elementari (vedi il Capitolo dedicato ai metodi di integrazione):

$$\begin{aligned} a &= -\frac{1}{2} + e^{-2R} + \frac{e^{-2R} - 1}{R} < 0 \\ b &= -\frac{s}{2} + e^{-R}(1 + R) < 0 \\ s &= e^{-R} \left(1 + R + \frac{R^2}{3} \right) > 0 \end{aligned} \quad (3.33)$$

Analizziamo la soluzione del problema (3.31); il sistema di due equazioni lineari si può riscrivere nella forma già familiare

$$(\mathbf{H} - E\mathbf{S})\mathbf{c} = \mathbf{0} \quad (3.34)$$

Come abbiamo visto in precedenza, il sistema, omogeneo con un numero di equazioni pari al numero di incognite, ammette soluzioni non triviali solo se il determinante della matrice dei coefficienti è zero, quindi si trova l'equazione secolare

$$\det(\mathbf{H} - E\mathbf{S}) = 0 \quad (3.35)$$

Si tratta di una semplice equazione di secondo grado in E , che si risolve trovando le seguenti espressioni per i due autovalori $E_1 < E_2$ e relativi autostati

Stato fondamentale

$$E_1 = \frac{a+b}{1+s} + \frac{1}{R} \quad |\Psi_1\rangle = \frac{1}{\sqrt{2(1+s)}}(|1\rangle + |2\rangle) = \frac{1}{\sqrt{2(1+s)\pi}}(e^{-rA} + e^{-rA}) \quad (3.36)$$

Stato eccitato

$$E_2 = \frac{a-b}{1-s} + \frac{1}{R} \quad |\Psi_2\rangle = \frac{1}{\sqrt{2(1-s)}}(|1\rangle - |2\rangle) = \frac{1}{\sqrt{2(1-s)\pi}}(e^{-rA} - e^{-rA}) \quad (3.37)$$

La scelta della base utilizzata per il calcolo approssimato dei primi due stati elettronici della molecola di idrogeno ionizzata è stata effettuata in base al comportamento al limite di $R \rightarrow \infty$ della molecola, cioè assumendo che sostanzialmente gli orbitali molecolari siano ottenibili come combinazioni lineari degli orbitali atomici idrogenoidi centrati sui nuclei (metodo LCAO, *Linear Combination of Atomic Orbitals*).

3.2.3 Il metodo di Hückel per i polieni coniugati

Un esempio che dovrebbe essere molto noto agli studenti alle prese con i rudimenti della meccanica quantistica applicata alla chimica (organica) e l'approccio semiempirico di Hückel per il calcolo degli orbitali molecolari π . Proposto nel 1930 da Erich Hückel, un chimico-fisico tedesco, il metodo omonimo presuppone che *i*) in un poliene coniugato esistano insiemi di elettroni che si possono trattare separatamente, detti appunto elettroni π che formano orbitali molecolari distribuiti su tutta la molecola, *ii*) sia possibile esprimere questi orbitali molecolari come una combinazione lineare di alcuni specifici orbitali atomici, di solito orbitali idrogenoidi di tipo $2p_z$, *iii*) si possano semplificare in modo molto semplice gli elementi di matrice che derivano dalla soluzione diretta dell'equazione di Schrödinger sulla base delle ipotesi *i*) e *ii*). Di per sé il metodo di Hückel può essere visto come una versione estremamente semplificata del metodo LCAO, oppure anche come il risultato di un calcolo Hartree-Fock-Roothaan semi-empirico con un'estrema semplificazione della matrice di Fock. Il metodo di Hückel è applicabile a molecole organiche, nelle quali sia possibile separare una classe di elettroni relativamente localizzati ed inattivi, tipicamente impegnati in legami σ , e un'insieme di elettroni delocalizzati e reattivi, impegnati in legami π . La natura di questa identificazione può non essere affatto chiara, anche se nei casi più comuni (p. es. l'insieme di legami coniugati nei sistemi aromatici) risulta essere abbastanza ovvia. Le classi più comuni di molecole coinvolte sono:

- Polieni lineari: es. 1,3-butadiene;
- idrocarburi aromatici: benzene, naftalene;
- etero-molecole aromatiche: piridina, furano, acroleina;

Consideriamo il caso di polieni organici contenenti solo C e H, senza la presenza di eteroatomi. In un linguaggio pseudo-algebrico, il metodo di Hückel assume che gli n orbitali molecolari di tipo π , dai quali si può poi costruire la funzione d'onda multielettronica del sottoinsieme di n elettroni π , sono ottenuti come combinazioni lineari degli orbitali $|1\rangle = |2p_{z1}\rangle, \dots, |2p_{zn}\rangle$, dove ogni $|i\rangle$ è un orbitale atomico idrogenoide $2p_z$ dell'atomo i -esimo. Si definisce l'elemento diagonale della matrice di Fock, $\alpha = \langle i|\mathcal{F}|i\rangle$ e come l'elemento di matrice $\beta = \langle i|\mathcal{F}|i+1\rangle = \langle i+1|\mathcal{F}|i\rangle$ tra orbitali di atomi adiacenti. Si pone infine $\langle i|\mathcal{F}|j\rangle = 0$ per $|i-j| > 1$ e $\langle i|j\rangle = \delta_{ij}$ (ortonormalità degli orbitali atomici). Secondo la terminologia di Hartree-Fock, α è l'integrale di Coulomb, β è l'integrale di risonanza (entrambi *negativi*). Un tentativo di sistematizzazione dell'approssimazione di Hückel è stato dato da Likos e Parr nell'ambito di un formalismo di partenza di tipo Hartree-Fock-Roothaan. Sotto le condizioni di separabilità $\sigma - \pi$ si può separare il calcolo della parte della funzione d'onda riferita ai soli orbitali π . Per esempio nel caso dell'etilene C_2H_4 otteniamo che il calcolo degli orbitali molecolari si traduce in un problema agli autovalori del tipo

$$\begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = E \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \quad (3.38)$$

dove l'orbitale molecolare è dato da $c_1|1\rangle + c_2|2\rangle$. In questo caso la soluzione è triviale, come nell'esempio precedente si tratta di un sistema a due livelli, con un'equazione secolare semplicissima $(\alpha - E)^2 - \beta^2 = 0$, $E_{1,2} = \alpha \pm \beta$

Stato fondamentale

$$E_1 = \alpha + \beta \quad |\Psi_1\rangle = \frac{1}{\sqrt{2}}(|1\rangle + |2\rangle) \quad (3.39)$$

Stato eccitato

$$E_2 = \alpha - \beta \quad |\Psi_2\rangle = \frac{1}{\sqrt{2}}(|1\rangle - |2\rangle) \quad (3.40)$$

In generale avremo

$$\mathbf{F}\mathbf{c} = E\mathbf{c} \quad (3.41)$$

dove \mathbf{F} è costruita secondo la struttura della molecola con elementi diversi da zero solo per gli elementi diagonali e relativi ad atomi adiacenti. Ponendo $-x = (\alpha - E)/\beta$, il calcolo è ricondotto alla diagonalizzazione della *matrice topologica* \mathbf{A} :

$$|\mathbf{A} - x\mathbf{1}| = 0 \quad (3.42)$$

dove A_{ij} è 1 per atomi adiacenti (primi vicini), 0 altrimenti.

Applicazione 2 . Consideriamo p. es. il caso dell'1,3-butadiene. La matrice topologica è data da:

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (3.43)$$

la soluzione del problema agli autovalori (3.42) porta all'equazione in $x^4 - 3x^2 + 1 = 0$ una biquadratica che si risolve facilmente trovando le quattro radici reali: $\pm\sqrt{\frac{3-\sqrt{5}}{2}}$, $\pm\sqrt{\frac{3+\sqrt{5}}{2}}$, da cui si ottengono i quattro

livelli (in ordine di energia crescente) $E_1 = \alpha + \beta\sqrt{\frac{3+\sqrt{5}}{2}}$, $E_2 = \alpha + \beta\sqrt{\frac{3-\sqrt{5}}{2}}$, $E_3 = \alpha - \beta\sqrt{\frac{3-\sqrt{5}}{2}}$,

$$E_4 = \alpha - \beta\sqrt{\frac{3+\sqrt{5}}{2}}.$$

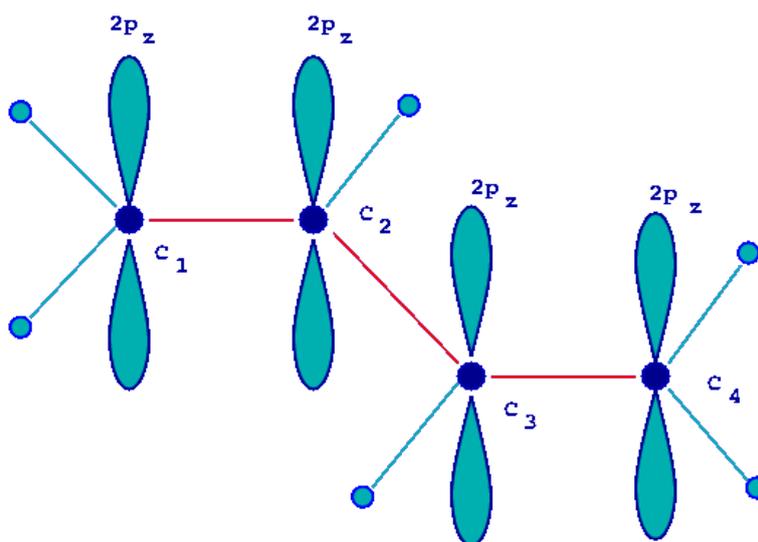


Figura 3.9: Esempio: 1,3-butadiene.

Per un *poliene lineare coniugato* formato da n atomi, la matrice topologica ha la forma *tridiagonale*:

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 1 & \dots & 0 \\ & & \dots & & \\ 0 & 1 & 0 & \dots & 1 \\ 0 & 0 & 1 & \dots & 0 \end{pmatrix} \quad (3.44)$$

mentre per un *poliene ciclico* la matrice assume la forma

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & \dots & 1 \\ 1 & 0 & 1 & \dots & 0 \\ & & \dots & & \\ 0 & 1 & 0 & \dots & 1 \\ 1 & 0 & 1 & \dots & 0 \end{pmatrix} \quad (3.45)$$

Approfondimento 6 . Una matrice tridiagonale generica ha la forma

$$\mathbf{T} = \begin{pmatrix} a_1 & b_1 & & & \\ c_1 & a_2 & b_2 & & \\ & c_2 & \ddots & \ddots & \\ & & \ddots & \ddots & b_{n-1} \\ & & & c_{n-1} & a_n \end{pmatrix} \quad (3.46)$$

Il determinante D_n della matrice è ottenibile dai determinanti D_{n-1} e D_{n-2} delle matrici di dimensioni $(n-1) \times (n-1)$ e $(n-2) \times (n-2)$ ottenute eliminando l'ultima riga e l'ultima colonna, o le ultime due righe e le ultime due colonne, mediante la relazione ricorsiva

$$D_n = a_n D_{n-1} - c_{n-1} b_{n-1} D_{n-2} \quad (3.47)$$

dove $D_0 = 1$ e $D_{-1} = 0$.

Se gli elementi di una matrice lungo una qualunque diagonale discendente sono uguali, la matrice si dice di Toeplitz. Nel caso delle matrici tridiagonali, ciò equivale ad avere $a_1 = a_2 = \dots = a_n = a$, $b_1 = b_2 = \dots = b_n = b$, $c_1 = c_2 = \dots = c_n = c$. In questo caso gli autovalori di \mathbf{T} , $\mathbf{T}\mathbf{c} = x\mathbf{c}$ sono dati dall'espressione analitica $x_k = a + 2\sqrt{bc} \cos\left(\frac{k\pi}{n+1}\right)$, con $k = 1, \dots, n$. Per un poliene lineare $a = 0$, $b = c = 1$, e quindi la matrice topologica è una matrice di Toeplitz, tridiagonale (e simmetrica), da cui segue $x_k = 2 \cos\left(\frac{k\pi}{n+1}\right)$.

3.2.4 Problemi agli autovalori di grandi dimensioni

La soluzione del problema generale del calcolo della funzione d'onda, soluzione del problema agli autovalori - e quindi la ricerca degli autostati della molecola o del sistema generico rappresentato dall'hamiltoniano generico \mathcal{H} - può essere affrontato in modo generale esprimendo la funzione (cioè l'autostato) come una combinazione lineare di funzioni note: nel linguaggio degli spazi di Hilbert: il vettore funzione d'onda è proiettato su una base dello spazio) e cercando i coefficienti dell'espansione risolvendo un problema agli autovalori in forma matriciale.

In teoria, dato che la dimensione dello spazio di Hilbert (cioè le possibili ipotetiche funzioni d'onda che descrivono uno stato del sistema) è infinita, l'espansione dovrebbe contenere infiniti termini; in pratica, si sceglie una base finita (troncata) che assicuri una accuratezza prestabilita alla funzione d'onda cercata. In generale si potrà esprimere lo stato (il vettore) come una combinazione (non necessariamente lineare) di vettori di base *noti*; i coefficienti dell'espansione possono essere determinati sulla base del *principio variazionale*, che sostanzialmente sfrutta le proprietà di hermiticità dell'operatore hamiltoniano, e giustifica formalmente molti metodi numerici applicati normalmente a problemi di chimica quantistica.

Approfondimento 7 . È dato un sistema descritto dall'hamiltoniano \mathcal{H} , con autofunzioni ψ , $\mathcal{H}\psi = E\psi$. Ci poniamo il seguente problema: sotto quali condizioni il funzionale di una generica funzione d'onda ξ

$$J[\xi] = \langle \xi | \mathcal{H} | \xi \rangle \quad (3.48)$$

è stazionario, cioè massimo o minimo? Assumiamo inoltre che ξ debba essere normalizzata

$$\langle \xi | \xi \rangle = 1 \quad (3.49)$$

Dal punto di vista del calcolo variazionale il problema può essere posto sotto forma della ricerca di un punto stazionario vincolato, ed è dunque equivalente ad individuare le condizioni di stazionarietà del funzionale non vincolato

$$J'[\xi] = \langle \xi | \mathcal{H} | \xi \rangle - \lambda(\langle \xi | \xi \rangle - 1) \quad (3.50)$$

Al primo ordine, la variazione di J' è data da

$$\delta J' = \langle \delta \epsilon | \mathcal{H} | \epsilon \rangle + \langle \epsilon | \mathcal{H} | \delta \epsilon \rangle - \delta \lambda (\langle \xi | \xi \rangle - 1) - \lambda \langle \delta \xi | \xi \rangle - \lambda \langle \xi | \delta \xi \rangle \quad (3.51)$$

che si può riscrivere come

$$\delta J' = (\langle \delta \xi | \mathcal{H} | \xi \rangle - \lambda \langle \delta \xi | \xi \rangle) + (\langle \delta \xi | \mathcal{H} | \xi \rangle - \lambda \langle \delta \xi | \xi \rangle)^* - \delta \lambda (\langle \xi | \xi \rangle - 1) \quad (3.52)$$

La soluzione si ottiene *i)* se $\langle \xi | \xi \rangle = 1$, dato che solo così il termine in $\delta \lambda$ si annulla e poi *ii)* se

$$\langle \delta \xi | \mathcal{H} | \xi \rangle = \lambda \langle \delta \xi | \xi \rangle \quad (3.53)$$

o in altri termini, dato che $\delta \xi$ è arbitraria, se

$$\mathcal{H} \xi = \lambda \xi \quad (3.54)$$

cioè se ξ è un'autofunzione normalizzata di \mathcal{H} con autovalore λ . In pratica, definiamo il *rapporto di Rayleigh* per un generica funzione d'onda ξ

$$\epsilon[\xi] = \frac{\langle \xi | \mathcal{H} | \xi \rangle}{\langle \xi | \xi \rangle} \quad (3.55)$$

$\epsilon[\xi] \equiv \epsilon[\xi]$ è un funzionale della funzione generica ξ . Se $J'[\xi]$ è reso stazionario da ξ anche ϵ è stazionario.

Il principio variazionale può essere applicato utilmente al calcolo degli autovalori di un sistema. Il *teorema di Eckart* afferma che se E_0 è il più basso autovalore del sistema, il rapporto di Rayleigh è sempre maggiore o uguale a E_0 .

$$\epsilon[\xi] \geq E_0 \quad (3.56)$$

La dimostrazione è la seguente: espandiamo $|\xi\rangle$ (usiamo la notazione di Dirac per convenienza) sulla base degli autostati *esatti* dell'hamiltoniano \mathcal{H}

$$\mathcal{H} |\psi_n\rangle = E_n |\psi_n\rangle \quad E_0 \leq E_1 \leq \dots \quad (3.57)$$

$$\langle \psi_n | \psi_{n'} \rangle = \delta_{nn'} \quad (3.58)$$

$$|\xi\rangle = \sum_n c_n |\psi_n\rangle \quad (3.59)$$

consideriamo la quantità $\langle \xi | \mathcal{H} - E_0 | \xi \rangle$:

$$\langle \psi_n | \mathcal{H} - E_0 | \psi_{n'} \rangle = \sum_{nn'} c_n^* c_{n'} (E_n - E_0) \langle \psi_n | \psi_{n'} \rangle = \sum_n |c_n|^2 (E_n - E_0) \geq 0 \quad (3.60)$$

infatti $|c_n|^2 \geq 0$, $E_n \geq E_0$; segue che:

$$\langle \xi | \mathcal{H} | \xi \rangle - E_0 \langle \xi | \xi \rangle \geq 0 \Rightarrow \epsilon[\xi] \geq E_0 \quad (3.61)$$

Data dunque una funzione generica normalizzabile, il rapporto di Rayleigh costituisce sempre un limite superiore dell'autovalore più basso dell'hamiltoniano del sistema.

Nella sua versione più comune, il teorema variazionale è implementato sotto forma di **metodo variazionale lineare**. Il vettore d'onda di prova $|\xi\rangle$ è cioè dato come una combinazione *lineare* di N vettori noti, che indichiamo semplicemente come $|1\rangle$, $|2\rangle$ etc.

$$|\xi\rangle = \sum_{n=1}^N c_n |n\rangle \quad (3.62)$$

Il rapporto di Rayleigh può essere scritto come

$$\epsilon = \frac{\mathbf{c}^\dagger \mathbf{H} \mathbf{c}}{\mathbf{c}^\dagger \mathbf{S} \mathbf{c}} \quad (3.63)$$

dove \mathbf{c} è il vettore dei coefficienti c_n , mentre \mathbf{H} e \mathbf{S} sono le matrici rappresentative di \mathcal{H} e dell'unità sui vettori $|n\rangle$, cioè le matrici $N \times N$ aventi elementi

$$H_{nn'} = \langle n | \mathcal{H} | n' \rangle \quad S_{nn'} = \langle n | n' \rangle \quad (3.64)$$

La variazione infinitesimale della funzione di prova diviene ora una variazione infinitesimale dei coefficienti c_n ; dal principio variazionale segue perciò

$$\delta\epsilon = \frac{\delta\mathbf{c}^\dagger \mathbf{H} \mathbf{c} - \epsilon \delta\mathbf{c}^\dagger \mathbf{S} \mathbf{c}}{\mathbf{c}^\dagger \mathbf{S} \mathbf{c}} + \text{complesso coniugato} = 0 \quad (3.65)$$

che è verificata solo se $\delta\mathbf{c}^\dagger (\mathbf{H} \mathbf{c} - \epsilon \mathbf{S} \mathbf{c}) = 0$ ovvero

$$\mathbf{H} \mathbf{c} = \epsilon \mathbf{S} \mathbf{c} \quad (3.66)$$

cioè la forma generalizzata di un problema agli autovalori, che si può risolvere calcolando le radici dell'equazione secolare

$$\det(\mathbf{H} - \epsilon \mathbf{S}) = 0 \quad (3.67)$$

Il problema agli autovalori diviene più semplice se si assumono delle funzioni di base ortonormali $\mathbf{S} = \mathbf{1}$:

$$\mathbf{H} \mathbf{c} = E \mathbf{c} \quad (3.68)$$

che è ora un problema agli autovalori in forma standard, con matrici e vettori di dimensioni finite, poiché la rappresentazione (3.62) è di solito arrestata ad un numero finito di termini.

Applicazione 3 . Come esempio conclusivo, discutiamo la definizione dei modi normali di vibrazione

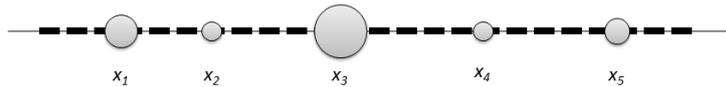


Figura 3.10: Molecola lineare

di una 'molecola' lineare. Consideriamo n atomi su una retta, con coordinate x_1, \dots, x_n ; descriviamo il sistema dal punto di vista quantistico come

$$\mathcal{H} \Psi(\mathbf{x}) = E \Psi(\mathbf{x}) \quad (3.69)$$

$$\mathcal{H} = -\frac{1}{2} \sum_{i=1}^n \frac{1}{m_i} \frac{\partial^2}{\partial x_i^2} + V(\mathbf{x}) = \mathcal{T} + V \quad (3.70)$$

dove $\mathbf{x} = (x_1, \dots, x_n)^\top$. \mathcal{T} è l'operatore energia cinetica. Scriviamo il potenziale come $V(\mathbf{x}) = \sum_{i=2}^n v_i(x_i - x_{i-1})$ (interazioni tra primi vicini). Assumiamo che le funzioni v_i siano caratterizzate da un minimo per $x_i - x_{i-1} = d_i$; sviluppando le funzioni v_i rispetto a $x_i - x_{i-1}$ intorno a d_i otteniamo $V(\mathbf{x}) = \text{cost} + \frac{1}{2} k_i (x_i - x_{i-1} - d_i)^2 + O[(x_i - x_{i-1} - d_i)^3]$ possiamo scegliere n costanti a_i tali che $a_i - a_{i-1} > 0$ e definire $\xi_i = m_i^{1/2} (x_i - a_i)$ ⁸, in modo tale che, rispetto alle nuove variabili ξ_i , l'operatore energia assuma una

⁸Le costanti a_i possono essere viste come posizioni di equilibrio degli atomi in un sistema di riferimento solidale con la molecola

forma particolarmente semplice, dato che $\frac{\partial}{\partial x_i} = \frac{\partial \xi_i}{\partial x_i} \frac{\partial}{\partial \xi} = m_i^{1/2} \frac{\partial}{\partial \xi}$, da cui $\mathcal{T} = -\frac{1}{2} \sum_{i=1}^n \frac{\partial^2}{\partial \xi_i^2}$. L'energia potenziale assume la forma (trascurando termini di ordine superiore al terzo) e scegliendo $a_1 = 0, a_2 = d_2, a_3 = d_2 + d_3, \dots, a_n = \sum_{i=2}^n d_i$:

$$V(\boldsymbol{\xi}) = \frac{1}{2} \boldsymbol{\xi}^T \mathbf{A} \boldsymbol{\xi} \quad (3.71)$$

dove la matrice \mathbf{A} è definita come

$$\mathbf{A} = \begin{pmatrix} \frac{k_2}{m_1} & -\frac{k_2}{\sqrt{m_1 m_2}} & 0 & \dots & 0 & 0 \\ -\frac{k_2}{\sqrt{m_1 m_2}} & \frac{k_2}{m_2} + \frac{k_3}{m_2} & \frac{k_3}{\sqrt{m_2 m_3}} & \dots & 0 & 0 \\ 0 & -\frac{k_3}{\sqrt{m_2 m_3}} & \frac{k_3}{m_3} + \frac{k_4}{m_3} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \frac{k_{n-1}}{m_{n-1}} + \frac{k_n}{m_{n-1}} & -\frac{k_n}{\sqrt{m_{n-1} m_n}} \\ 0 & 0 & 0 & \dots & -\frac{k_n}{\sqrt{m_{n-1} m_n}} & \frac{k_n}{m_n} \end{pmatrix} \quad (3.72)$$

che ha la forma di una matrice reale tridiagonale simmetrica. L'equazione di Schrödinger prende la forma

$$\mathcal{H}\Psi(\boldsymbol{\xi}) = E\Psi(\boldsymbol{\xi}) \quad (3.73)$$

$$\mathcal{H} = -\frac{1}{2} \sum_{i=1}^n \frac{\partial^2}{\partial \xi_i^2} + \frac{1}{2} \boldsymbol{\xi}^T \mathbf{A} \boldsymbol{\xi} \quad (3.74)$$

L'equazione descrive le autofunzioni di un sistema descritto da n coordinate accoppiate. Possiamo però definire gli autovettori e gli autovalori della matrice \mathbf{A}

$$\mathbf{A}\mathbf{C} = \mathbf{C}\boldsymbol{\Lambda} \quad (3.75)$$

dove $\mathbf{C}^T \mathbf{C} = \mathbf{1}$ e $\mathbf{C}^T \mathbf{A} \mathbf{C} = \boldsymbol{\Lambda}$, e le nuove coordinate, combinazioni lineari delle $\boldsymbol{\xi}$, $\mathbf{q} = \mathbf{C}^T \boldsymbol{\xi}$, dette *modi normali* della molecola. Nelle nuove coordinate possiamo verificare che l'equazione di Schrödinger diventa

$$\mathcal{H}\Psi(\mathbf{q}) = E\Psi(\mathbf{q}) \quad (3.76)$$

$$\mathcal{H} = -\frac{1}{2} \sum_{i=1}^n \frac{\partial^2}{\partial q_i^2} + \frac{1}{2} \mathbf{q}^T \boldsymbol{\Lambda} \mathbf{q} = \frac{1}{2} \sum_{i=1}^n \left(-\frac{\partial^2}{\partial q_i^2} + \lambda_i q_i^2 \right) = \sum_{i=1}^n \mathcal{H}_i \quad (3.77)$$

L'equazione ha una forma separabile, cioè l'hamiltoniano è dato dalla somma di n hamiltoniani, uno per ogni modo normale q_i . La soluzione può quindi essere scritta come il prodotto delle soluzioni per gli n modi normali.

Si può notare che: 1) lo spettro \mathbf{A} comprende autovalori non negativi; 2) di questi uno è nullo, e corrisponde al modo normale q che descrive il moto traslazionale (rigido) della molecola lungo la retta.

Per fissare meglio le idee, consideriamo il caso di una molecola lineare triatomica, con $k_2 = k_3 = k$ e $m_1 = m_2 = m_3 = m$. La matrice \mathbf{A} assume la forma:

$$\mathbf{A} = \begin{pmatrix} \frac{k}{m} & -\frac{k}{m} & 0 \\ -\frac{k}{m} & 2\frac{k}{m} & -\frac{k}{m} \\ 0 & -\frac{k}{m} & \frac{k}{m} \end{pmatrix} \quad (3.78)$$

Gli autovalori sono $\lambda_1 = 0$, $\lambda_2 = \frac{k}{m}$ e $\lambda_3 = 3\frac{k}{m}$; il primo autovalore, nullo, corrisponde al modo normale $q_1 = \frac{1}{\sqrt{3}}(\xi_1 + \xi_2 + \xi_3)$ (moto rigido traslazionale), mentre gli altri due autovalori corrispondono a modi normali di vibrazione caratterizzati da una frequenza propria rispettivamente pari a $\omega = \sqrt{\frac{k}{m}}$ e $\sqrt{3}\omega$.

3.2.5 Metodi numerici

Normalmente, i problemi agli autovalori tipici della chimica quantistica sono caratterizzati da matrici di dimensioni che vanno da poche decine a molti milioni. I metodi di risoluzione numerica implementati direttamente in molti applicativi, sia proprietari che open source, fanno uso di algoritmi complessi, che si basano su alcune tecniche generali volte alla riduzione della matrice analizzata in forme più maneggevoli (per esempio, tipicamente, in forma tridiagonale), per il calcolo diretto analitico o semianalitico degli autovalori e degli autovettori, o di osservabili dipendenti dagli autovalori ed autovettori. Per applicazioni avanzate, l'uso di librerie di calcolo specifiche all'interno di codici sviluppati *ad hoc* è di prammatica. Ricordiamo tra tutte **EISPACK** (in Fortran), e **LAPACK** (in C++, Fortran90) disponibili presso www.netlib.org, le librerie proprietarie **IMSL**, **NAG** etc.

In pratica, la selezione della routine di soluzione per il problema di interesse dipende dal problema in esame, cioè in ultima analisi dalla natura della matrice che deve essere diagonalizzata: dal caso più semplice di una matrice reale simmetrica definita sino al caso più complicato di una matrice complessa non hermitiana. Anche il tipo di esito può essere diverso da problema a problema: esistono algoritmi, per esempio, per il calcolo di soli autovalori, ed anche di soli autovalori estremi (minimi o massimi). Per esempio, possiamo essere interessati alle sole proprietà di stato fondamentale di una molecola, nel qual caso il calcolo dell'autovalore minimo è sufficiente. Come per la soluzione dei sistemi di equazioni lineari (ma ciò non è sorprendente, perchè i due problemi sono in fondo intimamente collegati), il caso delle matrici sparse dovrebbe poi essere sempre trattato in modo diverso dal caso delle matrici dense, per trarre vantaggio dalla loro struttura particolare e poter trattare problemi di dimensioni molto grandi.

Nel caso di matrici reali simmetriche gli algoritmi a disposizione sono numerosi, stabili e molto efficienti. Per matrici aventi dimensioni dell'ordine di qualche decina al massimo, può essere utile l'algoritmo di Jacobi, un approccio iterativo che trasforma la matrice originaria in una matrice diagonale mediante una serie di *trasformazioni di similitudine*

$$\mathbf{D}_k = \mathbf{G}_k \mathbf{D}_{k-1} \mathbf{G}_k^T \quad (3.79)$$

la matrice (ortogonale) \mathbf{G}_k ha la forma di una *matrice di Givens*

$$\mathbf{G}(i, j, \theta) = \begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ & & & & & & \\ 0 & \dots & \cos \theta & \dots & \sin \theta & \dots & 0 \\ & & & & & & \\ 0 & \dots & -\sin \theta & \dots & \cos \theta & \dots & 0 \\ & & & & & & \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \quad (3.80)$$

dove il coseno ed il seno di un angolo θ compaiono all'intersezione dell' i -esima riga e della j -esima colonna (in breve, nel seguito, posizione ij). Una matrice di Givens applicata ad un generico vettore \mathbf{x} lo ruota di un angolo θ nel piano (i, j) . Applicata alla matrice simmetrica \mathbf{D}_{k-1} la trasformazione produce una nuova matrice simmetrica \mathbf{D}_k con l'elemento nella posizione ij dato dall'espressione

$$(\mathbf{D}_k)_{ij} = (\mathbf{D}_{k-1})_{ij} \cos 2\theta + \frac{1}{2} [(\mathbf{D}_{k-1})_{ii} - (\mathbf{D}_{k-1})_{jj}] \sin 2\theta$$

se si sceglie $\tan 2\theta = \frac{2(\mathbf{D}_{k-1})_{ij}}{(\mathbf{D}_{k-1})_{ii} - (\mathbf{D}_{k-1})_{jj}}$, l'elemento $(\mathbf{D}_k)_{ij}$ va a zero. L'algoritmo quindi procede nel modo seguente

1. ad ogni nuova iterazione, si cerca l'elemento più grande in valore assoluto della matrice $(\mathbf{D}_{k-1})_{ij}$, detto (come in precedenza) *pivot*
2. si applica la trasformazione di similitudine per ottenere la nuova matrice \mathbf{D}_k con $(\mathbf{D}_k)_{ij}$ azzerato

Per $k \rightarrow \infty$ la matrice \mathbf{D}_k tende ad una matrice diagonale, cioè alla matrice degli autovalori (la matrice dei prodotti delle trasformazioni tende alla matrice degli autovettori) della matrice originaria.

Un approccio simile consiste nella riduzione, tramite una serie di trasformazioni, della matrice originaria ad una forma tridiagonale. Le matrici tridiagonali (vedi sopra) sono particolarmente 'facili' da diagonalizzare, con algoritmi ad hoc. Di conseguenza, la strategia numerica di elezione nel caso si debba risolvere un problema agli autovalori con una matrice di dimensioni superiore al centinaio, e come vedremo fra breve soprattutto nel caso di matrici molto grandi sparse, è quella 1) di portare la matrice ad una forma tridiagonale e 2) di risolvere il problema agli autovalori con la nuova matrice tridiagonale. Una delle tecniche più usate per la tridiagonalizzazione è la procedura di Householder, una serie di trasformazioni che sistematicamente annulla la parte di colonna i -esima e di riga i -esima corrispondente alla posizione diagonale ii , lasciando solo come elementi non nulli quelli appartenenti alla diagonale superiore ed inferiore, portando quindi alla fine alla formazione di una matrice tridiagonale.

Una metodologia per la tridiagonalizzazione di matrici *sparse* particolarmente efficiente è invece dato dalla famiglia di approcci basati sull'algoritmo di Lanczos, che per la sua facilità di implementazione è applicato con frequenza alla soluzione di problemi agli autovalori nell'ambito della chimica quantistica, direttamente o in forme modificate.

Approfondimento 8 . Supponiamo di voler trasformare la matrice generica \mathbf{A} (reale simmetrica per semplicità) in forma tridiagonale. Sia \mathbf{v}_1 un vettore che possiamo anche scegliere in modo arbitrario. Generiamo la sequenza di vettori \mathbf{v}_k come segue

$$\begin{aligned}\beta_{k+1}\mathbf{v}_{k+1} &= (\mathbf{A} - \alpha_k)\mathbf{v}_k - \beta_k\mathbf{v}_{k-1} \\ \alpha_k &= \mathbf{v}_k^T \mathbf{A} \mathbf{v}_k \\ \beta_k &= \mathbf{v}_k^T \mathbf{A} \mathbf{v}_{k-1}\end{aligned}\tag{3.81}$$

I vettori $\mathbf{v}_1, \mathbf{v}_2$ etc. sono ortonormali (almeno in precisione infinita, cioè in assenza di errori di arrotondamento) e formano una matrice ortogonale $\mathbf{V} = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)^T$ tale che $\mathbf{V}\mathbf{A}\mathbf{V}^T = \mathbf{T}$, dove \mathbf{T} è la matrice tridiagonale (simmetrica)

$$\mathbf{T} = \begin{pmatrix} \alpha_1 & \beta_1 & & & \\ \beta_1 & \alpha_2 & \beta_2 & & \\ & \beta_2 & \ddots & \ddots & \\ & & \ddots & \ddots & \beta_{n-1} \\ & & & \beta_{n-1} & \alpha_n \end{pmatrix}\tag{3.82}$$

Supponiamo di voler calcolare la grandezza scalare $j(s) = \mathbf{v}_1(s + \mathbf{A})^{-1}\mathbf{v}_1$ (si tratta di un'espressione, detta nel contesto *densità spettrale*, che si ottiene spesso nell'ambito della spettroscopia computazionale come risultato di elaborazione teoriche volte a calcolare sull'asse delle frequenze generiche s un segnale spettroscopico. In questo caso si ha che $j(s) = \mathbf{e}_1^T(s + \mathbf{A})^{-1}\mathbf{v}_1$ dove \mathbf{e}_1 è la colonna $\mathbf{e}_1 = (1, 0, \dots, 0)^T$. L'espressione si può scrivere direttamente sotto forma di frazione continua

$$j(s) = \frac{1}{s - \alpha_1 - \frac{\beta_1^2}{s - \alpha_2 - \frac{\beta_2^2}{s - \alpha_3 + \dots}}}\tag{3.83}$$

I coefficienti della matrice \mathbf{T} si ottengono direttamente dalla sequenza di operazioni di moltiplicazione matrice-vettore (3.81): opportunamente ottimizzati nel caso di matrici sparse, questi prodotti sono computazionalmente poco costosi. La matrice tridiagonale ottenuta può essere diagonalizzata impiegando algoritmi specifici, anch'essi a basso costo computazionale, come per esempio il metodo di decomposizione QR o QL. Il metodo di decomposizione QR (QL) permette di ridurre una matrice generica alla forma $\mathbf{A} = \mathbf{QR}$ ($\mathbf{A} = \mathbf{QL}$, dove \mathbf{Q} è ortogonale e \mathbf{R} (\mathbf{L}) è triangolare superiore (inferiore)). La procedura di decomposizione è accompagnata dalla generazione di una nuova matrice simile alla precedente usando \mathbf{Q} come matrice di trasformazione. La sequenza tende a generare la matrice diagonale degli autovalori. In pratica si procede come segue

1. si decompone $\mathbf{A}_0 = \mathbf{A} = \mathbf{QR}$
2. si genera $\mathbf{A}_1 = \mathbf{RQ} = \mathbf{Q}^T \mathbf{A}_0 \mathbf{Q}$
3. si continua decomponendo ad ogni iterazione $\mathbf{A}_k = \mathbf{Q}_k \mathbf{R}_k$ e generando $\mathbf{A}_{k+1} = \mathbf{R}_k \mathbf{Q}_k = \mathbf{Q}_k^T \mathbf{A}_k \mathbf{Q}_k$

Se per una matrice 'normale' di dimensioni n il costo computazionale cresce come $O(n^3)$, nel caso di una matrice tridiagonale il costo è solo $O(n)$ rendendo l'algoritmo molto conveniente.

Capitolo 4

Integrazione di funzioni

4.1 Definizioni

In questo (breve) capitolo saranno discussi i metodi di integrazione analitici e numerici più comuni, come sempre nel contesto di alcuni problemi di interesse generale per il chimico.

4.1.1 Integrali indefiniti, definiti, generalizzati, multipli

Richiamiamo prima di tutto alcuni concetti di base. L'integrale *indefinito*

$$F(x) = \int f(x)dx \quad (4.1)$$

è una possibile soluzione dell'equazione differenziale $\frac{dF}{dx} = f$, cioè una funzione la cui derivata prima è f (detta funzione integranda). L'*integrale definito*

$$I = \int_a^b f(x)dx \quad (4.2)$$

è un numero, definito come $F(b) - F(a)$, interpretabile geometricamente come l'area sottesa dalla curva $y = f(x)$ nell'intervallo $[a, b]$.

Esempio 25. Consideriamo una sostanza, per esempio acqua od anidride carbonica, racchiusa in un recipiente a pressione costante, ad una data temperatura T_i . vogliamo conoscere l'entropia della sostanza ad una temperatura T_f . Se il sistema subisce N transizioni di fase prima della temperatura T_f possiamo scrivere in generale

$$S(T_f) = S(T_i) + \sum_{n=1}^N \left[\int_{T_{n-1}}^{T_n} \frac{C_p(n)}{T} dT + \frac{\Delta H_n}{T_n} \right] + \int_{T_N}^{T_f} \frac{C_p(N+1)}{T} dT$$

dove T_n e ΔH_n sono la temperatura di transizione e la variazione di entalpia per la transizione n -esima, $C_p(n)$ è la capacità termica per la fase n -esima.

Esempio 26. Il coefficiente di fugacità di una gas reale è

$$\ln \gamma = \int_0^{p_{\text{mis}}} \left(\frac{Z-1}{p} \right) dp \quad T \text{ cost.}$$

dove γ è il coefficiente di fugacità a T e p_{mis} , e Z è il fattore di compressibilità. Una serie di misure di Z a T fissata variando p fino a p_{mis} fornisce il modo per valutare γ .

Un integrale *generalizzato* è definito come il limite di un integrale definito. Un primo caso è il seguente: siano a e b reali tali che $a < b$, e la funzione sia f definita in $]a, b[$ (cioè nell'intervallo che va da a a b , escluso a stesso); f sia invece illimitata in un intorno (destra) di a , cioè in pratica tende all'infinito per $x \rightarrow a+$; allora si dice che f è integrabile in senso generalizzato se esiste

$$I = \lim_{x \rightarrow a+} \int_a^b f(x) dx \quad (4.3)$$

se I è finito l'integrale converge, altrimenti se $I \rightarrow \infty$ l'integrale diverge. Un secondo caso è invece con estremi di integrazione infiniti: se f è per esempio definita in $] - \infty, b[$

$$I = \lim_{x \rightarrow -\infty} \int_a^b f(x) dx \quad (4.4)$$

in pratica di solito si scrive semplicemente $\int_a^b f(x) dx$, $\int_{-\infty}^b f(x) dx$.

Esempio 27. L'entropia assoluta alla temperatura T di una sostanza che tenda ad un cristallo perfetto per $T \rightarrow 0$ si scrive, sulla base del III principio della termodinamica, in assenza di transizioni di fase per temperature inferiori a T , come

$$S(T) = \int_0^T \frac{C_p(T)}{T} dT$$

l'espressione ha la forma di un integrale generalizzato, poichè apparentemente l'integrando diverge per $T \rightarrow 0$ (in pratica, C_p per $T \rightarrow 0$ va a zero in modo superlineare rispetto a T , quindi $C_p(T)/T$ tende a zero per $T \rightarrow 0$)

Infine, ricordiamo la definizione di *integrale multiplo*. Data una funzione $f(\mathbf{x})$ definita per semplicità in \mathcal{R}^n ed un dominio $D \subseteq \mathcal{R}^n$, definiamo l'integrale multiplo I come

$$I = \int_{\mathbf{x} \in D} f(\mathbf{x}) d\mathbf{x} \quad (4.5)$$

Gli integrali multipli in \mathcal{R}^2 e \mathcal{R}^3 sono detti rispettivamente integrali *doppi* e *tripli*

$$I = \int_{\mathbf{x} \in D} f(\mathbf{x}) d\mathbf{x} = \iint_{\mathbf{x} \in D} f(\mathbf{x}) dx_1 dx_2, \quad I = \int_{\mathbf{x} \in D} f(\mathbf{x}) d\mathbf{x} = \iiint_{\mathbf{x} \in D} f(\mathbf{x}) dx_1 dx_2 dx_3 \quad (4.6)$$

Esempio 28. Per una reazione all'equilibrio in fase gassosa abbiamo la relazione detta *equazione di van't Hoff*:

$$\frac{d \ln K}{dT} = \frac{\Delta_r H^\ominus(T)}{RT^2}$$

In forma integrale la relazione precedente assume la forma

$$\ln K(T_2) = \ln K(T_1) + \frac{1}{R} \int_{T_1}^{T_2} \frac{\Delta_r H^\ominus(T)}{T^2} dT$$

Sappiamo inoltre che vale la *legge di Kirkhoff*

$$\frac{d \Delta_r H^\ominus}{dT} = \frac{d}{dT} \sum_i \nu_i H_{m,i} = \sum_i \nu_i C_{p,m_i}^\ominus(T) = \Delta_r C_{p,m}^\ominus(T)$$

quindi

$$\Delta_r H^\ominus(T) = \Delta_r H^\ominus(T_0) + \int_{T_0}^T \Delta_r C_{p,m}^\ominus(T) dT$$

da cui si trova che la costante di equilibrio alla temperatura T_2 è ottenibile come

$$\ln K(T_2) = \ln K(T_1) + \frac{\Delta_r H^\ominus(T_0)}{R} \left(\frac{1}{T_1} - \frac{1}{T_2} \right) + \frac{1}{R} \int_{T_1}^{T_2} \left[\int_{T_0}^{T'} \frac{\Delta_r C_{p,m}^\ominus(T')}{T'^2} dT' \right] dT$$

4.1.2 Metodi di integrazione analitici

Ricordiamo ora alcuni metodi semplici che permettono in vari casi elementari di integrare esattamente una funzione. Diamo per scontato che siano già note i principali integrali elementari (nel seguito indichiamo il risultato di un integrale indefinito generico a meno di una costante arbitraria). In alcuni casi, semplici

$f(x)$	$F(x)$
x^a	$\frac{x^{a+1}}{a+1} \quad a \neq -1$
$\frac{1}{x}$	$\ln x$
a^x	$\frac{a^x}{\ln a}$
$\sin x$	$-\cos x$
$\cos x$	$\sin x$
$\frac{1}{\sin^2 x}$	$-\cot x$
$\frac{1}{\cos^2 x}$	$-\tan x$
$\frac{1}{1+x^2}$	$\arctan x$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin x$

Tabella 4.1: Alcuni integrali elementari

manipolazioni o sostituzioni di variabili consentono di ridurre un integrale ad una funzione di integrali noti. Ricordiamo prima di tutto la formula di *integrazione per parti*; se f e g sono due funzioni generiche e G è l'integrale di g :

$$\int f(x)g(x)dx = f(x)G(x) - \int f'(x)G(x)dx \quad (4.7)$$

che discende semplicemente dal fatto che $(fG)' = fg + f'G$.

Approfondimento 9 . L'integrazione per parti è utile con integrali del tipo

$$\int x^n e^{ax} dx \quad \int x^n \sin bx \quad \int x^n \cos bx$$

L'altra tecnica elementare utilizzata si basa su un cambio di variabile di integrazione. Per un generico integrando $f(x)$, supponiamo di cambiare la variabile indipendente esprimendola in funzione di un'altra variabile, $x = x(q)$, da cui $dx = \frac{dx}{dq} dq$. L'integrale indefinito generico diventa

$$I = \int f(x) dx = \int f[x(q)] \frac{dx}{dq} dq = \int g(q) dq \quad g(q) = f[x(q)] \frac{dx}{dq} \quad (4.8)$$

Esempio 29. Consideriamo una serie di misure sperimentali per un osservabile fisico, affette da errori casuali dimostrare che se la distribuzione di valori misurati segue una distribuzione gaussiana centrata sul valore μ e di ampiezza σ , il valore medio misurato è μ . Si scrive:

$$\bar{x} = \int_{-\infty}^{+\infty} \frac{x}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

ponendo $q = (x - \mu)/\sigma$, cioè $dq = dx/\sigma$

$$\begin{aligned} \bar{x} &= \int_{-\infty}^{+\infty} (x - \mu + \mu) \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx = \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} q \exp\left(-\frac{q^2}{2}\right) dq + \mu \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left(-\frac{q^2}{2}\right) dq = 0 + \mu = \mu \end{aligned}$$

dove si abbiamo tenuto conto del fatto che il primo dei due integrali in q è nullo (integrando dispari ...) e che il secondo è un integrale generalizzato noto

$$\int_{-\infty}^{+\infty} \exp\left(-\frac{x^2}{2}\right) = \sqrt{\pi}$$

Nel caso di integrali multipli, la trasformazione di variabili richiede l'introduzione del *jacobiano*. Dato l'integrale multiplo generico (4.5) poniamo $x_1 = x_1(q_1, \dots, q_n), \dots, x_n = x_n(q_1, \dots, q_n)$ ovvero $\mathbf{x} = \mathbf{x}(\mathbf{q})$. L'integrale diventa

$$I = \int_{\mathbf{q} \in D'} f[\mathbf{x}(\mathbf{q})] |J| d\mathbf{q} \quad (4.9)$$

dove D' è il dominio corrispondente a D , ma nello spazio \mathbf{q} , mentre J è il determinante della matrice delle derivate prime di $\mathbf{x} = \mathbf{x}(\mathbf{q})$, detto appunto jacobiano

$$J = \frac{\partial \mathbf{x}}{\partial \mathbf{q}} = \begin{vmatrix} \frac{\partial x_1}{\partial q_1} & \frac{\partial x_2}{\partial q_1} & \cdots & \frac{\partial x_n}{\partial q_1} \\ \frac{\partial x_1}{\partial q_2} & \frac{\partial x_2}{\partial q_2} & \cdots & \frac{\partial x_n}{\partial q_2} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial x_1}{\partial q_n} & \frac{\partial x_2}{\partial q_n} & \cdots & \frac{\partial x_n}{\partial q_n} \end{vmatrix} \quad (4.10)$$

Concludiamo questa sezione con alcuni metodi analitici di integrazione per classi di funzioni integrande.

Dati due generici polinomi $A(x)$ e $B(x)$, consideriamo l'integrale della funzione razionale formata dal rapporto A/B

$$I = \int_a^b \frac{A(x)}{B(x)} dx \quad (4.11)$$

prima di tutto notiamo che se il grado di A è maggiore di quello di B possiamo sempre scrivere la funzione integranda come la somma di un polinomio e di una funzione razionale di due polinomi, in cui il polinomio al numeratore ha un grado minore di quello al denominatore. Quindi l'integrazione della funzione A/B si può ricondurre ad un integrale triviale (integrando polinomiale) sommato all'integrale di una funzione razionale con grado del numeratore inferiore al grado del denominatore. Assumeremo perciò d'ora in poi che il grado di A sia minore di quello di B . Ora è possibile dimostrare che

$$B(x) = b_0(x - \alpha_1)^{p_1} \dots (x - \alpha_m)^{p_m} (x^2 + \beta_1 x + \gamma_1)^{q_1} \dots (x^2 + \beta_n x + \gamma_n)^{q_n} \quad (4.12)$$

dove $p_1 + \dots + p_m + 2q_1 + \dots + 2q_n$ è uguale al grado di B . Da ciò discende che la funzione A/B si può scrivere

$$\begin{aligned} \frac{A(x)}{B(x)} &= \frac{A_1^{(1)}}{x - \alpha_1} + \frac{A_2^{(1)}}{(x - \alpha_1)^2} + \dots + \frac{A_{p_1}^{(1)}}{(x - \alpha_1)^{p_1}} + \dots \\ &+ \frac{A_1^{(m)}}{x - \alpha_m} + \frac{A_2^{(m)}}{(x - \alpha_m)^2} + \dots + \frac{A_{p_m}^{(m)}}{(x - \alpha_m)^{p_m}} + \\ &+ \frac{B_1^{(1)}x + C_1^{(1)}}{x^2 + \beta_1x + \gamma_1} + \frac{B_2^{(1)}x + C_2^{(1)}}{(x^2 + \beta_1x + \gamma_1)^2} + \dots + \frac{B_{q_1}^{(1)}x + C_{q_1}^{(1)}}{(x^2 + \beta_1x + \gamma_1)^{q_1}} + \dots \\ &+ \frac{B_1^{(n)}x + C_1^{(n)}}{x^2 + \beta_nx + \gamma_n} + \frac{B_2^{(n)}x + C_2^{(n)}}{(x^2 + \beta_nx + \gamma_n)^2} + \dots + \frac{B_{q_n}^{(n)}x + C_{q_n}^{(n)}}{(x^2 + \beta_nx + \gamma_n)^{q_n}} + \dots \end{aligned} \quad (4.13)$$

Le costanti $A_1^{(1)}, \dots, C_{q_n}^{(n)}$ si trovano facendo il minimo comune denominatore dell'espressione a secondo membro, e confrontando il denominatore ottenuto con $A(x)$. L'integrale di A/B è quindi ricondotto ad una somma di integrali di funzioni relativamente semplici, tutti calcolabili per parti o per sostituzione di variabili.

Approfondimento 10 . In molti casi è possibile ricondurre forme anche complicate dell'integrando a funzioni razionali. Se per esempio l'integrando è una funzione razionale $f(x, x^{m_1/n_1}, \dots, x^{m_r/n_r})$ con m_1, n_1, \dots numeri interi, basta effettuare la sostituzione $q = x^\rho$, dove ρ è il minimo comune multiplo di n_1, \dots, n_r per ottenere un integrale con un integrando in forma di funzione razionale in q . In altri casi, una combinazione di sostituzione di variabili ed integrazione per parti può essere d'aiuto: per esempio, se l'integrando ha la forma generica di una funzione razionale con argomenti dati da seni e coseni, $f(\sin x \cos x)$, la trasformazione $q = \tan x/2$ permette di ottenere un integrale di una funzione razionale in q ; se l'integrando è una funzione razionale di e^{nx} , si può usare la trasformazione $q = e^{nx}$ etc.

4.2 Metodi numerici di integrazione

Accenniamo infine al problema del calcolo numerico di integrali definiti. In generale se vogliamo calcolare l'integrale definito (4.2), possiamo dividere l'intervallo $[a, b]$ in sottointervalli $[x_1 = a, x_2], [x_2, x_3], \dots, [x_{n-1}, x_n = b]$ ciascuno della stessa lunghezza h , ed approssimare l'arco della funzione $f(x)$ in ciascun intervallo con una retta che va dal punto $(x_i, f(x_i) = f_i)$ al punto $(x_{i+1}, f(x_{i+1}) = f_{i+1})$; l'integrale può essere approssimato alla somma delle aree dei trapezi adiacenti così formati. Il risultato molto noto che si ottiene è la cosiddetta formula dei trapezi

$$I = h \left(\frac{f_1}{2} + f_2 + \dots + f_{n-1} + \frac{f_n}{2} \right) \quad (4.14)$$

questa formula o espressioni simili ricavate su considerazioni analoghe, è alla base della maggior parte dei metodi di integrazione numerica di integrali semplici, almeno per funzioni relativamente 'lisce' di cui sia nota la forma funzionale o i valori distribuiti su una griglia regolare. Possiamo utilizzare schemi più complessi, per esempio considerare terne successive, invece di coppie successive di punti, trovare la parabola che passa per essi ed integrare in ciascun intervallo $[x_{i-1}, x_i, x_{i+1}]$, in questo caso si trova

$$I = h \left(\frac{f_1}{3} + \frac{4f_2}{3} + \frac{2f_3}{3} + \dots + \frac{2f_{n-2}}{3} + \frac{4f_{n-1}}{3} + \frac{f_n}{3} \right) \quad (4.15)$$

e così via. L'algoritmo più comune (detto di Romberg) usato in molte routine numeriche in effetti funziona costruendo approssimazioni successive all'integrale (4.2) aumentando via via l'ordine del polinomio interpolante, fino al raggiungimento della convergenza numerica del valore calcolato dell'integrale rispetto a qualche parametro di accuratezza designato. Concludiamo questa sezione accennando in ordine sparso ad alcuni problemi comuni

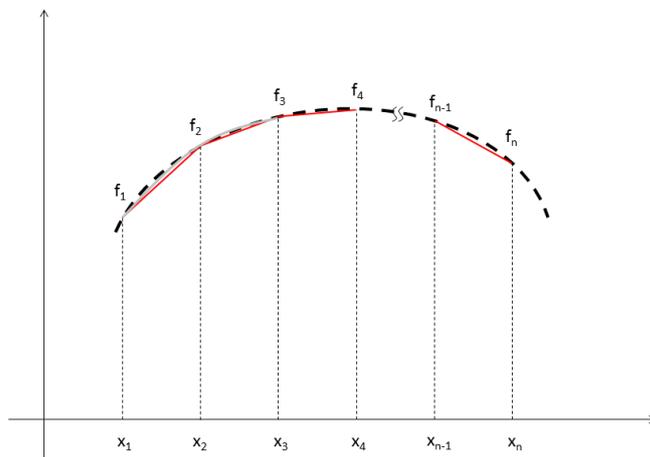


Figura 4.1: Schema di discretizzazione per la formula dei trapezi

1. altre tecniche di calcolo sono basate sull'idea di approssimare la funzione integranda ad una somma di funzioni note, i cui integrali siano semplici da calcolare (per esempio un'insieme di polinomi ortonormali, vedi l'Appendice); il lettore attento avrà compreso che qui la procedura è intimamente legata all'idea di spazio vettoriale di funzioni (f è proiettata su una base di funzioni, i coefficienti della proiezione sono i coefficienti di Fourier etc.)
2. nel caso di un integrale generalizzato, è meglio procedere se possibile con una trasformazione di variabile per rendere l'integrale proprio
3. il caso degli integrali multipli può essere, nel caso di una bassa dimensionalità, affrontato banalmente come l'applicazione dei metodi fin qui discussi per gli integrali semplici (l'integrale di un integrale di un integrale ...), ma la procedura diventa rapidamente proibitiva da un punto di vista computazionale; molto più utile è invece l'idea di scrivere l'integrando come somma di funzioni note facilmente integrabili; per esempio funzioni gaussiane. In effetti, una delle applicazioni principali di questa metodologia è ben nota al chimico computazionale: i complessi integrali tri ed esadimensionali di Coulomb e di scambio che intervengono nella determinazione della struttura elettronica di un sistema molecolare sono di solito ricondotti a somme di integrali di funzioni gaussiane, per esempio esprimendo gli orbitali atomici monoelettronici che intervengono nella definizione dell'integrando come somme di poche funzioni gaussiane (*Gaussian type orbitals*, GTO).¹

Applicazione 4 . Il calore molare del disolfuro di carbonio è stato misurato a varie temperature. Il problema presenta un classico esempio di misura sperimentale dell'entropia assoluta di una sostanza pura, basata sull'applicazione diretta del II e III principio. Alla temperatura finale il disolfuro di carbonio è liquido. Possiamo calcolare l'entropia assoluta come

$$S_m(298.15) = S_m(0) + \int_0^{161.11} \frac{C_{p,m}(s)}{T} dT + \frac{1049.0}{161.11} + \int_{161.11}^{298.15} \frac{C_{p,m}(l)}{T} dT$$

L'integrazione a partire da 0 fino alla temperatura di fusione è sostituita da una integrazione da 0 alla prima temperatura misurata, in cui il calore molare viene stimato in base alla legge di Debye

$$\int_0^{161.11} \frac{C_p(s)}{T} dT = \int_0^{T_1} \frac{C_{p,m}}{T} + \int_{T_1}^{161.11} \frac{C_p(s)}{T} dT$$

¹Sono possibili metodi ad hoc per problemi di grandi dimensioni (per esempio il metodo Monte Carlo)

$T \text{ K}^{-1}$	$C_{p,m} \text{ cal}^{-1} \text{ K mol}$
15.05	1.65
20.15	2.87
29.76	4.96
42.22	6.97
57.51	8.50
75.54	9.57
89.37	10.31
99.00	10.98
108.93	11.59
131.4	12.58
144.31	13.05
156.83	13.53
163.93	18.10
192.30	17.91
227.34	17.93
269.69	18.08
278.22	18.06
297.43	18.17

Tabella 4.2: Calore molare a pressione costante a varie temperature per il CS_2 .

Il contributo iniziale viene stimato come

$$\int_0^{T_1} \frac{C_{p,m}}{T} \approx \int_0^{T_1} aT^2 = a \frac{T_1^3}{3} \approx \frac{C_{p,m}(T_1)}{3} = \frac{1.65}{3} = 0.55$$

I contributi integrali restanti si possono ottenere interpolando i dati disponibili dei calori molari ed integrando numericamente. Infine in base al terzo principio si pone $S(0) = 0$. Si ottiene $S(298.15) = 36.0 \text{ cal K}^{-1}$.

Capitolo 5

Equazioni non lineari

5.1 Definizioni

Questo capitolo è dedicato ai metodi più comuni di soluzione del seguente problema: dato il sistema di n equazioni in n incognite (reali o complesse)

$$\mathbf{f}(\mathbf{x}) = \mathbf{0} \quad (5.1)$$

si vogliono trovare una o più soluzioni, cioè vettori \mathbf{x} che verifichino la (5.1). In generale, queste soluzioni, se esistono, si possono trovare esclusivamente con metodi numerici, a parte pochi casi, il principale dei quali è dato dai sistemi di equazioni lineari, descritti nei capitoli (2) e (3). Le difficoltà intrinseche del problema di trovare le soluzioni (o 'radici') del sistema (5.1) sono inoltre molto diverse se il problema è monodimensionale (una sola equazione) o multidimensionale. Nel primo caso è possibile praticamente sempre trovare soluzioni accurate in un dato intervallo $[a, b]$; nel secondo la ricerca di alcune delle possibili radici in un dominio generico D può essere molto complicata e computazionalmente 'difficile'.

I problemi riconducibili al sistema (5.1) che uno studente di chimica può incontrare durante il suo corso di studi sono molteplici. Nel campo della termodinamica ricordiamo per esempio la determinazione delle proprietà di una sostanza data una funzione di stato oppure, per i sistemi reattivi all'equilibrio, la ricerca delle attività delle specie chimiche in un reattore in cui siano presenti uno o più processi stechiometrici indipendenti; in entrambi i casi ci si trova quasi sempre nella necessità di risolvere sistemi di equazioni non lineari. In un contesto molto diverso, la determinazione dei luoghi dei punti nodali (punti, piani etc.) di orbitali elettronici o molecolari richiede il calcolo degli zeri delle funzioni che descrivono le funzioni d'onda orbitaliche.

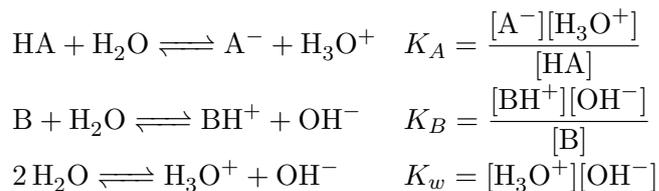
Esempio 30. Vogliamo calcolare il volume occupato da 1 mole di CO_2 , descritta dall'equazione di van der Waals, ad 1 atm e 300 K. Il problema è risolvibile analiticamente, ma richiede la soluzione di un'equazione cubica. Infatti, partendo dall'equazione di stato otteniamo che il volume molare di un gas di van der Waals verifica la relazione

$$V_m^3 - \frac{RT + pb}{p}V_m^2 + \frac{a}{p}V_m - \frac{ab}{p} = 0$$

che ammette *tre* soluzioni. Conviene risolvere l'equazione numericamente, ma è possibile anche (vedi oltre) scrivere le soluzioni esatte in termini di quadrature (operazioni elementari e radici; vedi oltre). Si verifica che delle tre soluzioni, una sola è reale e coincide con il volume molare. Per i parametri dati si ottiene $V_m = 23.98 \text{ L mol}^{-1}$.

Esempio 31. Consideriamo un esempio basato sugli equilibri chimici. Supponiamo di avere una soluzione acquosa contenente un acido debole monoprotico HA, di molarità m_A per un volume V_A ; aggiungiamo un volume V_B di una soluzione acquosa di una base debole B, di molarità m_B . Calcoliamo il pH della soluzione risultante, note la k_A dell'acido e la k_B della base.

Supponiamo di essere in soluzioni ideali diluite e di poter quindi descrivere gli equilibri di reazione in termini di molarità (coefficienti di attività unitari e volume di mescolamento nullo). Scriviamo gli equilibri di reazione



Cosa sappiamo inoltre del sistema all'equilibrio? Devono valere i bilanci di massa rispetto alle specie A e B,

$$\begin{aligned} (V_A + V_B)([\text{HA}] + [\text{A}^-]) &= V_A m_A \\ (V_A + V_B)([\text{B}] + [\text{BH}^+]) &= V_B m_B \end{aligned}$$

e la soluzione deve essere elettricamente neutra $[\text{H}_3\text{O}^+] + [\text{BH}^+] = [\text{A}^-] + [\text{OH}^-]$; abbiamo quindi un sistema di sei equazioni in sei incognite, che si può risolvere (vedi oltre) ottenendo $[\text{H}_3\text{O}^+]$ e quindi il $\text{pH} = -\log[\text{H}_3\text{O}^+]$.

5.2 Soluzioni esatte

Il calcolo delle soluzioni del sistema (5.1) per quadrature (cioè con un numero finito di operazioni elementari e radicali) è nella maggior parte impossibile. Anche quando una soluzione esatta analitica esista, spesso il suo impiego è, a tutti gli effetti, non pratico, e conviene ricorrere a soluzioni approssimate numeriche. Il caso più semplice ed emblematico da considerare è quello di una singola equazione polinomiale di grado n , della forma

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_2 x^2 + a_1 x + a_0 = 0 \quad (5.2)$$

Il *teorema fondamentale* dell'algebra stabilisce che la (5.2) ha un numero di radici complesse pari al grado del polinomio $p(x)$, contate con la loro molteplicità. È noto dal 1824, sulla base del *teorema di Abel*, che le soluzioni della (5.2) si possono trovare mediante quadrature in funzione dei coefficienti a_0, \dots, a_n solo se il grado è inferiore a cinque. È quindi possibili scrivere soluzioni analitiche esatte in termini di operazioni elementari per le equazioni lineari, quadratiche, cubiche e quartiche.

Approfondimento 11. Se i casi delle equazioni lineari e quadratiche sono ben noti, altrimenti non si può dire per il caso delle cubiche e delle quartiche. È istruttivo, per capire come la complessità di un problema possa crescere rapidamente in funzione di un solo parametro (in questo caso caso il grado), analizzare per curiosità insieme i casi $n = 2$ e $n = 3$

Quadratiche

Per $n = 2$, l'equazione è

$$ax^2 + bx + c = 0$$

dove a, b, c sono numeri reali; si definisce il discriminante

$$\Delta = b^2 - 4ac$$

tale che se

$\Delta > 0$: l'equazione ammette due soluzioni distinte reali

$\Delta = 0$: l'equazione ammette una soluzione doppia reale

$\Delta < 0$: l'equazione ammette due soluzioni complesse coniugate

Le radici in generale sono ottenute come

$$x_k = \frac{-b + u_k \sqrt{\Delta}}{2a}$$

con $k = 1, 2$; $u_1 = 1, u_2 = -1$ sono le due radici quadratiche dell'unità, e con $\sqrt{\Delta}$ si intende *una* delle due radici quadrate in senso complesso di Δ .

Cubiche

Per $n = 3$, l'equazione è

$$ax^3 + bx^2 + cx + d = 0$$

dove a, b, c, d sono numeri reali; si definisce il discriminante

$$\Delta = 18abcd - 4b^3d + b^2c^2 - 4ac^3 - 27a^2d^2$$

tale che se

$\Delta > 0$: l'equazione ammette tre soluzioni distinte reali

$\Delta = 0$: l'equazione ammette una soluzione tripla reale

$\Delta < 0$: l'equazione ammette una soluzione reale e due soluzioni complesse coniugate

Le radici in generale sono ottenute come

$$x_k = -\frac{1}{2a} \left(b + u_k C + \frac{\Delta_0}{u_k C} \right)$$

con $k = 1, 2, 3$; $u_1 = 1, u_2 = \frac{-1 + i\sqrt{3}}{2}, u_3 = \frac{-1 - i\sqrt{3}}{2}$ sono le tre radici cubiche dell'unità, $C =$

$\sqrt[3]{\frac{\Delta_1 + \sqrt{\Delta_1^2 - 4\Delta_0^3}}{2}}$ e $\Delta_0 = b^2 - 3ac, \Delta_1 = 2b^3 - 9abc + 27a^2d$ e i simboli di radice e radice cubica si intendono nel senso di *una* radice quadratica o cubica degli argomenti.

5.3 Soluzioni numeriche

Discutiamo in questa sezione alcuni metodi numerici per la soluzione di equazioni o sistemi di equazioni non lineari. Iniziamo dal caso di una equazione singola definita per una sola variabile x , vale a dire

$$f(x) = 0 \tag{5.3}$$

5.3.1 Una sola variabile

In generale la soluzione dell'equazione (5.3) procede sempre mediante la costruzione di una successione di approssimazioni successive, che possono non convergere necessariamente alla soluzione stessa; il metodo più semplice è il seguente. Supponiamo di scrivere $f(x) = g(x) - x$, cioè $x = g(x)$; sia a un punto di partenza e generiamo la successione

1. sia $x_1 = a$
2. sia $x_2 = g(x_1)$
3. torna allo step 2

quindi $x_{k+1} = g(x_k)$. Se $g(x)$ ha una derivata in modulo minore di 1 in un intorno della soluzione ξ ed a è scelto in questo intorno, la successione converge a ξ . In realtà questo semplicissimo approccio iterativo fallisce spesso, perchè di solito la condizione appena descritta non è verificata.

Metodo dicotomico

Un metodo lento ma sicuro è invece il *metodo dicotomico* o di bisezione. Ammettiamo di sapere che nell'intervallo $[a, b]$, l'equazione (5.3) ammetta una radice semplice ξ , cioè un'unica soluzione $f(\xi) = 0$, e che $f(a)f(b) < 0$. Il metodo più semplice per trovare una soluzione, detto dicotomico, è di costruire la successione

1. siano $x_1 = a, y_1 = b$
2. definiamo $z_1 = \frac{x_1 + y_1}{2}$, e scegliamo x_2, y_2 secondo la regola: se $f(x_1)f(z_1) < 0$ allora $x_2 = x_1, y_2 = z_1$, altrimenti se $f(z_1)f(y_1) < 0$ allora $x_2 = z_1, y_2 = y_1$
3. torna allo step 2.

La successione z_1, z_2, \dots converge a ξ (e naturalmente se ad un certo step $f(z_k) = 0, z_k = \xi$). il metodo dicotomico converge *sempre*, ma in modo relativamente lento (in effetti converge linearmente¹) ad una soluzione.

Metodi di Newton e *regula falsi*

Uno schema molto interessante è il metodo di Newton. Ancora una volta vogliamo generare una successione di punti x_k che convergano per $k \rightarrow \infty$ a ξ . Ad ogni iterazione possiamo supporre che $f(x_k + \delta x) = 0$ dove δx è una quantità incognita; al primo ordine in δx possiamo però espandere in serie di potenze la funzione f ottenendo

$$f(x_k + \delta x) = f(x_k) + \delta x f'(x_k)$$

dove $f'(x_k) = f'(x)|_{x=x_k}$; quindi una scelta corretta di δx è $-f(x_k)/f'(x_k)$;

1. sia $x_1 = a$
2. sia $x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}$
3. torna allo step 2

quindi ad ogni iterazione $x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$. Il metodo di Newton converge quasi sempre, ma non in ogni caso; la convergenza è *quadratica*. Il metodo richiede la valutazione della derivata prima della funzione oltre che della funzione stessa ad ogni iterazione. Un'alternativa è data dal metodo della secante variabile o *regula falsi*, in cui la derivata $f'(x_k)$ è calcolata come il rapporto incrementale nei punti x_k, x_{k-1} , $x_{k+1} = x_k - f(x_k) \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})}$. La *regula falsi* converge in modo superlineare (più lentamente del metodo di Newton, ma più velocemente del metodo dicotomico).

Applicazione 5. Nel Capitolo (7) si accenna alle funzioni d'onda che descrivono gli autostati dell'atomo di idrogeno a riposo. La parte radiale dipende dalle *funzioni associate di Laguerre* moltiplicate per l'esponenziale di $-r$. sono:

$$\Psi_{n,l,m_l}(r, \theta, \phi) = R_{nl}(r)Y_{lm_l}(\theta, \phi) \quad (5.4)$$

dove le funzioni radiali R_{nl} sono

$$R_{nl} = -\frac{2}{na} \frac{(n-l-1)!}{2n[(n+l)!]^{3/2}} \rho^l L_{n+l}^{2l+1}(\rho) \exp(-\rho/2) \quad (5.5)$$

¹Cioè la precisione di ogni nuova approssimazione è proporzionale alla precisione dell'approssimazione precedente

I punti nodali radiali dell'orbitale $\Psi_{n,l,m_l}(r, \theta, \phi)$ sono ottenuti per quei valori di r che rendono nulla la parte radiale dell'orbitale, quindi tali che $L_{n+l}^{2l+1}(\rho) = 0$. Le funzioni associate di Laguerre sono dei polinomi ottenuti ricorsivamente come

$$\begin{aligned} L_0^{(\alpha)}(x) &= 1 \\ L_1^{(\alpha)}(x) &= 1 + \alpha - x \\ L_{k+1}^{(\alpha)}(x) &= \frac{(2k+1+\alpha-x)L_k^{(\alpha)}(x) - (k+\alpha)L_{k-1}^{(\alpha)}(x)}{k+1} \end{aligned}$$

I polinomi di Laguerre semplici (vedi l'Appendice) sono casi speciali corrispondenti ad $\alpha = 0$ $L_n^{(0)}(x) = L_n(x)$.

Qual è l'approccio preferibile per la soluzione della (5.3)? In realtà, tutto considerato, la certezza della convergenza assicurata dal metodo dicotomico, che tra l'altro funziona bene anche con funzione aventi discontinuità nell'intervallo di partenza, lo rende un metodo molto apprezzabile. In una versione molto sofisticata, il metodo dicotomico è implementato in varie routine di calcolo di soluzione sotto forma di *algoritmo di Brent*. La maggior parte delle librerie di calcolo disponibili combinano spesso i vari metodi, accoppiando per esempio il metodo di Brent (per individuare un'approssimazione abbastanza vicina alla soluzione) con il metodo di Newton o regula falsi (per un raffinamento ulteriore dell'approssimazione); ricordiamo tra le altre le routine *brent* (in C), *dfzero* (in Fortran, del pacchetto NMS), *root* (in Fortran, del pacchetto Napack), tutte accessibili da gams.nist.gov.

5.3.2 Molte variabili

Come abbiamo già scritto sopra, il problema dell'individuazione delle soluzioni di un sistema di equazioni non lineari, (5.1), è molto più complesso del caso di una sola variabile. La complessità delle possibili soluzioni è tale che, se non si dispone di una buona approssimazione di partenza ad una specifica soluzione, ammesso che questa esista, non esistono metodi numerici sufficientemente robusti per risolvere il problema.

Tra le classi di metodi più utilizzati, con molta cautela, ricordiamo l'approccio *quasi-Newton*, che consiste sostanzialmente in una qualche generalizzazione a più dimensioni della regula falsi.

Possiamo generalizzare il metodo di Newton alla soluzione della (5.1) generando la successione di vettori

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{J}(\mathbf{x}_k)^{-1} \mathbf{f}(\mathbf{x}_k) \quad (5.6)$$

dove $\mathbf{J}(\mathbf{x}_k)$ è la matrice *giacobiana*

$$\mathbf{J}(\mathbf{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_2}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_1} \\ \frac{\partial f_1}{\partial x_2} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_2} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_1}{\partial x_n} & \frac{\partial f_2}{\partial x_n} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix} \quad (5.7)$$

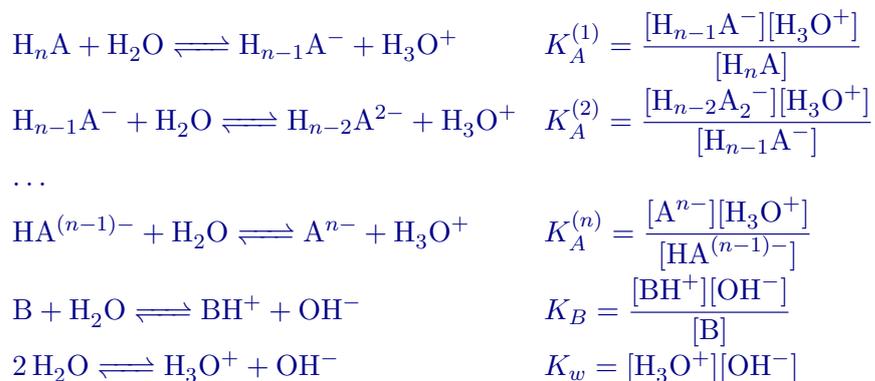
calcolata nel punto \mathbf{x}_k . Nei metodi quasi-Newton in pratica si sostituisce la matrice \mathbf{J} con qualche opportuna approssimazione, per evitare il calcolo delle derivate prime. Nelle routine di calcolo più usate (per esempio la *hybrid* del pacchetto *Minpack*) si impiegano metodi quasi-Newton con valutazioni basate su approcci alle differenze finite, vedi il Capitolo (7), per la valutazione della matrice giacobiana.

Applicazione 6. Consideriamo come applicazione il calcolo del pH di una soluzione di un acido debole poliprotico, a cui sia aggiunta una soluzione di una base debole.

Supponiamo perciò, come nell'esempio precedente, di avere una soluzione acquosa contenente un acido debole monoprotico H_nA , di molarità m_A per un volume V_A ; aggiungiamo un volume V_B di una soluzione

acquosa di una base debole B, di molarità m_B . Calcoliamo il pH della soluzione risultante, note le $k_A^{(i)}$, $i = 1, \dots, n$ dell'acido e la k_B della base.

Supponiamo di essere in soluzioni ideali diluite e di poter quindi descrivere gli equilibri di reazione in termini di molarità (coefficienti di attività unitari e volume di mescolamento nullo). Scriviamo gli equilibri di reazione



Ancora una volta devono valere i bilanci di massa rispetto alle specie A e B,

$$\begin{aligned} (V_A + V_B) \sum_{i=0}^n [\text{H}_{n-i}\text{A}^{i-}] &= V_A m_A \\ (V_A + V_B)([\text{B}] + [\text{BH}^+]) &= V_B m_B \end{aligned}$$

e la soluzione deve essere elettricamente neutra $[\text{H}_3\text{O}^+] + [\text{BH}^+] = [\text{OH}^-] + \sum_{i=1}^n i[\text{H}_{n-i}\text{A}^{i-}]$; abbiamo quindi un sistema di $n + 5$ equazioni in $n + 5$ incognite, che si può (!) risolvere ottenendo $[\text{H}_3\text{O}^+]$. *Suggerimento: impostate e risolvetevi le equazioni relative al caso dell'acido fosforico H_3PO_4 e dell'ammoniaca NH_3 .*

Capitolo 6

Equazioni differenziali

6.1 Introduzione

Un'equazione differenziale è un'equazione la cui soluzione è costituita da una funzione incognita in una o più variabili indipendenti, in cui compaiono anche le derivate (o i differenziali) della funzione incognita. Una generica equazione differenziale nella funzione incognita $y(x_1, \dots, x_n)$ che dipende dalle variabili indipendenti x_1, \dots, x_n ha quindi la forma

$$F \left(x_1, x_2, \dots, x_n, y, \dots, \frac{\partial^{(a_1+\dots+a_n)} y}{\partial x_1^{a_1} \dots \partial x_n^{a_n}}, \dots \right) = 0 \quad (6.1)$$

Il massimo ordine delle derivate che compaiono in F si dice *ordine* dell'equazione differenziale.

- Un sistema di equazioni differenziali è un insieme di equazioni differenziali con più di una funzione incognita.

$$F_1 \left(x_1, x_2, \dots, x_n, y_1, \dots, y_m, \dots, \frac{\partial^{(a_{11}+\dots+a_{1n})} y_1}{\partial x_1^{a_{11}} \dots \partial x_n^{a_{1n}}}, \dots, \frac{\partial^{(a_{m1}+\dots+a_{mn})} y_m}{\partial x_1^{a_{m1}} \dots \partial x_n^{a_{mn}}}, \dots \right) = 0$$

... (6.2)

$$F_m \left(x_1, x_2, \dots, x_n, y_1, \dots, y_m, \dots, \frac{\partial^{(a_{11}+\dots+a_{1n})} y_1}{\partial x_1^{a_{11}} \dots \partial x_n^{a_{1n}}}, \dots, \frac{\partial^{(a_{m1}+\dots+a_{mn})} y_m}{\partial x_1^{a_{m1}} \dots \partial x_n^{a_{mn}}}, \dots \right) = 0$$

Per esempio, le equazioni che descrivono la cinetica di un meccanismo di reazione costituiscono un sistema di equazioni differenziali che definiscono le concentrazioni delle varie specie chimiche in funzione del tempo (ed eventualmente di coordinate spaziali, vedi il Capitolo successivo).

- Un'equazione differenziale ordinaria è un'equazione differenziale in cui la variabile indipendente è una sola; se è a l'ordine

$$F \left(x, y, \frac{dy}{dx}, \dots, \frac{d^a y}{dx^a} \right) = 0 \quad (6.3)$$

Per esempio l'equazione di Schrödinger che descrive un autostato $\psi_n(x)$ di un oscillatore armonico monodimensionale è un'equazione differenziale ordinaria di secondo ordine nella variabile spaziale x .

- un'equazione differenziale alle derivate parziali è un'equazione differenziale in cui le variabili indipendenti sono almeno due. Per esempio l'equazione di Schrödinger dipendente dal tempo che descrive un stato generico $\psi(x, t)$ di un oscillatore armonico monodimensionale è un'equazione differenziale alle derivate parziali di secondo ordine nelle variabili spazio x , e tempo t .

Il concetto di "soluzione" di un'equazione differenziale deve essere chiarito. In generale un'equazione differenziale o un sistema di equazioni differenziali ammette infinite soluzioni, cioè funzioni y_1, \dots, y_m che soddisfano le condizioni espresse dalle F_1, \dots, F_m . L'univocità della soluzione è ottenuta solo specificando delle condizioni aggiuntive dette *condizioni al contorno* e/o *condizioni iniziali*. Riprendiamo l'argomento nella sezione (6.2), esplicitamente riservata alle più comuni classi di equazioni differenziali alle derivate parziali. Proseguiamo invece ora con l'esposizione di alcune proprietà di interesse delle equazioni differenziali ordinarie.

6.1.1 Equazioni differenziali ordinarie

Consideriamo prima di tutto le equazioni differenziali ordinarie; in forma implicita un'equazione differenziale ordinaria di ordine a è data da (6.3); se è possibile ricavare la derivata a -esima avremo la forma esplicita

$$\frac{d^a y}{dx^a} = f\left(x, y, \frac{dy}{dx}, \dots, \frac{d^{a-1}y}{dx^{a-1}}\right) \quad (6.4)$$

Una soluzione univoca si ottiene solo se sono date a condizioni iniziali:

$$y(x_0) = y_0, \left. \frac{dy}{dx} \right|_{x=x_0} = y'_0, \dots, \left. \frac{d^a y}{dx^a} \right|_{x=x_0} = y_0^{(a)} \quad (6.5)$$

oppure condizioni al contorno, che stabiliscono il comportamento della funzione incognita in più di un punto x_1, x_2, \dots ; per sempio, per un'equazione del secondo ordine $y(x_1) = y_1$ e $y(x_2) = y_2$. Sono anche definibili condizioni miste che definiscono il comportamento di funzioni della funzione incognita e di alcune sue derivate in uno o più valori della variabile indipendente.

Esempio 32. La concentrazione di una specie chimica $[A]$ in un reattore chiuso, che sia descritta da un cinetica di ordine r

$$\frac{d[A]}{dt} = -k[A]^r \quad [A]_{t=0} = [A]_0 \quad (6.6)$$

è un esempio di equazione differenziale del primo ordine¹ con una condizione iniziale

La termodinamica può fornirci vari esempi di equazioni differenziali. Supponiamo per esempio di avere una soluzione binaria a temperatura e pressione costante di un soluto (liquido se puro alla temperatura e pressione di lavoro) in acqua. Il potenziale chimico del soluto può esser ottenuto da quello dell'acqua tramite la relazione di Gibbs-Duhem. Se x è la frazione molare del soluto, e $\mu_{\text{H}_2\text{O}}(x)$ è una funzione nota che descrive il potenziale chimico dell'acqua in soluzione in funzione della frazione molare del soluto, abbiamo che

$$\frac{d\mu_{\text{soluto}}}{dx} = -\frac{1-x}{x} \frac{d\mu_{\text{H}_2\text{O}}}{dx} \quad \mu_{\text{soluto}}(1) = \mu_{\text{soluto}}^* \quad (6.7)$$

La condizione 'iniziale' è che il potenziale chimico ricercato è pari al potenziale chimico del solvente puro.

L'equazione di Schrödinger che definisce gli autostati di una particella che si muova in una scatola monodimensionale di lunghezza L (vedi il Capitolo successivo)

$$-\frac{\hbar^2}{2m} \frac{d^2 \Psi(x)}{dx^2} = E \Psi(x) \quad \Psi(0) = \Psi(L) = 0 \quad (6.8)$$

è un esempio di equazione differenziale ordinaria al secondo ordine con condizioni al contorno facilmente definite (zero probabilità di trovare la particella fuori dalla scatola).

Le equazioni differenziali ordinarie $F(x, y, y') = 0$ del primo ordine sono incontrate molto spesso nelle applicazioni più comuni, e vale la pena di analizzare brevemente alcune loro forme particolari. Ammettendo che valga la condizione iniziale $y(x_0) = y_0$

¹Attenzione: l'ordine dell'equazione differenziale è 1, perchè la derivata rispetto al tempo, variabile indipendente, è 1; l'ordine della cinetica è definito dalla potenza di $[A]$

Integrazione semplice $\frac{dy}{dx} = f(x)$ si risolve integrando $f(x)$:

$$y(x) = y(x_0) + \int_{x_0}^x f(u)du \quad (6.9)$$

Variabili separabili $M(x)dx + N(y)dy = 0$ si risolve ancora per integrazione semplice:

$$\int_{x_0}^x M(u)du + \int_{y_0}^y N(v)dv = 0 \quad (6.10)$$

Equazione lineare $y' + a(x)y + b(x) = 0$ si risolve mediante la definizione di una nuova funzione z tale che $y = ze^{-\int_{x_0}^x a(u)du}$ con $z(x_0) = y_0$; sostituendo nell'equazione lineare, si ottiene un'equazione in z risolvibile per integrazione semplice; in fine si ottiene una soluzione esplicita in $y(x)$

$$\begin{aligned} y' &= z'e^{-\int_{x_0}^x a(u)du} - za e^{-\int_{x_0}^x a(u)du} \\ z'e^{-\int_{x_0}^x a(u)du} + b &= 0 \Rightarrow z = y_0 - \int_{x_0}^x b(v)e^{\int_{x_0}^v a(u)du} dv \\ y &= e^{-\int_{x_0}^x a(u)du} \left[y_0 - \int_{x_0}^x b(v)e^{\int_{x_0}^v a(u)du} dv \right] \end{aligned} \quad (6.11)$$

Le equazioni differenziali di ordine superiore al primo sono caratterizzate da un numero di condizioni iniziali pari all'ordine dell'equazione stessa come abbiamo visto poco sopra. Il caso più semplice, $\frac{d^n y}{dx^n} = f(x)$ si risolve integrando ripetutamente $f(x)$:

$$y = \sum_{k=0}^{n-1} \frac{(x-x_0)^k}{k!} y_0^{(k)} + \int_{x_0}^x \int_{x_0}^{u_1} \dots \int_{x_0}^{u_{n-1}} f(u_{n-1}) \prod_{k=1}^{n-1} du_k \quad (6.12)$$

L'integrale multiplo però si può scrivere, grazie ad una proprietà dello sviluppo in serie di una funzione², come un integrale semplice, ottenendo la soluzione generale

$$y = \sum_{k=0}^{n-1} \frac{(x-x_0)^k}{k!} y_0^{(k)} + \frac{1}{(n-1)!} \int_{x_0}^x (x-u)^{n-1} f(u)du \quad (6.13)$$

ma in generale i metodi di soluzione analitici semplici sono rari e comunque riconducibili alla soluzione di sistemi di equazioni differenziali (oppure, nel caso di quasi tutte le applicazioni pratiche, possono essere sostituiti da *metodi numerici di soluzione*, vedi il Capitolo che segue). Le equazioni lineari sono la classe di equazioni differenziali ordinarie più studiata e meglio caratterizzata, riassumiamo quindi alcuni concetti fondamentali nella sezione successiva.

6.1.2 Equazioni lineari

Un'equazione differenziale ordinaria lineare si definisce come

$$y^{(n)} + a_1 y^{(n-1)} + \dots + a_{n-1} y' + a_n y = f \quad (6.14)$$

f, a_1, \dots, a_n sono funzioni in x , se $f = 0$ l'equazione si dice omogenea. Se a_1, \dots, a_n sono costanti abbiamo un'equazione lineare a coefficienti costanti. Poniamoci il consueto problema di voler risolvere l'equazione differenziale (6.14) con le condizioni iniziali $y(x_0) = y_0, \frac{dy}{dx} \Big|_{x=x_0} = y'_0 \equiv y_0^{(1)}, \dots, \frac{d^n y}{dx^n} \Big|_{x=x_0} = y_0^{(n)}$.

Si noti che il problema della soluzione dell'equazione (6.14) è equivalente alla soluzione di un sistema di

²Espansione di Taylor con resto sotto forma di integrale

n equazioni differenziali del primo ordine. Definiamo infatti $z_1 = y$, $z_2 = y'$, $z_3 = y''$ e così via fino a $z_n = y^{(n-1)}$; otteniamo

$$\begin{aligned} z_1' &= z_2 \\ z_2' &= z_3 \\ &\dots \\ z_{n-1}' &= z_n \\ z_n' &= -a_1 z_{n-1} - a_2 z_{n-2} - \dots - a_n z_1 + f \end{aligned} \tag{6.15}$$

Quindi, definendo i vettori $\mathbf{z} = (z_1, \dots, z_n)^T$, $\mathbf{f} = (0, \dots, f)^T$ e la matrice

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ -a_n & -a_{n-1} & -a_{n-2} & \dots & -a_1 & 0 \end{pmatrix}$$

otteniamo

$$\frac{d\mathbf{z}}{dx} = \mathbf{A}\mathbf{z} + \mathbf{f} \quad \mathbf{z}(x_0) = \mathbf{z}_0 \tag{6.16}$$

dove il vettore che definisce le condizioni iniziali, \mathbf{z}_0 è formato dai valori iniziali delle derivate di $y(x)$ in x_0 . Quindi la soluzione di una equazione differenziale ordinaria lineare di ordine n si può sempre ricondurre alla soluzione di un sistema lineare di n equazioni differenziali ordinarie lineari del primo ordine (il contrario è anche, generalmente, vero). Si può notare infine che l'equivalenza equazione differenziale ordinaria di ordine n / sistema di n equazioni differenziali ordinarie del primo ordine non è limitata alle sole equazioni differenziali ordinarie *lineari*. Se possiamo definire la generica equazione differenziale di ordine n in forma esplicita

$$y^{(n)} = F(x, y, y', \dots, y^{(n-1)}) \tag{6.17}$$

possiamo comunque definire $z_1 = y$, $z_2 = y'$, $z_3 = y''$ e così via fino a $z_n = y^{(n-1)}$, ottenendo

$$\begin{aligned} z_1' &= z_2 \\ z_2' &= z_3 \\ &\dots \\ z_{n-1}' &= z_n \\ z_n' &= F(x, z_1, z_2, \dots, z_{n-1}) \end{aligned} \tag{6.18}$$

In generale, per generiche funzioni a_1, \dots, a_n non è possibile trovare la soluzione dell'equazione (6.14) mediante operazioni semplici di integrazione, cambio di variabili etc. Valgono però alcune interessanti proprietà generali che rendono le equazioni lineari molto più 'facili' da trattare rispetto ad altre equazioni differenziali ordinarie.

Equazione omogenea Se $f = 0$ e sono note n soluzioni *linearmente indipendenti* y_i , con $i = 1, \dots, n$, cioè tali che non esistono coefficienti costanti α_i non tutti nulli tali che $\sum_{i=1}^n \alpha_i y_i = 0$, allora una soluzione generica si ottiene da una qualunque combinazione lineare delle n soluzioni

$$y = \sum_{i=1}^n c_i y_i \tag{6.19}$$

dove le costanti c_i si possono determinare sulla base delle condizioni iniziali, risolvendo un sistema lineare. La condizione di indipendenza lineare di n funzioni si può trovare calcolando il seguente determinante, detto *wronskiano*

$$\Delta(x) = \begin{pmatrix} y_1 & y_2 & \dots & y_n \\ y_1' & y_2' & \dots & y_n' \\ \dots & \dots & \dots & \dots \\ y_1^{(n-1)} & y_2^{(n-1)} & \dots & y_n^{(n-1)} \end{pmatrix} \equiv \Delta(x_0) e^{-\int_{x_0}^x a_1(u) du} \tag{6.20}$$

le n soluzioni sono linearmente indipendenti se e solo se il wronskiano non è identicamente nullo.

Equazione non omogenea Quando $f \neq 0$, se \bar{y} è una soluzione particolare dell'equazione (6.14) e y_i , con $i = 1, \dots, n$ sono n soluzioni *linearmente indipendenti* della corrispondente equazione omogenea, allora la soluzione generale si ottiene come

$$y = \bar{y} + \sum_{i=1}^n c_i y_i \quad (6.21)$$

Si noti che, se conosciamo le y_i è possibile trovare una soluzione particolare dell'equazione non omogenea mediante la seguente procedura

1. risolvere il sistema in n funzioni incognite \bar{y}'_i

$$\begin{aligned} \bar{y}'_1 y_1 + \dots + \bar{y}'_n y_n &= 0 \\ \bar{y}'_1 y'_1 + \dots + \bar{y}'_n y'_n &= 0 \\ &\dots \\ \bar{y}'_1 y_1^{(n-1)} + \dots + \bar{y}'_n y_n^{(n-1)} &= f \end{aligned} \quad (6.22)$$

2. calcolare per integrazione, note le derivate \bar{y}'_i , le funzioni \bar{y}_i
3. costruire la soluzione particolare \bar{y} come $\bar{y} = \sum_{i=1}^n \bar{y}_i y_i$

Oppure si può ricorrere ad un metodo di calcolo numerico ...

6.1.3 Sistemi lineari del primo ordine a coefficienti costanti

Consideriamo, alla fine di questa sezione, il caso, molto importante per un chimico, di un sistema di equazioni differenziali lineari del primo ordine a coefficienti costanti, che è come abbiamo già visto anche equivalente ad una equazione differenziale lineare a coefficienti costanti, di ordine pari alla dimensione del sistema. Per cambiare un po' la simbologia finora utilizzata, ed anche per ricordare al lettore che nella maggior parte dei casi incontrerà questo tipo di problema nel contesto di dipendenza dal tempo, scriviamo il sistema come

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{f}(t) \quad \mathbf{x}(t_0) = \mathbf{x}_0 \quad (6.23)$$

dove \mathbf{A} è una matrice a coefficienti costanti $n \times n$; si noti anche la scrittura alternativa per indicare la derivata, molto usata in problemi in cui la variabile indipendente è il tempo. I concetti base di algebra lineare esposti nel Capitolo (2) ci aiutano a trovare la soluzione generica della 6.23 in modo relativamente semplice.

Analizziamo dapprima il caso omogeneo, $\mathbf{f} = 0$. Possiamo scrivere subito la soluzione formale del problema come

$$\mathbf{x}(t) = e^{\mathbf{A}(t-t_0)} \mathbf{x}_0 \quad (6.24)$$

ricordiamo che la scrittura $e^{\mathbf{A}t}$ indica una matrice $n \times n$ tale che

$$e^{\mathbf{A}t} = \mathbf{1} + \mathbf{A}t + \frac{1}{2} \mathbf{A}^2 t^2 + \dots = \sum_{k=0}^{\infty} \frac{1}{k!} \mathbf{A}^k t^k \quad (6.25)$$

La dimostrazione che il vettore definito dalla (6.24) è effettivamente una soluzione procede semplicemente per sostituzione nella (6.23)

$$\frac{d}{dt} \left[\sum_{k=0}^{\infty} \frac{1}{k!} \mathbf{A}^k (t-t_0)^k \right] \mathbf{x}_0 = \left[\sum_{k=1}^{\infty} \frac{k}{k!} \mathbf{A}^k (t-t_0)^{k-1} \right] \mathbf{x}_0 = \mathbf{A} \left[\sum_{k=0}^{\infty} \frac{1}{k!} \mathbf{A}^k (t-t_0)^k \right] \mathbf{x}_0 = \mathbf{A}\mathbf{x} \quad (6.26)$$

Se invece $f \neq 0$, si può dimostrare facilmente che la soluzione generica è

$$\mathbf{x}(t) = e^{\mathbf{A}t} \left[\int_{t_0}^t e^{-\mathbf{A}\tau} \mathbf{f}(\tau) d\tau + e^{-\mathbf{A}t_0} \mathbf{x}_0 \right] \quad (6.27)$$

Naturalmente, la soluzione formale del problema, per quanto molto compatta, è di poca utilità nelle applicazioni. Ma come si ricorda dal Capitolo (2), se sono noti gli autovalori e gli autovettori della matrice \mathbf{A} , le cose cambiano. Supponiamo di avere risolto il problema agli autovalori $\mathbf{A}\mathbf{V} = \mathbf{V}\mathbf{\Lambda}$, dove $\mathbf{\Lambda}$ è la matrice diagonale degli autovalori di \mathbf{A} e \mathbf{V} la matrice degli autovettori, che supponiamo invertibile. Poiché per una funzione generica di matrice $f(\mathbf{A})$ possiamo scrivere $f(\mathbf{A}) = \mathbf{V}f(\mathbf{\Lambda})\mathbf{V}^{-1}$, otteniamo che la soluzione generica della (6.23) per $f = 0$ è semplicemente

$$\mathbf{x}(t) = \mathbf{V}e^{\mathbf{\Lambda}t}\mathbf{V}^{-1}\mathbf{x}_0 \quad (6.28)$$

mentre per il caso $f \neq 0$ otteniamo

$$\mathbf{x}(t) = \mathbf{V}e^{\mathbf{\Lambda}t} \left[\int_{t_0}^t e^{-\mathbf{\Lambda}\tau}\mathbf{V}^{-1}\mathbf{f}(\tau)d\tau + e^{-\mathbf{\Lambda}t_0}\mathbf{V}^{-1}\mathbf{x}_0 \right] \quad (6.29)$$

e il calcolo della matrice $e^{\mathbf{\Lambda}t}$ è triviale (è la matrice diagonale formata dai valori $e^{\lambda_1 t}, \dots, e^{\lambda_n t}$).

Approfondimento 12. Soprattutto nel contesto delle spettroscopie, capita spesso però anche di incontrare il problema della soluzione di un sistema di equazioni differenziali lineari del primo ordine con *coefficienti dipendenti dal tempo*. Il concetto di soluzione formale, opportunamente combinato con approcci di tipo perturbativo, può essere esteso alla soluzione almeno parziale di questo tipo di problemi, ed è stato studiato tra gli altri da Freeman Dyson e da Wilhelm Magnus negli anni '50 del XX secolo. Consideriamo per semplicità il sistema omogeneo

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) \quad \mathbf{x}(t_0) = \mathbf{x}_0 \quad (6.30)$$

La soluzione formale, che generalizza la (6.24), può essere derivata intuitivamente come segue; integriamo la (6.30), ottenendo

$$\mathbf{x}(t) = \mathbf{x}_0 + \int_{t_0}^t dt_1 \mathbf{A}(t_1)\mathbf{x}(t_1) \quad (6.31)$$

ora sostituiamo la precedente equazione in se stessa ottenendo

$$\mathbf{x}(t) = \mathbf{x}_0 + \int_{t_0}^t dt_1 \mathbf{A}(t_1) \left[\mathbf{x}_0 + \int_{t_0}^{t_1} dt_2 \mathbf{A}(t_2)\mathbf{x}(t_2) \right] = \left[\mathbf{1} + \int_{t_0}^t dt_1 \mathbf{A}(t_1) \right] \mathbf{x}_0 + \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \mathbf{A}(t_1)\mathbf{A}(t_2)\mathbf{x}(t_2) \quad (6.32)$$

possiamo ora sostituire la (6.31) nella (6.32) e continuare così si verifica che la seguente espansione in serie (detta di Dyson) è la soluzione formale della (6.30)

$$\begin{aligned} \mathbf{x}(t) = & \left[\mathbf{1} + \int_{t_0}^t dt_1 \mathbf{A}(t_1) + \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \mathbf{A}(t_1)\mathbf{A}(t_2)\mathbf{x}(t_2) + \right. \\ & \left. + \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \int_{t_0}^{t_2} dt_3 \mathbf{A}(t_1)\mathbf{A}(t_2)\mathbf{A}(t_3) + \dots \right] \mathbf{x}_0 \end{aligned} \quad (6.33)$$

si noti che $t_1 > t_2 > t_3 > \dots$. È utile definire l'*operatore di ordinamento temporale* \mathcal{T} , tale che

$$\mathcal{T}f(t)g(t') = \begin{cases} f(t)g(t') & t > t' \\ g(t')f(t) & t < t' \end{cases} \quad (6.34)$$

Si può dimostrare, con un po' di pazienza, che la soluzione (6.33) si scrive

$$\begin{aligned} \mathbf{x}(t) = & \left[\sum_{k=0}^{\infty} \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \dots \int_{t_0}^{t_{k-1}} dt_k \mathbf{A}(t_1)\mathbf{A}(t_2) \dots \mathbf{A}(t_k) \right] \mathbf{x}_0 = \\ = & \left[\sum_{k=0}^{\infty} \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \dots \int_{t_0}^{t_{k-1}} dt_k \mathcal{T}\mathbf{A}(t_1)\mathbf{A}(t_2) \dots \mathbf{A}(t_k) \right] \mathbf{x}_0 = \end{aligned} \quad (6.35)$$

$$\begin{aligned} = & \left[\sum_{k=0}^{\infty} \frac{1}{k!} \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \dots \int_{t_0}^{t_{k-1}} dt_k \mathcal{T}\mathbf{A}(t_1)\mathbf{A}(t_2) \dots \mathbf{A}(t_k) \right] \mathbf{x}_0 \\ = & \mathcal{T}e^{\int_{t_0}^t \mathbf{A}(\tau)d\tau} \mathbf{x}_0 = \mathbf{U}(t, t_0)\mathbf{x}_0 \end{aligned} \quad (6.36)$$

La matrice

$$\mathbf{U}(t, t_0) = \mathcal{T} e^{\int_{t_0}^t \mathbf{A}(\tau) d\tau} \quad (6.37)$$

è calcolabile quindi come una somma infinita di integrali multipli, e si riduce alla matrice esponenziale $e^{\int_{t_0}^t \mathbf{A}(\tau) d\tau}$ nel caso in cui $[\mathbf{A}(t), \mathbf{A}(t')] = 0$, cioè nel caso $\mathbf{A}(t)$ commuti con $\mathbf{A}(t')$ per una coppia di valori t, t' generici (come succede quando per esempio consideriamo il caso a coefficienti costanti).

Magnus propose per primo una scrittura alternativa, basata su alcune proprietà generali di funzioni di matrici, in cui $\mathbf{U}(t, t_0)$ è scritta come una funzione esponenziale di un'altra matrice $\mathbf{C}(t, t_0)$ a sua volta esprimibile come una serie

$$\mathbf{U}(t, t_0) = e^{\mathbf{C}(t, t_0)} \quad (6.38)$$

$$\mathbf{C}(t) = \sum_{k=1}^{\infty} \mathbf{C}_k(t) \quad (6.39)$$

I termini della serie sono di complessità crescente, e si scrivono in funzione di commutatori annidati della matrice originaria $\mathbf{A}(t)$:

$$\begin{aligned} \mathbf{C}_1(t) &= \int_0^t dt_1 \mathbf{A}(t_1) \\ \mathbf{C}_2(t) &= \frac{1}{2} \int_0^t dt_1 \int_0^{t_1} dt_2 [\mathbf{A}(t_1), \mathbf{A}(t_2)] \\ \mathbf{C}_3(t) &= \frac{1}{6} \int_0^t dt_1 \int_0^{t_1} dt_2 \int_0^{t_2} dt_3 ([\mathbf{A}(t_1), [\mathbf{A}(t_2), \mathbf{A}(t_3)]] + [\mathbf{A}(t_3), [\mathbf{A}(t_2), \mathbf{A}(t_1)]]) \\ \mathbf{C}_4(t) &= \frac{1}{12} \int_0^t dt_1 \int_0^{t_1} dt_2 \int_0^{t_2} dt_3 \int_0^{t_3} dt_4 ([\mathbf{A}(t_1), [[\mathbf{A}(t_2), \mathbf{A}(t_3)], \mathbf{A}(t_4)]] + \\ &+ [\mathbf{A}(t_1), [\mathbf{A}(t_2), [\mathbf{A}(t_3), \mathbf{A}(t_4)]]] + [\mathbf{A}(t_2), [\mathbf{A}(t_3), [\mathbf{A}(t_4), \mathbf{A}(t_1)]]]) \\ &\dots \end{aligned} \quad (6.40)$$

è possibile definire una procedura ricorsiva piuttosto complicata che consente di calcolare il k -esimo termine dell'espansione di Magnus dati i termini precedenti.

6.2 Equazioni alle derivate parziali

Passiamo ora ad una disamina, molto superficiale, di alcune classi di equazioni differenziali alle derivate parziali, un settore estremamente vasto e profondo dell'analisi matematica. Ci limiteremo nel seguito a discutere le *equazioni differenziali lineari alle derivate parziali di secondo ordine*, al cui interno si possono collocare molti degli strumenti formali usati per descrivere vari fenomeni della termodinamica chimica, cinetica e chimica quantistica. Siano $x_i, i = 1, \dots, n$ un insieme di coordinate indipendenti (cartesiane). La forma generale delle equazioni differenziali lineari alle derivate parziali di secondo ordine è la seguente:

$$\sum_{i,j=1}^n a_{ij}(\mathbf{x}) \frac{\partial^2 y}{\partial x_i \partial x_j} + \sum_{i=1}^n b_i(\mathbf{x}) \frac{\partial y}{\partial x_i} + c(\mathbf{x}) y = f(x) \quad (6.41)$$

È utile a questo punto ricordare la definizione di un operatore molto utile che semplifica la notazione in presenza di derivate parziali. Definiamo (nelle coordinate cartesiane x_i) l'operatore vettore *gradiente*:

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \dots \\ \frac{\partial}{\partial x_n} \end{pmatrix} \quad (6.42)$$

Il gradiente di una funzione ∇f è quindi il vettore formato dalle derivate prime di f . Si noti che la *divergenza* di un vettore $\mathbf{v}(\mathbf{x})$ di funzioni in \mathbf{x} , è il prodotto scalare del vettore gradiente con \mathbf{v} :

$$\nabla^T \mathbf{v} = \sum_{i=1}^n \frac{\partial v_i}{\partial x_i} \quad (6.43)$$

La (6.2) può essere quindi scritta, con una ridefinizione delle funzioni $b_i \leftarrow b_i - \sum_{j=1}^n \frac{\partial a_{ij}}{\partial x_j}$, nella forma compatta

$$\sum_{i,j=1}^n \frac{\partial}{\partial x_i} a_{ij}(\mathbf{x}) \frac{\partial y}{\partial x_i} + \sum_{i=1}^n b_i(\mathbf{x}) \frac{\partial y}{\partial x_i} + c(\mathbf{x})y = (\nabla^T \mathbf{A} \nabla + \mathbf{b}^T \nabla + c)y = f \quad (6.44)$$

dove \mathbf{A} è la matrice di funzioni a_{ij} , \mathbf{b} è il vettore di funzioni b_i (modificate).

L'esempio di maggiore importanza per il chimico è senz'altro l'equazione di Schrödinger, che descrive la funzione d'onda di un sistema rispetto alle coordinate \mathbf{x} ; si noti che l'equazione di Schrödinger è naturalmente definita per funzioni a valori complessi; in unità 'naturali' (per esempio SI) non scalate rispetto a grandezze atomiche:

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2} \sum_{i=1}^n \frac{1}{m_i} \frac{\partial^2 \Psi}{\partial x_i^2} + V(\mathbf{x})\Psi = \mathcal{H}\Psi \quad (6.45)$$

dove definendo $\mathbf{x} = (t = x_0, x_1, \dots, x_n)$ ritroviamo la forma canonica (6.41) ponendo $a_{0i} = 0$, $a_{ij} = -\hbar^2/2m_i$, $b_0 = -i\hbar$, $b_i = 0$, $c = V$, $f = 0$ con $i, j = 1, \dots, n$. Possiamo considerare anche l'equazione di diffusione, che descrive per esempio la concentrazione di un soluto, avente un tensore cartesiano di diffusione D_{ij} in una soluzione, in tre dimensioni, cioè con $i, j = 1, 2, 3$; la forma più conveniente è la (6.44)

$$\frac{\partial C}{\partial t} = \frac{\partial}{\partial x_i} D_{ij} \frac{\partial C}{\partial x_j} + S \quad (6.46)$$

S è una funzione che descrive i flussi in entrata e in uscita del soluto; definendo $\mathbf{x} = (t = x_0, x_1, x_2, x_3)$ ritroviamo la forma canonica (6.44) ponendo $a_{0i} = 0$, $a_{ij} = D_{ij}$, $b_0 = -1$, $b_i = 0$, $c = 0$, $f = S$ con $i, j = 1, 2, 3$. Si noti che in generale il tensore di diffusione e la funzione sorgente S possono dipendere dalla posizione e dal tempo. Nel caso sia presente un campo di velocità (flusso) \mathbf{v} dovuto a moti convettivi, l'equazione di diffusione diventa un'equazione di convezione-diffusione, con l'aggiunta del termine di trasporto dovuto al flusso

$$\frac{\partial C}{\partial t} = \sum_{i,j=1}^3 \frac{\partial}{\partial x_i} D_{ij} \frac{\partial C}{\partial x_j} - \sum_{i=1}^3 \frac{\partial v_i C}{\partial x_i} + S = \sum_{i,j=1}^3 \frac{\partial}{\partial x_i} D_{ij} \frac{\partial C}{\partial x_j} - \sum_{i=1}^3 v_i \frac{\partial C}{\partial x_i} - \left(\sum_{i=1}^3 \frac{\partial v_i}{\partial x_i} \right) C + S \quad (6.47)$$

dove la forma canonica (6.44) si ottiene, con $\mathbf{x} = (t = x_0, x_1, x_2, x_3)$, ponendo $a_{0i} = 0$, $a_{ij} = D_{ij}$, $b_0 = -1$, $b_i = -v_i$, $c = -\sum_{i=1}^3 \frac{\partial v_i}{\partial x_i}$, $f = S$ con $i, j = 1, 2, 3$. Se la soluzione è un fluido incompressibile $c = 0$, perchè la divergenza del vettore velocità di flusso di un fluido incompressibile è nulla.

Approfondimento 13 . Di interesse forse meno immediato per il chimico, ricordiamo l'equazione che descrive una dinamica vibrazionale (di una corda, membrana, volume); in tre dimensioni

$$\rho \frac{\partial^2 u}{\partial t^2} = \frac{\partial}{\partial x_i} p_{ij} \frac{\partial u}{\partial x_j} - qu + f \quad (6.48)$$

la vibrazione $u(\mathbf{x}, t)$ dipende dal tensore $p_{ij}(\mathbf{x}, t)$ e dalla funzione $q(\mathbf{x}, t)$ che dipendono dal mezzo oscillante, e dalla funzione $f(\mathbf{x}, t)$ che rappresenta una perturbazione esterna.

Le equazioni differenziali lineari alle derivate parziali di secondo ordine sono classificabili in un punto dato \mathbf{x}_0 sulla base delle proprietà della matrice \mathbf{A} che compare nella forma canonica (6.41); in particolare si parla di equazioni

ellittiche se gli autovalori di \mathbf{A} in \mathbf{x}_0 sono tutti positivi o tutti negativi

paraboliche se gli autovalori di \mathbf{A} in \mathbf{x}_0 sono tutti positivi o tutti negativi, meno uno che è nullo

iperboliche se gli autovalori di \mathbf{A} in \mathbf{x}_0 sono tutti positivi, meno uno che è negativo, o se sono tutti negativi, meno uno che è positivo

ultraiperboliche in tutti gli altri casi

L'equazione di diffusione e l'equazione di Schrödinger sono entrambe paraboliche (nel caso dell'equazione di diffusione, il tensore di diffusione è definito positivo, cioè ha solo autovalori maggiori di zero). Si noti però che l'equazione di Schrödinger indipendente dal tempo, per trovare gli autostati di un sistema quantistico

$$-\frac{\hbar^2}{2} \sum_{i=1}^n \frac{1}{m_i} \frac{\partial^2 \Psi}{\partial x_i^2} + V(\mathbf{x})\Psi = \mathcal{H}\Psi = E\Psi \quad (6.49)$$

è una equazione ellittica. Naturalmente la classificazione può cambiare da punto a punto. Per esempio l'equazione di Tricomi (usata nello studio dei fenomeni di flusso supersonico) $x_2 \frac{\partial^2 y}{\partial x_1^2} + \frac{\partial^2 y}{\partial x_2^2} = 0$ è iperbolica per $x_2 < 0$, parabolica per $x_2 = 0$ ed ellittica per $x_2 > 0$.

6.2.1 Condizioni al contorno

Per concludere questo riassunto minimo delle proprietà delle equazioni differenziali lineari alle derivate parziali di secondo ordine, discutiamo brevemente le condizioni aggiuntive che devono essere imposte per assicurare l'univocità della soluzione. Ancora una volta, non si pretende in questa esposizione di avere alcuna pretesa di completezza, quanto piuttosto di fornire alcune linee guida che possano costituire una base di partenza per lo studente interessato ad approfondire la materia.

In generale, dato un processo chimico o fisico descritto da un'equazione differenziale alle derivate parziali nelle n variabili (reali) indipendenti \mathbf{x} , si definirà una regione dello spazio $D \subset \mathcal{R}^n$ detta *dominio di definizione* dell'equazione. Al dominio D possiamo associare una frontiera F . Per la definizione di una soluzione univocamente definita è necessario stabilire delle condizioni aggiuntive in D che chiameremo nel seguito semplicemente *condizioni al contorno*. Le condizioni al contorno possono essere di varia natura, e nel caso delle equazioni differenziali lineari alle derivate parziali di secondo ordine si distinguono tre classi principali

Condizioni di Cauchy per le equazioni iperboliche e paraboliche; il problema di Cauchy si enuncia imponendo che la soluzione $y(\mathbf{x})$ verifichi le condizioni:

$$y|_{\mathbf{x} \in F} = \bar{y}_0(\mathbf{x}) \quad \left. \frac{\partial y}{\partial \mathbf{n}} \right|_{\mathbf{x} \in F} = \mathbf{n}^T \nabla y|_{\mathbf{x} \in F} = \bar{y}_1(\mathbf{x}) \quad (6.50)$$

si impone cioè la forma della soluzione e della sua derivata normale su F

Condizioni di Dirichlet per le equazioni ellittiche; il problema di Dirichlet si enuncia imponendo che la soluzione $y(\mathbf{x})$ verifichi la condizione:

$$y|_{\mathbf{x} \in F} = \bar{y}(\mathbf{x}) \quad (6.51)$$

si impone cioè la forma della soluzione su F

Condizioni di Neumann per le equazioni ellittiche; il problema di Neumann si enuncia imponendo che la soluzione $y(\mathbf{x})$ verifichi la condizione:

$$\left. \frac{\partial y}{\partial \mathbf{n}} \right|_{\mathbf{x} \in F} = \bar{y}(\mathbf{x}) \quad (6.52)$$

si impone cioè la forma della derivata normale su F

Condizioni miste per le equazioni ellittiche; il problema misto si enuncia imponendo che la soluzione $y(\mathbf{x})$ verifichi la condizione:

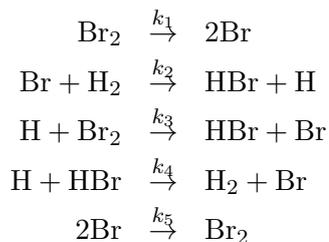
$$\left(\alpha y + \beta \frac{\partial y}{\partial \mathbf{n}} \right) \Big|_{\mathbf{x} \in F} = \bar{y}(\mathbf{x}) \quad (6.53)$$

dove α e β sono funzioni note.

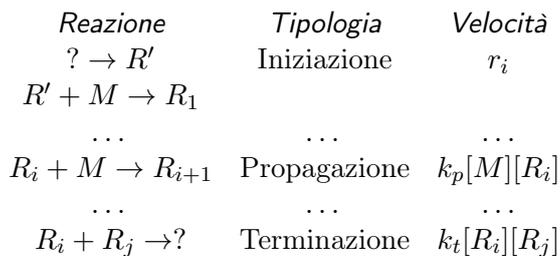
Sono possibili altri tipi di condizioni al contorno, ma sono meno comuni. In ogni caso è fondamentale verificare che le condizioni al contorno siano corrette, cioè che possano assicurare l'esistenza univoca di una soluzione continua nel dominio di definizione, e dipendente con continuità dai dati di frontiera.

6.3 Esercizi

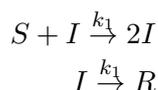
1. Un reattore contiene ^{222}Rn ; sapendo che il radon decade a ^{218}Po , che a sua volta decade a ^{218}Pb , che produce ^{214}Bi , che si trasforma in ^{214}Po , che genera ^{210}Pb , che decade a ^{210}Bi , che si trasforma in ^{210}Po , che finalmente forma l'isotopo stabile ^{206}Pb , scrivete il sistema di equazioni differenziali che descrive la catena di decadimento del ^{222}Rn
2. Qual è il sistema di equazioni differenziali che descrive questo meccanismo di reazione?



3. E questo?



4. Il modello di Kermack e Mckendrick (1927) descrive la diffusione di una malattia all'interno di una popolazione con un numero di individui fissato, $S_0 = S(t) + I(t) + R(t)$, dove $S(t)$ è il numero di individui sani, $I(t)$ il numero di individui malati e $R(t)$ il numero di individui 'rimossi' al tempo t .



Scrivete il sistema di equazioni differenziali che descrive il processo.

5. Un oscillatore armonico quantistico si muove lungo una retta; la massa della particella è m , la costante di forza dell'oscillatore è k . Si assume inoltre che la particella abbia una carica q e che sia presente un campo elettrico esterno oscillante $E = E_0 \cos(\Omega t)$. Scrivete l'equazione di Schrödinger dipendente dal tempo che descrive il sistema nella forma canonica (6.41)

6. Il potenziale elettrico $V(\vec{r})$, dove \vec{r} è un punto generico dello spazio, prodotto da una distribuzione di carica avente densità $\rho(\vec{r})$, obbedisce all'equazione di Poisson, che scriviamo

$$\nabla^T \epsilon \nabla V = -\rho$$

dove ϵ è il tensore, dipendente da \vec{r} che descrive la permittività dielettrica del sistema. Scrivete esplicitamente l'equazione di Poisson rispetto alle tre componenti di \vec{r} e classificatela secondo lo schema descritto nel testo.

Capitolo 7

Equazioni differenziali: applicazioni

7.1 Esempi: termodinamica e cinetica chimica

Dedichiamo questo capitolo ad illustrare alcuni esempi di applicazioni dei concetti generali sviluppati nel capitolo (6). In questa sezione ci occupiamo in particolare della definizione e della soluzione con metodi analitici di alcune equazioni differenziali per la descrizione di problemi tratti dalla termodinamica e cinetica chimica. la sezione successiva è dedicata a vari esempi di sistemi quantistici semplici, descritti da equazioni differenziali risolvibili analiticamente. L'ultima sezione è infine dedicata ad una discussione di alcune metodologie di soluzione approssimata e ad alcuni cenni di metodi numerici per la soluzione di equazioni differenziali, con uno o due esempi applicati a problemi generici di interesse per il chimico.

7.1.1 Termodinamica

Consideriamo una soluzione liquida contenente vari composti volatili, in equilibrio con il suo vapore. Se sono presenti n componenti, la regola delle fasi ci dice che il sistema è complessivamente n variante, cioè sono individuabili n variabili indipendenti che descrivono il sistema stesso. Per un sistema multicomponente, in generale, vale l'equazione di Gibbs-Duhem, che applichiamo alla soluzione stessa (fase condensata)

$$SdT - Vdp + \sum_{i=1}^n n_i d\mu_i = 0$$

qui S e V sono l'entropia ed il volume della soluzione; T e p sono la temperatura e la pressione del sistema. Se fissiamo la temperatura, ci rimane un'equazione che lega la pressione totale e i potenziali chimici dei componenti in soluzione

$$\sum_{i=1}^n n_i d\mu_i = Vdp$$

Il potenziale chimico i -esimo in soluzione è uguale al potenziale chimico i -esimo in fase vapore, ed assumendo che la fase vapore sia una miscela gassosa perfetta, si ha che $\mu_i = \mu_i^o(T) + RT \ln p_i \Rightarrow d\mu_i = RT d \ln p_i$ a T costante, quindi

$$RT \sum_{i=1}^n n_i d \ln p_i = Vdp \Rightarrow \sum_{i=1}^n x_i d \ln p_i = \frac{v}{RT} dp = \frac{v}{v_g} d \ln p$$

dove abbiamo diviso per il numero di moli n_{totali} in soluzione, ottenendo al secondo membro v , volume molare del liquido, e tenendo conto che $RT = pv_g$, dove v_g è il volume molare della fase gassosa perfetta. La relazione appena ricavata ha la forma di una equazione differenziale che lega le variazioni delle pressioni parziali alla variazione della pressione totale. Possiamo di solito approssimare il secondo membro a zero, dato $v/v_g \ll 1$:

$$\sum_{i=1}^n x_i d \ln p_i \approx 0$$

Avendo fissato la temperatura, possiamo scegliere $n-1$ frazioni molari in soluzione come variabili indipendenti del sistema, $d \ln p_i = \sum_{j=1}^{n-1} \frac{\partial \ln p_i}{\partial x_j} dx_j$ e perciò

$$\sum_{i=1}^n x_i \sum_{j=1}^{n-1} \frac{\partial \ln p_i}{\partial x_j} dx_j = \sum_{j=1}^{n-1} \left(\sum_{i=1}^n x_i \frac{\partial \ln p_i}{\partial x_j} \right) dx_j = 0$$

La relazione precedente deve essere sempre vera. Restano quindi $n-1$ equazioni differenziali alle derivate parziali nelle n incognite p_i , una per ogni valore di j

$$\sum_{i=1}^n x_i \frac{\partial \ln p_i}{\partial x_j} = 0 \quad j = 1, n \dots, n-1$$

Per un sistema binario, per esempio, abbiamo un'unica relazione (equazione di Duhem-Margules)

$$x_1 \frac{d \ln p_1}{dx_1} + x_2 \frac{d \ln p_2}{dx_1} = \frac{x_1 dp_1}{p_1 dx_1} + \frac{(1-x_1) dp_2}{p_2 dx_1} = 0 \quad (7.1)$$

Se per esempio $p_1 = p_1^* x_1$ (legge di Raoult):

$$\frac{dp_1}{dx_1} = p_1^* \Rightarrow 1 - \frac{x_2 dp_2}{p_2 dx_2} = 0 \Rightarrow \frac{dp_2}{dx_2} = \frac{p_2}{x_2}$$

dove si è fatto uso del fatto che $x_1 + x_2 = 1$; la semplice equazione differenziale $\frac{dp_2}{dx_2} = \frac{p_2}{x_2}$ è risolta imponendo come $p_2 = K x_2$, dove K è una costante indeterminata (legge di Henry).

Applicazione 7. Consideriamo la seguente applicazione. Vogliamo determinare la variazione dell'entalpia di una transizione di fase di una sostanza pura, presente nell'equazione di Clausius-Clapeyron

$$\frac{dp}{dT} = \frac{\Delta H}{T \Delta V}$$

dove $\Delta H = H_\alpha - H_\beta$ ($\Delta H = H_\alpha - H_\beta$) è la differenza di entalpia (volume) tra le fasi α e β . In generale sappiamo che, descrivendo il sistema in funzione di p e T (che NON sono indipendenti, data la presenza delle due fasi all'equilibrio)

$$\begin{aligned} d\Delta H &= \frac{\partial \Delta H}{\partial T} dT + \frac{\partial \Delta H}{\partial p} dp = \Delta C_p dT + \frac{\partial \Delta H}{\partial p} dp = \\ &= \Delta C_p dT + \left(\Delta V - T \frac{\partial \Delta V}{\partial T} \right) dp \end{aligned}$$

nella seconda espressione abbiamo usato la relazione $\frac{\partial H}{\partial p} = -T \frac{\partial V}{\partial T} + V$, nota dal corso di termodinamica¹.

Usando l'equazione di Clausius-Clapeyron otteniamo

$$\begin{aligned} \frac{d\Delta H}{dT} &= \Delta C_p + \left(\Delta V - T \frac{\partial \Delta V}{\partial T} \right) \frac{\Delta H}{T \Delta V} = \\ &= \Delta C_p + \frac{\Delta H}{T} - \Delta H \frac{d \ln V}{dT} \end{aligned} \quad (7.2)$$

l'equazione (7.2) è dovuta a Planck (1927). Chiamiamo $\Delta C_p = f(T)$, $\frac{d \ln V}{dT} = g(T)$; la formula (7.2) si scrive

$$\frac{d\Delta H}{dT} - \left[g(T) - \frac{1}{T} \right] \Delta H - f(T) = 0$$

¹Si ottiene dal differenziale dell'entalpia nelle sue variabili canoniche e da una relazione di Maxwell ...

che ha la forma di un'equazione differenziale lineare ordinaria del primo ordine omogenea, nell'incognita $\Delta H(T)$, risolvibile analiticamente, per ogni coppia di funzioni $f(T)$, $g(T)$.

7.1.2 Cinetica

la cinetica chimica dei sistemi omogenei è un settore della chimica dove le equazioni differenziali ordinarie trovano molte applicazioni immediate. Richiamiamo alcuni esempi semplici ed alcuni un po' più complicati.

Reazioni semplici

Una reazione chimica al *primo ordine* rispetto ad una specie chimica è caratterizzata da ordine pari ad uno. Consideriamo una reazione *monomolecolare* generica, che indichiamo con



in cui la molecola di reagente R genera i prodotti P_1 , P_2 etc. La reazione è del primo ordine. La legge cinetica è perciò

$$v = -\frac{d[R]}{dt} = \frac{1}{p_1} \frac{d[P_1]}{dt} = \frac{1}{p_2} \frac{d[P_2]}{dt} \dots = k[R] \quad (7.4)$$

Consideriamo una reazione generica del tipo $R \rightarrow$ prodotti che segue una legge al primo ordine $v = k[R]$. Siano per semplicità $x(t) = c_0 - [R]$ le moli per decimetro cubo di reagente che si sono consumate al tempo t , con la condizione iniziale che al tempo $t = 0$ la concentrazione del reagente valga $c_0 = [R]_{t=0}$. La legge cinetica può essere scritta come un'equazione differenziale lineare di primo ordine ai coefficienti costanti:

$$\frac{dx}{dt} = k(c_0 - x) \quad (7.5)$$

$$x(0) = 0 \quad (7.6)$$

L'equazione (7.5) si risolve separando le variabili x e t ed integrando

$$\begin{aligned} \frac{dx}{c_0 - x} = k dt &\Rightarrow \int_0^x (c_0 - u) du = k \int_0^t d\tau \\ &\Rightarrow -\ln\left(\frac{c_0}{c_0 - x}\right) = kt \Rightarrow x = c_0(1 - e^{-kt}) \Rightarrow [R] = [R]_0 e^{-kt} \end{aligned} \quad (7.7)$$

La dipendenza dal tempo, in forma integrata, della concentrazione di una specie che segue una legge cinetica del primo ordine, è esponenziale.

In una reazione cinetica caratterizzata da una legge al *secondo ordine* la somma degli esponenti è due. Consideriamo una reazione *bimolecolare* generica, che indichiamo con



in cui le molecole di reagente R_1 e R_2 generano i prodotti P_1 , P_2 etc. La reazione è del secondo ordine. La legge cinetica è perciò

$$v = -\frac{d[R_1]}{dt} = -\frac{d[R_2]}{dt} = \frac{1}{p_1} \frac{d[P_1]}{dt} = \frac{2}{p_2} \frac{d[P_2]}{dt} \dots = k[R_1][R_2] \quad (7.9)$$

R_1 ed R_2 possono anche ovviamente essere due molecole della stessa specie



in cui due molecole di reagente R generano i prodotti P_1 , P_2 etc. La legge cinetica è sempre del secondo ordine

$$v = -\frac{1}{2} \frac{d[R]}{dt} = \frac{1}{p_1} \frac{d[P_1]}{dt} = \frac{1}{p_2} \frac{d[P_2]}{dt} \dots = k[R]^2 \quad (7.11)$$

Proviamo a trovare le leggi cinetiche integrate di due reazioni chimiche che seguano rispettivamente la legge (7.9) e (7.10). Nel primo caso, $R_1 + R_2 \rightarrow$ prodotti, sia $x(t)$ le moli per decimetro cubico di reagente 1 che hanno reagito, pari evidentemente alle moli di reagente 2 che hanno reagito; come condizioni iniziali abbiamo $c_{10} = [R_1]_0$ e $c_{20} = [R_2]_0$ per i due reagenti. La legge cinetica è

$$\frac{dx}{dt} = k(c_{10} - x)(c_{20} - x) \quad (7.12)$$

$$x(0) = 0 \quad (7.13)$$

Separando le variabili x e t si ottiene la legge integrata

$$\frac{1}{c_{10} - c_{20}} = \ln \frac{c_{20}(c_{10} - x)}{c_{10}(c_{20} - x)} = kt \quad (7.14)$$

Nel secondo caso, un po' più semplice, siano $x(t) = c_0 - [R]$ le moli per decimetro cubo di reagente che si sono consumate al tempo t , con la condizione iniziale che al tempo $t = 0$ la concentrazione del reagente valga $c_0 = [R]_{t=0}$. La legge cinetica è (ponendo $k' = 2k$ per semplicità)

$$\frac{dx}{dt} = k(c_0 - x)^2 \quad (7.15)$$

$$x(0) = 0 \quad (7.16)$$

che risolta permette di ottenere l'andamento di x

$$\frac{x}{c_0(c_0 - x)} = kt \Rightarrow \frac{1}{[R]} - \frac{1}{[R]_0} = kt \quad (7.17)$$

Reazioni di *ordine superiore a due* sono piuttosto rare, e sono praticamente sempre il risultato di un meccanismo di reazione complesso, piuttosto che di una singola reazione elementare che comporti il contemporaneo 'incontro' reattivo di tre o più corpi. È piuttosto complicato, e di utilità minima, considerare esplicitamente tutti i casi possibili, che comunque sono risolvibili con i metodi semplici di analisi matematica accennati in precedenza. Consideriamo come esempio un'ipotetica reazione elementare del tipo $2R_1 + R_2 \rightarrow$ prodotti; siano $[R_1]_0 = c_{10}$ e $[R_2]_0 = c_{20}$ le concentrazioni iniziali dei reagenti; siano x le moli per decimetro cubo di reagente 2 che si sono consumate al tempo t (e quindi $2x$ le moli per decimetro cubo che si sono consumate di reagente 1); la legge cinetica in forma differenziale è

$$\frac{dx}{dt} = k(c_{10} - 2x)^2(c_{20}) \quad (7.18)$$

con la solita condizione $x(0) = 0$. Integrando, otteniamo

$$\frac{1}{(2c_{20} - c_{10})^2} \left[\frac{2x(2c_{20} - c_{10})}{c_{10}(c_{10} - 2x)} + \ln \frac{c_{20}(c_{10} - 2x)}{c_{10}(c_{20} - x)} \right] = kt \quad (7.19)$$

Per una reazione generica $R \rightarrow$ prodotti che segua la legge

$$\frac{d[R]}{dt} = -k[R]^n \quad (7.20)$$

con $[R]_0$ concentrazione iniziale ed $n \leq 2$, la legge cinetica integrata è semplicemente

$$\frac{1}{[R]^{n-1}} - \frac{1}{[R]_0^{n-1}} = (n-1)kt \quad (7.21)$$

cui corrisponde un tempo di dimezzamento pari a

$$t_{1/2} = \frac{1}{n-1} (2^{n-1} - 1) \frac{1}{k[R]_0^{n-1}} \quad (7.22)$$

Meccanismi di reazione

L'analisi della cinetica di una reazione chimica non elementare è uno studio complesso, che però fornisce molte informazioni indirette sul *meccanismo* della reazione stessa, (cioè sull'insieme di reazioni elementari che sono responsabili della reazione effettivamente osservata). Discutiamo brevemente alcuni esempi.

Il primo caso che intendiamo analizzare è quello di un semplice *equilibrio chimico*. Consideriamo due reazioni bimolecolari, una opposta all'altra, che coinvolgano per semplicità due reagenti e due prodotti



l'espressione precedente deve essere letta come l'affermazione che avvengano contemporaneamente sia la reazione diretta che la reazione inversa



con due costanti di velocità k_d e k_i . La velocità di formazione/scomparsa delle varie specie è calcolata, in un meccanismo di reazione, come la *somma delle velocità ottenute da ciascuna reazione elementare*. Così per esempio, le velocità relative alla concentrazione di R_1 ed R_2 sono

$$\frac{d[R_1]}{dt} = \frac{d[R_2]}{dt} = -k_d[R_1][R_2] + k_i[P_1][P_2] \quad (7.26)$$

la velocità relativa alla concentrazione di P_1 e P_2 sono uguali ed opposte etc. In condizioni di equilibrio le concentrazioni delle varie specie devono restare costanti, quindi le loro velocità di formazione/scomparsa devono essere pari a zero. ne consegue che

$$\frac{d[R_1]}{dt} = \frac{d[R_2]}{dt} = -k_d[R_1][R_2] + k_i[P_1][P_2] = 0 \Rightarrow \frac{[P_1][P_2]}{[R_1][R_2]} = \frac{k_d}{k_i} \quad (7.27)$$

Se consideriamo la reazione chimica (7.23) come un processo stechiometrico indipendente in una soluzione ideale in condizioni di equilibrio termodinamico, vale la relazione²

$$K = \frac{[P_1][P_2]}{[R_1][R_2]} \quad (7.28)$$

da cui consegue l'identità

$$K = \frac{k_d}{k_i} \quad (7.29)$$

che correla direttamente le costanti di velocità e la costante di equilibrio.

Uno schema di reazione molto importante è dato dalle *reazioni consecutive*. Discutiamo nel seguito in dettaglio il caso di tre specie chimiche R_1 , R_2 ed R_3 legate da due processi chimici elementari monomolecolari



Siano x_1 , x_2 ed x_3 le concentrazioni al tempo t delle tre sostanze. Valgono le tre equazioni differenziali

$$\begin{aligned} \frac{dx_1}{dt} &= -k_a x_1 \\ \frac{dx_2}{dt} &= k_a x_1 - k_b x_2 \\ \frac{dx_3}{dt} &= k_b x_2 \end{aligned} \quad (7.31)$$

²Esprimendo le molalità o le frazioni molari, impiegate per descrivere solitamente la soluzione, in termini di molarità

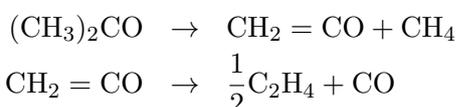
Si può notare che

$$\frac{dx_1}{dt} + \frac{dx_2}{dt} + \frac{dx_3}{dt} = 0 \quad (7.32)$$

che non è altro che l'affermazione che non si possono creare nuove molecole, ma solo, in base al meccanismo dato, trasformare 1 in 2 e 2 in 3. Come condizione iniziale, assumiamo che al tempo $t = 0$ sia presente solo la specie 1, con una concentrazione c_0 . La soluzione di questo insieme di equazioni lineari a coefficienti costanti è relativamente semplice. Si ottiene

$$\begin{aligned} x_1 &= e^{-k_a t} \\ x_2 &= \frac{k_a c_0}{k_b - k_a} \left[e^{(k_b - k_a)t} - 1 \right] e^{-k_b t} \\ x_3 &= c_0 \left(1 - \frac{k_b e^{-k_a t}}{k_b - k_a} - \frac{k_a e^{-k_b t}}{k_b - k_a} \right) \end{aligned} \quad (7.33)$$

Con un po' di attenzione si può notare come gli andamenti di x_1 , x_2 e x_3 corrispondano ad una scomparsa esponenziale di 1, alla formazione e successiva scomparsa dell'*intermedio* 2, ed alla formazione di 3. Tra le reazioni che seguono il meccanismo appena illustrato troviamo per esempio la pirolisi dell'acetone



La concentrazione del chetene intermedio aumenta fino ad un massimo e poi diminuisce fino a scompa-

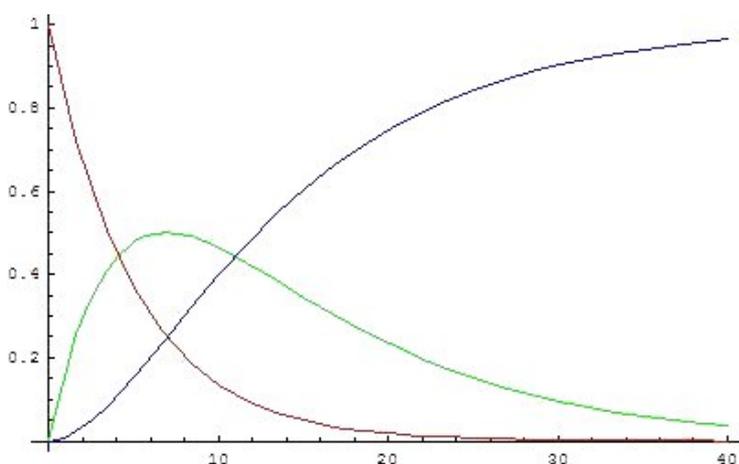
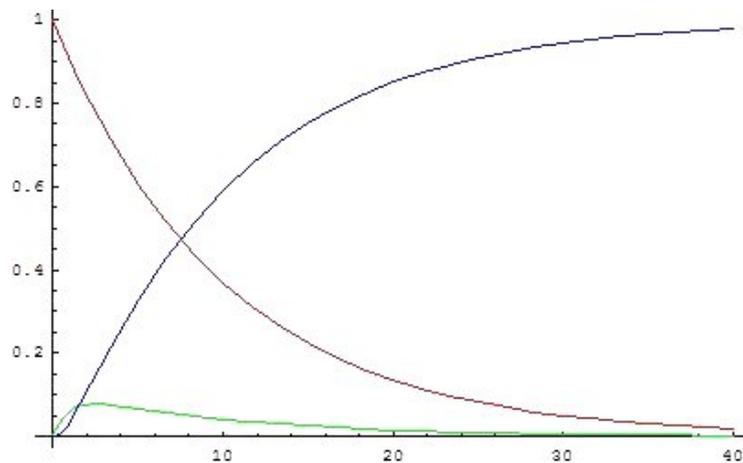


Figura 7.1: Reazioni consecutive: $k_b/k_a = 0.5$.

rire. In generale la concentrazione dell'intermedio è tanto più piccola quanto maggiore è il rapporto tra k_b (che misura la rapidità con cui l'intermedio scompare) e k_a (che misura la rapidità con cui l'intermedio si forma). Nelle due Figure è riportata la concentrazione delle tre specie per $c_0 = 1$ e $k_b/k_a = 0.5$ e 10, rispettivamente: nel secondo caso la concentrazione dell'intermedio è molto bassa. In questo caso la reazione può essere descritta in buona approssimazione da una reazione del primo ordine con costante di velocità pari al valore del passaggio più lento (k_a): la formazione dell'intermedio agisce da 'collo di bottiglia' della reazione complessiva e controlla la cinetica.

Consideriamo infine il caso di una sostanza che possa partecipare a varie reazioni chimiche elementari, generando diversi prodotti; esempi numerosi sono dati per esempio da reazioni di decomposizione di composti organici che generano più prodotti contemporaneamente. Per semplicità supponiamo di avere due prodotti ottenuti dallo stesso reagente mediante due reazioni monomolecolari:



Figura 7.2: Reazioni consecutive: $k_b/k_a = 10$.

Se inizialmente la concentrazione del reagente è c_0 , le concentrazioni dei prodotti, che indichiamo con x_1 e x_2 , seguono la legge

$$\frac{dx_1}{dt} = k_1(c_0 - x_1 - x_2) \quad (7.35)$$

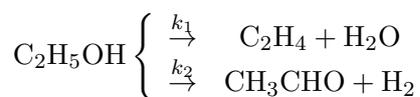
$$\frac{dx_2}{dt} = k_2(c_0 - x_1 - x_2) \quad (7.36)$$

e sotto la condizione che le concentrazioni iniziali dei due prodotti siano nulle otteniamo:

$$x_1 = \frac{k_1 c_0}{k_1 + k_2} \left[1 - e^{-(k_1 + k_2)t} \right] \quad (7.37)$$

$$x_2 = \frac{k_2 c_0}{k_1 + k_2} \left[1 - e^{-(k_1 + k_2)t} \right] \quad (7.38)$$

da cui possiamo osservare che $x_1/x_2 = k_1/k_2$: il rapporto tra le concentrazioni resta costante e pari al rapporto tra le costanti di velocità (ma nel caso di reazioni elementari di ordine superiore la relazione non è così semplice). Un esempio semplice è dato dalla disidratazione di alcol etilico ad etilene ed aldeide acetica



La composizione della miscela dei prodotti è determinata dal rapporto tra le costanti di velocità delle reazioni parallele, che competono fra loro: l'uso di opportune condizioni (temperatura, impiego di catalizzatori etc.) permette di aumentare la concentrazione dell'uno o dell'altro prodotto.

Analizziamo infine uno dei meccanismi di reazione più noti: il modello o **meccanismo di Michaelis-Menten**, utilizzato per l'interpretazione delle cinetiche di catalisi enzimatiche. Il modello spiega come inizialmente, all'aumentare anche di poco della concentrazione del substrato disponibile all'enzima, la velocità della reazione aumenti fino al raggiungimento di un massimo. In queste condizioni il substrato satura completamente l'enzima presente, e quindi un'ulteriore aggiunta di substrato non cambia apprezzabilmente la velocità. Indichiamo l'enzima libero con E il substrato S ed il complesso enzima-substrato con ES ; il complesso si scompone liberando l'enzima e generando il prodotto P . Lo schema MM è dunque dato da una coppia di reazioni opposte (*non* un pre-equilibrio) seguito da una reazione monomolecolare:



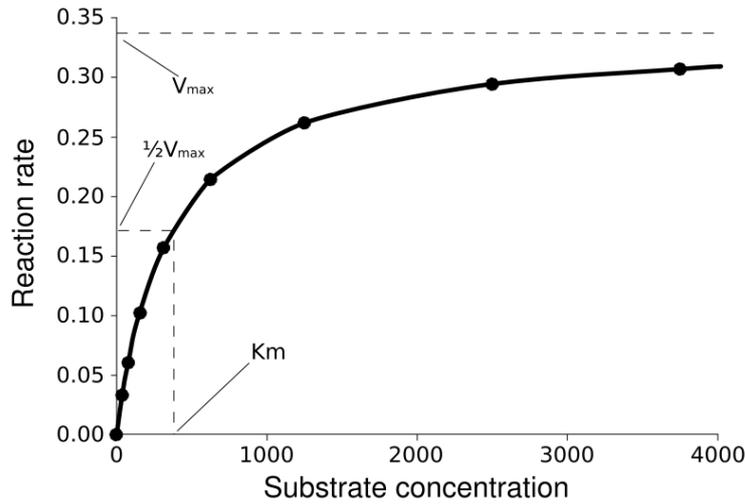


Figura 7.3: Velocità di reazione di una catalisi enzimatica.

Scriviamo il sistema di equazioni differenziali del sistema (7.39):

$$\begin{aligned}\frac{d[E]}{dt} &= -k_1[E][S] + (k_{-1} + k_2)[ES] \\ \frac{d[ES]}{dt} &= k_1[E][S] - (k_{-1} + k_2)[ES] \\ \frac{d[S]}{dt} &= -k_1[E][S] + k_1[ES] \\ \frac{d[P]}{dt} &= k_2[ES]\end{aligned}\tag{7.40}$$

valgono inoltre i bilanci di massa $[ES] + [E] = [E]_0$ e $[ES] + [S] + [P] = [S]_0$, dove $[E]_0$ e $[S]_0$ sono le concentrazioni iniziali di enzima e substrato; assumiamo che le concentrazioni iniziali di $[ES]$ e $[P]$ siano nulle. Solitamente si adottano le seguenti approssimazioni: i) la concentrazione di substrato è molto maggiore di tutte le altre specie, quindi si può considerare invariata nel corso della reazione $[S] \approx [S]_0$ e ii) la specie $[ES]$ è un intermedio labile, per il quale è applicabile l'ipotesi dello *stato stazionario*, secondo la quale la sua velocità di formazione/scomparsa è nulla (a parte pochi istanti iniziali); otteniamo perciò:

$$\frac{d[ES]}{dt} = k_1[E][S] - k_{-1}[ES] - k_2[ES] \approx 0\tag{7.41}$$

quindi la concentrazione di complesso enzima-substrato è data dall'espressione:

$$[ES] = \frac{k_1[E][S]}{k_{-1} + k_2} = \frac{[E][S]}{K_m}\tag{7.42}$$

dove si definisce la costante di Michaelis

$$K_m = \frac{k_{-1} + k_2}{k_1}\tag{7.43}$$

La velocità della reazione (formazione del prodotto) è ottenuta come

$$v = \frac{d[P]}{dt} = k_2[ES] = \frac{k_2[E][S]}{K_m}\tag{7.44}$$

Sperimentalmente è nota la sola concentrazione iniziale dell'enzima $[E]_0$, ed è conveniente esprimere la velocità in funzione di questo parametro. Dato che vale il bilancio di massa $[E] + [ES] = [E]_0$, si trova che:

$$[ES] = \frac{[E]_0[S]}{K_m + [S]}\tag{7.45}$$

e quindi si ottiene l'equazione di Michaelis-Menten

$$v = \frac{k_2[E]_0[S]}{K_m + [S]} \quad (7.46)$$

La velocità massima di reazione si raggiunge per $[S] \gg K_m$:

$$V_{max} = k_2[E]_0 \quad (7.47)$$

Applicazione 8 . Discutiamo la soluzione esatta del meccanismo di Michaelis-Menten, che almeno formalmente può essere ridotto ad una singola equazione *integro-differenziale non lineare*, la cui soluzione però è possibile solo con metodi numerici. Ricaviamo dai due bilanci di massa $[E] = [E]_0 - [ES]$ ed $[S] = [S]_0 - [ES] - [P]$, inoltre dall'ultima delle (7.39) troviamo $[P]$ in funzione di $[ES]$; chiamiamo $x(t) = [ES]$ per semplicità. Otteniamo $[P] = k_2 \int_0^t x d\tau$ e quindi

$$\frac{dx}{dt} = k_1([E]_0 - x) \left([S]_0 - x - k_2 \int_0^t x d\tau \right) - (k_{-1} + k_2)x$$

con la condizione iniziale $x(0) = 0$.

7.2 Esempi: sistemi quantistici risolvibili analiticamente

Discutiamo alcuni sistemi quantomeccanici risolvibili analiticamente. Tra i vari casi trattati sono compresi gli esempi che si trovano in tutti i libri di testo: particella nella scatola monodimensionale e bidimensionale, rotatore monodimensionale, oscillatore armonico, particella su una sfera, atomo idrogenoide. In più sono presentate "variazioni sul tema", come particelle in scatole monodimensionali di forme svariate, scatole tridimensionali, scatole cilindriche e set di oscillatori interagenti.

Consideriamo lo schema in Fig. (7.4), che rappresenta il potenziale confinante per una particella (in una dimensione) imprigionata in una scatola dalle pareti infinite. Qual è l'andamento nel tempo della funzione d'onda della particella, data una determinata *configurazione* iniziale? Vedremo che per rispondere a questo quesito sarà necessario definire le autofunzioni e gli autovalori dell'operatore hamiltoniano. Anzi, poichè il problema è di natura generale, consideriamo per cominciare un sistema generico definito nello spazio delle funzioni $f(\mathbf{x})$ dove \mathbf{x} è l'insieme delle coordinate del sistema. Al tempo t il sistema è definito dalla funzione d'onda $|\psi(\mathbf{x}, t)\rangle$ la cui evoluzione temporale è dettata dall'equazione di Schrödinger

$$i\hbar \frac{\partial}{\partial t} |\psi(\mathbf{x}, t)\rangle = \mathcal{H} |\psi(\mathbf{x}, t)\rangle \quad (7.48)$$

al tempo $t = 0$ la funzione è definita da una configurazione generica (che ci dice come il sistema è stato *preparato* inizialmente):

$$|\psi(\mathbf{x}, 0)\rangle = |\psi_0(\mathbf{x})\rangle \quad (7.49)$$

Sappiamo che l'hamiltoniano ammette autostati $|E\rangle$ con autoenergie E :

$$\mathcal{H}|E\rangle = E|E\rangle \quad (7.50)$$

che sono ortonormali $\langle E|E'\rangle = \delta_{E,E'}$ e costituiscono una base completa dello spazio. Quindi possiamo espandere $|\psi(\mathbf{x}, t)\rangle$

$$|\psi(\mathbf{x}, t)\rangle = \sum_E c_E(t) |E\rangle \quad (7.51)$$

i coefficienti $C_E(t)$ sono ottenute come proiezioni della funzione d'onda sulle autofunzioni

$$c_E(t) = \langle \psi(\mathbf{x}, t) | E \rangle \quad (7.52)$$

e al tempo $t = 0$ vale che

$$c_E(0) = \langle \psi_0(\mathbf{x}) | E \rangle \quad (7.53)$$

Sostituendo nell'Eq. (7.48) si ottiene, moltiplicando per un generico $\langle E |$ ed integrando

$$i\hbar \dot{c}_E(t) = E c_E(t) \quad (7.54)$$

Infine, con le condizioni iniziali (7.53) si trova

$$c_E(t) = \langle \psi_0(\mathbf{x}) | E \rangle \exp(Et/i\hbar) \quad (7.55)$$

e la funzione d'onda si può scrivere al tempo t come:

$$|\psi(\mathbf{x}, t)\rangle = \sum_E |E\rangle \langle \psi_0(\mathbf{x}) | E \rangle \exp(Et/i\hbar) \quad (7.56)$$

Dunque la funzione d'onda è una sovrapposizione degli autostati del sistema con 'pesi' che dipendono dal tempo in modo esponenziale, dove le 'costanti di decadimento' di ciascun esponenziale sono proporzionali agli autovalori, e dalle condizioni iniziali.

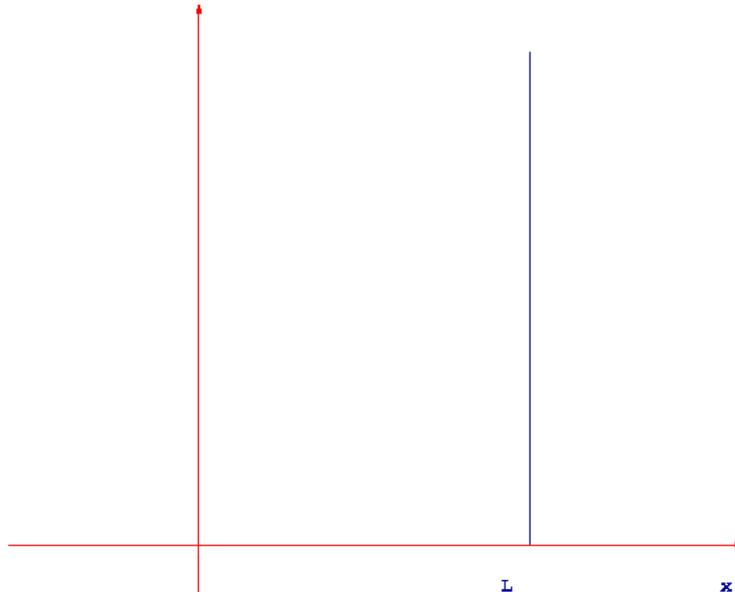


Figura 7.4: Potenziale per una particella in una scatola 1D.

In linea di principio possiamo distinguere lo spettro di autoenergie di un sistema in due parti: lo spettro *discreto*, ovvero l'insieme delle autoenergie separate da intervalli non accessibili, e lo spettro *continuo*. Consideriamo un sistema monodimensionale con un potenziale $U(x)$ che tenda a U_0 per $x \rightarrow +\infty$ e a $U_1 > U_0$ per $x \rightarrow -\infty$, vedi Fig. (7.5). Lo spettro discreto si ha per autostati con energie $E < U_0$: la particella è chiusa nella scatola. Se $E > U_0$ lo spettro diviene continuo, e la particella è libera. Il numero di autostati discreti può essere finito (se U_0 e U_1 sono quantità finite) od infinito. Nel seguito considereremo solo stati con autovalori discreti, cioè sistemi in stati *confinati*.

Torniamo ora al caso della particella in una scatola monodimensionale con pareti infinite. Assumiamo che l'energia potenziale sia nulla all'interno della scatola ed infinita all'esterno. La funzione d'onda è dunque definita solo per $0 \leq x \leq L$, vedi Fig. (7.4). L'equazione di Schrödinger indipendente dal tempo assume la forma³:

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi_E(x) = E \psi_E(x) \quad (7.57)$$

³Ci poniamo nella rappresentazione continua, in modo da lavorare direttamente con funzioni della variabile x

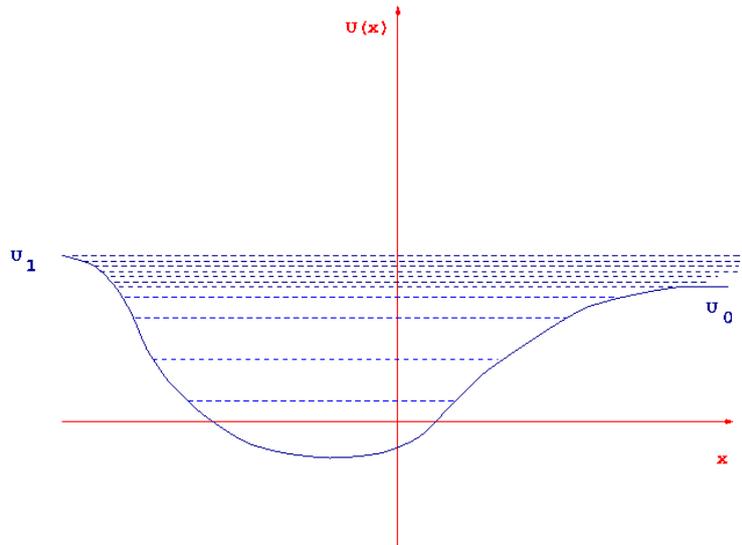


Figura 7.5: Un potenziale generico in una dimensione.

dove m è la massa della particella. Valgono le condizioni al contorno

$$\psi_E(0) = 0, \quad \psi_E(L) = 0 \quad (7.58)$$

L'Eq. (7.57) è un'equazione differenziale ordinaria a coefficienti costanti del secondo ordine. La sua soluzione generica ha la forma

$$\psi_E(x, t) = A_1 \exp(\alpha_1 x) + A_2 \exp(\alpha_2 x) \quad (7.59)$$

dove $\alpha_{1,2}$ sono le radici dell'equazione

$$-\frac{\hbar^2}{2m} \alpha^2 = E \quad (7.60)$$

vale a dire $\alpha_{1,2} = \pm i \sqrt{\frac{2mE}{\hbar^2}}$. Sostituendo nelle condizioni al contorno si ottiene

$$A_1 + A_2 = 0 \quad (7.61)$$

$$A_1 \exp\left(iL \sqrt{\frac{2mE}{\hbar^2}}\right) + A_2 \exp\left(-iL \sqrt{\frac{2mE}{\hbar^2}}\right) = 0 \quad (7.62)$$

che è un sistema lineare omogeneo in A_1 e A_2 , risolvibile solo se il determinante dei coefficienti è nullo. Si trova così la condizione in E :

$$\exp\left(2iL \sqrt{\frac{2mE}{\hbar^2}}\right) = 1 \quad (7.63)$$

le cui soluzioni sono

$$E = \frac{\hbar^2 \pi^2 k^2}{2mL^2}, \quad k \in \mathcal{N} \quad (7.64)$$

Gli autostati di una particella chiusa in un pozzo infinitamente profondo sono quantizzati come i quadrati dei numeri interi; sono anche inversamente proporzionali alla massa della particella e alla larghezza del pozzo. Vale a dire: per una massa molto grande, od un pozzo molto largo, gli autovalori sono molto vicini tra loro (lo spettro diviene quasi *continuo*), e il comportamento della particella si avvicina al comportamento previsto dalla meccanica classica. Le autofunzioni si trovano subito tenendo conto $A_1 = -A_2$; imponendo la condizione che siano normalizzate in $[0, L]$ si trova:

$$\psi_k(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{k\pi}{L} x\right) \quad (7.65)$$

Supponiamo ora che la particella sia inizialmente al centro del pozzo:

$$\psi(x, 0) = \delta(x - L/2) \quad (7.66)$$

Da cui segue che al tempo t la funzione d'onda è

$$\psi(x, t) = \frac{2}{L} \sum_{k \geq 1} \exp \left[\frac{k^2 \hbar \pi^2}{2mL^2} t \right] \sin \left(\frac{k\pi}{L} x \right) \sin \left(\frac{k\pi}{2} \right) \quad (7.67)$$

ma il seno di $k\pi/2$ vale zero per k pari; per $k = 2n + 1$ dispari, con $n \geq 0$, vale $(-1)^n$; perciò

$$\psi(x, t) = \frac{2}{L} \sum_{n \geq 0} (-1)^n \exp \left[\frac{(2n+1)^2 \hbar \pi^2}{2mL^2} t \right] \sin \left(\frac{k\pi}{L} x \right)$$

Naturalmente per una scatola con pareti infinite tutte le autoenergie sono discrete (non esistono stati "liberi" della particella).

Esempio 33. Calcolare e rappresentare in un grafico ai tempi t : 1 ps, 10 ps, 1 ns e 1 μ s la densità di probabilità $|\psi^2(x, t)|$ per un neutrone in un pozzo largo 1 Å, che sia posto al centro del pozzo a $t = 0$.



Figura 7.6: *Escape from New York*

7.2.1 Fuga dal pozzo

E se una le pareti del pozzo avessero un'altezza finita? Il potenziale è rappresentato in Fig. (7.7):

$$V(x) = \begin{cases} V & x \leq 0 \\ 0 & 0 < x < L \\ V & L \leq x \end{cases} \quad (7.69)$$

Consideriamo solo lo spettro discreto, vale a dire gli stati confinati della particella, $E < V$. La soluzione generica per l'autofunzione $\psi_E(x)$ con autoenergia E è definita nei singoli intervalli come

$$\psi(x) = \begin{cases} B_1 \exp(bx) & x \leq 0 \\ A_1 \exp(iax) + A_2 \exp(-iax) & 0 \leq x < L \\ B_2 \exp(-bx) & L \leq x \end{cases} \quad (7.70)$$

dove $a = \sqrt{2mE/\hbar^2}$ e $b = \sqrt{2m(V-E)/\hbar^2}$; si noti che per $x < 0$ si pone a zero il coefficiente dell'esponenziale $\exp(-bx)$ che altrimenti divergerebbe per $x \rightarrow -\infty$ (ed analogamente per $x \geq L$).

Valgono poi quattro condizioni relative alla continuità e derivabilità dell'autofunzione in $x = 0$ e $x = L$:

$$A_1 + A_2 = B_1 \quad (7.71)$$

$$iaA_1 - iaA_2 = bB_1 \quad (7.72)$$

$$A_1 \exp(iaL) + A_2 \exp(-iaL) = B_2 \exp(-bL) \quad (7.73)$$

$$iaA_1 \exp(iaL) - iaA_2 \exp(-iaL) = -bB_2 \exp(-bL) \quad (7.74)$$

Risolvendo il determinante 4×4 (scrivetelo), si ottiene una condizione in E . Saltando vari passaggi algebrici il risultato è:

$$\arcsin \frac{\hbar a}{\sqrt{2mV}} = \frac{n\pi - kL}{2} \quad (7.75)$$

con $n \in \mathcal{N}$, che si risolve in a . Si verifica subito che il numero degli autostati discreti è finito.

7.2.2 Una scatola con una barriera interna

L'esempio successivo è relativo al caso di un pozzo con una barriera interna finita, Fig. (7.7). Lo schema di soluzione è analogo ai casi precedenti: i) definizione delle forme generali delle autofunzioni; ii) imposizione di condizioni al contorno; iii) determinazione degli autostati dalla soluzione di un sistema lineare omogeneo. L'algebra è un po' più complicata, ma per geometrie elementari si semplifica considerevolmente. Il potenziale

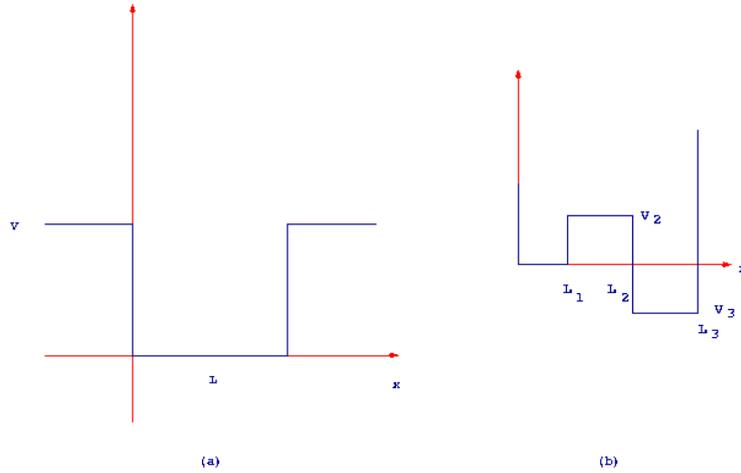


Figura 7.7: Potenziali per una particella in una scatola 1D con una barriera interna.

che descrive il nostro pozzo è dato in $[0, L_3]$:

$$V(x) = \begin{cases} V_1 & 0 \leq x < L_1 \\ V_2 & L_1 \leq x < L_2 \\ V_3 & L_2 < x \leq L_3 \end{cases} \quad (7.76)$$

La soluzione generica per l'autofunzione $\psi_E(x)$ con autoenergia E è definita nei singoli intervalli come

$$\psi(x) = \begin{cases} A_1^{(1)} \exp(\alpha_1^{(1)} x) + A_2^{(1)} \exp(\alpha_2^{(1)} x) & 0 \leq x < L_1 \\ A_1^{(2)} \exp(\alpha_1^{(2)} x) + A_2^{(2)} \exp(\alpha_2^{(2)} x) & L_1 \leq x < L_2 \\ A_1^{(3)} \exp(\alpha_1^{(3)} x) + A_2^{(3)} \exp(\alpha_2^{(3)} x) & L_2 < x \leq L_3 \end{cases} \quad (7.77)$$

dove $\alpha_{1,2}^{(i)} = \pm i \sqrt{\frac{2m(E-V_i)}{\hbar^2}}$. Le condizioni al contorno sono analoghe al primo caso: ai confini del pozzo l'autofunzione deve annullarsi, poichè la particella non può penetrare una barriera infinita. I coefficienti ignoti dell'autofunzione sono sei, e disponiamo di due condizioni al contorno più altre quattro condizioni: per ciascun dei punti interni, in L_1 ed L_2 la funzione deve essere continua e derivabile. In definitiva, valgono le sei relazioni lineari:

$$\begin{aligned} A_1^{(1)} + A_2^{(1)} &= 0 \\ A_1^{(1)} \exp(\alpha_1^{(1)} L_1) + A_2^{(1)} \exp(\alpha_2^{(1)} L_1) &= A_1^{(2)} \exp(\alpha_1^{(2)} L_1) + A_2^{(2)} \exp(\alpha_2^{(2)} L_1) \\ \alpha_1^{(1)} A_1^{(1)} \exp(\alpha_1^{(1)} L_1) + \alpha_2^{(1)} A_2^{(1)} \exp(\alpha_2^{(1)} L_1) &= \alpha_1^{(2)} A_1^{(2)} \exp(\alpha_1^{(2)} L_1) + \alpha_2^{(2)} A_2^{(2)} \exp(\alpha_2^{(2)} L_1) \\ A_1^{(2)} \exp(\alpha_1^{(2)} L_2) + A_2^{(2)} \exp(\alpha_2^{(2)} L_2) &= A_1^{(3)} \exp(\alpha_1^{(3)} L_2) + A_2^{(3)} \exp(\alpha_2^{(3)} L_2) \\ \alpha_1^{(2)} A_1^{(2)} \exp(\alpha_1^{(2)} L_2) + \alpha_2^{(2)} A_2^{(2)} \exp(\alpha_2^{(2)} L_2) &= \alpha_1^{(3)} A_1^{(3)} \exp(\alpha_1^{(3)} L_2) + \alpha_2^{(3)} A_2^{(3)} \exp(\alpha_2^{(3)} L_2) \\ A_1^{(3)} \exp(\alpha_1^{(3)} L_3) + A_2^{(3)} \exp(\alpha_2^{(3)} L_3) &= 0 \end{aligned} \quad (7.78)$$

da cui si ricava il determinante 6×6 che permette di determinare E .

Esempio 2. Scrivete il determinante relativo al sistema precedente.

Consideriamo il caso della buca simmetrica, Fig. (7.7). Il sistema precedente assume la forma, un poco più maneggevole:

$$\begin{aligned}
 A_1^{(1)} + A_2^{(1)} &= 0 \\
 A_1^{(1)} \exp(iaL) + A_2^{(1)} \exp(-iaL) &= A_1^{(2)} \exp[(ia - v)L] + A_2^{(2)} \exp[-(ia - v)L] \\
 iaA_1^{(1)} \exp(iaL) - iaA_2^{(1)} \exp(-iaL) &= (ia - v)A_1^{(2)} \exp[(ia - v)L] - (ia - v)A_2^{(2)} \exp[-(ia - v)L] \\
 A_1^{(2)} \exp[2(ia - v)L] + A_2^{(2)} \exp[-2(ia - v)L] &= A_1^{(3)} \exp(2iaL) + A_2^{(3)} \exp(-2iaL) \\
 (ia - v)A_1^{(2)} \exp[2(ia - v)L] - (ia - v)A_2^{(2)} \exp[-2(ia - v)L] &= iaA_1^{(3)} \exp(2iaL) - iaA_2^{(3)} \exp(-2iaL) \\
 A_1^{(3)} \exp(3iaL) + A_2^{(3)} \exp(-3iaL) &= 0
 \end{aligned} \tag{7.79}$$

dove $a = \sqrt{2mE/\hbar^2}$ e $v = \sqrt{2mV/\hbar^2}$

Esempio 3. Scrivete il determinante.

Dal calcolo del determinante si ottiene un'equazione trascendente (non polinomiale) in a (cioè in E), che si può risolvere per via numerica. A questo punto lo schema di soluzione per un potenziale generico monodimensionale che sia rappresentabile da una serie di segmenti orizzontali e verticali, con pareti infinite, vale a dire un pozzo monodimensionale infinitamente profonda con un "fondo" qualunque, dovrebbe risultare chiaro: 1) si imposta la soluzione generica in ciascun intervallo; 2) si impongono condizioni di continuità e derivabilità in ogni punto interno; 3) si annulla la soluzione alle due pareti. Il risultato è un sistema

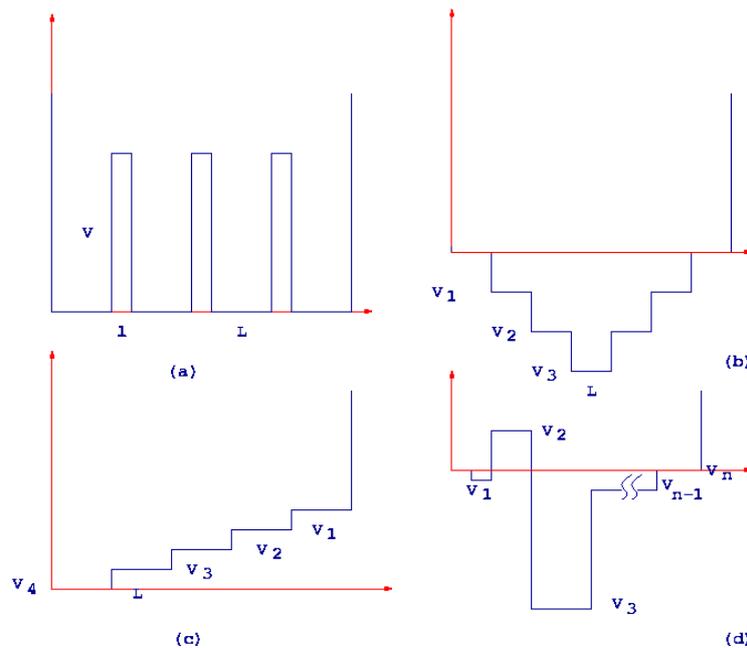


Figura 7.8: Varie scatole 1D.

omogeneo di equazioni per i coefficienti dell'autofunzione in ciascun intervallo, il cui determinante si deve annullare e fornisce un'equazione per calcolare le energie.

Esempio 34.

1. Impostare e risolvere numericamente il problema agli autostati relativo al pozzo (a) di Fig. (7.8). Cosa succede se $V \rightarrow 0$, o $l \rightarrow 0$?
2. Impostare e risolvere numericamente il problema agli autostati relativo al pozzo (b) di Fig. (7.8). Cosa succede se il numero di gradini diviene infinito, mentre la loro larghezza tende a zero (ovvero $N \rightarrow \infty$ con $NL \rightarrow \text{cost.}$)?
3. Impostare e risolvere numericamente il problema agli autostati relativo al pozzo (c) di Fig. (7.8). Cosa succede se il numero di gradini diviene infinito, mentre la loro larghezza tende a zero (ovvero $N \rightarrow \infty$ con $NL \rightarrow \text{cost.}$)?

4. Impostare e risolvere numericamente il problema agli autostati relativo ad un pozzo generico, come in (d) di Fig. (7.8).

7.2.3 Scatola tridimensionale

Una particella di massa m in tre dimensioni è descritto dall'hamiltoniano generico

$$\mathcal{H} = -\frac{\hbar^2}{2m} \hat{\nabla}^2 + U(\mathbf{x}) \quad (7.80)$$

dove $\hat{\nabla}$ è l'operatore *laplaciano*, che in coordinate cartesiane ha la forma

$$\hat{\nabla}^2 = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2} \quad (7.81)$$

Nel caso di una particella racchiusa in una scatola in tre dimensioni, con pareti di lato a_1 , a_2 e a_3 , il potenziale è nullo all'interno della scatola, ed infinito all'esterno. Il problema è ridotto allo studio di tre

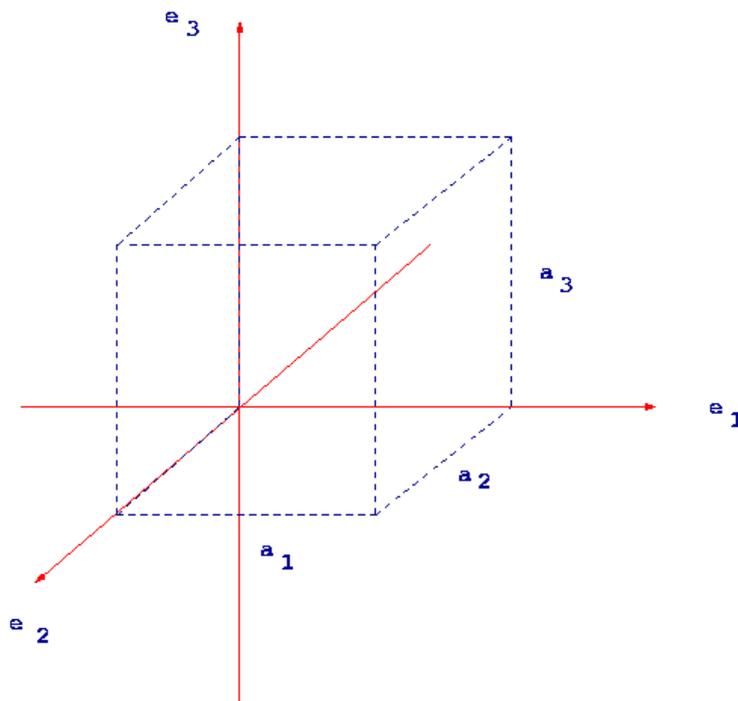


Figura 7.9: Scatola tridimensionale.

problemi separati lungo i tre assi; cioè l'autofunzione del sistema si può scrivere come:

$$\psi_E(\mathbf{x}) = \psi_E^{(1)}(x_1)\psi_E^{(2)}(x_2)\psi_E^{(3)}(x_3) \quad (7.82)$$

dove $\psi_E^{(1)}(x_1)$ è definita per $0 < x < a_1$ e così via. Le autoenergie sono date dall'espressione:

$$E = \frac{\hbar^2 \pi^2}{2m} \left(\frac{k_1^2}{a_1^2} + \frac{k_2^2}{a_2^2} + \frac{k_3^2}{a_3^2} \right) \quad (7.83)$$

dove k_1 , k_2 e k_3 sono interi. I corrispondenti autostati sono dati dall'equazione inserendo le forme esplicite di $\psi_E^{(1)}(x_1)$ etc. nella (7.82):

$$\psi_E(\mathbf{x}) = \sqrt{\frac{8}{a_1 a_2 a_3}} \sin\left(\frac{\pi k_1}{a_1}\right) \sin\left(\frac{\pi k_2}{a_2}\right) \sin\left(\frac{\pi k_3}{a_3}\right) \quad (7.84)$$

È importante notare che nel caso tridimensionale le autoenergie possono essere *degeneri*: per una scatola cubica ($a_1 = a_2 = a_3$) una coppia di autostati identificate da terne k_1, k_2 e k_3 che differiscano solo per una permutazione hanno lo stesso energia.

Esempio 35. Consideriamo il caso di una scatola bidimensionale quadrata: i) scrivere le espressioni generali per le autoenergie e gli autostati; ii) porre in grafico le autofunzioni corrispondenti ai primi cinque autovalori distinti.

7.2.4 Particella *in e su* un anello

Considereremo in questa Sezione e nelle Sezioni successive problemi agli autovalori relativi a sistemi che si possono risolvere analiticamente introducendo coordinate curvilinee, non cartesiane, che sono dettate naturalmente dalla geometria del potenziale. Come primo caso prendiamo in analisi una particella che sia racchiusa in una zona dello spazio bidimensionale in una circonferenza: $V = 0$ se $|\mathbf{x}| < a$, $V = \infty$ se $|\mathbf{x}| > a$, vedi Fig. (7.10) (a); a è il raggio della circonferenza; oppure che sia costretta a muoversi sulla circonferenza stessa (*rotatore planare*). L'hamiltoniano per il sistema nel piano è:

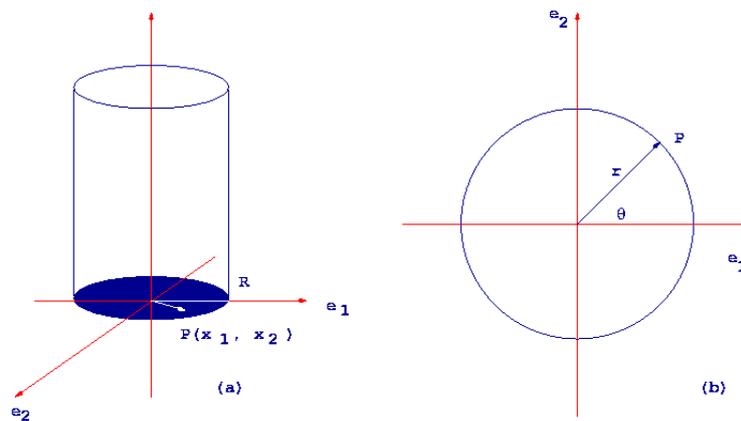


Figura 7.10: Un pozzo cilindrico ed un anello.

$$\mathcal{H} = -\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} \right) \quad (7.85)$$

Introduciamo le *coordinate polari*: r, ϕ , dove r è il modulo del vettore $\mathbf{x} = (x_1, x_2)$ e ϕ è l'angolo formato con l'asse delle ascisse. È facile dimostrare che l'hamiltoniano, proporzionale all'operatore laplaciano in due dimensioni, assume la forma:

$$\mathcal{H} = -\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \phi^2} \right) \quad (7.86)$$

Consideriamo prima il caso del rotatore planare. La particella si muove su una circonferenza a $r = a$, di conseguenza le derivate in r si annullano. L'hamiltoniano si riduce al solo termine dipendente da ϕ . L'equazione di Schrödinger assume la forma:

$$-\frac{\hbar^2}{2I} \frac{d^2}{d\phi^2} \Psi_E(\theta) = E \Psi_E(\phi) \quad (7.87)$$

dove $I = ma^2$ è il momento di inerzia. Risolvendo, si trova come nel caso della particella nella scatola monodimensionale:

$$\Psi_E(\theta) = A_1 \exp(im_1\phi) + A_2 \exp(-im_1\phi) \quad (7.88)$$

dove $m_l = \sqrt{2IE/\hbar^2}$. Le condizioni al contorno sono da ricercarsi nelle proprietà fondamentali delle funzioni d'onda di un sistema quantomeccanico, che devono sempre dar luogo ad una densità di probabilità univoca: cioè dato \mathbf{x} , un punto dello spazio delle coordinate del sistema, la funzione $|\phi(\mathbf{x})|$ deve essere univocamente determinata. Nel caso in questione questo implica che la funzione calcolata in ϕ o in $\phi + 2\pi$ (quindi nello stesso punto), deve essere uguale:

$$\Psi_E(\phi + 2\pi) = \Psi_E(\phi) \quad (7.89)$$

quindi m_l deve essere un numero intero. Gli autovalori sono dunque:

$$E = \frac{m_l^2 \hbar^2}{2I}, \quad m_l \in \mathcal{N} \quad (7.90)$$

Lo stato di una particella con autoenergia E è dunque descritto da una somma di due esponenziali complessi, che rappresentano in realtà due moti sovrapposti, in un senso (m_l) e nel senso opposto ($-m_l$). Supponiamo di aver preparato la particella con un unico senso di rotazione, $A_2 = 0$; la costante di normalizzazione dell'autofunzione rimanente si ottiene integrando $|\Psi_E(\phi)|$. Si ottiene infine:

$$\Psi_{m_l}(\phi) = \sqrt{\frac{1}{2\pi}} \exp(im_l \phi) \quad (7.91)$$

Ed ora consideriamo il caso di una particella racchiusa in un anello, che cioè sia costretta a muoversi all'interno della circonferenza. L'hamiltoniano è dato dalla (7.86). L'autofunzione generica si può scrivere in una forma fattorizzata, poichè le coordinate r e ϕ sono separabili:

$$\Psi_E(r, \phi) = R(r)\Psi(\phi) \quad (7.92)$$

Sostituendo nell'equazione di Schrödinger, $\mathcal{H}\Psi_E = E\Psi_E$, dividendo per Ψ_E e moltiplicando per r^2 , si ottiene:

$$-\frac{\hbar^2}{2mR} \left(r^2 \frac{d^2 R}{dr^2} + r \frac{dR}{dr} \right) - Er^2 = \frac{\hbar^2}{2m\Psi} \frac{d^2 \Psi}{d\phi^2} \quad (7.93)$$

che è un'equazione separabile; il primo e il secondo membro devono sempre essere uguali, indipendentemente dal valore di r e ϕ ; scriviamo perciò

$$\frac{d^2 \Psi}{d\phi^2} = -m_l^2 \Psi \quad (7.94)$$

dove m_l^2 è una costante; poichè la funzione deve essere univocamente definita e dunque Ψ periodica, m_l è un numero intero. Resta però da risolvere un'equazione in R :

$$r^2 \frac{d^2 R}{dr^2} + r \frac{dR}{dr} + \left(\frac{2mE}{\hbar^2} r^2 - m_l^2 \right) R = 0 \quad (7.95)$$

le condizioni al contorno sono: i) la funzione d'onda deve essere continua ovunque, anche all'origine $r = 0$; e ii) deve annullarsi sul bordo $r = a$. Le funzioni che verificano l'equazione radiale con queste condizioni sono dette funzioni di Bessel di ordine intero:

$$R_{m_l}(r) = J_{m_l} \left(\sqrt{\frac{2mE}{\hbar^2}} r \right) \quad (7.96)$$

con il vincolo:

$$J_{m_l} \left(\sqrt{\frac{2mE_{m_l,n}}{\hbar^2}} a \right) = 0 \quad (7.97)$$

risolvendo l'equazione precedente si trova un set di valori di $E_{m_l, n}$ per ogni valore di m_l , con $n = 1, 2, \dots$. In definitiva la soluzione generica è:

$$\Psi_{m_l, n}(r, \phi) = \exp(im_l \phi) J_{m_l} \left(\sqrt{\frac{2mE_{m_l, n}}{\hbar^2}} r \right) \quad (7.98)$$

con la condizione (7.97) sulle autoenergie.

Esempio 36. Fate una ricerca bibliografica sulle proprietà delle funzioni di Bessel e valutate con l'ausilio di tabelle i valori di $E_{0, l}$ con $l \leq 3$, se la massa della particella è quella di un protone e il raggio della circonferenza è pari a 2 \AA .

7.2.5 Particella in e su una sfera

Da due a tre dimensioni il passo è breve, almeno in questo caso. Generalizziamo le precedenti considerazioni al caso di una sfera: studiamo dunque il problema del moto di una particella di massa m su una sfera di raggio a e poi dentro la medesima sfera. L'hamiltoniano del sistema è in coordinate cartesiane, in analogia

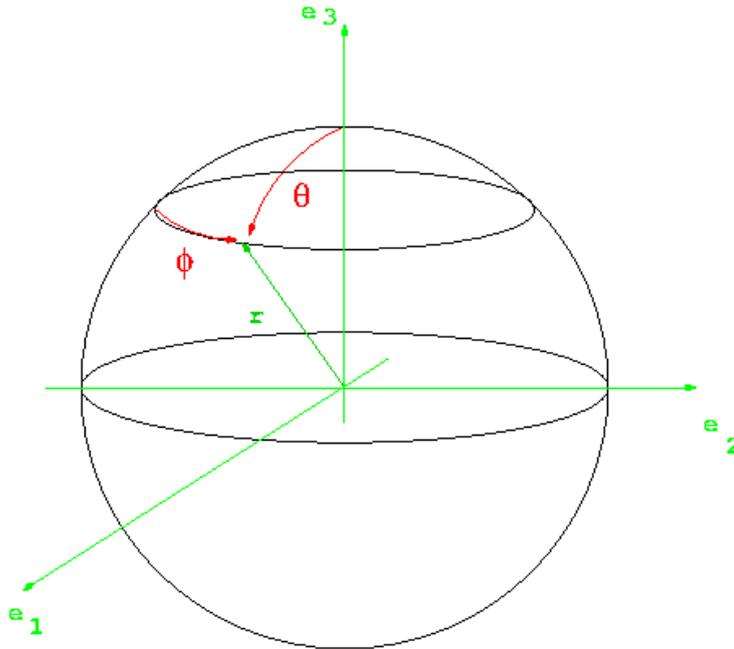


Figura 7.11: Coordinate polari sferiche.

con il caso dell'anello:

$$\mathcal{H} = -\frac{\hbar^2}{2m} \hat{\nabla}^2 \quad (7.99)$$

dove si è fatto uso dell'operatore laplaciano. Introduciamo ora le coordinate polari sferiche r, ϕ, θ , vedi Fig. (7.11):

$$\begin{aligned} x_1 &= r \sin \theta \cos \phi \\ x_2 &= r \sin \theta \sin \phi \\ x_3 &= r \cos \theta \end{aligned} \quad (7.100)$$

Il laplaciano assume una forma particolarmente semplice nelle coordinate sferiche:

$$\hat{\nabla}^2 = \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} + \frac{1}{r^2} \mathcal{M}^2 \quad (7.101)$$

$$\mathcal{M}^2 = \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} \quad (7.102)$$

dove \mathcal{M}^2 è anche detto operatore *legendriano*. Procediamo ora come abbiamo fatto per il caso del rotatore planare. Se la particella è costretta a muoversi sulla superficie della sfera, trascuriamo le derivate in r , e poniamo $r = R$. Dall'equazione di Schrödinger si ottiene un'equazione in $\Psi_E(\phi, \theta)$:

$$-\frac{\hbar^2}{2I} \mathcal{M}^2 \Psi_E = E \Psi_E \quad (7.103)$$

dove $I = ma^2$. La funzione Ψ_E dovrà essere periodica in ϕ e in θ . Inoltre deve essere finita a $\theta = \pi$. La soluzione generale si può scrivere in termini di funzioni speciali, note come funzioni *armoniche sferiche*, che altro non sono, da un certo punto di vista, che l'equivalente tridimensionale degli esponenziali complessi che sono stati usati nel caso bidimensionale. Varie loro proprietà sono note (e si rimanda lo studente ad opportuni testi specializzati). Tra le altre proprietà, fondamentale è il fatto che sono autofunzioni del legendriano:

$$\mathcal{M}^2 Y_{lm_l}(\theta, \phi) = -l(l+1) Y_{lm_l}(\theta, \phi) \quad (7.104)$$

dove $l = 0, 1, \dots$ e $m_l = -l, -l+1, \dots, l-1, l$. Segue dunque che le autoenergie di una particella che si muova su una sfera sono caratterizzate da:

$$E_{lm_l} = \frac{\hbar^2}{2I} l(l+1) \quad (7.105)$$

Si noti che per ogni autostato, caratterizzato dai numeri quantici l e m_l , esistono $2(l+1)$ autovalori degeneri, dato che E_{lm_l} non dipende da m_l .

Procediamo ora con il caso della particella confinata nella sfera. Ora dobbiamo considerare il laplaciano nella sua forma completa. Analogamente al caso precedente, si può assumere che la parte angolare debba essere quella determinata dal moto sulla sfera:

$$\Psi_E(r, \theta, \phi) = R(r) Y_{lm_l}(\theta, \phi) \quad (7.106)$$

Sostituendo nell'Eq. di Schrödinger si trova l'equazione per la parte radiale:

$$\frac{1}{r^2} \frac{d}{dr} r^2 \frac{dR}{dr} - \frac{l(l+1)}{r^2} R + \frac{2m}{\hbar^2} ER = 0 \quad (7.107)$$

con le condizioni al contorno che $R = 0$ per $r = a$ ed R finita a $r = 0$. In generale avremo dunque delle funzioni $R_{nl}(r)$ che si ottengono risolvendo la precedente equazione differenziale ordinaria. Consideriamo nel seguito il solo caso $l = 0$, cioè le soluzioni del problema che siano indipendenti dagli angoli polari ($Y_{0,0}$ è una costante). L'equazione in R si può porre nella forma ridotta

$$\frac{1}{r} \frac{d^2}{dr^2} rR - \kappa^2 R = 0 \quad (7.108)$$

dove $\kappa = \sqrt{2mE/\hbar^2}$. La soluzione finita a $r = 0$ è

$$R(r) = A \frac{\sin(\kappa r)}{r} \quad (7.109)$$

e la condizione al contorno impone che $ka = 0$, vale a dire ancora una volta

$$E = \frac{\hbar^2 n^2 \pi^2}{2I} \quad (7.110)$$

Esempio 37. Normalizzate l'autofunzione radiale. Ricordate che nel passaggio da coordinate cartesiane a coordinate sferiche si ha che:

$$\int dx_1 \int dx_2 \int dx_3 f(x_1, x_2, x_3) = \int dr r^2 \int d\theta \sin \theta \int d\phi f(r, \theta, \phi)$$

7.2.6 Oscillatore armonico

L'oscillatore lineare o *armonico* è un altro esempio di sistema risolvibile analiticamente, ovvero tale che le sue autofunzioni sono esprimibili in termini di funzioni note. Il potenziale è definito come una parabola, con una curvatura $m\omega^2$:

$$V(x) = \frac{m\omega^2}{2}x^2 \quad (7.111)$$

ω è la frequenza dell'oscillatore. Consideriamo l'hamiltoniano del sistema:

$$\left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{m\omega^2}{2}x^2 \right) \psi_E(x) = E\psi_E(x) \quad (7.112)$$

Introduciamo per comodità di calcolo la coordinata adimensionale $\xi = \sqrt{\frac{m\omega}{\hbar}}x$. L'equazione in $\psi(\xi)$ diviene:

$$\frac{d^2}{d\xi^2} \psi_E + \left(\frac{2E}{\hbar\omega} - \xi^2 \right) \psi_E = 0 \quad (7.113)$$

La funzione ψ deve essere finita per $\xi \rightarrow \infty$; per x molto grandi si può trascurare il termine in E rispetto a ξ^2 , quindi la funzione si deve comportare asintoticamente come $\exp(-\xi^2/2)$:

$$\psi_E(\xi) = \exp(-\xi^2/2)\chi(\xi) \quad (7.114)$$

dove $\chi(\xi)$ può anche divergere con $\xi \rightarrow \infty$, ma non più velocemente di $\exp(-\xi^2/2)$; possiamo perciò ottenere l'equazione in $\chi(\xi)$:

$$\frac{d^2\chi}{d\xi^2} - 2\xi \frac{d\chi}{d\xi} + \left(\frac{2E}{\hbar\omega} - 1 \right) \chi = 0 \quad (7.115)$$

La soluzione generica all'equazione differenziale precedente, con la condizione al contorno che per $\xi \rightarrow \infty$ la funzione χ non cresca più rapidamente di una potenza finita di ξ (in modo che comunque ψ_E vada a zero per $\xi \rightarrow \infty$, è data dai polinomi di Hermite, con il vincolo che

$$\frac{2E}{\hbar\omega} - 1 = 2n \quad (7.116)$$

ed n deve essere naturale. Le autofunzioni di un oscillatore armonico monodimensionale sono dunque proporzionali ai polinomi di Hermite, con autovalori quantizzati dall'indice n . La costante di proporzionalità è determinabile, come sempre, dalla condizione di normalizzazione della funzione d'onda. Sostituendo di nuovo x al posto di ξ infine si trova:

$$\psi_n(x) = \left(\frac{m\omega}{\pi\hbar} \right)^{1/4} \frac{1}{(2^n n!)^{1/2}} \exp\left(-\frac{m\omega}{2\hbar}x^2\right) H_n \left[\left(\frac{m\omega}{\hbar} \right)^{1/2} x \right] \quad (7.117)$$

cui corrisponde l'autovalore

$$E_n = \left(n + \frac{1}{2} \right) \hbar\omega \quad (7.118)$$

Esempio 38. Porre in grafico le prime tre autofunzioni di un oscillatore lineare con la massa di un elettrone e $\omega = 10^{-14} \text{ s}^{-1}$.

La generalizzazione ad un oscillatore armonico in tre dimensioni è semplice. L'equazione di Schrödinger è data da:

$$\left[-\frac{\hbar^2}{2m} \nabla^2 + \frac{1}{2} (\omega_1^2 x_1^2 + \omega_2^2 x_2^2 + \omega_3^2 x_3^2) \right] \Psi_E(x_1, x_2, x_3) = E\Psi_E(x_1, x_2, x_3) \quad (7.119)$$

analogamente al caso della particella in tre dimensioni (e in generale in tutti i casi in cui il potenziale sia scritto come una somma di potenziali separati agenti ciascuno su una coordinata), possiamo scrivere gli autostati come il prodotto degli autostati lungo ciascuna coordinata

$$\psi_{n_1, n_2, n_3}(\mathbf{x}) = \prod_{i=1}^3 \left(\frac{m\omega_i}{\pi\hbar} \right)^{1/4} \frac{1}{(2^{n_i} n_i!)^{1/2}} \exp\left(-\frac{m\omega_i}{2\hbar} x_i^2\right) H_{n_i} \left[\left(\frac{m\omega_i}{\hbar} \right)^{1/2} x_i \right] \quad (7.120)$$

e le autoenergie come somme:

$$E = \sum_i \left(n_i + \frac{1}{2} \right) \hbar\omega_i \quad (7.121)$$

Esempio 39. Qual è la degenerazione in funzione di $n = n_1 + n_2 + n_3$ dell'autostato $E = (n_1 + n_2 + n_3 + 3/2)\hbar\omega$ per un oscillatore armonico tridimensionale con $\omega_1 = \omega_2 = \omega_3 = \omega$?

7.2.7 Potenziale di Morse

Consideriamo il potenziale

$$V(x) = D(1 - e^{-\alpha x})^2 \quad (7.122)$$

detto *potenziale di Morse* ed è mostrato in Fig. (7.12). Per $x \rightarrow \infty$ il potenziale va a D ; per $x \rightarrow -\infty$ il

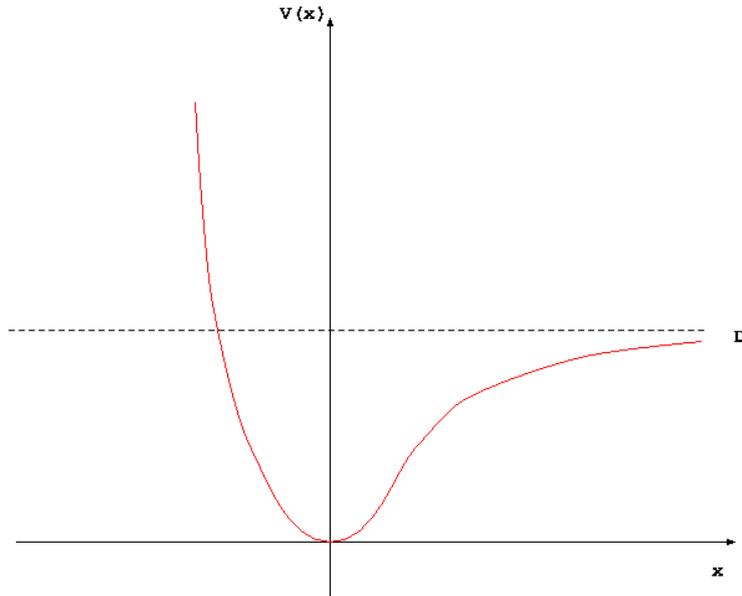


Figura 7.12: Potenziale di Morse.

potenziale va a $+\infty$. Evidentemente esisteranno autovalori discreti e positivi per $E < D$, ed uno spettro continuo per $E > D$. L'equazione di Schrödinger è data come al solito da

$$-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi_E(x) + D(1 - e^{-\alpha x})^2 \psi_E(x) = E \psi_E(x) \quad (7.123)$$

Introduciamo la variabile ausiliaria ξ che va da 0 a ∞

$$\xi = \frac{2\sqrt{2mD}}{\alpha\hbar} \exp(-\alpha x) \quad (7.124)$$

L'equazione di Schrödinger assume la forma ridotta:

$$\frac{\partial^2 \psi_E}{\partial \xi^2} + \frac{1}{\xi} \frac{\partial \psi_E}{\partial \xi} + \left(-\frac{1}{4} + \frac{n+s+1/2}{\xi} - \frac{s^2}{\xi^2} \right) \psi_E = 0 \quad (7.125)$$

dove $s = \frac{\sqrt{2m(D-E)}}{\alpha\hbar}$ e $n = \frac{\sqrt{2mD}}{\alpha\hbar} - s - \frac{1}{2}$. Per $\xi \rightarrow \infty$ la funzione deve evolvere come $\exp(-\xi/2)$ e per $\xi \rightarrow 0$ come ξ^s . La soluzione generica si può scrivere perciò come

$$\psi_E(\xi) = \exp(-\xi/2) \xi^s \chi(\xi) \quad (7.126)$$

Sostituendo, si trova un'equazione differenziale in $\chi(\xi)$ (come si vede la procedura è sostanzialmente analoga al caso dell'oscillatore armonico):

$$\xi \frac{\partial^2 \chi}{\partial \xi^2} + (2s+1-\xi) \frac{\partial \chi}{\partial \xi} + n\chi = 0 \quad (7.127)$$

La soluzione generica è detta *funzione ipergeometrica confluyente* e si scrive in generale $\chi = F(-n, 2s+1, \xi)$; sappiamo che χ deve essere finita per $\xi \rightarrow 0$; per $\xi \rightarrow \infty$ χ non può tendere all'infinito più rapidamente di una potenza finita di ξ . Queste condizioni si ottengono dallo studio delle proprietà delle funzioni ipergeometriche confluenti, solo nel caso n intero non negativo. Da ciò segue che le autoenergie discrete sono:

$$E_n = D \left\{ 1 - \left[1 - \frac{\alpha\hbar}{\sqrt{2mD}} \left(n + \frac{1}{2} \right) \right]^2 \right\} \quad (7.128)$$

Si noti che se $\alpha\hbar/\sqrt{2mD} > 2$ non esistono autovalori discreti. Il potenziale di interazione tra gli atomi di una molecola biatomica può essere rappresentato in modo qualitativo da una curva di Morse. In realtà possiamo interpretare un sistema rappresentato dalla curva di Morse come un esempio di oscillatore *anarmonico*.

7.2.8 Atomo idrogenoide

Concludiamo questa Sezione con un sistema 'chimico' semplice, risolvibile analiticamente: il calcolo dello spettro di autoenergie e delle relative autofunzioni per l'elettrone di un atomo di idrogeno con un nucleo a riposo. In generale, se il nucleo non è a riposo dobbiamo semplicemente sostituire alla massa dell'elettrone la *massa ridotta* del sistema protone/elettrone. L'hamiltoniano del sistema è perciò:

$$\mathcal{H} = -\frac{\hbar^2}{2\mu} \hat{\nabla}^2 - \frac{e^2}{4\pi\epsilon_0} \frac{1}{r} \quad (7.129)$$

dove μ è la massa ridotta e il secondo termine è il potenziale coulombiano attrattivo tra il nucleo con carica $+e$ e l'elettrone con carica $-e$ a distanza r . Il potenziale è a simmetria radiale. Possiamo perciò procedere senza perdere tempo come nel caso della particella racchiusa in una sfera: i) introduzione di coordinate sferiche; ii) fattorizzazione della funzione d'onda in una parte angolare dipendente da θ e ϕ e in una parte radiale dipendente da r ; iii) imposizione di condizioni al contorno per il calcolo delle autoenergie. La parte angolare dell'autofunzione non cambia, ed è data ancora dalle funzioni armoniche sferiche: in realtà ciò è vero per qualunque potenziale a simmetria sferica, cioè dipendente solo dal modulo r :

$$\Psi_E(r, \theta, \phi) = R(r) Y_{lm_l}(\theta, \phi) \quad (7.130)$$

Sostituendo nell'Eq. di Schrödinger si trova l'equazione per la parte radiale:

$$\frac{1}{r^2} \frac{d}{dr} r^2 \frac{dR}{dr} - \frac{l(l+1)}{r^2} R + \frac{2m}{\hbar^2} \left(E + \frac{e^2}{4\pi\epsilon_0} \frac{1}{r} \right) R = 0 \quad (7.131)$$

Lo spettro degli autovalori sarà discreto per $E < 0$, e continuo per $E > 0$ (il potenziale coulombiano tende a zero per $r \rightarrow \infty$ e a $-\infty$ per $r \rightarrow 0$). A questo punto possiamo procedere come in tutti i casi precedenti: l'equazione differenziale in R , con le condizioni al contorno R finita a $r \rightarrow 0$ e $R \rightarrow \infty$ è riconoscibile

com l'equazione che ha per soluzione delle funzioni note: le *funzioni associate di Laguerre* (moltiplicate per l'esponenziale di $-r$). In definitiva gli autostati e le autoenergie dell'atomo di idrogeno a riposo sono:

$$\Psi_{n,l,m_l}(r, \theta, \phi) = R_{nl}(r)Y_{lm_l}(\theta, \phi) \quad (7.132)$$

dove le funzioni radiali R_{nl} sono

$$R_{nl} = -\frac{2}{na} \frac{(n-l-1)!}{2n[(n+l)!]^3} \rho^l L_{n+l}^{2l+1}(\rho) \exp(-\rho/2) \quad (7.133)$$

e gli autovalori corrispondenti sono

$$E_{n,l,m_l} = -\frac{\mu e^4}{32\pi^2 \epsilon_0^2 \hbar^2} \frac{1}{n^2} \quad (7.134)$$

dove $\rho = 2r/na$ ed $a = 4\pi\epsilon_0\hbar^2/\mu e^2$. Gli autostati discreti sono negativi e dipendono solo da $1/n^2$.

Esempio 40. Qual è la degenerazione di ciascun livello energetico dell'atomo di idrogeno a riposo? E quanto vale il valore di attesa di r rispetto per l'autofunzione con $n = l = m_l = 0$.

7.3 Metodi numerici per le equazioni differenziali

Discutiamo qui alcuni dei metodi più comuni per la soluzione numerica dei sistemi di n equazioni differenziali ordinarie del primo ordine, a cui come abbiamo visto in precedenza si possono ridurre anche le equazioni differenziali ordinarie di ordine n .

7.3.1 Differenze finite

Consideriamo il sistema

$$\frac{d\mathbf{y}}{dx} = \mathbf{f}(x, \mathbf{y}) \quad (7.135)$$

con le condizioni iniziali $\mathbf{y}(x_0) = \mathbf{y}_0$. Risolvere numericamente il sistema (3.2) significa trovare un'approssimazione al vettore soluzione $\mathbf{y}(x)$ nota nel dominio di variazione di x . I metodi principali di soluzione sono di due tipi. I *metodi variazionali* esprimono, in termini semplici, il vettore soluzione come combinazione lineare di funzioni note $\bar{\mathbf{y}}$

$$\mathbf{y}(x) = \sum_i c_i \bar{\mathbf{y}}_i(x) \quad (7.136)$$

cercando sulla base di qualche criterio i coefficienti c_i che rendono la soluzione il più vicino possibile alla soluzione esatta, rispettando le condizioni iniziali. I *metodi alle differenze finite*⁴ ricercano la soluzione per una successione di valori discreti della variabile indipendente $x_1, x_2, \dots, x_k, \dots$, sostituendo all'equazione (7.135) un'equazione approssimata in cui le derivate prime sono sostituite da *rapporti incrementali* o comunque approssimazioni delle derivate stesse, ottenendo una *equazione alle differenze*.

Consideriamo nel seguito una griglia di punti equidistanti, $x_0, x_1, \dots, x_k, \dots$, dove $x_{k+1} = x_k + h$, dove h è un intervallo dato, e x_0 è il valore in cui è definita la condizione iniziale del sistema (7.135). Se è nota la soluzione \mathbf{y}_k in x_k , la soluzione in x_{k+1} si può trovare in linea di principio dall'espansione

$$\mathbf{y}_{k+1} = \mathbf{y}_k + h\mathbf{y}'(x_k) + \frac{h^2}{2}\mathbf{y}''(x_k) + \dots \quad (7.137)$$

⁴Esistono molti altri schemi di discretizzazione, e modalità diverse di cercare le soluzioni approssimate numericamente in punti, aree, volumi etc. del dominio di definizione dell'equazione; nel contesto delle equazioni differenziali ordinarie tuttavia, il metodo delle differenze finite resta il più generale, popolare e di facile implementazione.

da cui

$$\mathbf{y}'(x_k) = \frac{\mathbf{y}_{k+1} - \mathbf{y}_k}{h} - \frac{h}{2}\mathbf{y}''(x_k) + \dots \quad (7.138)$$

Sostituendo questa espressione nelle (7.135), trascurando tutti i termini di ordine superiore al primo in h , e calcolando la funzione \mathbf{f} in x_k , si ottiene il metodo alle differenze finite di propagazione approssimata più semplice, detto schema di Eulero

$$\frac{1}{h}(\mathbf{y}_{k+1} - \mathbf{y}_k) = \mathbf{f}(x_k, \mathbf{y}_k) \Rightarrow \mathbf{y}_{k+1} = \mathbf{y}_k + h\mathbf{f}(x_k, \mathbf{y}_k) \quad (7.139)$$

Lo schema di Eulero ha un interesse più che altro accademico, perchè è inaccurato e fortemente instabile, a meno di non scegliere valori estremamente piccoli di h tali da rendere lo schema praticamente inapplicabile. Con il termine *stabilità* si intende il fatto che un errore eventuale in \mathbf{y}_k non si propaghi, rendendo la soluzione divergente, per $k \rightarrow \text{infity}$) Però è utile per capire l'idea generale utilizzata per risolvere in modo numericamente approssimato mediante le differenze finite un sistema come (7.135): se h è sufficientemente piccolo (al limite per $h \rightarrow 0$) la successione $\mathbf{y}_1, \mathbf{y}_2, \dots$ è la soluzione corretta entro una precisione definita nei punti x_1, x_2, \dots . Lo schema di Eulero è uno schema del *primo ordine* (ad ogni passaggio, o *step*, l'errore è di ordine h^2 , ed *esplicito*).

7.3.2 Schemi di discretizzazione

Un vantaggio innegabile dello schema di Eulero è la sua 'economicità' dal punto di vista computazionale (ad ogni step richiede solo la valutazione esplicita di \mathbf{f} nel punto precedente). Un modo quindi molto usato, ed implementato in routine di calcolo utilizzabili anche con sistemi di equazioni differenziali abbastanza complicati, purchè con un vettore di funzioni \mathbf{f} ad elementi continui, e che in generale non siano caratterizzati da *stiffness*, rigidità (vedi oltre), è di mantenere lo schema esplicito, ma di aumentarne l'ordine. In pratica, si tratta di valutare la funzione \mathbf{f} non direttamente nel punto successivo a x_k , cioè x_{k+1} in un punto intermedio (o in più punti intermedi) tra x_k e x_{k+1} , scelti in modo da ridurre l'errore. Lo schema più comune è detto *schema del quarto ordine di Runge-Kutta*

$$\begin{aligned} \mathbf{b}_1 &= h\mathbf{f}(x_k, \mathbf{y}_k) \\ \mathbf{b}_2 &= h\mathbf{f}\left(x_k + \frac{h}{2}, \mathbf{y}_k + \frac{\mathbf{b}_1}{2}\right) \\ \mathbf{b}_3 &= h\mathbf{f}\left(x_k + \frac{h}{2}, \mathbf{y}_k + \frac{\mathbf{b}_2}{2}\right) \\ \mathbf{b}_4 &= h\mathbf{f}(x_k + h, \mathbf{y}_k + \mathbf{b}_2) \end{aligned} \quad (7.140)$$

$$\mathbf{y}_{k+1} = \mathbf{y}_k + \frac{\mathbf{b}_1}{6} + \frac{\mathbf{b}_2}{3} + \frac{\mathbf{b}_3}{3} + \frac{\mathbf{b}_4}{6} \quad (7.141)$$

lo schema è del quarto ordine perchè l'errore commesso è dell'ordine di h^5 ; lo schema di Runge-Kutta è comunque esplicito, ma richiede quattro valutazioni di \mathbf{f} ad ogni step.

Altri schemi di discretizzazione possono essere utilizzati, e sono correntemente implementati in librerie di calcolo numerico. In particolare, sono importanti i cosiddetti metodi multistep (per sempio il metodo Adams-Bashforth), in cui ad ogni step sono utilizzate informazioni relative non solo al punto precedente, ma anche a punti più lontani. Soprattutto però, di grande interesse per molti problemi chimici, si deve tenere presente che tutti gli schemi espliciti sono caratterizzati, in genere, da una instabilità intrinseca che si manifesta per h sufficientemente grande. Questo problema è particolarmente evidente nei sistemi di equazioni 'rigide', che tipicamente compaiono in quei problemi fisici in cui, se per sempio la variabile indipendente è il tempo, si hanno varie scale dei tempi, molto diverse, che caratterizzano il comportamento dei sistemi descritti (per esempio, un meccanismo di reazione con una reazione molto più veloce delle altre, vedi oltre). In questi casi l'unica ricetta che assicuri la possibilità di trovare una soluzione stabile (anche se magari poco accurata), è di ricorrere a metodi impliciti o semi-impliciti.

Il più semplice e più stabile schema implicito è appunto lo schema di Eulero implicito, ottenuto arrestando al primo ordine l'espansione e calcolando \mathbf{f} in x_{k+1}

$$\mathbf{y}_{k+1} = \mathbf{y}_k + h\mathbf{f}(x_{k+1}, \mathbf{y}_{k+1}) \quad (7.142)$$

la cui stabilità è massima anche per grandi valori di h (al limite, anche *incondizionatamente*, cioè indipendentemente dal valore di h). Il metodo (7.142) comporta però la soluzione di un sistema di equazione di equazioni non lineari ad ogni step, e quindi a meno di casi in cui la valutazione di \mathbf{f} e la soluzione dell'equazione non lineare (vedi il Capitolo dedicato in queste dispense) siano particolarmente semplici, di non facile applicazione. Si può però per esempio ricorrere a metodi di soluzione semi-impliciti, in cui si *linearizza* il secondo membro della (7.142), risolvendo ad ogni step un sistema di equazioni lineari in \mathbf{y}_{k+1} invece di un sistema di equazioni non lineari.

Tutte le delle metodologie finora discusse sono, ancora una volta, solitamente già implementate in librerie di calcolo numerico disponibili nei linguaggi di programmazione più comuni in ambito scientifico, come ad esempio ODEPACK, libreria Fortran disponibile presso <http://gams.nist.gov/>.

Infine, prima di proseguire, consideriamo il caso in cui il sistema (7.135) debba essere risolto con le condizioni al contorno $g_{1j_1}(x_1, \mathbf{y})$ con $j_1 = 1, \dots, n_1$ e $g_{2j_2}(x_2, \mathbf{y})$ con $j = 1, \dots, n_2$. In questo caso, la soluzione del sistema deve soddisfare due condizioni di bordo, definite nei punti x_1 e x_2 . In questo caso i metodi dell' differenze finite possono ancora essere applicati, ma il calcolo, piuttosto che progredire in modo esplicito od implicito a partire da un punto dato, mirerà a trovare insieme i valori della soluzione in una griglia di punti compresi tra x_1 e x_2 , imponendo le condizioni al contorno sotto forma di equazioni aggiuntive. Nella sezione successiva discuteremo a titolo di esempi alcune applicazioni.

7.3.3 Applicazioni

Tra le applicazioni più comuni in ambito chimico dei metodi di soluzione numerica dei sistemi di equazioni differenziali ordinarie (7.135) abbiamo sostanzialmente tre casi: il primo è l'insieme di applicazioni basate sui metodi della dinamica classica per lo studio delle proprietà di molecole e insiemi di molecole: la prossima sezione è interamente dedicata a questa applicazione. Il secondo è lo studio dei meccanismi di reazione, caratteristici della cinetica chimica. Il terzo è lo studio di problemi al contorno che risultino da equazioni di Schrödinger o di diffusione in una dimensione. Nel resto di questa sezione discutiamo alcuni esempi di soluzioni numeriche, basati su schemi alle differenze finite, del secondo e terzo caso.

Iniziamo dal problema generale dello studio della cinetica di un meccanismo di reazione complesso. In generale supponiamo che in reattore chiuso avvengano m reazioni chimiche elementari che scriviamo per n specie C_j ⁵

$$\sum_{j=1}^n \nu_{ij}^R C_j \xrightarrow{k_i} \sum_{j=1}^n \nu_{ij}^P C_j \quad i = 1, \dots, m \quad (7.143)$$

da cui possiamo ricavare le equazioni differenziali

$$\frac{dc_j}{dt} = \sum_{i=1}^m (\nu_{ij}^P - \nu_{ij}^R) \prod_{j=1}^n c_j^{\nu_{ij}^R} \quad (7.144)$$

dove $c_j = [C_j]$ con le condizioni iniziali

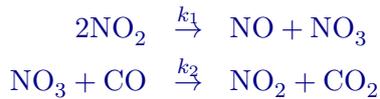
$$c_j(0) = c_j^0 \quad j = 1, \dots, n \quad (7.145)$$

Esempio 41. Verificate che i meccanismi di reazione descritti negli esercizi alla fine del capitolo precedente possono essere scritti nella forma generale (7.143), e che le equazioni differenziali che le descrivono sono descritte dalla forma generale (7.144).

La soluzione del problema (7.144) è dunque ridotto alla forma (7.135) che si può risolvere in modo numerico con uno dei metodi alle differenze finite appena descritti. Si noti che in generale sarà possibile ridurre il numero di equazioni indipendenti, utilizzando opportuni bilanci di massa.

⁵In forma del tutto generale, scriviamo sia a destra che a sinistra di ogni reazione i simboli di tutte le specie chimiche presenti nel sistema; quando una specie non partecipa effettivamente alla reazione, il coefficiente ad esso relativo (come reagente o prodotto) va a zero; in questo modo possiamo scrivere qualunque meccanismo, o insieme di reazioni

Applicazione 9 . Consideriamo il meccanismo



Siano $c_1 = [\text{NO}_2]$, $c_2 = [\text{NO}]$, $c_3 = [\text{NO}_3]$, $c_4 = [\text{CO}]$, $c_5 = [\text{CO}_2]$, da cui

$$\begin{aligned} \dot{c}_1 &= -2k_1c_1^2 + k_2c_3c_4 \\ \dot{c}_2 &= k_1c_1^2 \\ \dot{c}_3 &= k_1c_1^2 - k_2c_3c_4 \\ \dot{c}_4 &= -k_2c_3c_4 \\ \dot{c}_5 &= k_2c_3c_4 \end{aligned}$$

con le condizioni $c_1(0) = c_1^0$, $c_2(0) = 0$, $c_3(0) = 0$, $c_4 = c_4^0$ e $c_5(0) = 0$. La prima delle due reazioni è considerata solitamente il *rate determining step* del meccanismo, cioè $k_1 \ll k_2$; il problema quindi è caratterizzato da due scale dei tempi diverse, e si può considerare 'rigido'. Se riscriviamo le precedenti equazioni in forma vettoriale

$$\dot{\mathbf{c}} = \mathbf{f}(\mathbf{c}) \quad \mathbf{c}(0) = \mathbf{c}_0$$

possiamo risolvere con il metodo di Eulero implicito:

$$\mathbf{c}_{k+1} - h\mathbf{f}(\mathbf{c}_{k+1}) = \mathbf{c}_k$$

N.B.: 1) il numero di specie indipendenti si può ridurre a due introducendo i bilanci di massa rispetto all'azoto, ossigeno e carbonio; 2) la soluzione può essere semplificata considerevolmente applicando l'ipotesi dello stato stazionario all'intermedio NO_3 .

Come secondo esempio, consideriamo il problema al contorno della determinazione degli autovalori di un hamiltoniano quantistico monodimensionale, che descrive il moto di una particella di massa m in un potenziale generico $V(x)$, nell'intervallo $[0, L]$

$$-\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) = E\psi(x) \quad 0 \leq x \leq L \quad (7.146)$$

Introduciamo la griglia $x_0 = 0, x_1 = h, \dots, x_n = nh, x_{n+1} = L$ con $h = \frac{L}{n+1}$. Per definizione la funzione d'onda non potrà esistere per $x < 0$ e $x > L$, quindi le condizioni al contorno hanno la forma $\psi_0 = \psi(x_0) = \psi_{n+1} = \psi(x_{n+1}) = 0$. La derivata seconda che compare nella (7.146) si può discretizzare nel punto generico x_k con $k = 1, \dots, n$ come

$$\left. \frac{d^2\psi(x)}{dx^2} \right|_{x_k} = \frac{\psi_{k+1} - 2\psi_k + \psi_{k-1}}{h^2}$$

e quindi si può scrivere

$$-\frac{\hbar^2}{2mh^2}(\psi_{k+1} - 2\psi_k + \psi_{k-1}) + V_k\psi_k = E\psi_k$$

con $V_k = V(x_k)$, che può essere messo in forma matriciale tridiagonale

$$a_k\psi_{k+1} + (b_k - E)\psi_k + c_k\psi_{k-1} = 0$$

dove $a_k = -\frac{\hbar^2}{2mh^2} = c_k$, $b_k = \frac{\hbar^2}{mh^2} + V_k$. La soluzione è quindi ricondotta ai metodi di algebra lineare precedentemente descritti.

Applicazione 10 . Quali sono gli autovalori dell'hamiltoniano nel caso di un potenziale $V(x) = \Delta\sqrt{x^2 - L^2}$? Quale forma assume la funzione d'onda di stato fondamentale?

7.4 Dinamica molecolare: metodologie ed applicazioni

Questa sezione, dedicata ai metodi della dinamica molecolare, può essere letta indipendentemente dal resto del capitolo.

Le tecniche di calcolo di dinamica molecolare (*molecular dynamics*, MD) sono basate sulla soluzione "esatta" delle equazioni del moto di un insieme di N particelle e sul calcolo diretto delle varie osservabili fisiche:

simulazione matematica \rightarrow riproduzione di un sistema fisico

Da un punto di vista metodologico, possiamo distinguere le seguenti fasi nella messa a punto di una simulazione MD

- definizione delle condizioni del sistema (*insieme termodinamico di appartenenza*)
- definizione dei gradi di libertà e delle interazioni tra le componenti del sistema (*statica*)
- scelta della legge del moto (*dinamica*)
- condizioni al contorno e condizioni iniziali
- scelta del metodo di risoluzione numerica delle equazioni del moto
- calcolo delle osservabili fisiche ed implementazione computazionale: grandezze riscalate, condizioni periodiche al contorno, trattamento di gradi di libertà rotazionali

La definizione di un sistema di particelle interagenti, e la simulazione delle proprietà macroscopiche del sistema stesso, può essere effettuata nell'ambito di uno dei vari *insiemi termodinamici*: l'insieme *microcanonico* classico: un sistema all'equilibrio di N particelle (oppure $N_1, N_2 \dots$ per un sistema a più componenti) interagenti a energia E totale costante e volume V fisso $\rightarrow (E, N, V)$; l'insieme *canonico*: temperatura, numero di particelle e volume costanti $\rightarrow (T, N, V)$; l'insieme *isotermo-isobarico*: temperatura, numero di particelle e pressione costanti $\rightarrow (T, N, P)$; l'insieme *grandcanonico*: potenziale chimico, temperatura e volume costanti $\rightarrow (\mu, T, V)$. Nella maggior parte delle simulazioni MD si utilizza l'insieme microcanonico, si considera cioè un'insieme *isolato* di particelle interagenti.

Le equazioni del moto delle singole molecole sono, nella maggior parte delle applicazioni in letteratura, le equazioni della meccanica classica, nella loro forma standard di equazioni di Newton. Per un insieme di N particelle classiche puntiformi possiamo definire le coordinate e i momenti lineari in modo immediato come le posizioni (vettori cartesiani) $\{\vec{r}(t)_1, \vec{r}(t)_2, \dots, \vec{r}(t)_N\}$ e i vettori momenti lineari: $\{\vec{p}(t)_1, \vec{p}(t)_2, \dots, \vec{p}(t)_N\}$, dove

$$\vec{v}_i = m_i \frac{d\vec{r}_i(t)}{dt} = m_i \vec{v}_i(t) \quad (7.147)$$

Le proprietà statiche del sistema sono definite se è nota l'energia potenziale:

$$V = V(\{\vec{r}(t)_1, \vec{r}(t)_2, \dots, \vec{r}(t)_N\}) \quad (7.148)$$

che contribuisce all'energia totale del sistema

$$H = K + V = \sum_{i=1}^N \frac{\vec{p}_i^2}{2m_i} + V(\vec{r}) = \frac{1}{2} \vec{p} \cdot \mathbf{m}^{-1} \cdot \vec{p} + V(\vec{r}) \quad (7.149)$$

La presenza di eventuali campi esterni (e.g. campi gravitazionali, elettrici magnetici, stress meccanici imposti ai sistemi) permette di simulare sistemi fluidi in condizioni di non-equilibrio. Nel seguito faremo riferimento solo a sistemi isolati, aventi energia meccanica costante definita dalla (7.149).

7.4.1 Potenziali di interazione

Nel caso di N particelle (atomi) indistinguibili $m_1 = m_2 = \dots = m$, possiamo definire l'energia potenziale come risultante dalla somma di sole interazione tra coppie di particelle, $V(\vec{r}) = \sum_{i < j} v(\vec{r}_i, \vec{r}_j)$. In generale, l'energia potenziale si può scrivere come la somma di interazioni ad un corpo (interazioni con campi esterni), a due corpi, a tre corpi etc.:

$$V = \sum_i v_1(\vec{r}_i) + \sum_{i < j} v_2(\vec{r}_i, \vec{r}_j) + \sum_{i < j < k} v_3(\vec{r}_i, \vec{r}_j, \vec{r}_k) + \dots \quad (7.150)$$

È noto come fino al 10 % dell'energia reticolare di un cristallo di argon possa essere dovuta ai termini di interazione a tre corpi; nel seguito, in analogia alla maggioranza delle applicazioni MD, trascureremo termini di interazione a più di due corpi. I potenziali di coppia utilizzati per la simulazione MD di fluidi non-ionici

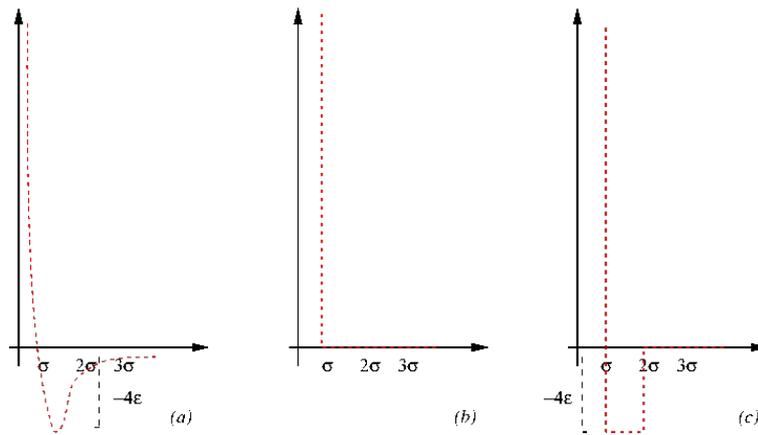


Figura 7.13: Potenziali: Lennard-Jones (a), hard-sphere (b), square-well (c)

sono di solito i seguenti

- 6-12 o Lennard-Jones (LJ):

$$v^{LJ}(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] \quad (7.151)$$

Il potenziale LJ è il più usato per la maggior parte delle applicazioni MD in cui si intendano simulare proprietà statiche e dinamiche di fluidi atomici (o molecolari, vedi oltre), poichè tiene conto in modo qualitativamente corretto sia delle interazioni repulsive che attrattive tra due particelle.

- potenziale hard-sphere e square-well:

$$v^{HS}(r) = \begin{cases} \infty & (r < \sigma) \\ 0 & (\sigma \leq r) \end{cases} \quad v^{SW}(r) = \begin{cases} \infty & (r < \sigma_1) \\ -\epsilon & (\sigma_1 \leq r < \sigma_2) \\ 0 & (\sigma_2 \leq r) \end{cases} \quad (7.152)$$

I potenziali hard-sphere e square-well sono potenziali semplificati che permettono una soluzione più rapida, dal punto di vista computazionale, delle equazioni del moto, rispetto ai potenziali continui come il potenziale di LJ.

- potenziali soft-sphere:

$$v^{SS}(r) = \epsilon \left(\frac{\sigma}{r} \right)^n \quad (7.153)$$

L'utilizzazione di potenziali soft-sphere è accettabili soprattutto nell'ambito di simulazioni di particelle non-atomiche (e.g. sospensioni di particelle macroscopiche in fasi fluide, colloidi etc.).

Per i sistemi ionici si deve considerare l'interazione elettrostatica:

$$v^z(r_{ij}) = \frac{z_i z_j}{4\pi\epsilon_0 r_{ij}} \quad (7.154)$$

I potenziali elettrostatici, che sono intrinsecamente diversi dai potenziali precedenti in quanto espressione di forze a *lungo raggio*, sono molto difficili da trattare nell'ambito di simulazioni MD, a causa della limitata dimensione del sistema trattabile nell'ambito delle consuete applicazioni MD.

7.4.2 Liquidi molecolari

Nel caso di sistemi formati da molecole non mono-atomiche, la definizione dei gradi di libertà del sistema è notevolmente complicata dalla presenza di coordinate rotazionali, ed eventualmente, per molecole non rigide, di coordinate interne. Da un punto di vista energetico, i sistemi molecolari possono essere agevolmente descritti da siti localizzati, che rappresentano singoli atomi o gruppi di atomi, interagenti mediante potenziali sito-sito del tipo descritto nella sezione precedente:

$$v(\vec{r}_{ij}, \mathbf{\Omega}_i, \mathbf{\Omega}_j) = \sum_a \sum_b v_{ab}(r_{ab}) \quad (7.155)$$

dove si indica con $\mathbf{\Omega}_i$ il set di angoli di Eulero che specifica l'orientazione della molecola i (vedi oltre), con $\vec{r}_{ij} = \vec{r}_i - \vec{r}_j$ la differenza tra i vettori che indicano il centro di massa delle molecole i e j e con $r_{ab} = |\vec{r}_{i_a} - \vec{r}_{j_b}|$ la differenza tra i vettori che indicano il sito a della molecola i e il sito b della molecola j . Applicazioni più sofisticate prevedono anche l'uso di *potenziali Lennard-Jones generalizzati*:

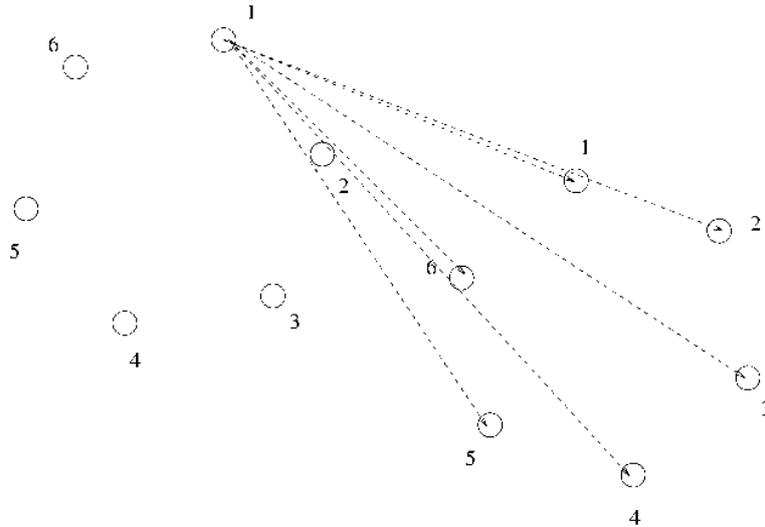


Figura 7.14: Due molecole interagenti a 6 siti

$$\epsilon(\mathbf{\Omega}_i, \mathbf{\Omega}_j) = \epsilon[1 - \chi^2(\mathbf{e}_i \cdot \mathbf{e}_j)^2]^{-1/2} \quad (7.156)$$

$$\sigma(\vec{r}_{ij}, \mathbf{\Omega}_i, \mathbf{\Omega}_j) = \sigma \left[1 - \frac{1}{2} \chi f(\vec{r}_{ij} \cdot \mathbf{e}_i, \vec{r}_{ij} \cdot \mathbf{e}_j, \mathbf{e}_i \cdot \mathbf{e}_j) \right] \quad (7.157)$$

e si possono poi considerare, soprattutto per sistemi a molti siti termini aggiuntivi di energia interna vibrazionale, torsionale ed interazioni dipolari, quadrupolari etc. Poichè il nostro studio sarà limitato a fluidi formati da *molecole lineari biatomiche rigide*, non considereremo nel dettaglio questo tipo di interazioni.

Leggi del moto

Le equazioni del moto definiscono al tempo t la posizione e la velocità di ciascuna particella. Per un sistema di N molecole interagenti, i gradi di libertà traslazionali \vec{r}_i , e le corrispondenti velocità lineari \vec{v}_i obbediscono alle equazioni del moto

$$m_i \dot{\vec{v}}_i = \vec{F}_i \quad (7.158)$$

dove \vec{F}_i è la forza totale agente sulla i -esima molecola, ottenuta come gradiente rispetto a \vec{r}_i del potenziale V ; indichiamo con \dot{a} la derivata temporale di una quantità a . Nel caso di molecole rigide non-monoatomiche dobbiamo considerare anche i gradi di libertà rotazionali di ciascuna molecola; in generale, nel caso di molecole rigide non lineari possiamo definire l'orientazione di un sistema di riferimento solidale con la molecola rispetto al laboratorio utilizzando come coordinate gli angoli di Eulero $\Omega = \phi, \theta, \psi$, oppure in termini di quaternioni $Q = q_0, q_1, q_2, q_3$. Alle coordinate rotazionali, definite in termini di angoli di Eulero o quaternioni, è associato il vettore di velocità angolare $\vec{\omega}$. La legge del moto rotazionale di una molecola rigida non-lineare i è in generale scritta in termini del momento della coppia $\vec{\tau}_i$ che è l'analogo rotazionale della forza agente sulle coordinate traslazionali del centro di massa \vec{F}_i . In generale, se le componenti di $\vec{\tau}_i$ e $\vec{\omega}_i$ sono riferite ad un sistema di riferimento solidale con la molecola i , le equazioni del moto rotazionale sono

$$\begin{aligned}\dot{\omega}_{ix} &= \frac{\tau_{ix}}{I_{ixx}} + \left(\frac{I_{iyy} - I_{izz}}{I_{ixx}} \right) \omega_{iy}\omega_{iz} \\ \dot{\omega}_{iy} &= \frac{\tau_{iy}}{I_{iyy}} + \left(\frac{I_{izz} - I_{ixx}}{I_{iyy}} \right) \omega_{iz}\omega_{ix} \\ \dot{\omega}_{iz} &= \frac{\tau_{iz}}{I_{izz}} + \left(\frac{I_{ixx} - I_{iyy}}{I_{izz}} \right) \omega_{ix}\omega_{iy}\end{aligned}\tag{7.159}$$

$$\tag{7.160}$$

dove $I_{ixx}, I_{iyy}, I_{izz}$ sono i valori principali del tensore di inerzia della molecola. Le leggi del moto rotazionale sono completate dalle relazioni geometriche che legano la velocità angolare con le derivate temporali delle coordinate rotazionali prescelte, siano esse gli angoli di Eulero o i quaternioni. Nel caso di molecole lineari, le equazioni del moto sono leggermente modificate, per tener conto della degenerazione del moto attorno all'asse di simmetria molecolare.

Sistemi termodinamici e condizioni di simulazione

In assenza di campi esterni, il sistema costituito da N molecole interagenti viene considerato un insieme microcanonico. In pratica, il numero limitato di molecole considerato ($N \sim O(10^2)$) e il tempo di simulazione anch'esso necessariamente limitato (dell'ordine di decine di picosecondi) implica che la condizione di equilibrio termodinamico conseguita dal sistema sia un'approssimazione. In pratica, il sistema è inizialmente preparato in una qualche condizione facilmente definibile (e.g. uno stato cristallino ordinato), e con una data densità ρ . Alle molecole è assegnata un'energia cinetica iniziale distribuendo in modo casuale, secondo una distribuzione di Maxwell, le velocità lineari ed angolari iniziali. Dopodichè si conducono uno o più lunghi *cicli di equilibrizzazione*, che ahnno lo scopo di portare il sistema nelle condizioni il più vicino possibile a quelle di equilibrio termodinamico. Solo a questo punto inizia la vera e propria simulazione, con la memorizzazione delle traiettorie delle particelle ed il calcolo di vari osservabili.

Osservabili

In generale, una proprietà osservabile \mathcal{A}_{obs} viene calcolata come media temporale della funzione che lega l'osservabile \mathcal{A} alle coordinate roto-traslazionali delle molecole costituenti il sistema:

$$\mathcal{A}_{obs} = \langle \mathcal{A} \rangle_{tempo} = \langle \mathcal{A}[\Gamma(t)] \rangle = \lim_{t \rightarrow \infty} \frac{1}{t_{obs}} \int_0^{t_{obs}} dt \mathcal{A}[\Gamma(t)]\tag{7.161}$$

Le funzioni termodinamiche sono perciò immediatamente calcolabili dalle loro definizioni microscopico-statistiche:

energia totale

$$E = \langle H \rangle = \langle K \rangle + \langle V \rangle\tag{7.162}$$

in un insieme isolato questa quantità deve essere costante, e dunque uguale all'energia assegnata al sistema a $t = 0$, e può quindi essere utilizzata per controllare l'accuratezza del metodo numerico di risoluzione delle equazioni del moto;

temperatura

$$3Nk_B T = 2\langle K \rangle = \left\langle \sum_{i=1}^N \frac{\vec{p}_i^2}{m_i} \right\rangle \quad (7.163)$$

la temperatura è calcolata in termini di energia cinetica media;

pressione

$$PV - Nk_B T = \langle W \rangle = \left\langle -\frac{1}{3} \sum_{i=1}^N \vec{r}_i \cdot \hat{\nabla}_i V \right\rangle \quad (7.164)$$

la pressione è calcolata come media della funzione **viriale** W .

Una funzione di grande utilità è la **funzione di correlazione di coppia**; per un fluido atomico, la funzione di correlazione è

$$g(\vec{r}_1, \vec{r}_2) = \frac{N(N-1)}{\rho^2 Z_{NVT}} \int d\vec{r}_3 d\vec{r}_4 \dots d\vec{r}_N \exp \left[-\frac{V(\vec{r}_1, \dots, \vec{r}_N)}{k_B T} \right] \quad (7.165)$$

In pratica la seguente definizione operativa (con media su tutte le coppie) è utilizzata

$$g(\vec{r}) = \frac{1}{\rho^2} \langle \sum_i \sum_{j \neq i} \delta(\vec{r}_i) \delta(\vec{r}_j - \vec{r}) \rangle = \frac{V}{N^2} \langle \sum_i \sum_{j \neq i} \delta(\vec{r} - \vec{r}_{ij}) \rangle \quad (7.166)$$

e per un sistema isotropo $g(\vec{r}) = g(r)$. Ogni media di una funzione di coppie di particelle è esprimibile in funzione di $g(r)$:

$$\langle a(\vec{r}_i, \vec{r}_j) \rangle = \frac{1}{V^2} \int d\vec{r}_i d\vec{r}_j g(\vec{r}_i, \vec{r}_j) a(\vec{r}_i, \vec{r}_j) \quad (7.167)$$

$$\langle \mathcal{A} \rangle = \langle \sum_i \sum_{j \neq i} a(r_{ij}) \rangle = \frac{N\rho}{2} \int_0^\infty dr 4\pi r^2 a(r) g(r) \quad (7.168)$$

Di conseguenza numerose osservabili possono essere calcolate indirettamente una volta che sia nota la funzione $g(r)$

energia

$$E = \frac{3}{2} Nk_B T + 2\pi N\rho \int_0^\infty dr r^2 v(r) g(r); \quad (7.169)$$

pressione

$$PV = Nk_B T - \frac{2}{3} \pi N\rho \int_0^\infty dr r^2 \frac{dv(r)}{dr} g(r); \quad (7.170)$$

potenziale chimico

$$\mu = k_B T \ln(\rho \Lambda^3) + 4\pi\rho \int_0^1 d\xi \int_0^\infty dr r^2 v(r) g(r; \xi) \quad (7.171)$$

dove $\Lambda = \sqrt{h^2/2\pi m k_B T}$ è la frequenza termica di De Broglie.

Funzioni termodinamiche più complesse quali l'energia libera di Gibbs, G , l'energia di Helmholtz, A , l'entropia, S . sono funzioni riconducibili al calcolo diretto della funzione di partizione Q :

$$\exp \left(\frac{A}{k_B T} \right) = \frac{1}{Q_{NVT}} = \langle \exp \left(\frac{V}{k_B T} \right) \rangle \quad (7.172)$$

il cui calcolo diretto nell'ambito del sistema microcanonico è difficoltoso. Si può perciò ricorrere a simulazioni MD che utilizzano insiemi termodinamici diversi come l'insieme NVT, ottenibile per esempio mediante una modifica stocastica della velocità di una o più particelle di un sistema NVE per simulare l'interazione con

una ipotetica particella del termostato; oppure coime l'insieme NPT, ottenibile con la definizione di un grado di libertà aggiuntivo che rappresenta un 'pistone' che esercita una pressione esterna P sul sistema.

Le proprietà del sistema dipendenti dal tempo sono descrivibili dallo *o quadratico mediodi* un'osservabile, che misura l'entità delle fluttuazioni del'oservabile dal valore medio

$$\sigma^2(\mathcal{A}) = \langle \mathcal{A}^2 \rangle = \langle (\mathcal{A} - \langle \mathcal{A} \rangle)^2 \rangle = \langle \mathcal{A}^2 \rangle - \langle \mathcal{A} \rangle^2 \quad (7.173)$$

e soprattutto dalle funzioni di correlazione temporale che misurano la correlazione media tra il valore di un osservabile al tempo t e al tempo $t = 0$:

$$C_{\mathcal{A}\mathcal{B}}(t) = \langle \delta \mathcal{A}(t) \delta \mathcal{B}(0) \rangle = \langle \delta \mathcal{A}[\Gamma(t)] \delta \mathcal{B}[\Gamma(0)] \rangle \quad (7.174)$$

Una funzione di correlazione si dice di autocorrelazione se $\mathcal{A} = \mathcal{B}$. Le funzioni di correlazioni hanno molte applicazioni:

- danno informazioni sul comportamento dinamico dei fluidi;
- i loro integrali sono legati a grandezze macroscopiche;
- le loro trasformate di Fourier sono legate a osservabili spettroscopiche.

Un esempio di una classe di osservabile legata al calcolo di funzioni di correlazione è dato dai **coefficienti di trasporto**. Il coefficiente traslazionale di diffusione di una particella sonda è per esempio definito in funzione della funzione di autocorrelazione della velocità lineare:

$$D = \frac{1}{3} \int_0^\infty dt \langle \vec{v}(t) \vec{v}(0) \rangle = \frac{1}{6} \lim_{t \rightarrow \infty} \frac{1}{t} \langle |\vec{r}_i(t) - \vec{r}_i(0)|^2 \rangle \quad (7.175)$$

Il calcolo di funzioni di correlazione rotazionali è utile per la comprensione della natura dei moti roto-traslazionali nei liquidi molecolari; inoltre opportune funzioni di correlazione rotazionale possono essere messe in relazione a vari osservabili spettroscopici

- spettri infrarosso

$$C_{1mm'}(t) = \langle Y_{1m}[\Omega(t)] Y_{1m'}[\Omega(0)] \rangle \quad (7.176)$$

- spettri Raman, NMR ed EPR

$$C_{2mm'}(t) = \langle Y_{2m}[\Omega(t)] Y_{2m'}[\Omega(0)] \rangle \quad (7.177)$$

Lo studio di liquidi rotazionali permette di verificare varie teorie per la descrizione della fase liquida. Un'osservabile importante è la funzione di autocorrelazione della velocità angolare:

$$c_{\omega\omega}(t) = \langle \omega(t) \cdot \omega(0) \rangle / \langle \omega^2 \rangle \quad (7.178)$$

che può essere messa in relazione con il coefficiente di diffusione rotazione di una singola molecola.

7.4.3 Metodi computazionali di MD

Come tutte le metodologie di simulazione, anche per la Dinamica Molecolare è necessario, al fine di implementare le potenziali applicazioni di questa tecnica allo studio delle proprietà dei fluidi, utilizzare una serie di accorgimenti pratici. Il primo aspetto che vogliamo sottolineare, anche per la comprensione dei risultati presentati nei Capitoli successivi è l'impiego di *grandezze scalate*: le varie grandezze fisiche sono di solito ricondotte e presentate in termini di unità di misura convenientemente scelte (vedi Tabella (7.4.3)). Così la funzione hamiltoniana scalata risulta essere:

$$H^* = \frac{1}{2} \sum_i v_i^{*2} + \sum_{i < j} V^*(|r_i^* - r_j^*|) \quad (7.179)$$

Grandezza	Grandezza riscalata	Coefficiente di scala (atomi di Ar)
Massa	$m^* = m/m_u$	$m_u = m_{Ar}$
Spazio	$l^* = l/l_u$	$l_u = \sigma$
Tempo	$t^* = t/t_u$	$t_u = \sqrt{\sigma^2 m_{Ar} / \epsilon}$
Energia	$E^* = E/E_u$	$E_u = \epsilon$
Forza	$F^* = F/F_u$	$F_u = \epsilon/\sigma$
Temperatura	$T^* = T/T_u$	$T_u = \epsilon/k_B$
Densità	$\rho^* = \rho/\rho_u$	$\rho_u = \sigma^{-3}$
Pressione	$P^* = P/P_u$	$P_u = \epsilon/\sigma^3$

Tabella 7.1: Grandezze riscalate in unità σ

dove il potenziale scalato è:

$$V^*(r^*) = 4 \left(\frac{1}{r^{*12}} - \frac{1}{r^{*6}} \right) \quad (7.180)$$

Un'altro aspetto pratico dell'implementazione dei metodi MD è costituito dall'uso di *condizioni periodiche*: nello studio dei fluidi, i sistemi trattati dalle simulazioni MD comprendono solitamente 10^2 , 10^3 particelle (nel caso per esempio di simulazione massive, altamente parallelizzate, di studio di stress meccanici in solidi atomici sono stati effettuati studi con 10^6 particelle). In presenza di una parete (sistema a volume costante), la maggior parte delle particelle interagirebbe solo con la parete e di conseguenza gli effetti di superficie sarebbero dominanti: a titolo di esempio per un sistema di 1000 particelle in un cristallo cubico ($10 \times 10 \times 10$), 488 giacciono sulle pareti. Nella maggior parte delle applicazioni, si preferisce perciò utilizzare condizioni periodiche, in cui il sistema è a contatto con un numero infinito di repliche, tali che se r_{x_i} , r_{y_i} , r_{z_i} è la posizione della particella i -esima, ogni sua replica ha posizione $r_{x_i} + n_x L_x$, $r_{y_i} + n_y L_y$, $r_{z_i} + n_z L_z$, dove n_x , n_y , n_z sono numeri interi. Ogni particella costituente il sistema interagisce con tutte le particelle

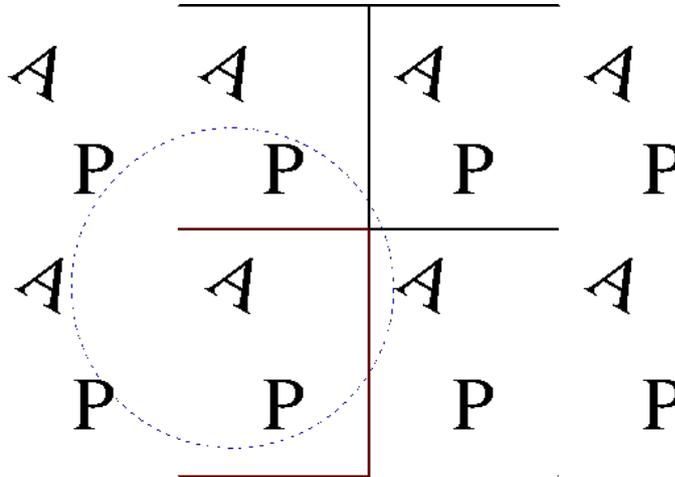


Figura 7.15: Condizioni al contorno periodiche e raggio di cut-off

delle repliche, il calcolo del potenziale agente sulla particella può essere effettuato in vari modi: utilizzando il metodo della *convenzione dell'immagine minima*: si considerano tutte le molecole che giacciono in una pseudoscatola centrata sulla molecola in esame; per un potenziale a coppie si calcolano $\frac{1}{2}N(N-1)$ termini; oppure usando il *raggio di cutoff*: si considerano solo le molecole giacenti in una sfera di raggio $r_c < \frac{1}{2}L$ (es. $r_c = 2.5\sigma$). Altre condizioni al contorno sono utilizzabili per studi MD specifici come le condizioni al contorno *ipersferiche*, in cui le molecole giacciono sulla superficie di un'ipersfera; e soprattutto l'impiego di *cluster* in cui si considerano veri cluster molecolari, con condizioni al contorno non-periodiche.

7.4.4 Metodi alle differenze finite

Il metodo principale di soluzione delle equazioni di Newton in MD consiste nella discretizzazione delle traiettorie rispetto al tempo usando uno dei vari schemi in cui le derivate temporali sono approssimate da

rapporti incrementali. Qualunque metodologia di calcolo dovrà ottimizzare i seguenti requisiti, imposti dalle limitate risorse computazionali

1. elevata velocità di calcolo;
2. minima richiesta di memoria su disco;
3. accuratezza nel calcolo delle traiettorie;
4. semplicità di implementazione;
5. conservazione dell'energia e della quantità di moto;
6. reversibilità temporale

Lo schema di discretizzazione più semplice delle equazioni del moto è basato, come già sappiamo, sull'approssimazione di Eulero. Se indichiamo con \vec{a} , \vec{v} , \vec{r} i vettori collettivi delle accelerazioni, velocità, vettori di posizione di tutte le particelle del sistema (e consideriamo per semplicità un fluido atomico), l'equazioni di Newton discretizzate secondo lo schema di Eulero sono:

$$\vec{a}(t + \delta t) \approx \vec{a}(t) \quad (7.181)$$

$$\vec{v}(t + \delta t) \approx \vec{v}(t) + \vec{a}(t)\delta t \quad (7.182)$$

$$\vec{r}(t + \delta t) \approx \vec{r}(t) + \vec{v}(t)\delta t + \frac{1}{2}\vec{a}(t)\delta t^2 \quad (7.183)$$

Come abbiamo visto sopra, il metodo di Eulero è molto semplice, e facilmente implementabile purtroppo, a meno di non usare un *passo temporale* δt estremamente piccolo non è né stabile né reversibile. Metodi di discretizzazione più sofisticati devono perciò essere utilizzati. Nella Sezione successiva dedicata allo studio di molecole lineari, lo schema di discretizzazione utilizzato in questo lavoro sarà descritto in dettaglio. Tra agli schemi di discretizzazione più utilizzati, soprattutto per fluidi atomici, possiamo ricordare il cosiddetto **algoritmo di Verlet** che consiste in una ridefinizione dello schema di discretizzazione di Eulero nel modo seguente:

$$\vec{r}(t + \delta t) = \vec{r}(t) + \vec{v}(t)\delta t + \frac{1}{2}\vec{a}\delta t^2 \quad (7.184)$$

$$\vec{r}(t - \delta t) = \vec{r}(t) - \vec{v}(t)\delta t + \frac{1}{2}\vec{a}\delta t^2 \quad (7.185)$$

sommando si elimina la velocità:

$$\vec{r}(t + \delta t) = 2\vec{r}(t) - \vec{r}(t - \delta t) + \vec{a}\delta t^2 \quad (7.186)$$

Poiché al tempo t sono noti $\vec{a}(t)$, $\vec{r}(t)$ e $\vec{r}(t - \delta t)$, al tempo $t + \delta t$ si ricava $\vec{r}(t + \delta t)$, mentre la velocità al tempo t è:

$$\vec{v}(t) = \frac{\vec{r}(t + \delta t) - \vec{r}(t - \delta t)}{2\delta t} \quad (7.187)$$

Il metodo di Verlet è stabile, permette di usare step fino a 0.01 ps, è reversibile e conserva energia e momento.

7.4.5 Molecole lineari

Nel caso di una molecola lineare, la definizione dell'orientazione molecolare è specificabile da solo due coordinate, che indichino la direzione e il versore di un versore unitario \vec{e} diretto lungo l'asse molecolare. È dunque possibile utilizzare una descrizione semplificata della geometria e delle equazioni del moto, anche se è possibile utilizzare descrizioni per esempio basate sulla geometria dei quaternioni, al prezzo comunque di una ridondanza nella definizione della geometria molecolare.

Il momento di inerzia relativo alla rotazione intorno all'asse di simmetria è nullo, mentre i due rimanenti valori principali di \mathbf{I} sono uguali. La velocità angolare e il momento della coppia $\vec{\tau}$ sono sempre perpendicolari all'asse molecolare. Di conseguenza, possiamo scrivere

$$\vec{\tau} = \vec{e} \times \vec{g} \quad (7.188)$$

dove \vec{g} è una funzione delle forze agenti sulla molecola. Per una molecola lineare formata da M siti puntiformi, per la quale cioè ciascun sito sia definito rispetto al centro di massa dall'identità vettoriale

$$\vec{d}_j = d_j \vec{e} \quad j = 1, \dots, M \quad (7.189)$$

il vettore \vec{g} è definito come

$$\vec{g} = \sum_{j=1}^{j=M} d_j \vec{f}_j \quad (7.190)$$

dove \vec{f}_j è la forza agente sul sito j . Il vettore \vec{g} può essere sostituito con la componente perpendicolare all'asse molecolare $\vec{g}^\perp = \vec{g} - (\vec{g} \cdot \vec{e})\vec{e}$. Le equazioni del moto rotazionale di una molecola lineare si possono scrivere nella seguente forma:

$$\dot{\vec{e}} = \vec{u} \quad (7.191)$$

$$\dot{\vec{u}} = \frac{\vec{g}^\perp}{I} + \lambda \vec{e} \quad (7.192)$$

La prima equazione definisce il momento generalizzato \vec{u} come derivata prima temporale di \vec{e} . La seconda descrive il moto rotazionale della molecola in presenza del momento di coppia generalizzato \vec{g} , con il termine aggiuntivo $\lambda \vec{e}$ che tiene conto del vincolo di lunghezza costante della molecola: λ è dunque un moltiplicatore indeterminato di Lagrange. Nel corso di una simulazione MD, λ viene determinato numericamente nell'ambito dello schema di discretizzazione impiegato. Tra gli schemi più comuni descriviamo nel dettaglio il metodo *leap-frog* che è anche stato impiegato nel corso delle simulazioni MD presentate in questo lavoro. Al tempo t siano noti i vettori $\vec{e}(t)$, $\vec{u}(t - \delta t/2)$ e $\vec{g}^\perp(t)$. Il vettore $\vec{u}(t)$ è calcolato dallo schema discreto al primo ordine

$$\vec{u}(t) = \vec{u}(t - \delta t/2) + \frac{1}{2} \delta t \left[\frac{\vec{g}^\perp(t)}{I} + \lambda(t) \vec{e}(t) \right] \quad (7.193)$$

se si prende il prodotto scalare con $\vec{e}(t)$ di entrambi i membri, si ottiene

$$\lambda(t) = -2\vec{u}(t - \delta t/2) \cdot \vec{e}(t) / \delta t \quad (7.194)$$

tenendo conto del fatto che $\vec{e} \cdot \vec{u} = \vec{e} \cdot \vec{g}^\perp = 0$. È ora possibile calcolare in modo diretto il vettore $\dot{\vec{u}}(t)$ dall'equazione (7.192) definita al tempo t . Noto $\dot{\vec{u}}$ si può ora valutare $\vec{u}(t + \delta t/2)$:

$$\vec{u}(t + \delta t/2) = \vec{u}(t - \delta t/2) + \delta t \dot{\vec{u}}(t) \quad (7.195)$$

Infine il versore di orientazione al tempo $t + \delta t$ è calcolato dall'equazione (discretizzata) (7.191)

$$\vec{e}(t + \delta t) = \vec{e}(t) + \delta t \vec{u}(t + \delta t/2) \quad (7.196)$$

7.5 Cenni di teoria delle perturbazioni

Dato un sistema descritto dall'equazione di Schrödinger, di solito è possibile conoscere in modo esatto, o quantomeno accurato, la soluzione esatta di un problema agli autovalori relativo ad un sistema fisico semplificato rispetto al sistema originario, di cui si vogliono effettivamente calcolare le proprietà. Se il sistema semplificato differisce dal sistema reale solo per un termine piccolo dell'hamiltoniano (*perturbazione*) è possibile cercare sistematicamente la correzione alle autoenergie ed autostati del sistema semplificato mediante la *teoria delle perturbazioni indipendenti dal tempo*.

Supponiamo perciò che l'hamiltoniano del sistema reale sia

$$\mathcal{H} = \mathcal{H}_0 + \epsilon \mathcal{V} \quad (7.197)$$

dove \mathcal{H}_0 è l'hamiltoniano del sistema semplificato (imperturbato) e \mathcal{V} è la perturbazione; ϵ è un parametro che definisce l'ordine di grandezza della perturbazione. Sia nota la soluzione al problema imperturbato (PI)

$$\mathcal{H}_0|n\rangle = E_n^{(0)}|n\rangle \quad (7.198)$$

Gli autovettori $|n\rangle$ costituiscono una base ortonormale e completa dello spazio hilbertiano. Cerchiamo la soluzione al problema perturbato (PP)

$$\mathcal{H}|\Phi\rangle = E|\Phi\rangle \quad (7.199)$$

Nel seguito, per semplicità considereremo inizialmente solo sistemi con spettri di autovalori discreti e *non degeneri*. Possiamo rappresentare le autofunzioni del PP nella base costituita dalle autofunzioni del PI:

$$|\Phi\rangle = \sum_n c_n |n\rangle \quad (7.200)$$

Sostituendo la precedente eguaglianza nell'espressione del PP si ottiene

$$\sum_n c_n \left[E_n^{(0)} + \epsilon \mathcal{V} \right] |n\rangle = E \sum_n c_n |n\rangle \quad (7.201)$$

ed integrando per un generico $\langle m|$

$$\left[E - E_m^{(0)} \right] c_m = \epsilon \sum_n V_{mn} c_n \quad (7.202)$$

dove $V_{mn} = \langle m|\mathcal{V}|n\rangle$. Si vuole calcolare la correzione all' m -esimo autostato, cioè si vuole che per $\epsilon \rightarrow 0$ (perturbazione nulla) il sistema perturbato tenda all'autostato m -esimo del sistema perturbato. Espandiamo ora i coefficienti c_n e l'energia E in serie di ϵ :

$$E = E_m^{(0)} + \epsilon E^{(1)} + \epsilon^2 E^{(2)} + \dots \quad (7.203)$$

$$c_n = \delta_{mn} + \epsilon c_n^{(1)} + \epsilon^2 c_n^{(2)} + \dots \quad (7.204)$$

come si vede, per $\epsilon \rightarrow 0$ l'energia diventa $E_m^{(0)}$, mentre i coefficienti c_n diventano δ_{mn} , e quindi 1 se $m = n$, e zero in caso contrario. Per calcolare le correzioni alle energie e ai coefficienti si procede sostituendo le espansioni in serie nelle equazioni (7.202), raccogliendo i termini nella stessa potenza di ϵ a primo e secondo membro, ed eguagliando. Così per esempio la correzione al primo ordine per le energie si ottiene conservando solo i termini di primo ordine nell'equazione (7.202) con $n = m$:

$$E^{(1)} = V_{mm} \quad (7.205)$$

mentre le correzioni al primo ordine per i coefficienti si ottengono con $n \neq m$:

$$c_n^{(1)} = \frac{V_{mn}}{E_m^{(0)} - E_n^{(0)}} \quad n \neq m \quad (7.206)$$

il coefficiente $c_m^{(1)}$ resta indeterminato. Si può specificare imponendo che l'autostato del PP al primo ordine sia normalizzato, a meno di termini al primo ordine: si può verificare che ciò avviene per $c_m^{(1)} = 0$. Le approssimazioni successive si calcolano in modo analogo.

Sistemi in ordine zero con autostati degeneri

Supponiamo che l'hamiltoniano del problema imperturbato ammetta un set di autofunzioni degeneri k volte $|m_1\rangle, |m_2\rangle, \dots, |m_k\rangle$ con autoenergia $E_m^{(0)}$. Possiamo definire un set di combinazioni lineari delle k autofunzioni, che sono ancora autofunzioni dell'operatore \mathcal{H}_0

$$|\phi\rangle = \sum_{j=1}^k |m_j\rangle a_j \quad (7.207)$$

Gli autovalori del sistema perturbato sono ancora scritte nella forma

$$E = E_m^{(0)} + \sum_{l=1}^{\infty} \epsilon^l E^{(l)} \quad (7.208)$$

mentre per la generica autofunzione perturbata possiamo scrivere

$$|\Phi\rangle = |\phi\rangle + \sum_{n \neq m} c_n |n\rangle \quad (7.209)$$

Per i coefficienti a_{ji} e c_n vale lo sviluppo perturbativo

$$a_j = \sum_{l=0}^{\infty} a_j^{(l)} \quad c_n = \sum_{l=0}^{\infty} c_n^{(l)} \quad (7.210)$$

Si tratta ora di trovare i termini correttivi $E^{(l)}$, e i coefficienti $a_j^{(l)}$ e $c_n^{(l)}$; possiamo limitarci a calcolare le correzioni degli autovalori al primo ordine, cioè calcolare solo $E^{(1)}$ e i coefficienti a_j, c_n in ordine zero. Sostituendo le varie espressioni precedenti nel PP e considerando i termini in ϵ^1 :

$$\left(\mathcal{H}_0 - E_m^{(0)}\right) \sum_{j=1}^k a_j^{(0)} |m_j\rangle = (E_m^{(1)} - \mathcal{V}) \sum_{j=1}^k a_j^{(0)} |m_j\rangle \quad (7.211)$$

che può essere moltiplicata per il generico elemento $\langle m_i |$. Il termine relativo al sistema imperturbato si annulla, e resta l'espressione

$$\sum_{j=1}^k \langle m_i | \mathcal{V} | m_j \rangle = E_m^{(1)} a_i^{(0)} \quad (7.212)$$

dove naturalmente le autofunzioni imperturbate sono ritenute ortonormali; in forma matriciale la precedente espressione è semplicemente $\mathbf{V}\mathbf{a} = E\mathbf{a}$, chiamando per semplicità \mathbf{V} , \mathbf{a} ed E la matrice rappresentativa di \mathcal{V} , il vettore dei coefficienti $a_j^{(0)}$ e la correzione al primo ordine all'autovalore degenero del PI. Si noti che se il sistema non è degenero, al posto di $a_j^{(0)}$ nella relazione abbiamo 1, e la somma è semplicemente $\langle m | \mathcal{V} | m \rangle$, e dunque si ritrova l'espressione della correzione al primo ordine per stati non-degeneri.

Esempio 42. Consideriamo ancora l'hamiltoniano dell'esempio precedente, che descrive una particella di massa m in un scatola di lato L . L'hamiltoniano del problema imperturbato è naturalmente

$$\mathcal{H} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2}$$

che, come vedremo nel capitolo successivo, ammette il set completo di autofunzioni

$$|k\rangle = \sqrt{\frac{2}{L}} \sin\left(\frac{k\pi x}{L}\right)$$

con autovalori non-degeneri $E_k^{(0)} = k^2 \hbar^2 / 2m$ e $k > 1$. Qual è la correzione al primo ordine dell'autovalore più basso del sistema, che definisce l'energia di punto zero del sistema imperturbato, $E_1^{(0)} = \hbar^2 / 2m$? Sostituendo troviamo

$$E_1^{(1)} = \langle 1 | \mathcal{V} | 1 \rangle = \frac{2V_0}{L} \int_0^{L/2} dx \sin^2\left(\frac{\pi x}{L}\right) = V_0 \frac{(\pi - 1)}{2\pi}$$

7.5.1 Perturbazioni dipendenti dal tempo

Supponiamo di aver definito un sistema mediante un hamiltoniano \mathcal{H}_0 . Al tempo $t = 0$ accendiamo in modo istantaneo un campo perturbativo dipendente dal tempo, $\mathcal{H}'(t)$, avente la forma

$$\mathcal{H}'(t) = \begin{cases} 0 & t < 0 \\ V(t) & 0 \leq t \leq \tau \\ 0 & t > \tau \end{cases} \quad (7.213)$$

Siano $|n\rangle$ le autofunzioni dell'hamiltoniano del sistema imperturbato \mathcal{H}_0 , con autovalori E_n . Per semplicità consideriamo solo il caso di uno spettro discreto.

Vogliamo determinare la funzione d'onda del sistema perturbato

$$i\hbar \frac{\partial \Psi}{\partial t} = [\mathcal{H}_0 + \mathcal{H}'(t)]\Psi \quad (7.214)$$

Poichè le funzioni $|n\rangle$ costituiscono un set completo o base ortonormale dello spazio hilbertiano delle funzioni d'onda del sistema, possiamo scrivere Ψ come

$$\Psi = \sum_n a_n(t) |n\rangle \exp(-iE_n t/\hbar) \quad (7.215)$$

Assumiamo che in assenza di perturbazione il sistema sia preparato nell'autostato $|i\rangle$

$$\Psi_{\text{in}} = a_i |i\rangle \exp(-iE_i t/\hbar) \quad t < 0 \quad (7.216)$$

cioè $a_n(t) = \delta_{ni}$ per $t < 0$. Durante la perturbazione, $0 \leq t \leq \tau$, i coefficienti cambiano fino a raggiungere i valori $a_{ni}(\tau)$. Al cessare della perturbazione, la funzione d'onda è perciò

$$\Psi_{\text{fin}} = \sum_n a_{ni}(\tau) |n\rangle \exp(-iE_n t/\hbar) \quad (7.217)$$

La probabilità che il sistema sia in uno stato stazionario $|f\rangle$ di energia E_f è perciò data dal coefficiente $a_{fi}(\tau)$:

$$w_{i \rightarrow f} = |a_{fi}(\tau)|^2 \quad (7.218)$$

Il calcolo dei coefficienti dipendenti dal tempo si può effettuare con un metodo perturbativo. Per moltiplicando l'equazione (7.214) per il bra $\langle f|$ ed integrando, otteniamo

$$i\hbar \dot{a}_f(t) = \sum_i V_{fi}(t) \exp(i\omega_{fi}t) a_i(t) \quad (7.219)$$

dove $V_{fi}(t) = \langle f|V(t)|i\rangle$ e $\hbar\omega_{fi} = E_f - E_i$. Integriamo con la condizione iniziale

$$a_f(0) = \delta_{fi} \quad (7.220)$$

Il sistema di equazioni (7.219) si può integrare numericamente noti i coefficienti $V_{fi}(t)$. In pratica, se τ non è grande, si può supporre che la variazione dei coefficienti sia piccola. Applicando il metodo delle approssimazioni successive ⁶ si ottiene *al primo ordine* per $f \neq i$

$$a_{fi}^{(1)} = \frac{1}{i\hbar} \int_0^t dt' V_{fi}(t') \exp(i\omega_{fi}t') \quad (7.221)$$

da cui segue che

$$w_{i \rightarrow f} = \frac{1}{\hbar^2} \left| \int_0^\tau dt V_{fi}(t) \exp(i\omega_{fi}t) \right|^2 \quad (7.222)$$

⁶È dato un sistema di equazioni differenziali al primo ordine, non autonomo $\dot{\mathbf{a}}(t) = \mathbf{f}[\mathbf{a}(t), t]$ con le condizioni iniziali $\mathbf{a}(0) = 0$. Allora l'approssimazione di ordine k , $\mathbf{a}^{(k)}(t)$, si ottiene dall'approssimazione di ordine $k - 1$ come

$$\mathbf{a}^{(k)}(t) = \int_0^t dt' \mathbf{f}[\mathbf{a}^{(k-1)}(t'), t']$$

nella tacita ipotesi che la successione così definita converga e converga alla soluzione del sistema per $k \rightarrow \infty$.

Appendice A

Funzioni speciali

Questo capitolo è dedicato al riassunto delle proprietà di alcune classi di funzioni che ritornano spesso nelle soluzioni di problemi della meccanica quantistica. Si tratta di un compendio molto breve, e senza pretese di completezza, che ha il solo scopo di fornire un 'formulario' allo studente.

A.1 Integrali ellittici

Consideriamo gli integrali indefiniti¹

$$I = \int dx R \left[x, \sqrt{P(x)} \right] \quad (\text{A.1})$$

dove R è una funzione razionale e $P(x)$ è un polinomio di terzo o quarto grado. Si dimostra che I è sempre riducibile ad una funzione elementare di

$$I_1 = \int dx \frac{1}{\sqrt{(1-x^2)(1-k^2x^2)}} \quad (\text{A.2})$$

$$I_2 = \int dx \frac{\sqrt{1-k^2x^2}}{\sqrt{1-x^2}} \quad (\text{A.3})$$

$$I_3 = \int dx \frac{1}{(1+n^2x^2)\sqrt{(1-x^2)(1-k^2x^2)}} \quad (\text{A.4})$$

che sono rispettivamente *integrali ellittici di prima, seconda e terza specie nella forma di Legendre*. La quantità k è detta *modulo* degli integrali. Altre espressioni utili sono le seguenti

$$\begin{aligned} F(\psi, k) &= \int_0^\psi d\alpha \frac{1}{\sqrt{(1-k^2 \sin^2 \alpha)}} \\ E(\psi, k) &= \int_0^\psi d\alpha \sqrt{(1-k^2 \sin^2 \alpha)} \\ D(\psi, k) &= \frac{F(\psi, k) - E(\psi, k)}{k^2} \\ \Pi(\psi, n, k) &= \int_0^\psi d\alpha \frac{1}{(1+n^2 \sin^2 \alpha)\sqrt{(1-k^2 \sin^2 \alpha)}} \end{aligned} \quad (\text{A.5})$$

Gli integrali per $\psi = \pi/2$ si dicono *integrali ellittici completi*.

¹Nel seguito, se non altrimenti specificato si usa x per una variabile reale e z per una variabile complessa.

A.2 Funzione gamma

Definiamo la funzione gamma sotto forma di integrale di Eulero:

$$\Gamma(z) = \int_0^{\infty} dt t^{z-1} e^{-t} \quad (\text{A.6})$$

Per $z = n + 1$ (n intero) la funzione gamma coincide con il fattoriale di n :

$$\Gamma(n + 1) = n! \quad (\text{A.7})$$

Vale la seguente espansione:

$$\frac{1}{\Gamma(z)} = z e^{\gamma z} \prod_{n=1}^{\infty} \left[\left(1 + \frac{z}{n} \right) e^{-z/n} \right] \quad (\text{A.8})$$

dove γ è la costante di Eulero:

$$\gamma = \lim_{m \rightarrow \infty} \left(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots + \frac{1}{m} - \ln m \right) = 0.5772156649 \dots \quad (\text{A.9})$$

Di interesse è l'espansione per $|z| \rightarrow \infty$ (formula di Stirling):

$$\Gamma(z) \sim e^{-z} z^{z-1/2} \sqrt{2\pi} \left(1 + \frac{1}{12z} + \frac{1}{288z^2} - \frac{139}{51840z^3} - \frac{571}{2488320z^4} + \dots \right) \quad (\text{A.10})$$

A.3 Funzione dell'errore

La funzione dell'errore è definita come:

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z dt e^{-t^2} \quad (\text{A.11})$$

La funzione complementare $\operatorname{erfc}(z)$ è $1 - \operatorname{erf}(z)$. Alcune relazioni utili sono le seguenti:

$$\frac{d^{n+1}}{dz^{n+1}} \operatorname{erf}(z) = (-1)^n \frac{2}{\sqrt{z}} \mathbf{H}_n(z) e^{-z^2} \quad (\text{A.12})$$

$$\operatorname{erf}(z) = \frac{2}{\sqrt{z}} \sum_{n=0}^{\infty} \frac{(-1)^n z^{2n+1}}{n!(2n+1)} \quad (\text{A.13})$$

$$\sqrt{\pi} z e^{z^2} \operatorname{erfc}(z) \sim 1 + \sum_{n=1}^{\infty} (-1)^n \frac{1 \cdot 3 \cdot \dots \cdot (2n-1)}{(2z^2)^n} \quad |z| \rightarrow \infty \quad (\text{A.14})$$

$$\int dx e^{-(ax^2+2bx+c)} = \frac{1}{2} \frac{\sqrt{\pi}}{a} e^{\frac{b^2-ac}{a}} \operatorname{erf} \left(\sqrt{a}x + \frac{b}{\sqrt{a}} \right) + \text{costante} \quad (\text{A.15})$$

A.4 Funzioni di Bessel

Sia ν un numero reale non-negativo. Le funzioni di Bessel del primo tipo $J_{\pm\nu}(z)$ e del secondo tipo $Y_{\pm\nu}(z)$ verificano la seguente equazione differenziale:

$$z^2 \frac{d^2 w}{dz^2} + z \frac{dw}{dz} + (z^2 - \nu^2) w = 0 \quad (\text{A.16})$$

Sono funzioni regolari, finite a $z = 0$; $J_{\pm\nu}(z)$ ha un estremo superiore finito per $\Re z \rightarrow \infty$. Le funzioni di Bessel modificate $I_{\pm\nu}(z)$ e $K_{\nu}(z)$ sono soluzioni dell'equazione differenziale

$$z^2 \frac{d^2 w}{dz^2} + z \frac{dw}{dz} - (z^2 + \nu^2) w = 0 \quad (\text{A.17})$$

$II_{\pm\nu}(z)$ tende ad un valore finito per $z \rightarrow 0$. Le funzioni di Bessel sono utili p.es. per la soluzione di equazioni differenziali in problemi con una simmetria cilindrica (vedi II Dispensa). Valgono le seguenti relazioni:

$$J_n(z) = \frac{1}{i^n \pi} \int_0^\pi d\theta e^{iz \cos \theta} \cos(n\theta) \quad (\text{A.18})$$

$$I_n(z) = \frac{1}{\pi} \int_0^\pi d\theta e^{z \cos \theta} \cos(n\theta) \quad (\text{A.19})$$

A.5 Polinomi ortogonali

È data una funzione reale non negativa $w(x)$ ed un intervallo $[a, b]$. Supponiamo che $\int_a^b dx x^n w(x)$ esista per n naturale $(0, 1, 2, \dots)$, e che sia positivo per $n = 0$. Allora esiste una successione di polinomi $p_0(x)$, $p_1(x)$, $p_2(x)$ etc. determinata univocamente dalle condizioni:

1. $p_n(x)$ è un polinomio di grado n ; il coefficiente q_n di x^n in $p_n(x)$ è positivo
2. i polinomi $p_n(x)$ sono ortogonali rispetto al prodotto scalare $\langle f|g \rangle = \int_a^b dx w(x) f(x) g(x)$:

$$\langle p_m|p_n \rangle = \delta_{mn} k_m \quad (\text{A.20})$$

dove k_m è la norma $k_m = \langle p_m|p_m \rangle$. La successione costituisce un set di polinomi ortogonali su $[a, b]$ rispetto alla funzione peso $w(x)$. Valgono le seguenti relazioni generali

1. *Relazione di Christoffel-Darboux*

$$\sum_{k=0}^n p_k(x) p_k(y) = \frac{q_n}{q_{n+1}} \frac{p_{n+1}(x) p_n(y) - p_n(x) p_{n+1}(y)}{x - y} \quad (\text{A.21})$$

2. *Relazione ricorsiva*

$$p_{n+1}(x) = (a_n x + b_n) p_n(x) - c_n p_{n-1}(x) \quad n \geq 1 \quad (\text{A.22})$$

A.6 Polinomi di Legendre

I polinomi di Legendre $P_n(x)$ sono definiti in $[-1, 1]$. Verificano l'equazione differenziale:

$$(1 - x^2) \frac{d^2 P_n(x)}{dx^2} - 2x \frac{dP_n(x)}{dx} + n(n+1) P_n(x) = 0 \quad (\text{A.23})$$

La funzione peso è $w = 1$, la norma è $k_n = 2/2n + 1$. La relazione ricorsiva è:

$$(n+1) P_{n+1}(x) = (2n+1)x P_n(x) - n P_{n-1}(x) \quad (\text{A.24})$$

dove $P_0(x) = 1$, $P_1(x) = x$.

A.7 Polinomi di Hermite

I polinomi di Hermite $H_n(x)$ sono definiti in $[-\infty, \infty]$. Verificano l'equazione differenziale:

$$\frac{d^2 H_n(x)}{dx^2} - 2x \frac{dH_n(x)}{dx} + 2n H_n(x) = 0 \quad (\text{A.25})$$

La funzione peso è $w = e^{-x^2}$, la norma è $k_n = \sqrt{\pi} 2^n n!$. La relazione ricorsiva è:

$$H_{n+1}(x) = 2nx H_n(x) - 2n H_{n-1}(x) \quad (\text{A.26})$$

dove $H_0(x) = 1$, $H_1(x) = 2x$.

A.8 Polinomi di Laguerre

I polinomi di Laguerre $L_n(x)$ sono definiti in $[0, \infty]$. Verificano l'equazione differenziale:

$$x \frac{d^2 L_n(x)}{dx^2} + (1-x) \frac{dL_n(x)}{dx} + nL_n(x) = 0 \quad (\text{A.27})$$

La funzione peso è $w = e^{-x}$, la norma è 1. La relazione ricorsiva è:

$$(n+1)L_{n+1}(x) = (2n+1-x)L_n(x) - nL_{n-1}(x) \quad (\text{A.28})$$

dove $L_0(x) = 1$, $L_1(x) = 1 - x$.

Appendice B

L'atomo idrogenoide

L'hamiltoniano in unità atomiche è:

$$\mathcal{H} = -\frac{1}{2m_Z} \hat{\nabla}_R^2 - \frac{1}{2} \hat{\nabla}_e^2 - \frac{Z}{|\vec{R} - \vec{r}_e|} \quad (\text{B.1})$$

dove indichiamo con \vec{R} e \vec{r}_e la posizione del nucleo e dell'elettrone, e con $\hat{\nabla}_R^2$ e $\hat{\nabla}_e^2$ i relativi gradienti. Possiamo procedere come nell'Esercizio 1 ed introdurre le variabili di spostamento relativo dell'elettrone e di centro di massa: $\vec{r} = \vec{r}_e - \vec{R}$ e $\vec{r}_c = \frac{\vec{r}_e + m_Z \vec{R}}{M}$, dove $M = 1 + m_Z$ è la massa totale del sistema. Rispetto alle nuove coordinate l'hamiltoniano risulta definito come la somma dell'hamiltoniano relativo al moto libero del baricentro con massa M e dell'hamiltoniano relativo al moto dell'elettrone rispetto ad un nucleo a riposo, ed avente massa ridotta $\mu = m_Z/M$ in unità atomiche. In unità atomiche l'hamiltoniano separato risulta essere:

$$\mathcal{H} = -\frac{1}{2\mu} \hat{\nabla}^2 - \frac{1}{2M} \hat{\nabla}_c^2 - \frac{Z}{r} \quad (\text{B.2})$$

dove ora r è la distanza tra il baricentro e l'elettrone. La parte relativa al baricentro determina il moto traslazionale libero dell'atomo, e non sarà considerata oltre. La parte elettronica può essere scritta come l'hamiltoniano di una particella di carica -1 che si muova nel campo di potenziale elettrostatico generato da un nucleo immobile di carica Z , e avente massa μ :

$$\mathcal{H} = -\frac{Z}{r} - \frac{1}{2\mu} \hat{\nabla}^2 \quad (\text{B.3})$$

Nel seguito considereremo $\mu \approx 1$: per $Z = 1$, assumendo per m_Z la massa di un protone avremo $\mu = 0.9994$. Il potenziale è a simmetria radiale. Possiamo perciò procedere risolvendo il sistema nel seguente modo i) introduzione di coordinate sferiche (r, θ, ϕ) ; ii) fattorizzazione della funzione d'onda in una parte angolare dipendente da θ e ϕ e in una parte radiale dipendente da r ; iii) imposizione di condizioni al contorno per il calcolo delle autoenergie. La parte angolare dell'autofunzione è data dalle funzioni armoniche sferiche (ciò è vero per qualunque potenziale a simmetria sferica)

$$\Psi_E(r, \theta, \phi) = R(r)Y_{lm_l}(\theta, \phi) \quad (\text{B.4})$$

Sostituendo nell'Eq. di Schrödinger si trova l'equazione per la parte radiale:

$$\frac{1}{r^2} \frac{d}{dr} r^2 \frac{dR}{dr} + 2 \left[E + \frac{Z}{r} - \frac{l(l+1)}{2r^2} \right] R = 0 \quad (\text{B.5})$$

Lo spettro degli autovalori sarà discreto per $E < 0$, e continuo per $E > 0$ (il potenziale coulombiano tende a zero per $r \rightarrow \infty$ e a $-\infty$ per $r \rightarrow 0$). L'equazione differenziale in R , con le condizioni al contorno R finita per $r \rightarrow 0$ e $r \rightarrow \infty$ è riconoscibile come l'equazione che ha per soluzione delle funzioni note: le *funzioni associate di Laguerre* (moltiplicate per l'esponenziale di $-r$). In definitiva gli autostati e le autoenergie (in Hartree) dell'atomo di idrogeno sono:

$$\Psi_{n,l,m_l}(r, \theta, \phi) = R_{nl}(r)Y_{lm_l}(\theta, \phi) \quad (\text{B.6})$$

$$R_{nl} = \sqrt{\frac{(n-l-1)!Z}{n^2[(n+l)!]^3}} \rho^{l+1} L_{n+l}^{2l+1}(\rho) \exp(-\rho/2) \quad (\text{B.7})$$

$$E_{n,l,m_l} = -\frac{Z^2}{2n^2} \quad (\text{B.8})$$

dove $\rho = 2Zr/na$; $L_n^\alpha(\rho)$ è un polinomio di Laguerre di grado n ed ordine α . Gli autostati discreti sono negativi e dipendono solo da $1/n^2$.

Appendice C

Problema generalizzato agli autovalori

Consideriamo l'equazione di Schrödinger indipendente dal tempo per le autofunzioni dell'operatore di Hamilton relativo ad un sistema definito rispetto alle coordinate \mathbf{q} . Una (auto)-funzione $\Psi(\mathbf{q})$ è un elemento dello spazio di Hilbert delle funzioni in \mathbf{q} , e sarà indicata nel seguito con il ket $|\Psi\rangle$

$$\mathcal{H}|\Psi\rangle = E|\Psi\rangle \quad (\text{C.1})$$

Introduciamo ora un insieme di funzioni linearmente indipendenti che costituiscano una base, non necessariamente ortonormale, nello spazio di Hilbert $|v_1\rangle, |v_2\rangle, \dots$; possiamo perciò esprimere $|\Psi\rangle$ come una combinazione lineare degli elementi della base:

$$|\Psi\rangle = \sum_j c_j |v_j\rangle \quad (\text{C.2})$$

Sostituendo nella (C.1) e premoltiplicando per il generico bra $\langle v_i|$:

$$\sum_j \langle v_i | \mathcal{H} | v_j \rangle c_j = E \sum_j \langle v_i | v_j \rangle c_j \quad (\text{C.3})$$

che in notazione matriciale si può scrivere come

$$\mathbf{H}\mathbf{c} = E\mathbf{S}\mathbf{c} \quad (\text{C.4})$$

dove $H_{ij} = \langle v_i | \mathcal{H} | v_j \rangle$, $S_{ij} = \langle v_i | v_j \rangle$. Se la base è ortonormale, la matrice di overlap \mathbf{S} coincide con la matrice unità, $S_{ij} = \delta_{ij}$, e dunque

$$\mathbf{H}\mathbf{c} = E\mathbf{c} \quad (\text{C.5})$$

In pratica, se la dimensione dello spazio hilbertiano considerato è infinita, il set di funzioni di base utilizzato viene troncato ad un numero finito N , introducendo un'approssimazione che dipende da N . Soluzioni non-triviali del sistema N -dimensionale (C.4) si ottengono dalle radici dell'equazione secolare:

$$\det(\mathbf{H} - E\mathbf{S}) = 0 \quad (\text{C.6})$$

Si ottengono N autovalori E_i , $i = 1, \dots, N$ cui corrispondono N autovettori \mathbf{c}_i . Se si definisce la matrice \mathbf{C} le cui colonne sono autovettori, $C_{ij} = (\mathbf{c}_j)_i$ la matrice diagonale i cui elementi sono diagonali sono gli autovalori, $E_{ij} = \delta_{ij}E_j$, risulta che:

$$\mathbf{H}\mathbf{C} = \mathbf{S}\mathbf{C}\mathbf{E} \quad (\text{C.7})$$

Appendice D

Studio MD di un fluido atomico

Descriviamo in questa appendice un'applicazione dei metodi della dinamica molecolare ad un sistema specifico. Partiremo dalla definizione del sistema di interesse; definiremo quindi le metodiche utilizzate per il calcolo; illustreremo un programma in linguaggio Fortran dedicato al calcolo stesso; ed infine commenteremo alcuni risultati. Il sistema considerato è un fluido atomico, formato cioè da particelle puntiformi interagenti mediante un potenziale a simmetria sferica.

D.1 Il sistema

Il primo sistema da noi considerato è costituito da N particelle puntiformi - atomi - interagenti mediante un potenziale a coppie Lennard-Jones. La semplicità di questo sistema è utile per introdurre le metodologie fondamentali della Dinamica Molecolare: le particelle costituenti sono dotate dei soli *gradi di libertà traslazionali*, e quindi non è necessaria un'implementazione complicata delle equazioni del moto. Il sistema considerato si può confrontare con un fluido costituito da atomi di gas rari, per esempio l'argon. In effetti, parametri ϵ e σ di un potenziale di Lennard-Jones per la riproduzione realistica di proprietà termodinamiche di fluidi atomici di Ar ed altri gas rari sono noti. Per fissare le idee, considereremo il caso di un insieme di N

	ϵ/k_B (K)	σ (nm)
He	10.2	0.228
Ne	47.0	0.272
Ar	119.8	0.341
Kr	164.0	0.383

Tabella D.1: Coefficienti LJ per i gas rari

atomi di Ar, cioè con $\epsilon/k_B = 119.8$ K e $\sigma = 0.341$ nm, contenuti in una scatola cubica, con condizioni al contorno periodiche, a densità e prefissata. Il sistema è isolato, cioè è trattato nell'ambito di una descrizione termodinamica microcanonica (NVE).

Come definiamo la configurazione iniziale del sistema? Solitamente si assegna una configurazione ordinata, basata su un qualche reticolo cristallino di facile costruzione. Assumeremo perciò nel seguito una configurazione iniziale basata su un reticolo cristallino cubico a facce centrate. (*face centered cubic lattice*, fcc).

Costruiamo cioè il sistema iniziale come un insieme di N_c celle elementari fcc per lato. Una cella elementare fcc contiene 4 unità molecolari ed il numero di celle in un volume cubico è N_c^3 , quindi il numero di particelle totali è $N = 4N_c^3$. In Fig. (D.1) è illustrata una cella elementare fcc (a) ed un sistema formato da 8 celle elementari ($N_c = 2$) per un totale di 32 atomi. Lo stato iniziale del sistema può essere definito in termini di una temperatura iniziale T_0 , che vedremo non corrispondere esattamente alla temperatura di equilibrio, e di una densità ρ . La temperatura iniziale è il parametro che definisce l'energia cinetica totale iniziale del sistema. In pratica le velocità iniziali delle particelle del sistema \vec{v}_i ($i = 1, \dots, N$) sono assegnate secondo una distribuzione gaussiana, secondo la legge di Maxwell, in modo tale che l'energia cinetica media di un atomo sia data da $3k_B T_0/2$. Nel corso della successiva evoluzione temporale dallo stato assegnato inizialmente, il sistema si porta in uno stato di equilibrio, caratterizzato

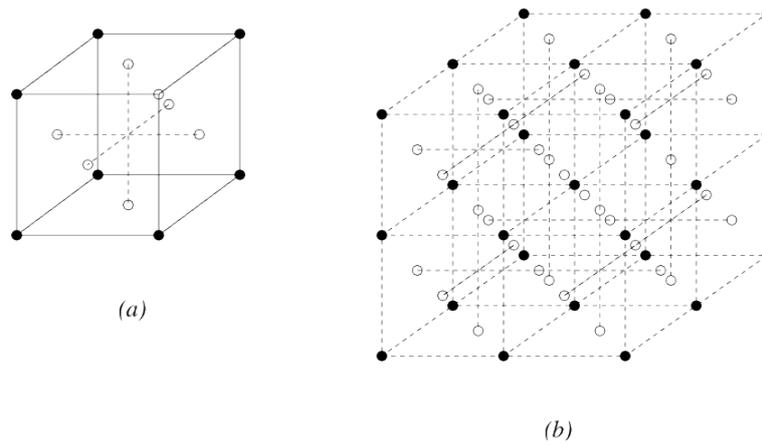


Figura D.1: Un reticolo cristallino fcc (a) ed un sistema di 4 celle elementari fcc (b)

da una temperatura di equilibrio T . Le condizioni al contorno sono, solitamente, di tipo periodico: ogni

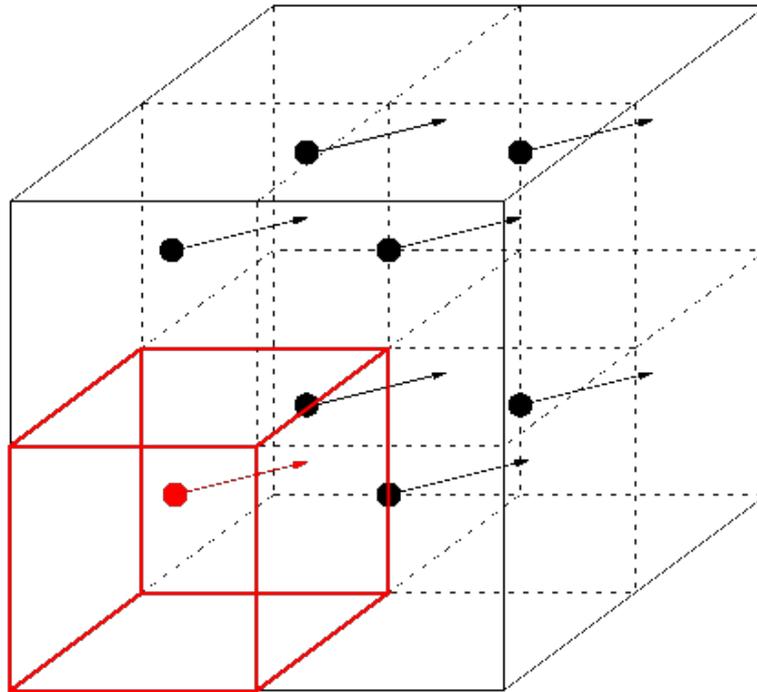


Figura D.2: Condizioni al contorno periodiche

particella risente perciò dell'interazione con le particelle entro il volume cubico, e di tutte le "repliche" o immagini negli infiniti volumi cubici circostanti. In pratica, dato che il potenziale di interazione decade rapidamente con la distanza, è sufficiente considerare una distanza massima o raggio di cut-off, r_c , che tipicamente è dell'ordine di 3, 4 volte σ .

D.2 Metodo di calcolo

Adottiamo come metodo di risoluzione numerica uno schema di discretizzazione di tipo leap-frog, in una versione semplificata per sole coordinate traslazionali. Ad ogni nuovo passo il vettore delle velocità del sistema è aggiornato secondo lo schema

$$\vec{v}(t + \delta t/2) = \vec{v}(t - \delta/2) + \delta t \vec{a}(t) \quad (\text{D.1})$$

dove il vettore delle accelerazioni è calcolato al tempo t usando la seconda legge di Newton e dunque calcolando le forze agenti sulle particelle. Il vettore delle posizioni è aggiornato al tempo $t + \delta t$ secondo lo

schema

$$\vec{r}(t + \delta t) = \vec{r}(t) + \delta t \vec{v}(t + \delta t/2) \quad (\text{D.2})$$

È necessario specificare dunque come ulteriore parametro di input il passo temporale dello schema di discretizzazione, δt . Una sessione di calcolo tipica è di solito completamente determinata specificando anche un lasso temporale t_{eq} , che specifichi il tempo assegnato al sistema per raggiungere lo stato di equilibrio termodinamico; ed un tempo di calcolo vero e proprio degli osservabili di interesse t_{calc} , che segue il tempo t_{eq} e che viene effettivamente considerato per il calcolo di funzioni di stato, funzioni di correlazione etc. Ri-capitolando, il sistema è definito dai seguenti parametri: N_c , n. celle elementari del reticolo fcc che specifica la configurazione iniziale; ϵ e σ che definiscono il potenziale di interazione tra gli atomi; T_0 , la temperatura iniziale, corrispondente ad uno stato di non equilibrio; ρ , la densità (intesa come numero di particelle su volume); r_c , raggio di cut-off del potenziale; δt , passo dello schema di discretizzazione delle equazioni di Newton; t_{eq} , tempo di equilibrizzazione e t_{calc} , tempo di calcolo. In pratica, i codici esistenti implementano le equazioni del moto utilizzando, per motivi di convenienza numerica e per evitare calcoli ridondanti, coordinate riscalate. È sufficiente definire il sistema, se si usano grandezze scalate, con il seguente insieme di coordinate di input: $N_c, T_0^*, \rho^*, r_c^*, \delta t^*, t_{\text{eq}}^*$ e t_{calc}^* .

Esercizio 7. Si vuole simulare un sistema preparato inizialmente secondo un reticolo fcc, formato da circa 100 atomi di Ar, con una temperatura iniziale di 100 K, una densità (massa su volume) di $\rho_m = 1.5 \times 10^3 \text{ Kg m}^{-3}$, con un passo di integrazione di 10 fs; fissiamo il tempo di equilibrizzazione in 30 ps ed il tempo di calcolo in 10 ps. I parametri ϵ e σ sono rispettivamente $1.653 \times 10^{-21} \text{ J}$ e $3.41 \times 10^{-10} \text{ m}$; assumiamo un raggio di cut-off pari a 3.5σ . Il sistema più vicino a quello desiderato corrisponde a $N_c = 3$, vale a dire 27 celle elementari, ovvero 108 atomi. La temperatura riscalata è $T_0^* = k_B T_0 / \epsilon = 0.834$. La densità riscalata è $\rho^* = \rho \sigma^3 / m_{\text{Ar}} = 0.815$; il raggio di cutoff è semplicemente 3.5; i parametri temporali sono $\delta t^* = 0.0046$, $t_{\text{eq}}^* = 13.8$, $t_{\text{calc}}^* = 4.62$. Per esercizio, calcolate la lunghezza riscalata del lato del volume cubico.

D.3 Codice

Può essere utile commentare un programma per la simulazione di un sistema atomico in modo dettagliato dal testo di Allen e Tildesley. Il programma principale è formato, come al solito da una sezione di dichiarazione delle variabili e di input dei parametri. Segue la definizione della configurazione iniziale (routines `fcc` e `comvel`). Il programma opera utilizzando un set di coordinate riferite ad una lunghezza unitaria del lato del volume cubico, quindi opera una conversione a partire dalle coordinate riscalate di input riferite a σ .

```
*****
** LOW-STORAGE MD PROGRAMS USING THE LEAPFROG VERLET ALGORITHM      **
*****
C *****
C ** TWO SEPARATE PARTS: FORTRAN AND BASIC VERSIONS.                **
C *****
C *****
C ** FICHE F.3 - PART A                                               **
C ** FORTRAN PROGRAM USING THE LEAPFROG ALGORITHM.                  **
C *****
C PROGRAM FROGGY
C include 'block.inc'
C *****
C ** FORTRAN PROGRAM TO CONDUCT MOLECULAR DYNAMICS OF ATOMS.        **
C **                                                                    **
C ** A SPECIAL LOW-STORAGE FORM OF THE LEAPFROG VERLET ALGORITHM     **
C ** IS USED AND WE TAKE THE LENNARD-JONES POTENTIAL AS AN EXAMPLE. **
C **                                                                    **
C ** REFERENCE:                                                         **
C **                                                                    **
C ** FINCHAM AND HEYES, CCP5 QUARTERLY, 6, 4, 1982.                 **
```

```

C      **
C      ** ROUTINES REFERENCED:
C      **
C      ** SUBROUTINE READCN ( CNFILE )
C      **   READS IN CONFIGURATION
C      ** SUBROUTINE FORCE ( DT, SIGMA, RCUT, NEWV, NEWVC, NEWW )
C      **   CALCULATES THE ACCELERATIONS AND ADVANCES THE VELOCITIES
C      **   FROM T - DT/2 TO T + DT/2. ALSO CALCULATES POTENTIAL
C      **   ENERGY AND VIRIAL AT TIMESTEP T.
C      ** SUBROUTINE MOVE ( DT )
C      **   ADVANCES POSITIONS FROM T TO T + DT.
C      ** SUBROUTINE KINET ( NEWK )
C      **   CALCULATES KINETIC ENERGY.
C      ** SUBROUTINE WRITCN ( CNFILE )
C      **   WRITES OUT CONFIGURATION
C      **
C      ** PRINCIPAL VARIABLES:
C      **
C      ** INTEGER N           NUMBER OF ATOMS
C      ** REAL   DT           TIMESTEP
C      ** REAL   RX(N),RY(N),RZ(N) ATOMIC POSITIONS
C      ** REAL   VX(N),VY(N),VZ(N) ATOMIC VELOCITIES
C      ** REAL   ACV,ACK ETC.   AVERAGE VALUE ACCUMULATORS
C      ** REAL   AVV,AVK ETC.   AVERAGE VALUES
C      ** REAL   ACVSQ, ACKSQ ETC. AVERAGE SQUARED VALUE ACCUMULATORS
C      ** REAL   FLV,FLK ETC.   FLUCTUATION AVERAGES
C      **
C      ** USAGE:
C      **
C      ** THE LEAPFROG ALGORITHM IS OF THE FORM
C      **  $VX(T + DT/2) = VX(T - DT/2) + DT * AX(T)$  (SIMILARLY FOR Y,Z)
C      **  $RX(T + DT) = RX(T) + DT * VX(T + DT/2)$  (SIMILARLY FOR Y,Z)
C      ** TO SAVE STORAGE IN THIS PROGRAM WE ACCUMULATE VALUES AX(T)
C      ** DIRECTLY ONTO THE VELOCITIES VX(T - DT/2) IN THE FORCE LOOP.
C      ** THIS MEANS THAT AN APPROXIMATION MUST BE USED FOR THE KINETIC
C      ** ENERGY AT EACH TIME STEP:
C      **  $K = ( OLDK + NEWK ) / 2 + ( OLDV - 2 * V + NEWV ) / 8$ 
C      ** WHERE  $K = K( T )$ ,  $V = V( T )$ 
C      **  $OLDK = K( T - DT/2 )$ ,  $OLDV = V( T - DT )$ 
C      **  $NEWK = K( T + DT/2 )$ ,  $NEWV = V( T + DT )$ 
C      ** AT THE START OF A STEP THE FOLLOWING VARIABLES ARE STORED:
C      ** R : R(STEP)           V : V(STEP+1/2)
C      ** OLDV : V(STEP-2)     V : V(STEP-1)     NEWV : V(STEP)
C      ** OLDK : K(STEP-1/2)   NEWK : K(STEP+1/2)
C      ** THIS PROGRAM USES UNIT 10 FOR CONFIGURATION INPUT AND OUTPUT
C      **
C      ** UNITS:
C      **
C      ** THE PROGRAM USES LENNARD-JONES REDUCED UNITS FOR USER INPUT
C      ** AND OUTPUT BUT CONDUCTS SIMULATION IN A BOX OF UNIT LENGTH.
C      ** SUMMARY FOR BOX LENGTH L, ATOMIC MASS M, AND LENNARD-JONES
C      ** POTENTIAL PARAMETERS SIGMA AND EPSILON:
C      **
C      ** OUR PROGRAM           LENNARD-JONES SYSTEM
C      ** LENGTH                 L           SIGMA
C      ** MASS                   M           M
C      ** ENERGY                EPSILON     EPSILON
C      ** TIME                    $SQRT(M*L**2/EPSILON)$   $SQRT(M*SIGMA**2/EPSILON)$ 
C      ** VELOCITY                $SQRT(EPSILON/M)$     $SQRT(EPSILON/M)$ 
C      ** PRESSURE                 $EPSILON/L**3$       $EPSILON/SIGMA**3$ 
C      *****
C      INTEGER   STEP, NSTEP, IPRINT

```

```

REAL      DENS, NORM, RCUT, DT, SIGMA, TEMP1
REAL      V, K, E, W
REAL      OLDK, NEWK, OLDV, NEWV, NEWW
REAL      OLDVC, NEWVC, VC, KC, EC
REAL      VN, KN, EN, ECN, PRES, TEMP
REAL      ACV, ACK, ACE, ACEC, ACP, ACT
REAL      AVV, AVK, AVE, AVEC, AVP, AVT
REAL      ACVSQ, ACKSQ, ACESQ, ACECSQ, ACPSQ, ACTSQ
REAL      FLV, FLK, FLE, FLEC, FLP, FLT
REAL      SR3, SR9, VLRC, WLRC, PI, SIGCUB
CHARACTER CNFILE*30, TITLE*80
REAL      FREE

PARAMETER ( FREE = 3.0 )
PARAMETER ( PI = 3.1415927 )
C *****
C ** READ IN INITIAL PARAMETERS **

WRITE(*,'(1H1, ' ' **** PROGRAM FROGGY **** ' '))
WRITE(*,'(//, ' ' MOLECULAR DYNAMICS OF LENNARD-JONES ATOMS ' '))
WRITE(*,'( ' ' LEAPFROG ALGORITHM WITH MINIMAL STORAGE ' '))
WRITE(*,'(' ' ENTER RUN TITLE ' '))
READ (*,'(A)') TITLE
WRITE(*,'(' ' ENTER NUMBER OF STEPS ' '))
READ (*,*) NSTEP
WRITE(*,'(' ' ENTER NUMBER OF STEPS BETWEEN OUTPUT LINES ' '))
READ (*,*) IPRINT
WRITE(*,'(' ' ENTER CONFIGURATION FILENAME ' '))
READ (*,'(A)') CNFILE
WRITE(*,'(' ' ENTER THE FOLLOWING IN LENNARD-JONES UNITS ' '))
WRITE(*,'(' ' ENTER DENSITY ' '))
READ (*,*) DENS
WRITE(*,'(' ' ENTER INIT. TEMPERATURE ' '))
READ (*,*) TEMP1
WRITE(*,'(' ' ENTER POTENTIAL CUTOFF DISTANCE ' '))
READ (*,*) RCUT
WRITE(*,'(' ' ENTER TIME STEP ' '))
READ (*,*) DT
WRITE(*,'(//1X,A)') TITLE
WRITE(*,'(' ' NUMBER OF ATOMS = ',I10 )') N
WRITE(*,'(' ' NUMBER OF STEPS = ',I10 )') NSTEP
WRITE(*,'(' ' OUTPUT FREQUENCY = ',I10 )') IPRINT
WRITE(*,'(' ' POTENTIAL CUTOFF = ',F10.4)') RCUT
WRITE(*,'(' ' DENSITY = ',F10.4)') DENS
WRITE(*,'(' ' TIME STEP = ',F10.6)') DT

C ** READ CONFIGURATION INTO COMMON / BLOCK1 / VARIABLES **
CALL FCC
CALL COMVEL( TEMP1 )
C ** CONVERT INPUT DATA TO PROGRAM UNITS **
SIGMA = ( DENS / REAL ( N ) ) ** ( 1.0 / 3.0 )
RCUT = RCUT * SIGMA
DT = DT * SIGMA
DENS = DENS / ( SIGMA ** 3 )

C ** CALCULATE LONG-RANGE CORRECTIONS **
C ** NOTE: SPECIFIC TO LENNARD-JONES **

SR3 = ( SIGMA / RCUT ) ** 3
SR9 = SR3 ** 3

```

```

SIGCUB = SIGMA ** 3
VLRC = ( 8.0 / 9.0 ) * PI * DENS * SIGCUB * REAL ( N )
:      * ( SR9 - 3.0 * SR3 )
WLRC = ( 16.0 / 9.0 ) * PI * DENS * SIGCUB * REAL ( N )
:      * ( 2.0 * SR9 - 3.0 * SR3 )

```

Segue il calcolo vero e proprio, effettuato come un ciclo su numero di passi temporali totali *nstep*; per semplicità non sono inclusi in questo codice passi preliminari di equilibratura. Ad ogni passo del ciclo vengono chiamate nell'ordine le routine *move* (che aggiorna le posizioni, date le velocità calcolate nel passo precedente), *force* che calcola le forze e aggiorna le velocità secondo l'algoritmo di *leap-frog* e *kinet* che calcola l'osservabile richiesto (in questo caso l'energia cinetica del sistema).

```
C  ** ZERO ACCUMULATORS **
```

```

ACV  = 0.0
ACK  = 0.0
ACE  = 0.0
ACEC = 0.0
ACP  = 0.0
ACT  = 0.0

ACVSQ = 0.0
ACKSQ = 0.0
ACESQ = 0.0
ACECSQ = 0.0
ACPSQ = 0.0
ACTSQ = 0.0

FLV  = 0.0
FLK  = 0.0
FLE  = 0.0
FLEC = 0.0
FLP  = 0.0
FLT  = 0.0

```

```
C  ** CALCULATE INITIAL VALUES **
```

```

CALL FORCE ( -DT, SIGMA, RCUT, NEWV, NEWVC, NEWW )
CALL MOVE ( -DT )
CALL FORCE ( -DT, SIGMA, RCUT, V, VC, W )
CALL FORCE ( DT, SIGMA, RCUT, V, VC, W )
CALL KINET ( OLDK )
CALL MOVE ( DT )
CALL FORCE ( DT, SIGMA, RCUT, NEWV, NEWVC, NEWW )
CALL KINET ( NEWK )

```

```
C  ** INCLUDE LONG-RANGE CORRECTIONS **
```

```

V    = V + VLRC
W    = W + WLRC
NEWV = NEWV + VLRC
NEWW = NEWW + WLRC

```

```
IF ( IPRINT .LE. 0 ) IPRINT = NSTEP + 1
```

```
WRITE(*, '(//1X, ''**** START OF DYNAMICS ****'')')
WRITE(*, 10001)
```

```

C *****
C ** MAIN LOOP BEGINS **
C *****

```

```

DO 1000 STEP = 1, NSTEP

C    ** IMPLEMENT ALGORITHM **
      CALL MOVE ( DT )
      OLDV = V
      V     = NEWV
      OLDVC = VC
      VC    = NEWVC
      W     = NEWW
      CALL FORCE ( DT, SIGMA, RCUT, NEWV, NEWVC, NEWW )
C    ** INCLUDE LONG-RANGE CORRECTIONS **
      NEWV = NEWV + VLRC
      NEWW = NEWW + WLRC

C    ** CALCULATE KINETIC ENERGY AT CURRENT STEP **
      K = ( NEWK + OLDK ) / 2.0
:      + ( NEWV - 2.0 * V + OLDV ) / 8.0
      KC = ( NEWK + OLDK ) / 2.0
:      + ( NEWVC - 2.0 * VC + OLDVC ) / 8.0
      OLDK = NEWK
      CALL KINET ( NEWK )
C    ** CALCULATE INSTANTANEOUS VALUES **
      E = K + V
      EC = KC + VC
      VN = V / REAL ( N )
      KN = K / REAL ( N )
      EN = E / REAL ( N )
      ECN = EC / REAL ( N )
      TEMP = 2.0 * KN / FREE
      PRES = DENS * TEMP + W
C    ** CONVERT TO LENNARD-JONES UNITS **
      PRES = PRES * SIGMA ** 3
C    ** INCREMENT ACCUMULATORS **
      ACE = ACE + EN
      ACEC = ACEC + ECN
      ACK = ACK + KN
      ACV = ACV + VN
      ACP = ACP + PRES
      ACESQ = ACESQ + EN ** 2
      ACECSQ = ACECSQ + ECN ** 2
      ACKSQ = ACKSQ + KN ** 2
      ACVSQ = ACVSQ + VN ** 2
      ACPSQ = ACPSQ + PRES ** 2
C    ** OPTIONALLY PRINT INFORMATION **
      IF ( MOD( STEP, IPRINT ) .EQ. 0 ) THEN
        WRITE(*, '(1X,I8,6(2X,F10.4))')
:      STEP, EN, ECN, KN, VN, PRES, TEMP
      ENDIF

1000  CONTINUE

C    *****
C    ** MAIN LOOP ENDS **
C    *****

      WRITE(*, '(/1X, '***** END OF DYNAMICS **** ' //)')

C    ** OUTPUT RUN AVERAGES **

      NORM = REAL ( NSTEP )
      AVE = ACE / NORM

```

```

AVEC = ACEC / NORM
AVK  = ACK  / NORM
AVV  = ACV  / NORM
AVP  = ACP  / NORM
ACESQ = ( ACESQ / NORM ) - AVE ** 2
ACECSQ = ( ACECSQ / NORM ) - AVEC ** 2
ACKSQ = ( ACKSQ / NORM ) - AVK ** 2
ACVSQ = ( ACVSQ / NORM ) - AVV ** 2
ACPSQ = ( ACPSQ / NORM ) - AVP ** 2
IF ( ACESQ .GT. 0.0 ) FLE = SQRT ( ACESQ )
IF ( ACECSQ .GT. 0.0 ) FLEC = SQRT ( ACECSQ )
IF ( ACKSQ .GT. 0.0 ) FLK = SQRT ( ACKSQ )
IF ( ACVSQ .GT. 0.0 ) FLV = SQRT ( ACVSQ )
IF ( ACPSQ .GT. 0.0 ) FLP = SQRT ( ACPSQ )
AVT = AVK * 2.0 / FREE
FLT = FLK * 2.0 / FREE

WRITE(*,'('' AVERAGES'',6(2X,F10.5))')
:      AVE, AVEC, AVK, AVV, AVP, AVT
WRITE(*,'('' FLUCTS  ''',6(2X,F10.5))')
:      FLE, FLEC, FLK, FLV, FLP, FLT

C  ** WRITE OUT FINAL CONFIGURATION **
CALL WRITCN ( CNFILE )

STOP

10001  FORMAT(/1X,'Timestep ..ENERGY.. CUTENERGY.'
:      ' ..KINETIC. ..POTENT..',
:      ' ..PRESSURE. ..TEMPER..')

END

```

Alcune routine sono riportate nel seguito. In particolare, si deve prestare attenzione al doppio ciclo presente nella routine `force` che costituisce la sezione del codice che utilizza la maggior parte del tempo di CPU richiesto dal programma (fino al 90%).

```

SUBROUTINE FORCE ( DT, SIGMA, RCUT, V, VC, W )

include 'block.inc'

C *****
C ** ROUTINE TO COMPUTE FORCES WITH CUTOFF & MINIMUM IMAGING. **
C ** ** **
C ** THE ROUTINE ACTUALLY RETURNS TWO DIFFERENT POTENTIAL ENERGIES.**
C ** V IS CALCULATED USING THE UNSHIFTED LENNARD-JONES POTENTIAL. **
C ** WHEN LONG-RANGE TAIL CORRECTIONS ARE ADDED, THIS MAY BE USED **
C ** TO CALCULATE THERMODYNAMIC INTERNAL ENERGY ETC. **
C ** VC IS CALCULATED USING THE SHIFTED LENNARD-JONES POTENTIAL **
C ** WITH NO DISCONTINUITY AT THE CUTOFF. THIS MAY BE USED TO **
C ** CHECK ENERGY CONSERVATION. **
C ** PROGRAM UNITS MAKE EPSILON = MASS = 1, BUT SIGMA IS NOT UNITY **
C ** SINCE WE TAKE A BOX OF UNIT LENGTH. **
C ** THE ROUTINE IS COMBINED WITH THE LEAPFROG ALGORITHM FOR **
C ** ADVANCING VELOCITIES, I.E. FORCES ARE ACCUMULATED DIRECTLY **
C ** ONTO VELOCITIES. **
C *****

REAL      V, RCUT, DT, SIGMA, W, VC

INTEGER   I, J, NCUT
REAL      RCUTSQ, VCUT, SIGSQ
REAL      RXI, RYI, RZI, VXI, VYI, VZI

```

```

REAL      RXIJ, RYIJ, RZIJ, RIJSQ
REAL      SR2, SR6, SR12
REAL      VIJ, WIJ, VELIJ, DVX, DVY, DVZ

```

```
C *****
```

```
C ** CALCULATE SQUARED DISTANCES FOR INNER LOOP **
```

```

SIGSQ = SIGMA ** 2
RCUTSQ = RCUT ** 2

```

```
C ** CONDUCT OUTER LOOP OVER ATOMS **
```

```

V = 0.0
W = 0.0
NCUT = 0

```

```

DO 1000 I = 1, N - 1
  RXI = RX(I)
  RYI = RY(I)
  RZI = RZ(I)
  VXI = VX(I)
  VYI = VY(I)
  VZI = VZ(I)

```

```
C ** CONDUCT INNER LOOP OVER ATOMS **
```

```

DO 900 J = I + 1, N
  RXIJ = RXI - RX(J)
  RXIJ = RXIJ - ANINT ( RXIJ )
  IF ( ABS ( RXIJ ) .LT. RCUT ) THEN
    RYIJ = RYI - RY(J)
    RYIJ = RYIJ - ANINT ( RYIJ )
    RIJSQ = RXIJ ** 2 + RYIJ ** 2
    IF ( RIJSQ .LT. RCUTSQ ) THEN
      RZIJ = RZI - RZ(J)
      RZIJ = RZIJ - ANINT ( RZIJ )
      RIJSQ = RIJSQ + RZIJ ** 2
      IF ( RIJSQ .LT. RCUTSQ ) THEN
        SR2 = SIGSQ / RIJSQ
        SR6 = SR2 ** 3
        SR12 = SR6 ** 2
        VIJ = 4.0 * ( SR12 - SR6 )
        WIJ = 24.0 * ( 2.0 * SR12 - SR6 )
        VELIJ = WIJ * DT / RIJSQ
        DVX = VELIJ * RXIJ
        DVY = VELIJ * RYIJ
        DVZ = VELIJ * RZIJ
        VXI = VXI + DVX
        VYI = VYI + DVY
        VZI = VZI + DVZ
        VX(J) = VX(J) - DVX
        VY(J) = VY(J) - DVY
        VZ(J) = VZ(J) - DVZ
        V = V + VIJ
        W = W + WIJ
        NCUT = NCUT + 1
      ENDIF
    ENDIF
  ENDIF

```

```
900
```

```
CONTINUE
```

```

C      ** END OF INNER LOOP **

      VX(I) = VXI
      VY(I) = VYI
      VZ(I) = VZI

1000  CONTINUE

C      ** END OF OUTER LOOP **
C      ** CALCULATE POTENTIAL SHIFT **

      SR2 = SIGSQ / RCUTSQ
      SR6 = SR2 * SR2 * SR2
      SR12 = SR6 * SR6
      VIJ = 4.0 * ( SR12 - SR6 )
      VC  = V - REAL ( NCUT ) * VIJ

      W   = W / 3.0

      RETURN
      END

```

la routine `move`, che aggiorna le posizioni ad ogni passo temporale (mentre le velocità sono aggiornate direttamente nella routine `force`).

```

      SUBROUTINE MOVE ( DT )

      include 'block.inc'

C      *****
C      ** LEAPFROG VERLET ALGORITHM FOR ADVANCING POSITIONS          **
C      *****

      REAL      DT

      INTEGER   I

C      *****

      DO 1000 I = 1, N

          RX(I) = RX(I) + VX(I) * DT
          RY(I) = RY(I) + VY(I) * DT
          RZ(I) = RZ(I) + VZ(I) * DT

1000  CONTINUE

      RETURN
      END

```

la routine `fcc`, che assegna la configurazione iniziale secondo un reticolo cubico a facce centrate

```

      SUBROUTINE FCC

C      *****
C      ** SETS UP THE ALPHA FCC LATTICE FOR N LINEAR MOLECULES.      **
C      **                                                              **
C      ** THE SIMULATION BOX IS A UNIT CUBE CENTRED AT THE ORIGIN.  **
C      ** N SHOULD BE AN INTEGER OF THE FORM ( 4 * ( NC ** 3 ) ),    **
C      ** WHERE NC IS THE NUMBER OF FCC UNIT CELLS IN EACH DIRECTION. **
C      *****

```

```

C   ** SEE FIGURE 5.10 FOR A DIAGRAM OF THE LATTICE AND A           **
C   ** DEFINITION OF THE FOUR ORIENTATIONAL SUBLATTICES.           **
C   **                                                                **
C   ** PRINCIPAL VARIABLES:                                         **
C   **                                                                **
C   ** INTEGER N                                                    NUMBER OF MOLECULES           **
C   ** REAL    RX(N),RY(N),RZ(N)  MOLECULAR POSITIONS        **
C   ** REAL    EX(N),EY(N),EZ(N)  UNIT VECTORS GIVING ORIENTATIONS **
C   ** REAL    RROOT3              1.0 / SQRT ( 3.0 )         **
C   ****
C
C   INCLUDE 'block.inc'
C
C   REAL      RROOT3
C
C   PARAMETER ( RROOT3 = 0.5773503 )
C
C   REAL      CELL, CELL2
C   INTEGER   I, IX, IY, IZ, IREF, M
C
C   ****
C
C   ** CALCULATE THE SIDE OF THE UNIT CELL **
C   CELL = 1.0 / REAL ( NC )
C   CELL2 = 0.5 * CELL
C
C   ** BUILD THE UNIT CELL **
C   ** SUBLATTICE A **
C   RX(1) = 0.0
C   RY(1) = 0.0
C   RZ(1) = 0.0
C
C   ** SUBLATTICE B **
C   RX(2) = CELL2
C   RY(2) = CELL2
C   RZ(2) = 0.0
C
C   ** SUBLATTICE C **
C   RX(3) = 0.0
C   RY(3) = CELL2
C   RZ(3) = CELL2
C
C   ** SUBLATTICE D **
C   RX(4) = CELL2
C   RY(4) = 0.0
C   RZ(4) = CELL2
C
C   ** CONSTRUCT THE LATTICE FROM THE UNIT CELL **
C
C   M = 0
C   DO 99 IZ = 1, NC
C     DO 98 IY = 1, NC
C       DO 97 IX = 1, NC
C         DO 96 IREF = 1, 4
C           RX(IREF+M) = RX(IREF) + CELL * REAL ( IX - 1 )
C           RY(IREF+M) = RY(IREF) + CELL * REAL ( IY - 1 )
C           RZ(IREF+M) = RZ(IREF) + CELL * REAL ( IZ - 1 )
C
C   96     CONTINUE
C         M = M + 4
C   97     CONTINUE
C   98     CONTINUE
C   99     CONTINUE
C
C   ** SHIFT CENTRE OF BOX TO THE ORIGIN **
C   DO 100 I = 1, N
C     RX(I) = RX(I) - 0.5

```

```

      RY(I) = RY(I) - 0.5
      RZ(I) = RZ(I) - 0.5
100  CONTINUE

      RETURN
      END

```

la routine comvel, che assegna le velocità iniziali, definite in modo casuale secondo una distribuzione gaussiana

```

C *****
c ** centre of mass and angular velocities for linear molecules **
c ** ** **
c ** principal variables: **
c ** ** **
c ** integer n the number of molecules **
c ** real rx(n),ry(n),rz(n) positions **
c ** real vx(n),vy(n),vz(n) velocities **
c ** real temp reduced temperature **
c ** ** **
c ** supplied routines: **
c ** ** **
c ** subroutine comvel ( temp ) **
c ** sets the centre of mass velocities for a configuration of **
c ** real function ranf ( dummy ) **
c ** returns a uniform random variate on the range zero to one **
c ** real function gauss ( dummy ) **
c ** returns a uniform random normal variate from a **
c ** distribution with zero mean and unit variance. **
c ** ** **
c ** units: **
c ** ** **
c ** we assume unit molecular mass and employ lennard-jones units **
c ** property units **
c ** rx, ry, rz (epsilon/m)**(1.0/2.0) **
c ** ox, oy, oz (epsilon/m*sigma**2)**(1.0/2.0) **
c ** inert m*sigma**2 **
c *****

      subroutine comvel ( temp )
      include 'block.inc'

c *****
c ** translational velocities from maxwell-boltzmann distribution **
c ** ** **
c ** the distribution is determined by temperature and (unit) mass.**
c ** this routine is general, and can be used for atoms, linear **
c ** molecules, and non-linear molecules. **
c ** ** **
c ** routine referenced: **
c ** ** **
c ** real function gauss ( dummy ) **
c ** returns a uniform random normal variate from a **
c ** distribution with zero mean and unit variance. **
c *****

      REAL temp
      REAL rtemp, sumx, sumy, sumz
      REAL gauss, dummy
      REAL ranf

```

```

external gauss
external ranf

c *****

      rtemp = dsqrt ( temp )
      do 100 i = 1, n
        vx(i) = rtemp * gauss ( dummy )
        vy(i) = rtemp * gauss ( dummy )
        vz(i) = rtemp * gauss ( dummy )
100    continue

c ** remove net momentum **
      sumx = zero
      sumy = zero
      sumz = zero
      do 200 i = 1, n
        sumx = sumx + vx(i)
        sumy = sumy + vy(i)
        sumz = sumz + vz(i)
200    continue

      sumx = sumx / real ( n )
      sumy = sumy / real ( n )
      sumz = sumz / real ( n )
      do 300 i = 1, n
        vx(i) = vx(i) - sumx
        vy(i) = vy(i) - sumy
        vz(i) = vz(i) - sumz
300    continue

      return
      end

      function gauss ( dummy )
      include 'block.inc'
c *****
c ** random variate from the standard normal distribution. **
c ** ** **
c ** the distribution is gaussian with zero mean and unit variance.**
c ** ** **
c ** reference: **
c ** ** **
c ** knuth d, the art of computer programming, (2nd edition **
c ** addison-wesley), 1978 **
c ** ** **
c ** routine referenced: **
c ** ** **
c ** real function ranf ( dummy ) **
c ** returns a uniform random variate on the range zero to one **
c *****

      REAL      a1, a3, a5, a7, a9
      parameter ( a1 = 3.949846138e0, a3 = 0.252408784e0 )
      parameter ( a5 = 0.076542912e0, a7 = 0.008355968e0 )
      parameter ( a9 = 0.029899776e0 )
      REAL      sum, r, r2
      REAL      ranf, dummy
      REAL      gauss
      external  ranf

```

```

C *****
      sum = zero
      do 10 i = 1, 12
        sum = sum + ranf ( dummy )
10    continue
      r = ( sum - 6.e0 ) / 4.e0
      r2 = r * r
      gauss = ((( ( a9 * r2 + a7 ) * r2 + a5 ) * r2 + a3 ) * r2 + a1 )
:      * r
      return
      end

      function ranf ( dummy )

C *****
C ** returns a uniform random variate in the range 0 to 1.      **
C **                                                              **
C **          *****                                          **
C **          ** warning **                                     **
C **          *****                                          **
C **                                                              **
C ** good random number generators are machine specific.      **
C ** please use the one recommended for your machine.        **
C *****

      integer      l, c, m
      parameter ( l = 1029, c = 221591, m = 1048576 )

      integer      seed
      REAL         dummy
      save         seed
      data         seed / 0 /
      REAL         ranf

C *****

      seed = mod ( seed * l + c, m )
      ranf = real ( seed ) / m
      return
      end

C *****
C ** common block                                             **
C *****

      integer nc,n

      parameter (nc=3,n=4*nc**3)
      real rx(n), ry(n), rz(n), vx(n), vy(n), vz(n)
      common / par / rx, ry, rz, vx, vy, vz

```

Ecco il risultato di una simulazione di dinamica molecolare effettuata sul sistema precedente. Il sistema è formato da 108 atomi (vedi, nel listato precedentemente descritto, la variabile *nc*). La simulazione corre su 5000 passi temporali, e ciascun passo corrisponde a 0.004 unità temporali riscalate. Non sono effettuati passi iniziali di equilibrizzazione. Ogni 500 passi il programma stampa l'energia totale del sistema, l'energia cinetica, l'energia potenziale, la pressione, la temperatura ed il coefficiente di compressibilità. Alla fine il programma calcola i valori medi di queste grandezze su tutta l'estensione temporale della simulazione, insieme alla deviazione quadratica media che dà una misura delle fluttuazioni presenti nel sistema.

```
1 **** PROGRAM FROGGY ****
```

```

MOLECULAR DYNAMICS OF LENNARD-JONES ATOMS
LEAPFROG ALGORITHM WITH MINIMAL STORAGE
ENTER RUN TITLE
test n. 1
ENTER NUMBER OF STEPS
5000
ENTER NUMBER OF STEPS BETWEEN OUTPUT LINES
500
ENTER CONFIGURATION FILENAME
save
ENTER THE FOLLOWING IN LENNARD-JONES UNITS
ENTER DENSITY
0.81
ENTER INIT. TEMPERATURE
1.5
ENTER POTENTIAL CUTOFF DISTANCE
3.5
ENTER TIME STEP
0.004

```

```

test n. 1
NUMBER OF ATOMS =      108
NUMBER OF STEPS =     5000
OUTPUT FREQUENCY =      500
POTENTIAL CUTOFF =    3.5000
DENSITY          =     0.8100
TIME STEP        =    0.004000

```

**** START OF DYNAMICS ****

TIMESTEP	..ENERGY..	..KINETIC..	..POTENT..	..PRESSURE..	..TEMPER..	..COMPRESS..
500	-4.4422	1.2232	-5.6656	0.4121	0.8155	-0.3761
1000	-4.4422	1.2445	-5.6865	0.4484	0.8297	-0.3327
1500	-4.4422	1.2990	-5.7413	-0.0802	0.8660	-1.1143
2000	-4.4422	1.2282	-5.6708	0.3816	0.8188	-0.4246
2500	-4.4422	1.2470	-5.6891	0.3459	0.8314	-0.4863
3000	-4.4422	1.3362	-5.7786	-0.3143	0.8908	-1.4356
3500	-4.4422	1.1561	-5.5984	0.6978	0.7707	0.1177
4000	-4.4422	1.1903	-5.6324	0.5920	0.7935	-0.0789
4500	-4.4422	1.1678	-5.6101	0.6005	0.7785	-0.0477
5000	-4.4422	1.1185	-5.5607	0.6746	0.7457	0.1169

**** END OF DYNAMICS ****

AVERAGES	-4.44221	1.23578	-5.67808	0.29658	0.82385	-0.52308
FLUCTS	0.00000	0.08129	0.08090	0.40580	0.05419	0.53147

Bibliografia

Complementi matematici

- [1] W.H. Press, B.P. Flannery, S.A. Teukolsky e W.T. Vetterling, *Numerical Recipes: the Art of Scientific Computing (FORTRAN Version)* (Cambridge University press, Cambridge, 1989)
- [2] M. Abramowitz e I. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1972).
- [3] H.D. Ikramov, *Linear Algebra: Problems Book* (Mir Publishers, Mosca, 1983).
- [4] V.I. Smirnov, *Corso di Matematica Superiore II* (Editori Riuniti, Roma, 1981).
- [5] C.L. Perrin, *Mathematics for Chemists*, (Wiley, New York, 1970).
- [6] M.D. Johnston, *Computational Chemistry* (Elsevier, Amsterdam, 1988)