

Funzioni periodiche, serie di Fourier e onde

1. Come vedrete, uno strumento matematico di cui faremo uso a un certo punto di questo corso e' la serie di Fourier. Per questo motivo, mi sembra opportuno introdurre l'argomento con un certo dettaglio.

Funzioni periodiche

2. Le serie di Fourier, per lo meno limitatamente all'uso che ne faremo in questo corso, sono definite per funzioni periodiche (o che possono essere trattate come tali). Cominciamo quindi a definire e analizzare le funzioni periodiche.

3. Una funzione di una variabile $f(x)$ e' periodica di periodo T se soddisfa:

$$\forall x \in \mathbb{R}, \forall k \in \mathbb{Z} : f(x + kT) = f(x) \quad (1)$$

La 1 significa questo: se ripartiamo l'asse x in intervalli di ampiezza T , il grafico della f si ripete esattamente uguale a se stesso in ciascun intervallo (figura 1).

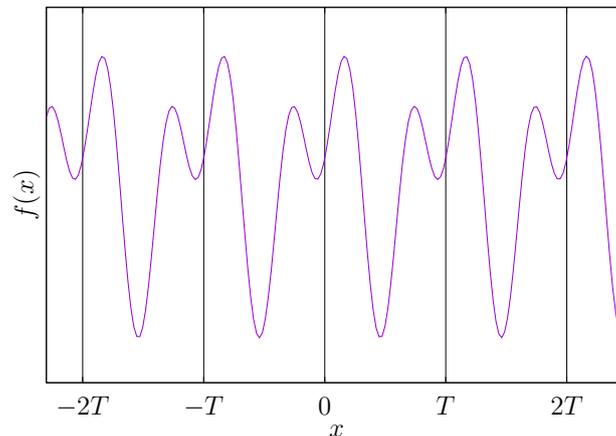


Figura 1:

Notate che la 1 si puo' scrivere in modo piu' semplice come:

$$\forall x : f(x + T) = f(x) \quad (2)$$

perche' la 2 implica la 1¹. Ma secondo me la 1 e' piu' "esplicativa".

4. Il periodo di una funzione periodica, T nella 1, e' il *minimo* numero per cui la 1 e' verificata: esso viene anche chiamato *periodo fondamentale*.

¹Supponiamo che valga la 2 e vogliamo provare che vale la 1.

$$f(x + kT) = f\left(\overbrace{x + (k-1)T}^{x'} + T\right) \stackrel{\text{per la 2}}{=} f(x' + T) = f(x') = f(x + (k-1)T)$$

A questo punto possiamo scrivere una serie di uguaglianze per ricorsione:

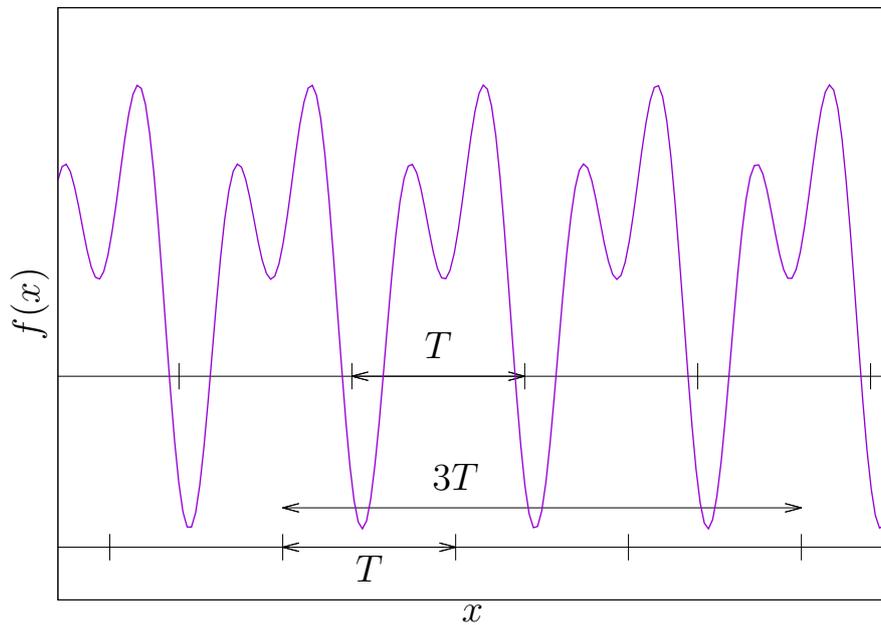


Figura 2: A una funzione periodica di periodo T e' associato un "reticolo" la cui cella primitiva e' l'intervallo di ampiezza T ; la figura mostra anche una supercella tripla (intervallo di ampiezza $3T$); inoltre, e' mostrato anche un reticolo traslato rispetto al primo: la funzione rimane periodica anche sul reticolo traslato.

E' chiaro che, come dice esplicitamente la 1, una funzione periodica di periodo T e' periodica anche di periodo $2T, 3T, 4T, \dots$

5. Facciamo una semplice osservazione che acquistera' maggior significato quando considereremo funzioni periodiche in 2 e 3 dimensioni.

Alla funzione periodica $f(x)$ avente periodo T rimane associato un "reticolo" monodimensionale sull'asse x costituito da punti equispaziati alla distanza T (figura 2). Per motivi che appariranno chiari piu' avanti, tale reticolo viene chiamato, in generale, *reticolo (o spazio) diretto*.

Se T e' il periodo fondamentale della f , chiamiamo *cella primitiva o elementare* del reticolo (monodimensionale) uno qualsiasi degli intervalli (di ampiezza T) in cui e' suddiviso l'asse x .

$$\begin{aligned}
 f(x + kT) &= f(x + (k - 1)T) \\
 f(x + (k - 1)T) &= f(x + (k - 2)T) \\
 f(x + (k - 2)T) &= f(x + (k - 3)T) \\
 &\dots = \dots \\
 f(x + T) &= f(x)
 \end{aligned}$$

da cui:

$$f(x + kT) = f(x)$$

Chiaramente, il grafico della funzione $f(x)$ si ripete identicamente in ogni cella primitiva del reticolo.

Possiamo ancora introdurre il concetto di *supercella*: un intervallo costituito dall'unione di piu' celle primitive. Una supercella doppia, tripla etc significa un intervallo di ampiezza $2T, 3T \dots$, rispettivamente. In base alla 1, il grafico della funzione $f(x)$ si ripete identicamente anche in ogni supercella.

Osservate ancora che non ha nessuna importanza il punto scelto come origine del reticolo associato alla f : qualsiasi sia l'origine fissata, le affermazioni fatte sopra riguardo alla periodicit  della f rimangono valide. La cosa essenziale   che la cella primitiva del reticolo sia un intervallo di ampiezza T .

6. Si definisce *frequenza* di una funzione periodica $f(x)$ di periodo T il numero di periodi per unit  della variabile x .

Siccome il numero di periodi contenuto in T unit  di x   1, la frequenza ν   data da:

$$\nu = \frac{[\text{n. di periodi}]}{[\text{n. di unit  dell'asse } x]} = \frac{1}{T} \quad (3)$$

Notate che l'unit  di misura della frequenza   l'inverso dell'unit  di misura del periodo.

La frequenza cos  definita viene anche detta *frequenza lineare*, per distinguerla dalla *frequenza angolare*, che definiremo fra un attimo.

7. Le funzioni periodiche per antonomasia sono il seno e il coseno.

Il periodo "naturale" di queste due funzioni   2π :

$$\begin{aligned} \sin(x + k(2\pi)) &= \sin(x) \\ \cos(x + k(2\pi)) &= \cos(x) \end{aligned}$$

E' comunque possibile imporre a queste due funzioni un periodo T qualsiasi moltiplicando l'argomento per $2\pi/T$; le funzioni:

$$\sin\left(\frac{2\pi}{T}x\right) \text{ e } \cos\left(\frac{2\pi}{T}x\right)$$

sono periodiche di periodo T :

$$\begin{aligned} \sin\left(\frac{2\pi}{T}(x + kT)\right) &= \sin\left(\frac{2\pi}{T}x + k(2\pi)\right) = \sin\left(\frac{2\pi}{T}x\right) \\ \cos\left(\frac{2\pi}{T}(x + kT)\right) &= \cos\left(\frac{2\pi}{T}x + k(2\pi)\right) = \cos\left(\frac{2\pi}{T}x\right) \end{aligned}$$

Siccome $\nu = 1/T$   la frequenza lineare (eq. 3), viene definita la *frequenza angolare* con:

$$\omega = \frac{2\pi}{T} = 2\pi\nu \quad (4)$$

In tal modo   possibile scrivere in forma piu' compatta le funzioni trigonometriche seno e coseno aventi un periodo qualsiasi. Le funzioni:

$$\sin(\omega x) \text{ e } \cos(\omega x) \quad (5)$$

hanno il periodo:

$$T = \frac{2\pi}{\omega} \quad (6)$$

e la frequenza lineare:

$$\nu = \frac{\omega}{2\pi} \quad (7)$$

8. Una identita' che useremo spesso in questo corso e' l'identita' di Eulero, che mette in relazione le funzioni seno e coseno con la funzione esponenziale:

$$\exp(i\theta) = \cos(\theta) + i \sin(\theta) \quad (8)$$

(i e' l'unita' immaginaria:² $i^2 = -1$)

Notate che se $\theta = -\alpha < 0$ l'identita' di Eulero continua a valere:

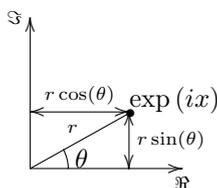
$$\exp(-i\alpha) = \cos(-\alpha) + i \sin(-\alpha) = \cos(\alpha) - i \sin(\alpha) \quad (9)$$

Per cui, piu' esplicitamente, si puo' scrivere:

$$\exp(\pm i\theta) = \cos(\theta) \pm i \sin(\theta) \quad (10)$$

9. Trovate la derivazione dell'identita' di Eulero su qualsiasi libro di testo. Riporto una delle derivazioni che potete trovare su Wikipedia.¹

Il numero complesso $\exp(ix)$ puo' essere sempre scritto in forma polare:



$$\exp(ix) = r (\cos(\theta) + i \sin(\theta)) \quad (11)$$

con $r = r(x)$ e $\theta = \theta(x)$.

Notate che deve essere: $r(x) > 0$ perche' l'esponenziale nella 11 e' sempre $\neq 0$ (per definizione il modulo della forma polare del numero complesso $a + ib$ e' sempre ≥ 0 , essendo dato da $r = \sqrt{a^2 + b^2}$).

Facendo la derivata prima di ambo i membri:

$$i \exp(ix) = (\cos(\theta) + i \sin(\theta)) \frac{dr}{dx} + r (-\sin(\theta) + i \cos(\theta)) \frac{d\theta}{dx}$$

Sostituiamo la forma polare a $\exp(ix)$ e isoliamo parte reale e immaginaria in ambo i membri:

$$i (r (\cos(\theta) + i \sin(\theta))) = \frac{dr}{dx} \cos(\theta) - r \frac{d\theta}{dx} \sin(\theta) + i \left(\frac{dr}{dx} \sin(\theta) + r \frac{d\theta}{dx} \cos(\theta) \right)$$

²Do per scontato che abbiate delle nozioni elementari sui numeri complessi.

$$-r \sin(\theta) + ir \cos(\theta) = \frac{dr}{dx} \cos(\theta) - r \frac{d\theta}{dx} \sin(\theta) + i \left(\frac{dr}{dx} \sin(\theta) + r \frac{d\theta}{dx} \cos(\theta) \right)$$

Dal confronto delle parti reali (o immaginarie) di entrambi i membri si ricava:

$$\frac{dr}{dx} = 0 \rightarrow r(x) = \text{costante} \quad (12)$$

$$\frac{d\theta}{dx} = 1 \rightarrow \theta(x) = x + \theta(0) \quad (13)$$

La 11 deve valere *per ogni* x , quindi anche per $x = 0$; in questo caso si ha:

$$1 = r(0) (\cos(\theta(0)) + i \sin(\theta(0)))$$

da cui segue:

$$\theta(0) = 0 \quad (\text{perche' (i) } r(0) \neq 0 \text{ e (ii) la parte immaginaria deve essere nulla})$$

$$\theta(0) = 0 \Rightarrow \theta = x \quad (\text{perche' } \theta(x) = x + \theta(0), \text{ eq. 13})$$

$$r(0) = 1 \Rightarrow r = 1 \quad \forall x \quad (\text{perche' } r = \text{costante})$$

e quindi la 11 diventa proprio la formula di Eulero 8.

10. Potete verificare molto facilmente che grazie all'identita' di Eulero le funzioni seno e coseno possono essere scritte come combinazione di funzioni esponenziali:

$$\cos(\theta) = \frac{1}{2} (\exp(i\theta) + \exp(-i\theta)) \quad (14)$$

$$\begin{aligned} \sin(\theta) &= \frac{1}{2i} (\exp(i\theta) - \exp(-i\theta)) \\ &= -\frac{i}{2} (\exp(i\theta) - \exp(-i\theta)) \end{aligned} \quad (15)$$

Sviluppo in serie di Fourier

11. Vediamo ora come qualsiasi funzione periodica³ possa essere scritta come una serie, che viene chiamata appunto *serie di Fourier*.

12. Questa introduzione non ha alcuna pretesa di rigore matematico, ma cerco di svolgere tutti i passaggi, perche' una volta nella vita e' bene farlo.

Faremo la derivazione dettagliata della serie di Fourier in una sola dimensione (cioe' per funzioni di una sola variabile) e poi estenderemo i risultati ottenuti (senza dettagli e in modo intuitivo) al caso bi e tridimensionale (cioe' per funzioni di due o tre variabili).

³In realta' non e' vero che *qualsiasi* funzione periodica puo' essere scritta come una serie di Fourier; tuttavia, in questo corso incontreremo solo funzioni "buone" che si lasciano scrivere come serie di Fourier.

13. Consideriamo una funzione reale di una variabile reale $f(x)$ periodica di periodo T e ci proponiamo di scriverla come una somma infinita di addendi. Siccome la funzione è periodica di periodo T , converrà svilupparla con funzioni aventi come periodo T o un sottomultiplo di T . Notate che, in base alla 1, una funzione $g(x)$ avente periodo $T' = T/k$ (con k intero) è ovviamente periodica anche di periodo T :

$$g(x+T) = g(x+kT') = g(x) \quad (16)$$

Non ci conviene prendere funzioni *non* periodiche o con periodo maggiore di T , perché con esse non potremmo riprodurre la periodicità della funzione f .

14. Potremmo quindi pensare di sviluppare la nostra $f(x)$ con delle funzioni seno e coseno di periodo $T, T/2, T/3, \dots$, ciascuna moltiplicata per un opportuno coefficiente da determinare (vedremo fra un attimo come):

$$\begin{aligned} f(x) &= \sum_{i=1}^{\infty} C_i \cos\left(\frac{2\pi}{T}x\right) + \sum_{i=1}^{\infty} S_i \sin\left(\frac{2\pi}{T}x\right) \\ &= \sum_{i=1}^{\infty} C_i \cos\left(\frac{2\pi i}{T}x\right) + \sum_{i=1}^{\infty} S_i \sin\left(\frac{2\pi i}{T}x\right) \end{aligned}$$

dove, come appena detto, i coefficienti C_i e S_i di ciascun addendo (indipendenti da x) andranno opportunamente determinati affinché l'uguaglianza sia soddisfatta.

C'è ancora un problema. Le funzioni seno e coseno oscillano sopra e sotto la retta $y=0$; ma, in generale, la funzione f da sviluppare potrebbe essere shiftata arbitrariamente verso l'alto o verso il basso lungo l'asse y : per tenere conto di ciò, dobbiamo aggiungere alla somma un termine costante che chiamiamo F_0 . In definitiva, la nostra "proposta" di sviluppo è la seguente:

$$f(x) = F_0 + \sum_{i=1}^{\infty} C_i \cos\left(\frac{2\pi i}{T}x\right) + \sum_{i=1}^{\infty} S_i \sin\left(\frac{2\pi i}{T}x\right) \quad (17)$$

15. La 17 può funzionare soltanto se esistono e sono determinabili i coefficienti F_0, C_i e S_i in modo tale che l'uguaglianza sia vera. Questo è quello che verificheremo ora.

Cominciamo con gli S_i . Per determinarli, moltiplichiamo ambo i membri della 17 per $\sin\left(\frac{2\pi j}{T}x\right)$ e integriamo su un periodo fra a e $a+T$ (a qualunque):

$$\begin{aligned} \int_a^{a+T} f(x) \sin\left(\frac{2\pi j}{T}x\right) dx &= F_0 \int_a^{a+T} \sin\left(\frac{2\pi j}{T}x\right) dx \\ &+ \sum_{i=1}^{\infty} C_i \int_a^{a+T} \cos\left(\frac{2\pi i}{T}x\right) \sin\left(\frac{2\pi j}{T}x\right) dx \\ &+ \sum_{i=1}^{\infty} S_i \int_a^{a+T} \sin\left(\frac{2\pi i}{T}x\right) \sin\left(\frac{2\pi j}{T}x\right) dx \quad (18) \end{aligned}$$

Fate attenzione: $j \neq i$; cioè j è un particolare valore preso fra tutti i possibili valori dell'indice i .

16. Per poter procedere ci servono alcuni risultati generali relativi agli integrali delle funzioni seno e coseno.

Dovreste essere in grado di verificare molto facilmente che:

$$\int_a^{a+T} \sin\left(\frac{2\pi j}{T}x\right) dx = 0 \quad \forall j \in \mathbb{Z} \quad (19)$$

$$\int_a^{a+T} \cos\left(\frac{2\pi j}{T}x\right) dx = \begin{cases} 0 & \forall j \neq 0 \\ T & j = 0 \end{cases} \quad (20)$$

Inoltre, utilizzando le formule di prostaferesi:⁴

$$\sin \alpha \sin \beta = \frac{1}{2} (\cos (\alpha - \beta) - \cos (\alpha + \beta)) \quad (21)$$

$$\sin \alpha \cos \beta = \frac{1}{2} (\sin (\alpha + \beta) + \sin (\alpha - \beta)) \quad (22)$$

si ha:

$$\begin{aligned} \int_a^{a+T} \cos\left(\frac{2\pi i}{T}x\right) \sin\left(\frac{2\pi j}{T}x\right) dx &= \frac{1}{2} \int_a^{a+T} \left(\sin\left(\frac{2\pi(i+j)}{T}x\right) \right. \\ &\quad \left. + \sin\left(\frac{2\pi(j-i)}{T}x\right) \right) dx \\ &= 0 \end{aligned} \quad (23)$$

per la 19.

Inoltre:

$$\begin{aligned} \int_a^{a+T} \sin\left(\frac{2\pi i}{T}x\right) \sin\left(\frac{2\pi j}{T}x\right) dx &= \frac{1}{2} \int_a^{a+T} \left(\cos\left(\frac{2\pi(i-j)}{T}x\right) \right. \\ &\quad \left. - \cos\left(\frac{2\pi(j+i)}{T}x\right) \right) dx \\ &= \begin{cases} 0 & \text{quando } i \neq j \\ T/2 & \text{quando } i = j \end{cases} \end{aligned} \quad (24)$$

in base alla 20.

17. Tornando alla 18 possiamo quindi vedere che il primo termine al secondo membro e' nullo per la 19; tutti i termini della sommatoria contenente i C_i sono

⁴Ecco un'occasione per vedere l'utilita' dell'identita' di Eulero.

Dimostriamo la 21 (la 22 si dimostra in modo analogo).

Utilizzando la 15 si ha:

$$\begin{aligned} \sin \alpha \sin \beta &= \frac{\exp(i\alpha) - \exp(-i\alpha)}{2i} \frac{\exp(i\beta) - \exp(-i\beta)}{2i} \\ &= - \frac{\overbrace{\exp(i(\alpha+\beta)) + \exp(-i(\alpha+\beta))}^{2 \cos(\alpha+\beta) \text{ per la 14}} - \overbrace{\exp(i(\alpha-\beta)) - \exp(-i(\alpha-\beta))}^{-2 \cos(\alpha-\beta) \text{ per la 14}}}{4} \\ &= \frac{1}{2} (\cos (\alpha - \beta) - \cos (\alpha + \beta)) \end{aligned}$$

nulli per la 23; dei termini nella sommatoria contenente gli S_i solo quello per $i = j$ “sopravvive” ed e’ uguale a $T/2$.

Otteniamo quindi:

$$\begin{aligned} \int_a^{a+T} f(x) \sin\left(\frac{2\pi j}{T}x\right) dx &= S_j\left(\frac{T}{2}\right) \\ S_j &= \frac{2}{T} \int_a^{a+T} f(x) \sin\left(\frac{2\pi j}{T}x\right) dx \end{aligned} \quad (25)$$

Cosa significa questo risultato? Siccome j e’ un qualsiasi valore di i fra tutti i possibili, la 25 definisce i coefficienti $S_j \forall j \geq 1$ e la 17 e’ vera se gli S_j hanno il valore prescritto dalla 25.

18. Passiamo ora ai C_i . In modo analogo a quanto fatto per gli S_i , moltiplichiamo ambo i membri della 17 per $\cos((2\pi j/T)x)$ e integriamo fra a e $a + T$:

$$\begin{aligned} \int_a^{a+T} f(x) \cos\left(\frac{2\pi j}{T}x\right) dx &= F_0 \int_a^{a+T} \cos\left(\frac{2\pi j}{T}x\right) dx \\ &+ \sum_{i=1}^{\infty} C_i \int_a^{a+T} \cos\left(\frac{2\pi i}{T}x\right) \cos\left(\frac{2\pi j}{T}x\right) dx \\ &+ \sum_{i=1}^{\infty} S_i \int_a^{a+T} \sin\left(\frac{2\pi i}{T}x\right) \cos\left(\frac{2\pi j}{T}x\right) dx \end{aligned} \quad (26)$$

Il primo termine al secondo membro e’ nullo per la 20 e il terzo termine (la sommatoria sugli S_i) e’ nullo per la 23.

Per quanto riguarda il secondo termine (la sommatoria sui C_i), utilizzando un’altra formula di prostaferesi:⁵

$$\cos \alpha \cos \beta = \frac{1}{2} (\cos (\alpha + \beta) + \cos (\alpha - \beta))$$

lo possiamo riscrivere come:

$$\begin{aligned} \sum_{i=1}^{\infty} C_i \int_a^{a+T} \cos\left(\frac{2\pi i}{T}x\right) \cos\left(\frac{2\pi j}{T}x\right) dx &= \\ \frac{1}{2} \sum_{i=1}^{\infty} C_i \int_a^{a+T} \cos\left(\frac{2\pi(i+j)}{T}x\right) + \cos\left(\frac{2\pi(i-j)}{T}x\right) dx & \end{aligned} \quad (27)$$

da cui si vede che, in base alla 20, tutti i termini della sommatoria sono nulli, tranne quando $i = j$: in questo caso, l’integrale vale T .

In definitiva, la 26 diventa:

⁵Anche questa si dimostra facilmente utilizzando l’identita’ di Eulero.

$$\int_a^{a+T} f(x) \cos\left(\frac{2\pi j}{T}x\right) dx = \frac{1}{2}C_j T$$

$$C_j = \frac{2}{T} \int_a^{a+T} f(x) \cos\left(\frac{2\pi j}{T}x\right) dx \quad (28)$$

Ci manca ancora da considerare F_0 . A questo scopo integriamo ambo i membri della 17 da a a $a+T$:

$$\int_a^{a+T} f(x) dx = \int_a^{a+T} F_0 dx + \sum_{i=1}^{\infty} C_i \int_a^{a+T} \cos\left(\frac{2\pi i}{T}x\right) dx + \sum_{i=1}^{\infty} S_i \int_a^{a+T} \sin\left(\frac{2\pi i}{T}x\right) dx$$

Chiaramente, tutti i termini delle sommatorie sono nulli per le 20 e 19, mentre il primo integrale al secondo membro e' uguale a $F_0 T$ (ricordate che F_0 non dipende da x). Quindi:

$$\int_a^{a+T} f(x) dx = F_0 T$$

$$F_0 = \frac{1}{T} \int_a^{a+T} f(x) dx \quad (29)$$

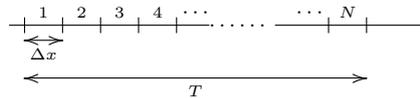
Come forse si poteva prevedere, il termine F_0 e' *il valore medio* della funzione $f(x)$ su un periodo⁶; cioe' e' lo shift che dobbiamo aggiungere alle somme di coseni e seni per fare in modo che le loro oscillazioni siano localizzate nello stesso intervallo dell'asse y nel quale e' localizzato il grafico della f .

19. Ricapitoliamo quanto fatto finora. Abbiamo proposto un possibile sviluppo in serie della funzione $f(x)$ (eq. 17). Tale sviluppo dipende da certi coefficienti (F_0, C_i, S_i) che, a priori, non e' detto sia possibile determinare. Ma abbiamo appena visto che la determinazione di questi coefficienti e' possibile e ne abbiamo trovato l'espressione. Quindi rimane "provato" che, data la funzione $f(x)$, e' sempre possibile svilupparla in serie nel modo seguente:

⁶In generale, il seguente integrale (anche per una funzione non periodica):

$$\frac{1}{T} \int_a^{a+T} f(x) dx$$

ha il significato di valore medio della funzione sull'intervallo T . Potete rendervene conto facilmente se scrivete l'integrale come limite della sua forma finita. Dividendo l'intervallo di integrazione T in N subintervalli di ampiezza comune Δx :



e detto x_i un valore di x interno all'intervallo i , si ha:

$$\frac{1}{T} \int_a^{a+T} f(x) dx = \lim_{N \rightarrow \infty} \frac{1}{N \Delta x} \sum_1^N f(x_i) \Delta x = \lim_{N \rightarrow \infty} \frac{\sum_1^N f(x_i)}{N} = \langle f(x) \rangle$$

dove $\langle f(x) \rangle$ e' proprio il valore medio della f sull'intervallo T .

$$\text{eq. 17: } f(x) = F_0 + \sum_{i=1}^{\infty} C_i \cos\left(\frac{2\pi i}{T}x\right) + \sum_{i=1}^{\infty} S_i \sin\left(\frac{2\pi i}{T}x\right)$$

$$\text{eq. 29: } F_0 = \frac{1}{T} \int_a^{a+T} f(x) dx$$

$$\text{eq. 28: } C_i = \frac{2}{T} \int_a^{a+T} f(x) \cos\left(\frac{2\pi i}{T}x\right) dx$$

$$\text{eq. 25: } S_i = \frac{2}{T} \int_a^{a+T} f(x) \sin\left(\frac{2\pi i}{T}x\right) dx$$

La serie che abbiamo ottenuto si chiama *serie di Fourier* ed e' difficile sopravvalutarne l'importanza.

20. Esistono molte versioni equivalenti delle equazioni che definiscono la serie di Fourier.

Quasi sempre, gli estremi di integrazione sono o $[0, T]$ (cioe' $a = 0$), oppure $[-T/2, T/2]$ (cioe' $a = -T/2$); dal procedimento che abbiamo seguito, dovrebbe tuttavia essere chiaro che a puo' essere un numero reale qualsiasi.⁷

In altri casi, il periodo della funzione $f(x)$ viene indicato con $2T$ (invece che con T). In tal modo le espressioni 17, 29, 28 e 25 diventano:

⁷A questo proposito, si puo' dimostrare facilmente che in generale, per una funzione periodica di periodo T , vale:

$$\int_a^{a+T} f(x) dx = \int_b^{b+T} f(x) dx \quad \forall a, b$$

cioe': l'integrale su un intervallo di ampiezza uguale al periodo e' invariante rispetto agli estremi di integrazione.

Definite la funzione:

$$H(z) = \int_z^{z+T} f(x) dx$$

Detta F una funzione primitiva di f , sara':

$$H(z) = F(z+T) - F(z)$$

Derivate ambo i membri rispetto a z :

$$\frac{dH}{dz} = \frac{dF(z+T)}{dz} - \frac{dF(z)}{dz} = f(z) - f(z) = 0$$

dove si sfrutta la definizione di funzione primitiva di una funzione data e la derivata di una funzione composta.

Ma quanto sopra equivale a dire che $H(z)$ e' una costante, ovvero:

$$H(a) = H(b) \quad \forall a, b \Rightarrow \int_a^{a+T} f(x) dx = \int_b^{b+T} f(x) dx \quad \forall a, b$$

(trovata sul web²)

$$f(x) = F_0 + \sum_{i=1}^{\infty} C_i \cos\left(\frac{\pi i}{T}x\right) + \sum_{i=1}^{\infty} S_i \sin\left(\frac{\pi i}{T}x\right) \quad (30)$$

$$F_0 = \frac{1}{2T} \int_a^{a+2T} f(x) dx \quad (31)$$

$$C_i = \frac{1}{T} \int_a^{a+2T} f(x) \cos\left(\frac{\pi i}{T}x\right) dx \quad (32)$$

$$S_i = \frac{1}{T} \int_a^{a+2T} f(x) \sin\left(\frac{\pi i}{T}x\right) dx \quad (33)$$

e ci sono meno fattori 2.

21. Notate che tutte le funzioni che compaiono all'interno degli integrali 29, 28 e 25 sono periodiche di periodo T .

$f(x)$, nella 29, lo è per ipotesi.

Detta $F(x)$ la funzione sotto il segno di integrale della 28, si ha:

$$\begin{aligned} F(x+T) &= f(x+T) \cos\left(\frac{2\pi i}{T}(x+T)\right) \\ &= f(x) \cos\left(\frac{2\pi i}{T}x + 2\pi i\right) \\ &= f(x) \cos\left(\frac{2\pi i}{T}x\right) \\ &= F(x) \end{aligned}$$

E analogamente per la funzione sotto il segno di integrale della 25.

22. Vi potrà capitare di trovare un'espressione più compatta della serie di Fourier:

$$f(x) = \sum_{i=-\infty}^{\infty} C'_i \cos\left(\frac{2\pi i}{T}x\right) + \sum_{i=-\infty}^{\infty} S'_i \sin\left(\frac{2\pi i}{T}x\right) \quad (34)$$

con:

$$C'_i = \frac{1}{2}C_i = \frac{1}{T} \int_a^{a+T} f(x) \cos\left(\frac{2\pi i}{T}x\right) dx \quad (35)$$

$$S'_i = \frac{1}{2}S_i = \frac{1}{T} \int_a^{a+T} f(x) \sin\left(\frac{2\pi i}{T}x\right) dx \quad (36)$$

Notate: le sommatorie vanno da $-\infty$ a ∞ , manca il termine singolo pre-sommatorie F_0 e nelle 35 e 36 c'è $1/T$ al posto di $2/T$.

E' semplice ricavare le 34, 35 e 36 dalle 17, 25 e 28.

Cominciamo a vedere i primi due addendi al secondo membro della 17:

$$F_0 + \sum_{i=1}^{\infty} C_i \cos\left(\frac{2\pi i}{T}x\right) \quad (37)$$

Sommo e sottraggo C_0 e osservo che:

$$C_0 = C_0 \cos\left(\frac{2\pi 0}{T}x\right) \quad (38)$$

per cui posso includere il termine C_0 positivo nella sommatoria, facendola partire da 0 (invece che da 1):

$$= F_0 - C_0 + \sum_{i=0}^{\infty} C_i \cos\left(\frac{2\pi i}{T}x\right) \quad (39)$$

In base alle definizioni dei coefficienti C_i (eq. 28) e di F_0 (eq. 29) vale:

$$C_0 = \frac{2}{T} \int_a^{a+T} f(x) \cos\left(\frac{2\pi 0}{T}x\right) dx = \frac{2}{T} \int_a^{a+T} f(x) dx = 2F_0 \quad (40)$$

e quindi possiamo continuare con:

$$= -F_0 + \sum_{i=0}^{\infty} C_i \cos\left(\frac{2\pi i}{T}x\right) \quad (41)$$

La sommatoria puo' essere scritta come differenza fra le seguenti due sommatorie:

$$= -F_0 + \sum_{i=-\infty}^{\infty} C_i \cos\left(\frac{2\pi i}{T}x\right) - \sum_{i=-\infty}^{-1} C_i \cos\left(\frac{2\pi i}{T}x\right) \quad (42)$$

Ora cambio il segno e l'ordine degli estremi nella seconda sommatoria e, contemporaneamente, cambio il segno dell'indice negli addendi:

$$= -F_0 + \sum_{i=-\infty}^{\infty} C_i \cos\left(\frac{2\pi i}{T}x\right) - \sum_{i=1}^{\infty} C_{-i} \cos\left(-\frac{2\pi i}{T}x\right) \quad (43)$$

Ora osservo che i coefficienti C_i sono funzioni *pari* di i , cioe': $C_{-i} = C_i$. Infatti, per definizione:

$$C_{-i} = \frac{2}{T} \int_a^{a+T} f(x) \cos\left(-\frac{2\pi i}{T}x\right) dx = \frac{2}{T} \int_a^{a+T} f(x) \cos\left(\frac{2\pi i}{T}x\right) dx = C_i \quad (44)$$

perche' il coseno e' una funzione pari.

Saltando le uguaglianze intermedie, siamo arrivati a:

$$F_0 + \sum_{i=1}^{\infty} C_i \cos\left(\frac{2\pi i}{T}x\right) = -F_0 + \sum_{i=-\infty}^{\infty} C_i \cos\left(\frac{2\pi i}{T}x\right) - \sum_{i=1}^{\infty} C_i \cos\left(\frac{2\pi i}{T}x\right) \quad (45)$$

ovvero:

$$2F_0 + 2 \sum_{i=1}^{\infty} C_i \cos\left(\frac{2\pi i}{T}x\right) = \sum_{i=-\infty}^{\infty} C_i \cos\left(\frac{2\pi i}{T}x\right) \quad (46)$$

da cui:

$$F_0 + \sum_{i=1}^{\infty} C_i \cos\left(\frac{2\pi i}{T}x\right) = \sum_{i=-\infty}^{\infty} \frac{1}{2} C_i \cos\left(\frac{2\pi i}{T}x\right) = \sum_{i=-\infty}^{\infty} C'_i \cos\left(\frac{2\pi i}{T}x\right) \quad (47)$$

con:

$$C'_i = \frac{1}{2} C_i = \frac{1}{T} \int_a^{a+T} f(x) \cos\left(\frac{2\pi i}{T}x\right) dx \quad (48)$$

che e' esattamente la 35.

Ora vediamo la seconda sommatoria della 17:

$$\sum_{i=1}^{\infty} S_i \sin\left(\frac{2\pi i}{T}x\right) \quad (49)$$

Prima di tutto osservo che posso aumentare la sommatoria con $i = 0$, visto che $\sin(0) = 0$ e quindi non cambia nulla:

$$= \sum_{i=0}^{\infty} S_i \sin\left(\frac{2\pi i}{T}x\right) \quad (50)$$

Come prima, scrivo la sommatoria come differenza fra le due seguenti:

$$= \sum_{i=-\infty}^{\infty} S_i \sin\left(\frac{2\pi i}{T}x\right) - \sum_{i=-\infty}^{-1} S_i \sin\left(\frac{2\pi i}{T}x\right) \quad (51)$$

Ora cambio il segno e l'ordine degli estremi nella seconda sommatoria e, contemporaneamente, cambio il segno dell'indice negli addendi:

$$= \sum_{i=-\infty}^{\infty} S_i \sin\left(\frac{2\pi i}{T}x\right) - \sum_{i=1}^{\infty} S_{-i} \sin\left(-\frac{2\pi i}{T}x\right) \quad (52)$$

Ora osservo che i coefficienti S_i sono funzioni *dispari* di i , cioe': $S_{-i} = -S_i$. Infatti, per definizione:

$$S_{-i} = \frac{2}{T} \int_a^{a+T} f(x) \sin\left(-\frac{2\pi i}{T}x\right) dx = -\frac{2}{T} \int_a^{a+T} f(x) \sin\left(\frac{2\pi i}{T}x\right) dx = -S_i \quad (53)$$

perche' il seno e' una funzione dispari.

Quindi deve valere:

$$S_{-i} \sin\left(-\frac{2\pi i}{T}x\right) = -S_i \left(-\sin\left(\frac{2\pi i}{T}x\right)\right) = S_i \sin\left(\frac{2\pi i}{T}x\right) \quad (54)$$

Sostituendo sopra possiamo continuare con:

$$= \sum_{i=-\infty}^{\infty} S_i \sin\left(\frac{2\pi i}{T}x\right) - \sum_{i=1}^{\infty} S_i \sin\left(\frac{2\pi i}{T}x\right) \quad (55)$$

Saltando le uguaglianze intermedie, siamo arrivati a:

$$\sum_{i=1}^{\infty} S_i \sin\left(\frac{2\pi i}{T}x\right) = \sum_{i=-\infty}^{\infty} S_i \sin\left(\frac{2\pi i}{T}x\right) - \sum_{i=1}^{\infty} S_i \sin\left(\frac{2\pi i}{T}x\right) \quad (56)$$

ovvero:

$$2 \sum_{i=1}^{\infty} S_i \sin\left(\frac{2\pi i}{T}x\right) = \sum_{i=-\infty}^{\infty} S_i \sin\left(\frac{2\pi i}{T}x\right) \quad (57)$$

da cui:

$$\sum_{i=1}^{\infty} S_i \sin\left(\frac{2\pi i}{T}x\right) = \sum_{i=-\infty}^{\infty} \frac{1}{2} S_i \sin\left(\frac{2\pi i}{T}x\right) = \sum_{i=-\infty}^{\infty} S'_i \sin\left(\frac{2\pi i}{T}x\right) \quad (58)$$

con:

$$S'_i = \frac{1}{2} S_i = \frac{1}{T} \int_a^{a+T} f(x) \sin\left(\frac{2\pi i}{T}x\right) dx \quad (59)$$

che e' la 36.

Ovviamente, sostituendo le 47 e 58 nella 17 si ottiene la 34.

23. La forma esponenziale complessa della serie di Fourier.

Come vedrete, noi non utilizzeremo la serie di Fourier nella forma che finora abbiamo visto (cioe' come somma di seni e coseni); invece, quella che utilizzeremo sara' la sua cosiddetta *forma esponenziale*.

Possiamo ricavarla a partire dalla 34 che, in base alle 35 e 36, puo' essere convenientemente scritta nella forma:

$$f(x) = \frac{1}{2} \sum_{n=-\infty}^{\infty} C_n \cos(\alpha_n) + S_n \sin(\alpha_n) \quad (60)$$

dove ho posto, per brevitaa':

$$\alpha_n = \frac{2\pi n}{T}x$$

(l'indice e' diventato n per non fare confusione con l'unitaa' immaginaria i ; C_n e S_n sono quelli definiti dalle 28 e 25, rispettivamente)

Scrivo il coseno e il seno come somma di esponenziali in base all'identita' di Eulero (eq. 14 e 15):

$$= \frac{1}{2} \sum_{n=-\infty}^{\infty} C_n \frac{\exp(i\alpha_n) + \exp(-i\alpha_n)}{2} - S_n \frac{i(\exp(i\alpha_n) - \exp(-i\alpha_n))}{2} \quad (61)$$

$$= \frac{1}{4} \sum_{n=-\infty}^{\infty} C_n (\exp(i\alpha_n) + \exp(-i\alpha_n)) - S_n (i(\exp(i\alpha_n) - \exp(-i\alpha_n))) \quad (62)$$

Svolgo i prodotti e raccolgo a fattor comune gli esponenziali:

$$= \frac{1}{4} \sum_{n=-\infty}^{\infty} (C_n - iS_n) \exp(i\alpha_n) + (C_n + iS_n) \exp(-i\alpha_n) \quad (63)$$

Osservo che, in base alle proprietà di parità 44 e 53, si ha:

$$C_n + iS_n = C_{-n} - iS_{-n} \quad (64)$$

e quindi continuiamo con:

$$= \frac{1}{4} \sum_{n=-\infty}^{\infty} (C_n - iS_n) \exp(i\alpha_n) + (C_{-n} - iS_{-n}) \exp(i\alpha_{-n}) \quad (65)$$

(notate che ho anche riscritto: $\exp(-i\alpha_n) = \exp(i\alpha_{-n})$)

Ora realizzate che, siccome la somma va da $-\infty$ a ∞ , posso cambiare tutti i $-n$ in n nel secondo addendo senza produrre alcun cambiamento: cambia solo l'ordine con cui i corrispondenti termini vengono sommati.

Quindi:

$$= \frac{1}{4} \sum_{n=-\infty}^{\infty} (C_n - iS_n) \exp(i\alpha_n) + (C_n - iS_n) \exp(i\alpha_n) \quad (66)$$

$$= \frac{1}{2} \sum_{n=-\infty}^{\infty} (C_n - iS_n) \exp(i\alpha_n) \quad (67)$$

$$= \sum_{n=-\infty}^{\infty} B_n \exp(i\alpha_n) \quad (68)$$

con:

$$B_n = \frac{1}{2} (C_n - iS_n) \quad (69)$$

$$= \frac{1}{2} \left(\frac{2}{T} \int_a^{a+T} f(x) \cos(\alpha_n) dx - i \frac{2}{T} \int_a^{a+T} f(x) \sin(\alpha_n) dx \right) \quad (70)$$

$$= \frac{1}{T} \int_a^{a+T} f(x) (\cos(\alpha_n) - i \sin(\alpha_n)) dx \quad (71)$$

$$= \frac{1}{T} \int_a^{a+T} f(x) \exp(-i\alpha_n) dx \quad (72)$$

In definitiva, la forma esponenziale della serie di Fourier è:

$$f(x) = \sum_{n=-\infty}^{\infty} B_n \exp(i\alpha_n) \quad (73)$$

$$B_n = \frac{1}{T} \int_a^{a+T} f(x) \exp(-i\alpha_n) dx \quad (74)$$

24. Se ora ritornate alla 63 e invece di operare sul secondo addendo, riscrivete il primo addendo come:

$$(C_n - iS_n) \exp(i\alpha_n) = (C_{-n} + iS_{-n}) \exp(-i\alpha_{-n}) \quad (75)$$

(sfruttando sempre le proprietà di parità dei coefficienti C_n e S_n) arrivate facilmente alla forma equivalente:

$$f(x) = \sum_{n=-\infty}^{\infty} B_n \exp(-i\alpha_n) \quad (76)$$

$$B_n = \frac{1}{T} \int_a^{a+T} f(x) \exp(i\alpha_n) dx \quad (77)$$

il che mostra che la forma esponenziale della serie di Fourier mantiene inalterata la sua validità quando si scambia il segno meno fra l'esponenziale nella sommatoria e quello nella definizione dei coefficienti B_n (confrontate le 76 e 77 con le 73 e 74).

25. Potreste a questo punto chiedervi come una serie che contiene termini complessi possa convergere a una funzione che finora abbiamo supposto reale.

Una prima risposta (ovvia) a questa domanda è fornita dal procedimento che abbiamo usato per ricavare la forma esponenziale complessa della serie di Fourier: il punto di partenza era la 34, che è chiaramente un'espressione reale (non complessa), e quindi rifacendo tutti i passaggi a ritroso ovviamente si deve tornare ad essa. Il che significa che se una funzione è reale, il suo sviluppo di Fourier, anche quando è scritto in forma complessa, deve necessariamente fornire un risultato reale. Vedremo questo fatto in modo formale fra non molto, alla luce di una proprietà della serie di Fourier che proveremo.

Si può inoltre dimostrare (ma noi non lo faremo) che la forma esponenziale complessa della serie di Fourier funziona perfettamente anche per funzioni complesse (di variabile reale): chiaramente, in questo caso, la parte immaginaria dello sviluppo non sarà nulla.

Serie di Fourier in più dimensioni

26. Ora dobbiamo estendere quanto abbiamo visto finora sullo sviluppo in serie di Fourier a funzioni definite nello spazio tridimensionale, cioè funzioni delle tre variabili x, y, z .

Come vedrete, l'estensione delle serie di Fourier a due o tre dimensioni è estremamente elegante, nel senso che, da un punto di vista formale, le espressioni che abbiamo ricavato per il caso monodimensionale rimangono praticamente inalterate, a parte dei cambiamenti dovuti all'aumentata dimensionalità.

27. Non ricaveremo la serie di Fourier in due o tre dimensioni con il dettaglio con cui l'abbiamo ricavata nel caso monodimensionale. Tuttavia, l'estensione di quanto abbiamo trovato in una sola dimensione richiede l'introduzione di alcuni concetti che saranno di importanza centrale per tutto il resto di questo corso, perché costituiscono uno strumento teorico fondamentale con cui studiare non solo gli sviluppi di Fourier in più dimensioni, ma anche una vasta classe di sistemi solidi normalmente chiamati *solidi cristallini*.

28. In questa parte introdurre' i concetti di *reticolo di Bravais* o *reticolo diretto* e *reticolo (o spazio) reciproco*. Ho gia' accennato al reticolo diretto nel punto 5 per una funzione periodica di una sola variabile; in quel caso, tuttavia, era difficile cogliere il senso concreto dell'affermazione: anche nel linguaggio comune, la parola *rete* o *reticolo* evoca in ogni caso un oggetto (come minimo) bidimensionale. Il reticolo monodimensionale a cui ho fatto cenno nel punto 5 e' perfettamente definito matematicamente, ma non ci dice molto da un punto di vista intuitivo.

Vedrete ora come il concetto di reticolo diventi molto piu' concreto in due o tre dimensioni.

29. Notate che in quasi tutti i testi reticolo diretto e reticolo reciproco vengono introdotti come strumenti per la descrizione e lo studio dei solidi cristallini. Il che ovviamente e' assolutamente lecito e corretto, visto che lo studio di quei sistemi non puo' prescindere da una presentazione dettagliata di questi concetti. Tuttavia, a me sembra una buona idea quella di introdurre la definizione dei due reticoli partendo dalla serie di Fourier (in piu' dimensioni) perche' in tal modo, come spero, la definizione del reticolo reciproco non sembra essere "calata dall'alto".

Voglio dire che il concetto di un reticolo reciproco sempre associato a un reticolo diretto, ancorche' di importanza fondamentale nello studio dei solidi cristallini, non e' assolutamente peculiare di quell'argomento, ma ha uno spettro di applicazioni estremamente vasto e variegato, di cui l'analisi di Fourier e' un esempio (e, in effetti, l'analisi di Fourier costituisce un pilastro concettuale dello studio dei solidi cristallini).

Posso ancora aggiungere che in almeno un testo, quello di Kittel,⁴ il reticolo reciproco viene introdotto a partire dagli sviluppi in serie di Fourier.

30. In questa parte faro' solo una prima introduzione, per poi ritornare con ulteriori approfondimenti nella parte riguardante la descrizione dei solidi cristallini.

31. Per cominciare, riprendiamo la forma esponenziale dello sviluppo in serie di Fourier di una funzione periodica 1D (eq. 73 e 74):

$$f(x) = \sum_{n=-\infty}^{\infty} B_n \exp\left(i\frac{2\pi n}{T}x\right)$$

$$B_n = \frac{1}{T} \int_a^{a+T} f(x) \exp\left(-i\frac{2\pi n}{T}x\right) dx$$

Come detto al punto 5, essendo periodica di periodo T , alla funzione e' associato un reticolo diretto (monodimensionale) la cui cella primitiva e' il segmento di lunghezza T (figura 2).

Ora, se guardate lo sviluppo in serie di Fourier della f vi rendete conto che esso e' completamente determinato dagli *infiniti valori equispaziati* $n(2\pi/T)$:

$$\dots -2\left(\frac{2\pi}{T}\right) \quad -1\left(\frac{2\pi}{T}\right) \quad 0\left(\frac{2\pi}{T}\right) \quad 1\left(\frac{2\pi}{T}\right) \quad 2\left(\frac{2\pi}{T}\right) \quad \dots \quad (78)$$

Evidentemente, questi valori definiscono un altro "reticolo" (in questo caso anch'esso monodimensionale) che chiamiamo *reticolo (o spazio) reciproco* (figura 3). La cella primitiva di questo reticolo reciproco e' il segmento di lunghezza $2\pi/T$.

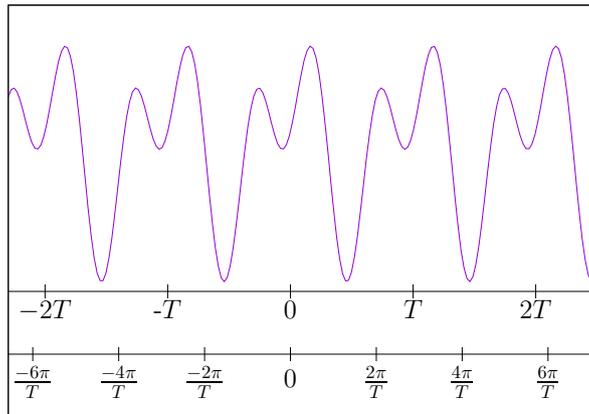


Figura 3:

E' anche evidente che esiste una corrispondenza biunivoca fra i punti del reticolo reciproco e i termini dello sviluppo di Fourier della funzione f . Ad esempio il punto $7 \times 2\pi/T$, individua univocamente il termine dello sviluppo in serie:

$$B_7 \exp\left(i \frac{2\pi}{T} 7 x\right) \quad (79)$$

con:

$$B_7 = \frac{1}{T} \int_a^{a+T} f(x) \exp\left(-i \frac{2\pi}{T} 7 x\right) dx \quad (80)$$

Una prima cosa da osservare e' che il reticolo reciproco ora introdotto e' definito su un asse la cui unita' di misura e' il reciproco dell'unita' di misura del periodo della funzione: se la variabile da cui dipende la funzione periodica $f(x)$ e' la distanza lineare e il suo periodo T e' misurato, ad esempio, in m, allora l'unita' di misura del reticolo reciproco ad essa associato sara' m^{-1} (cioe', per intenderci, la "lunghezza" della cella primitiva del reticolo reciproco sara' espressa in m^{-1}). Per inciso, l'aggettivo "reciproco" fa riferimento proprio a questa e a molte altre "relazioni di reciprocita'" che esistono fra i due reticoli e che metteremo in evidenza in piu' occasioni.

32. Da quanto detto segue che ad ogni funzione periodica sono sempre associati due reticoli:

il reticolo diretto i cui punti sono definiti da:

$$nT, n \in \mathbb{Z} \quad (81)$$

sull'asse che rappresenta la variabile da cui la funzione periodica dipende

il reticolo reciproco i cui punti sono definiti da:

$$n \frac{2\pi}{T}, n \in \mathbb{Z} \quad (82)$$

su un asse la cui unita' di misura e' il reciproco di quella dell'asse del reticolo diretto

E quindi, un modo importante e utile di interpretare il significato del reticolo reciproco e' il seguente:

Ad ogni funzione periodica e' associato un reticolo reciproco e ogni punto di tale reticolo fornisce un corrispondente termine dello sviluppo in serie di Fourier della funzione

Per questo motivo lo spazio reciproco viene anche chiamato *spazio di Fourier*.

33. Se indichiamo con k un generico punto del reticolo reciproco associato alla funzione periodica (di periodo T) $f(x)$ cioe':

$$k = \frac{2\pi n}{T} \quad (83)$$

allora possiamo scrivere il suo sviluppo in serie di Fourier come:

$$f(x) = \sum_k B_k \exp(ikx) \quad (84)$$

$$B_k = \frac{1}{T} \int_a^{a+T} f(x) \exp(-ikx) dx \quad (85)$$

dove la somma e' estesa a tutti gli infiniti punti k del reticolo reciproco.

Notate che k definito dalla 83 e' un multiplo di quella che abbiamo chiamato frequenza angolare al punto 7.

34. Quanto detto finora e' basato su una funzione di una sola variabile $f(x)$. Vediamo ora come cio' si possa estendere al caso multidimensionale.

Aperta parentesi: notazione.

Dovendo trattare funzioni definite in 2D o 3D (e non solo in questa parte del corso), useremo molto spesso il formalismo vettoriale, di cui assumo che conosciate gli aspetti elementari.

Rappresentero' una funzione f delle tre variabili x, y, z indifferentemente con:

$$f(x, y, z) \quad \text{oppure} \quad f(\vec{r}) \quad (86)$$

dove $\vec{r} = (x, y, z)$ e' il vettore le cui componenti sono (x, y, z) . Quindi \vec{r} parte dall'origine e arriva al punto di coordinate (x, y, z) . \vec{r} viene anche detto *vettore posizione*.

Ovviamente, usero' una notazione identica anche per funzioni di due sole variabili x, y :

$$f(x, y) \quad \text{oppure} \quad f(\vec{r}) \quad (87)$$

con $\vec{r} = (x, y)$.

Chiusa parentesi.

35. Conviene introdurre l'argomento considerando una funzione periodica di due variabili $f(x, y)$. Questo perche' il grafico di una funzione di due variabili si puo' rappresentare facilmente, mentre quello di una funzione di tre variabili no.

Un punto importante da comprendere nel passaggio da una a due (o piu') dimensioni riguarda il concetto di periodicita': cosa vuol dire che una funzione di due variabili e' periodica?

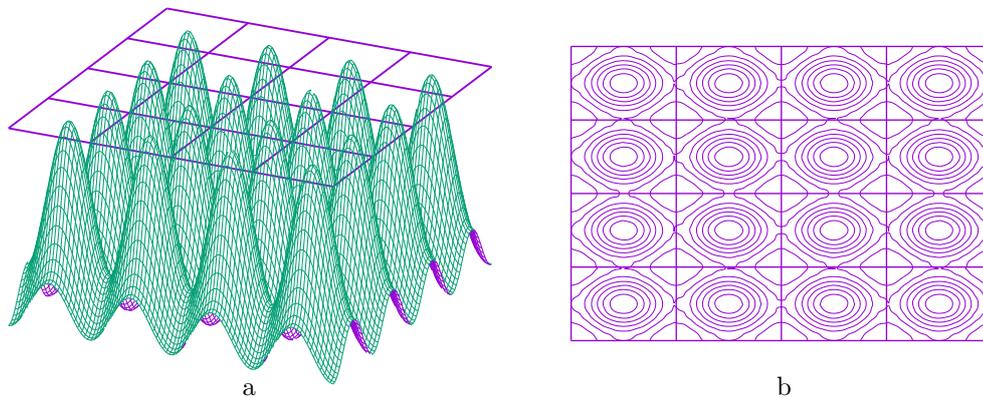


Figura 4:

Il dominio di una funzione di due variabili e' il piano cartesiano. Se una funzione periodica di una sola variabile $f(x)$ si ripete identica a se stessa a *intervalli* (intendo segmenti *lineari*, cioe' oggetti 1D) di ampiezza T sull'asse x , una funzione periodica di due variabili $f(x, y)$ si ripete identicamente su delle *regioni* (cioe' oggetti 2D) del piano xy .

Inoltre, senza dare di cio' nessuna prova formale, dovrebbe essere abbastanza intuitivo che tali regioni, come era il caso per gli intervalli di ampiezza T nel caso monodimensionale, devono essere *tutte uguali* e devono essere *dei parallelogrammi*.

In altre parole: una funzione periodica $f(x, y)$ individua sul piano un reticolo costituito da infinite rette le cui intersezioni determinano celle primitive identiche.

In ciascuna cella primitiva il grafico della funzione e' sempre lo stesso.

Il reticolo individuato dalla funzione periodica (e questa volta si tratta proprio di un reticolo, nel senso comune del termine) e' il reticolo diretto o reticolo di Bravais.

Nella figura 4a e' mostrato il grafico di una funzione periodica di due variabili ed e' evidenziato il reticolo diretto da essa individuato.

Nella figura 4b lo stesso grafico e' rappresentato con delle curve di livello o isocurve (cioe' curve che sono luoghi dei punti del piano xy su cui la funzione assume uno stesso valore): con questa rappresentazione si mette molto meglio in evidenza il reticolo di Bravais che definisce la periodicit  della funzione.

36. Dalla figura 4 appare evidente che la cella primitiva del reticolo diretto associato alla funzione e' rettangolare e i suoi lati sono paralleli agli assi cartesiani (voglio sottolineare che non e' quadrata e che gli angoli interni sono tutti retti: torneremo fra poco su questo punto).

Come si puo' esprimere matematicamente la periodicit  della funzione?

Non dovrebbe essere difficile comprendere che, se indichiamo con T_x e T_y le lunghezze dei lati della cella primitiva, allora possiamo esprimere il fatto che la funzione $f(x, y)$ e' periodica sul reticolo di Bravais mostrato nella figura con una condizione che e' l'estensione intuitiva della condizione vista per una funzione di una sola variabile (eq. 1):

$$f(x + n_x T_x, y + n_y T_y) = f(x, y) \quad \forall x, y \quad \forall n_x, n_y \in \mathbb{Z} \quad (88)$$

In altre parole (guardate la figura 4): se fisso un punto qualsiasi (x, y) in cui la funzione ha un certo valore e mi sposto a destra (o a sinistra) di un *multiplo intero*⁸ del lato di cella T_x e in alto (o in basso) di un *multiplo intero* del lato di cella T_y , il valore della funzione sul punto così determinato è lo stesso. E questo vale per *qualsiasi* punto di partenza e per multipli *qualsiasi* dei due lati di cella T_x e T_y .

37. Come sarà lo sviluppo in serie di Fourier della funzione $f(x, y)$?

Vi (e mi) risparmio la derivazione. Il risultato finale, come vi avevo preannunciato, è un'estensione molto intuitiva di quello monodimensionale:

$$f(x, y) = \sum_{n_x, n_y = -\infty}^{\infty} B_{n_x, n_y} \exp \left(i \left(\frac{2\pi n_x}{T_x} x + \frac{2\pi n_y}{T_y} y \right) \right) \quad (89)$$

con:

$$B_{n_x, n_y} = \frac{1}{T_x T_y} \int_a^{a+T_x} \int_a^{a+T_y} f(x, y) \exp \left(-i \left(\frac{2\pi n_x}{T_x} x + \frac{2\pi n_y}{T_y} y \right) \right) dx dy \quad (90)$$

Le modifiche dovute all'estensione in 2D sono le seguenti:

- ⇒ la somma è fatta su due indici
- ⇒ i coefficienti B dipendono da due numeri interi
- ⇒ gli esponenziali contengono ora la somma di due termini, uno per x e uno per y
- ⇒ l'integrale è diventato un integrale doppio
- ⇒ il fattore che moltiplica l'integrale contiene il prodotto delle dimensioni della cella primitiva (i due "periodi" T_x e T_y), cioè la sua area

Se confrontate le 89 e 90 con le 73 e 74 dovrete cogliere senz'altro la stretta analogia.

38. E come sarà il reticolo reciproco associato al reticolo diretto della $f(x, y)$ in questo caso?

Questo è abbastanza facile da vedere. Nel caso 2D ci sono *due* serie (infinite) di valori equispaziati:

$$\frac{2\pi n_x}{T_x} \quad n_x \in [-\infty, \infty] \quad (91)$$

$$\frac{2\pi n_y}{T_y} \quad n_y \in [-\infty, \infty] \quad (92)$$

che individuano il reticolo reciproco su due assi ortogonali le cui unità di misura sono il reciproco di quelle dei corrispondenti assi del reticolo diretto (figura 5). La cella primitiva del reticolo reciproco è anch'essa rettangolare, ma le lunghezze relative dei due lati sono invertite rispetto a quelle della cella primitiva del reticolo diretto: il lato più lungo della cella diretta è diventato il lato più corto della cella reciproca e viceversa (un'altra "relazione di reciprocità").

Il generico punto dello spazio reciproco:

⁸In tutto il corso parlerò di numeri interi intendendo anche numeri *relativi*, il preciso significato essendo sempre chiaro dal contesto.

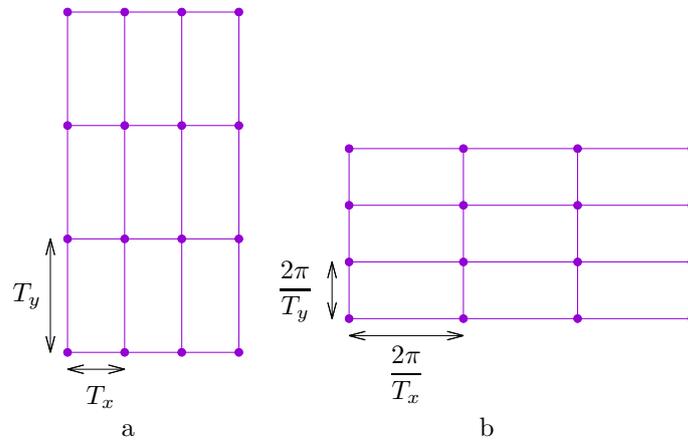


Figura 5: a: reticolo diretto di una funzione periodica 2D con periodi T_x e T_y ; b: reticolo reciproco associato al reticolo diretto in a.

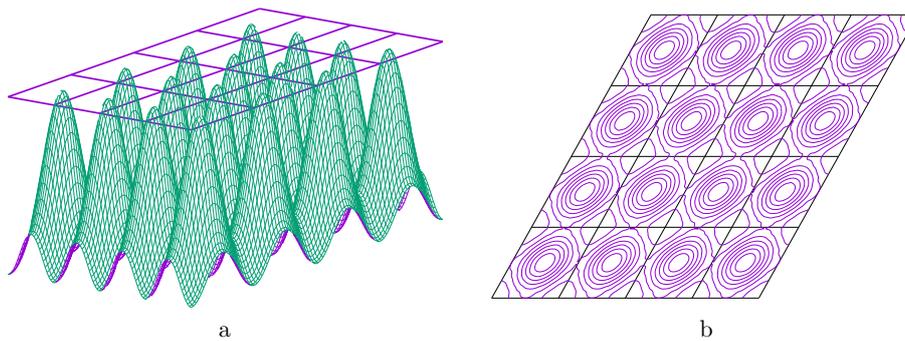


Figura 6:

$$\left(\frac{2\pi}{T_x} n_x, \frac{2\pi}{T_y} n_y \right) \quad (93)$$

individua un corrispondente termine dello sviluppo in serie di Fourier 89 della $f(x, y)$.

39. Ora però arriva una doccia fredda: purtroppo, quanto appena detto *non funziona in generale*.

La validità delle 88, 89 e 90 è ristretta al caso in cui i lati della cella primitiva del reticolo diretto sono fra loro ortogonali e paralleli agli assi x e y .

Ma questo non è sempre il caso.

Nella figura 6a ho costruito il grafico di una funzione periodica bidimensionale alla quale è associato un reticolo di Bravais la cui cella primitiva *non* è rettangolare, ma è un parallelogramma con angoli interni *non* retti. Questo si vede particolarmente bene nella figura 6b, dove il grafico della f è rappresentato con delle isocurve.

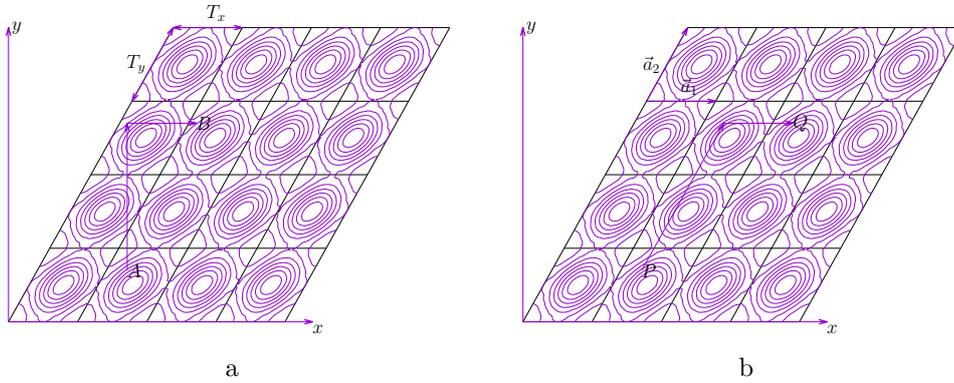


Figura 7: a: il punto A ha coordinate (x, y) e il punto B ha coordinate $(x + T_x, y + 2T_y)$. Siccome la cella primitiva del reticolo di Bravais non e' "retta", si ha: $f(x + T_x, y + 2T_y) \neq f(x, y)$. b: il punto P ha coordinate $(x, y) = \vec{r}$ e il punto Q ha coordinate $\vec{r} + \vec{a}_1 + 2\vec{a}_2$. In questo caso e' vero che: $f(\vec{r} + \vec{a}_1 + 2\vec{a}_2) = f(\vec{r})$.

Dovrebbe essere abbastanza semplice rendersi conto che in questo caso la 88 non vale.

Inoltre (anche se rendersene conto e' meno semplice) anche la validita' delle 89 e 90 (che dipendono dalla 88) viene meno.

Notate che un problema analogo si avrebbe anche se, pur essendo fra loro ortogonali, i lati della cella primitiva non fossero paralleli agli assi x e y del sistema di riferimento.

40. Per poter riformulare la serie di Fourier in piu' dimensioni, in un modo conciso, utile e veramente generale bisogna usare l'algebra vettoriale.

Cerchiamo di catturare l'idea in modo qualitativo e intuitivo.

Guardate la figura 7.

La funzione $f(x, y)$ e' periodica su un reticolo di Bravais la cui cella primitiva e', in generale, un parallelogramma qualsiasi. Per questo motivo, la periodicit  della funzione non puo' essere espressa *semplicemente* in base alle lunghezze *scalari* dei lati della cella primitiva (figura 7a). In altre parole la relazione 88:

$$f(x + n_x T_x, y + n_y T_y) = f(x, y) \quad \forall x, y \quad \forall n_x, n_y \in \mathbb{Z}$$

dove T_x e T_y sono le *lunghezze scalari* dei lati della cella primitiva, *non vale in generale*. Essa puo' funzionare *solo* se i lati della cella primitiva del reticolo di Bravais sono paralleli agli assi cartesiani x e y (e quindi fra loro ortogonali).

Tuttavia, se consideriamo i due *vettori* \vec{a}_1 e \vec{a}_2 che costituiscono i due lati "indipendenti" (cioe' non opposti) della cella primitiva (figura 7b), allora la periodicit  della f puo' essere espressa *in forma vettoriale* nel modo seguente:

$$f(\vec{r} + n_1 \vec{a}_1 + n_2 \vec{a}_2) = f(\vec{r}) \quad \forall \vec{r} \in \mathbb{R}^2 \text{ e } \forall n_1, n_2 \in \mathbb{Z} \quad (94)$$

ovvero:

$$f(\vec{r} + \vec{R}) = f(\vec{r}) \quad \forall \vec{r} \in \mathbb{R}^2 \text{ e } \forall \vec{R} \in \text{al reticolo di Bravais} \quad (95)$$

dove:

$$\vec{R} = n_1 \vec{a}_1 + n_2 \vec{a}_2 \quad \forall n_1, n_2 \in \mathbb{Z} \quad (96)$$

e' un generico punto del reticolo di Bravais.

In parole povere la 95 dice la cosa seguente: dato un punto del piano sul quale la funzione assume un certo valore, essa assumerà quello stesso valore su tutti i punti raggiungibili dal punto considerato sommando a quest'ultimo un vettore qualsiasi del reticolo di Bravais (dato dalla combinazione lineare dei vettori che definiscono la cella primitiva *con coefficienti interi*, eq. 96).

41. In algebre questo si chiama "cambiamento di base".

I punti del piano costituiscono uno spazio vettoriale di dimensione 2. Ogni punto del piano e' un elemento, cioe' un vettore, di questo spazio \mathbb{R}^2 . Tutti i vettori di \mathbb{R}^2 possono essere espressi come una combinazione lineare di due di essi che devono essere linearmente indipendenti⁹ e formano una cosiddetta *base*.

Per la rappresentazione di un qualsiasi elemento di \mathbb{R}^2 , non c'e' alcuna limitazione sulla scelta della base: qualsiasi coppia di vettori linearmente indipendenti puo' essere usata a questo scopo.

Ad esempio, in un sistema di riferimento cartesiano, la base del piano e' anche *ortonormale*, cioe' e' costituita da due vettori che, oltre ad essere linearmente indipendenti, sono anche fra loro ortogonali e hanno modulo unitario. Questi due vettori sono usualmente:

$$\hat{e}_1 = (1, 0) \text{ e } \hat{e}_2 = (0, 1) \quad (97)$$

La rappresentazione di un punto del piano \vec{r} rispetto alla base \hat{e}_1, \hat{e}_2 e':

$$\vec{r} = x \hat{e}_1 + y \hat{e}_2 \quad (98)$$

I coefficienti (x, y) della combinazione lineare vengono detti *coordinate* del punto e sono le usuali coordinate che tutti conoscete.

Il significato algebrico di quello che ho detto al punto 40 e' il seguente.

Le coordinate di un generico punto del reticolo di Bravais rispetto alla base "convenzionale" \hat{e}_1, \hat{e}_2 :

$$\vec{R} = x \hat{e}_1 + y \hat{e}_2 \quad x, y \in \mathbb{R} \quad \Leftarrow \text{NOTA} \quad (99)$$

non sono, in generale, numeri interi.

Tuttavia, se rappresentiamo i punti del reticolo diretto rispetto alla base costituita dai due vettori \vec{a}_1, \vec{a}_2 della cella primitiva:

$$\vec{R} = n_1 \vec{a}_1 + n_2 \vec{a}_2 \quad n_x, n_y \in \mathbb{Z} \quad \Leftarrow \text{NOTA} \quad (100)$$

allora i coefficienti della combinazione lineare *sono sempre* numeri interi e il vettore \vec{R} definito dalla 100 (ma non quello definito dalla 99!) puo' essere usato per esprimere matematicamente il fatto che la funzione $f(\vec{r})$ e' periodica sul reticolo di Bravais (eq. 95).

42. Notate che, ovviamente, la 88 e' un caso particolare della 95, quando \vec{a}_1 e \vec{a}_2 sono paralleli, rispettivamente, a \hat{e}_1 e \hat{e}_2 . In tal caso, infatti, la condizione di periodicit  diventa:

⁹Linearmente indipendenti vuol dire che se una loro combinazione lineare e' *nulla*, allora necessariamente i coefficienti di tale combinazione lineare devono essere *nulli*.

$$\begin{aligned}
f(\vec{r} + \vec{R}) &= f((x, y) + (n_1\vec{a}_1 + n_2\vec{a}_2)) \\
&= f((x, y) + (n_1T_x\hat{e}_1 + n_2T_y\hat{e}_2)) = f((x, y) + (n_1T_x(1, 0) + n_2T_y(0, 1))) \\
&= f((x, y) + ((n_1T_x, 0) + (n_2T_y, 1))) = f((x, y) + (n_1T_x, n_2T_y)) \\
&= f(x + n_1T_x, y + n_2T_y) = f(x, y)
\end{aligned}$$

che e' la condizione 88.

(sopra ho usato il fatto che nel caso della 88 T_x e' il modulo di \vec{a}_1 e T_y quello di \vec{a}_2)

43. Ci resta ancora da vedere un'espressione generale della serie di Fourier, valida per una qualsiasi cella primitiva (cioe' non soltanto rettangolare) del reticolo di Bravais sul quale la funzione f e' periodica.

Per ottenerla dobbiamo introdurre una definizione generale del reticolo reciproco: dovrebbe esservi chiaro a questo punto che quella vista al punto 38 non funziona quando la cella primitiva del reticolo di Bravais e' un parallelogramma generico, con angoli non retti.

Derivare questa definizione generale non e' semplice e noi non lo faremo; mi limitero' (come fanno tutti i testi di mia conoscenza) a darvela (cosi', calata dall'alto, sembra piuttosto astrusa, ma vedrete come grazie ad essa la serie di Fourier assume una forma estremamente concisa ed elegante).

Consideriamo un reticolo di Bravais la cui cella primitiva e' definita dai due vettori (non collineari) \vec{a}_1, \vec{a}_2 .

Introduciamo un terzo vettore \vec{a}_3 ortogonale al piano individuato da \vec{a}_1 e \vec{a}_2 . Come vedrete fra un attimo, questo terzo vettore e' necessario per la definizione che daremo: nel caso 2D, oltre ad essere necessario come vedremo fra un attimo, esso aggiunge un tocco di esoterismo (cioe': siamo nel piano, ma tiriamo in ballo un vettore che *non* appartiene al piano), mentre vedrete che nel caso 3D esso e' semplicemente il terzo vettore che definisce la cella primitiva (come potete intuire, la cella primitiva di un reticolo di Bravais tridimensionale e' definita da tre (non due) vettori non coplanari). Dovendo essere ortogonale al piano individuato da \vec{a}_1 e \vec{a}_2 , \vec{a}_3 e' dato dal loro prodotto vettoriale (cross product):

$$\vec{a}_3 = \vec{a}_1 \times \vec{a}_2 \quad (101)$$

Come vedrete fra un attimo, nel caso 2D la definizione del reticolo reciproco e' del tutto indipendente dal modulo di \vec{a}_3 .

44. Ora viene il woodoo.

Il reticolo reciproco associato al reticolo di Bravais dato e' definito dalla cella primitiva i cui vettori \vec{b}_1, \vec{b}_2 sono ottenuti nel modo seguente:

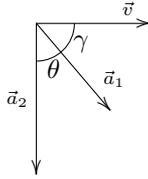
$$\vec{b}_1 = 2\pi \frac{\vec{a}_2 \times \vec{a}_3}{\vec{a}_1 \cdot \vec{a}_2 \times \vec{a}_3} \quad (102)$$

$$\vec{b}_2 = 2\pi \frac{\vec{a}_3 \times \vec{a}_1}{\vec{a}_1 \cdot \vec{a}_2 \times \vec{a}_3} \quad (103)$$

(\times indica il prodotto vettoriale, \cdot quello scalare)

Nonostante possa sembrare involuta e/o astratta, la definizione appena introdotta ha un'interpretazione geometrica molto semplice.

Innanzitutto osservate che l'espressione ai denominatori (che sono uguali) e' una quantita' scalare (e' il prodotto scalare di un vettore, \vec{a}_1 , per un altro vettore, $\vec{a}_2 \times \vec{a}_3$). Il suo valore e' ¹⁰ $a_1 a_2 \sin(\theta) a_3 = A a_3$, essendo θ l'angolo ($< \pi$ ¹¹) fra \vec{a}_1 e \vec{a}_2 e A l'area della cella primitiva del reticolo diretto. Infatti, siccome, per definizione, \vec{a}_3 e' ortogonale al piano individuato da \vec{a}_1, \vec{a}_2 , il vettore $\vec{v} = \vec{a}_2 \times \vec{a}_3$ (i) giace sul piano individuato da \vec{a}_1, \vec{a}_2 , (ii) e' perpendicolare a \vec{a}_2 e (iii) ha modulo $a_2 a_3$. Allora, detto $\gamma = 90 - \theta$ l'angolo fra \vec{v} e \vec{a}_1 ¹², si ha:



$$\begin{aligned} \vec{a}_1 \cdot \vec{a}_2 \times \vec{a}_3 &= \vec{a}_1 \cdot \vec{v} = a_1 v \cos(\gamma) = a_1 a_2 a_3 \cos(90 - \theta) \\ &= a_1 a_2 \sin(\theta) a_3 = A a_3 \end{aligned}$$

Al numeratore della 102 c'e' un vettore che giace sul piano individuato da \vec{a}_1, \vec{a}_2 , e' ortogonale a \vec{a}_2 e ha modulo $a_2 a_3$; analogamente, al numeratore della 103 c'e' un vettore che giace sul piano individuato da \vec{a}_1, \vec{a}_2 , e' ortogonale a \vec{a}_1 e ha modulo $a_1 a_3$.

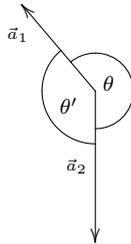
In definitiva, la ricetta "pratica" per costruire e visualizzare la cella primitiva reciproca a partire da quella diretta si riduce alla seguente (guardate la figura 8):

\Rightarrow \vec{b}_1 e' ottenuto tracciando sul piano individuato da \vec{a}_1 e \vec{a}_2 un vettore ortogonale a \vec{a}_2 di modulo:

$$b_1 = 2\pi \frac{a_2 a_3}{A a_3} = 2\pi \frac{a_2}{A} \quad (104)$$

¹⁰In tutto il corso indichero' un vettore con la freccia, ad esempio \vec{a} , e il suo modulo, sempre > 0 , col simbolo *senza* freccia, ad esempio a .

¹¹Non e' restrittivo assumere che l'angolo fra \vec{a}_1 e \vec{a}_2 sia sempre $< \pi$: se fosse $> \pi$ non si potrebbe costruire nessuna cella primitiva e si dovrebbe considerare come angolo fra \vec{a}_1 e \vec{a}_2 il complemento a 2π :



¹²Verificate da soli che il risultato non cambia se l'angolo θ fra \vec{a}_1 e \vec{a}_2 e' $> \pi/2$. In questo caso l'angolo γ fra \vec{v} e \vec{a}_1 e' $\gamma = \theta - \pi/2 > 0$ ed e' sempre vero che $\cos(\theta - \pi/2) = \sin(\theta)$. Verificate anche che il risultato e' invariante rispetto a tutte le possibili orientazioni relative di \vec{a}_1 e \vec{a}_2 .

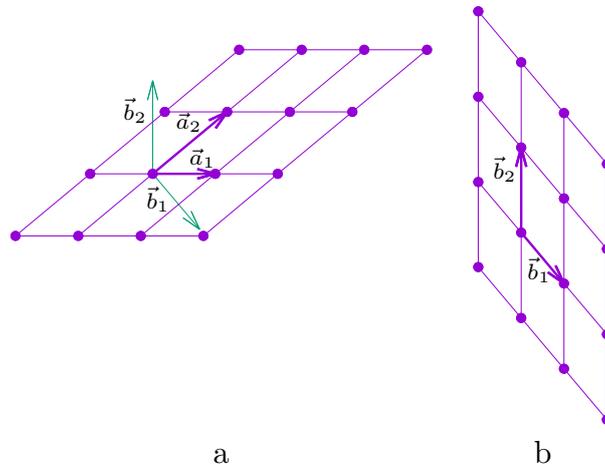


Figura 8: a: costruzione dei vettori della cella reciproca: \vec{b}_1 e' ortogonale a \vec{a}_2 e \vec{b}_2 e' ortogonale a \vec{a}_1 ; b: il reticolo reciproco associato al reticolo diretto in (a) e originato dai vettori \vec{b}_1 e \vec{b}_2 .

\Rightarrow \vec{b}_2 e' ottenuto tracciando sul piano individuato da \vec{a}_1 e \vec{a}_2 un vettore ortogonale a \vec{a}_1 di modulo:

$$b_2 = 2\pi \frac{a_1 a_3}{A a_3} = 2\pi \frac{a_1}{A} \quad (105)$$

Infine, l'intero reticolo reciproco e' costituito da tutti i punti \vec{k} che si ottengono dalle infinite combinazioni lineari:

$$\vec{k} = n_1 \vec{b}_1 + n_2 \vec{b}_2 \quad n_1, n_2 \in \mathbb{Z} \quad (106)$$

Come avevo preannunciato, nel caso 2D il vettore \vec{a}_3 scompare nel risultato finale (esso serve solo per determinare la direzione di \vec{b}_1 e \vec{b}_2 sul piano).

Potete verificare facilmente da soli che la definizione generale del reticolo reciproco appena introdotta (equazioni 102 e 103) si riduce alla 93 quando \vec{a}_1 e \vec{a}_2 sono paralleli agli assi cartesiani (e hanno modulo, rispettivamente, T_x e T_y).

Notate ancora che, come avevamo gia' osservato per il caso monodimensionale, l'unita' di misura dei vettori reciproci e' il reciproco di quella dei vettori diretti (nelle 104 e 105, detta q la generica unita' di misura del modulo di \vec{a}_i ($i = 1, 2$), l'unita' q al numeratore viene divisa per A , le cui dimensioni sono q^2).

45. Possiamo infine vedere l'espressione dello sviluppo di Fourier per una funzione periodica di due variabili.

Alla luce della definizione del reticolo reciproco data dalle equazioni 102–103, si puo' dimostrare che la funzione periodica $f(x, y) = f(\vec{r})$ puo' essere espressa con la serie di Fourier:

$$f(\vec{r}) = \sum_{\vec{k}} B_{\vec{k}} \exp(i\vec{k} \cdot \vec{r}) \quad (107)$$

dove la somma e' estesa a tutti i punti \vec{k} del reticolo reciproco e i coefficienti $B_{\vec{k}}$ sono dati da ¹³:

$$B_{\vec{k}} = \frac{1}{A} \int_A f(\vec{r}) \exp(-i\vec{k} \cdot \vec{r}) d\vec{r} \quad (108)$$

dove l'integrale (doppio) e' calcolato sull'area della cella primitiva del reticolo di Bravais.

Notate che usare il vettore reciproco \vec{k} come indice e' in realta' un modo per abbreviare l'indicazione dei 2 indici interi che sono i coefficienti della combinazione lineare dei vettori della cella reciproca con cui il vettore \vec{k} e' ottenuto (guardate la 106).

Voglio dire che la 107 e' una forma compatta di:

$$f(\vec{r}) = \sum_{\substack{n_1=-\infty \\ n_2=-\infty}}^{\infty} B_{n_1, n_2} \exp\left(i\left(n_1\vec{b}_1 + n_2\vec{b}_2\right) \cdot \vec{r}\right) \quad (109)$$

Useremo la forma esplicita del vettore reciproco \vec{k} quando sara' richiesto dalle circostanze.

46. Come gia' accennato a proposito della definizione generale del reticolo reciproco al punto 44, potete verificare facilmente da soli che le 107 e 108 si riducono alle 89 e 90 quando \vec{a}_1 e \vec{a}_2 sono paralleli agli assi cartesiani (e hanno modulo, rispettivamente, T_x e T_y).

47. L'estensione a tre dimensioni di quanto appena visto per due non comporta nessuna difficolta' addizionale.

Una funzione $f(x, y, z) = f(\vec{r})$ periodica in 3D individua un reticolo di Bravais tridimensionale, la cui cella primitiva e' definita da tre vettori di cella non coplanari \vec{a}_1 , \vec{a}_2 e \vec{a}_3 (approfondiremo ulteriormente il concetto di reticolo di Bravais nella parte riguardante i solidi cristallini).

La condizione di periodicita' della funzione viene espressa in modo generale da:

$$f(\vec{r} + \vec{R}) = f(\vec{r}) \quad \forall \vec{r} \in \mathbb{R}^3 \text{ e } \forall \vec{R} \in \text{al reticolo di Bravais} \quad (110)$$

con:

$$\vec{R} = n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3 \quad \forall n_1, n_2, n_3 \in \mathbb{Z} \quad (111)$$

¹³Introduco una notazione abbreviata che usero' altre volte e che viene usata praticamente in tutti i testi.

Con:

$$\int_Q \dots d\vec{r}$$

intendero' un integrale doppio o triplo (a seconda della dimensionalita' del sistema considerato) calcolato, rispettivamente, sulla superficie o il volume Q . Il differenziale $d\vec{r}$ e' un'abbreviazione, rispettivamente, per $dx dy$ o $dx dy dz$.

Quindi:

$$\int_Q \dots d\vec{r} \equiv \int_{x_1}^{x_2} \int_{y_1}^{y_2} \dots dx dy$$

oppure:

$$\int_Q \dots d\vec{r} \equiv \int_{x_1}^{x_2} \int_{y_1}^{y_2} \int_{z_1}^{z_2} \dots dx dy dz$$

(confrontate con la 95)

Al reticolo di Bravais della funzione f rimane associato un reticolo reciproco (ovviamente anch'esso tridimensionale) i cui vettori di cella sono definiti da:

$$\vec{b}_1 = 2\pi \frac{\vec{a}_2 \times \vec{a}_3}{\vec{a}_1 \cdot \vec{a}_2 \times \vec{a}_3} \quad (112)$$

$$\vec{b}_2 = 2\pi \frac{\vec{a}_3 \times \vec{a}_1}{\vec{a}_1 \cdot \vec{a}_2 \times \vec{a}_3} \quad (113)$$

$$\vec{b}_3 = 2\pi \frac{\vec{a}_1 \times \vec{a}_2}{\vec{a}_1 \cdot \vec{a}_2 \times \vec{a}_3} \quad (114)$$

(suggerimento: le formule sembrano complicate, ma in realta' memorizzarle non e' difficile: il denominatore e' sempre lo stesso e gli indici sono in sequenza: 1, 2, 3; gli indici di \vec{b}_i al primo membro e degli \vec{a}_i al numeratore sono permutati ciclicamente: (1, 2, 3), (2, 3, 1), (3, 1, 2))

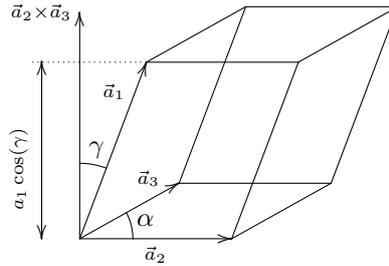
L'intero reticolo reciproco e' costituito da tutti i punti \vec{k} che si ottengono dalle infinite combinazioni lineari:

$$\vec{k} = n_1 \vec{b}_1 + n_2 \vec{b}_2 + n_3 \vec{b}_3 \quad n_1, n_2, n_3 \in \mathbb{Z} \quad (115)$$

Come avevo gia' accennato, vedete che le definizioni di \vec{b}_1 e \vec{b}_2 sono identiche a quelle viste in 2D (eq. 102 e 103); in questo caso il vettore \vec{a}_3 e' un vettore "reale", nel senso che esso e' il terzo vettore che definisce la cella primitiva del reticolo diretto (e quindi, in generale, *non* e' ortogonale a \vec{a}_1 e \vec{a}_2).

Ora la definizione dipende anche dal modulo di \vec{a}_3 . In particolare, potete rendervi conto che il valore assoluto del denominatore (comune) alle tre definizioni e' il volume V della cella primitiva del reticolo diretto.

Infatti, e' sempre possibile disporre la cella primitiva del reticolo diretto con la faccia individuata da \vec{a}_2 e \vec{a}_3 sul piano orizzontale. Allora:



$$|\vec{a}_1 \cdot \vec{a}_2 \times \vec{a}_3| = a_1 |\cos(\gamma)| |\vec{a}_2 \times \vec{a}_3| = \overbrace{a_1 |\cos(\gamma)|}^{\text{altezza}} \overbrace{(a_2 a_3 \sin(\alpha))}^{\text{area di base}} = V \quad (116)$$

Due osservazioni su questa espressione:

- \Rightarrow il valore assoluto per il termine $\cos(\gamma)$ e' necessario perche', se scambiate di posto \vec{a}_2 e \vec{a}_3 , allora il loro prodotto vettoriale inverte il suo verso e quindi l'angolo γ fra \vec{a}_1 e $\vec{a}_2 \times \vec{a}_3$ diventa $> \pi/2$, e quindi $\cos(\gamma) < 0$ (l'altezza del parallelepipedo rimane sempre uguale a $a_1 |\cos(\gamma)|$)

⇒ attenti a non fare confusione: l'espressione e' identica a quella vista al punto 44, ma in quel caso \vec{a}_3 era ortogonale (per definizione) al piano individuato da \vec{a}_1 e \vec{a}_2 e quindi $\alpha = 90 \Rightarrow \sin(\alpha) = 1$

48. L'interpretazione geometrica del reticolo reciproco in 3D e' la logica estensione di quella che abbiamo visto per il caso bidimensionale. Ora il vettore della cella reciproca \vec{b}_i e' ortogonale *al piano* individuato dai due vettori della cella diretta con indici diversi da i (cioe', ad esempio, \vec{b}_1 e' ortogonale al piano individuato da \vec{a}_2 e \vec{a}_3 , e cosi' via). Il modulo di \vec{b}_i e' dato da:

$$b_i = \frac{2\pi}{V} |\vec{a}_j \times \vec{a}_k| = \frac{2\pi}{V} a_j a_k \sin(\theta) \quad (117)$$

dove (i, j, k) sono una permutazione ciclica di $(1, 2, 3)$ e θ e' l'angolo ($< \pi$) fra \vec{a}_j e \vec{a}_k .

Nel caso piu' semplice possibile in cui i vettori primitivi del reticolo diretto sono fra loro ortogonali e hanno lo stesso modulo a , i vettori primitivi del reticolo reciproco saranno anch'essi ortogonali fra loro (con \vec{b}_1 ortogonale al piano $\vec{a}_2\vec{a}_3$ e cosi' via) e avranno tutti lo stesso modulo pari a $2\pi/a$ (un'altra relazione di reciprocita').

49. Dalle definizioni 112–114 segue una importante proprieta' sulle relazioni fra i vettori della cella diretta e quelli della cella reciproca. Vale infatti:

$$\vec{b}_i \cdot \vec{a}_j = 2\pi\delta_{ij} \quad (118)$$

dove δ_{ij} si chiama *delta di Kronecker* ed e' definito da:

$$\delta_{ij} = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases} \quad (119)$$

Le relazioni valide per $i \neq j$ seguono immediatamente dal fatto che \vec{b}_i e' ortogonale al piano individuato da \vec{a}_j e \vec{a}_k ($j, k \neq i$) e quindi ortogonale sia a \vec{a}_j che a \vec{a}_k .

Anche le relazioni valide per $i = j$ si ottengono facilmente. Quella per $i = j = 1$ e' immediata.

Quelle per $i = j = 2, 3$ sfruttano la proprieta' di permutazione ciclica di cui gode il cosiddetto prodotto triplo o prodotto misto: ¹⁴

¹⁴Per verificare la proprieta' di permutazione ciclica tenete presente che, in generale, le coordinate cartesiane del prodotto vettoriale dei due vettori \vec{a}_2 e \vec{a}_3 si ottengono (mnemonicamente) sviluppando in minori 2×2 lungo la prima riga il seguente determinante:

$$\vec{a}_2 \times \vec{a}_3 = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \vec{i}(a_{22}a_{33} - a_{23}a_{32}) - \vec{j}(a_{21}a_{33} - a_{23}a_{31}) + \vec{k}(a_{21}a_{32} - a_{22}a_{31})$$

dove a_{ij} e' la componente j -esima di \vec{a}_i e \vec{i}, \vec{j} e \vec{k} sono i versori dei tre assi cartesiani. Poi, realizzate che il prodotto triplo $\vec{a}_1 \cdot \vec{a}_2 \times \vec{a}_3$ e' uguale al seguente determinante:

$$\vec{a}_1 \cdot \vec{a}_2 \times \vec{a}_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

Do per scontato che tutti sappiate che scambiando due righe o due colonne di un determinate se ne cambia il segno, ma non il valore assoluto. Allora, la proprieta' 120 segue dal fatto che una permutazione ciclica delle tre righe del determinante consiste sempre in *due* scambi

$$\vec{a}_1 \cdot \vec{a}_2 \times \vec{a}_3 = \vec{a}_2 \cdot \vec{a}_3 \times \vec{a}_1 = \vec{a}_3 \cdot \vec{a}_1 \times \vec{a}_2 \quad (120)$$

Ad esempio:

$$\vec{b}_2 \cdot \vec{a}_2 = \vec{a}_2 \cdot \vec{b}_2 = 2\pi \vec{a}_2 \frac{\vec{a}_3 \times \vec{a}_1}{\vec{a}_1 \cdot \vec{a}_2 \times \vec{a}_3} = 2\pi \frac{\vec{a}_1 \cdot \vec{a}_2 \times \vec{a}_3}{\vec{a}_1 \cdot \vec{a}_2 \times \vec{a}_3} = 2\pi \quad (121)$$

50. Dalla proprieta' 118 ne segue un'altra che useremo in piu' occasioni. Detti \mathcal{R} e \mathcal{R}^* il reticolo diretto e quello reciproco, rispettivamente, se $\vec{k} = n_1 \vec{b}_1 + n_2 \vec{b}_2 + n_3 \vec{b}_3$ e' un punto di \mathcal{R}^* allora, per ogni punto $\vec{R} = m_1 \vec{a}_1 + m_2 \vec{a}_2 + m_3 \vec{a}_3$ di \mathcal{R} , vale:

$$\forall \vec{R} \in \mathcal{R} \quad \vec{k} \cdot \vec{R} = (n_1 m_1 + n_2 m_2 + n_3 m_3) 2\pi = 2\pi n \quad n \in \mathbb{Z} \quad (122)$$

cioe': il prodotto scalare di un qualunque vettore del reticolo reciproco \mathcal{R}^* e un qualunque vettore del reticolo diretto \mathcal{R} e' sempre uguale a un *multiplo intero* di 2π .

E' possibile dimostrare anche che la 122 vale *solo* per i punti del reticolo reciproco; cioe' se un vettore \vec{d} dello spazio continuo in cui \mathcal{R}^* "e' immerso" *non* appartiene al *reticolo* reciproco (cioe' non e' dato da una combinazione lineare dei vettori primitivi del reticolo reciproco *con coefficienti interi*), allora per esso la 122 non vale.

La dimostrazione e' ottenuta nel modo seguente.

Consideriamo un vettore $\vec{d} \notin \mathcal{R}^*$, ma appartenente allo spazio continuo di cui anche \mathcal{R}^* fa parte. Siccome $\vec{b}_1, \vec{b}_2, \vec{b}_3$ sono linearmente indipendenti, essi costituiscono una base (in senso algebrico) di tale spazio continuo e quindi \vec{d} puo' essere espresso come una loro combinazione lineare. Tuttavia, siccome \vec{d} non appartiene a \mathcal{R}^* , nella combinazione lineare *almeno un coefficiente non e' intero* (se tutti e tre i coefficienti fossero interi, il vettore \vec{d} sarebbe un vettore del reticolo reciproco). Sara' pertanto:

$$\vec{d} = x_1 \vec{b}_1 + x_2 \vec{b}_2 + x_3 \vec{b}_3 \quad (123)$$

con almeno uno fra x_1, x_2, x_3 non intero.

Il prodotto scalare di \vec{d} con un generico vettore $\vec{R} \in \mathcal{R}$ e':

$$\vec{d} \cdot \vec{R} = (x_1 m_1 + x_2 m_2 + x_3 m_3) 2\pi \quad (124)$$

Fate attenzione: dal fatto che nella 123 almeno uno degli x_i sia non intero, *non segue* che il secondo membro della 124 non e' *mai* un multiplo intero di 2π : ad esempio, se $(x_1, x_2, x_3) = (1/2, 1, 1)$ e $(m_1, m_2, m_3) = (2, 1, 1)$, allora si avrebbe: $x_1 m_1 + x_2 m_2 + x_3 m_3 = 3 \in \mathbb{Z}$.

Tuttavia, dal fatto che nella 123 almeno uno degli x_i sia non intero segue che esiste *sicuramente almeno un* vettore \vec{R} del reticolo diretto per il quale il coefficiente di 2π nella 124 *non e'* intero.

Infatti, detto j l'indice per cui $x_j \notin \mathbb{Z}$, il vettore \vec{a}_j e' certamente un vettore di \mathcal{R} e per esso vale:

$$\vec{d} \cdot \vec{a}_j = x_j 2\pi \quad x_j \notin \mathbb{Z} \quad (125)$$

(cioe' e' sempre una permutazione *pari*): ma questo significa che il determinante (cioe' il triplo prodotto) rimane identico (anche nel segno) is seguito ad una permutazione ciclica.

Ma allora e' dimostrato che la 122 non vale per \vec{d} , perche' essa deve valere *per ogni* vettore del reticolo diretto, mentre invece ne esiste almeno uno per cui essa non vale.

Quanto appena visto consente una definizione equivalente (piu' astratta, ma utile dal punto di vista matematico) di reticolo reciproco:

Dato un reticolo di Bravais \mathcal{R} con vettori reticolari \vec{R} , si definisce reticolo reciproco ad esso associato l'insieme di *tutti e soli* i punti \vec{k} che soddisfano la condizione:

$$\forall \vec{R} \in \mathcal{R} \quad \vec{k} \cdot \vec{R} = 2\pi n \quad n \in \mathbb{Z} \quad (126)$$

51. Infine, anche nel caso 3D si puo' dimostrare che, con la definizione del reticolo reciproco introdotta, lo sviluppo in serie di Fourier della funzione f assume la forma:

$$f(\vec{r}) = \sum_{\vec{k}} B_{\vec{k}} \exp(i\vec{k} \cdot \vec{r}) \quad (127)$$

dove la somma e' estesa a tutti i punti \vec{k} del reticolo reciproco e i coefficienti $B_{\vec{k}}$ sono dati da:

$$B_{\vec{k}} = \frac{1}{V} \int_V f(\vec{r}) \exp(-i\vec{k} \cdot \vec{r}) d\vec{r} \quad (128)$$

dove l'integrale (triplo) e' calcolato sul volume della cella primitiva del reticolo di Bravais.

Vi ricordo di nuovo che l'uso di \vec{k} come indice e' un'abbreviazione per tre indici interi. Cioe', ad esempio:

$$\sum_{\vec{k}} B_{\vec{k}} \exp(i\vec{k} \cdot \vec{r})$$

significa

$$\sum_{\substack{n_1=-\infty \\ n_2=-\infty \\ n_3=-\infty}}^{\infty} B_{n_1, n_2, n_3} \exp\left(i(n_1 \vec{b}_1 + n_2 \vec{b}_2 + n_3 \vec{b}_3) \cdot \vec{r}\right)$$

52. Notate come la definizione del reticolo reciproco (eq. 112–114) assieme alla proprieta' 122 che da essa discende, assicuri che lo sviluppo di Fourier 127 abbia la periodicita' del reticolo di Bravais (ovviamente questa e' una condizione necessaria, visto che la funzione $f(\vec{r})$ che compare al primo membro della 127 e' periodica sul reticolo di Bravais).

Infatti, dalla 122 segue:

$$\exp(i\vec{k} \cdot \vec{R}) = \exp(i 2\pi n) = 1 \quad (129)$$

e quindi, per ogni termine della sommatoria 127:

$$\exp\left(i\vec{k} \cdot (\vec{r} + \vec{R})\right) = \exp\left(i\vec{k} \cdot \vec{r} + i\vec{k} \cdot \vec{R}\right) = \exp\left(i\vec{k} \cdot \vec{r}\right) \exp\left(i\vec{k} \cdot \vec{R}\right) = \exp\left(i\vec{k} \cdot \vec{r}\right) \quad (130)$$

il che significa che tutto il secondo membro della 127 e' periodico sul reticolo di Bravais, come deve essere.

Alcune proprieta' delle serie di Fourier

Prima di procedere, dobbiamo vedere alcune semplici proprieta' delle serie di Fourier di cui faremo uso nel seguito.

53. Proprieta'. Se $f(\vec{r})$ e' reale, i coefficienti del suo sviluppo di Fourier:

$$f(\vec{r}) = \sum_{\vec{k}} B_{\vec{k}} \exp\left(i\vec{k} \cdot \vec{r}\right)$$

con $\vec{k} = n_1\vec{b}_1 + n_2\vec{b}_2 + n_3\vec{b}_3$ vettore del reticolo reciproco, sono legati dalla seguente relazione:

$$B_{-\vec{k}} = B_{\vec{k}}^*$$

($B_{\vec{k}}^*$ e' il complesso coniugato di $B_{\vec{k}}$)

Vale anche l'implicazione inversa, e quindi:

$$f(\vec{r}) = (f(\vec{r}))^* \quad (\text{cioe' } f(\vec{r}) \text{ reale}) \iff B_{-\vec{k}} = B_{\vec{k}}^* \quad (131)$$

Vediamo i due versi dell'implicazione.

$$\mathbf{f}(\vec{r}) = (\mathbf{f}(\vec{r}))^* \Rightarrow \mathbf{B}_{-\vec{k}} = \mathbf{B}_{\vec{k}}^*$$

Supponiamo che $f(\vec{r})$ sia reale.

L'espressione per il coefficiente di Fourier $B_{\vec{k}}$ e':

$$B_{\vec{k}} = \frac{1}{V} \int_V \exp\left(-i\vec{k} \cdot \vec{r}\right) f(\vec{r}) d\vec{r}$$

dove V e' il volume della cella primitiva del reticolo diretto e $d\vec{r} = dx dy dz$ e' l'elemento infinitesimo di volume.

Il complesso coniugato di $B_{\vec{k}}$ e':

$$B_{\vec{k}}^* = \left[\frac{1}{V} \int_V \exp\left(-i\vec{k} \cdot \vec{r}\right) f(\vec{r}) d\vec{r} \right]^*$$

Osservazione E' banale verificare che l'operatore di coniugazione e' lineare; cioe' dati due numeri reali a_1, a_2 e due numeri complessi z_1, z_2 , vale sempre:

$$[a_1 z_1 + a_2 z_2]^* = a_1 z_1^* + a_2 z_2^* \quad (132)$$

Osservazione E' banale verificare anche che, dati due numeri complessi z_1, z_2 , si ha:

$$[z_1 z_2]^* = z_1^* z_2^* \quad (133)$$

Osservazione Con $h(\vec{r})$ e $g(\vec{r})$ funzioni reali, vale:

$$\begin{aligned}
\left[\int h(\vec{r}) + ig(\vec{r}) d\vec{r} \right]^* &= \left[\int h(\vec{r}) d\vec{r} + i \int g(\vec{r}) d\vec{r} \right]^* \\
&= \int h(\vec{r}) d\vec{r} - i \int g(\vec{r}) d\vec{r} = \int h(\vec{r}) - ig(\vec{r}) d\vec{r} \\
&= \int [h(\vec{r}) + ig(\vec{r})]^* d\vec{r}
\end{aligned}$$

cioe' l'integrale e' "trasparente" rispetto all'operatore di coniugazione:

$$\left[\int \dots \right]^* = \int [\dots]^* \tag{134}$$

Quindi, in base alle 132, 133 e 134:

$$\begin{aligned}
B_{\vec{k}}^* &= \frac{1}{V} \left[\int_V \exp(-i\vec{k} \cdot \vec{r}) f(\vec{r}) d\vec{r} \right]^* = \frac{1}{V} \int_V \left[\exp(-i\vec{k} \cdot \vec{r}) f(\vec{r}) \right]^* d\vec{r} \\
&\text{(qua sfruttiamo l'ipotesi che } f(\vec{r}) \text{ e' reale)} \\
&= \frac{1}{V} \int_V f(\vec{r}) \left[\exp(-i\vec{k} \cdot \vec{r}) \right]^* d\vec{r} \\
&= \frac{1}{V} \int_V f(\vec{r}) \exp(i\vec{k} \cdot \vec{r}) d\vec{r} = \frac{1}{V} \int_V f(\vec{r}) \exp(-i(-\vec{k}) \cdot \vec{r}) d\vec{r} = B_{-\vec{k}}
\end{aligned}$$

e siamo a posto.

$$\mathbf{B}_{-\vec{k}} = \mathbf{B}_{\vec{k}}^* \Rightarrow \mathbf{f}(\vec{r}) = (\mathbf{f}(\vec{r}))^*$$

Scriviamo le due espressioni per $B_{-\vec{k}}$ e $B_{\vec{k}}^*$:

$$B_{-\vec{k}} = \frac{1}{V} \int_V \exp(i\vec{k} \cdot \vec{r}) f(\vec{r}) d\vec{r}$$

$$\begin{aligned}
B_{\vec{k}}^* &= \left[\frac{1}{V} \int_V \exp(-i\vec{k} \cdot \vec{r}) f(\vec{r}) d\vec{r} \right]^* = \frac{1}{V} \int_V \left[\exp(-i\vec{k} \cdot \vec{r}) f(\vec{r}) \right]^* d\vec{r} \\
&= \frac{1}{V} \int_V \exp(i\vec{k} \cdot \vec{r}) f(\vec{r})^* d\vec{r}
\end{aligned}$$

(abbiamo usato le osservazioni 132, 133 e 134)

Confrontando le espressioni per $B_{-\vec{k}}$ e $B_{\vec{k}}^*$ e ricordando che per ipotesi $B_{-\vec{k}} = B_{\vec{k}}^*$ si ricava:

$$f(\vec{r}) = f(\vec{r})^*$$

ovvero $f(\vec{r})$ e' reale.

54. Alla luce di questa proprieta' e' possibile vedere in modo semplice che la forma complessa della serie di Fourier monodimensionale si riduce alla forma 17

quando la funzione e' reale (in realta', come ho detto al punto 25, non c'e' bisogno di alcuna "dimostrazione", visto che abbiamo ricavato la forma complessa a partire dalla 34).

Partiamo dalla 73 e supponiamo che la funzione sviluppata in serie di Fourier sia reale:

$$\sum_{n=-\infty}^{\infty} B_n \exp(i\alpha_n) \quad (135)$$

con $\alpha_n = \frac{2\pi n}{T}x$ e (dalla 69):

$$B_n = \frac{1}{2}(C_n - iS_n) \quad (136)$$

Spezzo la sommatoria come segue:

$$= B_0 + \sum_{n=-\infty}^{-1} B_n \exp(i\alpha_n) + \sum_{n=1}^{\infty} B_n \exp(i\alpha_n) \quad (137)$$

Ora osservo che, in base alla 40:

$$B_0 = \frac{1}{2}(C_0 - iS_0) = \frac{1}{2}C_0 = F_0 \quad (138)$$

Quindi posso continuare, cambiando anche gli estremi della prima somma e, corrispondentemente, il segno dell'indice n :

$$= F_0 + \sum_{n=1}^{\infty} B_{-n} \exp(i\alpha_{-n}) + \sum_{n=1}^{\infty} B_n \exp(i\alpha_n) \quad (139)$$

Ora uso la proprieta' 131 e ottengo:

$$= F_0 + \sum_{n=1}^{\infty} B_n^* \exp(-i\alpha_n) + \sum_{n=1}^{\infty} B_n \exp(i\alpha_n) \quad (140)$$

dove ho anche riscritto in modo equivalente l'argomento dell'esponenziale: $i\alpha_{-n} = -i\alpha_n$.

Ora osservo che i termini della prima somma sono i complessi coniugati di quelli nella seconda somma. Quindi posso unire le sommatorie e usare il fatto che, dato il numero complesso $z = a + ib$, si ha:

$$z^* + z = a + ib + a - ib = 2a = 2\Re(z) \quad (141)$$

($\Re(\dots)$ indica la parte reale dell'argomento)

Quindi continuo con:

$$= F_0 + \sum_{n=1}^{\infty} (B_n \exp(i\alpha_n))^* + B_n \exp(i\alpha_n) = F_0 + \sum_{n=1}^{\infty} 2\Re(B_n \exp(i\alpha_n)) \quad (142)$$

Infine sostituisco a B_n la 69 ed esprimo l'esponenziale con la formula di Eulero 8:

$$= F_0 + \sum_{n=1}^{\infty} 2\Re \left(\frac{1}{2} (C_n - iS_n) (\cos(\alpha_n) + i \sin(\alpha_n)) \right) \quad (143)$$

$$= F_0 + \sum_{n=1}^{\infty} C_n \cos(\alpha_n) + S_n \sin(\alpha_n) \quad (144)$$

che e' l'espressione reale 17.

55. Proprieta'. I coefficienti di Fourier per lo sviluppo di una funzione *reale* e *pari* sono tutti reali e pari (rispetto a \vec{k}).

Cioe': dato lo sviluppo in serie di Fourier:

$$f(\vec{r}) = \sum_{\vec{k}} B_{\vec{k}} \exp(i\vec{k} \cdot \vec{r})$$

con $\vec{k} = n_1 \vec{b}_1 + n_2 \vec{b}_2 + n_3 \vec{b}_3$ vettore del reticolo reciproco, vale:

$$f(\vec{r}) = (f(\vec{r}))^* \text{ e } f(\vec{r}) = f(-\vec{r}) \implies B_{\vec{k}} = B_{\vec{k}}^* \text{ e } B_{\vec{k}} = B_{-\vec{k}} \quad (145)$$

Prova in una dimensione

Vediamo prima la dimostrazione in una sola dimensione.

In tali condizioni si ha:

$$\begin{aligned} \vec{k} &\rightarrow k = \frac{2\pi n}{T} \quad n \in \mathbb{Z} \quad (T = \text{periodo}) \\ \vec{r} &\rightarrow x \quad (\text{variabile scalare}) \end{aligned}$$

e quindi:

$$B_k = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(x) \exp(-ikx) dx$$

Operiamo il seguente cambio di variabile:

$$u = -x$$

Allora:

$$\begin{aligned} x &= -u \\ dx &= -du \\ x = -\frac{T}{2} &\rightarrow u = \frac{T}{2} \\ x = \frac{T}{2} &\rightarrow u = -\frac{T}{2} \end{aligned}$$

Quindi:

$$B_k = \frac{1}{T} \int_{\frac{T}{2}}^{-\frac{T}{2}} f(-u) \exp(iku) (-du)$$

Ma per ipotesi $f(x)$ e' pari e quindi:

$$B_k = -\frac{1}{T} \int_{\frac{T}{2}}^{-\frac{T}{2}} f(u) \exp(iku) du$$

Infine scambiamo gli estremi di integrazione (l'integrale cambia segno) e riarangiamo l'argomento dell'esponenziale:

$$B_k = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(u) \exp(iku) du = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(u) \exp(-i(-k)u) du = B_{-k} \quad (146)$$

e questo prova la parita'.

Infine, combinando il risultato appena ottenuto con la 131 (applicabile perche' la f e' reale per ipotesi), si ha:

$$\left. \begin{array}{l} B_k = B_{-k} \\ B_{-k} = B_k^* \end{array} \right\} \Rightarrow B_k = B_k^* \quad (147)$$

e questo prova che i coefficienti sono reali.

Prova nello spazio 3D

(questa parte della dimostrazione non e' inclusa nel programma del corso: e' qua solo per vostra (e mia) ulteriore informazione; se non vi interessa potete saltarla)

Partiamo dall'espressione del generico coefficiente di Fourier:

$$B_{\vec{k}} = \frac{1}{V} \int_V \exp(-i\vec{k} \cdot \vec{r}) f(\vec{r}) d\vec{r} \quad (148)$$

(V e' il volume della cella primitiva del reticolo diretto)

Per comodita' di notazione, poniamo:

$$\vec{r} = (x_1, x_2, x_3) \quad (149)$$

Ora, in modo analogo a quanto fatto in una sola dimensione, facciamo il cambio di variabile:

$$\vec{u} = -\vec{r} \quad (150)$$

Attenzione, pero': fare un cambio di variabile in un integrale multiplo (e quello sopra e' un integrale triplo) non e' la stessa cosa che farlo in un integrale su una sola variabile.⁵

In generale, dette (u_1, u_2, u_3) le componenti del vettore \vec{u} , il cambio di variabile e' definito dalle tre relazioni:

$$x_1 = X_1(u_1, u_2, u_3) \quad (151)$$

$$x_2 = X_2(u_1, u_2, u_3) \quad (152)$$

$$x_3 = X_3(u_1, u_2, u_3) \quad (153)$$

Allora l'integrale nelle variabili (x_1, x_2, x_3) e' legato a quello nelle variabili (u_1, u_2, u_3) dalla seguente relazione:

$$\int_V F(\vec{r}) d\vec{r} = \int_V F(x_1, x_2, x_3) dx_1 dx_2 dx_3 \quad (154)$$

$$= \int_{V'} F(X_1, X_2, X_3) |J| du_1 du_2 du_3 = \int_{V'} F(\vec{X}) |J| d\vec{u} \quad (155)$$

dove F e' la generica funzione da integrare espressa nelle variabili u_1, u_2, u_3 , V' e' la regione di spazio V espressa con le nuove coordinate (u_1, u_2, u_3) e $|J|$ e' il valore assoluto del determinante jacobiano delle vecchie coordinate rispetto alle nuove:

$$J = \begin{vmatrix} \frac{\partial X_1}{\partial u_1} & \frac{\partial X_1}{\partial u_2} & \frac{\partial X_1}{\partial u_3} \\ \frac{\partial X_2}{\partial u_1} & \frac{\partial X_2}{\partial u_2} & \frac{\partial X_2}{\partial u_3} \\ \frac{\partial X_3}{\partial u_1} & \frac{\partial X_3}{\partial u_2} & \frac{\partial X_3}{\partial u_3} \end{vmatrix} \quad (156)$$

Le relazioni che legano le vecchie variabili alle nuove sono :

$$x_1 = X_1(u_1, u_2, u_3) = -u_1 \quad (157)$$

$$x_2 = X_2(u_1, u_2, u_3) = -u_2 \quad (158)$$

$$x_3 = X_3(u_1, u_2, u_3) = -u_3 \quad (159)$$

e quindi il determinante jacobiano e':

$$J = \begin{vmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{vmatrix} = -1 \Rightarrow |J| = 1 \quad (160)$$

In definitiva, operando il cambio di variabile otteniamo:

$$\begin{aligned} B_{\vec{k}} &= \frac{1}{V} \int_V \exp(-i\vec{k} \cdot \vec{r}) f(\vec{r}) d\vec{r} = \frac{1}{V} \int_{V'} \exp(-i\vec{k} \cdot (-\vec{u})) f(-\vec{u}) d\vec{u} \\ &\text{(qua sfruttiamo l'ipotesi che la } f \text{ sia pari)} \\ &= \frac{1}{V} \int_{V'} \exp(-i(-\vec{k}) \cdot \vec{u}) f(\vec{u}) d\vec{u} = B_{-\vec{k}} \end{aligned} \quad (161)$$

e questo prova che i coefficienti sono pari.

Inoltre, siccome la f e' reale per ipotesi, possiamo invocare la 161 appena dimostrata e la proprieta' 131 e quindi:

$$\left. \begin{aligned} B_{\vec{k}} &= B_{-\vec{k}} \\ B_{-\vec{k}} &= B_{\vec{k}}^* \end{aligned} \right\} \Rightarrow B_{\vec{k}} = B_{\vec{k}}^* \quad (162)$$

e questo prova che i coefficienti sono reali.

56. Proprieta'. I coefficienti di Fourier per lo sviluppo di una funzione *reale* e *dispari* sono tutti immaginari e dispari.

Cioe': dato lo sviluppo in serie di Fourier:

$$f(\vec{r}) = \sum_{\vec{k}} f_{\vec{k}} \exp(i\vec{k} \cdot \vec{r})$$

con $\vec{k} = n_1\vec{b}_1 + n_2\vec{b}_2 + n_3\vec{b}_3$ vettore del reticolo reciproco, vale:

$$f(\vec{r}) = (f(\vec{r}))^* \text{ e } f(\vec{r}) = -f(-\vec{r}) \iff B_{\vec{k}} = -B_{-\vec{k}} \text{ e } B_{\vec{k}} = -B_{\vec{k}}^* \quad (163)$$

Anche di questa riporto la versione 1D e quella 3D della dimostrazione.

Prova in una sola dimensione

Come gia' visto al punto 55, in una sola dimensione si ha:

$$\begin{aligned} \vec{k} &\rightarrow k = \frac{2\pi n}{T} \quad n \in \mathbb{Z} \quad (T = \text{periodo}) \\ \vec{r} &\rightarrow x \quad (\text{variabile scalare}) \end{aligned}$$

e quindi:

$$B_k = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(x) \exp(-ikx) dx$$

Operiamo il seguente cambio di variabile:

$$u = -x$$

Allora:

$$\begin{aligned} x &= -u \\ dx &= -du \\ x = -\frac{T}{2} &\rightarrow u = \frac{T}{2} \\ x = \frac{T}{2} &\rightarrow u = -\frac{T}{2} \end{aligned}$$

Quindi:

$$B_k = \frac{1}{T} \int_{\frac{T}{2}}^{-\frac{T}{2}} f(-u) \exp(iku) (-du)$$

Ma per ipotesi $f(u) = -f(-u)$ (f e' dispari) e quindi:

$$B_k = \frac{1}{T} \int_{\frac{T}{2}}^{-\frac{T}{2}} f(u) \exp(iku) du$$

Scambiamo gli estremi di integrazione (l'integrale cambia segno) e riarrangiamo l'argomento dell'esponenziale:

$$B_k = -\frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(u) \exp(-i(-k)u) du = -B_{-k} \quad (164)$$

e questo prova la parita' dispari.

Infine, combinando il risultato appena ottenuto con la 131 (applicabile perche' la f e' reale per ipotesi), si ha:

$$\left. \begin{array}{l} B_k = -B_{-k} \\ B_{-k} = B_k^* \end{array} \right\} \Rightarrow B_k = -B_k^* \quad (165)$$

e questo prova che i coefficienti sono numeri immaginari puri.

Prova nello spazio 3D

(anche questa parte della dimostrazione non e' inclusa nel programma del corso: e' qua solo per chi sia interessato)

Fate riferimento a quanto abbiamo gia' visto al punto 55.

L'espressione del generico coefficiente di Fourier e':

$$B_{\vec{k}} = \frac{1}{V} \int_V \exp(-i\vec{k} \cdot \vec{r}) f(\vec{r}) d\vec{r} \quad (166)$$

(V e' il volume della cella primitiva del reticolo diretto)

Poniamo:

$$\vec{r} = (x_1, x_2, x_3) \quad (167)$$

Ora facciamo lo stesso cambio di variabile:

$$\vec{u} = -\vec{r} \quad (168)$$

Le relazioni che legano le variabili vecchie alle nuove sono esattamente le stesse:

$$x_1 = X_1(u_1, u_2, u_3) = -u_1 \quad (169)$$

$$x_2 = X_2(u_1, u_2, u_3) = -u_2 \quad (170)$$

$$x_3 = X_3(u_1, u_2, u_3) = -u_3 \quad (171)$$

e quindi anche il determinante jacobiano e' il medesimo:

$$J = \begin{vmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{vmatrix} = -1 \Rightarrow |J| = 1 \quad (172)$$

Operando il cambio di variabile otteniamo:

$$\begin{aligned} B_{\vec{k}} &= \frac{1}{V} \int_V \exp(-i\vec{k} \cdot \vec{r}) f(\vec{r}) d\vec{r} = \frac{1}{V} \int_{V'} \exp(-i\vec{k} \cdot (-\vec{u})) f(-\vec{u}) d\vec{u} \\ &\text{(qua sfruttiamo l'ipotesi che la } f \text{ sia dispari)} \\ &= -\frac{1}{V} \int_{V'} \exp(-i(-\vec{k}) \cdot \vec{u}) f(\vec{u}) d\vec{u} = -B_{-\vec{k}} \end{aligned} \quad (173)$$

e questo prova che i coefficienti sono dispari.

Inoltre, siccome la f e' reale per ipotesi, possiamo invocare la 173 appena dimostrata e la proprieta' 131 e quindi:

$$\left. \begin{aligned} B_{\vec{k}} &= -B_{-\vec{k}} \\ B_{-\vec{k}} &= B_{\vec{k}}^* \end{aligned} \right\} \Rightarrow B_{\vec{k}} = -B_{\vec{k}}^* \quad (174)$$

e questo prova che i coefficienti sono numeri immaginari puri.

57. In modo strettamente analogo a quanto abbiamo visto fin qua, si possono dimostrare altre due proprieta' simili, di cui pero' saltiamo la dimostrazione (perche' non le useremo).

Riassumo tutte queste proprieta' nella tabella seguente:

se la $f(\vec{r})$ e'...	i $B_{\vec{k}}$ sono...	ref: proprieta'...
reale	$B_{-\vec{k}} = B_{\vec{k}}^*$	131
reale e pari	reali e pari	145
reale e dispari	immaginari e dispari	163
immaginaria e pari	immaginari e pari	N/D
immaginaria e dispari	reali e dispari	N/D

Richiami sulle onde

58. Prima di procedere alla parte successiva, vediamo ancora qualche richiamo sulle *onde* e la loro rappresentazione matematica. Questo mi pare opportuno, visto che nel seguito avremo spesso a che fare con tipiche funzioni ondulatorie.

59. In generale, un'onda e' la variazione (rispetto ad un valore di riferimento) di una qualche grandezza fisica definita in ogni punto di una regione di spazio (la regione puo' essere anche infinita e coincidere con l'intero spazio). L'aspetto essenziale e peculiare di un'onda e' che la variazione della grandezza che la costituisce dipende dal tempo oltre che dalle coordinate spaziali. Cioe' fissato un punto della regione considerata, la variazione della grandezza che definisce l'onda in quel punto dipende dal tempo.

In generale, dunque, per quello che ci riguarda, definiamo matematicamente un'onda come una funzione di quattro coordinate: le tre coordinate spaziali e il tempo:

$$f(x, y, z, t) = f(\vec{r}, t) \quad (175)$$

60. Un esempio di onde che tutti conosciamo sono le onde del mare. In questo caso lo spazio in cui l'onda è definita è uno spazio bidimensionale, cioè la superficie del mare; la grandezza fisica che varia è l'altezza dell'acqua rispetto al livello medio della superficie marina; la dipendenza dal tempo è facilmente verificabile: se fissiamo lo sguardo su un determinato punto della superficie, la sua altezza rispetto al livello medio varia periodicamente in funzione del tempo. Le onde del mare sono un esempio di *onde progressive* (“traveling waves”). Se guardate la cresta di un'onda del mare vedete che essa si sposta (ad una velocità approssimativamente costante). Lo stesso vale per qualsiasi altra parte dell'onda (ma guardare la propagazione delle creste è più semplice). Naturalmente, tutti sapete che mentre le onde del mare “viaggiano”, il mare stesso non si muove: se così non fosse la riva del mare, intesa come confine tra il mare e la terraferma, non esisterebbe...

61. Un altro esempio di onda è quella che si genera quando pizzichiamo la corda di una chitarra: in questo caso lo spazio in cui è definita l'onda è monodimensionale e, in modo molto simile alle onde del mare, la variazione della grandezza fisica da considerare è lo spostamento della corda dalla sua posizione di riposo: tale spostamento varia con la posizione lungo la corda e, per una data posizione, varia col tempo. Questo tipo di onda, a differenza delle onde marine, è un'onda *stazionaria* (“standing wave”). La differenza consiste nel fatto che, in questo caso, l'onda non si sposta e ciò appare evidente se considerate che i due estremi a cui la corda è fissata (e che fanno parte di essa) non si muovono (ovviamente). Ogni punto della corda oscilla nel tempo, ma la posizione del massimo (e minimo) di oscillazione lungo la corda non cambia.

62. Cominciamo ad approfondire la descrizione matematica di un'onda progressiva.

In pratica, si tratta di descrivere matematicamente una funzione delle coordinate spaziali il cui grafico, però, “si sposta” col tempo.¹⁵

Per rendere le cose più semplici, consideriamo una funzione di una sola coordinata (spaziale): $f(x)$ e chiediamoci intanto come sia possibile traslare il grafico della f lungo l'asse x .

Questo è molto facile (guardate la figura 9): dato il grafico di $f(x)$, il grafico della funzione $f(x-t)$, con $t > 0$ è identico al grafico di $f(x)$, ma traslato di una quantità t nel verso positivo dell'asse x (cioè verso destra).

Infatti, la funzione $f(x-t)$ assume il valore $f(x')$ per $x = x' + t$:

$$f((x' + t) - t) = f(x') \quad (176)$$

e siccome questo vale $\forall x'$, ne segue che il grafico della $f(x-t)$ si ottiene trasladando verso destra di una quantità t quello della $f(x)$.

Ovviamente, con un ragionamento analogo, si comprende che il grafico della funzione $f(x+t)$ ($t > 0$) è identico a quello della $f(x)$, ma traslato della quantità t nel verso negativo dell'asse x cioè verso sinistra.

¹⁵Come ho già accennato a proposito delle onde del mare, in un'onda progressiva cioè che “viaggia” non è il mezzo (ad esempio l'acqua del mare) in cui l'onda si manifesta, ma la *descrizione* (cioè nella maggior parte dei casi *il grafico*) della variazione della grandezza fisica che origina l'onda. Quindi, nel seguito, quando dirò che un'onda progressiva si sposta o si muove, intenderò sempre lo spostamento o il movimento del suo grafico (la cosiddetta *forma d'onda*).

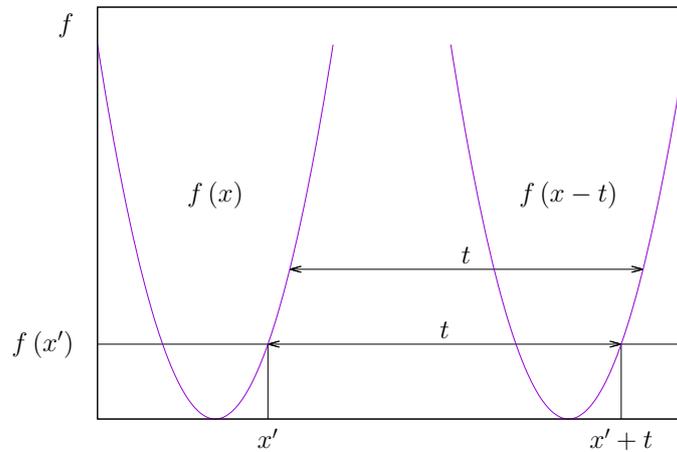


Figura 9:

63. Se ora pensate a t non piu' come a una costante, ma come a una seconda variabile da cui la f dipende, avete ottenuto un'onda progressiva: per t fisso, la funzione avra' un certo grafico (rappresentato sul piano cartesiano); all'aumentare di t il grafico si sposta ("viaggia") verso destra o verso sinistra (dipendentemente dal segno con cui viene preso t).

Per noi t indichera' sempre il tempo.

Ci manca ancora un dettaglio.

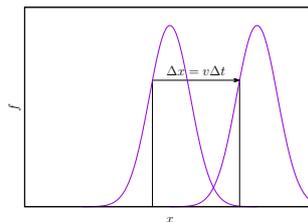
Se definiamo la nostra onda progressiva semplicemente con:

$$f(x \pm t) \quad (177)$$

allora, in un dato intervallo di tempo, lo spostamento del grafico e' sempre lo stesso. D'altro canto le onde progressive si spostano, ma lo fanno con velocita' diverse (pensate alla velocita' con cui si muove un'onda marina e a quella con cui si muove un'onda elettromagnetica, che e' anch'essa un'onda progressiva). Se vogliamo che l'onda possa muoversi a velocita' diverse, dobbiamo moltiplicare la variabile t per un fattore v che aumenti o diminuisca la traslazione subita dal grafico della funzione per uno stesso intervallo di tempo:

$$f(x \pm vt) \quad (178)$$

Dovrebbe essere evidente che il fattore v e' la velocita' con cui l'onda si muove lungo l'asse x . In un intervallo di tempo Δt un determinato punto del grafico della f si sposta di $\Delta x = v\Delta t$ lungo l'asse x e quindi:



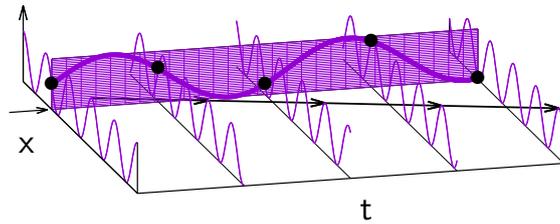


Figura 10: In questa figura ho rappresentato il grafico dell'onda: $f(x, t) = A \cos\left(\frac{2\pi}{\lambda}(x - vt)\right)$; le 5 sinusoidi parallele sono sezioni del grafico a 5 valori equispaziati di tempo; le 4 frecce collegano sempre lo stesso massimo per mettere in evidenza il movimento dell'onda; le frecce sono allineate su una retta perché la velocità dell'onda è costante. La sezione ombreggiata mostra l'andamento in funzione del tempo per un valore fissato di x (indicato dalla piccola freccia); i punti neri sono le intersezioni delle 5 sinusoidi col piano $x = \text{costante}$.

$$\Delta x = v\Delta t \Rightarrow v = \frac{\Delta x}{\Delta t} \quad (179)$$

64. Le onde che considereremo in questo corso saranno sempre delle *onde periodiche* sinusoidali di velocità costante, cioè onde in cui la funzione $f(\vec{r}, t)$ è una funzione periodica (sinusoidale) sia della variabile spaziale \vec{r} che del tempo t . Questo significa che, fissato un certo istante di tempo, il grafico di f in funzione della coordinata spaziale è sinusoidale; e, analogamente, fissato un certo punto dello spazio, il grafico di f in funzione del tempo (in quel punto) è anch'esso quello di una funzione sinusoidale.

Cio' è mostrato nella figura 10 dove ho rappresentato a diversi istanti di tempo un'onda periodica monodimensionale che si muove verso destra, mettendo in evidenza la periodicità sia spaziale che temporale.

65. L'espressione più semplice di un'onda sinusoidale progressiva in una dimensione è:¹⁶

$$f(x, t) = A \cos\left(\frac{2\pi}{\lambda}(x - vt)\right) \quad (180)$$

A viene detta *ampiezza* dell'onda ed è il massimo valore che assume la variazione della grandezza fisica rappresentata dall'onda (l'onda assume valori nell'intervallo $[-A, A]$ in corrispondenza ai valori -1 e 1 , rispettivamente, del coseno).

¹⁶In realtà, l'espressione 180 assume che l'ampiezza iniziale dell'onda, cioè a $x = t = 0$ sia A ; questo non è, in generale, necessario. L'espressione di un'onda sinusoidale con ampiezza iniziale qualsiasi (compresa fra $-A$ e A) è:

$$f(x, t) = A \cos\left(\frac{2\pi}{\lambda}(x - vt) + \Phi\right)$$

Φ viene anche detta *fase iniziale*.

L'argomento della funzione coseno viene detto *fase* dell'onda.¹⁷

Siccome l'onda sinusoidale e' periodica sia nella variabile x che nella variabile t , ci saranno due periodi (un periodo "spaziale" e uno "temporale"), due frequenze etc. Per questo motivo, si usa una nomenclatura che distingue le grandezze che si riferiscono alla variabile spaziale da quelle che invece sono definite per la variabile temporale (notate che questo problema non c'era quando abbiamo parlato delle funzioni periodiche in generale a p. 1).

Se riscrivete la 180 come:

$$f(x, t) = A \cos \left(\frac{2\pi}{\lambda} x - \frac{2\pi}{\frac{\lambda}{v}} t \right) \quad (181)$$

e tenete presente quanto abbiamo detto al punto 7, vedete subito che il periodo rispetto a x e' λ : nel caso di un'onda, il termine "periodo" viene pero' usato per indicare il periodo rispetto a t ; per questo, il periodo rispetto a x viene detto *lunghezza d'onda*. L'origine del nome e' che esso rappresenta la distanza (vera e propria, misurata in m, mm, Å etc) fra due punti corrispondenti e consecutivi (ad esempio due massimi) della rappresentazione spaziale dell'onda (cioe' f in funzione di x a t fissato).

La grandezza:

$$f = \frac{1}{\lambda} \quad (182)$$

e' detta *frequenza spaziale* (per distinguerla dalla frequenza temporale o semplicemente frequenza, che si usa per la dipendenza temporale dell'onda) ed e' la frequenza che avevamo definito in generale per una funzione periodica (eq. 3) riferita alla funzione f pensata come funzione periodica di x a t fissato.

Analogamente:

$$k = \frac{2\pi}{\lambda} \quad (183)$$

e' quella che avevamo definito "frequenza angolare" per una semplice funzione periodica (eq. 4). Come per il periodo, anche questo nome viene riservato alla corrispondente grandezza temporale e quindi, nel caso di un'onda, k viene detto *numero d'onda*. Quando x varia di λ (a t costante), la fase dell'onda varia di:

$$\frac{2\pi}{\lambda} (x + \lambda) - \frac{2\pi}{\frac{\lambda}{v}} t - \left(\frac{2\pi}{\lambda} x - \frac{2\pi}{\frac{\lambda}{v}} t \right) = 2\pi \quad (184)$$

Da cio' si ricava un significato fisico del numero d'onda: esso e' la variazione della fase espressa in radianti (a t costante) per variazione unitaria di x :

$$\frac{\text{variaz. della fase in rad}}{\text{variaz. di } x} = \frac{2\pi}{\lambda} = k \quad (185)$$

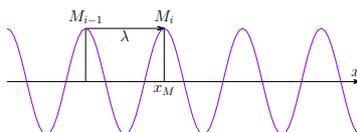
¹⁷Notate che esprimere un'onda sinusoidale con un coseno o un seno e' la stessa cosa, visto che una rappresentazione puo' essere convertita nell'altra tramite una semplice costante additiva nella fase:

$$\cos(\alpha) = \sin \left(\alpha + \frac{\pi}{2} \right)$$

66. Sempre dalla 181 si vede che il periodo temporale, detto semplicemente *periodo* (cioè il periodo dell'onda pensata come funzione di t a x costante) è:

$$T = \frac{\lambda}{v} \quad (186)$$

Questa relazione mostra che periodo e lunghezza d'onda sono legati e questo non dovrebbe stupire. Immaginate di fissare la x al valore x_M di un massimo M_i dell'onda per un certo istante.



Se l'onda viaggia a una velocità (costante) v , il tempo che impiega il massimo M_{i-1} (cioè il massimo immediatamente “a sinistra” di M_i) per raggiungere x_M deve essere λ/v (perché λ è la distanza che M_{i-1} deve percorrere). Ma questo è anche il tempo che impiega la funzione f (pensata come funzione del tempo per il valore fissato $x = x_M$) per compiere un ciclo completo, cioè il suo periodo T .

La grandezza:

$$\nu = \frac{1}{T} = \frac{v}{\lambda} \quad (187)$$

viene detta *frequenza* ed è la frequenza che avevamo definito in generale a proposito delle funzioni periodiche (eq. 3) riferita alla funzione f pensata come funzione periodica del tempo (a x costante).

Infine, si definisce *frequenza angolare* dell'onda la stessa grandezza che avevamo già definito in generale per una funzione periodica qualsiasi:

$$\omega = \frac{2\pi}{T} = 2\pi\nu \quad (188)$$

67. Riassumendo, le grandezze che definiscono la forma matematica di un'onda progressiva sinusoidale sono:

$$f(x, t) = A \cos\left(\frac{2\pi}{\lambda}(x - vt)\right)$$

A	ampiezza
λ	lunghezza d'onda
$f = \frac{1}{\lambda}$	frequenza spaziale
$k = \frac{2\pi}{\lambda}$	numero d'onda
v	velocità
$T = \frac{\lambda}{v}$	periodo
$\nu = \frac{1}{T}$	frequenza
$\omega = 2\pi\nu$	frequenza angolare

Utilizzando le definizioni di numero d'onda e frequenza angolare la 180 puo' esser riscritta in modo piu' compatto come:

$$f(x, t) = A \cos(kx - \omega t) \quad (189)$$

68. Onde stazionarie.

Un'onda stazionaria (le onde stazionarie sono sempre periodiche) e' un'onda per la quale la parte spaziale *non* "viaggia", cioe' il cui grafico spaziale non trasla col trascorrere del tempo.

La forma matematica piu' semplice per un'onda stazionaria monodimensionale e':

$$f(x, t) = A \cos(kx) \cos(\omega t) \quad (190)$$

dove A , k e ω sono quelli definiti sopra.

Notate la differenza rispetto a un'onda progressiva (eq. 189): la coordinata spaziale e quella temporale sono "separate", cioe' non compaiono nella stessa funzione sinusoidale. A p. 43 abbiamo visto che la condizione necessaria affinche' un'onda possa "viaggiare" e' proprio quella che x e t compaiano sempre insieme, combinate come $(x \pm vt)$ nella funzione che descrive matematicamente l'onda (eq. 178). In un'onda stazionaria questa condizione non e' verificata.

Potete rendervi conto facilmente che il grafico di un'onda stazionaria non subisce traslazione lungo l'asse x col trascorrere del tempo. I punti dell'asse x in cui la funzione si annulla vengono detti *nodi* dell'onda stazionaria. Essi sono dati da:

$$f(x, t) = 0 \iff kx = \frac{\pi}{2} + n\pi \Rightarrow x = \frac{1}{k} \left(\frac{\pi}{2} + n\pi \right) \quad n \in \mathbb{Z} \quad (191)$$

Dalla 190 e' chiaro che la f si annulla in corrispondenza ai valori di x mostrati sopra *per qualsiasi valore di t* e quindi i nodi della funzione non cambiano la loro posizione al variare di t ; ma se i nodi rimangono fissi, allora tutto il grafico della $f(x, t)$ in funzione di x rimane "localizzato" sull'asse x e non viaggia.

Cosa cambia nel grafico al variare del tempo?

Abbiamo appena visto che i nodi rimangono inalterati, cioe' nei punti dove x soddisfa la 191, il grafico e' indipendente dal tempo.

Considerate un punto x_0 che non sia un nodo. L'andamento della f in quel punto in funzione del tempo e':

$$f(x_0, t) = A \cos(kx_0) \cos(\omega t) \quad (192)$$

che e' un'oscillazione (armonica) di ampiezza $A \cos(kx_0)$.

Quindi in ogni punto della regione di spazio interessata da un'onda stazionaria la grandezza che la definisce varia periodicamente nel tempo, ma l'onda non si muove. Questo comportamento e' confrontato con quello di un'onda progressiva nella figura 11.

69. Onde periodiche in piu' dimensioni.

Per quello che riguarda questo corso, incontreremo un solo tipo di onda multidimensionale che e' l'estensione in due o tre dimensioni della semplice onda sinusoidale 189. Tale tipo di onda appartiene a una categoria di onde multidimensionali che vengono chiamate *onde piane*.

La forma matematica di un'onda piana sinusoidale progressiva e':

$$f(\vec{r}, t) = A \cos(k\vec{r} \cdot \vec{v} - \omega t) \quad (193)$$

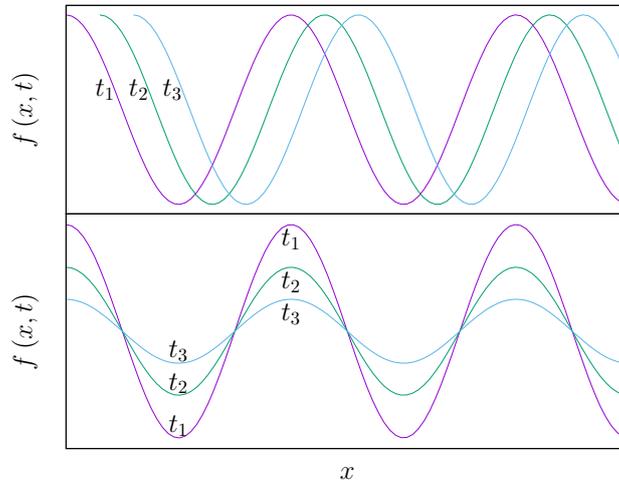


Figura 11: Un'onda progressiva (sopra) e un'onda stazionaria (sotto) a tre istanti di tempo successivi t_1, t_2, t_3 .

In un'onda piana bidimensionale, \vec{r} e' un vettore di due coordinate spaziali $\vec{r} = (x, y)$; in un'onda piana tridimensionale, \vec{r} e' un vettore di tre coordinate spaziali $\vec{r} = (x, y, z)$.

Come vedete, la 193 e' formalmente molto simile all'eq. 189, l'unica differenza essendo la natura vettoriale della parte spaziale $\vec{r} \cdot \vec{v}$.

I parametri A, k e ω hanno lo stesso significato visto nel caso monodimensionale (con una precisazione riguardante k che faremo fra un attimo).

\vec{v} e' un vettore di modulo unitario o *versore* e determina, come vedremo, la direzione di propagazione dell'onda piana sinusoidale; k e' il numero d'onda *nella direzione di propagazione*, cioe' nella direzione individuata dal versore \vec{v} .

Il significato di questa affermazione si comprende nel modo seguente (guardate la figura 12).

E' sempre possibile decomporre il vettore posizione \vec{r} in una componente parallela, \vec{r}_{\parallel} , e una componente perpendicolare, \vec{r}_{\perp} , a \vec{v} :

$$\vec{r} = \vec{r}_{\parallel} + \vec{r}_{\perp} \quad (194)$$

E allora:

$$f(\vec{r}, t) = A \cos(k\vec{r} \cdot \vec{v} - \omega t) = A \cos(k(\vec{r}_{\parallel} + \vec{r}_{\perp}) \cdot \vec{v} - \omega t) \quad (195)$$

$$= A \cos \left(\overbrace{k(\vec{r}_{\parallel} \cdot \vec{v})}^{=|\vec{r}_{\parallel}|=r_{\parallel}} + \overbrace{k(\vec{r}_{\perp} \cdot \vec{v})}^{=0} - \omega t \right) \quad (196)$$

$$= A \cos(kr_{\parallel} - \omega t) = f(r_{\parallel}, t)$$

cioe': nel generico punto \vec{r} il valore assunto dalla $f(\vec{r}, t)$ e' uguale a quello che assume la funzione (monodimensionale) $f(r_{\parallel}, t)$, dove r_{\parallel} e' la coordinata

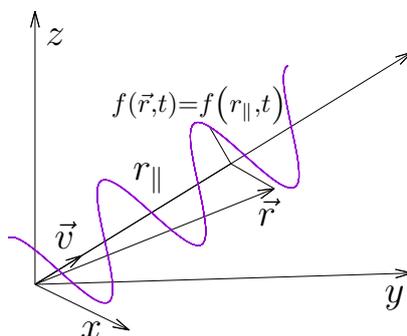


Figura 12: Il valore dell'onda piana $f(\vec{r}, t)$ nel generico punto \vec{r} a t fissato e' ottenuto proiettando \vec{r} lungo la direzione di propagazione \vec{v} e calcolando $f(r_{\parallel}, t)$, dove r_{\parallel} e' il modulo della proiezione.

scalare lungo l'asse individuato dal versore \vec{v} . Detto in altri termini: il valore della funzione $f(\vec{r}, t)$ nel generico punto \vec{r} e' ottenuto proiettando \vec{r} lungo la direzione di \vec{v} (il prodotto scalare $\vec{r} \cdot \vec{v}$ fa esattamente questo) e calcolando $f(r_{\parallel}, t)$, dove r_{\parallel} e' il modulo della proiezione. Questo aspetto e' mostrato nella figura 12.

70. Per comprendere meglio come e' fatta un'onda piana (nel seguito sottintendero' l'aggettivo "sinusoidale" per brevitaa) consideriamo un piano Π ortogonale a \vec{v} (guardate la figura 13). Chiaramente, tutti i vettori \vec{r} che si appoggiano su Π hanno la stessa proiezione lungo la direzione di \vec{v} , cioe' per tutti i punti del piano Π vale:

$$\vec{r} \cdot \vec{v} = d = \text{costante} \quad (197)$$

Notate che d e' anche la distanza di Π dall'origine (la distanza di un piano da un punto e' la lunghezza del segmento che va dal punto al piano ed e' ortogonale a quest'ultimo).

In base alla 197, per tutti i punti di Π si ha:

$$f(\vec{r}, t) = A \cos(k\vec{r} \cdot \vec{v} - \omega t) = A \cos(kd - \omega t) \quad (198)$$

E questo significa che, a t fissato, la fase dell'onda e' la stessa su tutti i punti di Π , ovvero il valore assunto da $f(\vec{r}, t)$ su tutti i punti di Π e' lo stesso (a un certo istante). Questo ragionamento e' applicabile a qualsiasi piano ortogonale a \vec{v} (ci sono infiniti piani ortogonali a \vec{v} , e quindi fra loro paralleli); ovviamente il particolare valore assunto dalla funzione $f(\vec{r}, t)$ su piani diversi sara', in generale, diverso, ma il valore e' lo stesso su tutti i punti di uno stesso piano. Il luogo dei punti in cui un'onda multidimensionale assume uno stesso valore viene chiamato *fronte d'onda* (wave front).

Quindi abbiamo appena ottenuto il seguente risultato: i fronti d'onda delle onde piane sono dei piani fra loro paralleli e ortogonali alla direzione di propagazione (cioe' la direzione del versore \vec{v}) (questo e' anche il motivo per cui questo tipo di onde multidimensionali sono dette "piane").

Siccome le onde piane che stiamo considerando sono periodiche nello spazio, e' chiaro che, a t fissato e fissato un certo valore dell'onda (ad esempio quello di

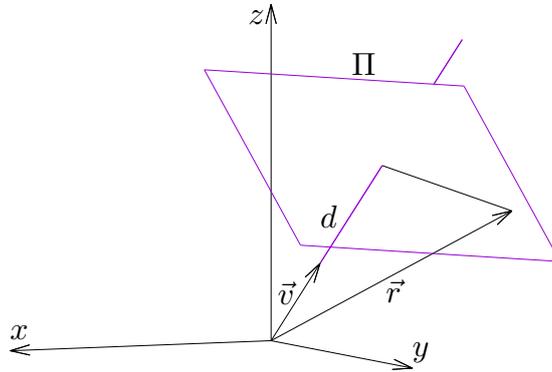


Figura 13: Tutti i vettori che si appoggiano sul piano Π ortogonale al versore \vec{v} hanno la stessa proiezione di modulo d lungo la direzione di \vec{v} . d e' anche la distanza di Π dall'origine.

un suo massimo), esisteranno infiniti fronti d'onda, tutti fra loro paralleli ed equispaziati, sui quali l'onda piana assume quel particolare valore. Si puo' vedere facilmente che la spaziatura di tali fronti d'onda e' uguale alla lunghezza d'onda λ . Infatti, detti Π_1 e Π_2 due fronti d'onda *consecutivi* appartenenti alla famiglia su cui l'onda assume uno stesso valore a un certo istante di tempo t , dovra' essere:

$$\overbrace{A \cos(kd_1 - \omega t)}^{\text{su } \Pi_1} = \overbrace{A \cos(kd_2 - \omega t)}^{\text{su } \Pi_2} \quad (199)$$

con:

$$d_1 = \vec{r} \cdot \vec{v} \quad \forall \vec{r} \in \Pi_1 \quad (200)$$

$$d_2 = \vec{r} \cdot \vec{v} \quad \forall \vec{r} \in \Pi_2 \quad (201)$$

Assumendo, per semplicita' che Π_2 sia piu' lontano dall'origine (e quindi $d_2 > d_1$), la 199 e' verificata quando gli argomenti dei coseni differiscono di 2π , cioe':

$$kd_2 - \omega t - (kd_1 - \omega t) = 2\pi \iff k(d_2 - d_1) = 2\pi \quad (202)$$

Ricordate che $d_2 - d_1$ non e' altro che la distanza fra i due fronti d'onda consecutivi, cioe' la spaziatura cercata.

Allora, sostituendo $k = 2\pi/\lambda$, si ottiene:

$$\text{spaziatura} = d_2 - d_1 = \lambda \quad (203)$$

71. Per rendervi conto che un'onda piana si sposta nella direzione del versore \vec{v} considerate l'onda 193 all'istante di tempo t_0 . Sia d_0 la distanza dall'origine di un certo fronte d'onda individuato da un vettore \vec{r}_0 :

$$d_0 = \vec{r}_0 \cdot \vec{v} \quad (204)$$

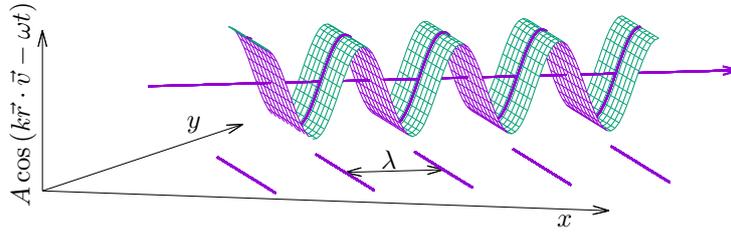


Figura 14: Un'onda piana bidimensionale a un istante di tempo fissato. L'onda si propaga nella direzione $\vec{v} = (\cos(\pi/6), \sin(\pi/6))$, evidenziata dalla freccia. Sulla superficie che rappresenta il grafico dell'onda piana ho anche evidenziato la sua proiezione sul piano normale a xy e contenente la direzione di propagazione: tale proiezione e' una semplice onda sinusoidale monodimensionale con la stessa lunghezza d'onda. Sul piano xy sono anche mostrati i fronti d'onda corrispondenti ai massimi.

Ci poniamo ora la seguente domanda: quando e' trascorso un intervallo di tempo Δt dove si trovera' quel particolare fronte d'onda?

Evidentemente dovra' trovarsi in un punto \vec{r}_1 tale che:

$$A \cos(k\vec{r}_1 \cdot \vec{v} - \omega(t_0 + \Delta t)) = A \cos(k\vec{r}_0 \cdot \vec{v} - \omega t_0) \quad (205)$$

perche' l'onda deve avere, a $t_0 + \Delta t$, lo stesso valore che aveva a t_0 (stiamo considerando uno stesso fronte d'onda).

L'uguaglianza appena scritta puo' essere vera soltanto se:

$$k\vec{r}_1 \cdot \vec{v} - \omega(t_0 + \Delta t) = k\vec{r}_0 \cdot \vec{v} - \omega t_0 + 2\pi n \quad n \in \mathbb{N} \quad (206)$$

(consideriamo solo valori positivi di n perche' $\Delta t > 0$ e quindi la fase e' aumentata)

Risolvendo rispetto a $\vec{r}_1 \cdot \vec{v}$ si ottiene:

$$\vec{r}_1 \cdot \vec{v} = d_1 = \vec{r}_0 \cdot \vec{v} + \frac{\omega\Delta t + 2\pi n}{k} = d_0 + \frac{\omega\Delta t + 2\pi n}{k} > d_0 \quad (207)$$

Ma questo significa che nell'intervallo di tempo Δt il fronte d'onda considerato si e' allontanato dall'origine ($d_1 > d_0$); e siccome tutti i fronti d'onda sono ortogonali a \vec{v} , ne segue che il fronte d'onda deve essersi spostato *lungo la direzione e nel verso del versore \vec{v}* .

72. Un modo per visualizzare un'onda piana tridimensionale progressiva e' quello di immaginare tutti i suoi fronti d'onda corrispondenti, ad esempio, ai massimi, che si spostano nella direzione del versore \vec{v} restando paralleli, equispaziati fra loro e ortogonali a \vec{v} stesso.

Dovrebbe esservi chiaro che per un'onda piana *bidimensionale* i fronti d'onda non sono dei piani, ma delle rette sul piano xy : nella figura 14 ho rappresentato un'onda piana bidimensionale mettendo in evidenza i fronti d'onda corrispondenti ai massimi e la direzione di movimento dell'onda.

73. Come abbiamo visto, la parte spaziale di un'onda piana e' una funzione periodica monodimensionale con numero d'onda k nella direzione di propagazione (la direzione di \vec{v}).

In questo punto voglio mostrarvi che la parte spaziale puo' essere anche vista come una funzione periodica tridimensionale su un reticolo di Bravais la cui cella primitiva e' costituita da vettori paralleli ai tre assi cartesiani.

Per vederlo, supponete per semplicita' $t = 0$ e inoltre:

$$\vec{r} = (x \ y \ z) \quad (208)$$

$$\vec{v} = (v_1 \ v_2 \ v_3) \quad (209)$$

Allora si ha:

$$\begin{aligned} f(\vec{r}, 0) &= f(x, y, z, 0) = A \cos(k(x \ y \ z) \cdot (v_1 \ v_2 \ v_3)) \\ &= A \cos\left(\frac{2\pi}{\lambda} v_1 x + \frac{2\pi}{\lambda} v_2 y + \frac{2\pi}{\lambda} v_3 z\right) \\ &= A \cos\left(\frac{2\pi}{\lambda/v_1} x + \frac{2\pi}{\lambda/v_2} y + \frac{2\pi}{\lambda/v_3} z\right) \\ &= A \cos\left(\frac{2\pi}{\lambda_x} x + \frac{2\pi}{\lambda_y} y + \frac{2\pi}{\lambda_z} z\right) \\ &= A \cos\left(2\pi \left(\frac{1}{\lambda_x} \ \frac{1}{\lambda_y} \ \frac{1}{\lambda_z}\right) \cdot \vec{r}\right) \end{aligned} \quad (210)$$

Potete verificare molto facilmente che la funzione 210 e' periodica su un reticolo di Bravais generato dai tre vettori:

$$\vec{a}_1 = \lambda_x (1 \ 0 \ 0)$$

$$\vec{a}_2 = \lambda_y (0 \ 1 \ 0)$$

$$\vec{a}_3 = \lambda_z (0 \ 0 \ 1)$$

Infatti, posto $\vec{R} = n_x \vec{a}_1 + n_y \vec{a}_2 + n_z \vec{a}_3$ un generico vettore del reticolo di Bravais, per essa vale:

$$\begin{aligned} f(\vec{r} + \vec{R}, 0) &= A \cos\left(2\pi \left(\frac{1}{\lambda_x} \ \frac{1}{\lambda_y} \ \frac{1}{\lambda_z}\right) \cdot (\vec{r} + \vec{R})\right) = \\ &= A \cos\left(2\pi \left(\frac{1}{\lambda_x} \ \frac{1}{\lambda_y} \ \frac{1}{\lambda_z}\right) \cdot (\vec{r} + n_x \vec{a}_1 + n_y \vec{a}_2 + n_z \vec{a}_3)\right) = \\ &= A \cos\left(2\pi \left(\frac{1}{\lambda_x} \ \frac{1}{\lambda_y} \ \frac{1}{\lambda_z}\right) \cdot ((xyz) + (n_x \lambda_x \ n_y \lambda_y \ n_z \lambda_z))\right) = \\ &= A \cos\left(\frac{2\pi}{\lambda_x} (x + n_x \lambda_x) + \frac{2\pi}{\lambda_y} (y + n_y \lambda_y) + \frac{2\pi}{\lambda_z} (z + n_z \lambda_z)\right) = \\ &= A \cos\left(\frac{2\pi}{\lambda_x} x + \frac{2\pi}{\lambda_y} y + \frac{2\pi}{\lambda_z} z + (n_x + n_y + n_z) 2\pi\right) = \\ &= A \cos\left(\frac{2\pi}{\lambda_x} x + \frac{2\pi}{\lambda_y} y + \frac{2\pi}{\lambda_z} z\right) = f(\vec{r}, 0) \quad \forall \vec{r}, \forall \vec{R} \end{aligned} \quad (211)$$

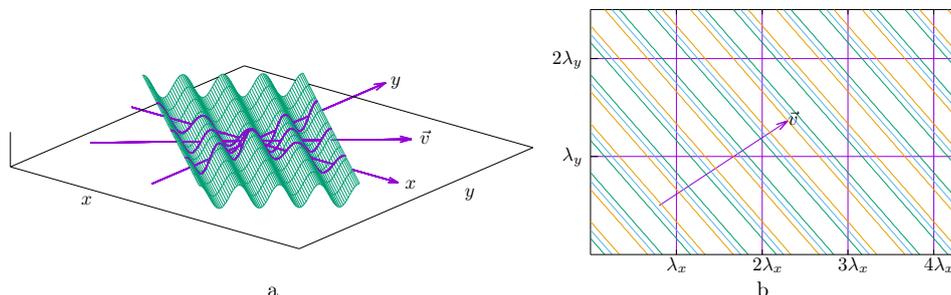


Figura 15: a: sul grafico di un'onda piana bidimensionale (a t fissato) ho indicato le “tracce” sinusoidali su delle sezioni verticali nella direzione di propagazione \vec{v} e nelle direzioni degli assi cartesiani x e y ; tali sinusoidi hanno lunghezza d'onda λ , λ_x e λ_y , rispettivamente; b: il fatto che l'onda piana possa essere descritta in modo equivalente come una funzione periodica bidimensionale su un reticolo di Bravais la cui cella primitiva e' un rettangolo di lati λ_x e λ_y e' messo in evidenza con una rappresentazione mediante isocurve. Dovrebbe esservi evidente, guardando la parte a della figura, che in questo caso le isocurve sono in realta' delle “isorette” sul piano xy . Notate come il “contenuto” di ciascuna cella primitiva sia sempre lo stesso; ho rappresentato anche la direzione del vettore di propagazione \vec{v} , che e' ortogonale alle “isorette”.

Notate che $\lambda = 2\pi/k$ e' la lunghezza d'onda nella direzione di propagazione (cioe' la direzione del vettore \vec{v}) mentre λ_i , $i = x, y, z$ sono le lunghezze d'onda nelle direzioni dei tre assi cartesiani. Tutte queste lunghezze d'onda sono, in generale, diverse tra loro.

Nella figura 15 ho messo in evidenza queste periodicita' per un'onda piana bidimensionale. Dalla figura dovrete rendervi conto che sezionando l'onda piana con un piano ortogonale al piano xy si ottiene *sempre* una sinusoide (tranne quando il piano e' anche ortogonale alla direzione di propagazione). Quando il piano e' parallelo alla direzione di propagazione \vec{v} , la sinusoide risultante ha la lunghezza d'onda minima (che e' $\lambda = 2\pi/k$); quando il piano e' parallelo all'asse x , la lunghezza d'onda della sinusoide ottenuta e' λ_x , e cosi' via.

Non vorrei crearvi confusione: la 210 non e' null'altro che una rappresentazione matematica dell'onda piana equivalente alla 193, ma quest'ultima e' molto piu' semplice e conveniente.

74. Spesso per scrivere l'espressione di un'onda piana in forma piu' compatta si definisce il *vettore d'onda* con:

$$\vec{k} = k\vec{v} = \frac{2\pi}{\lambda}\vec{v} \quad (212)$$

e quindi:

$$f(\vec{r}, t) = A \cos(\vec{k} \cdot \vec{r} - \omega t) \quad (213)$$

75. La parte temporale di un'onda piana e' una funzione periodica del tempo con frequenza angolare ω , come potete rendervi conto facilmente se considerate \vec{r} fissato nella 213:

$$A \cos \left(\overbrace{\vec{k} \cdot \vec{r}}^{\text{costante}} - \omega t \right) \quad (214)$$

76. Onde piane stazionarie.

Analogamente a quanto visto per il caso monodimensionale (punto 68), anche in 2D o 3D possono esistere onde piane stazionarie, la cui forma matematica piu' semplice e' del tipo:

$$f(\vec{r}, t) = A \cos(k\vec{r} \cdot \vec{v}) \cos(\omega t) \quad (215)$$

Di nuovo, notate come la parte spaziale e quella temporale sono separate, ed e' questo che determina la "stazionarieta'" dell'onda.

Anche in piu' dimensioni valgono le stesse considerazioni viste per il caso monodimensionale.

In particolare, i nodi dell'onda piana (che in piu' dimensioni non sono punti, ma rette del piano (2D) o piani nello spazio (3D)) sono indipendenti dal tempo: cioe' i luoghi dei punti dove l'onda ha valore nullo rimangono localizzati nelle stesse regioni del piano o dello spazio indipendentemente dall'istante di tempo considerato.

Nel punto \vec{r}_0 non appartenente a un nodo, l'onda piana oscilla armonicamente in funzione del tempo, con un'ampiezza pari a $A \cos(k\vec{r}_0 \cdot \vec{v})$ (intendo dire che i valori assunti dall'onda oscillano mantenendosi compresi nell'intervallo $[-A \cos(k\vec{r}_0 \cdot \vec{v}), A \cos(k\vec{r}_0 \cdot \vec{v})]$).

77. Forma esponenziale complessa delle onde.

Un formalismo estremamente utile per trattare le onde di tipo sinusoidale e' la loro rappresentazione con la funzione esponenziale complessa.

Considerate l'onda piana:

$$f(\vec{r}, t) = A \cos(\vec{k} \cdot \vec{r} - \omega t) \quad (216)$$

Grazie all'identita' di Eulero (eq. 8) e' evidente che tale onda piana puo' essere considerata come la parte reale dell'esponenziale:

$$A \exp(i(\vec{k} \cdot \vec{r} - \omega t)) = A \cos(\vec{k} \cdot \vec{r} - \omega t) + i \sin(\vec{k} \cdot \vec{r} - \omega t) \quad (217)$$

Quindi un'onda piana puo' essere scritta nella forma:

$$f(\vec{r}, t) = A \exp(i(\vec{k} \cdot \vec{r} - \omega t)) \quad (218)$$

essendo sottinteso che ai fini della rappresentazione del comportamento di un'onda, solo la parte reale ha significato fisico.

L'utilita' di questa rappresentazione sta nel fatto che le manipolazioni algebriche con la funzione esponenziale sono in genere estremamente piu' comode di quelle con le funzioni trigonometriche (questo e' anche il motivo per cui lo sviluppo in serie di Fourier si usa quasi sempre nella sua forma esponenziale complessa, eq. 73-74).

Quello che normalmente si fa e' scrivere un'onda sinusoidale in forma esponenziale complessa, svolgere l'algebra richiesta per risolvere un certo problema, e poi riottenere la funzione trigonometrica finale come la parte reale del risultato complesso:

$$A \cos(\vec{k} \cdot \vec{r} - \omega t) = \Re\left(A \exp\left(i(\vec{k} \cdot \vec{r} - \omega t)\right)\right) \quad (219)$$

Useremo spesso la forma esponenziale complessa nel seguito. Inoltre, usero' il termine "onda piana" anche quando mi riferisco in realta' alla sola parte spaziale. E' chiaro che l'esponenziale:

$$\exp(i\vec{k} \cdot \vec{r}) \quad (220)$$

puo' essere sempre considerato come la parte spaziale di un' onda piana che si ottiene moltiplicandolo per $\exp(-i\omega t)$ (e per un fattore di ampiezza A):

$$A \exp(i\vec{k} \cdot \vec{r}) \exp(-i\omega t) = A \exp\left(i(\vec{k} \cdot \vec{r} - \omega t)\right) \quad (221)$$

A questo proposito, se riguardate la 127, vedete che lo sviluppo in serie di Fourier di una funzione periodica e' fatto in termini della parte spaziale di onde piane per le quali il vettore d'onda \vec{k} ha i valori corrispondenti ai punti del reticolo reciproco associato al reticolo di Bravais che definisce la periodicita' della funzione.

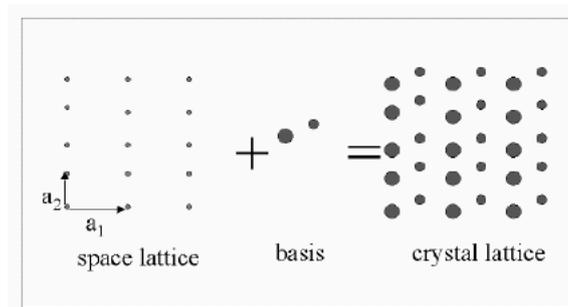


Figura 16:

La struttura dei solidi cristallini

78. Come avrete modo di rendervi conto, gli argomenti del corso che seguono richiedono alcuni dei concetti che stanno alla base della descrizione dei solidi cristallini. Per questo motivo introdurrò quelli di cui avremo bisogno in questa sezione.

Sottolineo che mi limiterò alle parti strettamente funzionali al corso (che *non* è un corso di cristallografia).

79. Do per scontato che tutti abbiate la nozione di un solido cristallino come di un sistema la cui caratteristica peculiare è quella di essere costituito, a livello microscopico, da un raggruppamento di atomi (o ioni o molecole o comunque “oggetti di tipo chimico”) che viene ripetuto *periodicamente* e (in teoria) *indefinitamente* nello spazio.

80. Il raggruppamento atomico che, ripetendosi per traslazione lungo tre direzioni, genera l'intero cristallo viene chiamato *base*. Nel caso più semplice possibile (e nel seguito ci limiteremo a questo) la base è costituita da un solo atomo (e quindi il cristallo macroscopico sarà in generale quello di un elemento, nella maggior parte dei casi, ma non sempre, metallico).

81. La ripetizione periodica della base nello spazio viene descritta in modo astratto con un reticolo tridimensionale detto *reticolo di Bravais* o *reticolo diretto* (o ancora *spazio reale*).

Ad ogni punto del reticolo di Bravais è associata una replica della base che genera il cristallo e in tal modo viene realizzata la sua ripetizione indefinita nello spazio (figura 16).

Nel caso più semplice, ad ogni punto del reticolo di Bravais è associato un singolo atomo e questo sarà l'unico caso che considereremo in questo corso.

Tuttavia, e lo dico perché vedrete che ne faremo uso, l'oggetto astratto associato ad ogni punto reticolare può essere qualsiasi cosa (anche un insieme di punti reticolari ordinati, come vedremo fra breve).

82. Ricorderete che abbiamo già introdotto il concetto di reticolo di Bravais a proposito dello sviluppo in serie di Fourier di funzioni periodiche in più dimensioni (p. 16). Ovviamente stiamo parlando esattamente della stessa cosa e potete rendervene conto alla luce della seguente considerazione.

Siccome gli atomi che costituiscono un cristallo sono ripetuti periodicamente, qualsiasi proprietà fisica che sia legata alla posizione degli atomi *deve essere una*

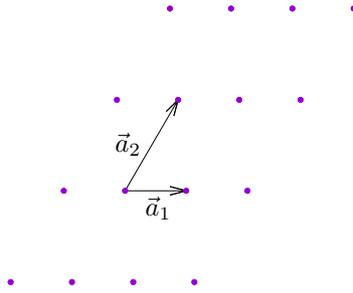


Figura 17:

funzione periodica tridimensionale delle coordinate spaziali e la sua periodicit  deve essere determinata dal reticolo di Bravais su cui la struttura cristallina   basata. Una propriet  fisica di questo tipo  , banalmente, la densit  elettronica: essa   una funzione di x, y, z e chiaramente in un cristallo avr  un andamento periodico nelle tre dimensioni (con massimi in corrispondenza ai nuclei); tale periodicit  deve necessariamente coincidere con quella del cristallo, cio  deve essere basata sul suo reticolo di Bravais. Un altro esempio di una grandezza fisica che ha un andamento periodico in un cristallo   il potenziale generato dalla presenza degli atomi (nuclei ed elettroni).

Dovrebbe pertanto essere chiaro che il reticolo di Bravais che serve a descrivere una struttura cristallina, costituisce anche la base per la descrizione matematica di tutte quelle propriet  fisiche che dalla struttura cristallina dipendono.

In questa parte riprenderemo i concetti di reticolo di Bravais e reticolo reciproco per precisarli e approfondirli ulteriormente.

83. Possiamo dare due definizioni formali equivalenti di reticolo di Bravais.

1. Un reticolo di Bravais   un insieme di punti discreti dello spazio tali che la loro disposizione e orientazione appaiono esattamente le stesse quando l'insieme venga osservato da uno qualsiasi dei suoi elementi
2. Un reticolo di Bravais   l'insieme dei punti dello spazio individuati dai vettori \vec{R} definiti da:

$$\vec{R} = n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3 \quad (222)$$

dove $\vec{a}_1, \vec{a}_2, \vec{a}_3$ sono tre vettori non coplanari (cio  linearmente indipendenti) e $n_1, n_2, n_3 \in \mathbb{Z}$.

I vettori $\vec{a}_1, \vec{a}_2, \vec{a}_3$ vengono detti vettori *primitivi* o *generatori* del reticolo.

Il punto individuato dal vettore \vec{R} si raggiunge spostandosi di n_1 passi di lunghezza a_1 nella direzione di \vec{a}_1 , n_2 passi di lunghezza a_2 nella direzione di \vec{a}_2 e n_3 passi di lunghezza a_3 nella direzione di \vec{a}_3 .

La figura 17 mostra una parte di un reticolo di Bravais bidimensionale con i due vettori primitivi (che, nel caso 2D, devono essere non paralleli).

La figura 18 mostra il pi  semplice reticolo di Bravais tridimensionale: il *reticolo cubico semplice*, i cui vettori primitivi hanno lo stesso modulo e sono orientati lungo i tre assi cartesiani.

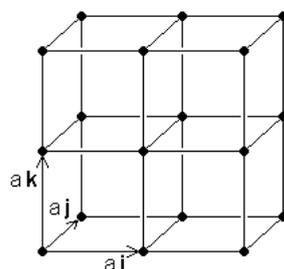


Figura 18:

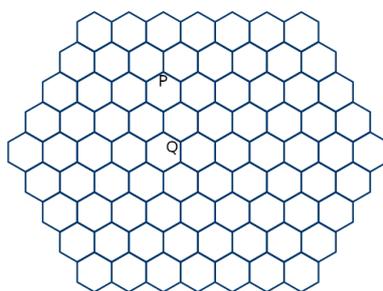


Figura 19:

84. Notate che non tutti i reticoli regolari di punti sono reticoli di Bravais. L'esempio piu' citato e' quello del reticolo bidimensionale "a nido d'ape" mostrato nella figura 19. Tutti i punti di questo reticolo sono connessi ai punti piu' vicini nello stesso modo, ma l'*orientazione* di queste connessioni non e' la stessa per tutti. Il punto P ha tre punti ad esso piu' vicini e alla stessa distanza: uno "in alto", uno "in basso a destra" e uno "in basso a sinistra"; il punto Q e' circondato da tre punti alla stessa distanza di quelli che circondano il punto P , ma visti da Q , questi tre punti sono uno "in basso", uno "in alto a destra" e uno "in alto a sinistra" (cioe' sono l'immagine speculare dei punti visti da P). Siccome il reticolo non e' lo stesso se visto dai punti P o Q , il reticolo a nido d'ape non soddisfa la definizione 1 e quindi non e' un reticolo di Bravais.

85. La scelta dei vettori primitivi di un reticolo di Bravais non e' univoca; in realta', esistono infinite terne di vettori primitivi che generano lo stesso reticolo. Questo e' mostrato nella figura 20 per un reticolo bidimensionale.

86. Fissato un certo punto P di un reticolo di Bravais, i punti alla minima distanza da P vengono detti *primi vicini* di P . Quelli alla distanza immediatamente maggiore vengono detti *secondi vicini* e cosi' via.

Chiaramente, per definizione di reticolo di Bravais, ogni punto del reticolo ha esattamente gli stessi primi, secondi etc vicini, orientati nello stesso modo.

Tutti i primi vicini di un determinato punto toccano la superficie di una sfera (in 3D) o una circonferenza (in 2D) il cui raggio e' uguale alla distanza comune dal punto di riferimento. Lo stesso vale per i secondi vicini e cosi' via. Le sfere dei primi, secondi etc vicini sono concentriche e ordinabili sulla base del loro

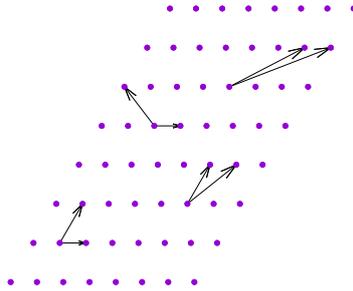


Figura 20: La scelta dei vettori primitivi di un reticolo di Bravais non e' univoca. In questa figura sono mostrate 4 diverse scelte equivalenti dei vettori generatori di un reticolo bidimensionale.

FIXME

Figura 21:

raggio.

Generalmente, il numero e la distanza (comune) dei primi, secondi etc vicini dipende dal tipo di reticolo di Bravais.

Inoltre, il numero dei vicini aumenta all'aumentare del raggio della corrispondente sfera e la differenza fra i raggi di due sfere successive diminuisce all'aumentare della distanza dal punto di riferimento.

Questo e' mostrato nella figura 21 per un reticolo cubico semplice.

87. Cella primitiva.

Come avevo gia' accennato (punto 47) i tre vettori primitivi $\vec{a}_1, \vec{a}_2, \vec{a}_3$ di un reticolo di Bravais definiscono una *cella primitiva* o *elementare*, che e' il parallelepipedo i cui (12) lati sono costituiti dai tre vettori, opportunamente traslati.

Gli 8 vertici della cella primitiva si ottengono combinando i tre vettori primitivi secondo (fig. 22):

$$n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3 \quad n_1, n_2, n_3 = 0, 1 \quad (223)$$

Con *parametri di cella* si intende normalmente l'insieme delle lunghezze e angoli che definiscono univocamente una certa cella elementare.

La cella primitiva di un reticolo di Bravais gode delle seguenti proprieta':

1. quando viene traslata secondo tutti i vettori \vec{R} del reticolo essa riempie tutto lo spazio senza mai sovrapporsi a se' stessa ne' lasciare vuoti; quindi ogni replica della cella primitiva del reticolo e' associata in modo univoco ad un particolare vettore reticolare \vec{R} ; un reticolo di Bravais puo' essere pensato in modo equivalente come generato dalla traslazione della sua cella primitiva tramite gli infiniti vettori $\vec{R} = n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3$ con $n_1, n_2, n_3 \in \mathbb{Z}$
2. una volta costruita secondo l'eq. 223, la cella primitiva puo' essere traslata (ma non ruotata!) di un vettore qualsiasi (non necessariamente facente parte del reticolo di Bravais) continuando a godere della proprieta' 1

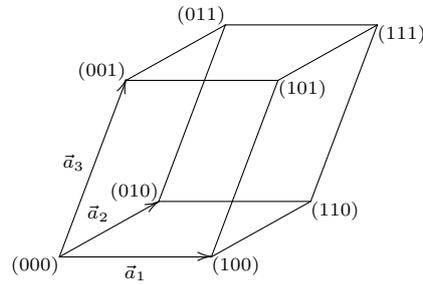


Figura 22: Il parallelepipedo formato dai tre vettori primitivi combinati secondo la 223. Per ogni vertice e' mostrata la terna di valori interi (n_1, n_2, n_3) con cui sono stati combinati i vettori primitivi.

3. la cella primitiva contiene sempre e solo un punto reticolare; questo puo' essere verificato trasladandola opportunamente (grazie alla proprieta' 2) in modo che non ci sia alcun punto reticolare sulle sue facce o spigoli o vertici; alternativamente, il numero di punti contenuti nella cella primitiva si puo' calcolare tenendo conto che:

- ⇒ un punto su una faccia conta solo per $1/2$, essendo condiviso fra 2 celle
- ⇒ un punto su uno spigolo conta solo per $1/4$, essendo condiviso fra 4 celle
- ⇒ un punto su un vertice conta solo per $1/8$, essendo condiviso fra 8 celle

(naturalmente, nel caso di un reticolo bidimensionale (che usero' spesso per fare degli esempi perche' e' molto piu' semplice da visualizzare e disegnare), la cella primitiva e' un parallelogramma e quindi non ha facce; inoltre, in questo caso, le frazioni da considerare saranno: lato $\rightarrow 1/2$, vertice $\rightarrow 1/4$)

Nella figura 23 ho mostrato i due metodi per il conteggio del numero di punti reticolari contenuti nella cella primitiva di un reticolo bidimensionale.

4. tutte le infinite possibili celle primitive di un reticolo di Bravais hanno lo stesso volume. Infatti, detta ρ la densita' (costante!) di punti reticolari per unita' di volume del reticolo, il volume v di *qualsiasi* cella primitiva (che contiene un solo punto reticolare per ipotesi) e':

$$v = \frac{1}{\rho} \quad (\text{costante}) \quad (224)$$

5. come gia' accennato al punto 1 piu' sopra, siccome la cella primitiva contiene un solo punto reticolare, esiste una relazione biunivoca fra le celle primitive e i vettori del reticolo di Bravais: ogni vettore \vec{R} individua una e una sola cella primitiva e, corrispondentemente, per ogni cella primitiva esiste uno e un solo vettore \vec{R} che la individua.

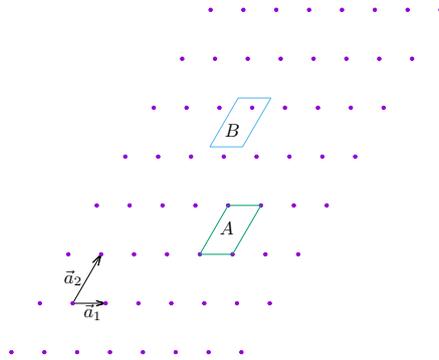


Figura 23: I vettori primitivi \vec{a}_1 e \vec{a}_2 del reticolo di Bravais bidimensionale mostrato determinano una cella primitiva di cui A e B sono due repliche identiche. Il numero di punti reticolari contenuti nella cella A si ottiene tenendo conto che ciascuno dei 4 punti ai vertici e' condiviso da 4 celle e quindi conta solo per $1/4$: n. di punti = $4 \times 1/4 = 1$. La cella B e' stata traslata di $0.4\vec{a}_1 + 0.2\vec{a}_2$ in modo che non abbia punti reticolari sui lati o vertici: in tal modo si vede subito che essa contiene solo un punto reticolare.

88. Supercelle.

Per alcune applicazioni e' conveniente definire e utilizzare una cella *non primitiva* ottenuta, nel caso piu' semplice, "unendo" n_1 celle primitive contigue lungo la direzione di \vec{a}_1 , n_2 celle primitive contigue lungo la direzione di \vec{a}_2 ed n_3 celle primitive contigue lungo la direzione di \vec{a}_3 . Chiameremo una cella ottenuta in questo modo "*supercella* $n_1 \times n_2 \times n_3$ " o "cella unitaria" (avevo gia' accennato al concetto di supercella in una sola dimensione, al punto 5).

Ovviamente, una supercella contiene piu' di un punto reticolare; in particolare, siccome e' costituita da $N = n_1 n_2 n_3$ celle primitive, ciascuna contenente un singolo punto reticolare, una supercella $n_1 \times n_2 \times n_3$ conterra' N punti reticolari. I tre vettori che definiscono una supercella (del tipo semplice che stiamo considerando) sono evidentemente dei multipli dei vettori che definiscono la cella primitiva da cui la supercella e' stata ottenuta. In particolare, i vettori $\vec{A}_1, \vec{A}_2, \vec{A}_3$ che definiscono una supercella $n_1 \times n_2 \times n_3$ sono legati ai vettori della cella primitiva $\vec{a}_1, \vec{a}_2, \vec{a}_3$ dalle relazioni:

$$\vec{A}_1 = n_1 \vec{a}_1 \quad (225)$$

$$\vec{A}_2 = n_2 \vec{a}_2 \quad n_1, n_2, n_3 \in \mathbb{N} \quad (226)$$

$$\vec{A}_3 = n_3 \vec{a}_3 \quad (227)$$

E' importante realizzare che i vettori \vec{A}_i che costituiscono una supercella generano, *da soli*, un reticolo i cui punti sono un *sottoinsieme* di quelli del reticolo generato dai vettori primitivi \vec{a}_i . Tuttavia, il reticolo di Bravais "piu' denso" viene generato dalla supercella quando essa viene replicata per traslazione *assieme al suo contenuto di punti reticolari*. Questi possono quindi essere visti come

FIGME

Figura 24:

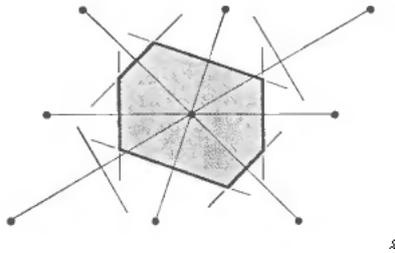


Figura 25: Costruzione della cella primitiva di Wigner–Seitz per un reticolo bidimensionale.

la base (punto 81) associata ad ogni punto del reticolo “meno denso” (cioè quello generato dai vettori \vec{A}_i).

Dovrebbe altresì essere chiaro che il reticolo generato dai vettori \vec{A}_i è anch'esso un reticolo di Bravais secondo le definizioni che abbiamo dato (p. 57).

Ho mostrato la relazione fra cella primitiva e supercella per un reticolo bidimensionale nella figura 24.

In definitiva, un certo reticolo di Bravais può essere generato o per traslazione di una sua cella primitiva contenente un solo punto reticolare, oppure per traslazione di una supercella, contenente un certo numero (> 1) di punti reticolari.

89. Cella primitiva di Wigner–Seitz.

Abbiamo detto che la scelta dei vettori primitivi e quindi della cella primitiva di un certo reticolo di Bravais non è univoca.

Esiste tuttavia una procedura per generare una cella primitiva “standard”, detta *cella di Wigner–Seitz*.

La procedura di Wigner–Seitz è mostrata nella figura 25 per un reticolo bidimensionale.

In pratica: scelto un punto reticolare, lo si connette a tutti i suoi primi, secondi etc vicini e si tracciano delle rette che: (i) sono ortogonali a ciascun segmento di connessione e (ii) lo intersecano nel punto medio. La superficie di area minima racchiusa dalle rette così tracciate è la cella primitiva di Wigner–Seitz.

Passando a un reticolo 3D la procedura è esattamente la stessa, con le dovute sostituzioni di termini; si connette un punto reticolare a tutti i suoi primi, secondi etc vicini e si tracciano dei piani che: (i) sono ortogonali a ciascun segmento di connessione e (ii) lo intersecano nel punto medio. Il poliedro di volume minimo racchiuso dai piani così tracciati è la cella primitiva di Wigner–Seitz.

Ovviamente, per come è costruita, la cella di Wigner–Seitz contiene un solo punto reticolare. Inoltre, sempre per costruzione, ogni lato della cella di Wigner–Seitz centrata su un dato punto è condiviso col lato della stessa cella centrata su

un altro punto (guardate la figura): ma questo significa che la cella di Wigner–Seitz riempie tutto lo spazio senza mai sovrapporsi a se' stessa ne' lasciare vuoti, cioe' e' una cella primitiva (in base alla proprieta' 1 di p. 59).

La cella di Wigner–Seitz puo' essere definita in modo equivalente come la regione di spazio i cui punti sono piu' vicini al punto reticolare contenuto nella cella di quanto lo siano rispetto a qualsiasi altro punto reticolare.

90. Classificazione dei reticoli di Bravais.

Tutti i possibili reticoli di Bravais sono classificabili in base alle loro caratteristiche di simmetria.

Trattero' molto superficialmente questo argomento, visto che nel corso non useremo quasi mai questa classificazione (se non in modo estremamente intuitivo e informale). Tenete tuttavia presente che la simmetria dei reticoli di Bravais occupa un posto molto importante nello studio della struttura dei solidi cristallini.

Un'operazione di simmetria e' tale per un reticolo di Bravais se lo trasforma in se' stesso. Tutte le operazioni di simmetria valide per i reticoli di Bravais possono essere sempre decomposte in una traslazione pari ad un vettore reticolare \vec{R} seguita da una operazione di simmetria (rotazione, riflessione o inversione) che lascia un punto del reticolo fisso (notate che una traslazione non lascia alcun punto reticolare fisso).

Quando si considerano tutte le possibili operazioni di simmetria di un dato reticolo, si puo' dimostrare (in modo non banale) che ci sono solo 14 gruppi di operazioni di simmetria possibili. Questi 14 gruppi di operazioni di simmetria definiscono altrettanti "prototipi" di reticoli di Bravais, che esauriscono tutte le possibilita'.

Quindi ci sono solo 14 "prototipi" di reticoli di Bravais. Essi sono definiti dal tipo di cella e dai suoi parametri.

Nella figura 26 sono mostrati i 14 reticoli di Bravais con le caratteristiche in base alle quali sono classificati.

Vale la pena di sottolineare che, come potete vedere dalla tabella 26, i particolari valori dei parametri di cella non hanno alcuna importanza; sono le loro *relazioni reciproche* che definiscono un particolare tipo di reticolo di Bravais. Ad esempio due reticoli cubici semplici che hanno un diverso valore del lato della cella elementare appartengono allo stesso tipo di reticolo di Bravais (cubico semplice), e cosi' via.

91. Cella convenzionale.

Ai fini della classificazione, puo' succedere che nessuna cella primitiva di un dato reticolo possieda *tutti* i suoi elementi di simmetria (ad esempio il reticolo cubico a facce centrate). In questi casi si considera una cosiddetta *cella convenzionale*, costruita in modo tale che possieda *la simmetria completa* del reticolo. Una cella convenzionale e' una supercella (cioe' contiene piu' di un punto reticolare e genera l'intero reticolo per traslazione). Tuttavia, una cella convenzionale e' ottenuta in un modo piu' complesso di quello che abbiamo visto al punto 88, mediante un'opportuna combinazione lineare dei vettori primitivi:

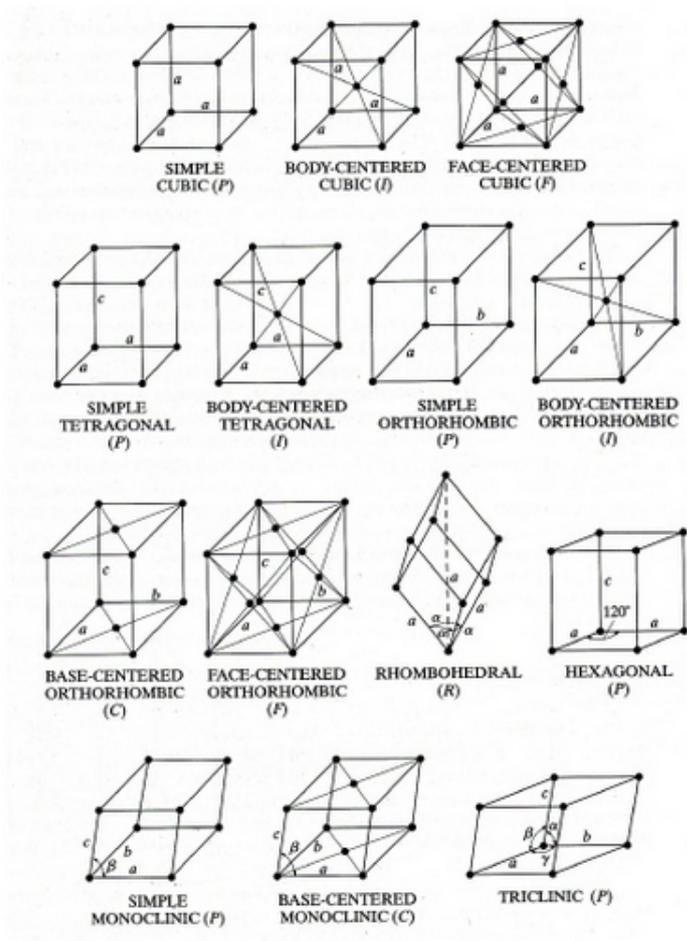


Figura 26:

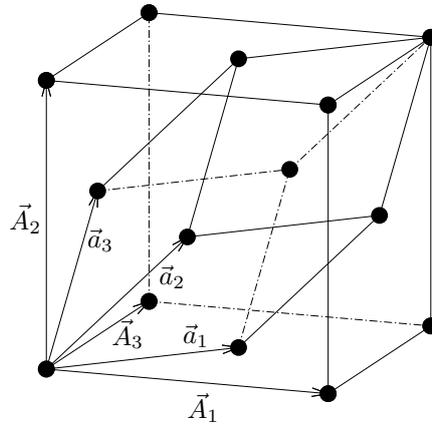


Figura 27: Relazione fra cella primitiva (definita dai vettori \vec{a}_i) e cella convenzionale (definita dai vettori \vec{A}_i) per il reticolo cubico a facce centrate.

$$\vec{A}_1 = c_{11}\vec{a}_1 + c_{12}\vec{a}_2 + c_{13}\vec{a}_3 \quad (228)$$

$$\vec{A}_2 = c_{21}\vec{a}_1 + c_{22}\vec{a}_2 + c_{23}\vec{a}_3 \quad (229)$$

$$\vec{A}_3 = c_{31}\vec{a}_1 + c_{32}\vec{a}_2 + c_{33}\vec{a}_3 \quad (230)$$

Notate che le 225-227 sono un caso particolare delle 228-230 con:

$$c_{ij} = 0 \quad \text{quando } i \neq j$$

$$c_{ij} = n_i \in \mathbb{N} \quad \text{quando } i = j$$

Nella figura 27 sono mostrate la cella primitiva e quella convenzionale per il reticolo cubico a facce centrate. La cella convenzionale e' una supercella che contiene 4 punti reticolari.

Detto a il lato della cella cubica (convenzionale), i vettori della cella primitiva hanno coordinate cartesiane:

$$\vec{a}_1 = a(1/2 \ 1/2 \ 0) \quad (231)$$

$$\vec{a}_2 = a(1/2 \ 0 \ 1/2) \quad (232)$$

$$\vec{a}_3 = a(0 \ 1/2 \ 1/2) \quad (233)$$

e quelli della cella convenzionale:

$$\vec{A}_1 = a(100) \quad (234)$$

$$\vec{A}_2 = a(001) \quad (235)$$

$$\vec{A}_3 = a(010) \quad (236)$$

Le due terne di vettori sono legate dalla trasformazione:

$$\begin{pmatrix} \vec{A}_1 \\ \vec{A}_2 \\ \vec{A}_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} \vec{a}_1 \\ \vec{a}_2 \\ \vec{a}_3 \end{pmatrix} \quad (237)$$

dove gli elementi della matrice al secondo membro sono i c_{ij} delle 228–230.

92. Il reticolo reciproco.

Data una struttura cristallina basata su un certo reticolo di Bravais, ad essa rimane associato un reticolo reciproco definito dalle eq 112-114 e i cui punti reticolari individuano i vettori d'onda che definiscono lo sviluppo in serie di Fourier di qualsiasi funzione periodica (densità elettronica, energia potenziale etc) sul reticolo di Bravais (eq. 127 e 128).

Il reticolo reciproco ha molte altre proprietà e in questa parte ne vedremo alcune.

1. Il reticolo reciproco è esso stesso un reticolo di Bravais.

La dimostrazione segue immediatamente dalla definizione del reticolo reciproco (eq 112-114). Siccome i vettori primitivi del reticolo di Bravais non sono coplanari (per ipotesi) e i vettori primitivi del reticolo reciproco sono ortogonali ai tre piani individuati dai vettori primitivi del reticolo di Bravais presi a coppie (punto 48), anche i vettori primitivi del reticolo reciproco non sono coplanari.

Quindi il reticolo reciproco soddisfa la definizione 2 di p. 57 ed è un reticolo di Bravais.

2. Siccome abbiamo appena visto che il reticolo reciproco di un dato reticolo di Bravais è a sua volta un reticolo di Bravais, ha senso porsi la domanda di quale sia il reticolo reciproco del reticolo reciproco. Ebbene, detto \mathcal{R} un reticolo di Bravais e \mathcal{R}^* il suo reticolo reciproco, vale:

$$(\mathcal{R}^*)^* = \mathcal{R} \quad (238)$$

cioè: il reticolo reciproco di \mathcal{R}^* è il reticolo di Bravais di cui \mathcal{R}^* è il reciproco.

La dimostrazione è la seguente.

Essendo \mathcal{R}^* un reticolo di Bravais, ogni funzione con la sua periodicità può essere sviluppata in serie di Fourier su $(\mathcal{R}^*)^*$, che è il suo reticolo reciproco (eq. 127):

$$f(\vec{q}) = \sum_{\vec{G}} B_{\vec{G}} \exp(i\vec{G} \cdot \vec{q}) \quad (239)$$

dove \vec{q} è una variabile *continua* di \mathcal{R}^* e \vec{G} è un vettore di $(\mathcal{R}^*)^*$.

Gli esponenziali complessi dello sviluppo devono avere la stessa periodicità della f (punto 13) e quindi, detto $\vec{k} = \sum_j n_j \vec{b}_j$ un generico vettore di \mathcal{R}^* , deve essere:

$$\begin{aligned} \exp(i\vec{G} \cdot \vec{q}) = \exp(i\vec{G} \cdot (\vec{q} + \vec{k})) &\iff \exp(i\vec{G} \cdot \vec{k}) = 1 \\ \iff \vec{G} \cdot \vec{k} = n(2\pi) \quad n \in \mathbb{Z} \quad \forall \vec{k} \in \mathcal{R}^* & \end{aligned} \quad (240)$$

(per l'ultima doppia implicazione usate l'identita' di Eulero, eq. 10)

Ora, e' senz'altro vero che ogni vettore $\vec{G} = \sum_j g_j \vec{a}_j$ di \mathcal{R} (con $g_j \in \mathbb{Z}$) soddisfa la 240 in base alla 122 (tenete presente che $\vec{G} \in \mathcal{R}$ e $\vec{k} \in \mathcal{R}^*$).

Quindi tutti i vettori di \mathcal{R} fanno parte di $(\mathcal{R}^*)^*$.

D'altro canto, nessun altro vettore \vec{v} dello spazio diretto puo' fare parte di $(\mathcal{R}^*)^*$. Infatti, se cosi' fosse, siccome i vettori primitivi di \mathcal{R} sono una base (in senso algebrico) dello spazio diretto, \vec{v} dovrebbe essere del tipo:

$$\vec{v} = x_1 \vec{a}_1 + x_2 \vec{a}_2 + x_3 \vec{a}_3 \quad (241)$$

con almeno un coefficiente non intero. Ma allora, con un ragionamento identico a quello fatto al punto 50, questo implicherebbe l'esistenza di un vettore \vec{k} per il quale:

$$\vec{v} \cdot \vec{k} = 2\pi \sum_j x_j n_j \quad \text{con} \quad \sum_j x_j n_j \notin \mathbb{Z} \quad (242)$$

e quindi la 240 non sarebbe soddisfatta.

In conclusione, tutti i vettori di \mathcal{R} fanno parte di $(\mathcal{R}^*)^*$, nessun altro vettore puo' far parte di $(\mathcal{R}^*)^*$, e quindi:

$$(\mathcal{R}^*)^* \equiv \mathcal{R} \quad (243)$$

3. Il volume v della cella primitiva del reticolo diretto e quello v^* della cella primitiva del reticolo reciproco sono legati dalla relazione:

$$v^* = \frac{(2\pi)^3}{v} \quad (244)$$

(un'altra relazione di reciprocita')

Per dimostrare questa proprieta' bisogna conoscere un'identita' che riguarda la combinazione di prodotto scalare e vettoriale, e cioe':

$$(\vec{a} \times \vec{b}) \cdot (\vec{q} \times \vec{r}) = (\vec{a} \cdot \vec{q})(\vec{b} \cdot \vec{r}) - (\vec{a} \cdot \vec{r})(\vec{b} \cdot \vec{q}) \quad (245)$$

Allora: il volume della cella reciproca e' (rivedete l'eq. 116):

$$v^* = \left| \vec{b}_1 \cdot (\vec{b}_2 \times \vec{b}_3) \right| \quad (246)$$

dove i \vec{b}_i sono i vettori primitivi del reticolo reciproco.

Esprimiamo \vec{b}_1 in funzione dei vettori primitivi \vec{a}_i del reticolo diretto (eq. 112), usiamo la proprieta' 245 e le relazioni di ortogonalita' 118:

$$= \left| \overbrace{2\pi \frac{\vec{a}_2 \times \vec{a}_3}{\vec{a}_1 \cdot \vec{a}_2 \times \vec{a}_3}}^{\vec{b}_1} \cdot (\vec{b}_2 \times \vec{b}_3) \right| \quad (247)$$

$$= \left| 2\pi \frac{\overbrace{(\vec{a}_2 \cdot \vec{b}_2)}^{=2\pi} \overbrace{(\vec{a}_3 \cdot \vec{b}_3)}^{=2\pi} - \overbrace{(\vec{a}_2 \cdot \vec{b}_3)}^{=0} \overbrace{(\vec{a}_3 \cdot \vec{b}_2)}^{=0}}{\vec{a}_1 \cdot \vec{a}_2 \times \vec{a}_3} \right| \quad (248)$$

$$= \frac{(2\pi)^3}{|\vec{a}_1 \cdot \vec{a}_2 \times \vec{a}_3|} = \frac{(2\pi)^3}{v} \quad (249)$$

Nel caso bidimensionale, come potete facilmente intuire, la relazione e':

$$A^* = \frac{(2\pi)^2}{A} \quad (250)$$

dove A^* e A sono, rispettivamente, l'area della cella primitiva reciproca e quella della cella primitiva diretta.

Il risultato si ottiene nel modo seguente.

Siccome \vec{b}_1 e \vec{b}_2 sono ortogonali ad \vec{a}_2 e \vec{a}_1 , rispettivamente, l'angolo θ (sempre assunto $< \pi$) fra \vec{b}_1 e \vec{b}_2 e' uguale a quello fra \vec{a}_1 e \vec{a}_2 .

Allora:

$$A^* = b_1 b_2 \sin(\theta) \quad (251)$$

Sostituiamo a b_1 e b_2 le espressioni 104 e 105, rispettivamente:

$$= 2\pi \frac{a_2}{A} 2\pi \frac{a_1}{A} \sin(\theta) = \frac{(2\pi)^2}{A^2} \overbrace{a_1 a_2 \sin(\theta)}^{=A} = \frac{(2\pi)^2}{A} \quad (252)$$

Infine, nel caso monodimensionale, abbiamo gia' visto che la relazione e':

$$a^* = \frac{2\pi}{a} \quad (253)$$

dove a^* e a sono, rispettivamente, la lunghezza della "cella" primitiva reciproca e quella della "cella" primitiva diretta (punto 31).

4. Zone di Brillouin.

Definiamo ora le cosiddette *zone di Brillouin*, perche' ci serviranno nel seguito.

Le zone di Brillouin si possono considerare una sorta di generalizzazione della cella primitiva di Wigner–Seitz (punto 89). Tradizionalmente esse sono sempre riferite al reticolo reciproco, ma, come vedrete fra un attimo, la procedura per la definizione e la costruzione delle zone di Brillouin prescinde completamente dalla natura del reticolo considerato.

Si parla di zone (cioè al plurale) perché ce ne sono infinite.

La prima zona di Brillouin per il reticolo reciproco è la sua cella primitiva costruita in modo identico a quello con cui si costruisce la cella primitiva di Wigner–Seitz per il reticolo diretto.

Detto in altri termini: se costruite la cella primitiva di Wigner–Seitz per il reticolo reciproco (con la procedura descritta al punto 89) avete ottenuto la prima zona di Brillouin.

Le zone successive si costruiscono allo stesso modo, continuando a connettere il punto centrale preso come riferimento con i punti che gli stanno attorno e tracciando dei piani (nel caso 3D) o delle rette (nel caso 2D) che siano ortogonali alle connessioni e le intersechino nel punto medio.

Le regioni del reticolo reciproco così ottenute sono ordinabili in base al loro volume (nel caso 3D) o area (nel caso 2D). Indichiamo tali volumi ordinati con:

$$V_1 V_2 V_3 \dots$$

La regione di volume (o area) V_1 è, come già detto, la prima zona di Brillouin.

Per le zone successive fate riferimento alla figura 28 che mostra il procedimento per un reticolo bidimensionale.

La seconda zona di Brillouin è la regione di volume (o area) V_2 da cui, però, si deve “sottrarre” V_1 (uso il termine “sottrarre” non in senso algebrico, ma nel senso che la regione corrispondente a V_1 non fa parte della seconda zona di Brillouin).

La terza zona di Brillouin è la regione di volume (o area) V_3 , da cui si sottraggono le regioni corrispondenti a V_1 e V_2 . E così via.

Potete trovare una rappresentazione “esagerata” e coloratissima di una lunga successione di zone di Brillouin per alcuni reticoli bidimensionali all’indirizzo web del riferimento 6.

Come vi dicevo, la procedura funzionerebbe ugualmente bene anche nel reticolo diretto, ma, come vedrete, sono solo le zone di Brillouin del reticolo reciproco ad essere importanti per le applicazioni.

Siccome ci capiterà di considerare le zone di Brillouin per un “reticolo reciproco” monodimensionale, diamone una descrizione a questo punto (guardate la figura 29).

Il caso monodimensionale è particolarmente semplice: la procedura è sempre la stessa, ma in questo caso i primi, secondi etc vicini sono molto facili da individuare.

I primi vicini del punto all’origine, P_0 , sono il punto alla sua sinistra (P_{-1}) e quello alla sua destra (P_1), entrambi a distanza $2\pi/a$ da P_0 (a è la

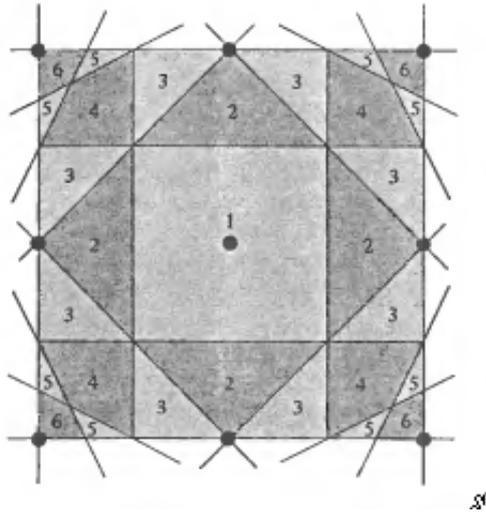


Figura 28: Costruzione delle prime zone di Brillouin per un reticolo quadrato bidimensionale. Le zone sono numerate; i punti del reticolo sono i nove tondi neri. Solo le prime tre zone di Brillouin sono completamente contenute nella figura; delle altre (4, 5, 6) e' rappresentata solo la parte contenuta nel quadrato delimitato dai punti reticolari diversi da quello al centro.

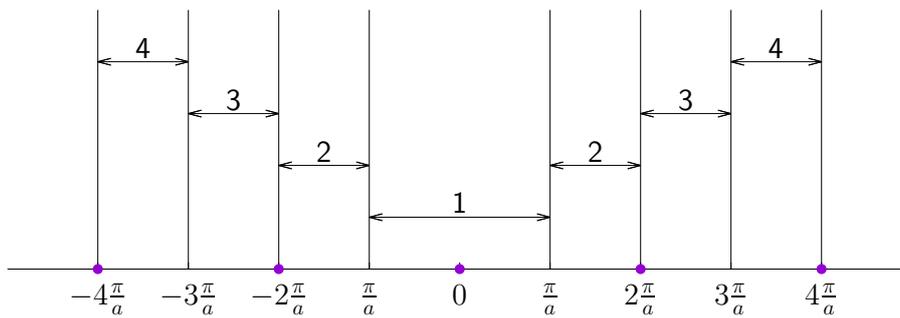


Figura 29: Le prime quattro zone di Brillouin per un reticolo reciproco monodimensionale (a e' la lunghezza della cella diretta).

FIGURE

Figura 30:

lunghezza della “cella primitiva” del reticolo diretto, punto 31): la prima zona di Brillouin e' l'intervallo dell'asse reciproco delimitato dal punto medio del segmento $P_{-1} - P_0$ e dal punto medio del segmento $P_0 - P_1$, cioe': $[-\pi/a, \pi/a]$. I secondi vicini di P_0 sono P_{-2} , alla distanza $4\pi/a$ nel verso negativo, e P_2 , alla stessa distanza, ma nel verso positivo: la seconda zona di Brillouin e' costituita quindi dai due intervalli $[-2\pi/a, -\pi/a]$ e $[\pi/a, 2\pi/a]$ (come gia' detto, la prima zona di Brillouin *non* e' inclusa nella seconda). E cosi' via.

Quindi i confini fra le zone di Brillouin (che nel caso monodimensionale si riducono a dei punti) sono individuati da:

$$n \frac{\pi}{a} \quad n \in \mathbb{Z} - \{0\} \quad (254)$$

5. Piani reticolari e vettori reciproci.

Dato un reticolo di Bravais, si definisce *piano reticolare* un piano individuato da tre punti del reticolo non collineari. Data la simmetria traslazionale del reticolo di Bravais, un piano reticolare contiene infiniti punti del reticolo (che costituiscono, a loro volta, un reticolo bidimensionale). Inoltre, sempre per la simmetria del reticolo di Bravais, dato un particolare piano reticolare, esistono infiniti piani reticolari che sono paralleli ad esso ed equispaziati. L'insieme di questi piani viene detto *famiglia* di piani reticolari. I piani di una famiglia contengono tutti i punti di un reticolo di Bravais, cioe': un reticolo di Bravais puo' essere completamente ripartito fra i piani di una famiglia.

La figura 30 mostra alcuni piani reticolari per un reticolo di Bravais cubico semplice.

Evidentemente, per un dato reticolo di Bravais, esistono infinite famiglie di piani che differiscono per la loro orientazione.

Queste famiglie possono essere classificate sulla base dei vettori del reticolo reciproco grazie al seguente teorema.

Per ogni famiglia di piani reticolari equispaziati alla distanza d esistono vettori (fra loro paralleli) del reticolo reciproco che sono ortogonali ai piani della famiglia e il piu' corto di tali vettori ha modulo $2\pi/d$.

Vale anche il viceversa: per ogni vettore reciproco \vec{k} esiste una famiglia di piani reticolari ortogonali a \vec{k} ed equispaziati alla distanza d , dove $2\pi/d$ e' il modulo del piu' corto vettore reciproco parallelo a \vec{k} .

La dimostrazione della prima parte e' la seguente. Chiamiamo \mathcal{R} il reticolo diretto, \vec{R} un generico vettore del reticolo diretto e \mathcal{R}^* il reticolo reciproco.

Sia inoltre \vec{v} il versore perpendicolare alla famiglia di piani reticolari equispaziati alla distanza d .

(a) Preso:

$$\vec{k} = \frac{2\pi}{d}\vec{v} \quad (255)$$

l'onda piana¹⁸ $\exp(i\vec{k} \cdot \vec{r})$ si propaga nella direzione di \vec{v} e i piani della famiglia sono suoi fronti d'onda, perche' la loro spaziatura e' uguale alla lunghezza d'onda, $\lambda = 2\pi/k = d$ (punto 70)

(b) Siccome qualsiasi famiglia di piani reticolari contiene tutti i punti di \mathcal{R} , un piano della famiglia considerata passa sicuramente per l'origine (c'e' sempre un punto reticolare all'origine). Per tutti i vettori \vec{R} di tale piano si ha:

$$\exp(i\vec{k} \cdot \vec{R}) = 1 \quad (256)$$

perche' tutti i vettori \vec{R} contenuti nel piano reticolare passante per l'origine sono ortogonali a \vec{k} .

(c) Ma allora si avra':

$$\exp(i\vec{k} \cdot \vec{R}) = 1 \quad (257)$$

su *tutti i piani della famiglia* perche' l'onda piana $\exp(i\vec{k} \cdot \vec{r})$ assume lo stesso valore su tutti i suoi fronti d'onda

(d) Quindi:

$$\exp(i\vec{k} \cdot \vec{R}) = 1 \quad \forall \vec{R} \in \mathcal{R} \quad (258)$$

perche' la famiglia di piani contiene tutti i punti di \mathcal{R} .

(e) Ma la 258 implica:

$$\vec{k} \cdot \vec{R} = 2\pi n \quad n \in \mathbb{Z} \quad \forall \vec{R} \in \mathcal{R} \quad (259)$$

e quindi $\vec{k} \in \mathcal{R}^*$ in base alla definizione di p. 32.

(f) \vec{k} e' il piu' corto vettore di \mathcal{R}^* ortogonale ai piani della famiglia. Se esistesse un vettore \vec{k}' di \mathcal{R}^* con modulo $k' < k$ e ortogonale alla stessa famiglia di piani reticolari, la corrispondente onda piana $\exp(i\vec{k}' \cdot \vec{r})$ avrebbe una lunghezza d'onda $\lambda = 2\pi/k' > d$ e quindi ci sarebbero piani reticolari della famiglia che non sarebbero fronti dell'onda piana e sui quali, quindi, la 258 (ovvero la 259) non sarebbe valida; ma allora \vec{k}' non sarebbe un punto di \mathcal{R}^* (perche' non soddisferebbe la definizione di p. 32).

¹⁸Molto spesso parlero' di onde piane considerandone solo la parte spaziale. Ovviamente e' sottinteso che l'onda piana di cui sto parlando e' la parte spaziale moltiplicata per la corrispondente parte temporale:

$$\exp(i\vec{k} \cdot \vec{r}) \exp(-i\omega t) = \exp(i(\vec{k} \cdot \vec{r} - \omega t))$$

Per dimostrare la seconda parte del teorema si procede come segue.

- (a) Prendiamo un vettore $\vec{k} \in \mathcal{R}^*$ che abbia il modulo minimo fra tutti i vettori di \mathcal{R}^* ad esso paralleli e consideriamo tutti i fronti dell'onda piana $\exp(i\vec{k} \cdot \vec{r})$ definiti da:

$$\exp(i\vec{k} \cdot \vec{r}) = 1 \quad (260)$$

- (b) I piani in questione sono tutti ortogonali a \vec{k} ed equispaziati alla distanza pari alla lunghezza d'onda $d = 2\pi/k$
- (c) Siccome $\vec{k} \in \mathcal{R}^*$ per esso vale (in base alla definizione di p.50):

$$\vec{k} \cdot \vec{R} = 2\pi n \quad n \in \mathbb{Z} \quad \forall \vec{R} \in \mathcal{R} \quad (261)$$

ovvero:

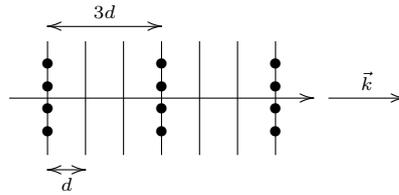
$$\exp(i\vec{k} \cdot \vec{R}) = 1 \quad \forall \vec{R} \in \mathcal{R} \quad (262)$$

Questo significa che *tutti* i punti \vec{R} del reticolo diretto sono contenuti nei fronti dell'onda piana definiti dalla 260 e quindi esiste una famiglia di piani reticolari ortogonali a \vec{k} .

- (d) Questo non ci assicura ancora che la spaziatura dei piani reticolari sia uguale a d .

Sicuramente la spaziatura dei piani reticolari non puo' essere $< d$, perche' altrimenti ci sarebbero dei punti reticolari *non* contenuti nei fronti d'onda, contrariamente a quanto abbiamo appena appurato sopra.

Tuttavia, la spaziatura dei piani reticolari potrebbe essere un multiplo di d . Ad esempio, se i punti reticolari fossero presenti solo su un fronte d'onda ogni tre, la spaziatura dei piani reticolari sarebbe uguale a $3d$:



- (e) Che la spaziatura dei piani reticolari sia uguale a quella dei fronti d'onda (cioe' $= d$) segue dal fatto che il vettore \vec{k} e' quello di modulo minimo fra tutti i vettori reciproci ad esso paralleli.

Infatti, se la spaziatura fosse $d' > d$, in base alla prima parte del teorema, esisterebbe un vettore $\vec{k}' \in \mathcal{R}^*$ ortogonale alla famiglia di piani reticolari (e quindi parallelo a \vec{k}) e avente modulo $k' = 2\pi/d' < 2\pi/d = k$, contrariamente all'ipotesi che \vec{k} abbia il modulo minimo fra tutti i vettori di \mathcal{R}^* ad esso paralleli.

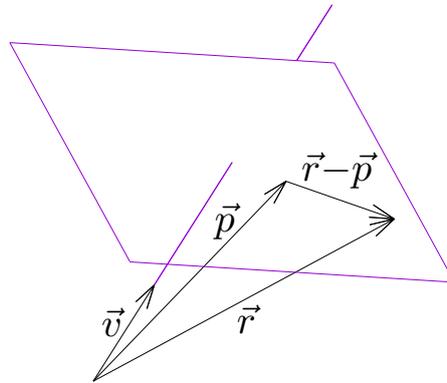


Figura 31: Condizione di appartenenza al piano passante per il punto \vec{p} e avente una normale \vec{v} .

6. Indici di Miller.

Quanto visto al punto precedente offre un modo per specificare l'orientazione dei piani reticolari.

In generale, dal punto di vista geometrico, l'orientazione di un piano nello spazio è completamente specificata da un vettore ad esso ortogonale.

L'equazione generale di un piano:

$$ax + by + cz = d \quad (263)$$

può essere infatti ottenuta su questa base. Guardate la figura 31.

Detto $\vec{v} = (v_1 v_2 v_3)$ il versore ortogonale al piano e $\vec{p} = (p_1 p_2 p_3)$ un suo punto qualsiasi, per tutti i punti $\vec{r} = (xyz)$ del piano deve valere:

$$(\vec{r} - \vec{p}) \cdot \vec{v} = 0 \quad (264)$$

ovvero:

$$\vec{r} \cdot \vec{v} = \vec{p} \cdot \vec{v} \quad (265)$$

$$v_1 x + v_2 y + v_3 z = v_1 p_1 + v_2 p_2 + v_3 p_3 \quad (266)$$

da cui si vede che:

$$a = v_1 \quad (267)$$

$$b = v_2 \quad (268)$$

$$c = v_3 \quad (269)$$

$$d = v_1 p_1 + v_2 p_2 + v_3 p_3 \quad (270)$$

Sulla base di cio', si puo' dire che l'equazione di un piano avente normale \vec{v} e':

$$\vec{r} \cdot \vec{v} = C \quad (271)$$

dove la costante C (che dipende dalle coordinate cartesiane del versore \vec{v} e di un punto qualsiasi del piano) fissa la posizione del piano lungo la direzione di \vec{v} (sara' $C = 0$ per un piano passante per l'origine).

Grazie a quanto visto al punto 5 precedente (p. 71), e' possibile individuare una data famiglia di piani reticolari con un vettore del reticolo reciproco ad essa ortogonale. Per rendere univoca la corrispondenza (visto che ci sono infiniti vettori reciproci fra loro paralleli e ortogonali alla famiglia di piani) si prende il vettore di modulo minimo. Tale vettore avra' la forma:

$$h\vec{b}_1 + k\vec{b}_2 + l\vec{b}_3 \quad (272)$$

dove i \vec{b}_i sono i vettori primitivi del reticolo reciproco e $h, k, l \in \mathbb{Z}$.

Quindi una data famiglia di piani reticolari sara' individuata univocamente dai tre numeri interi h, k, l , che vengono detti *indici di Miller* della famiglia di piani. Dal punto di vista della notazione, la famiglia di piani viene indicata ponendo i suoi indici di Miller fra parentesi rotonde: (hkl) . Se qualche indice e' negativo, il segno meno viene posto sopra di esso, come ad esempio in $(1\bar{2}2)$.

Notate che siccome il vettore reciproco che individua la famiglia di piani e' quello di modulo minimo, i tre numeri interi hkl non hanno un fattore comune.

Notate ancora che hkl dipendono dai particolari vettori primitivi scelti per definire il reticolo reciproco.

7. Interpretazione geometrica degli indici di Miller.

E' possibile dare una semplice interpretazione geometrica degli indici di Miller che consente di visualizzare subito l'orientazione dei piani da essi individuati.

Un qualsiasi piano della famiglia (hkl) e' ortogonale al vettore reciproco $\vec{K} = h\vec{b}_1 + k\vec{b}_2 + l\vec{b}_3$. Quindi tale piano reticolare sara' contenuto nel piano *continuo* che ha equazione:

$$\vec{K} \cdot \vec{r} = C \quad (273)$$

per un'opportuna costante C .

Indichiamo le intersezioni di questo piano con gli assi individuati dai vettori primitivi diretti $\vec{a}_1, \vec{a}_2, \vec{a}_3$ con $x_1\vec{a}_1, x_2\vec{a}_2, x_3\vec{a}_3$ (guardate la figura 32)

Chiaramente i tre vettori $x_i\vec{a}_i$ soddisfano la condizione per appartenere al piano.

Quindi, considerando ad esempio $x_1\vec{a}_1$, si ha:

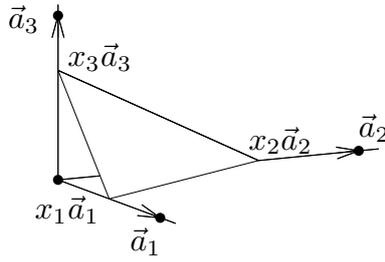


Figura 32: Un piano reticolare appartenente alla famiglia (hkl) interseca gli assi individuati dai vettori primitivi diretti $\vec{a}_1, \vec{a}_2, \vec{a}_3$ nei punti $x_1\vec{a}_1, x_2\vec{a}_2, x_3\vec{a}_3$, rispettivamente.

$$\begin{aligned} \vec{K} \cdot x_1\vec{a}_1 &= C \\ (h\vec{b}_1 + k\vec{b}_2 + l\vec{b}_3) \cdot x_1\vec{a}_1 &= C \\ 2\pi hx_1 &= C \\ x_1 &= \frac{C}{2\pi} \frac{1}{h} \end{aligned}$$

(nel penultimo passaggio ho sfruttato le relazioni di ortogonalita' 118)

In definitiva si ha:

$$x_1 = \frac{C}{2\pi} \frac{1}{h} \quad (274)$$

$$x_2 = \frac{C}{2\pi} \frac{1}{k} \quad (275)$$

$$x_3 = \frac{C}{2\pi} \frac{1}{l} \quad (276)$$

cioe': gli indici di Miller di una famiglia di piani sono inversamente proporzionali alle intersezioni di un qualsiasi piano della famiglia con gli assi individuati dai vettori primitivi del reticolo diretto.

Così, ad esempio, un piano della famiglia (111) intersecherà gli assi $\vec{a}_1\vec{a}_2\vec{a}_3$ a uguali distanze dall'origine; un piano della famiglia (121) intersecherà gli assi $\vec{a}_1\vec{a}_2\vec{a}_3$ a distanze dall'origine pari a $d_1d_2d_3$ tali che $d_1 : d_2 : d_3 = 1 : 2 : 1$, e così via.

Alcuni piani con i rispettivi indici di Miller e le intersezioni sugli assi del reticolo diretto sono mostrati nella figura 33.

Notate che se un indice di Miller è nullo, la corrispondente intersezione è infinita, cioè il piano è parallelo al corrispondente asse.

FIGURE

Figura 33:

8. Anche se non è direttamente legata al reticolo reciproco, ma siccome abbiamo introdotto la notazione per i piani reticolari, da questo punto la notazione per specificare le direzioni nel reticolo diretto.

La direzione individuata dal vettore diretto:

$$n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3 \quad (277)$$

viene indicata coi tre numeri interi $n_1n_2n_3$ posti fra parentesi quadrate, per non fare confusione con la notazione usata per i piani. Ad esempio: $[120]$ oppure $[12\bar{3}]$.

93. Cenni alla teoria della diffrazione.

Per *diffrazione* si intende generalmente il fenomeno per cui quando un'onda incontra un ostacolo (o una fenditura), interagisce con esso in modo tale che l'ostacolo (o la fenditura) si comporta come una sorgente (secondaria) dell'onda stessa.

In altre parole, tutto funziona come se l'ostacolo "colpito" dall'onda la riemetta a sua volta in tutte le direzioni. Il processo attraverso cui ciò avviene è in generale chiamato *scattering*.¹⁹

Nel seguito supporremo che il processo di scattering sia *elastico*: questo significa che nell'interazione con lo scatterer (cioè l'ostacolo puntiforme) l'onda non perde né guadagna energia. La conseguenza di ciò è che l'onda generata dallo scatterer ha la stessa lunghezza d'onda di quella incidente (notate che non sempre questo è il caso).

Supporremo inoltre che il processo di scattering non cambi la fase dell'onda o la cambi di una quantità definita e costante (anche questa caratteristica del processo di scattering non è garantita in generale).

Se un'onda piana incide su due ostacoli, che supporremo puntiformi per semplicità, ciascuno dei due diventa una sorgente. Le onde riemesse dai due scatterers danno *interferenza*.

L'interferenza è una diretta conseguenza del *principio di sovrapposizione*, secondo cui se due (o più) onde attraversano uno stesso punto dello spazio, la variazione nel tempo della grandezza che definisce tali onde nel punto considerato è la somma delle variazioni dovute alle singole onde. In altre parole, l'ampiezza dell'onda risultante nel punto considerato è la somma delle ampiezze delle onde componenti.

L'ampiezza risultante nel dato punto dipende dalla fase relativa delle due onde componenti. I due casi estremi sono i seguenti.

Se le onde componenti (che supponiamo di uguale ampiezza massima A) raggiungono lo stesso punto *in fase* (cioè nel punto considerato la fase delle onde incidenti ha lo stesso valore, a meno di un multiplo intero di 2π), allora si ha

¹⁹Mi rifiuto di tradurre; la traduzione letterale sarebbe *dispersione*, *spargimento*, *sparpagliamento*; cioè: l'onda proveniente da una direzione precisa, viene riemessa dall'ostacolo in tutte le direzioni, viene cioè dispersa, sparpagliata nello spazio.

una cosiddetta *interferenza costruttiva*, perché l'ampiezza risultante è ad ogni istante maggiore di quella delle singole onde componenti ed ha il massimo valore possibile (siccome la fase è la stessa, le due ampiezze componenti sono ad ogni istante uguali e hanno, ovviamente, sempre lo stesso segno).

Se le onde componenti raggiungono lo stesso punto *in opposizione di fase* (cioè nel punto considerato quando una fase "è al massimo", l'altra "è al minimo"), allora si ha una *interferenza distruttiva*, e l'ampiezza risultante è nulla ad ogni istante (quando il coseno di una delle due onde vale $+a$, il coseno dell'altra vale $-a$ e quindi la somma vale sempre 0).

Per sfasamenti intermedi si avrà ovviamente una corrispondente situazione intermedia in cui l'ampiezza risultante è compresa tra 0 e il valore massimo possibile.

Come vedremo meglio più avanti, lo sfasamento fra le due onde difratte dipende dalla differenza del loro cammino ottico, cioè, in parole povere, dalla differenza della lunghezza del percorso che ciascuna di esse deve compiere per raggiungere il punto considerato. Dovrebbe essere evidente che per avere interferenza costruttiva la differenza di cammino ottico deve essere uguale a un *numero intero di lunghezze d'onda*, mentre per avere interferenza distruttiva, la differenza di cammino ottico deve essere uguale a un *numero dispari di mezze lunghezze d'onda*.

La differenza di cammino ottico dipende, a sua volta, dalla distanza fra i due scatterers e dall'angolo formato dalla direzione delle due onde difratte con la direzione ortogonale alla retta contenente i due scatterers (vedremo questo aspetto in dettaglio fra poco; l'idea è illustrata nella figura 34). In generale, dunque, se si pone a una certa distanza dagli scatterers uno "schermo" (o qualcosa di equivalente) capace di rivelare l'ampiezza delle onde incidenti in ogni suo punto, si vedranno dei punti in cui tale ampiezza è massima e altri in cui essa è nulla. La distribuzione dell'ampiezza sullo schermo viene detta "figura o spettro di interferenza" o "figura o spettro di diffrazione" (siccome l'interferenza è una conseguenza della diffrazione, molto spesso i termini diffrazione e interferenza vengono usati come sinonimi).

Affinché due onde rimesse da due scatterers possano dare interferenza nel modo che ho appena descritto è necessario che i due scatterers siano a una distanza dell'ordine di grandezza della lunghezza d'onda dell'onda incidente.

94. La radiazione elettromagnetica così come le "onde materiali" associate ad oggetti quantomeccanici come gli elettroni o i neutroni possono venire difratte dagli atomi. Un solido cristallino è costituito da atomi disposti in modo regolare e periodico e quindi costituisce un *reticolo di diffrazione* (esattamente come una serie di fenditure allineate ed equispaziate, che forse avrete già incontrato in qualche corso).

Le distanze interatomiche nei solidi cristallini sono dell'ordine degli Å ($1 \text{ Å} = 1 \times 10^{-8} \text{ cm}$). I raggi X (radiazione elettromagnetica) hanno lunghezze d'onda di quell'ordine di grandezza e quindi se si fa incidere un fascio di raggi X su un cristallo, gli atomi in esso contenuti diffrangono la radiazione incidente e i raggi X diffratti danno interferenza.

Come vedremo, anche le onde associate agli elettroni hanno lunghezze d'onda dell'ordine degli Å e questo ha delle importanti conseguenze sulla struttura elettronica dei solidi cristallini.

95. In questa parte presenterò alcuni aspetti elementari della diffrazione di onde piane da parte degli atomi di un solido cristallino. In particolare, otterremo

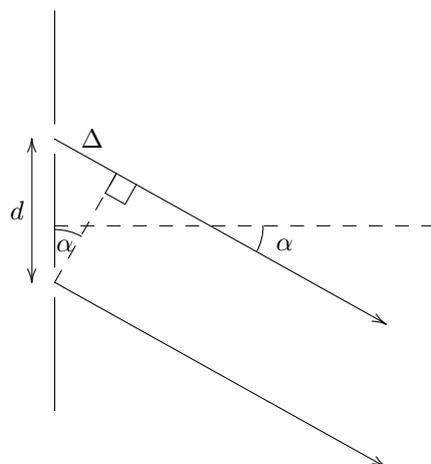


Figura 34: Due onde diffratte generate da due fenditure poste alla distanza d e che si propagano lungo una direzione inclinata di un angolo α rispetto all'asse orizzontale. La differenza di cammino ottico e' $\Delta = d \sin \alpha$.

la condizione affinché le onde diffratte dagli atomi contenuti nel solido possano interferire costruttivamente fra loro.

Questo argomento e' il nucleo centrale della cristallografia. La disposizione degli atomi in un cristallo puo' essere determinata inviando sul cristallo un fascio di raggi X e misurando l'intensita' dei raggi X diffratti (l'intensita' in ogni punto dello spazio e' una funzione dell'ampiezza dovuta alle onde diffratte) in funzione della posizione su uno schermo posto dietro il campione (figura 35). Dalla posizione e dall'intensita' dei massimi dovuti all'interferenza costruttiva dei raggi X diffratti si puo' risalire alla disposizione degli atomi nel cristallo.

La cristallografia e' un argomento molto vasto e articolato e in questa parte verremo in contatto con essa molto superficialmente: l'intento e' quello di stabilire alcuni risultati che, come vedrete, sono essenziali per la comprensione della struttura elettronica dei solidi.

96. In cio' che viene di seguito supporremo che un'onda piana (avente lunghezza d'onda tale da essere diffratta e dare interferenza) incida su un cristallo la cui struttura, per semplicita', e' costituita da atomi tutti uguali e localizzati sui punti di un reticolo di Bravais.²⁰

Come abbiamo detto, la radiazione elettromagnetica nel range di lunghezze d'onda tipiche dei raggi X e' in grado di interagire con un cristallo dando delle figure di interferenza definite e per fissare le idee supporremo che l'onda incidente sul cristallo sia un fascio di raggi X (anche se, come vedremo, il tipo di onda a cui saremo interessati e' in realta' quella associata a un elettrone).

97. Il modello della diffrazione di Laue. Come ho accennato sopra, se un fascio di raggi X viene focalizzato su un cristallo, gli atomi contenuti in esso costituiscono un reticolo di diffrazione per la radiazione incidente. I raggi X diffratti danno interferenza e se si pone uno schermo rivelatore dietro il cristal-

²⁰Cioe', non considereremo il caso di una struttura costituita da un reticolo di Bravais a ciascun punto del quale sia associata una base poliatomica (punto 81).

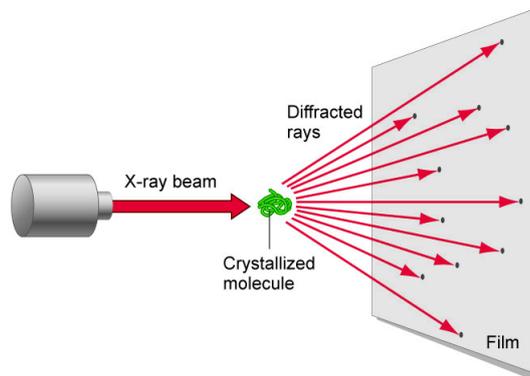


Figura 35:

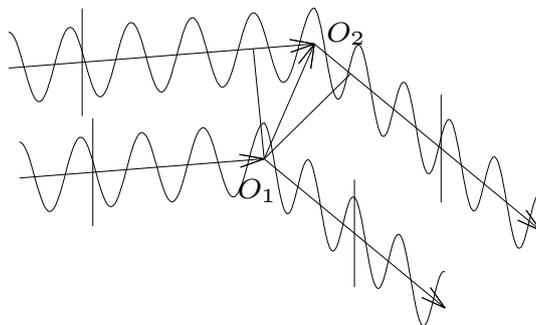


Figura 36:

lo (ad esempio una lastra fotografica), nelle direzioni per cui tale interferenza e' costruttiva si osserveranno dei massimi di intensita' sotto forma di macchie localizzate e definite sullo schermo (figura 35).

Cio' che faremo ora e' ricavare la condizione affinche' in una certa direzione della radiazione emergente dal cristallo si abbia un massimo di intensita' diffratta. L'approccio che useremo fu elaborato dal fisico tedesco Max von Laue nei primi anni del '900.

Iniziamo col considerare due soli atomi posizionati a O_1 e O_2 (gli scatterers) la cui posizione reciproca e' definita da un vettore \vec{d} (fate riferimento alle figure 36 e 37).

Un'onda piana (raggi X, per fissare le idee) incide sui due atomi. La direzione dell'onda incidente e' definita da un versore \vec{n} ; la lunghezza d'onda e' λ e quindi il vettore d'onda e':

$$\vec{k} = \frac{2\pi}{\lambda} \vec{n} \quad (278)$$

In seguito allo scattering, ciascuno dei due atomi diventa una sorgente puntiforme e riemette l'onda in tutte le direzioni. A rigori, l'onda riemessa non e' un'onda piana ma e' un'onda sferica (i cui fronti d'onda non sono dei piani,

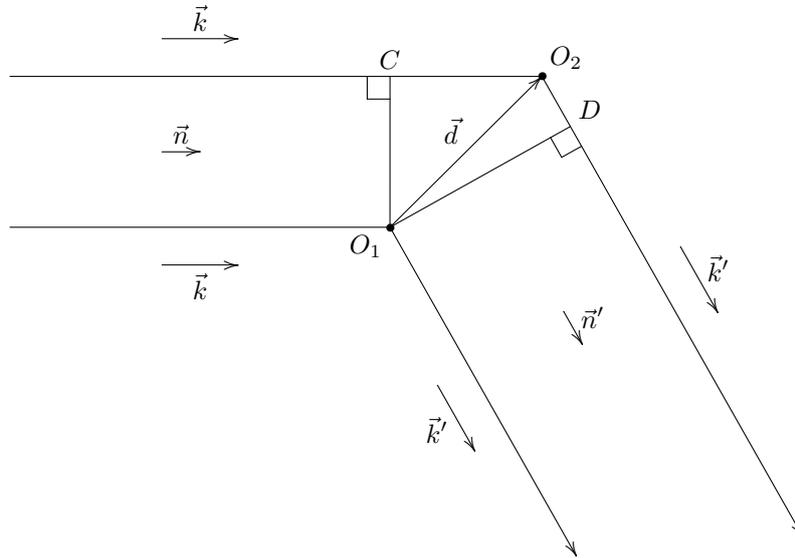


Figura 37:

ma delle sfere); tuttavia, se assumiamo, come e' sempre il caso, di osservare gli effetti dell'interferenza a una distanza molto maggiore della lunghezza d'onda della radiazione emessa, possiamo legittimamente assumere che la curvatura dei fronti d'onda sia molto piccola e che, quindi, l'onda riemessa dagli scatterers sia a tutti gli effetti un'onda piana.

Siccome facciamo l'ipotesi che lo scattering sia elastico, la lunghezza d'onda della radiazione diffratta da ciascun atomo e' uguale a quella della radiazione incidente. Inoltre, come detto, assumiamo anche che nel processo di scattering, le onde riemesse da ciascun atomo abbiano subito, nel punto dove e' avvenuto lo scattering, la stessa variazione di fase (che, per semplicita', possiamo assumere uguale a zero).

In un punto situato dietro il cristallo lungo una certa direzione si otterra' un massimo di intensita' diffratta se le onde diffratte dai due atomi danno interferenza costruttiva in quel punto. Cio' puo' verificarsi solo se le onde giungono a quel punto *in fase*.

Indichiamo con \vec{k}' il vettore d'onda delle onde riemesse dai due atomi in una certa direzione definita da un versore \vec{n}' ; sotto le ipotesi che abbiamo fatto, \vec{k} e \vec{k}' hanno lo stesso modulo ($2\pi/\lambda$, perche' la lunghezza d'onda non e' cambiata), ma *direzione* diversa.²¹

Ci chiediamo ora quale condizione deve essere verificata affinche' le onde diffratte raggiungano un certo punto in fase.

La risposta a questa domanda e' evidentemente la seguente: le due onde raggiungeranno il punto in fase se la differenza di cammino ottico fra esse e' uguale a un multiplo intero della (comune) lunghezza d'onda.

²¹ Assumiamo che il punto sul quale sono dirette le onde diffratte sia sufficientemente distante in modo tale che le loro direzioni di propagazione possano essere considerate parallele

Dalla figura 37 si vede che la differenza di cammino fra l'onda riemessa da O_1 e quella riemessa da O_2 e':

$$\overline{CO_2} + \overline{O_2D} \quad (279)$$

Tale differenza di cammino deve essere uguale a un multiplo intero di λ :

$$\overline{CO_2} + \overline{O_2D} = m\lambda \quad m \in \mathbb{N} + \{0\} \quad (280)$$

Possiamo esprimere il primo membro della 280 in funzione del vettore \vec{d} e dei due versori \vec{n} e \vec{n}' .

La lunghezza del segmento $\overline{CO_2}$ e' uguale a quella della proiezione di \vec{d} lungo la direzione di \vec{n} ; siccome \vec{n} ha modulo unitario, quest'ultima e' data semplicemente dal prodotto scalare $\vec{d} \cdot \vec{n}$:

$$\overline{CO_2} = \vec{d} \cdot \vec{n} \quad (281)$$

In modo analogo, la lunghezza del segmento $\overline{O_2D}$ e' uguale a quella della proiezione di \vec{d} lungo la direzione di \vec{n}' ; anche ora, quest'ultima e' il prodotto scalare $\vec{d} \cdot \vec{n}'$, ma in questo caso il prodotto scalare e' negativo (perche' l'angolo fra \vec{d} e \vec{n}' e' $> \pi/2$) e quindi bisogna prenderlo con il segno meno per avere la lunghezza (positiva) del segmento $\overline{O_2D}$:

$$\overline{O_2D} = -\vec{d} \cdot \vec{n}' \quad (282)$$

Sostituendo sopra si ha:²²

$$\vec{d} \cdot \vec{n} - \vec{d} \cdot \vec{n}' = m\lambda \quad (283)$$

$$\vec{d} \cdot (\vec{n} - \vec{n}') = m\lambda \quad (284)$$

(moltiplico ambo i membri per $2\pi/\lambda$)

$$\vec{d} \cdot \left(\frac{2\pi}{\lambda} \vec{n} - \frac{2\pi}{\lambda} \vec{n}' \right) = 2\pi m \quad (285)$$

$$\vec{d} \cdot (\vec{k} - \vec{k}') = 2\pi m \quad (286)$$

Notate che \vec{k} e \vec{k}' hanno lo stesso modulo, ma direzione diversa e quindi la loro differenza (vettoriale) *non* e' nulla.

Ora passiamo a considerare il reticolo di Bravais con un atomo in ogni suo punto; se vogliamo che tutte le onde diffratte da tutti gli atomi del reticolo raggiungano un certo punto in fase (producendo cosi' un massimo di intensita' in quel punto), l'eq. 286 deve essere soddisfatta contemporaneamente per *tutte le possibili coppie di punti del reticolo*. In questo caso, il vettore che definisce la posizione relativa di una coppia qualsiasi di punti reticolari e' ovviamente un vettore reticolare $\vec{R} = n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3$, dove n_1, n_2, n_3 sono numeri interi a $\vec{a}_1, \vec{a}_2, \vec{a}_3$ sono i vettori primitivi del reticolo di Bravais.

La condizione diventa pertanto:

$$\vec{R} \cdot (\vec{k} - \vec{k}') = 2\pi m \quad (287)$$

²²Potete verificare da soli che allo stesso risultato si arriva non solo per la geometria mostrata nella figura 37, ma per qualsiasi altra orientazione relativa dell'onda incidente e di quella emergente.

Ma in base all'eq. 122 cio' significa che il vettore $(\vec{k} - \vec{k}')$ deve essere un vettore del reticolo reciproco.

Arriviamo cosi' al seguente risultato: *le onde diffratte daranno interferenza costruttiva nella direzione tale per cui la differenza fra il vettore d'onda incidente e quello dell'onda emergente e' uguale a un vettore del reticolo reciproco.*

La 287 e' una delle forme in cui puo' essere scritta la cosiddetta *condizione di Laue*.

Possiamo riscrivere la 287 in un modo che ne consente un'interpretazione geometrica semplice.

Poniamo:

$$\vec{K} = \vec{k} - \vec{k}' \quad (288)$$

dove \vec{K} e' il vettore del reticolo reciproco che soddisfa la condizione di Laue 287. Siccome l'onda incidente e quella diffratta hanno lo stesso modulo del vettore d'onda, deve essere:

$$|\vec{k}| = |\vec{k}'| \quad (289)$$

$$k = |\vec{k} - \vec{K}| \quad (290)$$

(elevo al quadrato ambo i membri)

$$k^2 = (\vec{k} - \vec{K})^2 \quad (291)$$

$$\left(\text{qua uso: } \vec{v}^2 = \vec{v} \cdot \vec{v} = |\vec{v}|^2 = v^2\right)$$

$$k^2 = k^2 + K^2 - 2\vec{k} \cdot \vec{K} \quad (292)$$

$$= k^2 + K^2 - 2K\vec{k} \cdot \vec{n}_K \quad (293)$$

dove \vec{n}_K e' un versore nella direzione di \vec{K} .

Arriviamo cosi' al risultato finale:

$$\vec{k} \cdot \vec{n}_K = \frac{1}{2}K \quad (294)$$

Innanzitutto vedete che abbiamo riscritto la condizione di Laue 287 in termini del solo vettore d'onda della radiazione incidente e di un corrispondente vettore del reticolo reciproco.

Il significato geometrico della 294 e' illustrato nella figura 38.

Il termine $\vec{k} \cdot \vec{n}_K$ non e' altro che il modulo della proiezione di \vec{k} lungo la direzione di \vec{K} . Allora la 294 dice che si avra' un massimo nell'intensita' dell'onda diffratta ogni volta che il modulo della proiezione del vettore d'onda dell'onda incidente nella direzione di un vettore del reticolo reciproco e' uguale alla meta' del modulo di quel vettore reciproco.

Come mostra la figura 38 questa condizione e' soddisfatta da tutti i vettori d'onda \vec{k} che si appoggiano su un piano ortogonale al vettore reciproco \vec{K} e lo interseca nel suo punto medio.

Notate anche che $\vec{k}' = \vec{k} - \vec{K}$ e' l'immagine simmetrica di \vec{k} rispetto al piano ortogonale a \vec{K} .

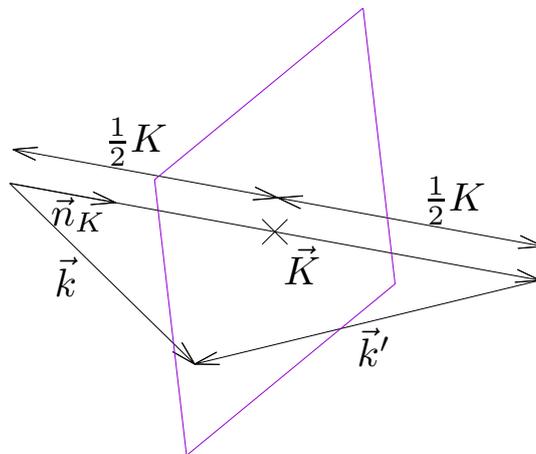


Figura 38:

Ora: se riguardate il punto 4 a p. 68 realizzate immediatamente che i piani per cui la 294 e' soddisfatta *sono i confini delle zone di Brillouin.*

Abbiamo cosi' messo in evidenza una proprieta' molto importante delle zone di Brillouin: tutte le onde piane incidenti su un cristallo e aventi un vettore d'onda che si appoggia su qualsiasi punto del confine di una zona di Brillouin verranno diffratte dando interferenza costruttiva.

Il modello del gas di elettroni

98. In questa parte prenderemo in esame il modello piu' semplice per trattare il comportamento degli elettroni nei metalli; tale modello fu introdotto e sviluppato da Arnold Sommerfeld a partire dal 1927.²³

99. Il modello viene anche chiamato "*electron gas model*", ovvero "modello del gas di elettroni" e il motivo e' il seguente. Un solido metallico viene considerato come costituito da ioni metallici (nuclei piu' elettroni di core) in posizioni fisse, mentre gli elettroni di valenza sono completamente liberi di muoversi in tutto il volume occupato dal metallo. Tali elettroni di valenza sono quindi assimilati a un gas, nel senso che il loro moto e' casuale ed essi tendono ad occupare tutto il volume disponibile del metallo.

Come vedremo, un aspetto essenziale per il trattamento del modello del gas di elettroni a temperature maggiori di 0 K e' la legge di distribuzione di Fermi-Dirac: per questo motivo, spesso questo modello viene anche chiamato *gas di Fermi*.

Un'aspetto essenziale del modello e' che gli elettroni che costituiscono il "gas" sono *indipendenti*, cioe' non interagiscono fra loro, e *liberi*, cioe' non interagiscono neanche con gli ioni. Questa assunzione puo' sembrare non realistica (e in effetti lo e'); cio' non di meno, le previsioni che si possono dedurre dal modello del gas di elettroni sono in molti casi in buonissimo accordo con i dati sperimentali; e questo si spiega col fatto che molte proprieta' dei metalli dipendono principalmente dalla mobilita' dei loro elettroni di valenza e molto poco da come questa mobilita' viene influenzata dalle loro interazioni reciproche e/o da quelle con i nuclei.

100. A tutti gli effetti, il modello del gas di elettroni indipendenti e liberi equivale a considerare gli elettroni di valenza di un metallo confinati nel suo volume V e come se ciascuno di essi si trovasse da solo, in assenza di interazioni con altri elettroni e/o con nuclei.

In queste condizioni, gli stati accessibili agli elettroni che costituiscono il gas si trovano risolvendo l'equazione di Schrödinger *per un singolo elettrone*. Lo stato fondamentale del gas di elettroni (cioe' l'occupazione degli stati a 0 K) si ottiene assegnando gli elettroni agli stati in ordine di energia crescente e nel rispetto del principio di Pauli.

Tutti i possibili stati monoelettronici si ottengono risolvendo l'equazione di Schrödinger per un singolo elettrone, che nella sua forma piu' generale applicata al caso presente e':²⁴

$$-\frac{\hbar^2}{2m}\nabla^2\xi(\vec{r},t) = i\hbar\frac{\partial}{\partial t}\xi(\vec{r},t) \quad (295)$$

²³In realta', prima della scoperta della meccanica quantistica, Paul Drude sviluppo' un modello per il comportamento degli elettroni nei metalli ancora piu' semplificato, detto appunto "Drude model". In tale modello gli elettroni sono considerati come particelle classiche, soggette alla statistica di Boltzmann. Nonostante i presupposti (e anche grazie a degli errori), il modello di Drude riusciva a spiegare alcune osservazioni sperimentali molto importanti e quindi godette di una certa fortuna fino a che venne sostituito dal modello di Sommerfeld. Il modello di Drude e' presentato nel capitolo 1 del libro di Ashcroft e Mermin.⁷

²⁴In tutto il corso assumeremo sempre che la funzione di stato di un elettrone possa essere fattorizzata in una parte dipendente dalle coordinate spaziali, una parte dipendente dal tempo e una parte dipendente dallo spin; se, come sara' sempre il caso per noi, l'hamiltoniano e' indipendente dallo spin, la parte della funzione di stato che dipende dallo spin non compare nell'equazione di Schrödinger e le soluzioni di quest'ultima saranno doppiamente degeneri rispetto allo spin.

La 295 e' l'equazione di Schrödinger monoelettronica dipendente dal tempo. E' un'equazione differenziale a derivate parziali. La soluzione $\xi(\vec{r}, t)$ e' una funzione di 4 variabili: le 3 coordinate spaziali ($\vec{r} = (x, y, z)$) e il tempo t . m e' la massa dell'elettrone, $\hbar = h/2\pi$ e h e' la costante di Plank, $i = \sqrt{-1}$ e' l'unita' immaginaria.

L'operatore che opera sulla funzione $\xi(\vec{r}, t)$ al primo membro e' l'operatore hamiltoniano (normalmente indicato con H): per il sistema che stiamo considerando (gas di elettroni indipendenti e liberi) l'operatore hamiltoniano contiene solo il termine dell'energia cinetica:

$$H = -\frac{\hbar^2}{2m}\nabla^2 = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) \quad (296)$$

101. Siccome nel nostro caso l'operatore hamiltoniano al primo membro dell'equazione 295 non dipende dal tempo, le sue soluzioni possono essere ottenute assumendo che la funzione $\xi(\vec{r}, t)$ si possa scrivere come prodotto di una funzione $\Psi(\vec{r})$ dipendente solo dalla posizione e di una funzione $\Theta(t)$ dipendente solo dal tempo:

$$\xi(\vec{r}, t) = \Psi(\vec{r})\Theta(t) \quad (297)$$

Sostituendo nella 295 si ottiene:

$$\Theta(t)\left(-\frac{\hbar^2}{2m}\nabla^2\Psi(\vec{r})\right) = \Psi(\vec{r})\left(i\hbar\frac{\partial}{\partial t}\Theta(t)\right) \quad (298)$$

perche' $\Theta(t)$ si comporta come una costante rispetto all'hamiltoniano al primo membro, che dipende solo da \vec{r} , mentre $\Psi(\vec{r})$ si comporta come una costante rispetto alla derivata del secondo membro, che e' fatta rispetto al tempo t .

Dividendo ambo i membri per $\xi(\vec{r}, t) = \Psi(\vec{r})\Theta(t)$ si arriva a:

$$\frac{1}{\Psi(\vec{r})}\left(-\frac{\hbar^2}{2m}\nabla^2\Psi(\vec{r})\right) = \frac{1}{\Theta(t)}\left(i\hbar\frac{\partial}{\partial t}\Theta(t)\right) \quad (299)$$

Il primo membro e' funzione *solo* della posizione \vec{r} dell'elettrone, mentre il secondo membro e' funzione *solo* del tempo t ; l'equazione ottenuta richiede l'uguaglianza del primo e secondo membro *per qualsiasi valore di \vec{r} e t* e questo puo' avvenire solo se essi sono separatamente uguali alla stessa *costante*, che chiamiamo ϵ .

102. L'equazione di partenza e' dunque equivalente alle due equazioni separate per $\Psi(\vec{r})$ ed $\Theta(t)$:

$$-\frac{\hbar^2}{2m}\nabla^2\Psi(\vec{r}) = \epsilon\Psi(\vec{r}) \quad (300)$$

$$i\hbar\frac{d}{dt}\Theta(t) = \epsilon\Theta(t) \quad (301)$$

(nella seconda equazione la derivata parziale e' diventata una derivata totale perche' ora agisce su una funzione della sola variabile t)

La costante ϵ , essendo l'autovalore dell'operatore hamiltoniano nella prima equazione, e' un possibile valore di energia dell'elettrone.

103. L'equazione temporale 301 si risolve facilmente mediante separazione delle variabili.

Si ha:

$$\frac{d\Theta}{\Theta} = \frac{\epsilon}{i\hbar} dt \quad (302)$$

$$\int \frac{d\Theta}{\Theta} = \frac{\epsilon}{i\hbar} \int dt \quad (303)$$

$$\ln \Theta = \frac{\epsilon}{i\hbar} t + C' \quad (304)$$

$$\Theta(t) = \exp\left(\frac{\epsilon}{i\hbar} t + C'\right) \quad (305)$$

$$\Theta(t) = C \exp\left(\frac{\epsilon}{i\hbar} t\right) = C \exp\left(-i\frac{\epsilon}{\hbar} t\right) \quad (306)$$

con $C = \exp(C')$ una costante.

Facendo uso della relazione di Plank-Einstein che lega l'energia ϵ alla frequenza (lineare) ν di una particella:

$$\epsilon = h\nu \quad (307)$$

si ottiene:

$$\Theta(t) = C \exp(-i\omega t) \quad (308)$$

dove $\omega = 2\pi\nu$ e' la frequenza angolare 4.

Siccome la funzione $\xi(\vec{r}, t)$ e' il prodotto delle due funzioni $\Psi(\vec{r})$ ed $\Theta(t)$, si puo' assumere $C = 1$ "scaricando" tutti gli aggiustamenti necessari sulla costante che comparira' nella soluzione spaziale $\Psi(\vec{r})$.

In definitiva, per l'equazione temporale di Schrödinger possiamo assumere la soluzione:

$$\Theta(t) = \exp(-i\omega t) \quad (309)$$

104. Il risultato ottenuto e' di portata generale: per tutti i sistemi per cui e' possibile la fattorizzazione 297, la dipendenza dal tempo della soluzione dell'equazione di Schrödinger si ottiene semplicemente moltiplicando la soluzione spaziale (la $\Psi(\vec{r})$ nella 297) per l'esponenziale 309:

$$\xi(\vec{r}, t) = \Psi(\vec{r}) \exp(-i\omega t) \quad (310)$$

105. Ci occupiamo ora della soluzione dell'equazione di Schrödinger indipendente dal tempo 300:

$$-\frac{\hbar^2}{2m} \nabla^2 \Psi(\vec{r}) = \epsilon \Psi(\vec{r})$$

Si tratta di un'equazione differenziale d'autovalore a derivate parziali (perche' $\Psi(\vec{r}) = \Psi(x, y, z)$ e' una funzione di 3 variabili). Risolvere la 300 significa trovare tutti i possibili valori di ϵ e le corrispondenti funzioni $\Psi(\vec{r})$ (ad ogni valore di ϵ corrisponde una o piu' funzioni $\Psi(\vec{r})$) che soddisfano l'equazione.

106. Siccome stiamo considerando il caso di un elettrone libero che possiede solo energia cinetica, l'energia che compare nella 300 sara' sempre ≥ 0 (prendendo

come valore nullo dell'energia quello di un elettrone fermo). Conviene allora raggruppare tutti i termini costanti dell'equazione nel modo seguente:

$$k^2 = \frac{2m\epsilon}{\hbar^2} \quad (311)$$

La 300 puo' cosi' essere riscritta come:

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \Psi(x, y, z) + k^2 \Psi(x, y, z) = 0 \quad (312)$$

107. In modo simile a quanto fatto al punto 101 per la fattorizzazione della $\xi(\vec{r}, t)$, anche ora possiamo assumere che la soluzione $\Psi(\vec{r})$ cercata si possa scrivere come il prodotto di tre funzioni, ciascuna dipendente da una sola delle tre coordinate spaziali:

$$\Psi(\vec{r}) = \Psi(x, y, z) = X(x) Y(y) Z(z) \quad (313)$$

Sostituendo nella 312 e osservando di nuovo che ciascuna derivata parziale "lascia passare" come costanti le due funzioni che non dipendono dalla coordinata sulla quale essa opera, con passaggi analoghi a quelli visti al punto 101 si arriva facilmente a:

$$\frac{1}{X} \frac{\partial^2 X}{\partial x^2} + \frac{1}{Y} \frac{\partial^2 Y}{\partial y^2} + \frac{1}{Z} \frac{\partial^2 Z}{\partial z^2} = -k^2 \quad (314)$$

Siccome ciascuna delle tre funzioni X, Y, Z dipende da una sola coordinata e siccome il secondo membro e' una costante, l'equazione 314 puo' essere verificata *per qualsiasi valore di x, y, z* solo se ciascun addendo al primo membro e' separatamente uguale a una costante. Questo significa che la 314 equivale alle tre equazioni separate:

$$\frac{1}{X} \frac{d^2 X}{dx^2} + k_x^2 = 0 \quad (315)$$

$$\frac{1}{Y} \frac{d^2 Y}{dy^2} + k_y^2 = 0 \quad (316)$$

$$\frac{1}{Z} \frac{d^2 Z}{dz^2} + k_z^2 = 0 \quad (317)$$

con il vincolo che:

$$k_x^2 + k_y^2 + k_z^2 = k^2 \quad (318)$$

(le derivate parziali sono diventate totali perche' ciascuna equazione contiene ora una sola variabile)

108. Ovviamente, essendo le tre equazioni appena scritte formalmente identiche, e' sufficiente risolverne una. Prendiamo ad esempio quella per la variabile x , che riscriviamo nel modo seguente:

$$\frac{d^2 X}{dx^2} + k_x^2 X = 0 \quad (319)$$

Questa e' un'equazione differenziale ordinaria lineare e omogenea del secondo ordine e potete trovare il metodo di risoluzione su qualsiasi testo riguardante l'argomento (ad esempio il testo di Harris e Stocker⁵).

Assumendo una soluzione del tipo $X(x) = C_x \exp(\lambda x)$ e sostituendo nella 319 si ottiene:

$$C_x \exp(\lambda x) (\lambda^2 + k_x^2) = 0 \quad (320)$$

Siccome l'esponenziale non si annulla mai, deve per forza essere:

$$\lambda^2 + k_x^2 = 0 \quad (321)$$

ovvero:

$$\lambda = \pm i k_x \quad (322)$$

L'equazione 319 ammette dunque due diverse soluzioni per ogni valore di k (ovvero per ogni valore dell'energia ϵ , in base alla 311):

$$X(x) = C_x \exp(i k_x x) \quad e \quad X(x) = C_x \exp(-i k_x x) \quad (323)$$

La generica soluzione puo' essere scritta come:

$$X(x) = C_x \exp(i k_x x) \quad (324)$$

con $k_x \in [-\infty; \infty]$ e C_x una costante (per il momento) ancora da determinare.

109. Naturalmente, le altre due equazioni hanno soluzioni formalmente identiche:

$$Y(y) = C_y \exp(i k_y y) \quad (325)$$

$$Z(z) = C_z \exp(i k_z z) \quad (326)$$

e quindi, sostituendo nella 313, si ottiene:

$$\Psi(\vec{r}) = X(x) Y(y) Z(z) = C_x C_y C_z \exp(i(k_x x + k_y y + k_z z)) \quad (327)$$

k_x, k_y e k_z possono essere pensati come componenti di un vettore \vec{k} in un corrispondente spazio 3D: $\vec{k} = (k_x, k_y, k_z)$ e quindi la 327 puo' essere riscritta in modo piu' compatto:

$$\Psi(\vec{r}) = C \exp(i \vec{k} \cdot \vec{r}) \quad (328)$$

con $C = C_x C_y C_z$ e $\vec{r} = (x, y, z)$.

La soluzione dell'equazione di Schrödinger monoelettronica indipendente dal tempo ha quindi la forma della parte spaziale di un'onda piana 220. Il vettore d'onda \vec{k} e' legato all'energia dell'elettrone dalle 318 e 311:

$$k_x^2 + k_y^2 + k_z^2 = |\vec{k}|^2 = k^2 = \frac{2m\epsilon}{\hbar^2} \quad (329)$$

110. La soluzione dell'equazione di Schrödinger dipendente dal tempo e' ottenuta sostituendo la 328 e la 309 nella 297:

$$\xi(\vec{r}, t) = \Psi(\vec{r}) \Theta(t) = C \exp\left(i \left(\vec{k} \cdot \vec{r} - \omega t\right)\right) \quad (330)$$

da cui vediamo che la funzione di stato che descrive un elettrone libero ha la forma matematica di un'onda piana progressiva rappresentata con un esponenziale complesso 218.

Vale la pena di sottolineare a questo proposito che il significato della funzione di stato *non* e' quello di un'onda *in senso fisico*; cioe' mentre la funzione che descrive un'onda del mare rappresenta l'altezza dell'acqua rispetto al suo livello medio (quindi i valori assunti dalla funzione si riferiscono a una grandezza fisica ben precisa e definita), l'interpretazione fisica della funzione di stato che descrive un elettrone non e' definita. Cio' che ha significato fisico e' il modulo quadrato di tale funzione, il cui integrale calcolato su una regione di spazio fornisce la probabilita' che l'elettrone si trovi in quella regione ad un certo istante.

111. Abbiamo ricavato la 328 assumendo che l'elettrone non sia in alcun modo confinato; cio' ha due implicazioni: in primo luogo osservate che l'energia dell'elettrone e' continua, cioe' non quantizzata (l'energia e' legata al vettore d'onda \vec{k} dalla 311 e nella derivazione nulla impedisce alle componenti di \vec{k} di variare in modo continuo); in secondo luogo, la costante C e' ancora indeterminata (come gia' detto, tale costante include anche un termine che si riferisce alla parte temporale della soluzione).

Ora vedremo come questi due aspetti dello stato dell'elettrone cambiano se esso e' confinato nel volume V del solido.

112. Il confinamento dell'elettrone consente innanzitutto di determinare la costante C che compare nella 328. La densita' di probabilita' associata alla $\Psi(\vec{r})$ e':

$$\Psi^*(\vec{r})\Psi(\vec{r}) = |\Psi(\vec{r})|^2 \quad (331)$$

Tale densita' di probabilita', integrata in una regione di spazio, fornisce la probabilita' di trovare l'elettrone in quella regione.

Siccome l'elettrone non puo' trovarsi al di fuori del volume del solido V , la probabilita' di trovarlo in un punto *qualsiasi* interno a tale volume deve essere unitaria. Cio' consente di normalizzare la $\Psi(\vec{r})$:

$$\int_V \Psi^*(\vec{r})\Psi(\vec{r}) d\vec{r} = 1 \quad (332)$$

Sostituendo (e assumendo per semplicita' che C sia reale):

$$\int_V C \exp(-i\vec{k} \cdot \vec{r}) C \exp(i\vec{k} \cdot \vec{r}) d\vec{r} = C^2 \int_V d\vec{r} = C^2 V = 1 \quad (333)$$

da cui segue:

$$C = \frac{1}{\sqrt{V}} \quad (334)$$

Quindi l'espressione della $\Psi(\vec{r})$ normalizzata rispetto al volume del solido V e':

$$\Psi(\vec{r}) = \frac{1}{\sqrt{V}} \exp(i\vec{k} \cdot \vec{r}) \quad (335)$$

113. Una seconda conseguenza del confinamento dell'elettrone nel volume V e' la necessita' che la funzione di stato $\Psi(\vec{r})$ debba soddisfare delle corrispondenti condizioni al contorno.

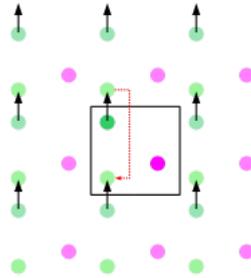


Figura 39: Funzionamento delle condizioni al contorno periodiche in 2D: un elettrone esce dal campione solido (costituito dal quadrato mostrato) verso l'alto, rientrando con la stessa direzione dal basso.

A questo punto dobbiamo essere piu' precisi nella definizione del volume del solido V . Siccome siamo interessati a descrivere il comportamento degli elettroni *all'interno* del solido, la scelta della sua forma esteriore non ha alcuna importanza. Conviene quindi scegliere la forma piu' semplice, che e' quella di un cubo di lato $L = \sqrt[3]{V}$.

Inoltre, per rendere trascurabile qualsiasi effetto dovuto alla superficie del solido, imponiamo delle cosiddette *condizioni al contorno periodiche* (dette anche condizioni di Born e von Karman).

In pratica, l'elettrone e' libero di uscire dal volume V , ma quando lo fa, esso rientra automaticamente dalla parte opposta mantenendo inalterata la sua velocita' (sia come direzione che come modulo) (avrete sicuramente visto un effetto di questo genere in qualche videogioco). La logica delle condizioni al contorno periodiche per il caso bidimensionale e' mostrata nella figura 39.

E' importante realizzare che, con le condizioni al contorno periodiche, tutto funziona come se il cubo di volume V che costituisce il solido fosse *la cella primitiva di un "cristallo"* (notate che tale cella primitiva e' enorme rispetto a quella di un cristallo "reale", visto che le sue dimensioni sono quelle di un campione solido macroscopico (*cm*), mentre quelle della cella primitiva di un vero cristallo sono dell'ordine degli ångström).

Notate come le condizioni al contorno periodiche rimuovano completamente qualsiasi riferimento alla superficie del campione: la facce del cubo che costituisce il campione non sono esposte all'ambiente, ma alle repliche del campione stesso traslate per periodicitá.

Dal punto di vista matematico, le condizioni di Born e van Karman sono espresse nel modo seguente:

$$\Psi(x + L, y, z) = \Psi(x, y, z) \quad (336)$$

$$\Psi(x, y + L, z) = \Psi(x, y, z) \quad (337)$$

$$\Psi(x, y, z + L) = \Psi(x, y, z) \quad (338)$$

Il che significa semplicemente che la $\Psi(\vec{r})$ deve essere *periodica* su un reticolo cubico semplice con costante di cella pari a $L = \sqrt[3]{V}$.

114. La conseguenza della richiesta di tale periodicità e', come ora vedremo facilmente, la quantizzazione del vettore d'onda \vec{k} . Per essere periodica di periodo L lungo la direzione x , la funzione Ψ deve soddisfare:

$$\begin{aligned} \exp(i(k_x(x+L) + k_y y + k_z z)) &= \exp(ik_x L) \exp(i(k_x x + k_y y + k_z z)) \\ &= \exp(i(k_x x + k_y y + k_z z)) \end{aligned} \quad (339)$$

Ma ciò può avvenire solo se:

$$\exp(ik_x L) = 1 \quad (340)$$

Riscrivendo l'esponenziale con la formula di Eulero vi rendete conto che:

$$\exp(ik_x L) = 1 \iff k_x = \frac{2\pi n_x}{L} \quad (341)$$

con $n_x \in \mathbb{Z}$.

Ovviamente, una condizione analoga deve valere per la periodicità lungo y e z . Le condizioni al contorno periodiche richiedono quindi che le componenti del vettore d'onda \vec{k} (che in loro assenza possono assumere con continuità qualsiasi valore), debbano soddisfare:

$$\begin{aligned} k_x &= \frac{2\pi n_x}{L}, & n_x &\in \mathbb{Z} \\ k_y &= \frac{2\pi n_y}{L}, & n_y &\in \mathbb{Z} \\ k_z &= \frac{2\pi n_z}{L}, & n_z &\in \mathbb{Z} \end{aligned} \quad (342)$$

e siccome n_x, n_y, n_z sono numeri interi, i valori di k_x, k_y, k_z sono *quantizzati*.

115. Al reticolo di Bravais cubico semplice del quale il campione solido macroscopico che stiamo considerando costituisce la cella primitiva e' associato un reticolo reciproco, anch'esso cubico semplice (punto 48) e la cui costante di cella e' data da $2\pi/L$.

Ma allora le equazioni 342 dicono che i vettori \vec{k} permessi (cioe' compatibili con le condizioni al contorno periodiche) sono tutti e soli i punti di tale reticolo reciproco.

Detti $\hat{x}, \hat{y}, \hat{z}$ i vettori unitari nelle tre direzioni ortogonali degli spigoli della cella primitiva (cubica) del reticolo reciproco appena introdotto, un generico punto di esso e' dato dalla combinazione lineare:

$$n_x \frac{2\pi}{L} \hat{x} + n_y \frac{2\pi}{L} \hat{y} + n_z \frac{2\pi}{L} \hat{z} = k_x \hat{x} + k_y \hat{y} + k_z \hat{z} = \vec{k} \quad (343)$$

116. Naturalmente, la quantizzazione del vettore \vec{k} implica la quantizzazione dell'energia ϵ , legata al quadrato del modulo di \vec{k} (eq. 311):

$$\epsilon = \frac{\hbar^2 k^2}{2m} \quad (344)$$

La relazione fra ϵ e k viene anche detta "relazione di dispersione".

Vediamo ora alcune osservazioni riguardanti il risultato ottenuto.

117. Conteggio degli stati permessi.

L'equazione 335 mostra che gli stati permessi all'elettrone indipendente e libero sono indicizzati dal vettore \vec{k} (ovvero dai tre numeri interi n_x, n_y, n_z che determinano le sue componenti tramite le equazioni 342): ad ogni vettore \vec{k} permesso corrisponde uno stato accessibile all'elettrone. Abbiamo anche stabilito che i vettori d'onda permessi sono tutti e soli i punti del reticolo reciproco associato al reticolo di Bravais che realizza le condizioni al contorno periodiche che abbiamo utilizzato (eq. 342).

Ne segue che il numero totale di stati accessibili a ogni elettrone che costituisce il gas e' uguale al numero di punti del reticolo reciproco. Questo numero e' ovviamente infinito, ma i punti (ovvero gli stati accessibili) sono discreti.

Notiamo inoltre che nel ricavare la funzione di stato 335 abbiamo implicitamente assunto che essa non dipenda dallo spin elettronico. Ma allora, per il principio di Pauli, possiamo dire che ogni stato del tipo 335 puo' descrivere *due* elettroni con spin antiparallelo. Detto in termini equivalenti, un numero N di vettori \vec{k} permessi individua un numero doppio ($2N$) di stati *monoelettronici* accessibili.

118. Superfici isoenergetiche e degenerazione.

L'equazione 344 fornisce la relazione fra l'energia di un elettrone e il vettore d'onda che indicizza il suo stato.

Tale relazione nello spazio \vec{k} e' di tipo parabolico:

$$\epsilon = \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2 (k_x^2 + k_y^2 + k_z^2)}{2m} \quad (345)$$

Siccome l'energia dipende dal solo modulo del vettore d'onda, ne segue che tutti i punti \vec{k} che giacciono su una sfera centrata nell'origine individuano stati alla stessa energia, cioe' degeneri.

L'equazione di tale sfera isoenergetica e':

$$k_x^2 + k_y^2 + k_z^2 = \frac{2m\epsilon}{\hbar^2} \quad (346)$$

Lo stato elettronico a energia minore e' quello corrispondente al punto $\vec{k} = (0, 0, 0)$ che sta all'origine (per il quale si ha $\epsilon = 0$). Lo stato a energia immediatamente superiore e' quello corrispondente ai vettori \vec{k} che sono i primi vicini dell'origine (guardate la fig. 40). Per il reticolo reciproco cubico semplice che stiamo considerando questi primi vicini sono 6 e il valore del loro modulo comune e' pari alla costante di cella del reticolo reciproco, cioe' $2\pi/L$; questo e' anche il raggio della sfera isoenergetica che li contiene. Le componenti di questi vettori, in unita' di $2\pi/L$, sono:

$$(\pm 1, 0, 0) \quad (0, \pm 1, 0) \quad (0, 0, \pm 1) \quad (347)$$

Questi 6 punti individuano altrettanti stati elettronici degeneri a energia:

$$\epsilon = \frac{\hbar^2 \left(\frac{2\pi}{L}\right)^2}{2m} \quad (348)$$

La sfera isoenergetica immediatamente successiva corrisponde ai 12 vettori che costituiscono la diagonale di 3 facce di ciascuna delle 8 celle cubiche in contatto

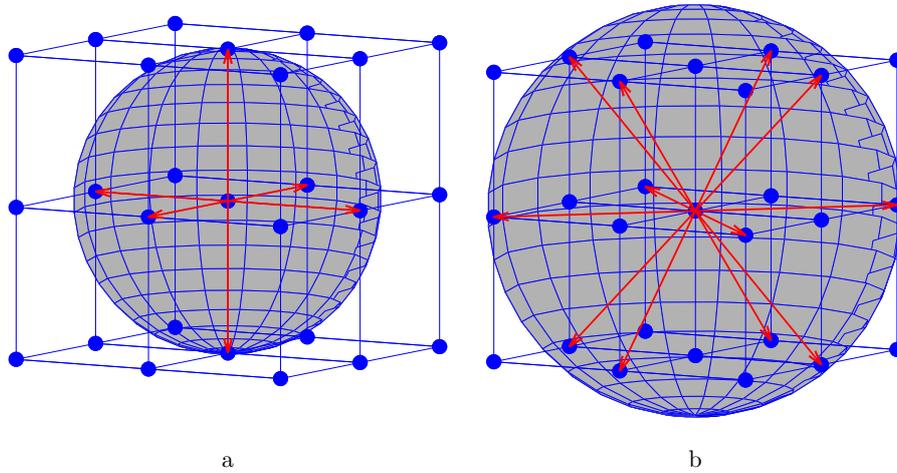


Figura 40: Per un reticolo cubico semplice: a: sfera isoenergetica corrispondente ai primi vicini dell'origine; b: sfera isoenergetica corrispondente ai secondi vicini dell'origine.

con l'origine (in totale ci sono $3 \times 8 = 24$ facce, ma ciascuna di esse è condivisa fra due delle 8 celle e quindi le facce distinte sono solo $24/2 = 12$). Il modulo comune di questi vettori è:

$$\sqrt{\left(\frac{2\pi}{L}\right)^2 + \left(\frac{2\pi}{L}\right)^2} = \frac{2\pi}{L}\sqrt{2} \quad (349)$$

La seconda sfera isoenergetica contiene quindi 12 stati degeneri a energia:

$$\epsilon = \frac{\hbar^2 \left(\frac{2\pi}{L}\sqrt{2}\right)^2}{2m} \quad (350)$$

Dovreste rendervi conto facilmente che all'aumentare dell'energia, cioè all'aumentare della distanza dall'origine, il numero dei punti \vec{k} degeneri, cioè i punti \vec{k} permessi giacenti su una data sfera isoenergetica, aumenta molto rapidamente.

119. Spaziatura dei livelli energetici.

Per avere un'idea della spaziatura dei livelli energetici possiamo calcolare la differenza di energia fra le prime due sfere isoenergetiche, di raggio $2\pi/L$ e $\frac{2\pi}{L}\sqrt{2}$:

$$\frac{\hbar^2 \left(\frac{2\pi}{L}\sqrt{2}\right)^2}{2m} - \frac{\hbar^2 \left(\frac{2\pi}{L}\right)^2}{2m} = \frac{\hbar^2}{2m} \left(\frac{2\pi}{L}\right)^2 = 1.5 \times 10^{-14} \text{ eV} \quad (351)$$

dove ho usato: $\hbar = 1.05 \times 10^{-34} \text{ J s}$, $m = 9.11 \times 10^{-31} \text{ kg}$, $L = 1 \times 10^{-2} \text{ m}$ e $1 \text{ J} = 6.24 \times 10^{18} \text{ eV}$.

Questo è un valore estremamente piccolo (e la spaziatura fra i livelli successivi è ancora minore). Per farvi un'idea, la differenza di energia fra il livello $2s$ e il livello $1s$ dell'atomo di idrogeno è 10.2 eV .

In altre parole, i livelli elettronici nel modello del gas di elettroni sono praticamente continui.

120. Conteggio dei punti \vec{k} contenuti in una certa regione dello spazio di Fourier. Molto spesso per il calcolo di proprietà dei solidi e in generale per la descrizione della loro struttura elettronica serve conoscere quanti punti \vec{k} permessi sono contenuti all'interno di una certa regione dello spazio reciproco. Inoltre, normalmente la regione di interesse contiene un numero molto grande di punti. Sotto questa ipotesi, il numero cercato può essere calcolato in modo praticamente esatto come:

$$\frac{\text{volume della regione}}{\text{volume per punto } \vec{k}} \quad (352)$$

Siccome i punti \vec{k} permessi sono tutti e soli i punti del reticolo reciproco, che nel modello del gas di elettroni che stiamo considerando, è un reticolo cubico semplice, il volume per punto \vec{k} è semplicemente quello della cella primitiva del reticolo reciproco, ovvero:

$$\left(\frac{2\pi}{L}\right)^3 = \frac{8\pi^3}{V} \quad (353)$$

(allo stesso risultato si arriva in base alla 244)

121. Riempimento dei livelli elettronici nello stato fondamentale.

Lo stato fondamentale (ground state) di un metallo è lo stato del metallo a 0 K. A 0 K i livelli elettronici del metallo sono occupati a partire da quello a minore energia ($\epsilon = 0$, corrispondente al punto \vec{k} nell'origine) e in ordine di energia crescente, nel rispetto del principio di Pauli. Come abbiamo detto, ad ogni punto \vec{k} corrispondono due stati monoelettronici che differiscono solo per lo spin.

Siccome il numero N di elettroni presenti in un campione metallico di dimensioni macroscopiche è enorme (dell'ordine del numero di Avogadro) e siccome l'energia dei livelli dipende dal solo modulo del vettore d'onda, la regione dello spazio \vec{k} che contiene tutti i punti corrispondenti a stati occupati nel ground state è a tutti gli effetti una sfera. Il raggio di tale sfera (detta *sfera di Fermi*) è uguale al modulo del più lungo vettore \vec{k} corrispondente a uno stato occupato e viene normalmente indicato con k_F e chiamato *vettore d'onda di Fermi*.

Il numero di punti \vec{k} contenuti nella sfera di Fermi si ottiene come detto al punto 204:

$$\frac{\text{volume della sfera di Fermi}}{\text{volume per punto } \vec{k}} = \frac{\frac{4}{3}\pi k_F^3}{\frac{8\pi^3}{V}} = \frac{k_F^3}{6\pi^2} V \quad (354)$$

Il numero di punti \vec{k} contenuti nella sfera di Fermi è uguale alla metà del numero di elettroni N (perché ogni stato indicizzato da \vec{k} può contenere 2 elettroni). Quindi:

$$\frac{1}{2}N = \frac{k_F^3}{6\pi^2} V \Rightarrow \frac{N}{V} = n = \frac{k_F^3}{3\pi^2} \Rightarrow k_F^3 = n3\pi^2 \quad (355)$$

$n = N/V$ non è altro che la densità di elettroni del campione solido e si può ricavare facilmente nota la densità ρ del metallo (espressa in unità di massa

per unita' di volume), la sua massa molare M e il numero di elettroni di valenza Z :

$$n = \frac{N}{V} = \frac{\text{moli} \times N_o \times Z}{V} = \frac{\text{massa}}{M} \times N_o Z = N_o \frac{Z}{M} \rho \quad (356)$$

(N_o e' il numero di Avogadro)

Ad esempio per il sodio ($Z = 1$, $M = 22.99$ g/mol, $\rho = 0.968$ g/cm³) si ha:

$$n = N_o \frac{Z}{M} \rho = 2.53 \times 10^{22} \text{ e/cm}^3 \quad (357)$$

Inserendo la 355 nella 344, si ricava:

$$\epsilon_F = \frac{\hbar^2 k_F^2}{2m} = \frac{\hbar^2}{2m} (k_F^3)^{\frac{2}{3}} = \frac{\hbar^2}{2m} (n3\pi^2)^{\frac{2}{3}} \quad (358)$$

ϵ_F e' l'energia del piu' alto livello occupato a 0 K e viene detta *energia (o livello) di Fermi*.

Per il sodio, utilizzando il valore della densita' di elettroni calcolato sopra e i valori delle costanti, l'energia di Fermi calcolata nel modello del gas di elettroni ha il valore:

$$\epsilon_F = \frac{\hbar^2}{2m} (n3\pi^2)^{\frac{2}{3}} = 3.14 \text{ eV} \quad (359)$$

122. Un significato fisico del vettore d'onda \vec{k} .

L'operatore che rappresenta il momento (lineare) di una particella e':

$$\frac{\hbar}{i} \nabla = \frac{\hbar}{i} \left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y} + \frac{\partial}{\partial z} \right) \quad (360)$$

E' facile verificare che tale operatore commuta con l'hamiltoniano e quindi la funzione $\Psi(\vec{r})$ e' autofunzione anche dell'operatore del momento. Applicando tale operatore alla $\Psi(\vec{r})$ si ottiene:

$$\frac{\hbar}{i} \nabla \Psi(\vec{r}) = \frac{\hbar}{i} \nabla \frac{1}{\sqrt{V}} \exp(i\vec{k} \cdot \vec{r}) = \hbar \vec{k} \frac{1}{\sqrt{V}} \exp(i\vec{k} \cdot \vec{r}) = \hbar \vec{k} \Psi(\vec{r}) \quad (361)$$

Troviamo quindi che $\hbar \vec{k}$ e' l'autovalore della $\Psi(\vec{r})$ rispetto all'operatore del momento; in altre parole, un elettrone descritto dalla $\Psi(\vec{r})$ ha un momento $\vec{p} = \hbar \vec{k}$.

Come abbiamo detto, nel gas di Fermi gli elettroni hanno solo energia cinetica (essi non interagiscono fra loro ne' con gli ioni); il risultato appena ottenuto ci consente di esprimere cio' anche nel modo usuale della fisica classica. Infatti, avendo appena visto che $\hbar \vec{k} = \vec{p}$ e' il momento dell'elettrone, la 344 si puo' riscrivere come:

$$\epsilon = \frac{\hbar^2 k^2}{2m} = \frac{p^2}{2m} = \frac{(mv)^2}{2m} = \frac{1}{2} mv^2 \quad (362)$$

Un elettrone che possiede l'energia di Fermi ϵ_F si muove con una velocita' v_F (detta *velocita' di Fermi*), data da:

$$v_F = \sqrt{\frac{2\epsilon_F}{m}} \quad (363)$$

Di nuovo, per il sodio ($\epsilon_F = 3.14 \text{ eV}$):

$$v_F = \sqrt{\frac{2\epsilon_F}{m}} = 1051 \text{ km/s} \quad (364)$$

E' una velocita' molto elevata: circa 0.3% della velocita' della luce. Nella cornice del modello del gas di Fermi, questa e' la velocita con cui si trasmette un segnale elettrico attraverso un cavo metallico.

123. La distribuzione di Fermi–Dirac.

Nella parte precedente abbiamo discusso l'occupazione dei livelli energetici nel gas di elettroni nello stato fondamentale (cioe' a 0 K). In tali condizioni N elettroni occupano i primi $N/2$ stati \vec{k} in ordine di energia (modulo di \vec{k}) crescente.

Se la temperatura e' $> 0 \text{ K}$, alcuni livelli elettronici che non erano occupati a 0 K vengono occupati e quindi dei corrispondenti livelli che erano occupati a 0 K rimangono vuoti.

Una grandezza particolarmente importante a questo riguardo e' la probabilita' che, a una certa temperatura, un certo livello elettronico sia occupato. Una funzione che ad ogni possibile valore di una variabile associa la corrispondente probabilita' viene detta in generale *funzione di distribuzione*.

Cio' che faremo in questa sezione e' ricavare la funzione di distribuzione che associa all'energia di ciascun livello monoelettronico la probabilita' che quel livello sia occupato.

Per livello monoelettronico intendo, in questo contesto, un livello energetico che puo' "ospitare" un singolo elettrone; esso sara' pertanto univocamente determinato da un punto \vec{k} permesso e da una particolare orientazione ($+\hbar/2$ o $-\hbar/2$) dello spin.

124. E' facile rendersi conto che la funzione di distribuzione di cui stiamo parlando deve dipendere, oltre che dall'energia del livello considerato, anche dalla temperatura. Ad esempio, per $T = 0 \text{ K}$, la probabilita' di occupazione dei primi N livelli monoelettronici (cioe' i primi $N/2$ punti \vec{k} permessi) e' esattamente 1 e la probabilita' di occupazione di tutti i livelli monoelettronici a energia superiore e' esattamente 0.

Per temperature superiori a 0 K la probabilita' di occupazione dei primi N livelli monoelettronici diminuira' (rispetto al valore unitario) e quella dei livelli a energia superiore non sara' piu' esattamente nulla.

La funzione di distribuzione che ricaveremo si chiama *distribuzione di Fermi–Dirac*. Una caratteristica fondamentale di questa distribuzione e' che, come vedremo nella derivazione, essa tiene conto del principio di Pauli. E' proprio questo aspetto che ha determinato un drastico passo avanti nella descrizione del comportamento degli elettroni nei solidi, rispetto alle teorie iniziali, in cui gli elettroni erano considerati come delle particelle classiche (quindi *non* soggette al principio di Pauli).

125. Potete apprezzare la differenza che fa la considerazione del principio di Pauli nel calcolo dell'occupazione dei livelli energetici accessibili agli elettroni considerando il gas di Fermi a 0 K: se gli elettroni non seguissero il principio di Pauli, in tali condizioni *tutti* occuperebbero lo stesso livello, ovvero il piu'

basso possibile a $\epsilon = 0$ e l'energia del ground state sarebbe nulla. Invece, il fatto che ciascun elettrone debba essere descritto da una serie di numeri quantici differente da quella di tutti gli altri (principio di Pauli), implica che anche a 0 K N elettroni debbano occupare stati a energia diversa da zero (e cioè' gli stati corrispondenti ai primi $N/2$ vettori \vec{k} permessi, in ordine di modulo crescente). Questo, ad esempio, fa sì che, come abbiamo visto (punto 121) l'energia di Fermi del sodio sia 3.14 eV oppure che i metalli siano dei conduttori elettrici anche a 0 K (perché' la velocità' degli elettroni non è' nulla, eq. 363)

Il principio di Pauli si applica non solo agli elettroni, ma a tutte le particelle quantistiche con spin *semiintero* ($\pm\hbar(n + 1/2)$). Tali particelle vengono dette in generale *fermioni* e vengono distinte dalle particelle con spin intero che non sono soggette al principio di Pauli (e vengono chiamate *bosoni*).

126. Per la derivazione della distribuzione di Fermi-Dirac seguìo il libro di Ashcroft e Mermin.⁸

Consideriamo un sistema in equilibrio termico alla temperatura T costituito da N elettroni, ciascuno avente a disposizione un numero (infinito) di livelli energetici monoelettronici accessibili. Vogliamo determinare la funzione $f_i^N(\epsilon_i)$ che fornisce la probabilità' di occupazione dello stato monoelettronico i -esimo a energia ϵ_i nel sistema di N elettroni.

In un generico stato di questo sistema gli elettroni occupano N stati mono-elettronici diversi perché' devono soddisfare il principio di esclusione di Pauli. Chiameremo *stato N -elettronico* un particolare stato del sistema in cui ciascun elettrone occupa un determinato stato mono-elettronico.

Cominciamo con la seguente osservazione.

La probabilità' $f_i^N(\epsilon_i)$ che lo stato mono-elettronico i sia occupato si può' scrivere come somma delle probabilità' (indipendenti) $P_N(E_\alpha^N)$, essendo $P_N(E_\alpha^N)$ la probabilità' che il sistema si trovi nell' α -esimo stato N -elettronico in cui lo stato mono-elettronico i è' occupato da un elettrone:

$$f_i^N(\epsilon_i) = \sum_{\alpha} P_N(E_\alpha^N) \quad (365)$$

(α corre su tutti gli stati N -elettronici in cui lo stato mono-elettronico i è' occupato)

Un'espressione analitica per la probabilità' $P_N(E_\alpha^N)$ è' fornita dalla meccanica statistica: in generale, la probabilità' che un sistema di N particelle in equilibrio termico alla temperatura T si trovi in un certo stato a energia E_α^N è' data da:

$$P_N(E_\alpha^N) = \frac{\exp\left(-\frac{E_\alpha^N}{k_B T}\right)}{\sum_p \exp\left(-\frac{E_p^N}{k_B T}\right)} \quad (366)$$

(k_B è' la costante di Boltzmann)

La somma al denominatore si chiama *funzione di partizione* ed è' estesa a *tutti* gli stati accessibili al sistema.

La funzione di partizione è' legata all'energia libera di Helmholtz F , definita in termodinamica classica come:

$$F = U - TS \quad (367)$$

dove U ed S sono, rispettivamente, l'energia interna e l'entropia del sistema. La relazione che lega la funzione di partizione all'energia libera di Helmholtz e':¹⁰

$$F_N = -k_B T \ln \left(\sum_p \exp \left(-\frac{E_p^N}{k_B T} \right) \right) \quad (368)$$

ovvero:

$$\sum_p \exp \left(-\frac{E_p^N}{k_B T} \right) = \exp \left(-\frac{F_N}{k_B T} \right) \quad (369)$$

Alla luce della 369, la 366 puo' essere riscritta in modo piu' compatto:

$$P_N (E_\alpha^N) = \exp \left(-\frac{E_\alpha^N - F_N}{k_B T} \right) \quad (370)$$

Torniamo ora alla 365. In base al principio di Pauli, per uno stato monoelettronico fissato possono verificarsi solo due casi: o esso e' occupato da un elettrone, oppure esso e' vuoto. Da cio' segue che la probabilita' che lo stato i sia occupato si puo' anche scrivere come il complemento a 1 della probabilita' che lo stato i non sia occupato:

$$f_i^N (\epsilon_i) = 1 - [\text{prob. che lo stato } i \text{ non sia occupato}] \quad (371)$$

Sottolineo che questa equazione e' una *diretta conseguenza del principio di Pauli*, che quindi, come dicevo, e' uno dei punti chiave che stanno alla base della distribuzione di Fermi-Dirac.

Analogamente a quanto gia' visto, la probabilita' che lo stato i non sia occupato deve essere la somma delle probabilita' $P_N (E_\gamma^N)$, essendo $P_N (E_\gamma^N)$ la probabilita' che il sistema si trovi nel γ -esimo stato N -elettronico in cui lo stato monoelettronico i non e' occupato da un elettrone.

Quindi:

$$f_i^N (\epsilon_i) = 1 - \sum_\gamma P_N (E_\gamma^N) \quad (372)$$

(γ corre su tutti gli stati N -elettronici in cui lo stato monoelettronico i non e' occupato)

Ora: sempre grazie all'assunzione della validita' del principio di Pauli, l'energia di uno stato N -elettronico in cui lo stato monoelettronico i non e' occupato si puo' sempre scrivere come differenza fra l'energia di uno stato $(N+1)$ -elettronico in cui lo stato i e' occupato e l'energia ϵ_i dello stato i stesso, cioe':

$$E_\gamma^N = E_\alpha^{N+1} - \epsilon_i \quad (373)$$

Inoltre, il numero degli stati N -elettronici in cui lo stato i non e' occupato e' esattamente uguale al numero degli stati $(N+1)$ -elettronici in cui lo stato i e' occupato.

Queste due osservazioni permettono di scrivere la seguente uguaglianza:

$$\sum_{\gamma} P_N(E_{\gamma}^N) = \sum_{\alpha} P_N(E_{\alpha}^{N+1} - \epsilon_i) \quad (374)$$

γ corre su tutti gli stati N -elettronici in cui lo stato monoelettronico i non e' occupato; α corre su tutti gli stati $(N + 1)$ -elettronici in cui lo stato monoelettronico i e' occupato: e' importante realizzare che al variare dei due indici γ e α vengono individuati esattamente gli stessi stati N -elettronici.

La 372 puo' quindi essere riscritta come:

$$f_i^N(\epsilon_i) = 1 - \sum_{\alpha} P_N(E_{\alpha}^{N+1} - \epsilon_i) \quad (375)$$

Ora utilizziamo la 370 per riscrivere i termini della sommatoria:

$$\begin{aligned} P_N(E_{\alpha}^{N+1} - \epsilon_i) &= \exp\left(-\frac{(E_{\alpha}^{N+1} - \epsilon_i) - F_N}{k_B T}\right) \\ &\quad (\text{sommo e sottraggo } F_{N+1}) \\ &= \exp\left(-\frac{(E_{\alpha}^{N+1} - \epsilon_i) - F_{N+1} + F_{N+1} - F_N}{k_B T}\right) \\ &= \exp\left(-\frac{E_{\alpha}^{N+1} - F_{N+1} - \epsilon_i + (F_{N+1} - F_N)}{k_B T}\right) \\ &= \exp\left(-\frac{E_{\alpha}^{N+1} - F_{N+1}}{k_B T}\right) \exp\left(\frac{\epsilon_i - (F_{N+1} - F_N)}{k_B T}\right) \\ &= \exp\left(\frac{\epsilon_i - \mu}{k_B T}\right) \exp\left(-\frac{E_{\alpha}^{N+1} - F_{N+1}}{k_B T}\right) \end{aligned} \quad (376)$$

$\mu = F_{N+1} - F_N$ e' la variazione che subisce l'energia libera di Helmholtz del sistema quando viene aggiunto un elettrone a T e V costanti: la termodinamica classica ci dice che μ e' il potenziale chimico del sistema di N elettroni.

Dalla definizione $F = U - TS$ si ha:

$$dF = dU - TdS - SdT \quad (377)$$

Introducendo il primo principio per un sistema aperto contenente un solo tipo di particelle:

$$dU = TdS - PdV + \mu dN \quad (378)$$

si ha:

$$dF = TdS - PdV + \mu dN - TdS - SdT = -SdT - PdV + \mu dN \quad (379)$$

A T e V costanti (il nostro caso) otteniamo:

$$dF = \mu dN \quad (380)$$

Essendo il numero di particelle (gli elettroni del gas) molto grande, il differenziale dN puo' essere scritto in modo praticamente esatto come:

$$dN = (N + 1) - N = 1 \quad (381)$$

(cioe': assumo che variare N di un'unita' costituisca una variazione infinitesima)
Allora:

$$dF = F_{N+1} - F_N = \mu((N+1) - N) = \mu \quad (382)$$

Tenendo sempre presente la 370, si vede che il secondo esponenziale della 376 non e' altro che $P_{N+1}(E_\alpha^{N+1})$. Quindi, sostituendo la 376 nella 372 arriviamo a:

$$f_i^N(\epsilon_i) = 1 - \exp\left(\frac{\epsilon_i - \mu}{k_B T}\right) \sum_\alpha P_{N+1}(E_\alpha^{N+1}) \quad (383)$$

(α corre su tutti gli stati $(N+1)$ -elettronici in cui lo stato monoelettronico i e' occupato)

Se ora riguardate la 365 realizzate subito che:

$$\sum_\alpha P_{N+1}(E_\alpha^{N+1}) = f_i^{N+1}(\epsilon_i) \quad (384)$$

e quindi arriviamo a:

$$f_i^N(\epsilon_i) = 1 - \exp\left(\frac{\epsilon_i - \mu}{k_B T}\right) f_i^{N+1}(\epsilon_i) \quad (385)$$

Se, come sempre e' il caso per un solido reale, N e' un numero enorme (dell'ordine del numero di Avogadro), sarebbe assurdo pensare che l'aggiunta di un elettrone al sistema possa modificare la probabilita' che lo stato i sia occupato per piu' di una quantita' a tutti gli effetti insignificante. Ma allora possiamo lecitamente assumere che sia:

$$f_i^{N+1}(\epsilon_i) = f_i^N(\epsilon_i) \quad (386)$$

e quindi:

$$f_i^N(\epsilon_i) = 1 - \exp\left(\frac{\epsilon_i - \mu}{k_B T}\right) f_i^N(\epsilon_i) \quad (387)$$

da cui si arriva all'espressione finale per la distribuzione di Fermi-Dirac:

$$f(\epsilon) = \frac{1}{\exp\left(\frac{\epsilon - \mu}{k_B T}\right) + 1} \quad (388)$$

dove ho dismesso gli indici N e i , essendo sottinteso che ϵ e' l'energia (quantizzata secondo la 344) di uno stato permesso e il numero di elettroni e' costante (determinato dal potenziale chimico $\mu = F_{N+1} - F_N$).

Nella 388 $f(\epsilon)$ e' la probabilita' che in un sistema di fermioni in equilibrio termico alla temperatura T e confinato in un volume V (come l'insieme degli elettroni che costituisce il gas di Fermi che stiamo considerando) lo stato di particella singola a energia ϵ sia occupato.

Facciamo alcune osservazioni sul risultato ottenuto.

127. Il principio di esclusione di Pauli.

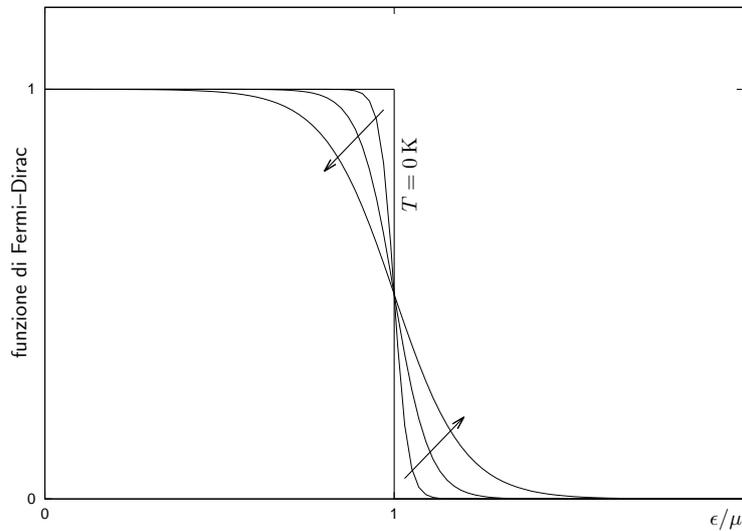


Figura 41: Andamento della funzione di Fermi-Dirac in funzione dell'energia e per diverse temperature; per $T = 0$ K la funzione diventa uno scalino; le frecce indicano la direzione in cui la temperatura aumenta.

Se ripercorrete la derivazione della distribuzione di Fermi-Dirac vi rendete conto che l'ipotesi di validità del principio di Pauli è cruciale. Gli elettroni possono distribuirsi in qualsiasi modo sui livelli permessi, ma il risultato ottenuto è valido solo sotto l'ipotesi che ciascuno stato monoelettronico permesso descriva *al più* un singolo elettrone (o nessuno).

Questa è un'ipotesi drasticamente diversa da quella su cui si basa la distribuzione di Maxwell-Boltzmann valida per un gas di particelle classiche, in cui lo stesso stato può essere occupato da un numero qualsiasi di particelle.

128. L'andamento della funzione di Fermi-Dirac.

La funzione di Fermi-Dirac è una funzione di due variabili: l'energia ϵ e la temperatura T . Tuttavia, le sue caratteristiche principali si possono osservare in un grafico bidimensionale in cui si riporta la funzione al variare dell'energia per diversi valori fissati di temperatura.

Questo è mostrato nella figura 41, dove ho diagrammato la funzione nella forma:

$$f\left(\frac{\epsilon}{\mu}, \frac{k_B T}{\mu}\right) = \frac{1}{\exp\left(\frac{\epsilon/\mu - 1}{k_B T/\mu}\right) + 1} \quad (389)$$

in funzione di ϵ/μ per vari valori di $k_B T/\mu$, in modo da rendere il grafico indipendente dal particolare valore di μ .

Quando $\epsilon/\mu - 1 \ll 0$, ovvero quando $\epsilon \ll \mu$, la funzione tende a 1; all'opposto, quando $\epsilon/\mu - 1 \gg 0$, ovvero quando $\epsilon \gg \mu$, la funzione tende a 0; ciò è quello che possiamo aspettarci: a una data temperatura è molto probabile che stati a energia molto bassa siano occupati mentre è poco probabile che vengano occupati stati a energia molto alta.

Quando la temperatura tende a 0 K la funzione tende ad assumere l'andamento di uno scalino: vale 1 per $\epsilon < \mu$ e 0 per $\epsilon > \mu$ (guardate la figura). Questo è

esattamente quello che ci aspettiamo sulla base di quanto detto al punto 121. A 0 K gli stati a energia \leq dell'energia di Fermi sono tutti occupati, mentre quelli a energia $>$ dell'energia di Fermi sono tutti vuoti. In termini della probabilita' di occupazione deve pertanto essere:

$$f_{T=0\text{K}}(\epsilon) = \begin{cases} 1 & \epsilon \leq \mu \\ 0 & \epsilon > \mu \end{cases} \quad (390)$$

Da cio' segue anche che, a $T = 0\text{K}$, il potenziale chimico degli elettroni e' uguale all'energia di Fermi:

$$\mu = \epsilon_F \quad \text{a } 0\text{K} \quad (391)$$

Notate tuttavia che il potenziale chimico μ dipende dalla temperatura (mentre non e' cosi' per l'energia di Fermi) e quindi, come ho indicato esplicitamente, l'uguaglianza appena scritta vale solo a 0K. Per temperature superiori a 0K, l'energia di Fermi rimane inalterata (perche' essa e' *definita* a 0K), mentre il potenziale chimico varia con la temperatura.

La dipendenza del potenziale chimico dalla temperatura si vede dalla sua definizione $\mu = F_{N+1} - F_N$ nell'equazione 376 e dalla dipendenza dell'energia di Helmholtz F dalla temperatura (equazione 368).

In generale, in un range di temperature molto ampio, la variazione del potenziale chimico rispetto al suo valore a 0K (l'energia di Fermi) e' molto piccola: questo consente di assumere la validita' dell'uguaglianza 391 anche a temperature > 0 , in contesti in cui non sia richiesta una elevata precisione.

129. Per temperature maggiori di 0K la distribuzione di Fermi-Dirac si "smussa" progressivamente (guardate la figura 41). In termini di probabilita' di occupazione degli stati, questo smussamento significa che l'occupazione di alcuni stati che erano occupati a 0K diventa meno probabile e contemporaneamente quella di alcuni stati che non erano occupati a 0K diventa piu' probabile. In termini equivalenti, a temperatura maggiore una frazione di elettroni viene eccitata dagli stati che occupava nel ground state a stati a piu' alta energia, *sempre nel rispetto del principio di esclusione*.

In ogni caso, come potete rendervi conto facilmente guardando l'equazione 388, la probabilita' di occupazione degli stati che hanno un'energia pari al potenziale chimico μ rimane sempre uguale al 50% ad ogni temperatura.

130. L'occupazione degli stati a $T > 0\text{K}$.

E' istruttivo ottenere una stima un po' piu' quantitativa di come la distribuzione degli elettroni sui livelli accessibili cambia con la temperatura.

A questo scopo possiamo ricavare l'intervallo di energia centrato sul potenziale chimico μ entro cui e' contenuto lo smussamento della distribuzione di Fermi-Dirac per una data temperatura T .

Basatevi sulla figura 42.

L'allargamento dello scalino puo' essere misurato dall'intervallo di energia individuato dalle intersezioni della retta tangente al flesso della funzione con la retta orizzontale $y = 1$ e con l'asse delle ascisse ($y = 0$).

Il flesso della funzione ha ascissa pari a μ e quindi la pendenza della retta tangente ad esso si ottiene dalla derivata della funzione calcolata per $\epsilon = \mu$.

Si ha:

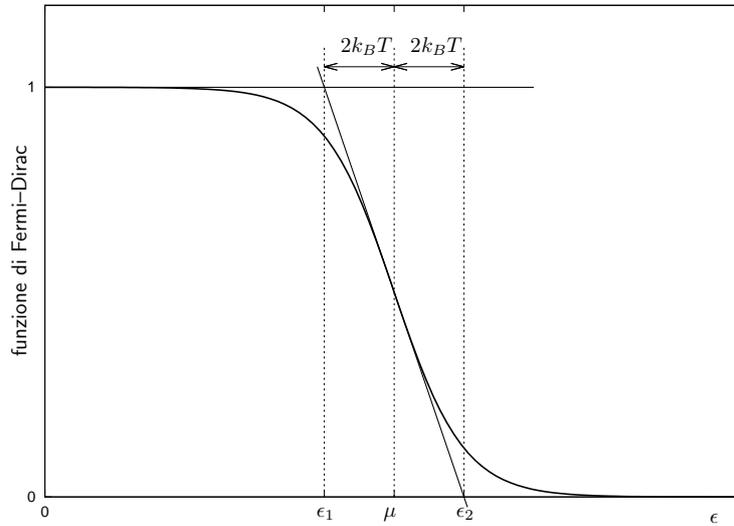


Figura 42: Una possibile misura dello “smussamento” che subisce la distribuzione di Fermi-Dirac a temperatura > 0 K: tracciate la tangente al flesso e individuate le intersezioni di tale tangente con le rette orizzontali $y = 1$ e $y = 0$. L’intervallo di energia $\epsilon_2 - \epsilon_1$ individuato da tali intersezioni sull’asse delle ascisse vi da’ un’idea dell’allargamento subito dalla curva.

$$\frac{df}{d\epsilon} = \left(-\frac{1}{\left(\exp\left(\frac{\epsilon - \mu}{k_B T}\right) + 1\right)^2} \right) \exp\left(\frac{\epsilon - \mu}{k_B T}\right) \left(\frac{1}{k_B T}\right)$$

$$\left. \frac{df}{d\epsilon} \right|_{\epsilon=\mu} = -\frac{1}{4} \frac{1}{k_B T} \quad (392)$$

L’equazione della retta con pendenza data dalla 392 e passante per il punto di coordinate $(\mu, 1/2)$ e’:

$$y(\epsilon) = -\frac{1}{4k_B T} (\epsilon - \mu) + \frac{1}{2} \quad (393)$$

Le ascisse ϵ_1 ed ϵ_2 delle intersezioni di questa retta con la retta orizzontale $y = 1$ e con l’asse delle ascisse sono, rispettivamente:

$$-\frac{1}{4k_B T} (\epsilon_1 - \mu) + \frac{1}{2} = 1 \Rightarrow \epsilon_1 = \mu - 2k_B T \quad (394)$$

$$-\frac{1}{4k_B T} (\epsilon_2 - \mu) + \frac{1}{2} = 0 \Rightarrow \epsilon_2 = \mu + 2k_B T \quad (395)$$

(come ci si poteva aspettare, la funzione di Fermi-Dirac e’ simmetrica rispetto a μ)

In conclusione:

$$\Delta\epsilon = \epsilon_2 - \epsilon_1 = 4k_B T \quad (396)$$

cioè: la deformazione che subisce lo scalino (perfetto a 0 K) in seguito all'aumento della temperatura da 0 K a T è contenuta in un intervallo di energia centrato attorno a μ e pari a $4k_B T$.

A temperatura ambiente ($T = 300$ K) il valore di $k_B T$ è:

$$k_B T = 1.38 \times 10^{-23} \frac{\text{J}}{\text{K}} \times 300 \text{ K} = 0.026 \text{ eV} \quad (397)$$

Questo valore è molto piccolo (ad esempio è solo lo 0.8 % dell'energia di Fermi del sodio, eq. 359).

Il risultato appena ottenuto ci dice quindi che in seguito ad un aumento di temperatura, per temperature ordinarie, solo una frazione molto piccola degli elettroni di un metallo è in grado di assorbire energia compiendo transizioni verso livelli che restano molto vicini all'energia corrispondente al potenziale chimico.

Questa è una diretta conseguenza del principio di Pauli: per energie di eccitazione piccole, solo gli elettroni che si trovano in stati "confinanti" con stati vuoti possono assorbire energia, mentre quelli degli stati a bassa energia non possono essere eccitati poiché dovrebbero compiere transizioni a stati che sono già occupati (e quindi indisponibili, per il principio di esclusione).

Un modo equivalente per rendersi conto di quanto piccola sia l'influenza della temperatura sulla distribuzione di Fermi-Dirac è il seguente.

Abbiamo detto che, per temperature non esagerate, l'energia dovuta all'agitazione termica, $k_B T$, è molto piccola rispetto all'energia di Fermi. Potremmo chiederci a quale temperatura l'energia dell'agitazione termica diventerebbe confrontabile con l'energia di Fermi. Evidentemente la temperatura cercata T_F deve soddisfare l'equazione:

$$k_B T_F = \epsilon_F \Rightarrow T_F = \frac{\epsilon_F}{k_B} \quad (398)$$

T_F è detta *temperatura di Fermi*. Come è facile prevedere sulla base di quanto abbiamo detto sopra, il valore di T_F è in genere molto elevato.

Ad esempio, per il sodio si ha:

$$T_F = \frac{\epsilon_F}{k_B} = \frac{3.14}{8.62 \times 10^{-5}} = 36427 \text{ K} \quad (399)$$

(8.62×10^{-5} è il valore della costante di Boltzmann in eV/K)

131. Un'utile interpretazione alternativa della distribuzione di Fermi-Dirac.

La distribuzione di Fermi-Dirac fornisce la probabilità che, a una certa temperatura, uno stato monoelettronico a energia ϵ sia occupato.

Proprio per questo, il suo significato è anche quello di *occupazione media* ovvero *numero medio di elettroni* presenti nel livello a energia ϵ (alla temperatura T).

La giustificazione di ciò è molto semplice. Per il principio di Pauli, uno stato monoelettronico a energia ϵ può contenere 1 oppure 0 elettroni.

Ora, immaginate di avere tutti i possibili stati N -elettronici accessibili al sistema in un sacchetto, come i numeri di una tombola (siccome gli elettroni si possono distribuire su un numero infinito di stati monoelettronici, gli stati N -elettronici

accessibili al sistema sono anch'essi infiniti); se eseguite $\mathcal{N} (\rightarrow \infty)$ estrazioni di uno stato N -elettronico dal sacchetto e indicate con \mathcal{N}_ϵ il numero di estrazioni in cui lo stato monoelettronico a energia ϵ e' occupato, allora il valore *medio* dell'occupazione dello stato monoelettronico a energia ϵ sara':

$$\begin{aligned} \left[\begin{array}{l} \text{occupazione media del} \\ \text{lo stato monoelettroni-} \\ \text{co a energia } \epsilon \end{array} \right] &= \lim_{\mathcal{N} \rightarrow \infty} \frac{\overbrace{1 + 1 + 1 + \dots + 1}^{\mathcal{N}_\epsilon \text{ termini}} + \overbrace{0 + 0 + 0 + \dots + 0}^{\mathcal{N} - \mathcal{N}_\epsilon \text{ termini}}}{\mathcal{N}} \\ &= \lim_{\mathcal{N} \rightarrow \infty} \frac{\mathcal{N}_\epsilon \times 1 + (\mathcal{N} - \mathcal{N}_\epsilon) \times 0}{\mathcal{N}} \\ &= \lim_{\mathcal{N} \rightarrow \infty} \frac{\mathcal{N}_\epsilon}{\mathcal{N}} = f(\epsilon) \end{aligned} \quad (400)$$

dove l'ultima uguaglianza deriva dall'interpretazione "frequentistica" della probabilita' come limite della frequenza relativa di un evento al tendere all'infinito del numero di esperimenti in cui quel dato evento puo' verificarsi.

Dall'interpretazione della distribuzione di Fermi-Dirac come numero medio di elettroni in un dato livello monoelettronico segue che essa deve soddisfare una *condizione di normalizzazione*.

Infatti, il numero totale di elettroni N di un campione deve essere dato semplicemente dalla somma su tutti gli stati monoelettronici del numero medio di elettroni contenuto in ciascuno stato, ovvero:

$$N = \sum_i f_i = \sum_i \frac{1}{\exp\left(\frac{\epsilon_i - \mu}{k_B T}\right) + 1} \quad (401)$$

dove ho reintrodotta l'indicizzazione dell'energia degli stati per ricordarci che i livelli elettronici del gas di Fermi sono discreti (anche se la loro spaziatura e' molto piccola: punto 119). La condizione di normalizzazione su scritta consente (almeno in linea di principio) di esprimere il potenziale chimico in funzione di N (o preferibilmente della densita' di elettroni $n = N/V$) e T .

132. Confronto fra la distribuzione di Fermi-Dirac e la distribuzione di Maxwell-Boltzmann.

Per ribadire e comprendere meglio l'importanza del principio di Pauli nella distribuzione di Fermi-Dirac e' interessante confrontare il suo andamento con quello della distribuzione di Maxwell-Boltzmann, che descrive un insieme di particelle classiche. La cosa essenziale, nel caso della seconda, e' che non c'e' alcun vincolo sul numero di particelle che possono occupare un dato (intendo *lo stesso*) livello energetico.

L'espressione per l'occupazione media di uno stato a energia ϵ di un sistema che obbedisce alla statistica di Maxwell-Boltzmann e':²⁵

$$b(\epsilon) = \exp\left(-\frac{\epsilon - \mu}{k_B T}\right) \quad (402)$$

Il confronto fra le due distribuzioni e' mostrato nella figura 43.

²⁵Non deriviamo la distribuzione di Maxwell-Boltzmann. Se vi interessa vedere come viene derivata, trovate il procedimento in qualsiasi libro di chimica fisica o termodinamica, ad esempio il libro di Schroeder.⁹

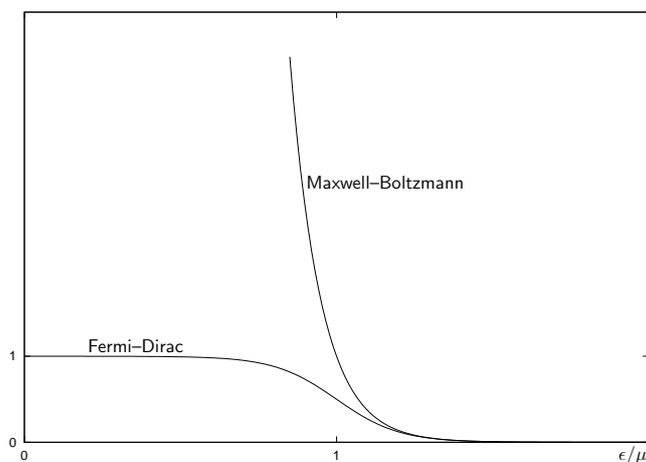


Figura 43: Confronto dell'andamento dell'occupazione media dei livelli di particella singola secondo la distribuzione di Fermi-Dirac e quella di Maxwell-Boltzmann, a parità del parametro $k_B T/\mu$.

In modo analogo alla figura 41, ho diagrammato anche la funzione di Maxwell-Boltzmann nella forma:

$$b\left(\frac{\epsilon}{\mu}, \frac{k_B T}{\mu}\right) = \exp\left(-\frac{\epsilon/\mu - 1}{k_B T/\mu}\right) \quad (403)$$

Notate come i due andamenti divergono al diminuire dell'energia (per una temperatura fissata): questo è dovuto al fatto che la distribuzione di Maxwell-Boltzmann non pone alcun vincolo sul numero massimo di particelle che possono occupare un dato livello e quindi tale numero può essere maggiore di 1; l'aumento esponenziale dell'occupazione media prevista dalla statistica di Maxwell-Boltzmann al diminuire dell'energia si spiega col fatto che, a qualsiasi temperatura, la probabilità di occupazione dei livelli a bassa energia è sempre maggiore di quella dei livelli a energia più alta e questo effetto viene esaltato via via che l'energia dei livelli è più bassa.

Notate ancora che alle alte energie, le due distribuzioni hanno lo stesso andamento. Questo è facile da giustificare in modo formale se riguardate l'espressione della distribuzione di Fermi-Dirac 388; quando:

$$\frac{\epsilon - \mu}{k_B T} \gg 1 \quad (404)$$

essa si riduce proprio alla forma 402 della distribuzione di Maxwell-Boltzmann.

133. Calcolo di proprietà legate agli stati occupati di un campione solido.

Molte proprietà dei solidi di interesse pratico sono legate all'occupazione degli stati monoelettronici. In generale, detto $Q(\vec{k})$ il valore che una proprietà assume per un singolo stato monoelettronico individuato dal vettore d'onda \vec{k} , il valore Q di quella proprietà per l'intero campione solido è ottenuto sommando su tutti gli stati il valore della proprietà in ciascuno stato moltiplicato per la sua occupazione (perché, ovviamente, solo gli stati occupati possono dare un

contributo). Inoltre, siccome il numero dei livelli occupati e' uguale al numero di elettroni del campione e questo e' direttamente proporzionale al volume del campione stesso, normalmente si calcola o misura il valore della proprieta' normalizzata per il volume V del campione:

$$Q = \frac{2}{V} \sum_{\vec{k}} Q(\vec{k}) \times [\text{occupazione dello stato monoelettronico } \vec{k}] \quad (405)$$

Come abbiamo visto piu' volte, il fattore 2 tiene conto del fatto che ad ogni punto \vec{k} permesso corrispondono due stati monoelettronici degeneri (in assenza di campi esterni, ad esempio campi magnetici, che rimuovano tale degenerazione). In base a quanto detto al punto precedente l'espressione diventa:

$$Q = \frac{2}{V} \sum_{\vec{k}} Q(\vec{k}) f(\epsilon(\vec{k})) \quad (406)$$

Notate che la somma 406 contiene un numero infinito di termini (tutti i punti \vec{k} permessi). Vedremo come sia possibile trasformarla in un oggetto piu' facilmente gestibile (punto 175).

La densita' degli stati

134. Normalmente, gli stati accessibili agli elettroni di un campione solido sono un insieme infinito (anche se numerabile). Anche gli stati elettronici accessibili agli elettroni di un atomo o una molecola sono infiniti, tuttavia il numero di elettroni effettivamente presenti in un campione solido e' dell'ordine del numero di Avogadro, cioe' enormemente maggiore di quello di un atomo o una molecola. Questo significa che gli elettroni e i livelli energetici da essi effettivamente occupati in un solido non possono essere trattati singolarmente (come si fa nello studio degli atomi o delle molecole).

135. Per questo motivo, una grandezza molto importante nell'analisi della struttura elettronica dei solidi e' la cosiddetta **Densita' Degli Stati** (DOS: Density Of States). Il suo significato e':

numero di stati elettronici permessi per unita' di energia

Siccome quasi sempre, come vedremo, il numero di stati elettronici permessi in un determinato intervallo di energia dipende (linearmente) anche dal volume del campione solido, la densita' degli stati viene normalmente riferita ad un volume di campione unitario e definita come:

numero di stati elettronici permessi per unita' di energia e per unita' di volume di campione solido

136. In generale, il numero di stati elettronici permessi non e' una funzione lineare dell'energia e quindi la densita' degli stati (numero di stati permessi per unita' di energia) *dipende* dall'energia; cioe' detta g la densita' degli stati ed ϵ l'energia, si avra':

$$g = g(\epsilon) \quad (407)$$

137. Ci sono almeno tre modi equivalenti di ricavare un'espressione matematica per la densita' degli stati.

Nel caso piu' semplice, se si conosce la funzione $N(\epsilon)$ dove N e' il numero *cumulativo* di stati permessi, cioe' il numero totale di stati permessi con energia $\leq \epsilon$, allora la densita' degli stati $g(\epsilon)$ e' semplicemente la derivata di $N(\epsilon)$ rispetto ad ϵ (normalizzata per il volume V del campione solido):

$$g(\epsilon) = \frac{1}{V} \frac{dN(\epsilon)}{d\epsilon} \quad (408)$$

138. Da questa espressione si ricava un altro modo utile di interpretare la densita' degli stati:

$g(\epsilon) d\epsilon$ e' il numero di stati permessi (dN) contenuti nel range di energia $[\epsilon, \epsilon + d\epsilon]$ (per unita' di volume del campione solido)

Dalla 408 segue che il numero di stati permessi contenuto nell'intervallo *finito* di energia compreso fra ϵ ed $\epsilon + \Delta\epsilon$ e' dato da:

$$N(\epsilon + \Delta\epsilon) - N(\epsilon) = V \int_{\epsilon}^{\epsilon + \Delta\epsilon} g(\epsilon) d\epsilon \quad (409)$$

139. Notate come il concetto di densita' degli stati sia identico sotto molti aspetti a quello, ad esempio, di velocita' spaziale: detta $s(t)$ la funzione che lega al tempo lo spazio *cumulativo* percorso (cioe' lo spazio totale percorso a partire dall'istante iniziale), la velocita' spaziale $v(t)$ (nel caso monodimensionale) e' definita come:

$$v(t) = \frac{ds(t)}{dt} \quad (410)$$

La velocita' spaziale puo' essere pensata come una "densita' dello spazio" (rispetto al tempo): grande velocita' vuol dire "tanto spazio in poco tempo", cioe' grande "densita' di spazio"; piccola velocita' vuol dire "poco spazio in tanto tempo", cioe' piccola "densita' di spazio".

Identicamente alla densita' degli stati, $v(t) dt$ e' lo spazio (ds) percorso nell'intervallo di tempo $[t, t + dt]$ e lo spazio percorso in un intervallo di tempo finito, $[t, t + \Delta t]$ si ottiene con l'integrale:

$$s(t + \Delta t) - s(t) = \int_t^{t + \Delta t} v(t) dt \quad (411)$$

140. Essendo $N(\epsilon)$ il numero *cumulativo* di stati permessi, il suo andamento non puo' essere decrescente: all'aumentare di ϵ $N(\epsilon)$ puo' solo crescere o rimanere costante: questo e' ragionevole, visto che $g(\epsilon)$, essendo una densita', deve essere sempre ≥ 0 . Questa e' una (ovvia) diversita' rispetto all'analogia con la velocita', dove l'oggetto in movimento puo' "tornare indietro", determinando una "densita'" (ie: velocita') negativa.

141. Una differenza importante fra la densita' degli stati e la velocita' spaziale (intesa come densita' dello spazio rispetto al tempo) e' che, mentre il tempo e' una variabile continua, le energie permesse in un sistema quantomeccanico sono, in generale, quantizzate, cioe' *discrete*. Ad ogni energia permessa corrispondono uno o piu' stati distinti (il numero di stati corrispondenti a un particolare livello

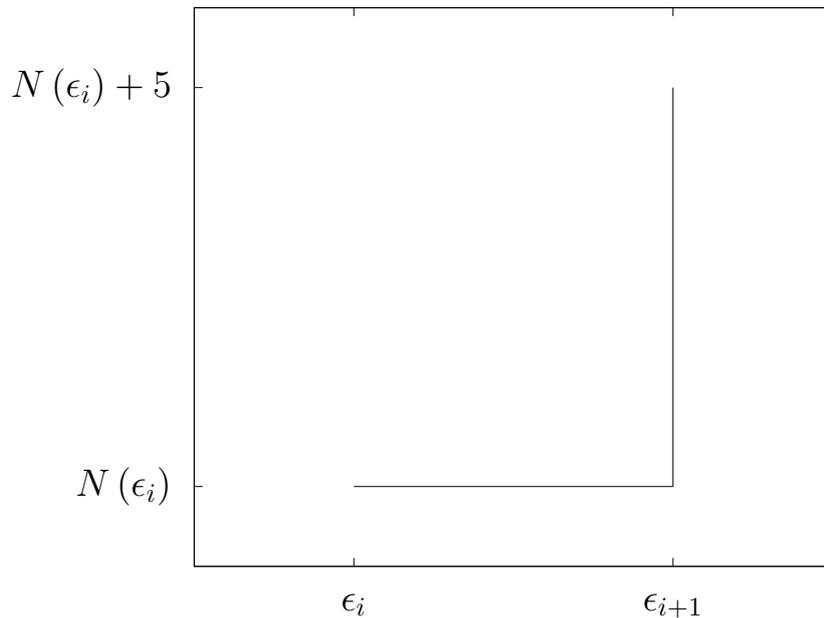


Figura 44:

energetico $e' > 1$ se il livello energetico e' degenerare); da cio' segue che il grafico di $N(\epsilon)$ in funzione di ϵ e' fatto a scalini. Infatti, supponete che ϵ_i ed ϵ_{i+1} siano due valori permessi di energia successivi (fate riferimento alla figura 44). Tutte le energie comprese fra ϵ_i ed ϵ_{i+1} sono proibite e quindi ad esse non corrisponde alcuno stato permesso. $N(\epsilon_i)$ e' il numero totale di stati permessi con energia $\leq \epsilon_i$. Supponete che per $\epsilon = \epsilon_{i+1}$ il sistema abbia 5 stati permessi: come sara' la parte del grafico $N(\epsilon)$ vs ϵ nell'intervallo $[\epsilon_i, \epsilon_{i+1}]$? $N(\epsilon)$ vale $N(\epsilon_i)$ per $\epsilon = \epsilon_i$, poi *rimane costante* per $\epsilon_i < \epsilon < \epsilon_{i+1}$ e *cresce verticalmente* di 5 per $\epsilon = \epsilon_{i+1}$: Cioe': il grafico presenta un gradino per $\epsilon = \epsilon_{i+1}$. Ovviamente lo stesso ragionamento vale per qualsiasi coppia di energie successive e quindi si conclude che il grafico di $N(\epsilon)$ presenta dei salti discontinui ad ogni valore di energia permesso. In linea di principio questo pone un problema nella valutazione della 408, che diventa infinita in corrispondenza ad ogni gradino. In realta', per tutti i casi di interesse, la spaziatura dei valori di energia permessi e' molto piccola e quindi si puo' assumere che le energie permesse siano continue; inoltre, sempre nei casi di interesse, $N(\epsilon)$ assume subito valori molto grandi e quindi i tratti verticali in corrispondenza a ciascun gradino diventano infinitesimi rispetto al valore di $N(\epsilon)$.

Per questi motivi, l'espressione 408 rimane valida e operativa se per $N(\epsilon)$ si usa una sua approssimazione continua (cioe' *non* a gradini).

142. Ad esempio, la figura 45 mostra l'andamento di $N(\epsilon)$ e della sua approssimazione continua (che ricaveremo fra un attimo) per i primi 15 livelli energetici di un gas di elettroni liberi e indipendenti in un campione solido di forma cubica. Il prolungamento dei tratti verticali del grafico a scalini individua sull'asse delle ascisse i valori di energia permessi; l'altezza dei gradini rappresenta il numero di stati (la degenerazione) corrispondenti a ciascun livello energetico: gradino piu'

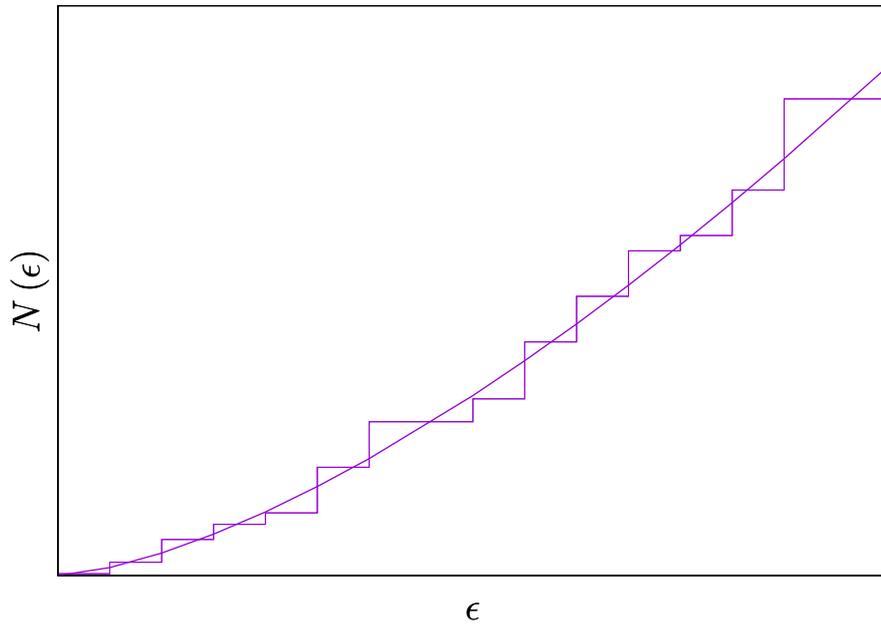


Figura 45:

alto significa maggiore degenerazione, cioè maggior numero di stati distinti, ma alla stessa energia.

143. Vediamo ora come si può ricavare la densità degli stati applicando la 408 ad un sistema costituito da un gas di elettroni liberi di Fermi confinato in un volume cubico.

Dobbiamo ricavare $N(\epsilon)$, cioè il numero cumulativo di stati permessi per energie $\leq \epsilon$ e farne la derivata rispetto ad ϵ .

Come abbiamo visto (punto 118), per questo sistema le superfici isoenergetiche nello spazio \vec{k} sono delle sfere. Allora, il numero totale di stati \vec{k} permessi per energie $\leq \epsilon$ può essere ricavato nel modo visto al punto 204:

$$N(\epsilon) = \frac{[\text{volume della sfera corrispondente all'energia } \epsilon]}{[\text{volume corrispondente a un singolo punto } \vec{k} \text{ permesso}]} = \frac{\frac{4}{3}\pi k_\epsilon^3}{\left(\frac{2\pi}{L}\right)^3} \quad (412)$$

dove k_ϵ è il modulo del vettore \vec{k} tale che l'energia ad esso corrispondente è ϵ e il denominatore è il volume della cella primitiva del reticolo reciproco.

Dalla relazione di dispersione per questo sistema (eq. 344):

$$\epsilon = \frac{\hbar^2 k_\epsilon^2}{2m}$$

possiamo ricavare k_ϵ^3 :

$$k_\epsilon^3 = (k_\epsilon^2)^{\frac{3}{2}} = \left(\frac{2m}{\hbar^2} \epsilon\right)^{\frac{3}{2}} = \left(\frac{2m}{\hbar^2}\right)^{\frac{3}{2}} \epsilon^{\frac{3}{2}} \quad (413)$$

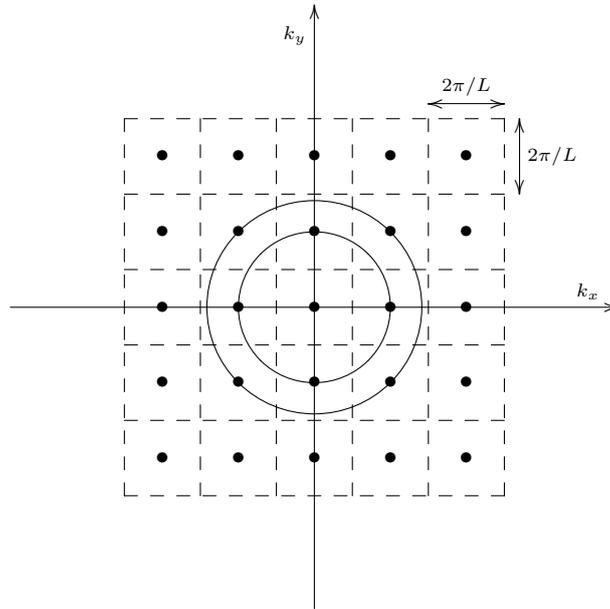
Sostituendo sopra si ottiene:

$$N(\epsilon) = \frac{\frac{4}{3}\pi \left(\frac{2m}{\hbar^2}\right)^{\frac{3}{2}} \epsilon^{\frac{3}{2}}}{\left(\frac{2\pi}{L}\right)^3} = \frac{V}{6\pi^2} \left(\frac{2m}{\hbar^2}\right)^{\frac{3}{2}} \epsilon^{\frac{3}{2}} \quad (414)$$

dove $V = L^3$ e' il volume del campione solido.

144. Come avevamo accennato a p. 109, l'espressione ottenuta e' un'approssimazione *continua* a una grandezza, $N(\epsilon)$, che e' intrinsecamente *discontinua*. L'approssimazione nasce dal modo in cui abbiamo ricavato $N(\epsilon)$ dividendo il volume della sfera isoenergetica a energia ϵ per il volume *cubico* della cella corrispondente a un punto \vec{k} permesso: dovrete realizzare facilmente che il risultato di questa divisione non sara', in generale, un numero intero (come dovrebbe essere in corrispondenza a valori permessi di ϵ) e cio' sostanzialmente a causa del fatto che il volume di una sfera puo' essere ottenuto *esattamente* sommando il volume di n cubi solo per poche combinazioni del raggio della sfera, dello spigolo dei cubi e di n .

145. Potete rendervi conto di cio' in modo piu' concreto facendo riferimento alla figura di seguito, che si riferisce a un sistema bidimensionale in cui il reticolo dei punti \vec{k} permessi e' contenuto nel piano e la cella unitaria contenente un singolo punto \vec{k} e' un quadrato di lato $2\pi/L$. Naturalmente in questo caso le "superfici" isoenergetiche sono delle circonferenze centrate nell'origine.



Innanzitutto, per $\epsilon = 0$, il numero di stati permessi e' 1 (lo stato $(k_x, k_y) = (0, 0)$), mentre il risultato che fornisce la funzione continua (eq. 414) e' 0.

Per l'energia corrispondente all'isocirconferenza piu' interna il numero cumulativo di punti \vec{k} permessi e' 5 (contateli), ma il rapporto fra l'area di tale circonferenza e l'area della cella unitaria corrispondente a un singolo punto \vec{k} non e' un numero intero ed e' *minore* di 5, come potete facilmente rendervi conto confrontando visivamente l'area occupata dalle 5 celle primitive dei punti di coordinate (espresse in unita' $2\pi/L$) $(0, 0)$, $(1, 0)$, $(0, 1)$, $(-1, 0)$ e $(0, -1)$ con l'area dell'isocirconferenza.

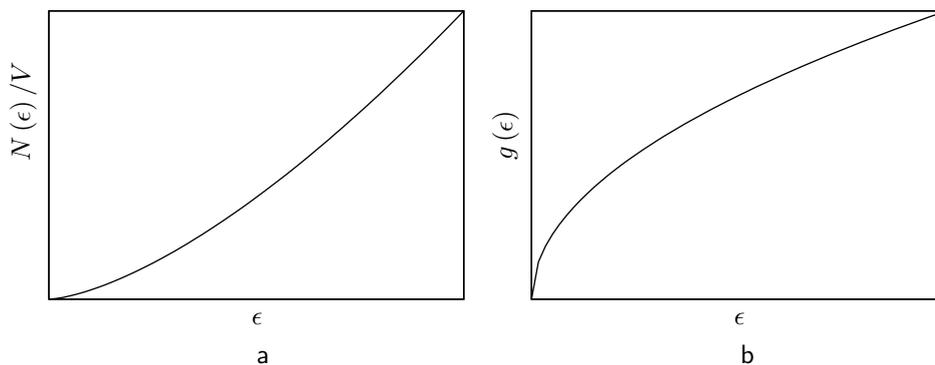


Figura 46: Grafico del numero cumulativo degli stati permessi (diviso per il volume del campione) (a) e della densita' degli stati (b) per un sistema costituito da un gas di Fermi confinato in un volume cubico. Il grafico in b e' la derivata di quello in a.

Lo stesso vale anche per l'energia corrispondente all'isocirconferenza piu' esterna, che contiene 9 punti \vec{k} permessi (anche in questo caso il risultato che si ottiene con l'equazione 414 e' minore di 9).

146. Nonostante sia solo un'approssimazione, l'eq. 414 funziona molto bene quasi sempre. Inoltre essa diventa sempre piu' accurata all'aumentare dell'energia (i cubi sono sempre piu' piccoli rispetto all'isosfera e l'errore diminuisce di conseguenza).

147. Avendo ricavato il numero cumulativo di stati permessi, la densita' degli stati $g(\epsilon)$ si ottiene semplicemente facendo la derivata e dividendo per il volume del solido V (eq. 408):

$$g(\epsilon) = \frac{1}{V} \frac{d}{d\epsilon} N(\epsilon) = \frac{1}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{\frac{3}{2}} \epsilon^{\frac{1}{2}} \quad (415)$$

L'andamento delle funzioni $N(\epsilon)/V$ e $g(\epsilon)$ e' mostrato nella figura 46.

148. Una prima considerazione sul risultato ottenuto e' che la densita' degli stati per un campione cubico contenente un gas di elettroni di Fermi e' una funzione crescente dell'energia. Questo significa che, dato un intervallo infinitesimo di energia $d\epsilon$, il numero di stati permessi $g(\epsilon)d\epsilon$ con energia compresa fra ϵ e $\epsilon + d\epsilon$ e' maggiore ad energia maggiore.

149. Un'altra cosa da notare e' che quella ottenuta e' la densita' dei punti \vec{k} permessi. Ogni punto \vec{k} permesso individua un corrispondente stato che *non dipende dallo spin elettronico*. Quindi, per il principio di Pauli, ad ogni punto \vec{k} permesso corrispondono 2 stati monoelettronici che differiscono per lo spin. In altre parole, la densita' *degli stati monoelettronici* (per unita' di volume del campione) e' il doppio della densita' dei punti \vec{k} permessi:

$$\frac{1}{2\pi^2} \left(\frac{2m}{\hbar^2} \right)^{\frac{3}{2}} \epsilon^{\frac{1}{2}} \quad (416)$$

150. Una terza considerazione e' che non dovete fare confusione fra la densita' degli stati permessi (espressa in termini di punti \vec{k} o in termini di stati

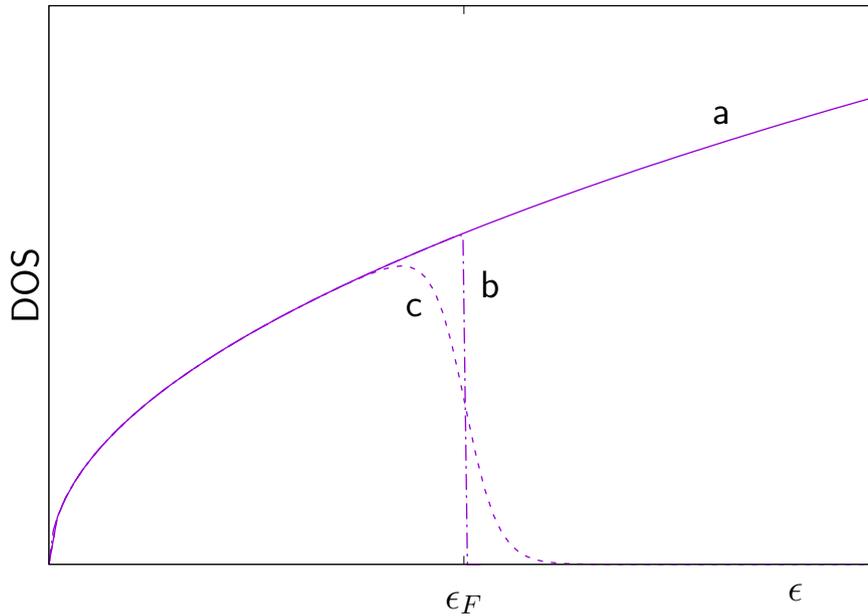


Figura 47: a: densita' degli stati accessibili; b: densita' degli stati occupati a $T = 0$ K; c: densita' degli stati occupati a $T > 0$ K.

monoelettronici: c'è un fattore 2 di differenza) con la densita' degli stati *occupati*. Cioè: un conto sono gli stati accessibili (permessi), e un conto sono gli stati effettivamente occupati dagli elettroni. Questi ultimi, e la loro corrispondente densita' rispetto all'energia, dipendono prima di tutto dalla temperatura del campione. A 0 K solo gli stati a più bassa energia sono occupati fino al raggiungimento del numero di elettroni del campione (e dell'energia di Fermi); a temperature > 0 una frazione degli elettroni occupa stati a maggiore energia lasciando “vuoti” gli stati che occupava a 0 K, secondo la distribuzione di Fermi-Dirac (eq. 388). Conseguentemente, la densita' degli stati accessibili è indipendente dalla temperatura, mentre quella degli stati occupati dipende dalla temperatura.

La densita' degli stati occupati si ottiene moltiplicando la densita' degli stati accessibili per la funzione di Fermi-Dirac. Ad ogni energia, la funzione di Fermi-Dirac fornisce l'occupazione media di uno stato monoelettronico avente quell'energia e quindi, moltiplicando la densita' degli stati accessibili per la funzione di Fermi-Dirac, in pratica ogni stato viene conteggiato sulla base della sua occupazione: uno stato completamente occupato conterà per 1, uno parzialmente occupato conterà per una frazione < 1 e uno non occupato (o con un'occupazione media $\rightarrow 0$) non verrà conteggiato (o quasi). A $T = 0$ K, la densita' degli stati occupati segue la 415 fino all'energia di Fermi e poi diventa nulla con un gradino verticale per energie maggiori; a temperature > 0 K l'andamento sarà smussato riflettendo il corrispondente smussamento della funzione di Fermi-Dirac. La densita' degli stati accessibili è confrontata con quella degli stati occupati, a $T = 0$ e $T > 0$, nella figura 47.

151. Un'altra espressione (equivalente) per la densita' degli stati, quando,

come spesso succede, non si dispone di un'espressione analitica per $N(\epsilon)$, puo' essere ottenuta nel modo seguente.

Dalla definizione 408 abbiamo:

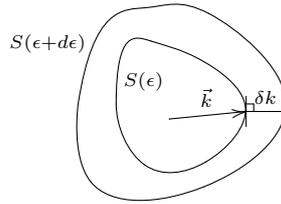
$$g(\epsilon) d\epsilon = \frac{1}{V} dN(\epsilon) \quad (417)$$

$dN(\epsilon)$ e' il numero di stati permessi nel range di energia compreso fra ϵ e $\epsilon + d\epsilon$. Come ora vedremo, e' possibile ricavare un'espressione piuttosto generale per $dN(\epsilon)$: uguagliando tale espressione a $g(\epsilon) d\epsilon$ sara' poi possibile ottenere la corrispondente espressione per $g(\epsilon)$.

152. Abbiamo visto che per un gas di elettroni di Fermi in un campione cubico ci sono delle superfici sferiche nello spazio \vec{k} tali che tutti i punti su di esse individuano stati elettronici alla stessa energia (degeneri).

In generale, per un campione qualunque (non cubico) e, come vedremo, anche per il caso in cui gli elettroni non siano (completamente) liberi (cioe' risentano di un potenziale dovuto alla presenza dei nuclei), ci sono ancora delle superfici isoenergetiche nello spazio \vec{k} , ma questa volta non sono delle semplici sfere, bensì delle superfici qualunque.

153. Consideriamo allora due di queste superfici isoenergetiche nello spazio \vec{k} a energie ϵ ed $\epsilon + d\epsilon$ (nella figura di seguito e' mostrata la loro proiezione sul piano):



In modo simile a quanto abbiamo fatto per ricavare la forma analitica di $N(\epsilon)$ per il gas di elettroni di Fermi in un campione cubico, detto δV il volume di spazio \vec{k} compreso fra le due superfici isoenergetiche, dovra' essere:

$$dN(\epsilon) = \frac{\delta V}{\left[\text{volume di una cella primitiva contenente un singolo punto } \vec{k} \right]} \quad (418)$$

Come abbiamo visto in generale (eq. 244) e anche al punto 204, il volume della cella primitiva dello spazio \vec{k} e' $8\pi^3/V$; quindi l'espressione per $dN(\epsilon)$ puo' essere riscritta come:

$$dN(\epsilon) = \frac{\delta V}{\frac{8\pi^3}{V}} = \frac{V}{8\pi^3} \delta V \quad (419)$$

Vediamo ora come possiamo esprimere δV , il volume compreso fra le due superfici isoenergetiche.

Detto $d\vec{k}$ un elemento infinitesimo di volume dello spazio \vec{k} , dovra' chiaramente essere:

$$\delta V = \int_{\delta V} d\vec{k} \quad (420)$$

Ora: prendiamo il volume infinitesimo $d\vec{k}$ usato per l'integrazione come un "cilindretto" infinitesimo avente per base un elemento infinitesimo dS della superficie interna $S(\epsilon)$ e per altezza la distanza δk , misurata perpendicolarmente rispetto alla superficie $S(\epsilon)$, fra le due superfici isoenergetiche (guardate la figura):

$$d\vec{k} = \delta k dS \quad (421)$$

δk e' normale in ogni punto alla superficie $S(\epsilon)$ e dipende da \vec{k} , cioe' dal particolare punto di $S(\epsilon)$ in cui ci troviamo (perche' le due superfici sono di forma qualunque).

Quindi, a questo punto, l'integrale calcolato sul volume δV puo' essere sostituito da un integrale calcolato sulla superficie $S(\epsilon)$ in cui il differenziale per l'integrazione e' l'elemento infinitesimo di superficie dS :

$$\delta V = \int_{S(\epsilon)} \delta k_{\vec{k}} dS \quad (422)$$

Ora esprimiamo $\delta k_{\vec{k}}$ in funzione di $d\epsilon$ nel modo seguente.

Sappiamo che, in termini del tutto generali, l'energia degli stati di un sistema solido e' funzione della posizione nello spazio \vec{k} , cioe':

$$\epsilon = \epsilon(\vec{k}) = \epsilon(k_x, k_y, k_z) \quad (423)$$

ϵ e' una funzione di tre variabili.

Consideriamo il gradiente di tale funzione: $\vec{\nabla}\epsilon(\vec{k})$.

154. Aperta parentesi: breve richiamo sul gradiente di una funzione.

Per una funzione scalare di tre coordinate $f(\vec{r}) = f(x, y, z)$ il gradiente e' il vettore le cui componenti sono le derivate parziali della funzione rispetto alle tre coordinate:

$$\vec{\nabla}f = \hat{e}_x \frac{\partial f}{\partial x} + \hat{e}_y \frac{\partial f}{\partial y} + \hat{e}_z \frac{\partial f}{\partial z} \quad (424)$$

dove \hat{e}_x , \hat{e}_y e \hat{e}_z sono i versori (vettori di modulo unitario) nella direzione dei tre assi cartesiani.

Il prodotto scalare del gradiente della funzione $f(x, y, z)$ e di un versore \vec{u} in una direzione arbitraria fornisce la *derivata direzionale* della funzione, cioe' la variazione della funzione *per unita' di spostamento nella direzione di \vec{u}* :

$$D_{f, \vec{u}} = \vec{\nabla}f \cdot \vec{u} \quad (425)$$

Alcune proprieta' del gradiente di cui faremo uso sono:

1. Il gradiente della funzione $f(x, y, z)$ si annulla in un punto di massimo o minimo. Questo segue banalmente dal fatto che in un punto di massimo o minimo le derivate parziali della f sono tutte nulle.
2. In un certo punto dello spazio (x, y, z) , il gradiente della funzione $f(x, y, z)$ e' orientato *nella direzione di massimo incremento* della funzione e il suo modulo, $|\vec{\nabla}f|$, e' uguale *alla variazione della funzione per unita' di spostamento* (nel punto (x, y, z) e nella direzione del gradiente). Cio' si puo'

vedere in modo semplice. La derivata direzionale nella direzione di un arbitrario versore \vec{u} e':

$$D_{f,\vec{u}} = \vec{\nabla} f \cdot \vec{u} = \left| \vec{\nabla} f \right| \cos \theta \quad (\text{il modulo di } \vec{u} \text{ e' unitario}) \quad (426)$$

La derivata direzionale (che e' la variazione della funzione per unita' di spostamento nella direzione di \vec{u}) assume il suo valore massimo positivo quando $\theta = 0$, cioe' quando \vec{u} e' *nella stessa direzione del gradiente*; ma questo equivale a dire che il gradiente e' nella direzione del massimo incremento della funzione f . Inoltre, in queste condizioni, si ha:

$$D_{f,\vec{u}} = \left| \vec{\nabla} f \right| \quad (427)$$

cioe' il modulo del gradiente e' uguale *alla variazione della funzione per unita' di spostamento* (nel punto (x, y, z) e nella direzione del gradiente)

3. Il gradiente della funzione $f(x, y, z)$ e' normale in ogni punto ad una sua qualsiasi isosuperficie. Anche questa proprieta' e' giustificata facilmente.

Consideriamo un punto arbitrario P della isosuperficie e un versore \vec{v} ad essa parallelo in P (cioe' tangente alla isosuperficie nel punto P). Per la derivata direzionale della funzione nel punto P rispetto al versore \vec{v} deve valere:

$$D_{f,\vec{v}} = \vec{\nabla} f \cdot \vec{v} \quad (428)$$

Ma siccome \vec{v} e' parallelo alla isosuperficie e il valore della funzione e' ovviamente costante su tutti i punti di essa, dovra' essere:

$$D_{f,\vec{v}} = 0 = \vec{\nabla} f \cdot \vec{v} \quad (429)$$

ovvero, $\vec{\nabla} f$ deve essere normale a \vec{v} , cioe' normale alla isosuperficie. E siccome P era scelto arbitrariamente, questo deve valere per qualsiasi punto della isosuperficie.

Chiusa parentesi.

155. Tornando alle due superfici isoenergetiche e al gradiente della funzione $\epsilon(\vec{k})$, alla luce di quanto ricordato sopra in generale riguardo al gradiente di una funzione, possiamo dire che:

- $\implies \vec{\nabla} \epsilon(\vec{k})$ e' normale a $S(\epsilon)$ in ogni suo punto
- \implies la variazione $d\epsilon$ della funzione $\epsilon(\vec{k})$ nel passare dalla isosuperficie $S(\epsilon)$ a quella $S(\epsilon + d\epsilon)$ nella direzione del gradiente (che coincide col cammino lungo δk nella figura precedente) deve essere data da:

$$d\epsilon = \left| \vec{\nabla} \epsilon(\vec{k}) \right| \delta k_{\vec{k}} \quad (430)$$

cioe':

$$\overbrace{\left[\text{variaz. di } \epsilon(\vec{k}) \right]}^{d\epsilon} = \overbrace{\left[\text{variaz. di } \epsilon(\vec{k}) \text{ per unita' di spostamento} \right]}^{|\vec{\nabla}\epsilon(\vec{k})|} \times \underbrace{\left[\text{spostamento} \right]}_{\delta k_{\vec{k}}}$$

156. Notate che $d\epsilon$, essendo la variazione della funzione $\epsilon(\vec{k})$ nel passaggio da $S(\epsilon)$ a $S(\epsilon + d\epsilon)$, e' indipendente da \vec{k} , cioe' dal particolare punto considerato sulla isosuperficie $S(\epsilon)$ (quando si passa da un *qualsiasi* punto di $S(\epsilon)$ a un *qualsiasi* punto di $S(\epsilon + d\epsilon)$ la variazione subita dalla funzione $\epsilon(\vec{k})$ e' sempre la stessa, $d\epsilon$, per definizione di isosuperficie!). Al contrario, sia $|\vec{\nabla}\epsilon(\vec{k})|$ che $\delta k_{\vec{k}}$ dipendono dal punto di $S(\epsilon)$ considerato, cioe' da \vec{k} : infatti, essendo le due isosuperfici di forma qualunque, $|\vec{\nabla}\epsilon(\vec{k})|$ e $\delta k_{\vec{k}}$ devono variare in ogni punto di $S(\epsilon)$ proprio per fare in modo che il loro prodotto ($d\epsilon$) resti costante.

157. A questo punto possiamo ricavare $\delta k_{\vec{k}}$ dalla 430 e sostituire nell'integrale 422, facendo uso del fatto appena rimarcato sopra che $d\epsilon$ non dipende dal punto considerato sulla isosuperficie $S(\epsilon)$:

$$\delta V = \int_{S(\epsilon)} \frac{d\epsilon}{|\vec{\nabla}\epsilon(\vec{k})|} dS = d\epsilon \int_{S(\epsilon)} \frac{1}{|\vec{\nabla}\epsilon(\vec{k})|} dS \quad (431)$$

158. Infine, sostituendo δV nella 419 e riarrangiando otteniamo il risultato finale:

$$dN(\epsilon) = \frac{V}{8\pi^3} d\epsilon \int_{S(\epsilon)} \frac{1}{|\vec{\nabla}\epsilon(\vec{k})|} dS \quad (432)$$

$$g(\epsilon) = \frac{1}{V} \frac{dN(\epsilon)}{d\epsilon} = \frac{1}{8\pi^3} \int_{S(\epsilon)} \frac{1}{|\vec{\nabla}\epsilon(\vec{k})|} dS \quad (433)$$

159. Osservazioni.

- Notate che la dipendenza di $g(\epsilon)$ da ϵ nell'eq. 433 sta nell'isovalore ϵ della superficie su cui viene calcolato l'integrale: cioe', la "ricetta" per calcolare la densita' degli stati g al valore di energia ϵ e': "(i) trova la isosuperficie il cui valore di energia costante e' ϵ ; (ii) calcola l'integrale dell'eq. 433 su tale superficie; (iii) il valore ottenuto dell'integrale e' $g(\epsilon)$ ".
- Come gia' detto, quella ottenuta e' la densita' di punti \vec{k} permessi; ad ogni punto \vec{k} corrispondono 2 stati monoelettronici che differiscono per lo spin; quindi la densita' degli stati monoelettronici e' l'espressione scritta sopra, ma con un 4 al posto dell'8 (ad esempio l'eq. 8.63 del libro di Mermin e Ashcroft¹¹).
- Il risultato ottenuto mette in relazione la densita' degli stati con il gradiente (ovvero la rapidita' di variazione) della loro energia: la densita' degli stati e' alta nelle regioni dello spazio \vec{k} dove l'energia degli stati permessi varia lentamente con \vec{k} , e viceversa.

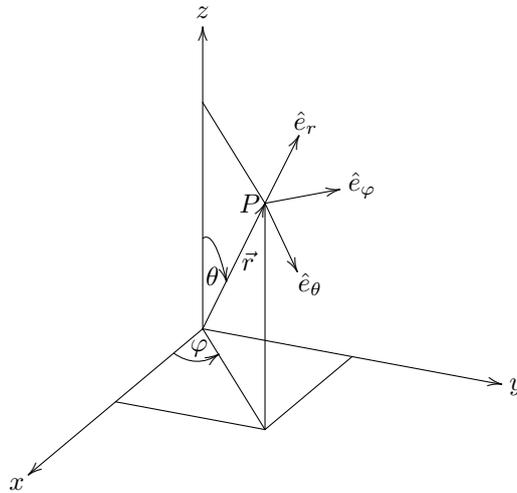


Figura 48: Coordinate sferiche.

- E' possibile che l'energia degli stati permessi attraversi punti di massimo o minimo nello spazio \vec{k} : in corrispondenza di tali punti il gradiente di $\epsilon(\vec{k})$ diventa nullo e il valore dell'argomento dell'integrale nella 433 diventa infinito (si puo' comunque dimostrare che la presenza di un numero finito di questi punti singolari non pregiudica l'esistenza, in senso matematico, dell'integrale). Questo effetto e' osservabile come la comparsa di cuspidi (con una tangente verticale) nella densita' degli stati. Tali cuspidi possono essere osservate sperimentalmente e vengono dette *singularita' di Van Hove*.

160. Possiamo verificare che l'eq. 433 fornisce la stessa espressione per la densita' degli stati che avevamo gia' trovato per il gas di elettroni di Fermi confinato in un campione di forma cubica (equazione 415).

In pratica, si tratta di risolvere l'integrale di superficie per questo particolare sistema.

Sappiamo gia' che in questo caso le superfici isoenergetiche sono delle sfere concentriche centrate nell'origine dello spazio \vec{k} :

$$\epsilon(\vec{k}) = \frac{\hbar^2}{2m} k^2$$

Data la simmetria delle superfici isoenergetiche, conviene usare coordinate sferiche.

161. Aperta parentesi: richiamo sulle coordinate sferiche.

Le coordinate sferiche di un punto nello spazio P sono definite come (r, θ, φ) , dove (guardate la figura 48):

- r e' il modulo del vettore posizione \vec{r} del punto P . Quindi: $0 \leq r \leq +\infty$
- θ e' l'angolo fra il vettore \vec{r} e l'asse z positivo; per definizione: $0 \leq \theta \leq 180$. Questo angolo viene anche detto "angolo di zenith" o "angolo polare".
- φ e' l'angolo fra la proiezione del vettore \vec{r} sul piano xy e l'asse x positivo, misurato in senso antiorario. Per definizione: $0 \leq \varphi \leq 360$. Questo angolo viene anche detto "angolo di azimuth" o "angolo equatoriale".

Le relazioni per la trasformazione fra coordinate cartesiane e sferiche sono (rendetevene conto guardando la figura):

$$x = r \sin \theta \cos \varphi \quad (434)$$

$$y = r \sin \theta \sin \varphi \quad (435)$$

$$z = r \cos \theta \quad (436)$$

$$r = \sqrt{x^2 + y^2 + z^2} \quad (437)$$

$$\theta = \begin{cases} \arctan \frac{\sqrt{x^2+y^2}}{z} & \text{se } z > 0 \\ 180 - \arctan \frac{\sqrt{x^2+y^2}}{|z|} & \text{se } z < 0 \end{cases} \quad (438)$$

$$\varphi = \begin{cases} \arctan \frac{y}{x} & \text{se } x, y > 0 \\ 180 - \arctan \frac{y}{x} & \text{se } x < 0, y > 0 \\ 180 + \arctan \frac{y}{x} & \text{se } x < 0, y < 0 \\ 360 - \arctan \frac{y}{x} & \text{se } x > 0, y < 0 \end{cases} \quad (439)$$

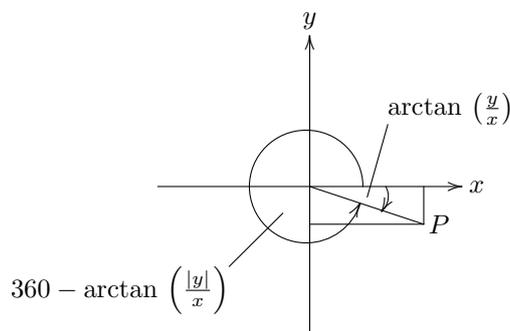
Notate come nelle espressioni per θ e φ si debbano distinguere diversi casi a seconda del segno di x e/o y e/o z ; cio' e' dovuto al fatto che, per definizione[‡], $0 \leq \theta \leq 180$ e $0 \leq \varphi \leq 360$.

Ad esempio, guardando la proiezione sul piano xy (figura in basso), se un punto P ha $x > 0$ e $y < 0$, allora per esso si avrebbe:

$$\varphi = \arctan \left(\frac{y}{x} \right) < 0 \quad (440)$$

cioe' $\varphi \notin [0, 360]$. Per avere il valore dell'angolo misurato in senso antiorario rispetto all'asse x positivo e interno all'intervallo $[0, 360]$, bisogna esprimerlo come:

$$\varphi = 360 - \left| \arctan \left(\frac{y}{x} \right) \right| = 360 - \arctan \left(\frac{|y|}{x} \right)$$



[‡]In realta' esistono molte convenzioni e definizioni diverse per le coordinate sferiche, e in particolare per la misura e l'intervallo di variazione degli angoli θ e φ

I versori di un sistema di coordinate sferiche sono \hat{e}_r , \hat{e}_θ and \hat{e}_φ (guardate la figura all'inizio). Come i versori cartesiani, essi sono una terna di vettori ortonormali (cioè che hanno modulo unitario e sono mutuamente ortogonali) che costituiscono la base (nel senso dell'algebra lineare) rispetto a cui qualsiasi altro elemento (vettore) di \mathbb{R}^3 può essere espresso. La definizione dei tre versori è:

- \hat{e}_r e' nella direzione di \vec{r} , orientato nel verso di un allontanamento dall'origine
- \hat{e}_θ e' normale a \hat{e}_r , giace nel piano definito da \vec{r} e dall'asse z ed e' orientato nella direzione in cui θ aumenta (θ e' misurato rispetto all'asse z positivo)
- \hat{e}_φ e' normale a \hat{e}_r e \hat{e}_θ ed e' orientato nella direzione in cui φ (che e' misurato in senso antiorario rispetto all'asse x positivo) aumenta

Notate che, diversamente da quanto accade per un sistema cartesiano, \hat{e}_r , \hat{e}_θ and \hat{e}_φ sono funzioni di \vec{r} , θ e φ : la loro orientazione (ma non il loro modulo) cambia con \vec{r} , θ and φ . Cio' e' ragionevole: i tre versori devono essere sempre ortogonali; quindi, se il punto dello spazio considerato cambia, anche il suo vettore posizione \vec{r} cambia e cio', a sua volta, fa cambiare l'orientazione di \hat{e}_r ; ma allora \hat{e}_θ and \hat{e}_φ devono cambiare orientazione anch'essi per mantenere l'ortogonalita' reciproca. L'espressione del gradiente di una funzione delle tre coordinate sferiche $f(r, \theta, \varphi)$ e' [1]†:

$$\vec{\nabla} f = \hat{e}_r \frac{\partial f}{\partial r} + \hat{e}_\theta \frac{1}{r} \frac{\partial f}{\partial \theta} + \hat{e}_\varphi \frac{1}{r \sin \theta} \frac{\partial f}{\partial \varphi} \quad (441)$$

Chiusa parentesi.

162. Tornando alla relazione di dispersione:

$$\epsilon(\vec{k}) = \frac{\hbar^2}{2m} k^2$$

dovreste rendervi conto facilmente che essa esprime ϵ proprio in coordinate sferiche, dove abbiamo usato il modulo del vettore \vec{k} , k , per indicare il modulo del vettore posizione del punto dello spazio \vec{k} (in pratica, il simbolo k che compare nella relazione di dispersione così' come l'abbiamo usata finora non e' altro che il simbolo r che abbiamo usato nel richiamo sulle coordinate sferiche).

Tanto per chiarire ulteriormente, la relazione di dispersione in coordinate cartesiane dovrebbe essere scritta come:

$$\epsilon(\vec{k}) = \epsilon(k_x, k_y, k_z) = \frac{\hbar^2}{2m} (k_x^2 + k_y^2 + k_z^2) \quad (442)$$

con k_x , k_y e k_z le componenti del vettore \vec{k} lungo i tre assi cartesiani.

163. Data la simmetria sferica, in coordinate sferiche l'energia dipende solo dal modulo (al quadrato) del vettore \vec{k} , ma *non* dagli angoli θ e/o φ :

†Non riporto la derivazione perché e' troppo lunga; nella maggior parte dei libri si fa altrettanto, ma nel libro di McQuarrie [1] l'autore scrive: “[la conversione del gradiente in coordinate sferiche] is a long, tedious, standard problem that arises in every quantum-chemistry course and should be done once (and only once!) by any serious student”

[1] D.A. McQuarrie. Quantum chemistry. University Science Books, 1983.

$$\epsilon(\vec{k}) = \epsilon(k, \theta, \varphi) = \frac{\hbar^2}{2m} k^2 \quad (443)$$

E allora il suo gradiente e':

$$\vec{\nabla}\epsilon(\vec{k}) = \hat{e}_r \frac{\partial\epsilon}{\partial k} + \hat{e}_\theta \frac{1}{r} \overbrace{\frac{\partial\epsilon}{\partial\theta}}^{=0} + \hat{e}_\varphi \frac{1}{r \sin\theta} \overbrace{\frac{\partial\epsilon}{\partial\varphi}}^{=0} \quad (444)$$

$$= \hat{e}_r \frac{\partial\epsilon}{\partial k} \quad (445)$$

e per il modulo del gradiente si ha:

$$|\vec{\nabla}\epsilon(\vec{k})| = \left| \frac{\partial\epsilon}{\partial k} \right| = 2 \left(\frac{\hbar^2}{2m} \right) k \quad (446)$$

(k , essendo il modulo di \vec{k} , e' sempre > 0)

Sostituendo nella 433 si ha:

$$g(\epsilon) = \frac{1}{8\pi^3} \int_{S(\epsilon)} \frac{1}{2 \left(\frac{\hbar^2}{2m} \right) k} dS \quad (447)$$

Sempre per il fatto che le isosuperfici sono sferiche, k e' uguale in tutti i punti di $S(\epsilon)$ e quindi tutto il denominatore si puo' portare fuori dal segno di integrale:

$$g(\epsilon) = \frac{1}{8\pi^3} \frac{1}{2 \left(\frac{\hbar^2}{2m} \right) k} \int_{S(\epsilon)} dS \quad (448)$$

A questo punto l'integrale e' diventato semplicemente la superficie di una sfera di raggio k :

$$g(\epsilon) = \frac{1}{8\pi^3} \frac{1}{2 \left(\frac{\hbar^2}{2m} \right) k} 4\pi k^2 = \frac{1}{4\pi^2} \frac{2m}{\hbar^2} k = \frac{1}{4\pi^2} \frac{2m}{\hbar^2} \left(\frac{2m}{\hbar^2} \right)^{\frac{1}{2}} \epsilon^{\frac{1}{2}} = \frac{1}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{\frac{3}{2}} \epsilon^{\frac{1}{2}} \quad (449)$$

che e' proprio l'eq. 415.

164. La terza (e piu' rigorosa) definizione della densita' degli stati e' la seguente:

$$g(\epsilon(\vec{k}')) = \frac{1}{V} \sum_{\vec{k}} \delta(\epsilon(\vec{k}') - \epsilon(\vec{k})) \quad (450)$$

V e' il volume del campione solido, la somma e' eseguita su tutti gli stati \vec{k} permessi, \vec{k}' e' un punto generico (non necessariamente permesso) dello spazio \vec{k} ed ϵ e' l'energia (funzione della posizione nello spazio \vec{k}). Il simbolo δ indica la funzione delta di Dirac.

Prima di discutere la definizione 450 apriamo una parentesi sulla funzione delta di Dirac.

165. Nonostante venga comunemente chiamata “funzione” (e anche noi la chiameremo così), in realtà δ non è una funzione “ordinaria” in senso matematico (in matematica, δ è una *distribuzione* o *funzione generalizzata*). Ciò appare chiaro dalla sua definizione,¹² che è:

$$\begin{cases} \int_{-\infty}^{\infty} \delta(x) dx = 1 \\ \delta(x) = 0 \quad \text{per } x \neq 0 \end{cases} \quad (451)$$

Come funzione, δ è piuttosto strana. Essa è nulla ovunque, tranne che per $x = 0$, dove però il suo valore è indefinito! Inoltre, il suo integrale su tutto l'asse reale vale 1, nonostante che essa sia diversa da zero solo in un punto ($x = 0$). Notate che per qualsiasi funzione “normale” che sia nulla ovunque tranne che in un numero finito di punti (in questo caso il solo $x = 0$) lo stesso integrale vale 0.

Nonostante queste “anomalie”, la funzione δ è utilissima. Come vedremo, la sua principale utilità è quella di consentire di trattare variabili discrete con strumenti matematici che si usano per variabili continue.

166. Il modo migliore per visualizzare la funzione δ è quello di realizzare che essa rappresenta il limite di alcune funzioni o della loro derivata.

Considerate ad esempio la seguente funzione:

$$f(x) = \begin{cases} 0 & x < -a \\ 1/(2a) & -a \leq x \leq a \\ 0 & x > a \end{cases} \quad (452)$$

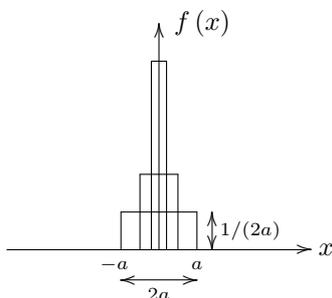
con $a > 0$.

Si tratta di una funzione che descrive un impulso rettangolare di altezza $1/(2a)$ e larghezza $2a$. Il suo integrale è ovviamente 1 $\forall a > 0$. δ rappresenta il limite di questa funzione quando $a \rightarrow 0$:

$$\lim_{a \rightarrow 0} f(x) = \delta(x) \quad (453)$$

Man mano che a diminuisce tendendo a zero, l'impulso diventa sempre più stretto, mantenendo però il valore del suo integrale unitario: al limite la funzione vale 0 ovunque tranne che per $x = 0$, dove il suo valore tende ad infinito, e il suo integrale vale 1, che è esattamente la definizione di δ .

È semplice realizzarlo se guardate la figura seguente, dove sono mostrati tre degli impulsi rettangolari in corrispondenza a tre valori sempre più piccoli di a .



Lo stesso procedimento può essere impiegato con qualsiasi altra funzione che dia un picco centrato a $x = 0$ e il cui integrale sia unitario. Ad esempio, la funzione di Gauss normalizzata:

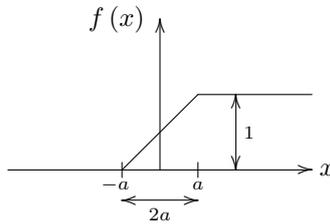
$$\lim_{\sigma \rightarrow 0} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2\right) = \delta(x) \quad (454)$$

167. Un altro modo equivalente di visualizzare la funzione δ e' quello di ottenerla come limite della derivata di funzioni "a gradino" centrate a $x = 0$. Considerate ad esempio la seguente funzione:

$$f(x) = \begin{cases} 0 & x < -a \\ \frac{x+a}{2a} & -a \leq x \leq a \\ 1 & x > a \end{cases} \quad (455)$$

con $a > 0$.

Il grafico di questa funzione e':



Per la sua derivata si ha:

$$f'(x) = \begin{cases} 0 & x < -a \\ 1/(2a) & -a \leq x \leq a \\ 0 & x > a \end{cases} \quad (456)$$

Notate che la derivata di questa funzione non e' altro che la funzione ad impulso rettangolare considerata prima. Ma allora deve essere:

$$\lim_{a \rightarrow 0} f'(x) = \delta(x) \quad (457)$$

Quando $a \rightarrow 0$, la funzione $f(x)$ tende a diventare la cosiddetta funzione di Heaviside, $H(x)$, in cui il gradino centrato a $x = 0$ diventa perfettamente verticale. In tali condizioni la sua derivata e' nulla ovunque, tranne che per $x = 0$, dove tende ad infinito, ($\lim_{a \rightarrow 0} 1/(2a)$), essendo il tratto di curva in quel punto perfettamente verticale.

Questo consente una definizione equivalente della funzione δ :

$$\delta(x) = \frac{dH(x)}{dx} \quad (458)$$

o, simmetricamente, una definizione della funzione di Heaviside:

$$H(x) = \int_{-\infty}^x \delta(x) dx \quad (459)$$

168. Al di la' del formalismo, la descrizione "intuitiva" della funzione δ e' quella di una funzione nulla ovunque, tranne che a $x = 0$, dove essa "spara un colpo", o "compie un salto" o "compie un guizzo" (genera uno spike di larghezza nulla e altezza tendente ad infinito). Per questo motivo uno dei suoi impieghi e' quello di "contare" oggetti, facendole "sparare un colpo" ogni volta

che ne incontra uno (chiariremo meglio fra un attimo il significato di questa affermazione).

169. Prima di chiudere questa parentesi sulla funzione δ ne vediamo alcune proprietà che useremo. Mostriamo qui le proprietà in una sola dimensione, ma esse valgono inalterate (con ovvie modifiche) anche nello spazio 3D.

1. Il punto in cui la funzione δ “batte un colpo” può essere fatto diventare un qualunque punto dell’asse x applicando una traslazione (punto 62). Quindi:

$$\delta(x - a) \tag{460}$$

è nulla ovunque tranne che per $x = a$, dove il suo valore tende ad infinito.

2. Una delle più importanti proprietà della funzione δ è la cosiddetta “sifting property” (to sift=setacciare, vagliare); data una funzione qualunque $f(x)$ vale:

$$\int_{-\infty}^{\infty} f(x) \delta(x - a) dx = f(a) \tag{461}$$

cioè: l’effetto di moltiplicare l’argomento di un integrale per la funzione δ è quello di far sì che il risultato dell’integrale sia semplicemente il suo argomento calcolato nel punto speciale della funzione δ .

La funzione δ “passa al setaccio” i valori della funzione f facendo passare oltre il setaccio il solo valore $f(a)$.

In generale, come vedremo, la funzione δ diventa utile solo quando si trova sotto il segno di integrale; questo è intuitivo: mentre il valore non nullo della funzione δ è indefinito ($\rightarrow \infty$), il valore del suo integrale su tutto l’asse reale è definito ($= 1$).

La dimostrazione di questa proprietà è semplice.

Partiamo dall’integrale:

$$\int_{-\infty}^{\infty} f(x) \delta(x - a) dx \tag{462}$$

Siccome $\delta(x - a)$ vale 0 su tutto l’asse reale tranne che per $x = a$, il valore dell’integrale deve rimanere lo stesso se a $f(x)$ sostituiamo $f(a)$:

$$= \int_{-\infty}^{\infty} f(a) \delta(x - a) dx$$

ma $f(a)$ non dipende da x e quindi lo possiamo portare fuori dal segno di integrale; quest’ultimo vale 1 per definizione e quindi la proprietà è dimostrata:

$$= f(a) \int_{-\infty}^{\infty} \delta(x - a) dx = f(a) \times 1 = f(a) \tag{463}$$

3. Scaling property:

$$\delta(ax) = \frac{1}{a} \delta(x) \quad (464)$$

La dimostrazione e':

$$\int_{-\infty}^{\infty} \delta(ax) dx = \int_{-\infty}^{\infty} \delta(ax) \frac{1}{a} d(ax) = \int_{-\infty}^{\infty} \frac{1}{a} \delta(y) dy \quad (465)$$

Dal confronto fra il primo e il terzo integrale segue la proprieta' (ovviamente, il nome della variabile y e' completamente arbitrario e avrebbe potuto tranquillamente essere x : ho usato y solo per non fare confusione con la variabile x usata nei passaggi precedenti).

Chiusa parentesi sulla funzione δ di Dirac.

170. Torniamo ora alla 450:

$$g(\epsilon(\vec{k}')) = \frac{1}{V} \sum_{\vec{k}} \delta(\epsilon(\vec{k}') - \epsilon(\vec{k}))$$

Il significato della somma e' il seguente. Ogni addendo e' una funzione δ che "spara un colpo" quando l'energia $\epsilon(\vec{k}')$, pensata come variabile continua della variabile (a sua volta) continua \vec{k}' (cioe' un *qualsiasi* punto dello spazio \vec{k} , non necessariamente appartenente al reticolo dei punti permessi), e' uguale a quella di un livello permesso $\epsilon(\vec{k})$.

Il grafico del singolo addendo e' dunque il seguente (indichiamo lo spike della funzione δ con una freccia diretta verso l'alto):



Siccome ogni addendo e' nullo ovunque tranne che in corrispondenza a un particolare valore di energia ($\epsilon(\vec{k})$), l'intera sommatoria sara' nulla ovunque, tranne che in corrispondenza alle energie permesse (che si ottengono in corrispondenza ai valori permessi di \vec{k} , cioe' ai punti del reticolo reciproco):

$$\sum_{\vec{k}} \delta(\epsilon(\vec{k}') - \epsilon(\vec{k})) : \dots \begin{array}{c} \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \quad \dots \\ \epsilon(\vec{k}_i) \quad \epsilon(\vec{k}_{i+1}) \quad \epsilon(\vec{k}_{i+2}) \quad \epsilon(\vec{k}_{i+3}) \quad \epsilon(\vec{k}_{i+4}) \end{array} \epsilon(\vec{k}') \quad (466)$$

Come avevamo gia' accennato (p. 124) la sommatoria *conta* gli stati permessi: il suo grafico in funzione dell'energia mostra uno spike (di larghezza zero) ogni volta che si incontra un valore di energia permessa.

171. Ora che abbiamo capito cosa significa la sommatoria introdotta nella 450, vogliamo assicurarci del fatto che essa esprima effettivamente la densita' degli stati, cosi' come l'abbiamo definita finora (p. 108) (in effetti, a prima vista,

sembra difficile poter interpretare la 450 come il numero di stati permessi per unita' di energia).

A questo scopo riprendiamo la 409:

$$N(\epsilon + \Delta\epsilon) - N(\epsilon) = V \int_{\epsilon}^{\epsilon + \Delta\epsilon} g(\epsilon) d\epsilon$$

Se la definizione 450 e' corretta, allora sostituendola a $g(\epsilon)$ nella 409 dobbiamo ottenere il numero di stati permessi contenuti nel range di energia $[\epsilon, \epsilon + \Delta\epsilon]$. Ma questo e' precisamente cio' che succede e lo possiamo verificare molto facilmente.

Sostituiamo $\frac{1}{V} \sum_{\vec{k}} \delta(\epsilon(\vec{k}') - \epsilon(\vec{k}))$ al posto di $g(\epsilon)$:

$$\int_{\epsilon}^{\epsilon + \Delta\epsilon} \frac{1}{V} \sum_{\vec{k}} \delta(\epsilon(\vec{k}') - \epsilon(\vec{k})) d\epsilon = \frac{1}{V} \int_{\epsilon}^{\epsilon + \Delta\epsilon} \sum_{\vec{k}} \delta(\epsilon(\vec{k}') - \epsilon(\vec{k})) d\epsilon \quad (467)$$

Ora, grazie alla linearita' dell'operatore di integrazione, scambiamo di posto l'integrale con la sommatoria ("l'integrale di una somma e' uguale alla somma degli integrali"):

$$= \frac{1}{V} \sum_{\vec{k}} \int_{\epsilon}^{\epsilon + \Delta\epsilon} \delta(\epsilon(\vec{k}') - \epsilon(\vec{k})) d\epsilon$$

Ora: siccome gli integrali sono calcolati fra ϵ ed $\epsilon + \Delta\epsilon$, tutti i termini della sommatoria in cui $\epsilon(\vec{k})$ e' *al di fuori dell'intervallo di integrazione* sono identicamente nulli (pensate a come e' fatto il grafico della funzione δ). Possiamo percio' restringere la sommatoria a quei soli valori permessi di \vec{k} tali che $\epsilon(\vec{k}) \in [\epsilon, \epsilon + \Delta\epsilon]$:

$$= \frac{1}{V} \sum_{\vec{k}: \epsilon(\vec{k}) \in [\epsilon, \epsilon + \Delta\epsilon]} \int_{\epsilon}^{\epsilon + \Delta\epsilon} \delta(\epsilon(\vec{k}') - \epsilon(\vec{k})) d\epsilon$$

Ora osserviamo che, per definizione ciascun integrale che compare nella sommatoria e' = 1 e quindi:

$$= \frac{1}{V} \sum_{\vec{k}: \epsilon(\vec{k}) \in [\epsilon, \epsilon + \Delta\epsilon]} 1 \quad (468)$$

Nell'espressione a cui siamo arrivati si somma 1 per ogni punto \vec{k} tale che l'energia ad esso corrispondente sia inclusa nell'intervallo $[\epsilon, \epsilon + \Delta\epsilon]$: ma questo e' come dire che la sommatoria e' uguale al numero di stati \vec{k} permessi contenuti nell'intervallo finito di energia $[\epsilon, \epsilon + \Delta\epsilon]$, cioe' esattamente la 409.

172. Osservazioni:

- Delle tre definizioni che abbiamo introdotto per la densita' degli stati (eq. 408, 433, 450), quella basata sulla funzione δ di Dirac e' la piu' rigorosa. In essa, infatti, la natura discreta delle energie permesse per gli elettroni (un fatto squisitamente quantomeccanico) viene considerata in modo "naturale" grazie all'uso della funzione di Dirac, che, come abbiamo gia' accennato (p. 123), nasce proprio per questo.

- Come già sottolineato più volte, la 450 è la densità dei punti \vec{k} permessi e siccome ad ogni punto \vec{k} permesso corrispondono due stati monoelettronici con spin antiparallelo, la densità di *stati monoelettronici* è semplicemente la 450 con un fattore 2 addizionale.

173. Come abbiamo fatto con le definizioni 408 e 433, verifichiamo che anche la 450 fornisce lo stesso risultato (eq. 415) per un gas di elettroni di Fermi confinato in un campione cubico di lato L .

Partiamo dalla definizione 450:

$$g(\epsilon(\vec{k}')) = \frac{1}{V} \sum_{\vec{k}} \delta(\epsilon(\vec{k}') - \epsilon(\vec{k}))$$

Moltiplichiamo e dividiamo per il volume della cella primitiva dello spazio \vec{k} : $\Delta\vec{k}' = \frac{8\pi^3}{V}$:

$$= \frac{1}{V} \frac{1}{\frac{8\pi^3}{V}} \sum_{\vec{k}} \delta(\epsilon(\vec{k}') - \epsilon(\vec{k})) \Delta\vec{k}' = \frac{1}{8\pi^3} \sum_{\vec{k}} \delta(\epsilon(\vec{k}') - \epsilon(\vec{k})) \Delta\vec{k}'$$

Ora: siccome la densità degli stati e' *normalizzata* per il volume V del campione, non dipende da esso e quindi non perdiamo in generalità se consideriamo un campione avente un volume sufficientemente grande ($V \rightarrow \infty$) da far sì che il volume della cella primitiva del reticolo reciproco, ($\Delta\vec{k}' = 8\pi^3/V$), possa essere considerato infinitesimo nella sommatoria (ricordate che il volume della cella primitiva diretta e quello della cella primitiva reciproca sono inversamente proporzionali, eq. 244). Sotto questa ipotesi, la sommatoria può essere riscritta come il corrispondente integrale su tutto lo spazio \vec{k} :

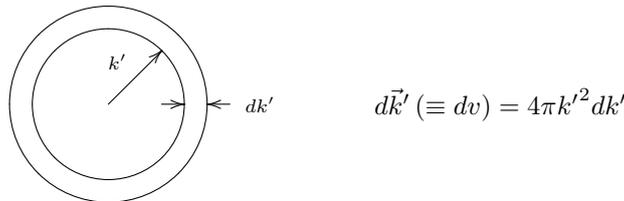
$$= \frac{1}{8\pi^3} \int_{\text{all } \vec{k} \text{ space}} \delta(\epsilon(\vec{k}') - \epsilon(\vec{k})) d\vec{k}'$$

dove $d\vec{k}'$ è l'elemento infinitesimo di volume dello spazio reciproco.

Ora esprimiamo le due energie argomento del δ con la relazione di dispersione valida per il sistema che stiamo considerando:

$$= \frac{1}{8\pi^3} \int_{\text{all } \vec{k} \text{ space}} \delta\left(\frac{\hbar^2 k'^2}{2m} - \frac{\hbar^2 k^2}{2m}\right) d\vec{k}'$$

Siccome l'energia di questo sistema dipende solo dal modulo del vettore \vec{k} (simmetria sferica), ci conviene esprimere il volume infinitesimo $d\vec{k}'$ come il volume di un guscio sferico di raggio k' e spessore dk' :



$$\begin{aligned}
d\vec{k}' (\equiv dv) &= \frac{4}{3}\pi (k' + dk')^3 - \frac{4}{3}\pi k'^3 \\
&= \frac{4}{3}\pi \left(\cancel{k'^3} + \overbrace{3k' dk'^2}^{\approx 0} + 3k'^2 dk' + \overbrace{dk'^3}^{\approx 0} - \cancel{k'^3} \right) = 4\pi k'^2 dk'
\end{aligned}$$

(gli infinitesimi dk'^2 e dk'^3 sono trascurabili rispetto a dk')
In tal modo l'integrale diventa:

$$= \frac{1}{8\pi^3} \int_0^\infty 4\pi k'^2 \delta \left(\frac{\hbar^2 k'^2}{2m} - \frac{\hbar^2 k^2}{2m} \right) dk' = \frac{1}{2\pi^2} \int_0^\infty k'^2 \delta \left(\frac{\hbar^2 k'^2}{2m} - \frac{\hbar^2 k^2}{2m} \right) dk'$$

(notate anche che gli estremi di integrazione sono cambiati, essendo ora la variabile di integrazione il modulo del generico vettore \vec{k}' dello spazio reciproco)
Per risolvere l'integrale, operiamo il seguente cambio di variabile:

$$y = \frac{\hbar^2 k'^2}{2m} \quad (469)$$

Si avra':

$$k'^2 = \frac{2m}{\hbar^2} y \quad (470)$$

$$k' = \left(\frac{2m}{\hbar^2} \right)^{\frac{1}{2}} y^{\frac{1}{2}} \quad (471)$$

$$dk' = \frac{1}{2} \left(\frac{2m}{\hbar^2} \right)^{\frac{1}{2}} y^{-\frac{1}{2}} dy \quad (472)$$

Sostituendo si ottiene:

$$\begin{aligned}
&= \frac{1}{2\pi^2} \int_0^\infty \frac{2m}{\hbar^2} y \delta \left(y - \frac{\hbar^2 k^2}{2m} \right) \frac{1}{2} \left(\frac{2m}{\hbar^2} \right)^{\frac{1}{2}} y^{-\frac{1}{2}} dy \\
&= \frac{1}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{\frac{3}{2}} \int_0^\infty y^{\frac{1}{2}} \delta \left(y - \frac{\hbar^2 k^2}{2m} \right) dy
\end{aligned}$$

Per la proprieta' di sifting della funzione δ (proprieta' 2, p. 125), l'integrale e' immediato:

$$\int_0^\infty y^{\frac{1}{2}} \delta \left(y - \frac{\hbar^2 k^2}{2m} \right) dy = \left(\frac{\hbar^2 k^2}{2m} \right)^{\left(\frac{1}{2}\right)} = \epsilon^{\frac{1}{2}} \quad (473)$$

dove, nell'ultimo passaggio, abbiamo sfruttato la relazione di dispersione (eq. 344).
Si ottiene pertanto la 415:

$$= \frac{1}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{\frac{3}{2}} \epsilon^{\frac{1}{2}}$$

174. OK, ora sappiamo bene cos'è la densità degli stati: ma cosa ce ne facciamo?

La risposta in termini del tutto generali è:

La densità degli stati (che si può ottenere per via computazionale o sperimentale) ci consente di trasformare in integrali delle somme discrete che sarebbero altrimenti intrattabili.

175. Come abbiamo visto (eq. 406), qualsiasi proprietà Q legata agli stati monoelettronici occupati in un campione solido è ottenibile moltiplicando il valore della proprietà in un determinato livello $Q(\epsilon(\vec{k}))$ per l'occupazione di tale livello (l'occupazione è data dalla funzione di Fermi-Dirac (punto 131), $f(\epsilon(\vec{k}))$), sommando su tutti i livelli e normalizzando la somma rispetto al volume V del campione:

$$Q = \frac{2}{V} \sum_{\vec{k}} Q(\epsilon(\vec{k})) f(\epsilon(\vec{k})) \quad (474)$$

176. La cosa importante in questo contesto è che i livelli elettronici occupati in un campione solido sono dell'ordine del numero di Avogadro e quindi affrontare in modo diretto una somma come quella mostrata è improponibile (come ho già notato, questo è uno degli aspetti più drasticamente diversi nel confronto dell'analisi della struttura elettronica dei solidi con quella delle molecole).

Ed è a questo punto che entra in gioco la densità degli stati.

Per la proprietà di sifting della funzione δ (proprietà 2, p. 125) vale:

$$Q(\epsilon(\vec{k})) f(\epsilon(\vec{k})) = \int_{\text{all } \vec{k} \text{ space}} Q(\epsilon(\vec{k}')) f(\epsilon(\vec{k}')) \delta(\epsilon(\vec{k}') - \epsilon(\vec{k})) d\vec{k}' \quad (475)$$

e quindi possiamo riscrivere la somma come:

$$\sum_{\vec{k}} \int_{\text{all } \vec{k} \text{ space}} Q(\epsilon(\vec{k}')) f(\epsilon(\vec{k}')) \delta(\epsilon(\vec{k}') - \epsilon(\vec{k})) d\vec{k}' \quad (476)$$

Scambiando di posto la sommatoria e l'integrale (l'integrale è un operatore lineare) si ha:

$$= \int_{\text{all } \vec{k} \text{ space}} \sum_{\vec{k}} Q(\epsilon(\vec{k}')) f(\epsilon(\vec{k}')) \delta(\epsilon(\vec{k}') - \epsilon(\vec{k})) d\vec{k}'$$

$Q(\epsilon(\vec{k}'))$ e $f(\epsilon(\vec{k}'))$ non dipendono da \vec{k} e quindi:

$$= \int_{\text{all } \vec{k} \text{ space}} Q(\epsilon(\vec{k}')) f(\epsilon(\vec{k}')) \sum_{\vec{k}} \delta(\epsilon(\vec{k}') - \epsilon(\vec{k})) d\vec{k}'$$

E ora riconosciamo nella sommatoria proprio la densità degli stati (moltiplicata per V) (eq. 450):

$$= \int_{\text{all } \vec{k} \text{ space}} Q(\epsilon(\vec{k}')) f(\epsilon(\vec{k}')) V g(\epsilon(\vec{k}')) d\vec{k}' \quad (477)$$

Ricapitolando, come preannunciato, abbiamo trasformato una somma intrattabile in un integrale molto piu' maneggevole (anche se in realta' gli integrali di questo tipo sono piuttosto complicati, ma sempre meno complicati della somma diretta).

In definitiva, dunque, vale:

$$\frac{2}{V} \sum_{\vec{k}} Q(\epsilon(\vec{k})) f(\epsilon(\vec{k})) = 2 \int_{\text{all } \vec{k} \text{ space}} Q(\epsilon(\vec{k}')) f(\epsilon(\vec{k}')) g(\epsilon(\vec{k}')) d\vec{k}' \quad (478)$$

177. Una possibile interpretazione dell'integrale:

$$\int_{\text{all } \vec{k} \text{ space}} Q(\epsilon(\vec{k}')) f(\epsilon(\vec{k}')) g(\epsilon(\vec{k}')) d\vec{k}' \quad (479)$$

e' la seguente.

Dovremmo sommare il valore di $Q(\epsilon(\vec{k}')) f(\epsilon(\vec{k}'))$ su tutti i punti \vec{k}' permessi:

$$\sum_{\vec{k}'} Q(\epsilon(\vec{k}')) f(\epsilon(\vec{k}')) \quad (480)$$

Per semplificare la somma, potremmo pensare di ripartire lo spazio reciproco in elementi di volume $\Delta\vec{k}'$, ciascuno centrato sul vettore reciproco \vec{k}' . Ad ogni elemento di volume $\Delta\vec{k}'$ corrisponde un determinato range di energia $[\epsilon(\vec{k}'), \epsilon(\vec{k}') + \Delta\epsilon_{\Delta\vec{k}'}]$ ($\epsilon(\vec{k}')$ e $\epsilon(\vec{k}') + \Delta\epsilon_{\Delta\vec{k}'}$ sono, rispettivamente, il minimo e il massimo valore di energia corrispondenti a un punto interno all'elemento di volume $\Delta\vec{k}'$; quindi $\Delta\epsilon_{\Delta\vec{k}'}$ dipende da $\Delta\vec{k}'$). Supponendo che $Q(\epsilon(\vec{k}')) f(\epsilon(\vec{k}'))$ sia *approssimativamente costante* nell'intervallo di energia $[\epsilon(\vec{k}'), \epsilon(\vec{k}') + \Delta\epsilon_{\Delta\vec{k}'}]$ possiamo dire:

$$\left[\begin{array}{l} \text{somma dei valori} \\ \text{di } Q(\epsilon(\vec{k}')) f(\epsilon(\vec{k}')) \text{ su tutti} \\ \text{i punti permessi nell'elemento} \\ \text{di volume } \Delta\vec{k}' \text{ centrato a } \vec{k}' \end{array} \right] \approx Q(\epsilon(\vec{k}')) f(\epsilon(\vec{k}')) \times \underbrace{\left[\begin{array}{l} \text{n. di punti reciproci} \\ \text{permessi nell'elemento} \\ \text{di volume } \Delta\vec{k}' \end{array} \right]}_{\approx g(\epsilon(\vec{k}')) \Delta\epsilon_{\Delta\vec{k}'}}$$

ovvero:

$$\sum_{\vec{k}'} Q(\epsilon(\vec{k}')) f(\epsilon(\vec{k}')) \approx \sum_{\Delta\vec{k}'} Q(\epsilon(\vec{k}')) f(\epsilon(\vec{k}')) g(\epsilon(\vec{k}')) \Delta\epsilon_{\Delta\vec{k}'} \quad (481)$$

cioe', in pratica, abbiamo raggruppato i punti \vec{k}' in gruppi $\Delta\vec{k}'$ (ciascun gruppo e' formato dai punti permessi interni all'elemento di volume $\Delta\vec{k}'$) e assunto che:

- \Rightarrow il valore di $Q(\epsilon(\vec{k}')) f(\epsilon(\vec{k}'))$ sia costante in ciascun gruppo di punti, cioè in ciascun elemento di volume $\Delta\vec{k}'$
 \Rightarrow il numero di punti \vec{k}' in ciascun gruppo sia ottenuto dalla densità degli stati moltiplicata per l'intervallo di energia corrispondente all'elemento di volume dello spazio reciproco $\Delta\vec{k}'$

Queste due assunzioni sono solo approssimativamente vere se $\Delta\vec{k}'$ è finito, ma diventano esatte se $\Delta\vec{k}' \rightarrow 0$. Ma in questo caso, la somma al secondo membro dell'uguaglianza scritta sopra diventa l'integrale 478.

178. Come esempio di applicazione della 478, per il calcolo dell'energia del gas di elettroni di Fermi, si ha semplicemente:

$$Q(\epsilon(\vec{k})) = \epsilon(\vec{k}) \quad (482)$$

e quindi, detta u l'energia per unità di volume del solido:

$$u = \frac{2}{V} \sum_{\vec{k}} \epsilon(\vec{k}) f(\epsilon(\vec{k})) = 2 \int_{\text{all } \vec{k} \text{ space}} \epsilon(\vec{k}) f(\epsilon(\vec{k})) g(\epsilon(\vec{k})) d\vec{k} \quad (483)$$

179. In modo analogo, per la densità di elettroni n (numero di elettroni del campione per unità di volume), nella 478 si ha:

$$Q(\epsilon(\vec{k})) = 1 \quad (484)$$

(in questo caso la proprietà è l'occupazione dello stato monoelettronico a energia ϵ , cioè semplicemente la funzione di Fermi-Dirac $f(\epsilon(\vec{k}))$)

e quindi:

$$n = \frac{2}{V} \sum_{\vec{k}} f(\epsilon(\vec{k})) = 2 \int_{\text{all } \vec{k} \text{ space}} f(\epsilon(\vec{k})) g(\epsilon(\vec{k})) d\vec{k} \quad (485)$$

Notate che questa non è altro che la "riedizione" della condizione di normalizzazione 401, in cui la somma è fatta sui punti \vec{k} permessi (invece che sugli stati monoelettronici) e si divide per il volume V per ottenere la densità di elettroni invece che il loro numero totale.

La struttura elettronica in presenza di un potenziale periodico ²⁶

180. Nella parte precedente abbiamo visto il comportamento degli elettroni in un solido *in assenza di un potenziale* (ovvero in presenza di un potenziale costante e nullo). Cioe' l'equazione d'onda che abbiamo studiato era del tipo:

$$-\frac{\hbar^2}{2m}\nabla^2\Psi(\vec{r}) = \epsilon\Psi(\vec{r})$$

In pratica, nel gas di elettroni di Fermi si assume che gli elettroni possiedano soltanto energia cinetica: vengono ignorate le interazioni fra gli elettroni e i nuclei atomici carichi positivamente o le interazioni degli elettroni fra loro stessi.

181. In questa parte affineremo la descrizione di un solido studiando l'equazione di Schrödinger monoelettronica nel caso in cui essa contenga anche un termine di energia potenziale (e non solo l'operatore dell'energia cinetica), cioe':

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + U(\vec{r})\right)\Psi(\vec{r}) = \epsilon\Psi(\vec{r}) \quad (486)$$

182. La prima cosa da dire e' che non studieremo il caso di un potenziale qualsiasi, ma quello di un potenziale *periodico*. In altre parole, il solido che studieremo sara' un solido *crystallino*, cioe' un campione solido in cui e' presente una cella primitiva dalla cui ripetizione periodica in tre direzioni (non necessariamente fra loro perpendicolari) e' ottenuto l'intero campione.

Il potenziale $U(\vec{r})$ che compare nell'equazione d'onda che consideriamo e' dovuto ai nuclei degli atomi che costituiscono il solido cristallino: siccome tali atomi sono posizionati in modo periodico, il potenziale da essi generato deve necessariamente essere periodico (punto 82).

183. La figura 49 mostra in modo qualitativo come potrebbe essere verosimilmente la funzione $U(\vec{r})$ per il caso di un "cristallo" bidimensionale con una cella primitiva quadrata: l'energia potenziale di un elettrone dovuta all'attrazione nucleare e' costante e vicina a 0 nelle regioni lontane dai nuclei, mentre vicino ad essi diventa rapidamente negativa:

184. Prima di continuare e' opportuno rimarcare che il problema generale del trattamento di un sistema di elettroni in un solido (sia pure periodico come quello a cui ci interessiamo) e' ben piu' complesso di come lo definiamo in questa sezione; in termini rigorosi e' un problema a molte particelle in cui si dovrebbero considerare non solo le forze esercitate dai nuclei sugli elettroni, ma anche le forze che gli elettroni esercitano fra loro (e quindi, usare un operatore hamiltoniano *monoelettronico* e' gia' una grossa approssimazione). Un'altra complicazione e' che i nuclei non sono fermi, ma vibrano attorno alle loro posizioni di equilibrio. Possiamo pensare che tutte queste difficolta' addizionali vengano in qualche modo superate definendo un potenziale "efficace" che ne tenga conto: la cosa essenziale nel presente contesto e' che, qualunque sia la specifica definizione del potenziale $U(\vec{r})$, esso sia periodico della stessa periodicit  del reticolo cristallino del solido considerato.

185. Definizione del sistema.

Il solido che consideriamo ha una struttura cristallina basata su un reticolo di Bravais la cui cella primitiva e' determinata dai tre vettori:

²⁶Trovate il materiale di questa parte nel libro di Ashcroft e Mermin¹³ e/o nel libro di Kittel¹⁴ e/o nel libro di Ibach e L th.¹⁶

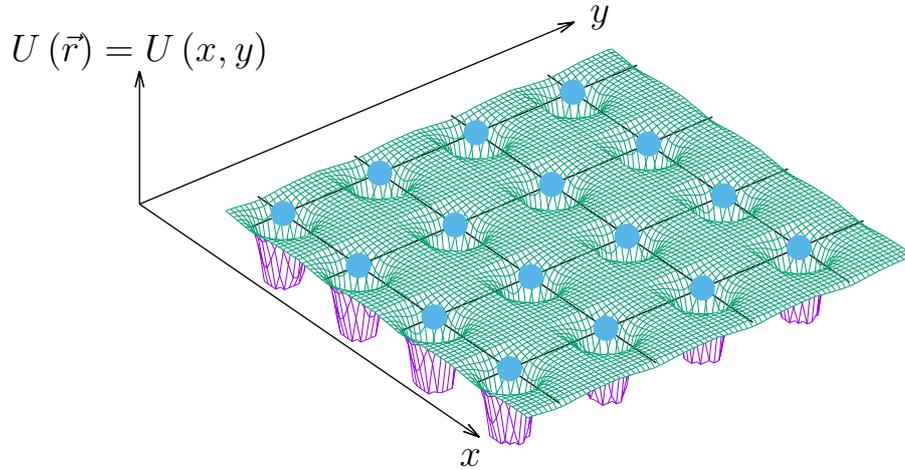


Figura 49:

$$\vec{a}_i, \quad i = 1, 3 \quad (487)$$

Indicheremo un generico vettore del reticolo di Bravais con:

$$\vec{R} = n_1 \vec{a}_1 + n_2 \vec{a}_2 + n_3 \vec{a}_3 = \sum_{i=1}^3 n_i \vec{a}_i, \quad n_i \in \mathbb{Z} \quad (488)$$

Supporremo che la base della struttura cristallina sia monoatomica: cioè c'è un solo atomo associato ad ogni punto del reticolo di Bravais.

L'assunzione che il potenziale $U(\vec{r})$ sia periodico, significa la cosa seguente:

$$U(\vec{r} + \vec{R}) = U(\vec{r}) \quad \forall \vec{R} \in \text{al reticolo di Bravais} \quad (489)$$

cioè $U(\vec{r})$ (ri-)assume lo stesso valore quando ci spostiamo (nello spazio) di un qualsiasi vettore traslazione del reticolo di Bravais (guardate anche la figura 49). Siamo interessati alle soluzioni dell'equazione di Schrödinger stazionaria e mono-elettronica seguente:

$$\left(-\frac{\hbar^2}{2m} \nabla^2 + U(\vec{r}) \right) \Psi(\vec{r}) = \epsilon \Psi(\vec{r}) \quad (490)$$

186. Trattare l'equazione differenziale 490 non è semplice; esiste tuttavia un teorema che fornisce un grosso aiuto a questo proposito: è noto come teorema di Bloch (dal nome del fisico svizzero Felix Bloch che lo dimostrò) e ha determinato un passo avanti fondamentale in tutta la fisica dello stato solido. Esso restringe il range delle possibili soluzioni dell'equazione 490 a una classe definita di funzioni.

187. Il teorema di Bloch può essere enunciato in due modi equivalenti.

1. Le soluzioni dell'equazione 490 possono essere espresse come il prodotto fra un'onda piana²⁷ e una funzione avente la periodicità del reticolo di Bravais. Cioè²⁸:

$$\Psi_{\vec{k}n}(\vec{r}) = \exp\left(\overbrace{i\vec{k} \cdot \vec{r}}^{\text{onda piana}}\right) u_{\vec{k}n}(\vec{r}) \quad (491)$$

con:

$$u_{\vec{k}n}(\vec{r} + \vec{R}) = u_{\vec{k}n}(\vec{r}) \quad \forall \vec{R} \in \text{al reticolo di Bravais} \quad (492)$$

2. Ogni soluzione $\Psi(\vec{r})$ dell'equazione 490 è tale che il suo valore dopo una traslazione pari a un vettore del reticolo di Bravais è ottenuto moltiplicando il valore prima della traslazione per il "fattore di fase" $\exp(i\vec{k} \cdot \vec{R})$:

$$\Psi(\vec{r} + \vec{R}) = \exp(i\vec{k} \cdot \vec{R}) \Psi(\vec{r}) \quad (493)$$

188. Vedete come il teorema di Bloch pone dei vincoli ben precisi alla forma funzionale delle soluzioni dell'equazione di Schrödinger contenente un potenziale periodico: esse possono essere espresse come prodotto fra un'onda piana e una funzione che abbia la stessa periodicità del reticolo di Bravais.

Notate che il teorema di Bloch non dice nulla sulla funzione $u(\vec{r})$ oltre al fatto che essa ha la stessa periodicità del reticolo di Bravais: in effetti, ottenere la funzione $u(\vec{r})$ significa risolvere l'equazione di Schrödinger (visto che poi è sufficiente moltiplicare la funzione $u(\vec{r})$ per l'onda piana $\exp(i\vec{k} \cdot \vec{r})$; notate però che anche l'onda piana contiene un parametro che va determinato: il numero d'onda \vec{k} : avremo parecchio da dire su di esso nel seguito).

L'esatta forma della funzione $u(\vec{r})$ dipende dalla forma del potenziale periodico $U(\vec{r})$ e in generale non è affatto semplice da trovare; tuttavia, il teorema di Bloch afferma che, qualunque sia la forma specifica della funzione $u(\vec{r})$, essa *deve* essere una funzione periodica con la stessa periodicità del reticolo di Bravais.

189. Attenti a non fare confusione: la soluzione dell'equazione di Schrödinger, $\Psi(\vec{r})$, *non* ha la periodicità del reticolo di Bravais (questo appare evidente anche dalla 493); in effetti non c'è alcun motivo per cui ciò debba avvenire: l'autofunzione dell'hamiltoniano 490 descrive lo stato di un elettrone nel campione solido e quindi è una funzione che si estende *in tutto il campione*.

Il fatto che il campione sia cristallino e quindi costituito dalla ripetizione di un'unica cella primitiva (e, quindi, che il potenziale che compare nell'hamiltoniano sia periodico) restringe la classe delle possibili soluzioni a quella delle funzioni di Bloch: funzioni *non* aventi la stessa periodicità del reticolo, ma date dal prodotto di un'onda piana e una funzione con la stessa periodicità del reticolo di Bravais.

Chiarito quanto sopra, vedrete che, in effetti, la funzione $\Psi(\vec{r})$ è periodica, ma *non* sul reticolo di Bravais su cui è basata la struttura cristallina del campione;

²⁷Come già detto in precedenza, parliamo di onda piana anche quando in realtà intendiamo la sola parte spaziale di un'onda piana.

²⁸L'autofunzione $\Psi(\vec{r})$ e la funzione periodica $u(\vec{r})$ sono indicizzate dal vettore d'onda \vec{k} , ma anche da un secondo indice n che definiremo in seguito (punto 219)

la periodicità della $\Psi(\vec{r})$ deriva dalle condizioni al contorno che imporreemo per la soluzione dell'equazione di Schrödinger (punto 196).

190. Prima di continuare, proviamo l'equivalenza dei due enunciati del teorema di Bloch.

491 e 492 \Rightarrow 493 Usando la 491 valutiamo $\Psi(\vec{r} + \vec{R})$:

$$\Psi(\vec{r} + \vec{R}) = \exp(i\vec{k} \cdot (\vec{r} + \vec{R})) u(\vec{r} + \vec{R}) \quad (494)$$

Svolgiamo l'esponenziale:

$$\Psi(\vec{r} + \vec{R}) = \exp(i\vec{k} \cdot \vec{R}) \exp(i\vec{k} \cdot \vec{r}) u(\vec{r} + \vec{R}) \quad (495)$$

In base alla 492:

$$\Psi(\vec{r} + \vec{R}) = \exp(i\vec{k} \cdot \vec{R}) \exp(i\vec{k} \cdot \vec{r}) u(\vec{r}) \quad (496)$$

ovvero:

$$\Psi(\vec{r} + \vec{R}) = \exp(i\vec{k} \cdot \vec{R}) \Psi(\vec{r}) \quad (497)$$

che è la 493

493 \Rightarrow 491 e 492 Partiamo dalla 493 e moltiplichiamo il secondo membro per $1 = \exp(i\vec{k} \cdot \vec{r}) \exp(-i\vec{k} \cdot \vec{r})$:

$$\begin{aligned} \Psi(\vec{r} + \vec{R}) &= \exp(i\vec{k} \cdot \vec{R}) \exp(i\vec{k} \cdot \vec{r}) \exp(-i\vec{k} \cdot \vec{r}) \Psi(\vec{r}) \\ &= \exp(i\vec{k} \cdot (\vec{r} + \vec{R})) \exp(-i\vec{k} \cdot \vec{r}) \Psi(\vec{r}) \\ &= \exp(i\vec{k} \cdot (\vec{r} + \vec{R})) u(\vec{r}) \end{aligned} \quad (498)$$

dove abbiamo definito:

$$u(\vec{r}) = \exp(-i\vec{k} \cdot \vec{r}) \Psi(\vec{r}) \quad (499)$$

Riscriviamo l'espressione appena ottenuta:

$$\Psi(\vec{r} + \vec{R}) = \exp(i\vec{k} \cdot (\vec{r} + \vec{R})) u(\vec{r}) \quad (500)$$

Innanzitutto, $u(\vec{r})$ ha la periodicità del reticolo di Bravais. Infatti:

$$u(\vec{r} + \vec{R}) = \exp(-i\vec{k} \cdot (\vec{r} + \vec{R})) \Psi(\vec{r} + \vec{R})$$

Sostituendo a $\Psi(\vec{r} + \vec{R})$ la sua espressione 493:

$$\begin{aligned}
&= \exp\left(-i\vec{k} \cdot (\vec{r} + \vec{R})\right) \exp\left(i\vec{k} \cdot \vec{R}\right) \Psi(\vec{r}) \\
&= \exp\left(-i\vec{k} \cdot \vec{r}\right) \Psi(\vec{r}) \\
&= u(\vec{r})
\end{aligned}$$

e questo dimostra 493 \Rightarrow 492.

Infine, proprio perche' abbiamo appena dimostrato che $u(\vec{r})$ ha la periodicita' del reticolo di Bravais, possiamo riscrivere la 500 nel modo seguente:

$$\Psi(\vec{r} + \vec{R}) = \exp\left(i\vec{k} \cdot (\vec{r} + \vec{R})\right) u(\vec{r} + \vec{R}) \quad (501)$$

che e' proprio la 491, col cambio di variabile $\vec{r}' \leftarrow \vec{r} + \vec{R}$.

191. Dimostrazione del teorema di Bloch.²⁹

Dimostreremo il teorema di Bloch nella forma data dall'eq. 493.

Per ogni vettore \vec{R} del reticolo di Bravais su cui e' basata la struttura del campione solido, definiamo un operatore di traslazione $T_{\vec{R}}$ che, applicato ad una funzione della posizione nello spazio $f(\vec{r})$, ne trasla il "grafico" di \vec{R} (punto 62):

$$T_{\vec{R}}f(\vec{r}) = f(\vec{r} + \vec{R}) \quad (502)$$

Cio' che vogliamo fare per prima cosa e' mostrare che tutti gli operatori $T_{\vec{R}}$ (ce ne sono tanti quanti sono i vettori \vec{R} , cioe' infiniti) commutano³⁰ sia con l'operatore hamiltoniano che fra loro stessi.

Vediamo la commutazione con l'hamiltoniano (definito nell'eq. 490).

Data una funzione *qualsiasi* $\Psi(\vec{r})$, dobbiamo verificare che vale:

$$T_{\vec{R}}(H(\vec{r})\Psi(\vec{r})) = H(\vec{r})(T_{\vec{R}}\Psi(\vec{r})) \quad (503)$$

dove ho messo in evidenza che anche l'operatore hamiltoniano dipende da \vec{r} (eq. 490).

L'operatore di traslazione $T_{\vec{R}}$ trasla del vettore \vec{R} il suo argomento e quindi:

$$T_{\vec{R}}(H(\vec{r})\Psi(\vec{r})) = H(\vec{r} + \vec{R})\Psi(\vec{r} + \vec{R}) \quad (504)$$

E' semplice verificare che:

$$H(\vec{r} + \vec{R}) = H(\vec{r}) \quad (505)$$

Infatti:

$$H(\vec{r} + \vec{R}) = -\frac{\hbar^2}{2m}\nabla_{\vec{r}+\vec{R}}^2 + U(\vec{r} + \vec{R}) \quad (506)$$

²⁹Questa e' la dimostrazione a p. 134 del libro di Ashcroft e Mermin.⁷

³⁰Si dice che due operatori O e P commutano se l'ordine in cui vengono applicati e' indifferente, cioe' se $OP = PO$

Per ipotesi $U(\vec{r})$ e' periodico sul reticolo di Bravais e quindi:

$$U(\vec{r} + \vec{R}) = U(\vec{r}) \quad (507)$$

Inoltre, posti:

$$\begin{aligned} \vec{r} &= (xyz) \\ \vec{R} &= (R_x R_y R_z) \end{aligned}$$

per una generica funzione $f(\vec{r}) = f(xyz)$ si ha:

$$\nabla_{\vec{r} + \vec{R}}^2 f = \frac{\partial^2 f}{\partial (x + R_x)^2} + \frac{\partial^2 f}{\partial (y + R_y)^2} + \frac{\partial^2 f}{\partial (z + R_z)^2} \quad (508)$$

Prendiamo il primo termine al secondo membro; posto:

$$u(x) = x + R_x \quad (509)$$

per la regola della catena (chain rule) si ha ³¹:

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial u} \frac{\partial u}{\partial x} \quad (510)$$

Dato un certo operatore di traslazione $T_{\vec{R}}$, il vettore \vec{R} e' fissato e quindi le sue componenti scalari sono delle costanti; cio' significa che:

$$\frac{\partial u}{\partial x} = \frac{\partial}{\partial x} (x + R_x) = 1 \quad (511)$$

e quindi:

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial u} = \frac{\partial f}{\partial (x + R_x)} \quad (512)$$

Ma allora:

$$\frac{\partial^2 f}{\partial (x + R_x)^2} = \frac{\partial}{\partial (x + R_x)} \left(\frac{\partial f}{\partial (x + R_x)} \right) = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x} \right) = \frac{\partial^2 f}{\partial x^2} \quad (513)$$

³¹Nel caso piu' generale, se $f = f(u, v, w \dots)$ con $u = u(x, y, z \dots)$, $v = v(x, y, z \dots)$, $w = w(x, y, z \dots) \dots$ la regola della catena ha la forma;

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial f}{\partial v} \frac{\partial v}{\partial x} + \frac{\partial f}{\partial w} \frac{\partial w}{\partial x} + \dots$$

e lo stesso per le derivate parziali rispetto a $y, z \dots$.

Ma nel nostro caso $f = f(u, v, w)$ con u, v, w funzioni, rispettivamente, solo di x, y, z e quindi, ad esempio:

$$\frac{\partial v}{\partial x} = \frac{\partial w}{\partial x} = 0$$

e cosi' via.

Ovviamente lo stesso vale per gli altri due termini al secondo membro della 508 e quindi rimane provata la validita' della 505. E allora:

$$\begin{aligned} T_{\vec{R}}(H(\vec{r})\Psi(\vec{r})) &= H(\vec{r} + \vec{R})\Psi(\vec{r} + \vec{R}) = H(\vec{r})\Psi(\vec{r} + \vec{R}) \\ &= H(\vec{r})(T_{\vec{R}}\Psi(\vec{r})) \end{aligned} \quad (514)$$

che e' la commutativita' dei due operatori $T_{\vec{R}}$ e $H(\vec{r})$.

Dati due vettori di traslazione \vec{R} ed \vec{R}' , si vede banalmente che i due corrispondenti operatori commutano. Per una funzione qualsiasi $\Psi(\vec{r})$:

$$\begin{aligned} T_{\vec{R}}(T_{\vec{R}'}\Psi(\vec{r})) &= T_{\vec{R}}\Psi(\vec{r} + \vec{R}') = \Psi(\vec{r} + \vec{R}' + \vec{R}) \\ &= \Psi(\vec{r} + \vec{R} + \vec{R}') = T_{\vec{R}'}\Psi(\vec{r} + \vec{R}) = T_{\vec{R}'}(T_{\vec{R}}\Psi(\vec{r})) \end{aligned} \quad (515)$$

Abbiamo quindi dimostrato che tutti gli (infiniti) operatori $T_{\vec{R}}$ commutano sia con l'operatore hamiltoniano che fra loro stessi.

Alla luce di questo risultato, possiamo utilizzare un importante teorema della meccanica quantistica: per un insieme di operatori che commutano fra loro e' possibile trovare un insieme di funzioni che sono *contemporaneamente* autofunzioni di tutti gli operatori dell'insieme.

Quindi e' possibile trovare le autofunzioni dell'hamiltoniano H in modo tale che esse siano simultaneamente autofunzioni di tutti gli operatori di traslazione $T_{\vec{R}}$; cioe': detta $\Psi(\vec{r})$ la generica autofunzione dell'hamiltoniano, e' possibile sceglierla in modo tale che sia:

$$H\Psi(\vec{r}) = \epsilon\Psi(\vec{r}) \quad (516)$$

$$T_{\vec{R}}\Psi(\vec{r}) = C(\vec{R})\Psi(\vec{r}) \quad \forall \vec{R} \quad (517)$$

($\Psi(\vec{r})$ e' autofunzione di H e di tutti i $T_{\vec{R}}$, ma con autovalori diversi; in particolare, gli autovalori $C(\vec{R})$ degli operatori $T_{\vec{R}}$ dipenderanno dal particolare operatore, cioe' saranno funzioni di \vec{R})

Gli autovalori dei $T_{\vec{R}}$ sono legati fra loro, come si puo' vedere facilmente. Per due vettori di traslazione \vec{R} ed \vec{R}' si ha infatti:

$$T_{\vec{R}}(T_{\vec{R}'}\Psi(\vec{r})) = T_{\vec{R}}C(\vec{R}')\Psi(\vec{r}) = C(\vec{R})C(\vec{R}')\Psi(\vec{r}) \quad (518)$$

ma anche:

$$\begin{aligned} T_{\vec{R}}(T_{\vec{R}'}\Psi(\vec{r})) &= T_{\vec{R}}\Psi(\vec{r} + \vec{R}') = \Psi(\vec{r} + \vec{R}' + \vec{R}) \\ &= T_{\vec{R}+\vec{R}'}\Psi(\vec{r}) = C(\vec{R} + \vec{R}')\Psi(\vec{r}) \end{aligned} \quad (519)$$

(ovviamente, $\vec{R} + \vec{R}'$ e' sempre un vettore del reticolo di Bravais)
e quindi, per confronto:

$$C(\vec{R} + \vec{R}') = C(\vec{R}) C(\vec{R}') \quad (520)$$

I tre vettori \vec{a}_j che definiscono la cella primitiva del reticolo di Bravais sono anch'essi vettori di traslazione del reticolo e quindi anche per essi vale:

$$T_{\vec{a}_j} \Psi(\vec{r}) = C(\vec{a}_j) \Psi(\vec{r}) \quad (521)$$

Nel caso piu' generale, gli autovalori degli operatori che stiamo considerando sono numeri complessi. Inoltre non consideriamo il caso in cui un autovalore sia nullo: in tale circostanza, la corrispondente autofunzione sarebbe identicamente nulla e quindi non interessante:

$$T_{\vec{R}} \Psi(\vec{r}) = \Psi(\vec{r} + \vec{R}) = 0 \cdot \Psi(\vec{r}) \Rightarrow \Psi(\vec{r} + \vec{R}) = 0 \Rightarrow \Psi(\vec{r}) = 0 \quad \forall \vec{r} \quad (522)$$

Ora: per qualsiasi numero complesso non nullo e' sempre possibile trovare un numero complesso x_j tale che:

$$C(\vec{a}_j) = \exp(i2\pi x_j) \quad (523)$$

Infatti, esprimendo $C(\vec{a}_j)$ in forma polare:

$$C(\vec{a}_j) = \rho (\cos(\theta) + i \sin(\theta)) \quad (\rho > 0) \quad (524)$$

si ha, con l'identita' di Eulero:

$$\rho (\cos(\theta) + i \sin(\theta)) = \rho \exp(i\theta) = \exp(\ln(\rho)) \exp(i\theta) \quad (525)$$

$$= \exp(i\theta + \ln(\rho)) = \exp(i2\pi x_j) \quad (526)$$

in cui il valore di x_j si ricava dall'equazione:

$$i\theta + \ln(\rho) = i2\pi x_j \Rightarrow x_j = \frac{i\theta + \ln(\rho)}{i2\pi} = \frac{\theta - i \ln(\rho)}{2\pi} \quad (527)$$

(vedrete, tuttavia, che grazie all'uso di opportune condizioni al contorno, saremo in grado di assicurare che gli x_j siano numeri reali (punto 200))

Consideriamo ora il generico vettore di traslazione del reticolo di Bravais \vec{R} ; esso puo' essere espresso come una combinazione lineare con coefficienti interi dei tre vettori di cella a_j (eq. 222):

$$\vec{R} = n_1 \vec{a}_1 + n_2 \vec{a}_2 + n_3 \vec{a}_3$$

Vediamo allora come possiamo esprimere convenientemente l'autovalore corrispondente a \vec{R} basandoci sulla 520:

$$\begin{aligned}
C(\vec{R}) &= C(n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3) \\
&= C(\vec{a}_1 + (n_1 - 1)\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3) \\
&= C(\vec{a}_1) C((n_1 - 1)\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3) \text{ (grazie alla 520)} \\
&= C(\vec{a}_1) C(\vec{a}_1 + (n_1 - 2)\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3) \\
&= C(\vec{a}_1)^2 C((n_1 - 2)\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3) \text{ (grazie alla 520)} \\
&= \dots \\
&= C(\vec{a}_1)^{n_1} C(n_2\vec{a}_2 + n_3\vec{a}_3) \\
&= C(\vec{a}_1)^{n_1} C(\vec{a}_2 + (n_2 - 1)\vec{a}_2 + n_3\vec{a}_3) \\
&= C(\vec{a}_1)^{n_1} C(\vec{a}_2) C((n_2 - 1)\vec{a}_2 + n_3\vec{a}_3) \text{ (grazie alla 520)} \\
&= \dots \\
&= C(\vec{a}_1)^{n_1} C(\vec{a}_2)^{n_2} C(\vec{a}_3)^{n_3}
\end{aligned} \tag{528}$$

ovvero, in base alla 523:

$$\begin{aligned}
C(\vec{R}) &= \exp(i2\pi n_1 x_1) \exp(i2\pi n_2 x_2) \exp(i2\pi n_3 x_3) \\
&= \exp(i2\pi (n_1 x_1 + n_2 x_2 + n_3 x_3))
\end{aligned} \tag{529}$$

Lo step successivo consiste nel realizzare che il termine $2\pi (n_1 x_1 + n_2 x_2 + n_3 x_3)$ nell'esponenziale puo' essere riscritto semplicemente come:

$$n_1 x_1 + n_2 x_2 + n_3 x_3 = \vec{k} \cdot \vec{R} \tag{530}$$

con:

$$\vec{R} = n_1 \vec{a}_1 + n_2 \vec{a}_2 + n_3 \vec{a}_3 \tag{531}$$

$$\vec{k} = x_1 \vec{b}_1 + x_2 \vec{b}_2 + x_3 \vec{b}_3 \tag{532}$$

dove i b_j sono i tre vettori che definiscono la cella primitiva del reticolo reciproco (e quindi: $a_j \cdot b_j = 2\pi\delta_{ij}$, eq. 118). Notate che \vec{k} non e' (per il momento) un vettore del reticolo reciproco, perche' i coefficienti nella 532 sono numeri complessi qualsiasi.

Quindi:

$$T_{\vec{R}}\Psi(\vec{r}) = C(\vec{R})\Psi(\vec{r}) = \exp(i\vec{k} \cdot \vec{R})\Psi(\vec{r}) \tag{533}$$

In conclusione, abbiamo dimostrato che possiamo scegliere le autofunzioni $\Psi(\vec{r})$ dell'hamiltoniano 490 in modo tale che, per ogni traslazione \vec{R} del reticolo di Bravais, sia:

$$T_{\vec{R}}\Psi(\vec{r}) = \Psi(\vec{r} + \vec{R}) = \exp(i\vec{k} \cdot \vec{R})\Psi(\vec{r}) \tag{534}$$

che e' precisamente la tesi del teorema di Bloch, espressa nella forma 493.

192. Il teorema di Bloch e' un aiuto fondamentale per il problema che stiamo trattando:

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + U(\vec{r})\right)\Psi(\vec{r}) = \epsilon\Psi(\vec{r})$$

$$\left[U(\vec{r} + \vec{R}) = U(\vec{r}) \quad \forall \vec{R} \in \text{al reticolo di Bravais}\right]$$

Tale problema, tuttavia, necessita ancora della specifica delle condizioni al contorno: queste condizioni, come abbiamo gia' visto trattando il gas di Fermi (elettroni indipendenti e liberi) p. 91, fissano il comportamento che deve avere la soluzione (l'autofunzione dell'hamiltoniano) ai "confini" del sistema considerato.

193. Nel caso del gas di Fermi avevamo assunto che il sistema (il campione solido) fosse un cubo di lato L all'interno del quale gli elettroni si muovevano liberamente (senza risentire di alcun potenziale).

Nel caso che stiamo considerando ora, invece, il sistema e' un solido cristallino costituito dalla ripetizione lungo tre direzioni (non coplanari) di una cella primitiva. Per questo motivo, ora conviene considerare un campione macroscopico la cui forma e il cui volume siano *commensurati* a quelli della cella primitiva.

In altre parole, il sistema macroscopico che considereremo ora e' una *supercella* (punto 88) della cella primitiva della struttura cristallina del solido.

Detti $\vec{a}_j, (j=1,3)$ i tre vettori che definiscono la cella primitiva del reticolo di Bravais ed N_j tre corrispondenti numeri interi, la supercella che costituisce il campione macroscopico che consideriamo e' definita dai tre vettori:

$$\vec{A}_1 = N_1\vec{a}_1 \quad (535)$$

$$\vec{A}_2 = N_2\vec{a}_2 \quad (536)$$

$$\vec{A}_3 = N_3\vec{a}_3 \quad (537)$$

Notate che la supercella di cui stiamo parlando contiene un numero *enorme* di celle primitive; cioe' gli N_j sono numeri dell'ordine del numero di Avogadro (pensate a quante celle primitive sono contenute in un cristallo di dimensioni ordinarie, ad esempio un cristallo di glucosio o $NaCl$ che tutti potete trovare a casa vostra).

194. La situazione e' schematizzata in due dimensioni nella figura 50.

Ovviamente la supercella rappresentata nella figura non e' realistica per quanto riguarda le sue dimensioni rispetto a quelle della cella primitiva: quest'ultima e' stata replicata solo 10 volte lungo le due direzioni \vec{a}_1 e \vec{a}_2 .

Tuttavia guardando la figura dovrebbe risultare chiaro a tutti cosa intendo quando dico che la forma e il volume del solido macroscopico che consideriamo sono commensurati a quelli della cella primitiva.

195. Se il campione solido macroscopico e' una supercella ottenuta replicando la cella primitiva N_i volte lungo ciascun vettore di cella \vec{a}_i , il numero totale di celle primitive contenuto nel campione macroscopico e':

$$N = N_1N_2N_3 \quad (538)$$

Come gia' detto, l'ordine di grandezza di N e' quello del numero di Avogadro (10^{23}).

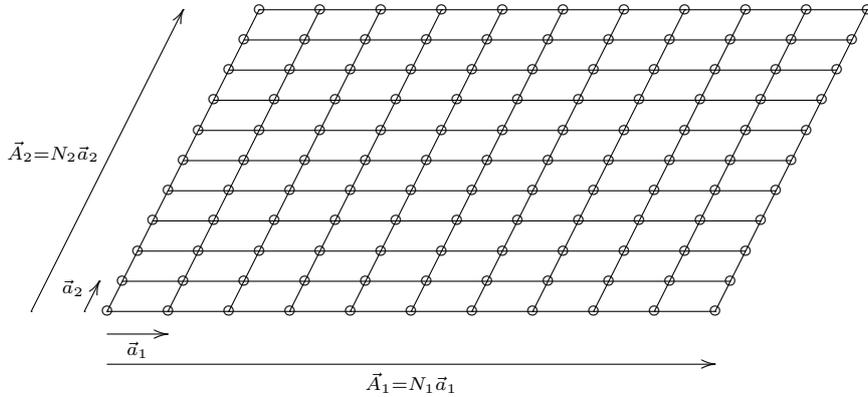


Figura 50:

Il volume v della cella primitiva determinata dai vettori \vec{a}_i e quello V della supercella sono legati dalla relazione:

$$V = N_1 N_2 N_3 v = Nv \quad (539)$$

196. Veniamo ora alle condizioni al contorno. Come già detto nel caso del gas di Fermi (p. 336), facciamo l'ipotesi che il sistema oggetto di studio sia sufficientemente esteso da poter considerare trascurabili tutti gli effetti "di bordo".

Sotto questa ipotesi le condizioni al contorno possono essere scelte sulla base di un criterio di pura convenienza matematica (proprio perché assumiamo che le proprietà del bulk del campione non siano influenzate in modo apprezzabile dalla configurazione della superficie).

E allora, in modo analogo a quanto avevamo già fatto per il gas di Fermi, utilizzeremo le condizioni al contorno periodiche di Born e Von Karman.

Siccome però ora il campione non è, in generale, di forma cubica (la forma del campione è quella, "scalata" di un fattore dell'ordine del numero di Avogadro, della cella primitiva della sua struttura cristallina, guardate la figura 50), le condizioni al contorno periodiche devono essere generalizzate come abbiamo visto al punto 47:

$$\Psi(\vec{r} + \vec{Q}) = \Psi(\vec{r}) \quad \forall \vec{Q} \quad (540)$$

con:

$$\vec{Q} = p_1 \vec{A}_1 + p_2 \vec{A}_2 + p_3 \vec{A}_3 \quad p_j \in \mathbb{Z} \quad (541)$$

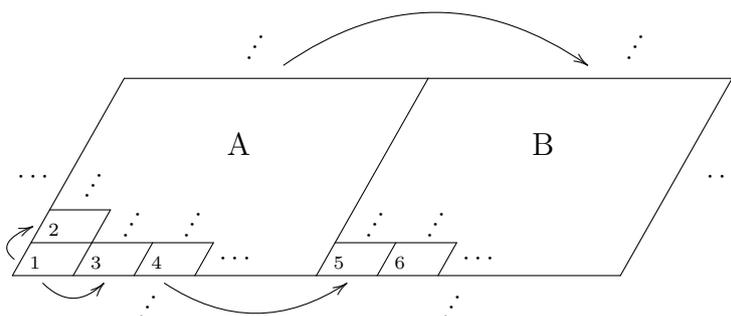
un generico vettore del reticolo di Bravais corrispondente alla *supercella*.

In altre parole, stiamo imponendo che la soluzione dell'equazione di Schrödinger $\Psi(\vec{r})$ sia periodica sul reticolo di Bravais corrispondente alla *supercella* (e questo è il senso dell'anticipazione che avevamo fatto a p. 135).

197. Attenti a non fare confusione: la $\Psi(\vec{r})$ non è periodica sul reticolo di Bravais della struttura cristallina del campione (definito dai tre vettori \vec{a}_i , $i = 1, 2, 3$): tale periodicità è posseduta dall'operatore dell'energia potenziale $U(\vec{r})$,

eq. 490, e dal fattore $u(\vec{r})$ che compare nel teorema di Bloch, eq. 493. Invece, ed e' puramente una conseguenza delle condizioni al contorno che stiamo adottando, la $\Psi(\vec{r})$ ha la periodicita' del reticolo di Bravais della supercella che costituisce il campione macroscopico, definita dai tre vettori \vec{A}_i , $i = 1, 3$, eq. 535–537.

Per chiarire ulteriormente, nella figura seguente sono rappresentate due repliche (A e B) della supercella che costituisce il campione macroscopico (assunto bidimensionale per semplicita').



Naturalmente, il campione e' uno solo (ad esempio la supercella A), ma l'applicazione delle condizioni al contorno periodiche di Born e Von Karman fa si che tutto funzioni come se il campione fosse ripetuto periodicamente all'infinito (la ripetizione delle repliche e' suggerita dai puntini nella figura).

All'interno di ciascuna delle due repliche sono rappresentate alcune celle primitive del reticolo di Bravais su cui e' basata la struttura cristallina del solido (1, ..., 6).

L'autofunzione dell'hamiltoniano $\Psi(\vec{r})$ e' definita su un'intera supercella (ad esempio A) e si ripete periodicamente quando passiamo in un'altra supercella (ad esempio B).

Il potenziale $U(\vec{r})$ e' definito su una cella primitiva del reticolo di Bravais (ad esempio fissatevi sulla cella primitiva 1) e si ripete periodicamente, come indicato dalle frecce, in tutte le altre celle primitive (2, ..., 6...); il "grafico"³² della $\Psi(\vec{r})$ e' diverso (nel senso che non si ripete) quando passiamo da una cella primitiva ad una qualsiasi altra cella primitiva.

Ovviamente, la porzione del "grafico" di $\Psi(\vec{r})$ contenuta, ad esempio, nella cella primitiva 1 e' identica a quella contenuta nella cella primitiva 5 perche' le due celle primitive sono immagini periodiche nel reticolo della supercella. Tuttavia, la porzione del "grafico" di $\Psi(\vec{r})$ contenuta nella cella primitiva 2 e' diversa, ad esempio, da quella contenuta nella cella primitiva 6.

198. Nel seguito, parleremo di *reticolo microscopico* per riferirci al reticolo definito dai vettori \vec{a}_i , e di *reticolo macroscopico* per riferirci al reticolo definito dai vettori \vec{A}_i . La situazione e' riassunta nella seguente tabella:

³²Nel caso 3D non si puo' parlare del grafico della $\Psi(\vec{r})$, per lo meno non nel senso usuale del termine: si tratta di una funzione di 3 variabili, e quindi il suo "grafico" richiederebbe uno spazio a 4 dimensioni.

	reticolo microscopico	reticolo macroscopico
vettori di cella	$\vec{a}_1, \vec{a}_2, \vec{a}_3$	$\vec{A}_1 = N_1\vec{a}_1, \vec{A}_2 = N_2\vec{a}_2, \vec{A}_3 = N_3\vec{a}_3$
volume di cella	v	$V = N_1N_2N_3v = Nv$
funz. periodiche	$U(\vec{r}), u(\vec{r})$	$\Psi(\vec{r})$

Dovrebbe essere chiaro che tutti i punti del reticolo macroscopico sono contemporaneamente punti del reticolo microscopico, mentre *non* tutti i punti del reticolo microscopico sono contemporaneamente punti del reticolo macroscopico (guardate la figura al punto 194).

199. Conseguenze delle condizioni al contorno periodiche sul vettore d'onda del teorema di Bloch (eq. 493).

Abbiamo visto che, per il teorema di Bloch, la $\Psi(\vec{r})$ deve soddisfare la condizione (eq. 493):

$$\Psi(\vec{r} + \vec{R}) = \exp(i\vec{k} \cdot \vec{R}) \Psi(\vec{r})$$

$\forall \vec{R} \in$ al reticolo di Bravais microscopico.

Il teorema di Bloch non pone alcuna limitazione al vettore d'onda \vec{k} , nel senso che esso puo' variare con continuita' sia in modulo che in direzione. Abbiamo infatti visto (eq. 532) che esso e' una combinazione lineare, con coefficienti in generale complessi, dei tre vettori che definiscono la cella primitiva del reticolo reciproco del reticolo microscopico (nel seguito: reticolo reciproco microscopico per brevitaa):

$$\vec{k} = x_1\vec{b}_1 + x_2\vec{b}_2 + x_3\vec{b}_3 \quad (542)$$

Tuttavia, come ora vedremo, l'adozione delle condizioni al contorno periodiche 540 fa si che solo certi vettori d'onda \vec{k} siano permessi (il numero dei vettori d'onda permessi rimarra' comunque infinito, ma sara' numerabile; cioe': i vettori d'onda permessi sono discreti).

Siccome i vettori $\vec{A}_j = N_j\vec{a}_j$, ($j=1,3$) sono vettori sia del reticolo macroscopico che di quello microscopico, la 493 e la 540 devono valere contemporaneamente:

$$\begin{aligned} \text{teorema di Bloch: eq. 493:} \quad \Psi(\vec{r} + N_j\vec{a}_j) &= \exp(i\vec{k} \cdot N_j\vec{a}_j) \Psi(\vec{r}) \\ \text{periodicitaa': eq. 540:} \quad \Psi(\vec{r} + N_j\vec{a}_j) &= \Psi(\vec{r}) \\ j &= 1, 3 \end{aligned}$$

Per confronto, se ne deduce che deve essere:

$$\exp(i\vec{k} \cdot N_j\vec{a}_j) = 1 \quad (543)$$

Sostituendo a \vec{k} l'espressione 532 si ha:

$$\exp(i\vec{k} \cdot N_j\vec{a}_j) = \exp(iN_j(x_1\vec{b}_1 + x_2\vec{b}_2 + x_3\vec{b}_3) \cdot \vec{a}_j) = \exp(iN_j 2\pi x_j) = 1 \quad (544)$$

(ricordate che $\vec{a}_i \cdot \vec{b}_j = \delta_{ij}$, eq. 118)

L'ultima uguaglianza puo' essere soddisfatta solo se l'argomento dell'esponenziale e' uguale a un multiplo intero di 2π , cioe':

$$N_j x_j = m_j \Rightarrow x_j = \frac{m_j}{N_j} \quad m_j \in \mathbb{Z} \quad (545)$$

200. Come preannunciato, le condizioni al contorno periodiche limitano i possibili vettori d'onda che compaiono nel teorema di Bloch a quelli dati da:

$$\vec{k} = \frac{m_1}{N_1} \vec{b}_1 + \frac{m_2}{N_2} \vec{b}_2 + \frac{m_3}{N_3} \vec{b}_3 \quad m_1, m_2, m_3 \in \mathbb{Z} \quad (546)$$

Vediamo anche che, come avevamo accennato a p. 140, i coefficienti che compaiono nell'espressione di \vec{k} devono essere numeri reali (perche' rapporto fra due numeri interi).

201. Osservazione: i vettori d'onda \vec{k} permessi sono tutti e soli i punti del *reticolo reciproco del reticolo macroscopico* (nel seguito: reticolo reciproco macroscopico, per brevità').

Al punto 193 e seguenti abbiamo detto che il campione solido macroscopico che consideriamo e' una supercella (definita dai vettori $\vec{A}_1 = N_1 \vec{a}_1, \vec{A}_2 = N_2 \vec{a}_2, \vec{A}_3 = N_3 \vec{a}_3$) della cella primitiva del reticolo microscopico. I vettori \vec{B}_i della cella primitiva del reticolo reciproco macroscopico (cioe' il reticolo reciproco del reticolo di Bravais definito dai vettori della supercella) sono dati da (eq. 112–114):

$$\vec{B}_1 = 2\pi \frac{N_2 \vec{a}_2 \times N_3 \vec{a}_3}{N_1 \vec{a}_1 \cdot (N_2 \vec{a}_2 \times N_3 \vec{a}_3)} = \frac{1}{N_1} 2\pi \frac{\vec{a}_2 \times \vec{a}_3}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} = \frac{1}{N_1} \vec{b}_1 \quad (547)$$

$$\vec{B}_2 = 2\pi \frac{N_3 \vec{a}_3 \times N_1 \vec{a}_1}{N_2 \vec{a}_1 \cdot (N_2 \vec{a}_2 \times N_3 \vec{a}_3)} = \frac{1}{N_2} 2\pi \frac{\vec{a}_3 \times \vec{a}_1}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} = \frac{1}{N_2} \vec{b}_2 \quad (548)$$

$$\vec{B}_3 = 2\pi \frac{N_1 \vec{a}_1 \times N_2 \vec{a}_2}{N_3 \vec{a}_1 \cdot (N_2 \vec{a}_2 \times N_3 \vec{a}_3)} = \frac{1}{N_3} 2\pi \frac{\vec{a}_1 \times \vec{a}_2}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} = \frac{1}{N_3} \vec{b}_3 \quad (549)$$

dove i \vec{b}_i sono i vettori della cella primitiva del reticolo reciproco microscopico. Se riguardate l'eq. 546, vi rendete conto che essa puo' essere riscritta come:

$$\vec{k} = m_1 \vec{B}_1 + m_2 \vec{B}_2 + m_3 \vec{B}_3 \quad m_1, m_2, m_3 \in \mathbb{Z} \quad (550)$$

il che significa proprio che i vettori d'onda permessi sono tutti e soli i punti del reticolo reciproco macroscopico.

202. Facciamo il punto della situazione.

Per un campione solido cristallino che e' una supercella ($N_1 \times N_2 \times N_3$) della cella primitiva del corrispondente reticolo di Bravais, stiamo studiando l'equazione di Schrödinger stazionaria e monoelettronica, con condizioni al contorno periodiche:

$$\left(-\frac{\hbar^2}{2m} \nabla^2 + U(\vec{r}) \right) \Psi(\vec{r}) = \epsilon \Psi(\vec{r})$$

$$U(\vec{r} + \vec{R}) = U(\vec{r}) \quad \forall \vec{R} \in \text{al reticolo microscopico}$$

$$\Psi(\vec{r} + \vec{Q}) = \Psi(\vec{r}) \quad \forall \vec{Q} \in \text{al reticolo macroscopico}$$

Grazie al teorema di Bloch e alle condizioni al contorno periodiche possiamo dire che la soluzione deve essere del tipo:

$$\Psi_{\vec{k}n}(\vec{r}) = \exp(i\vec{k} \cdot \vec{r}) u_{\vec{k}n}(\vec{r}) \quad (551)$$

con $u_{\vec{k}n}(\vec{r})$ avente la stessa periodicit  del reticolo microscopico e \vec{k} appartenente al reticolo reciproco macroscopico.

Notate come si   parzialmente chiarita l'indicizzazione delle soluzioni (che avevamo anticipato nell'equazione 491): ogni soluzione   indicizzata da un vettore d'onda permesso \vec{k} , ovvero dai tre numeri interi che lo definiscono (eq. 550). Resta ancora da chiarire la natura e il significato dell'ulteriore indice n : per questo dovrete aspettare ancora un po' (punto 219).

203. Confronto dei due reticoli reciproci.

Abbiamo introdotto il reticolo microscopico e il reticolo macroscopico; il primo   il reticolo di Bravais determinato dai vettori \vec{a}_i della cella primitiva su cui   basata la struttura cristallina del solido; il secondo   il reticolo di Bravais determinato dai vettori $\vec{A}_i = N_i \vec{a}_i$ della supercella che costituisce il campione macroscopico. I due reticoli diretti sono in stretta relazione, nel senso che il reticolo macroscopico (il meno denso)   "immerso" in quello microscopico (il pi  denso): cio  tutti i punti del reticolo macroscopico fanno contemporaneamente parte di quello microscopico (mentre non   vero il viceversa).

A ciascuno dei due reticoli diretti   associato un reticolo reciproco.

Siccome il volume della cella "primitiva" del reticolo macroscopico (cio  la supercella)   un multiplo pari a $N = N_1 N_2 N_3$ di quello della cella primitiva del reticolo microscopico, i volumi delle celle primitive dei corrispondenti reticoli reciproci stanno fra loro nel rapporto inverso; cio  detto V^* il volume della cella primitiva del reticolo reciproco macroscopico e v^* quello della cella primitiva del reticolo reciproco microscopico, si ha:

$$V^* = \frac{v^*}{N} \quad (552)$$

Cio' si puo' vedere anche in modo pi  formale.

Come abbiamo visto in generale (eq. 244), il volume v_{dir} della cella primitiva del reticolo diretto e quello v_{rec} della cella primitiva del corrispondente reticolo reciproco sono legati da:

$$v_{rec} = \frac{(2\pi)^3}{v_{dir}} \quad (553)$$

Per il reticolo microscopico e quello macroscopico di cui stiamo parlando sara':

$$v^* = \frac{(2\pi)^3}{v} \quad (554)$$

e

$$V^* = \frac{(2\pi)^3}{V} \quad (555)$$

Ma al punto 195 abbiamo visto che:

$$V = Nv \quad (556)$$

e quindi, sostituendo, si ricava la relazione mostrata sopra.
 Notate come:

cella diretta piu' grande \Rightarrow cella reciproca piu' piccola

e viceversa (n -esima relazione di reciprocita').

204. Per il fatto che una cella primitiva contiene un solo punto reticolare e in base alla relazione:

$$v^* = NV^* \quad (557)$$

si conclude che la cella primitiva del reticolo reciproco microscopico contiene $N = N_1N_2N_3$ punti del reticolo reciproco macroscopico.

Siccome i vettori d'onda permessi sono i punti del reticolo reciproco macroscopico, quanto sopra equivale a dire che la cella primitiva del reticolo reciproco microscopico contiene N vettori d'onda \vec{k} permessi.

Ricordate che $N = N_1N_2N_3$ e' il numero di celle primitive del reticolo microscopico contenute nel campione solido macroscopico.

Il volume dello spazio reciproco per punto \vec{k} permesso e' chiaramente il volume della cella primitiva del reticolo reciproco macroscopico, cioe':

$$\Delta\vec{k} = \frac{(2\pi)^3}{V} \quad (558)$$

dove V e' il volume del campione.

Questo e' esattamente lo stesso risultato che avevamo gia' trovato per il gas di Fermi (p. 353).

205. Le relazioni fra i vettori delle celle primitive dei due reticoli reciproci (quello microscopico e quello macroscopico) viste al punto 201 possono essere riscritte nel modo seguente:

$$\vec{b}_1 = N_1\vec{B}_1 \quad (559)$$

$$\vec{b}_2 = N_2\vec{B}_2 \quad (560)$$

$$\vec{b}_3 = N_3\vec{B}_3 \quad (561)$$

Da cio' si vede che la cella primitiva del reticolo reciproco microscopico e' una *supercella* ($N_1 \times N_2 \times N_3$) di quella del reticolo reciproco macroscopico.

Notate la simmetria: al punto 193 avevamo visto che la cella primitiva del reticolo macroscopico e' una supercella ($N_1 \times N_2 \times N_3$) di quella del reticolo microscopico; quando si considerano le celle primitive dei corrispondenti reticoli reciproci, i ruoli vengono esattamente invertiti.

La situazione in due dimensioni e' illustrata nella figura 51, dove sono mostrate 6 celle primitive del reticolo reciproco microscopico ($A \dots F$) i cui vettori generatori sono \vec{b}_1 e \vec{b}_2 ; e' delineata anche una cella primitiva del reticolo reciproco macroscopico (in alto a sinistra) i cui vettori generatori sono $\vec{b}_1/4$ e $\vec{b}_2/3$ ³³.

³³Come gia' osservato per la figura riguardante le celle dei reticoli diretti (punto 194), il rapporto fra le dimensioni delle due celle reciproche in questa figura non e' realistico (tenete presente che $N = N_1N_2N_3$ e' un numero dell'ordine del numero di Avogadro).

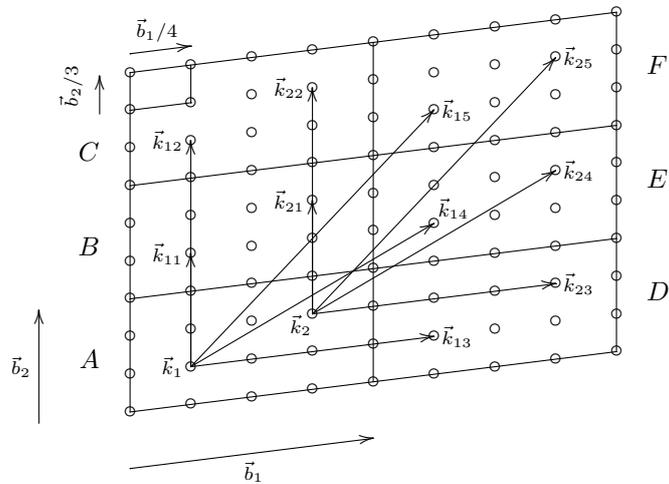


Figura 51:

Quindi nell'esempio bidimensionale della figura, la cella primitiva del reticolo reciproco microscopico e' una supercella (4×3) della cella primitiva del reticolo reciproco macroscopico; la supercella contiene quindi $4 \times 3 = 12$ celle primitive del reticolo reciproco macroscopico, ovvero 12 punti \vec{k} permessi.

Come gia' osservato per i reticoli diretti (punto 198), tutti i punti del reticolo reciproco microscopico (cella primitiva piu' grande) sono anche punti del reticolo reciproco macroscopico (cella primitiva piu' piccola); non e' vero il contrario: cioe' *non tutti* i punti del reticolo reciproco macroscopico sono anche punti del reticolo reciproco microscopico.

206. A questo proposito e' opportuno fare la seguente semplice osservazione (che ci servira' in seguito, punto 216): tutti i punti \vec{k} permessi (cioe' tutti i punti del reticolo reciproco macroscopico, che e' il piu' denso), possono essere raggiunti a partire da quelli contenuti in un'unica cella primitiva del reticolo reciproco microscopico (il meno denso, generato dalla supercella reciproca) sommando a ciascuno di essi tutti i vettori del reticolo reciproco microscopico.

Nella figura e' esemplificato questo procedimento per i due vettori d'onda permessi \vec{k}_1 e \vec{k}_2 nella cella A.

Sommando al punto \vec{k}_1 il vettore del reticolo reciproco microscopico \vec{b}_2 si ottiene la sua immagine \vec{k}_{11} nella cella B. Sommando il vettore del reticolo reciproco microscopico $2\vec{b}_2$ si ottiene la sua immagine \vec{k}_{12} nella cella C e cosi' via. Riporto di seguito gli esempi mostrati nella figura:

$$\begin{aligned}
\vec{k}_1 + (\vec{b}_2) &\rightarrow \vec{k}_{11} \\
\vec{k}_1 + (2\vec{b}_2) &\rightarrow \vec{k}_{12} \\
\vec{k}_1 + (\vec{b}_1) &\rightarrow \vec{k}_{13} \\
\vec{k}_1 + (\vec{b}_1 + \vec{b}_2) &\rightarrow \vec{k}_{14} \\
\vec{k}_1 + (\vec{b}_1 + 2\vec{b}_2) &\rightarrow \vec{k}_{15}
\end{aligned}$$

Ovviamente lo stesso vale per il punto \vec{k}_2 :

$$\begin{aligned}
\vec{k}_2 + (\vec{b}_2) &\rightarrow \vec{k}_{21} \\
\vec{k}_2 + (2\vec{b}_2) &\rightarrow \vec{k}_{22} \\
\vec{k}_2 + (\vec{b}_1) &\rightarrow \vec{k}_{23} \\
\vec{k}_2 + (\vec{b}_1 + \vec{b}_2) &\rightarrow \vec{k}_{24} \\
\vec{k}_2 + (\vec{b}_1 + 2\vec{b}_2) &\rightarrow \vec{k}_{25}
\end{aligned}$$

e per qualsiasi altro punto \vec{k} della cella A .

207. Un altro modo di vedere la stessa cosa e' il seguente.

Il reticolo reciproco microscopico e' un reticolo di Bravais; la sua cella primitiva contiene $N = N_1 N_2 N_3$ punti \vec{k} permessi (punti del reticolo reciproco macroscopico); tali punti possono essere considerati come la *base* associata ad ogni punto del reticolo reciproco microscopico (punto 80): ogni elemento della base ha pertanto un numero infinito di immagini ad esso associate dalle infinite traslazioni del reticolo reciproco microscopico. L'unica "peculiarita'" di questo scenario e' che, mentre in generale gli atomi che costituiscono la base associata ai punti di un reticolo di Bravais sono distribuiti in modo qualsiasi, i punti \vec{k} che costituiscono la base associata ai punti del reticolo reciproco microscopico costituiscono a loro volta un reticolo (il reticolo reciproco macroscopico) commensurato al primo.

208. Da quanto detto segue che l'insieme infinito dei punti del reticolo reciproco macroscopico (cioe' i vettori d'onda permessi) puo' essere ripartito in $N = N_1 N_2 N_3$ sottoinsiemi (anch'essi infiniti) indicizzati dai punti \vec{k} contenuti in una singola cella del reticolo reciproco microscopico; ogni sottoinsieme e' costituito da un punto \vec{k} della cella di riferimento e da tutte le sue infinite immagini (repliche) ottenute per traslazione nel reticolo reciproco microscopico.

In altre parole, introducendo i seguenti simboli:

\mathcal{R}_m reticolo reciproco microscopico
 \mathcal{C} cella primitiva fissata di \mathcal{R}_m
 \vec{G} generico vettore di \mathcal{R}_m

i punti \vec{k} del sottoinsieme j -esimo di questa ripartizione sono definiti da:

$$\vec{k} : \vec{k} = \vec{k}_j + \vec{G} \quad \vec{k}_j \in \mathcal{C} \text{ e } \vec{G} \in \mathcal{R}_m \quad (562)$$

Inoltre, i sottoinsiemi così ottenuti sono disgiunti, cioè nessuno di essi ha punti in comune con alcun'altro sottoinsieme. In altre parole, se un punto \vec{k} appartiene a un certo sottoinsieme, esso non può appartenere contemporaneamente ad alcun altro sottoinsieme (l'intersezione di due qualsiasi sottoinsiemi è nulla). Ad esempio, con riferimento alla figura 51, il punto \vec{k}_{24} appartiene al sottoinsieme generato a partire dal punto \vec{k}_2 nella cella A : esso può essere raggiunto sommando la traslazione $(\vec{b}_1 + \vec{b}_2)$ a \vec{k}_2 ; esso *non* può essere raggiunto mediante traslazioni del reticolo reciproco microscopico a partire da alcun altro punto nella cella A (ad esempio è impossibile raggiungere \vec{k}_{24} con traslazioni del reticolo reciproco microscopico a partire dal punto \vec{k}_1).

209. Un'altra osservazione a questo riguardo è la seguente: per ogni coppia di punti \vec{k} appartenente ad uno *stesso* sottoinsieme esiste sempre un vettore traslazione del reticolo reciproco microscopico che connette i due membri della coppia; viceversa, dati due punti \vec{k} appartenenti a due sottoinsiemi *diversi*, *non esiste* alcun vettore traslazione del reticolo reciproco microscopico che li connetta.

L'equazione centrale

210. Ricaveremo ora un'equazione, che Kittel³ chiama *equazione centrale*, e che, almeno in linea di principio, fornisce la soluzione dell'equazione di Schrödinger che stiamo studiando.

Inoltre, come vedremo, la derivazione dell'equazione centrale offre la possibilità di una seconda dimostrazione del teorema di Bloch (punto 222).

Il layout del procedimento è il seguente.

Siccome il potenziale $U(\vec{r})$ ha la stessa periodicità del reticolo microscopico, lo possiamo sviluppare in serie di Fourier sui punti del reticolo reciproco microscopico (riguardate la sezione sulle serie di Fourier e il punto 32 in particolare).

In modo analogo, siccome la $\Psi(\vec{r})$ ha la stessa periodicità del reticolo macroscopico, possiamo svilupparla in serie di Fourier sui punti del reticolo reciproco macroscopico (che sono i vettori d'onda \vec{k} permessi, punto 201).

Sostituiremo i due sviluppi di Fourier nell'equazione di Schrödinger e ricaveremo un'equazione (l'equazione centrale, appunto) dalla cui soluzione è possibile ricavare i coefficienti dello sviluppo in serie della $\Psi(\vec{r})$, che è il nostro obiettivo. Oltre ai coefficienti della funzione d'onda, l'equazione centrale conterrà anche il corrispondente autovalore, che costituisce un'altra incognita da trovare.

In realtà l'equazione che otterremo è un sistema omogeneo di dimensione (rigorosamente) infinita, quindi un oggetto matematico non facile da trattare. Tuttavia l'equazione centrale rappresenta uno strumento importantissimo sia dal punto di vista pratico (nel senso che rappresenta il punto di partenza per molti metodi computazionali) sia dal punto di vista teorico.

211. Come detto, essendo periodico sul reticolo microscopico, possiamo sviluppare il potenziale $U(\vec{r})$ in serie di Fourier nel modo seguente:³⁴

³⁴In questa parte userò la notazione di Kittel e Ibach e Lüth, secondo la quale i punti del reticolo reciproco microscopico sono indicati con \vec{G} e quelli del reticolo reciproco macroscopico con \vec{k} .

$$U(\vec{r}) = \sum_{\vec{G}} U_{\vec{G}} \exp(i\vec{G} \cdot \vec{r}) \quad (563)$$

dove la somma e' estesa a tutti i vettori \vec{G} del reticolo reciproco microscopico:

$$\vec{G} = m_1 \vec{b}_1 + m_2 \vec{b}_2 + m_3 \vec{b}_3 \quad m_1, m_2, m_3 \in \mathbb{Z} \quad (564)$$

($\vec{b}_1, \vec{b}_2, \vec{b}_3$ sono i vettori della cella primitiva del reticolo reciproco microscopico)
Gli $U_{\vec{G}}$ sono i coefficienti dello sviluppo in serie di Fourier (eq. 128):

$$U_{\vec{G}} = \frac{1}{v} \int_{cell} U(\vec{r}) \exp(-i\vec{G} \cdot \vec{r}) d\vec{r} \quad (565)$$

e l'integrale e' calcolato sul volume della cella primitiva del reticolo microscopico. In generale, i coefficienti $U(\vec{G})$ sono dei dati noti del problema, visto che si suppone che il potenziale $U(\vec{r})$ sia noto³⁵.

212. Essendo periodica sul reticolo macroscopico, possiamo scrivere anche per la funzione d'onda uno sviluppo di Fourier:

$$\Psi(\vec{r}) = \sum_{\vec{k}} C_{\vec{k}} \exp(i\vec{k} \cdot \vec{r}) \quad (566)$$

dove la somma e' estesa a tutti i vettori \vec{k} del reticolo reciproco macroscopico. I coefficienti dello sviluppo $C_{\vec{k}}$ sono delle incognite da trovare: una volta trovati i coefficienti, abbiamo trovato la funzione d'onda tramite la sommatoria scritta sopra.

213. Ora sostituiamo gli sviluppi di Fourier nell'equazione di Schrödinger:

$$\left(-\frac{\hbar^2}{2m} \nabla^2 + U(\vec{r}) \right) \Psi(\vec{r}) = \epsilon \Psi(\vec{r}) \quad (567)$$

Sara':

$$\left(-\frac{\hbar^2}{2m} \nabla^2 + \sum_{\vec{G}} U_{\vec{G}} \exp(i\vec{G} \cdot \vec{r}) \right) \sum_{\vec{k}} C_{\vec{k}} \exp(i\vec{k} \cdot \vec{r}) = \epsilon \sum_{\vec{k}} C_{\vec{k}} \exp(i\vec{k} \cdot \vec{r}) \quad (568)$$

Per quanto riguarda l'azione del primo termine dell'hamiltoniano sulla sommatoria su \vec{k} si ha:

$$-\frac{\hbar^2}{2m} \nabla^2 \left(\sum_{\vec{k}} C_{\vec{k}} \exp(i\vec{k} \cdot \vec{r}) \right) = \sum_{\vec{k}} -\frac{\hbar^2}{2m} C_{\vec{k}} \nabla^2 \exp(i\vec{k} \cdot \vec{r}) \quad (569)$$

Ora:

³⁵In realta', anche la definizione del potenziale $U(\vec{r})$ e' tutt'altro che semplice, visto che normalmente si tratta, come ho accennato al punto 184, di un potenziale efficace che deve cercare di compensare molte delle approssimazioni senza le quali il trattamento di questo problema sarebbe semplicemente impossibile; tuttavia, da un punto di vista di principio, $U(\vec{r})$ si assume noto in questo contesto.

$$\begin{aligned}
\nabla^2 \exp(i\vec{k} \cdot \vec{r}) &= \\
\frac{\partial^2}{\partial x^2} \exp(i(k_x x + k_y y + k_z z)) + \frac{\partial^2}{\partial y^2} \exp(i(k_x x + k_y y + k_z z)) + \\
\frac{\partial^2}{\partial z^2} \exp(i(k_x x + k_y y + k_z z)) &= \\
ik_x \frac{\partial}{\partial x} \exp(i(k_x x + k_y y + k_z z)) + ik_y \frac{\partial}{\partial y} \exp(i(k_x x + k_y y + k_z z)) + \\
ik_z \frac{\partial}{\partial z} \exp(i(k_x x + k_y y + k_z z)) &= \\
(ik_x)^2 \exp(i(k_x x + k_y y + k_z z)) + (ik_y)^2 \exp(i(k_x x + k_y y + k_z z)) + \\
(ik_z)^2 \exp(i(k_x x + k_y y + k_z z)) &= \\
-(k_x^2 + k_y^2 + k_z^2) \exp(i\vec{k} \cdot \vec{r}) &= -k^2 \exp(i\vec{k} \cdot \vec{r}) \quad (570)
\end{aligned}$$

In definitiva, sostituendo sopra:

$$= \sum_{\vec{k}} \frac{\hbar^2}{2m} k^2 C_{\vec{k}} \exp(i\vec{k} \cdot \vec{r}) \quad (571)$$

Ora vediamo l'azione del secondo termine dell'hamiltoniano sulla funzione d'onda.

Osservazione

Vale:

$$\Psi(\vec{r}) = \sum_{\vec{k}} C_{\vec{k}} \exp(i\vec{k} \cdot \vec{r}) = \sum_{\vec{k}} C_{\vec{k}-\vec{G}} \exp(i(\vec{k}-\vec{G}) \cdot \vec{r}) \quad (572)$$

dove \vec{G} e' un vettore (fissato) del reticolo reciproco microscopico.

La giustificazione e' semplice.

Siccome l'indice della sommatoria varia su tutti gli infiniti punti \vec{k} e siccome il vettore \vec{G} connette sempre due punti \vec{k} (appartenenti allo stesso sottoinsieme generato da un punto \vec{k} in una cella primitiva del reticolo reciproco microscopico, per quanto visto al punto 209), le due sommatorie contengono esattamente gli stessi addendi.

In altre parole, l'addendo j -esimo:

$$C_{\vec{k}_j} \exp(i\vec{k}_j \cdot \vec{r}) \quad (573)$$

si ottiene quando $\vec{k} = \vec{k}_j$ nella prima sommatoria oppure, equivalentemente, quando $k = \vec{k}_j + \vec{G}$ nella seconda.

Allora, per il motivo che apparira' chiaro fra un attimo, nel valutare l'azione del secondo termine dell'hamiltoniano sulla funzione d'onda conviene scrivere il suo sviluppo di Fourier come il secondo membro dell'uguaglianza 572.

Quindi:

$$\begin{aligned}
& \sum_{\vec{G}} U_{\vec{G}} \exp(i\vec{G} \cdot \vec{r}) \left(\sum_{\vec{k}} C_{\vec{k}-\vec{G}} \exp(i(\vec{k}-\vec{G}) \cdot \vec{r}) \right) \\
&= \sum_{\vec{k}} \sum_{\vec{G}} U_{\vec{G}} \exp(i\vec{G} \cdot \vec{r}) C_{\vec{k}-\vec{G}} \exp(i(\vec{k}-\vec{G}) \cdot \vec{r}) \\
&= \sum_{\vec{k}} \sum_{\vec{G}} U_{\vec{G}} C_{\vec{k}-\vec{G}} \exp(i\vec{k} \cdot \vec{r})
\end{aligned} \tag{574}$$

Quindi l'equazione di Schrödinger diventa:

$$\sum_{\vec{k}} \frac{\hbar^2}{2m} k^2 C_{\vec{k}} \exp(i\vec{k} \cdot \vec{r}) + \sum_{\vec{k}} \sum_{\vec{G}} U_{\vec{G}} C_{\vec{k}-\vec{G}} \exp(i\vec{k} \cdot \vec{r}) = \epsilon \sum_{\vec{k}} C_{\vec{k}} \exp(i\vec{k} \cdot \vec{r}) \tag{575}$$

ovvero:

$$\sum_{\vec{k}} \exp(i\vec{k} \cdot \vec{r}) \left\{ \left(\frac{\hbar^2}{2m} k^2 - \epsilon \right) C_{\vec{k}} + \sum_{\vec{G}} U_{\vec{G}} C_{\vec{k}-\vec{G}} \right\} = 0 \tag{576}$$

(Vedete che l'aver rimaneggiato la sommatoria come mostrato sopra consente ora di raccogliere a fattor comune il termine $\exp(i\vec{k} \cdot \vec{r})$)

Notate che quella appena ottenuta non e' altro che una "riedizione" dell'equazione di Schrödinger "nello spazio dei coefficienti di Fourier"; cioe' abbiamo trasformato il problema dell'integrazione di un'equazione differenziale in cui l'incognita da trovare e' una funzione, nel problema algebrico della risoluzione di un'equazione in cui le incognite sono i coefficienti $C_{\vec{k}}$ dello sviluppo in serie della $\Psi(\vec{r})$ e il corrispondente autovalore ϵ .

Il primo membro dell'equazione e' costituito da una somma (infinita) su tutti i vettori \vec{k} del reticolo reciproco macroscopico.

Ciascun termine della sommatoria nell'eq. 576 e' il prodotto fra un'onda piana dello sviluppo di Fourier della $\Psi(\vec{r})$ e il termine fra parentesi.

Vediamo ora come l'eq. 576 implichi che tutti i fattori che moltiplicano le onde piane:

$$\left\{ \left(\frac{\hbar^2}{2m} k^2 - \epsilon \right) C_{\vec{k}} + \sum_{\vec{G}} U_{\vec{G}} C_{\vec{k}-\vec{G}} \right\} \tag{577}$$

debbano essere necessariamente nulli.

Fissato un particolare vettore \vec{k}' del reticolo reciproco macroscopico, moltiplico ambo i membri della 576 per $\exp(-i\vec{k}' \cdot \vec{r})$ ed integro sul volume V della cella primitiva del reticolo macroscopico (che e' semplicemente il volume del campione):

$$\int_V \sum_{\vec{k}} \exp(i\vec{k} \cdot \vec{r}) \exp(-i\vec{k}' \cdot \vec{r}) \left\{ \left(\frac{\hbar^2}{2m} k^2 - \epsilon \right) C_{\vec{k}} + \sum_{\vec{G}} U_{\vec{G}} C_{\vec{k}-\vec{G}} \right\} d\vec{r} = 0 \quad (578)$$

Ora scambio di posto l'integrale e la sommatoria e osservo che il termine fra parentesi non dipende da \vec{r} :

$$\sum_{\vec{k}} \left\{ \left(\frac{\hbar^2}{2m} k^2 - \epsilon \right) C_{\vec{k}} + \sum_{\vec{G}} U_{\vec{G}} C_{\vec{k}-\vec{G}} \right\} \int_V \exp(i(\vec{k} - \vec{k}') \cdot \vec{r}) d\vec{r} = 0 \quad (579)$$

Quando $\vec{k} = \vec{k}'$ l'integrale e' banalmente uguale a V ; posso allora "estrarre" dalla sommatoria il termine corrispondente a $\vec{k} = \vec{k}'$:

$$\begin{aligned} & V \left\{ \left(\frac{\hbar^2}{2m} k'^2 - \epsilon \right) C_{\vec{k}'} + \sum_{\vec{G}} U_{\vec{G}} C_{\vec{k}'-\vec{G}} \right\} \\ & + \sum_{\vec{k} \neq \vec{k}'} \left\{ \left(\frac{\hbar^2}{2m} k^2 - \epsilon \right) C_{\vec{k}} + \sum_{\vec{G}} U_{\vec{G}} C_{\vec{k}-\vec{G}} \right\} \int_V \exp(i(\vec{k} - \vec{k}') \cdot \vec{r}) d\vec{r} = 0 \end{aligned} \quad (580)$$

E ora dimostriamo che:

$$\int_V \exp(i(\vec{k} - \vec{k}') \cdot \vec{r}) d\vec{r} = \int_V \exp(i\vec{K} \cdot \vec{r}) d\vec{r} = 0 \quad \forall \vec{K} \neq 0 \quad (581)$$

con $\vec{K} = \vec{k} - \vec{k}' \neq 0$.

Per prima cosa realizzate che $\vec{K} = \vec{k} - \vec{k}'$ e' sempre un vettore del reticolo reciproco macroscopico (in generale, e' banale rendersi conto che qualsiasi combinazione lineare con coefficienti interi di vettori di un reticolo di Bravais e' ancora un vettore del reticolo di Bravais).

La dimostrazione della 581 in una sola dimensione e' molto semplice.

In tal caso, infatti, la cella primitiva del reticolo macroscopico e' un segmento dell'asse reale di lunghezza L ; la variabile \vec{r} diventa una variabile scalare x e il generico punto non nullo \vec{K} del reticolo reciproco macroscopico diventa lo scalare $2\pi n/L$ (punto 31), con $n \neq 0$:

$$V \longrightarrow L \quad (582)$$

$$\vec{r} \longrightarrow x \quad (583)$$

$$\vec{K} \longrightarrow \frac{2\pi}{L} n \quad (n \neq 0) \quad (584)$$

Allora si ha:

$$\begin{aligned} \int_V \exp(i\vec{K} \cdot \vec{r}) d\vec{r} &\longrightarrow \int_0^L \exp\left(i\frac{2\pi}{L}nx\right) dx \\ &= \int_0^L \cos\left(\frac{2\pi}{L}nx\right) dx + i \int_0^L \sin\left(\frac{2\pi}{L}nx\right) dx = 0 \end{aligned} \quad (585)$$

dove l'ultima uguaglianza dovrebbe risultarvi banale (ricordate che $n \neq 0$). La dimostrazione in 3D e' quella offerta da Ashcroft e Mermin:¹⁷ per ammissione degli autori non e' rigorosa, ma e' convincente. Essendo un termine dello sviluppo in serie di Fourier della funzione $\Psi(\vec{r})$ periodica sul reticolo di Bravais macroscopico, l'onda piana $\exp(i\vec{K} \cdot \vec{r})$ e' essa stessa periodica su tale reticolo (punto 52). Da cio' segue che l'integrale:

$$\int_V \exp(i\vec{K} \cdot \vec{r}) d\vec{r} \quad (586)$$

e' indipendente dalla particolare cella primitiva su cui viene calcolato (ricordate che un reticolo di Bravais e' descrivibile con infinite celle primitive equivalenti e aventi tutte lo stesso volume V). Allora, detta C una cella primitiva e C' un'altra cella primitiva ottenuta per traslazione della cella C di un vettore non nullo qualsiasi \vec{d} (non necessariamente un vettore del reticolo di Bravais macroscopico), dovra' essere (uso C e C' come indici degli integrali per sottolineare che sono effettuati sulle due celle primitive: il volume e' ovviamente sempre V):

$$\int_C \exp(i\vec{K} \cdot \vec{r}) d\vec{r} = \int_{C'} \exp(i\vec{K} \cdot \vec{r}) d\vec{r} \quad (587)$$

(il fatto che una cella primitiva possa essere traslata *arbitrariamente* lo abbiamo visto in generale alla proprieta' 2 di p. 59)

Ora realizzate che, se C' e' stata ottenuta con una traslazione $\pm\vec{d}$ di C , allora l'integrale su C' deve essere uguale all'integrale su C dell'onda piana traslata di $\mp\vec{d}$ (notate l'inversione dei segni: \pm e \mp), cioe':

$$\int_{C'} \exp(i\vec{K} \cdot \vec{r}) d\vec{r} = \int_C \exp(i\vec{K} \cdot (\vec{r} \mp \vec{d})) d\vec{r} \quad (588)$$

(in altre parole: invece che traslare la cella di d in una certa direzione, traslo la funzione da integrare di d nella direzione *opposta*)

Combinando l'ultima equazione con la 587 si ottiene:

$$\begin{aligned} \int_C \exp(i\vec{K} \cdot \vec{r}) d\vec{r} &= \int_C \exp(i\vec{K} \cdot (\vec{r} \mp \vec{d})) d\vec{r} \\ &= \int_C \exp(i\vec{K} \cdot \vec{r}) \exp(i\vec{K} \cdot (\mp\vec{d})) d\vec{r} \\ &= \exp(i\vec{K} \cdot (\mp\vec{d})) \int_C \exp(i\vec{K} \cdot \vec{r}) d\vec{r} \end{aligned}$$

ovvero:

$$\left(1 - \exp\left(i\vec{K} \cdot (\mp\vec{d})\right)\right) \int_C \exp\left(i\vec{K} \cdot \vec{r}\right) d\vec{r} = 0 \quad (589)$$

Siccome $\vec{K} \neq 0$ e $\vec{d} \neq 0$, il fattore che moltiplica l'integrale non puo' essere nullo, e quindi deve essere nullo l'integrale.

Questo significa che, come annunciato, tutti i termini della sommatoria nella 580 sono nulli e quindi:

$$\left\{ \left(\frac{\hbar^2}{2m} k'^2 - \epsilon \right) C_{\vec{k}'} + \sum_{\vec{G}} U_{\vec{G}} C_{\vec{k}' - \vec{G}} \right\} = 0 \quad (590)$$

e siccome k' era arbitrario, si ottiene il risultato finale:

$$\forall \vec{k} : \left(\frac{\hbar^2}{2m} k^2 - \epsilon \right) C_{\vec{k}} + \sum_{\vec{G}} U_{\vec{G}} C_{\vec{k} - \vec{G}} = 0 \quad (591)$$

Questa e' l'equazione centrale.

214. Come potete vedere, l'equazione centrale non e' un'equazione singola, ma un sistema lineare omogeneo costituito da un numero infinito di equazioni: esso contiene tante equazioni quanti sono i vettori \vec{k} (notate il $\forall \vec{k}$ nella 591).

Le incognite da trovare nell'equazione centrale sono i coefficienti $C_{\vec{k}}$ e gli autovalori ϵ .

Da un punto di vista di principio, l'approccio alla soluzione e' il seguente.

Il sistema omogeneo ammette soluzioni non banali (cioe' soluzioni in cui i $C_{\vec{k}}$ non siano tutti nulli) se e solo se il determinante costruito con i coefficienti che moltiplicano i $C_{\vec{k}}$ e' nullo. Imponendo questa condizione si ottiene un'equazione (detta *equazione secolare o caratteristica*) nell'incognita ϵ : la soluzione dell'equazione caratteristica fornisce tutti gli autovalori del problema. Successivamente, per ciascun autovalore, la soluzione del sistema omogeneo fornisce il corrispondente set di coefficienti $C_{\vec{k}}$, ovvero l'autofunzione corrispondente a quel particolare autovalore tramite l'eq. 566.

215. Osservate subito come l'equazione centrale fornisca correttamente l'energia degli stati che avevamo gia' trovato per gli elettroni liberi nel caso in cui il potenziale periodico sia identicamente nullo. In questo caso, infatti, tutti i coefficienti $U_{\vec{G}}$ sono nulli e l'equazione centrale si riduce a:

$$\forall \vec{k} : \left(\frac{\hbar^2}{2m} k^2 - \epsilon \right) C_{\vec{k}} = 0 \quad (592)$$

che e' un sistema di infinite equazioni *disaccoppiate* in una sola incognita ($C_{\vec{k}}$). Ciascuna di queste equazioni ammette una soluzione non banale (cioe' diversa da $C_{\vec{k}} = 0$) se e solo se:

$$\frac{\hbar^2}{2m} k^2 - \epsilon = 0 \quad (593)$$

ovvero:

$$\epsilon = \frac{\hbar^2}{2m} k^2 \quad (594)$$

che e' esattamente lo stesso risultato che avevamo ottenuto per gli elettroni liberi del gas di Fermi (eq. 344).

216. E' importante comprendere bene la struttura dell'equazione centrale. Ogni equazione del sistema e' indicizzata dal vettore \vec{k} : ci sono tante equazioni quanti sono i vettori \vec{k} cioe' quanti sono i punti del reticolo reciproco macroscopico (infiniti).

Fissato un certo \vec{k} , la corrispondente equazione contiene una somma (infinita) di termini $U_{\vec{G}} C_{\vec{k}-\vec{G}}$, indicizzati da \vec{G} ; la somma contiene tanti addendi quanti sono i punti \vec{G} del reticolo reciproco microscopico (anche questi sono in numero infinito).

Possiamo ordinare le equazioni del sistema 591 tenendo presente quanto abbiamo detto al punto 206. I punti \vec{k} possono essere ripartiti in N sottoinsiemi ($N = N_1 N_2 N_3$ e' il numero totale di celle primitive contenute nel campione macroscopico).

Prendiamo una cella primitiva di riferimento del reticolo reciproco microscopico: quasi sempre, anche se cio' non e' assolutamente necessario, si considera come cella primitiva la prima zona di Brillouin (punto 4 a p. 68). Tale cella contiene N punti \vec{k} ; ciascuno di tali punti individua un sottoinsieme (infinito) di punti \vec{k} che sono le sue immagini generate per traslazione tramite i vettori \vec{G} del reticolo reciproco microscopico.

Ciascuno degli N sottoinsiemi cosi' generati non ha punti in comune con alcuno degli altri; in altre parole, gli N sottoinsiemi sono *disgiunti*.

Torniamo al sistema 591: siccome le sue infinite equazioni sono mappate sui punti \vec{k} , e siccome ogni equazione contiene solamente termini corrispondenti a punti \vec{k} che sono connessi da un vettore \vec{G} , *il sistema puo' essere ripartito in N sottosistemi indipendenti, ciascuno contenente solo i punti \vec{k} connessi da traslazioni \vec{G} .*

217. Cerchiamo di "vedere" questa ripartizione del sistema 591. Prendiamo la prima zona di Brillouin (da qui in poi quando parleremo della prima zona di Brillouin sara' sempre sottinteso che ci riferiamo al reticolo reciproco microscopico) e numeriamo, da 1 a N , i punti \vec{k} in essa contenuti:³⁶ $\vec{k}_1, \vec{k}_2, \dots, \vec{k}_N$. Supponiamo inoltre di numerare anche tutti i possibili vettori di traslazione del reticolo reciproco microscopico: $\vec{G}_1, \vec{G}_2, \dots$ ³⁷.

A questo punto possiamo ripartire le equazioni del sistema 591 in N gruppi ciascuno dei quali contiene un punto \vec{k}_i interno alla prima zona di Brillouin e tutti i punti ad esso collegati da traslazioni \vec{G} .

Per uniformita' di notazione e maggiore semplicita', conviene scrivere il punto \vec{k} collegato dalla traslazione \vec{G}_j al punto \vec{k}_i interno alla prima zona di Brillouin come $\vec{k}_i - \vec{G}_j$ ³⁸.

³⁶Notate che, mentre nel caso monodimensionale la numerazione dei punti si puo' immaginare molto facilmente (ad esempio li possiamo numerare in ordine crescente), nel caso 3D un metodo di enumerazione non e' immediatamente evidente; tuttavia possiamo supporre di numerare i punti \vec{k} *in qualche modo*, i cui dettagli non sono importanti in questo contesto fin tanto che rimane stabilita una corrispondenza biunivoca fra i punti \vec{k} e l'insieme dei numeri naturali

³⁷Mentre i vettori \vec{k} contenuti nella prima zona di Brillouin sono in numero finito, N , i vettori di traslazione \vec{G} sono in numero infinito

³⁸Realizzate che, siccome le immagini del punto \vec{k}_i sono generate da *tutti* i possibili vettori \vec{G}_j , non fa alcuna differenza scrivere la generica immagine come $\vec{k}_i - \vec{G}_j$ oppure $\vec{k}_i + \vec{G}_j$

equazioni dell'insieme $\vec{k}_1 - \vec{G}_i$:

$$\begin{aligned} & \left(\frac{\hbar^2 k_1^2}{2m} - \epsilon \right) C_{\vec{k}_1} + U_{\vec{G}_1} C_{\vec{k}_1 - \vec{G}_1} + U_{\vec{G}_2} C_{\vec{k}_1 - \vec{G}_2} + \dots = 0 \\ & \left(\frac{\hbar^2 |\vec{k}_1 - \vec{G}_1|^2}{2m} - \epsilon \right) C_{\vec{k}_1 - \vec{G}_1} + U_{\vec{G}_1} C_{\vec{k}_1 - \vec{G}_1 - \vec{G}_1} + U_{\vec{G}_2} C_{\vec{k}_1 - \vec{G}_1 - \vec{G}_2} + \dots = 0 \\ & \left(\frac{\hbar^2 |\vec{k}_1 - \vec{G}_2|^2}{2m} - \epsilon \right) C_{\vec{k}_1 - \vec{G}_2} + U_{\vec{G}_1} C_{\vec{k}_1 - \vec{G}_2 - \vec{G}_1} + U_{\vec{G}_2} C_{\vec{k}_1 - \vec{G}_2 - \vec{G}_2} + \dots = 0 \\ & \dots \end{aligned}$$

equazioni dell'insieme $\vec{k}_2 - \vec{G}_i$:

$$\begin{aligned} & \left(\frac{\hbar^2 k_2^2}{2m} - \epsilon \right) C_{\vec{k}_2} + U_{\vec{G}_1} C_{\vec{k}_2 - \vec{G}_1} + U_{\vec{G}_2} C_{\vec{k}_2 - \vec{G}_2} + \dots = 0 \\ & \left(\frac{\hbar^2 |\vec{k}_2 - \vec{G}_1|^2}{2m} - \epsilon \right) C_{\vec{k}_2 - \vec{G}_1} + U_{\vec{G}_1} C_{\vec{k}_2 - \vec{G}_1 - \vec{G}_1} + U_{\vec{G}_2} C_{\vec{k}_2 - \vec{G}_1 - \vec{G}_2} + \dots = 0 \\ & \left(\frac{\hbar^2 |\vec{k}_2 - \vec{G}_2|^2}{2m} - \epsilon \right) C_{\vec{k}_2 - \vec{G}_2} + U_{\vec{G}_1} C_{\vec{k}_2 - \vec{G}_2 - \vec{G}_1} + U_{\vec{G}_2} C_{\vec{k}_2 - \vec{G}_2 - \vec{G}_2} + \dots = 0 \\ & \dots \end{aligned}$$

equazioni dell'insieme $\vec{k}_N - \vec{G}_i$:

$$\begin{aligned} & \left(\frac{\hbar^2 k_N^2}{2m} - \epsilon \right) C_{\vec{k}_N} + U_{\vec{G}_1} C_{\vec{k}_N - \vec{G}_1} + U_{\vec{G}_2} C_{\vec{k}_N - \vec{G}_2} + \dots = 0 \\ & \left(\frac{\hbar^2 |\vec{k}_N - \vec{G}_1|^2}{2m} - \epsilon \right) C_{\vec{k}_N - \vec{G}_1} + U_{\vec{G}_1} C_{\vec{k}_N - \vec{G}_1 - \vec{G}_1} + U_{\vec{G}_2} C_{\vec{k}_N - \vec{G}_1 - \vec{G}_2} + \dots = 0 \\ & \left(\frac{\hbar^2 |\vec{k}_N - \vec{G}_2|^2}{2m} - \epsilon \right) C_{\vec{k}_N - \vec{G}_2} + U_{\vec{G}_1} C_{\vec{k}_N - \vec{G}_2 - \vec{G}_1} + U_{\vec{G}_2} C_{\vec{k}_N - \vec{G}_2 - \vec{G}_2} + \dots = 0 \end{aligned}$$

218. Il sottosistema i -esimo e' indicizzato dal punto \vec{k}_i della prima zona di Brillouin e accoppia fra loro solo coefficienti del tipo $C_{\vec{k}_i - \vec{G}_j}$ ³⁹.

I coefficienti che compaiono in un dato sottosistema, non compaiono in nessuno degli altri sottosistemi. Quindi, ciascun sottosistema puo' essere risolto indipendentemente dagli altri.

219. Abbiamo quindi suddiviso il problema di partenza (il sistema 591) in N sistemi lineari omogenei indipendenti. Ciascuno di essi ha una dimensione infinita e quindi ha infinite soluzioni (ricordate che ϵ e' un'incognita da trovare e

³⁹Notate che qualsiasi combinazione lineare dei vettori \vec{G} con coefficienti interi continua ad essere un vettore di traslazione del reticolo reciproco microscopico, per cui, ad esempio, il coefficiente $C_{\vec{k}_2 - \vec{G}_2 - \vec{G}_1}$ e' sempre del tipo $C_{\vec{k}_i - \vec{G}_j}$

quindi ci sarà una soluzione per ciascuno degli infiniti valori di ϵ che soddisfano l'equazione caratteristica). Notate che una singola soluzione significa un set di coefficienti $C_{\vec{k}_i - \vec{G}_j}$ e il corrispondente autovalore ϵ .

Siccome un dato sottosistema, corrispondente a un certo punto \vec{k} nella prima zona di Brillouin, ha infinite soluzioni, una singola soluzione e il suo corrispondente autovalore non possono essere indicizzati dal solo punto \vec{k} : fissato \vec{k} servirà un secondo indice n che enumera tutte le soluzioni (e gli autovalori) di un dato sottosistema. Tale indice è lo stesso che indicizza i vettori \vec{G} : ad ogni vettore del reticolo reciproco microscopico \vec{G} corrisponde una particolare soluzione del sottosistema individuato da un certo punto \vec{k} .

In altre parole, individuiamo una particolare soluzione dell'equazione di Schrödinger e il suo autovalore con:

$$\Psi_{\vec{k}n}(\vec{r}) \text{ e } \epsilon_{\vec{k}n}$$

dove \vec{k} è un vettore d'onda permesso interno alla prima zona di Brillouin ed $n = 1, 2, 3, \dots$

Normalmente, gli autovalori $\epsilon_{\vec{k}n}$ sono indicizzati da n in ordine di energia crescente: $\epsilon_{\vec{k}1} \leq \epsilon_{\vec{k}2} \leq \epsilon_{\vec{k}3} \leq \dots$

Come vedete, si è finalmente chiarita anche la natura del secondo indice (n) che avevamo anticipato nell'enunciato del teorema di Bloch a p. 135.

220. Siccome una soluzione dell'equazione centrale 591 corrispondente al vettore d'onda permesso \vec{k}_i contiene solo coefficienti del tipo $C_{\vec{k}_i - \vec{G}_j}$, lo sviluppo in serie di Fourier della $\Psi(\vec{r})$ può essere convenientemente scritto come:

$$\Psi_{\vec{k}n}(\vec{r}) = \sum_{\vec{G}} C_{(\vec{k} - \vec{G})n} \exp\left(i(\vec{k} - \vec{G}) \cdot \vec{r}\right) \quad (595)$$

(di nuovo, abbiamo ommesso gli indici per \vec{k} e \vec{G} per alleggerire e abbiamo aggiunto l'indice n al set di coefficienti che definiscono $\Psi_{\vec{k}n}(\vec{r})$)

Notate che $\vec{k} - \vec{G}$ è comunque un punto \vec{k} .

La 595 mette in risalto il fatto che, nello sviluppo in serie di Fourier della particolare soluzione indicizzata dal punto \vec{k} nella prima zona di Brillouin (e dall'indice n), non compaiono *tutti* i punti \vec{k} , ma solo le (infinite) immagini del particolare punto \vec{k} generate dalle traslazioni \vec{G} del reticolo reciproco microscopico⁴⁰.

221. A questo punto è opportuna una precisazione.

Come vi sarete resi conto, l'equazione centrale rimane un oggetto intrattabile: è vero che l'abbiamo ripartita in un numero finito (anche se enorme, dell'ordine di grandezza del numero di Avogadro) di sottosistemi lineari omogenei indipendenti, ma ciascun sottosistema è pur sempre costituito da un numero infinito di equazioni (il numero di equazioni di un sottosistema è uguale al numero dei vettori \vec{G}).

D'altro canto l'equazione centrale ha un enorme valore teorico: grazie ad essa siamo in grado di indicizzare le soluzioni dell'equazione di Schrödinger e

⁴⁰Attenti a non fare confusione con l'eq. 572. Quella è solo un "cambiamento di origine" dello sviluppo in serie di Fourier della $\Psi(\vec{r})$: l'indice della sommatoria nella 572 corre su tutti i vettori \vec{k} e il vettore \vec{G} è unico; viceversa, nella 595, la sommatoria corre su tutti i vettori \vec{G} mentre \vec{k} è un unico punto nella prima zona di Brillouin.

discuterne alcune proprietà'. Inoltre, come vedremo fra poco (punto 222) l'equazione centrale e la 595 da essa ricavata costituiscono un'altra dimostrazione del teorema di Bloch.

Aggiungo a questo punto che l'equazione centrale diventa anche uno strumento di utilità pratica nel momento in cui sia possibile (e opportunamente giustificabile):

- ⇒ troncare gli sviluppi in serie del potenziale e della funzione $\Psi(\vec{r})$ a un numero *finito* (e gestibile) di termini
- ⇒ limitare il numero di punti \vec{k} da considerare nella prima zona di Brillouin

Sotto queste ipotesi, sia il numero di sottosistemi da risolvere che la dimensione di ciascuno di essi possono diventare effettivamente trattabili da un punto di vista pratico.

222. Un'altra dimostrazione del teorema di Bloch.

L'equazione centrale 591 e l'espressione 595 per $\Psi_{\vec{k}n}(\vec{r})$ sono state ottenute *senza utilizzare il teorema di Bloch*; in effetti la 595 costituisce una dimostrazione alternativa di questo teorema.

Infatti essa può essere riscritta come:

$$\Psi_{\vec{k}n}(\vec{r}) = \exp(i\vec{k} \cdot \vec{r}) \sum_{\vec{G}} C_{(\vec{k}-\vec{G})n} \exp(-i\vec{G} \cdot \vec{r}) = \exp(i\vec{k} \cdot \vec{r}) u_{\vec{k}n}(\vec{r}) \quad (596)$$

con:

$$u_{\vec{k}n}(\vec{r}) = \sum_{\vec{G}} C_{(\vec{k}-\vec{G})n} \exp(-i\vec{G} \cdot \vec{r}) \quad (597)$$

E' facile verificare che $u_{\vec{k}n}(\vec{r})$ ha la periodicità del reticolo diretto microscopico; per ogni vettore di traslazione \vec{R} del reticolo diretto microscopico si ha:

$$\begin{aligned} u_{\vec{k}n}(\vec{r} + \vec{R}) &= \sum_{\vec{G}} C_{(\vec{k}-\vec{G})n} \exp(-i\vec{G} \cdot (\vec{r} + \vec{R})) \\ &= \sum_{\vec{G}} C_{(\vec{k}-\vec{G})n} \exp(-i\vec{G} \cdot \vec{R}) \exp(-i\vec{G} \cdot \vec{r}) \\ &= \sum_{\vec{G}} C_{(\vec{k}-\vec{G})n} \exp(-i\vec{G} \cdot \vec{r}) = u_{\vec{k}n}(\vec{r}) \end{aligned} \quad (598)$$

perché $\exp(-i\vec{G} \cdot \vec{R}) = 1$ (p. 32).

Ma allora la 596 afferma proprio la tesi del teorema di Bloch, nella forma 491. Osservate che il vettore \vec{k} che indicizza la $\Psi(\vec{r})$ nella 491 e' lo stesso vettore \vec{k} che individua il sottosistema \vec{k} -esimo nell'equazione centrale (punto 217).

223. Periodicità di $\Psi_{\vec{k}n}(\vec{r})$ ed $\epsilon_{\vec{k}n}$ nello spazio \vec{k} .

Da quanto abbiamo detto sull'equazione centrale emerge che tutte le soluzioni dell'equazione di Schrödinger 490 con le condizioni al contorno 540 sono indicizzate dai vettori \vec{k} contenuti nella prima zona di Brillouin. Implicitamente, questo suggerisce che tutti gli altri punti dello spazio \vec{k} (cioè i punti del reticolo

reciproco macroscopico) siano in qualche modo “ridondanti”, cioè non aggiungano nulla a quanto è già stato possibile ricavare restringendosi ai soli punti \vec{k} della prima zona di Brillouin.

Che questo sia effettivamente il caso possiamo confermarlo nel modo seguente. Riprendiamo l'espressione 595 della soluzione dell'equazione di Schrödinger corrispondente a un punto \vec{k} della prima zona di Brillouin per un n fissato:

$$\Psi_{\vec{k}n}(\vec{r}) = \sum_{\vec{G}} C_{(\vec{k}-\vec{G})n} \exp\left(i(\vec{k}-\vec{G}) \cdot \vec{r}\right) \quad (599)$$

e chiediamoci quale sarebbe l'autofunzione corrispondente al punto $\vec{k} + \vec{G}'$ (se \vec{k} è interno alla prima zona di Brillouin, $\vec{k} + \vec{G}'$ è senz'altro esterno ad essa⁴¹). Sostituendo sopra $\vec{k} + \vec{G}'$ al posto di \vec{k} si ha:

$$\Psi_{(\vec{k}+\vec{G}')n}(\vec{r}) = \sum_{\vec{G}} C_{(\vec{k}+\vec{G}'-\vec{G})n} \exp\left(i(\vec{k} + \vec{G}' - \vec{G}) \cdot \vec{r}\right) \quad (600)$$

Ma $\vec{G}' - \vec{G}$ è sempre un vettore di traslazione del reticolo reciproco microscopico; allora, ponendo:

$$\vec{G}'' = \vec{G} - \vec{G}' \quad (601)$$

possiamo riscrivere la sommatoria come:

$$\Psi_{(\vec{k}+\vec{G}')n}(\vec{r}) = \sum_{\vec{G}''} C_{(\vec{k}-\vec{G}'')n} \exp\left(i(\vec{k} - \vec{G}'') \cdot \vec{r}\right) = \Psi_{\vec{k}n}(\vec{r}) \quad (602)$$

Come avevamo intuito dalla derivazione dell'equazione centrale, tutti i punti \vec{k} esterni alla prima zona di Brillouin (o qualsiasi sia la cella primitiva presa come riferimento) sono ridondanti nel senso che la funzione d'onda ad essi corrispondente è esattamente la stessa indicizzata da un corrispondente punto \vec{k} all'interno della prima zona di Brillouin.

In altre parole, due autofunzioni i cui vettori d'onda \vec{k} differiscono per un vettore traslazione \vec{G} sono identiche.

Quanto appena trovato ha una corrispondente implicazione sulla dipendenza degli autovalori da \vec{k} .

Siccome $\Psi_{\vec{k}n}(\vec{r})$ e $\Psi_{(\vec{k}+\vec{G})n}(\vec{r})$ sono identiche, i due problemi di Schrödinger:

$$H\Psi_{\vec{k}n}(\vec{r}) = \epsilon_{\vec{k}n} \Psi_{\vec{k}n}(\vec{r}) \quad (603)$$

e

$$H\Psi_{(\vec{k}+\vec{G})n}(\vec{r}) = \epsilon_{(\vec{k}+\vec{G})n} \Psi_{(\vec{k}+\vec{G})n}(\vec{r}) \quad (604)$$

devono necessariamente avere lo stesso autovalore. Ma questo significa che:

$$\epsilon_{(\vec{k}+\vec{G})n} = \epsilon_{\vec{k}n} \quad (605)$$

cioè: per un valore di n fissato, ϵ è una funzione periodica di \vec{k} con la periodicità del reticolo reciproco microscopico.

⁴¹Tranne che nel caso banale, che escludiamo, $\vec{G} = 0$

224. La struttura a bande.

Abbiamo visto che, per ognuno degli N punti \vec{k} interni alla prima zona di Brillouin, l'equazione di Schrödinger 490 ammette un insieme (teoricamente infinito) di soluzioni e corrispondenti autovalori indicizzati da n : $\Psi_{\vec{k}n}(\vec{r})$ e $\epsilon_{\vec{k}n}$. Fissato un certo valore di n , l'energia $\epsilon_{\vec{k}n}$ degli stati monoelettronici e' una funzione del vettore d'onda \vec{k} ; per mettere meglio in evidenza questo fatto, possiamo indicarla con $\epsilon_n(\vec{k})$.

Detti ϵ_n^{min} ed ϵ_n^{max} i valori minimo e massimo, rispettivamente, che la funzione $\epsilon_n(\vec{k})$ assume sui punti \vec{k} della prima zona di Brillouin (per un certo n fissato), si definisce *banda n-esima* l'intervallo di energia $\epsilon_n^{max} - \epsilon_n^{min}$.

L'insieme delle bande che si ottengono per i valori successivi di $n = 1, 2, 3 \dots$ costituiscono quella che viene chiamata *la struttura a bande di un solido*.

Approfondiremo questo argomento a piu' riprese. Per il momento e' importante che fissiate il concetto di banda come un intervallo di energia (una banda, appunto) che delimita superiormente e inferiormente l'andamento della funzione $\epsilon_n(\vec{k})$ per un particolare valore fissato di n .

Siccome abbiamo visto al punto 223 che $\epsilon_n(\vec{k})$ e' periodica nello spazio \vec{k} (con la periodicitá' del reticolo reciproco microscopico), il suo andamento su tutto lo spazio \vec{k} si ottiene semplicemente come replica periodica di quello all'interno della prima zona di Brillouin: quindi il concetto di banda rimane esattamente lo stesso anche quando si considera l'intero spazio \vec{k} : infatti, se la funzione $\epsilon_n(\vec{k})$ assume valori interni all'intervallo $[\epsilon_n^{min}, \epsilon_n^{max}]$ nella prima zona di Brillouin, lo stesso deve valere, a causa della periodicitá', non solo nella prima zona di Brillouin, ma in tutto lo spazio \vec{k} .

Siccome le bande di un solido sono indicizzate da n , esso viene anche detto *indice di banda*.

225. La rappresentazione della struttura a bande (nel modello "empty lattice").

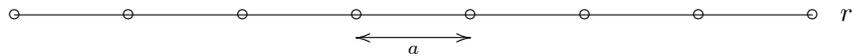
La rappresentazione grafica della funzione $\epsilon_n(\vec{k})$ nel caso di un solido cristallino "reale" (intendendo con cio' "tridimensionale") non e' affatto banale: si tratta pur sempre di una funzione di 3 variabili (k_x, k_y, k_z), il cui "grafico" richiederebbe uno spazio a 4 dimensioni.

Per questo motivo, e' molto utile ed istruttivo iniziare con l'analisi di un sistema monodimensionale (mettendo da parte, quindi, ogni pretesa di realta' fisica).

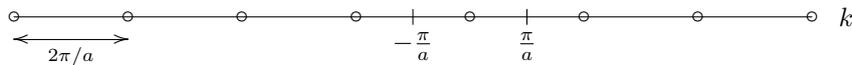
Per semplificare ulteriormente le cose, supporremo che il potenziale periodico (monodimensionale) sia *sufficientemente piccolo* da poter essere completamente trascurato nella risoluzione dell'equazione di Schrödinger, ma *sufficientemente grande* da indurre la periodicitá' della funzione $\epsilon_n(\vec{k})$ nello spazio \vec{k} che abbiamo messo in evidenza al punto 223. Questo modello semplificato viene chiamato "empty lattice approximation" (ad esempio nel libro di Kittel¹⁸).

In altre parole, questo modello consente di assumere che le autofunzioni e gli autovalori dell'hamiltoniano siano quelli degli elettroni liberi del gas di Fermi (eq. 335 e 344), ma introduce comunque la periodicitá' nello spazio \vec{k} della relazione di dispersione $\epsilon_n(\vec{k})$.

Consideriamo dunque un "cristallo" monodimensionale avente una "cella" primitiva di lunghezza a .



Il reticolo reciproco di questa struttura e' una griglia monodimensionale con spaziatura pari a $2\pi/a$. La prima zona di Brillouin e' l'intervallo $[-\pi/a, \pi/a]$. Questo reticolo reciproco corrisponde al reticolo reciproco microscopico che avevamo introdotto al punto 198 e seguenti.



Siccome assumiamo che il potenziale sia talmente piccolo da non influenzare le soluzioni dell'equazione di Schrödinger, gli autovalori saranno quelli che avevamo trovato a suo tempo per il gas di Fermi (in cui il potenziale e' rigorosamente nullo), cioe':

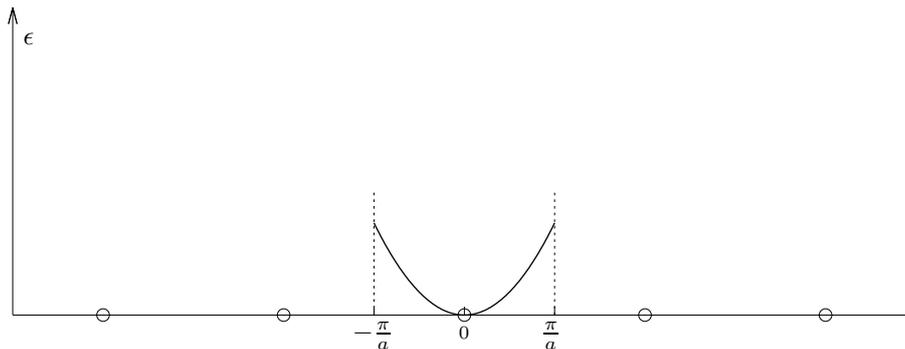
$$\epsilon(k) = \frac{\hbar^2 k^2}{2m} \quad (606)$$

Notate che nel modello monodimensionale che stiamo considerando il vettore " \vec{k} " e' in realta' uno scalare. In una dimensione, questa e' l'equazione di una parabola con vertice nell'origine.

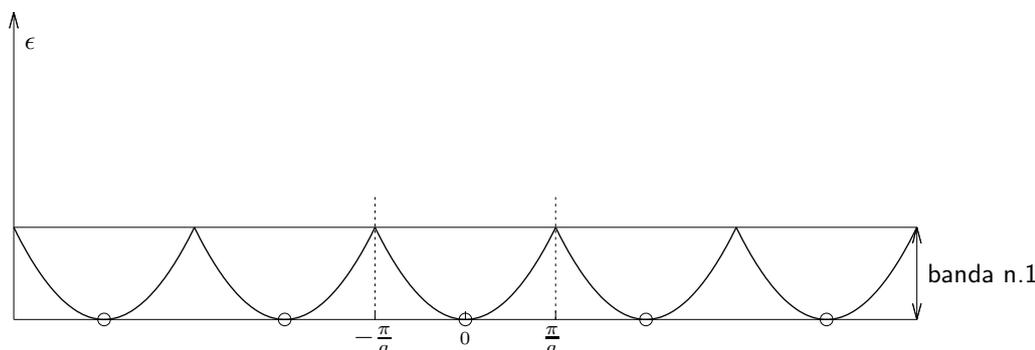
Ora: tenendo presente la relazione di dispersione scritta sopra, per ogni punto k interno alla prima zona di Brillouin, l'autovalore piu' basso, cioe' quello con $n = 1$, e' dato da:

$$\epsilon_1(k) = \frac{\hbar^2 k^2}{2m} \quad k \in \left[-\frac{\pi}{a}, \frac{\pi}{a}\right] \quad (607)$$

cioe':



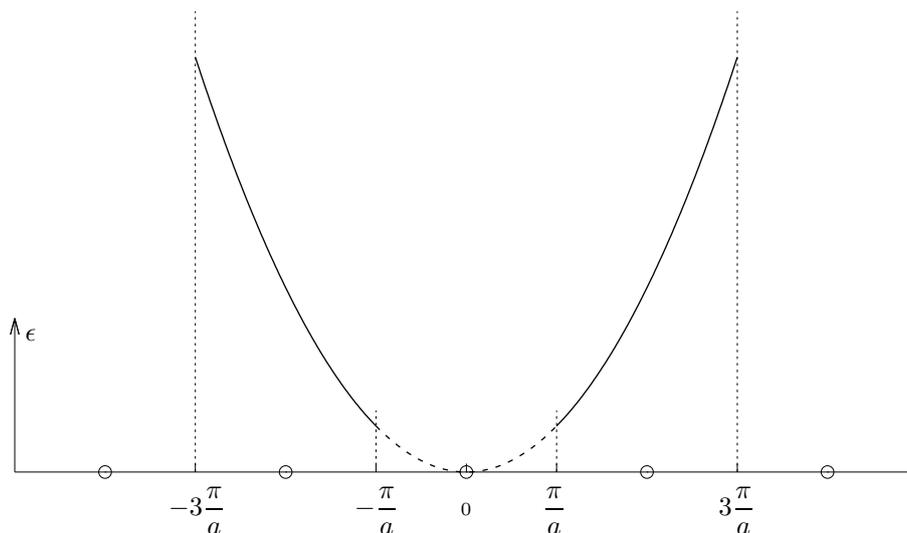
Ora sfruttiamo l'ipotesi che il potenziale sia sufficientemente grande da indurre la periodicita' della relazione di dispersione (punto 223): l'andamento di $\epsilon_1(k)$ nella prima zona di Brillouin si ripete in tutte le celle primitive del reticolo reciproco; l'intervallo di energia entro il quale e' contenuto il grafico di $\epsilon_1(k)$ e' la banda n.1:



Ora vogliamo trovare $\epsilon_2(k)$, cioè l'autovalore immediatamente superiore al più basso, per tutti i punti k nella prima zona di Brillouin. L'intervallo di energia entro cui il grafico di $\epsilon_2(k)$ è limitato costituirà la banda n.2.

Qua apparentemente abbiamo un problema: il grafico della parabola che descrive la relazione di dispersione è *esterno* alla prima zona di Brillouin per valori di energia maggiori dell'estremo superiore della prima banda (guardate la prima figura): come facciamo a calcolare $\epsilon_2(k)$ *all'interno* della prima zona di Brillouin? Ci viene in aiuto la periodicità della funzione $\epsilon_n(k)$ dovuta all'ipotesi che il potenziale periodico non sia (esattamente) nullo (punto 223); allora, possiamo valutare i valori della relazione di dispersione $\epsilon(k)$ in una cella primitiva *qualsiasi* per poi *traslarli* nella prima zona di Brillouin.

Sorge però un secondo problema. I valori di $\epsilon(k)$ immediatamente superiori a quelli che abbiamo calcolato per la banda n.1 si ottengono calcolando la parabola $\hbar^2 k^2 / 2m$ nelle due celle adiacenti alla prima zona di Brillouin; ma il grafico di $\epsilon(k)$ nelle due celle è diverso; si tratta sempre di due rami della parabola, ma essi sono immagini speculari l'uno dell'altro (per maggior chiarezza, nel grafico seguente, ho tracciato anche il tratto di parabola che abbiamo già considerato per la prima banda):



In realtà questo non è un problema.

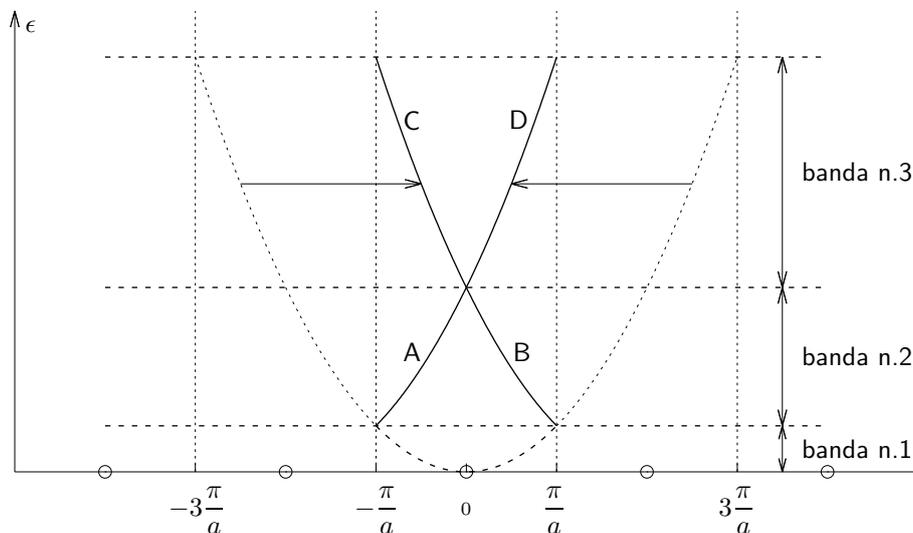


Figura 52:

Dovete tenere presente che stiamo cercando, per ogni punto k nella prima zona di Brillouin, gli autovalori (*qualsiasi essi siano*) immediatamente maggiori di quelli che definiscono la prima banda.

Se trasliamo entrambi i rami della parabola nella prima zona di Brillouin, otteniamo ciò che è mostrato nella figura 52.

Dalla figura vediamo che gli autovalori immediatamente maggiori di quelli che definiscono la prima banda sono quelli del tratto A nell'intervallo $[-\pi/a, 0]$, e quelli del tratto B nell'intervallo $[0, \pi/a]$. Quindi la seconda banda di energia è definita dall'unione dei tratti A e B .

Inoltre, sempre dalla figura, si vede che i tratti C e D assieme rappresentano gli autovalori immediatamente successivi a quelli della seconda banda e quindi definiscono la terza banda: abbiamo trovato simultaneamente la seconda e la terza banda di energia per questo sistema.

226. Questo procedimento si può generalizzare.

Riguardate la figura 52: dal punto di vista matematico, la traslazione del ramo di parabola $B C$ dalla cella $[-3\pi/a, -\pi/a]$ all'interno della prima zona di Brillouin significa calcolare la relazione di dispersione per gli elettroni liberi in $k - 2\pi/a$ per $k \in [-\pi/a, \pi/a]$, cioè:

$$\frac{\hbar^2}{2m} \left(k - \frac{2\pi}{a} \right)^2 \quad k \in [-\pi/a, \pi/a] \quad (608)$$

In modo analogo, la traslazione del ramo di parabola $A D$ dalla cella $[\pi/a, 3\pi/a]$ all'interno della prima zona di Brillouin significa calcolare la relazione di dispersione per gli elettroni liberi in $k + 2\pi/a$ per $k \in [-\pi/a, \pi/a]$:

$$\frac{\hbar^2}{2m} \left(k + \frac{2\pi}{a} \right)^2 \quad k \in [-\pi/a, \pi/a] \quad (609)$$

Le bande successive possono essere generate traslando rami di parabola da celle via via piu' lontane dalla prima zona di Brillouin. In generale, il ramo di parabola della cella j viene traslato all'interno della prima zona di Brillouin (considerata cella 0) calcolando:

$$\frac{\hbar^2}{2m} \left(k - j \frac{2\pi}{a} \right)^2 \quad k \in [-\pi/a, \pi/a] \quad (610)$$

Per $j = 0$ otteniamo la banda n.1, per $j = \pm 1$ otteniamo le bande n.2 e 3, per $j = \pm 2, \pm 3, \dots$ otterremo le bande successive.

Osservando che $j \frac{2\pi}{a}$ non e' altro che il generico vettore G del reticolo reciproco, l'espressione vista sopra puo' essere riscritta come:

$$\frac{\hbar^2}{2m} (k - G)^2 \quad k \in [-\pi/a, \pi/a] \quad \forall G \quad (611)$$

dove $G = 0$ corrisponde alla prima zona di Brillouin, $G = \pm 2\pi/a$ genera le bande n.2 e 3 e cosi' via.

Questo procedimento e' completamente generale nel senso che rimane valido anche per un sistema 3D e per qualsiasi espressione della relazione di dispersione $\epsilon(\vec{k})$. Quindi la struttura a bande di un sistema qualsiasi avente una relazione di dispersione qualsiasi puo' essere generata calcolando:

$$\epsilon(\vec{k} - \vec{G}) \quad (612)$$

per $\vec{k} \in$ alla prima zona di Brillouin e $\forall \vec{G}$.

227. Nella figura 53 sono mostrate le prime tre bande e la ripetizione periodica della prima zona di Brillouin.

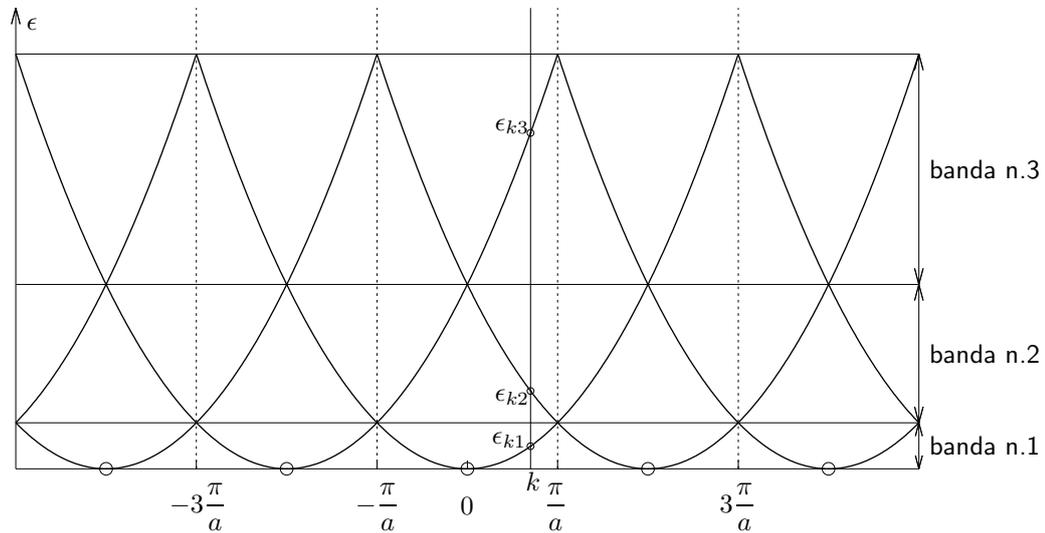


Figura 53: Rappresentazione "repeated zone" delle prime 3 bande del modello "empty lattice" monodimensionale.

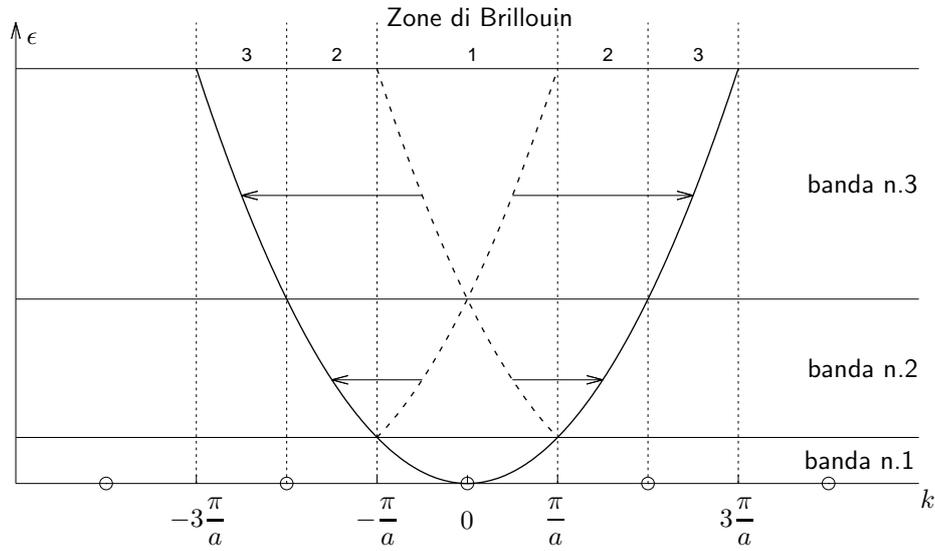


Figura 55: Rappresentazione “extended zone” delle prime 3 bande per un cristallo monodimensionale contenente elettroni liberi.

e quindi:

$$\epsilon(-k) = \frac{\hbar^2}{2m} \left(|k| - \frac{2\pi}{a} \right)^2 = \frac{\hbar^2}{2m} \left(-|k| + \frac{2\pi}{a} \right)^2 = \epsilon(k) \quad (614)$$

Analogamente:

$$\text{quando: } 0 \leq k \leq \frac{\pi}{a}, \text{ allora: } -k = -|k| \in \left[-\frac{\pi}{a}, 0 \right]$$

e quindi:

$$\epsilon(-k) = \frac{\hbar^2}{2m} \left(-|k| + \frac{2\pi}{a} \right)^2 = \frac{\hbar^2}{2m} \left(|k| - \frac{2\pi}{a} \right)^2 = \epsilon(k) \quad (615)$$

Un terzo modo di rappresentare la struttura a bande e' detto “extended zone representation” e consiste nel riportare i rami di ciascuna banda in una differente zona di Brillouin. Come potete immaginare, nel caso di un cristallo monodimensionale contenente elettroni liberi, questo procedimento equivale al grafico della parabola $\epsilon = \hbar^2 k^2 / (2m)$ ed e' mostrato nella figura 55.

228. Agli stessi risultati si poteva giungere per una via equivalente. Sempre utilizzando il fatto che la presenza del potenziale periodico induce una periodicitá della relazione di dispersione nello spazio \vec{k} , la relazione parabolica:

$$\epsilon(k) = \frac{\hbar^2 k^2}{2m}$$

viene replicata in tutte le celle del reticolo reciproco. In tal modo si ottiene direttamente la figura 53, dove le bande sono individuate dalle intersezioni fra le varie immagini della parabola.

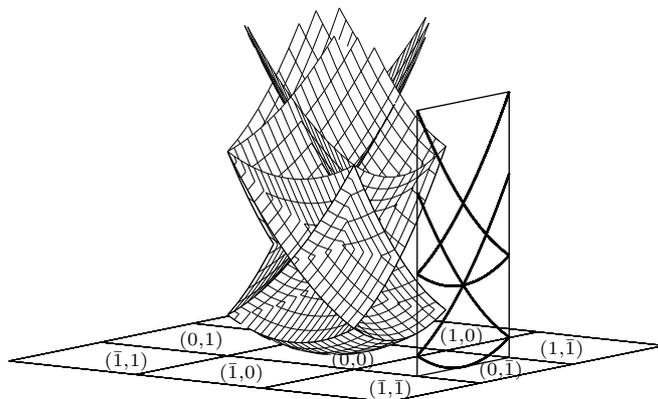


Figura 56: Struttura a bande nella rappresentazione “reduced zone” per un cristallo bidimensionale con reticolo di Bravais cubico semplice contenente elettroni liberi. La spiegazione della figura e’ dettagliata nel testo.

229. La rappresentazione grafica della struttura a bande diventa piu’ complicata quando si passa da un sistema monodimensionale a uno bi- o tridimensionale. A titolo di esempio vediamo le prime bande in un cristallo bidimensionale basato su un reticolo di Bravais cubico semplice (con costante di cella a) contenente elettroni liberi (cioe’ nel modello “empty lattice”). Applichiamo il metodo generale visto al punto 226. Le infinite bande del sistema si ottengono calcolando:

$$\epsilon(\vec{k} - \vec{G})$$

per tutti i punti \vec{k} interni alla prima zona di Brillouin e per tutti i possibili vettori \vec{G} .

Il reticolo reciproco di un reticolo cubico semplice e’ anch’esso un reticolo cubico semplice, con costante di cella $2\pi/a$; la prima zona di Brillouin e’ un quadrato di lato $2\pi/a$ centrato nell’origine. Il generico vettore del reticolo reciproco e’:

$$\vec{G} = \frac{2\pi}{a} (n_x, n_y) \quad n_x, n_y \in \mathbb{Z} \quad (616)$$

Siccome ci limitiamo alle sole prime bande, cioe’ quelle a piu’ bassa energia, consideriamo solo un piccolo numero di vettori \vec{G} .

Fate riferimento alla figura 56, dove ho rappresentato la prima zona di Brillouin e le 8 celle primitive che la “circondano”. Ci sono in totale 9 celle primitive, ciascuna individuata da un vettore \vec{G} distinto: ho usato gli interi n_x ed n_y definiti sopra per etichettare le celle.

La relazione di dispersione e’ quella per gli elettroni liberi, che, per un sistema bidimensionale, e’ la parabola:

$$\begin{aligned}
\epsilon(\vec{k}) &= \frac{\hbar^2}{2m} (\vec{k} - \vec{G})^2 = \frac{\hbar^2}{2m} |\vec{k} - \vec{G}|^2 = \frac{\hbar^2}{2m} \left| (k_x, k_y) - \frac{2\pi}{a} (n_x, n_y) \right|^2 \\
&= \frac{\hbar^2}{2m} \left(\left(k_x - \frac{2\pi}{a} n_x \right)^2 + \left(k_y - \frac{2\pi}{a} n_y \right)^2 \right) \tag{617}
\end{aligned}$$

Ho ottenuto la figura calcolando i rami della parabola su una griglia di punti \vec{k} nella prima zona di Brillouin (cella (0,0)) per i valori di n_x ed n_y mostrati, cioè:

n_x	n_y	$\epsilon(\vec{k}) = \frac{\hbar^2}{2m} \vec{k} - \vec{G} ^2$
-1	-1	$\frac{\hbar^2}{2m} \left(\left(k_x + \frac{2\pi}{a} \right)^2 + \left(k_y + \frac{2\pi}{a} \right)^2 \right)$
-1	0	$\frac{\hbar^2}{2m} \left(\left(k_x + \frac{2\pi}{a} \right)^2 + k_y^2 \right)$
-1	1	$\frac{\hbar^2}{2m} \left(\left(k_x + \frac{2\pi}{a} \right)^2 + \left(k_y - \frac{2\pi}{a} \right)^2 \right)$
0	-1	$\frac{\hbar^2}{2m} \left(k_x^2 + \left(k_y + \frac{2\pi}{a} \right)^2 \right)$
0	0	$\frac{\hbar^2}{2m} (k_x^2 + k_y^2)$
0	1	$\frac{\hbar^2}{2m} \left(k_x^2 + \left(k_y - \frac{2\pi}{a} \right)^2 \right)$
1	-1	$\frac{\hbar^2}{2m} \left(\left(k_x - \frac{2\pi}{a} \right)^2 + \left(k_y + \frac{2\pi}{a} \right)^2 \right)$
1	0	$\frac{\hbar^2}{2m} \left(\left(k_x - \frac{2\pi}{a} \right)^2 + k_y^2 \right)$
1	1	$\frac{\hbar^2}{2m} \left(\left(k_x - \frac{2\pi}{a} \right)^2 + \left(k_y - \frac{2\pi}{a} \right)^2 \right)$

Nel caso bidimensionale la struttura a bande e' ancora rappresentabile graficamente in modo completo, ma, come potete vedere dalla figura, la rappresentazione e' notevolmente complicata rispetto al caso monodimensionale.

Per questo motivo, normalmente la struttura a bande viene rappresentata in modo "parziale"; cioè, invece di rappresentare la relazione di dispersione su tutti i punti \vec{k} interni alla prima zona di Brillouin, si riporta solo una sezione del grafico ottenuta lungo una *traiettoria* (quasi sempre costituita da uno o piu' segmenti rettilinei).

Nella figura 56 e' mostrata la sezione del grafico bidimensionale ottenuta lungo la direzione $[k_x, 0]$ (la sezione dovrebbe stare al centro della cella (0,0), ma l'ho traslata nella cella (0,1) per renderla visibile).

Il calcolo e' fatto sempre nello stesso modo, ma questa volta $k_y = 0$ e quindi $\epsilon(\vec{k})$ diventa una funzione di una sola variabile:

$$\begin{aligned}
\epsilon(\vec{k}) &= \frac{\hbar^2}{2m} (\vec{k} - \vec{G})^2 = \frac{\hbar^2}{2m} |\vec{k} - \vec{G}|^2 = \frac{\hbar^2}{2m} \left| (k_x, 0) - \frac{2\pi}{a} (n_x, n_y) \right|^2 \\
&= \frac{\hbar^2}{2m} \left| \left(k_x - \frac{2\pi}{a} n_x, -\frac{2\pi}{a} n_y \right) \right|^2 \\
&= \frac{\hbar^2}{2m} \left(\left(k_x - \frac{2\pi}{a} n_x \right)^2 + \left(-\frac{2\pi}{a} n_y \right)^2 \right) = \epsilon(k_x)
\end{aligned} \tag{618}$$

Ci sono inoltre due aspetti “nuovi”, rispetto al caso monodimensionale.

230. Il primo e' che le bande possono “sovrapporsi” parzialmente o totalmente. Ho cercato di mettere in evidenza questo aspetto nella figura 57.

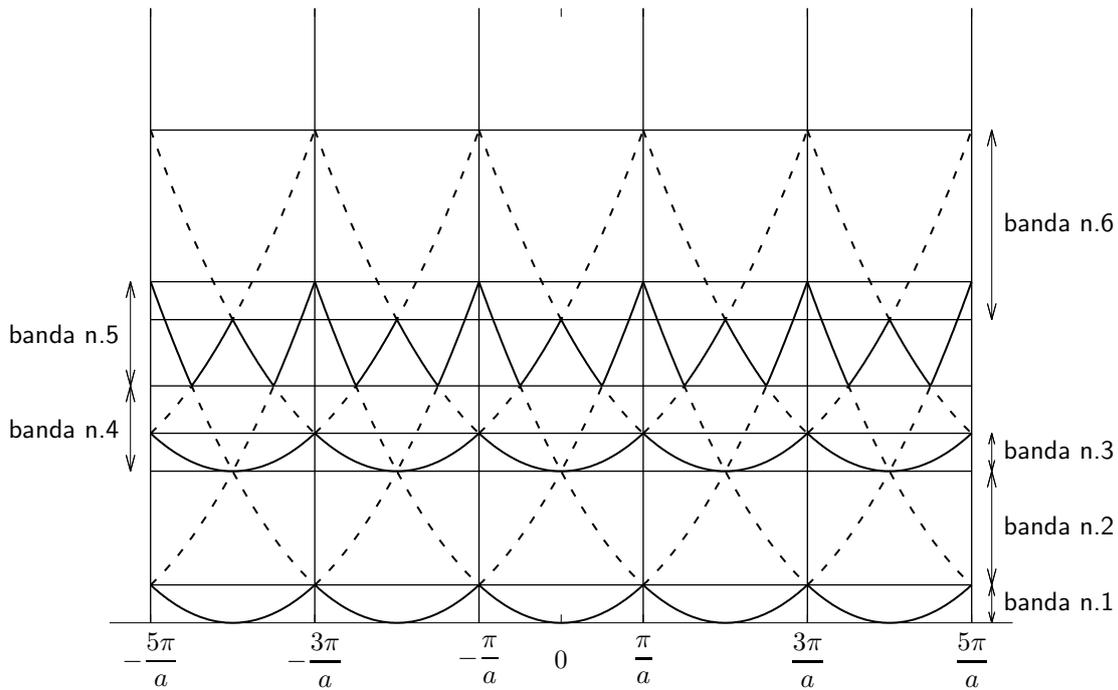


Figura 57: Struttura a bande (“repeated zone”) lungo la direzione $[k_x, 0]$ per un cristallo bidimensionale con reticolo di Bravais cubico semplice (costante di cella a) contenente elettroni liberi. Ho cercato di mettere in evidenza le bande alternando linee continue e tratteggiate; l’intervallo di energia in cui ciascuna banda e’ inclusa e’ indicato ai lati.

Tale figura e’ la versione “repeated zone” della sezione mostrata nella figura 56 (convincetevi che il contenuto della prima zona di Brillouin in figura 57 e’ identico al contenuto della sezione mostrata nella figura 56). Ho cercato di mettere in evidenza le bande successive alternando linee continue e tratteggiate; inoltre, l’ampiezza di ciascuna banda e’ indicata ai lati del grafico. Per capire come e’ fatto l’andamento di ciascuna banda, partite da un punto dell’asse delle ascisse

(che non sia il confine di una zona di Brillouin) e individuate gli autovalori in ordine di energia crescente: l'autovalore n. 1 individua la banda n. 1, l'autovalore n. 2 individua la banda n. 2 e così via; se fate questo per vari punti lungo l'asse delle ascisse vi dovrebbe diventare chiaro che le bande sono quelle mostrate.

Dalla figura si vede che:

- ⇒ le bande 1 e 2 non danno sovrapposizione con alcuna altra banda, ne' tra loro
- ⇒ la banda 3, invece, e' completamente sovrapposta alla banda 4, che e' piu' ampia (la banda 4 e' rappresentata con una linea tratteggiata, mentre la banda 3 e' a tratto intero)
- ⇒ le bande 5 e 6 sovrappongono parzialmente: la parte superiore della banda 5 sovrappone con la parte inferiore della banda 6

Come avremo modo di vedere, la sovrapposizione fra due o piu' bande puo' giocare un ruolo fondamentale nel determinare la natura di conduttore o isolante elettrico di un cristallo.

In pratica, la sovrapposizione di due bande successive n ed $n + 1$ significa che stati elettronici della banda n possono avere energia *maggiore* di quella di stati della banda $n + 1$. Naturalmente cio' puo' avvenire solo in corrispondenza a valori *diversi* del vettore d'onda (guardate la figura 57 e ricordate che, fissato un vettore d'onda, cioe' un punto nella prima zona di Brillouin, gli autovalori sono indicizzati dall'indice di banda in ordine di energia crescente e quindi si ha sempre: $\epsilon_n(\vec{k}) \leq \epsilon_{n+1}(\vec{k})$).

231. Il secondo aspetto legato alla dimensionalita' maggiore e' la degenerazione dei livelli energetici. Gia' nel caso monodimensionale, se guardate la figura 53, vi rendete subito conto che gli autovalori presentano una doppia degenerazione in corrispondenza ai confini delle zone di Brillouin $(0, \pm\pi/a, \pm 2\pi/a \dots)$, dove le parabole $\epsilon(k)$ si intersecano.

Questo effetto e' amplificato nei sistemi a dimensionalita' due o tre.

Se guardate la figura 57 potete vedere, ad esempio, che per $k_x = i2\pi/a$, $i = 0, \pm 1, \pm 2 \dots$ si ha una tripla degenerazione dell'energia rispetto all'indice di banda: le bande 2, 3 e 4 hanno un punto in comune.

Inoltre, lungo la sezione $k_y = 0$ della prima zona di Brillouin (cioe' sempre nella figura 57), la banda n. 3 e' data dall'eq. 618 nella forma:

$$\epsilon(k_x) = \frac{\hbar^2}{2m} \left(k_x^2 + \left(\frac{2\pi}{a} \right)^2 \right) \quad (619)$$

(cioe': la parabola $\epsilon(k_x) = (\hbar^2/(2m)) k_x^2$ traslata verso l'alto della quantita' $(\hbar^2/(2m)) (2\pi/a)^2$)

La forma 619 dell'eq. 618 si ottiene *sia* per $(n_x, n_y) = (0, -1)$, *che* per $(n_x, n_y) = (0, 1)$.

Ma allora questo significa che, nella figura 57, la banda n. 3 e' costituita in realta' da due bande degeneri sovrapposte, ottenute per $(n_x, n_y) = (0, \pm 1)$ nell'eq. 618:

$$\epsilon(\vec{k} - \vec{G}) = \epsilon\left((k_x, 0) - \frac{2\pi}{a}(0, -1)\right) = \epsilon\left((k_x, 0) - \frac{2\pi}{a}(0, 1)\right) \quad (620)$$

Come ulteriore esempio consideriamo, sempre lungo la sezione $k_y = 0$ della prima zona di Brillouin, la banda n. 4 in figura 57. Essa e' costituita da quattro

tratti ciascuno dei quali si ottiene con l'eq. 618 per opportuni valori di n_x ed n_y , secondo la seguente tabella:

nell'intervallo:	l'eq. 618 diventa:	possibili $(n_x n_y)$
$-\frac{\pi}{a} \leq k_x \leq -\frac{1}{2} \frac{\pi}{a}$	$\frac{\hbar^2}{2m} \left(\left(k_x + \frac{2\pi}{a} \right)^2 + \left(\frac{2\pi}{a} \right)^2 \right)$	$(-1, 1), (-1, -1)$
$-\frac{1}{2} \frac{\pi}{a} \leq k_x \leq 0$	$\frac{\hbar^2}{2m} \left(k_x - \frac{2\pi}{a} \right)^2$	$(1, 0)$
$0 \leq k_x \leq \frac{1}{2} \frac{\pi}{a}$	$\frac{\hbar^2}{2m} \left(k_x + \frac{2\pi}{a} \right)^2$	$(-1, 0)$
$\frac{1}{2} \frac{\pi}{a} \leq k_x \leq \frac{\pi}{a}$	$\frac{\hbar^2}{2m} \left(\left(k_x - \frac{2\pi}{a} \right)^2 + \left(\frac{2\pi}{a} \right)^2 \right)$	$(1, 1), (1, -1)$

Ad esempio: il primo tratto e' ottenuto dalla parabola che rappresenta la relazione di dispersione traslata verso sinistra di $2\pi/a$ e traslata verso l'alto di $(\hbar^2/(2m))(2\pi/a)^2$, e cosi' via. Dalla tabella potete vedere che i due tratti piu' esterni della banda n. 4 sono dati dalla sovrapposizione di due bande degeneri corrispondenti a due vettori \vec{G} diversi: $(-1, 1)$, $(-1, -1)$ e $(1, 1)$, $(1, -1)$, rispettivamente.

Tenete presente che in questo esempio ci siamo limitati a un numero molto piccolo di vettori \vec{G} (9). La degenerazione delle bande (parziale o totale) diventa sempre piu' frequente all'aumentare del numero di vettori \vec{G} che si considerano (e questo numero e' teoricamente infinito).

Di nuovo, come gia' osservato per il caso monodimensionale, la forma quadratica della relazione di dispersione per gli elettroni liberi induce una simmetria speculare rispetto all'asse delle ordinate.

232. Come gia' detto, la rappresentazione completa della relazione di dispersione per sistemi 3D non e' possibile e quindi in questo caso si ricorre quasi invariabilmente a rappresentazioni parziali su sezioni o traiettorie della prima zona di Brillouin.

A questo proposito va detto che per ognuna delle prime zone di Brillouin corrispondenti a ciascuno dei 14 reticoli di Bravais esiste una nomenclatura tradizionale per indicare punti "speciali" (di elevata simmetria) al loro interno oppure sulle loro facce. Ad esempio, in tutti i casi si indica con Γ il punto centrale della prima zona di Brillouin; per la prima zona di Brillouin di un reticolo cubico semplice vengono indicati coi simboli M , X e R , rispettivamente, i punti determinati dall'intersezione con le facce della zona (che e' un cubo) delle direzioni (reciproche) $[110]$, $[010]$ e $[111]$ ⁴² (figura 58). Quindi, ad esempio, potreste trovare la struttura a bande per un cristallo con reticolo cubico semplice calcolata lungo la traiettoria $\Gamma XM\Gamma RX$ e cosi' via.

Come esempio, vediamo l'estensione di quanto gia' fatto in due dimensioni al caso di un cristallo tridimensionale avente un reticolo di Bravais di tipo cubico semplice (lato di cella: a), nel modello "empty lattice".

Costruiamo la rappresentazione "reduced zone" per le bande "empty lattice" lungo la traiettoria $k(1, 1, 1)$ della prima zona di Brillouin (in termini di punti

⁴²Trovate la rappresentazione di tutte le (prime) zone di Brillouin con la specifica dei punti speciali per ciascuna di esse a questo indirizzo web: https://en.wikipedia.org/wiki/Brillouin_zone

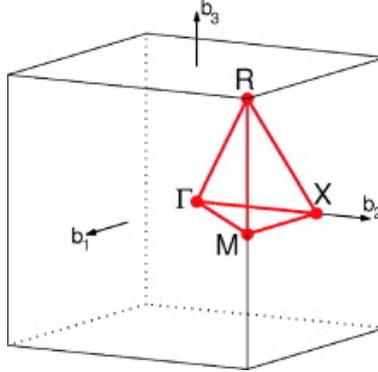


Figura 58:

speciali questo e' il segmento ΓR , fig. 58) (un esempio simile e' riportato per la direzione $k [1, 0, 0]$ nel libro di Kittel,¹⁹ e in quello di Ibach e Lüth²⁰).

Anche in questo caso ci limitiamo a considerare l'insieme di vettori \vec{G} che individua le celle prime vicine della prima zona di Brillouin; convincetevi che questi vettori sono in tutto 26: se considerate la prima zona di Brillouin al centro, ci sono 9 celle primitive "sopra" e "sotto" di essa e 8 celle primitive "attorno", per un totale di $9 + 9 + 8 = 26$ celle, ovvero 26 vettori \vec{G} , del tipo $2\pi/a (n_x, n_y, n_z)$, con $n_x, n_y, n_z = 0, \pm 1$. Contando anche il vettore nullo (corrispondente alla prima zona di Brillouin) ci sono in tutto 27 vettori \vec{G} .

Dobbiamo calcolare:

$$\begin{aligned}
 \frac{\hbar^2}{2m} (\vec{k} - \vec{G})^2 &= \frac{\hbar^2}{2m} |\vec{k} - \vec{G}|^2 = \frac{\hbar^2}{2m} \left| k(1, 1, 1) - \frac{2\pi}{a} (n_x, n_y, n_z) \right|^2 \\
 &= \frac{\hbar^2}{2m} \left| \left(k - \frac{2\pi}{a} n_x, k - \frac{2\pi}{a} n_y, k - \frac{2\pi}{a} n_z \right) \right|^2 \\
 &= \frac{\hbar^2}{2m} \left(\left(k - \frac{2\pi}{a} n_x \right)^2 + \left(k - \frac{2\pi}{a} n_y \right)^2 + \left(k - \frac{2\pi}{a} n_z \right)^2 \right) \quad (621)
 \end{aligned}$$

su tutti i punti del segmento $k(1, 1, 1)$ interno alla prima zona di Brillouin per $n_x, n_y, n_z = 0, \pm 1$.

Il segmento in questione e' la diagonale che unisce due vertici opposti della prima zona di Brillouin (per un reticolo cubico semplice questa e' un cubo di lato $2\pi/a$ centrato nell'origine). Le coordinate dei due vertici sono:

$$\frac{2\pi}{a} \left(-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2} \right) = -\frac{\pi}{a} (1, 1, 1) \quad (622)$$

e

$$\frac{2\pi}{a} \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2} \right) = \frac{\pi}{a} (1, 1, 1) \quad (623)$$

Quindi i punti del segmento $k(1, 1, 1)$ interno alla prima zona di Brillouin si ottengono quando k varia nell'intervallo:

$$-\frac{\pi}{a} \leq k \leq \frac{\pi}{a} \quad (624)$$

Calcolando l'espressione vista sopra per $-\pi/a \leq k \leq \pi/a$ e per $n_x, n_y, n_z = 0, \pm 1$ si ottiene la figura 59.

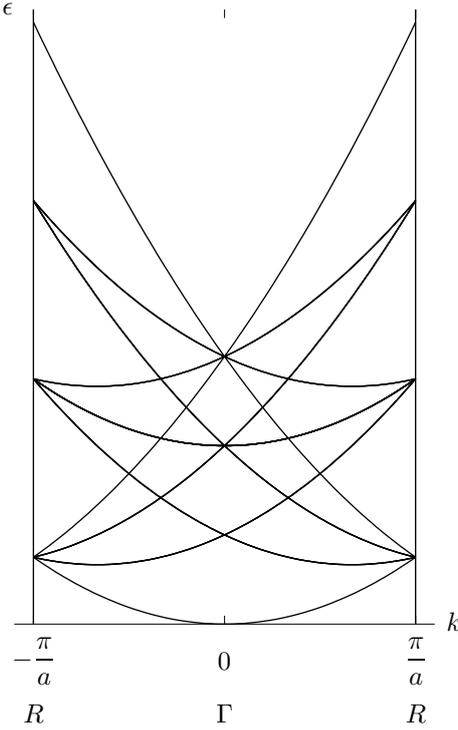


Figura 59: Struttura a bande calcolata lungo la direzione $[1, 1, 1]$ per un cristallo con reticolo di Bravais cubico semplice (costante di cella a) contenente elettroni liberi.

Oltre alla simmetria speculare rispetto all'asse delle ordinate, come nei casi precedenti, notate che le bande sono solo 10 (mettetevi su un punto generico dell'asse k , salite in verticale e contate quante volte intersecate le curve). Avendo fatto il calcolo per 27 vettori \vec{G} , dovrebbero esserci 27 bande; ne troviamo solo 10 perché le rimanenti sono degeneri con l'una o l'altra di queste 10.

Ad esempio:

$$\begin{aligned} \epsilon_{(n_x, n_y, n_z)=(1,0,0)}(k) &= \epsilon_{(n_x, n_y, n_z)=(0,1,0)}(k) = \epsilon_{(n_x, n_y, n_z)=(0,0,1)}(k) \\ &= \frac{\hbar^2}{2m} \left(2k^2 + \left(k - \frac{2\pi}{a} \right)^2 \right) \end{aligned} \quad (625)$$

233. Un altro modo di rappresentare la struttura a bande consiste nella costruzione della *superficie di Fermi*. Abbiamo già definito la superficie di Fermi per il caso degli elettroni liberi (punto 121): in quel caso, la superficie di Fermi è una sfera il cui raggio è uguale al modulo del più lungo vettore \vec{k} corrispondente a uno stato occupato. In generale, la superficie di Fermi è definita dall'equazione:

$$\epsilon(\vec{k}) = \epsilon_F \quad (626)$$

dove l'energia di Fermi ϵ_F è la massima energia fra quelle dei livelli elettronici occupati nel ground state.

Tutti gli stati corrispondenti ai punti sulla superficie di Fermi hanno la stessa energia, che è la massima energia posseduta dagli elettroni nel ground state. Nel caso della presenza di un potenziale periodico abbiamo visto che l'energia degli stati elettronici è ripartita in bande. La posizione delle bande rispetto all'energia di Fermi può dare luogo ad almeno tre scenari diversi:

1. se l'estremo inferiore ϵ_{min} di una banda è maggiore di ϵ_F :

$$\epsilon_F < \epsilon_{min} \quad (627)$$

gli stati della banda sono tutti vuoti e si dice che la banda è *vuota*.

2. se l'estremo superiore ϵ_{max} di una banda è minore di ϵ_F :

$$\epsilon_{max} < \epsilon_F \quad (628)$$

tutti gli stati della banda sono occupati e si dice che la banda è *completamente occupata o piena*

3. se ϵ_F è compreso nell'intervallo di energia che definisce una banda:

$$\epsilon_{min} < \epsilon_F < \epsilon_{max} \quad (629)$$

allora gli stati della banda a energia minore di ϵ_F saranno occupati, mentre quelli a energia maggiore di ϵ_F saranno vuoti. Si dice in questo caso che la banda è *parzialmente occupata*

Per ogni banda n parzialmente occupata è possibile rappresentare graficamente la superficie nello spazio \vec{k} definita da:

$$\epsilon_n(\vec{k}) = \epsilon_F \quad (630)$$

L'insieme delle superfici ottenute per tutte le bande parzialmente occupate è la superficie di Fermi per il caso del potenziale periodico. Le parti della superficie di Fermi originate da ciascuna banda vengono dette *rami* della superficie di Fermi.

Siccome la funzione $\epsilon_n(\vec{k})$ è periodica nello spazio \vec{k} , anche la superficie di Fermi è periodica e quindi anche per essa è invalsa la consuetudine di rappresentarla, sulla base di criteri di opportunità, in uno dei tre schemi visti: "reduced zone", "extended zone" o "repeated zone".

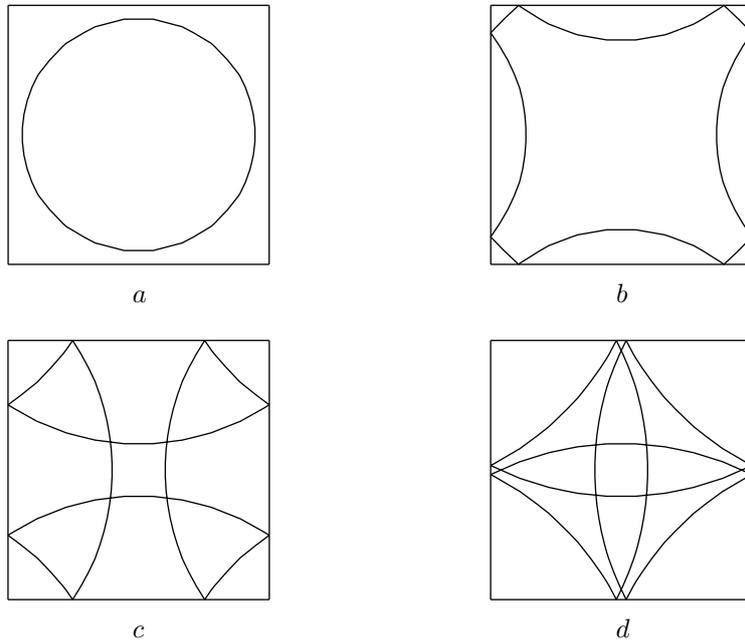


Figura 60: Superficie di Fermi ricavata dalla figura 56 per valori crescenti dell'energia di Fermi nella sequenza a, b, c, d . In ogni caso i rami della superficie sono stati riportati all'interno della prima zona di Brillouin ("reduced zone representation").

Per esemplificare, nella figura 60 ho rappresentato la superficie di Fermi per il cristallo bidimensionale con cella primitiva quadrata nel modello "empty lattice" che abbiamo già visto al punto 229 .

Essendo il sistema 2D, la superficie di Fermi si riduce a una combinazione di una o più traiettorie sul piano xy . I grafici della figura 60 sono ottenuti sezionando il grafico della figura 56 con il piano $z = \epsilon_F$, per diversi valori dell'energia di Fermi ϵ_F . Così, ad esempio, nella figura 60a l'energia di Fermi è minore dell'estremo superiore della prima banda, il cui grafico è quello di una parabola bidimensionale (guardate la figura 56): ne segue che la superficie di Fermi è una circonferenza; per valori dell'energia di Fermi maggiori (figura 60b,c,d) la superficie di Fermi diventa più complicata.

Siccome le bande possono sovrapporsi (punto 229) e quindi l'attribuzione di un ramo della superficie di Fermi ad una particolare banda può non essere univoco, si usa anche rappresentare la superficie di Fermi in relazione a successive zone di Brillouin; cioè, ad esempio nello schema "reduced zone", si rappresentano separatamente all'interno della cella primitiva del reticolo reciproco (prima zona di Brillouin) i rami della superficie di Fermi per la seconda, terza etc zona di Brillouin (una rappresentazione di questo tipo è mostrata nel libro di Kittel²¹). A questo indirizzo web:

http://lampx.tugraz.at/~hadley/ss2/fermisurface/2d_fermisurface/2dsquare.php

trovate un'applicazione interattiva che vi permette di vedere come varia la superficie di Fermi al variare dell'energia di Fermi per lo stesso sistema a cui si riferiscono le figure 56 e 60; l'applicazione mostra anche la rappresentazione periodica ("repeated zone") dei rami della superficie di Fermi corrispondenti alle prime quattro zone di Brillouin.

234. Densità degli stati.

Il concetto di densità degli stati che abbiamo introdotto a proposito degli elettroni liberi (p. 133) si estende senza complicazioni anche al caso della presenza di un potenziale periodico. Siccome in questo caso gli stati elettronici sono organizzati in bande, ha senso definire la densità degli stati $g_n(\epsilon(\vec{k}'))$ per una singola banda n . Cioè: usando ad esempio la definizione generale 450, per la banda n si avrà:

$$g_n(\epsilon(\vec{k}')) = \frac{1}{V} \sum_{\vec{k}} \delta(\epsilon(\vec{k}') - \epsilon_n(\vec{k})) \quad (631)$$

dove $\epsilon_n(\vec{k})$ è l'energia dello stato corrispondente al vettore d'onda permesso \vec{k} nella banda n .

Chiaramente, la densità degli stati totale è ottenuta come:

$$g(\epsilon(\vec{k}')) = \sum_n g_n(\epsilon(\vec{k}')) = \frac{1}{V} \sum_{n\vec{k}} \delta(\epsilon(\vec{k}') - \epsilon_n(\vec{k})) \quad (632)$$

Siccome in presenza di un potenziale periodico l'energia degli stati elettronici è una funzione periodica nello spazio \vec{k} (punto 223), per il calcolo della densità degli stati in tali condizioni si usa l'espressione:

$$g_n(\epsilon(\vec{k}' - \vec{G}_n)) \quad (633)$$

dove \vec{k}' è un punto interno alla prima zona di Brillouin e \vec{G}_n è il vettore del reticolo reciproco che bisogna sommare o sottrarre a \vec{k}' per ottenere il valore di energia del corrispondente stato nella banda n .

Nel caso del semplice modello "empty lattice", la densità degli stati coincide con quella che avevamo già ottenuto per gli elettroni liberi (cioè in assenza di un potenziale periodico).

Infatti in questo caso il numero cumulativo di stati permessi, $N(\epsilon)$ per la banda n è dato da (riguardate l'eq. 414):

$$N(\epsilon(\vec{k} - \vec{G}_n)) = \frac{V}{6\pi^2} \left(\frac{2m}{\hbar^2}\right)^{\frac{3}{2}} (\epsilon(\vec{k} - \vec{G}_n))^{\frac{3}{2}} \quad (634)$$

con $\vec{k} \in$ prima zona di Brillouin e \vec{G}_n opportuno vettore del reticolo reciproco. La densità degli stati per la banda n è ottenuta derivando $N(\epsilon(\vec{k} - \vec{G}_n))$ rispetto a $\epsilon(\vec{k} - \vec{G}_n)$ (eq. 415) e tale derivata è ovviamente identica a quella già vista a suo tempo per questo sistema.

Il band gap⁴³

⁴³Mi rifiuto di tradurre in italiano questo termine inglese, che è universalmente utilizzato; "gap" vuol dire apertura, vuoto, intervallo, interruzione; quindi il band gap è, come avremo modo di spiegare, un intervallo di energia proibito che separa due bande successive.

235. Abbiamo preso in considerazione il modello “empty lattice” e abbiamo visto che la presenza del potenziale periodico determina l’organizzazione degli autovalori dell’hamiltoniano in bande di energia indicizzate dall’indice di banda n (figura 53).

Tuttavia, nel modello “empty lattice” il potenziale periodico $U(\vec{r})$ e’ supposto infinitesimo, cioe’ talmente piccolo da essere sempre trascurabile rispetto all’energia cinetica posseduta dagli elettroni (e rappresentata dal termine: $-(\hbar^2/2m)\nabla^2$ nell’hamiltoniano).

Vedremo ora che, quando il potenziale non puo’ essere trascurato (il che, ovviamente, e’ la norma quando si vuole considerare qualsiasi caso che abbia realta’ fisica), si verifica un altro fenomeno che gioca un ruolo fondamentale nella determinazione delle proprieta’ dei solidi cristallini.

Nella figura 53 potete vedere che le bande di energia che avevamo trovato col semplicistico modello “empty lattice” sono *contigue*, cioe’ al crescere dell’energia si passa con continuita’ da una banda alla successiva.

Ebbene, se il potenziale periodico nel cristallo diventa apprezzabile, le bande si *discostano*, cioe’ fra una banda e la successiva compare una regione di energia in cui *non* sono presenti autovalori; queste zone di energia proibite vengono dette *band gaps* e, come vedremo, sono di importanza fondamentale nella discussione delle proprieta’ dei solidi cristallini.

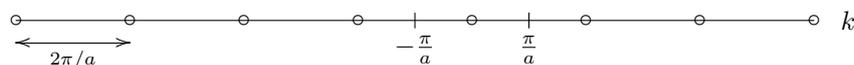
Il modello degli elettroni quasi liberi

236. E’ possibile dare una giustificazione sia matematica che fisica della formazione del band gap utilizzando un modello un po’ piu’ realistico di quello “empty lattice”; tale modello viene detto “nearly free electron model” o modello degli elettroni quasi liberi. Si intende con cio’ un modello in cui il potenziale periodico e’ apprezzabile (quindi non e’ infinitesimo come nel modello “empty lattice”), ma e’ comunque sufficientemente piccolo da rendere valida l’assunzione che i livelli energetici degli elettroni siano *solo di poco* diversi da quelli degli elettroni liberi nel gas di Fermi (eq. 344). Le parole “*solo di poco*” assumeranno un significato piu’ chiaro nel seguito.

237. Una giustificazione matematica del band gap.⁴⁴

Ora daremo una giustificazione di tipo matematico per la generazione del band gap nel modello degli elettroni quasi liberi e in un sistema costituito da un cristallo monodimensionale avente una cella primitiva di lunghezza a (e quindi una cella primitiva reciproca di “lunghezza” $2\pi/a$).

Il sistema e’ identico a quello che avevamo considerato per discutere la struttura a bande al punto 225: il reticolo reciproco e’ una griglia monodimensionale e la prima zona di Brillouin e’ l’intervallo $[-\pi/a, \pi/a]$:



Siccome il sistema e’ monodimensionale, il “vettore” d’onda e i “vettori” del reticolo reciproco sono in realta’ degli scalari. Il loro segno individua una delle due sole possibili orientazioni: nel verso positivo oppure in quello negativo dell’asse k .

Posto:

⁴⁴Il materiale di questa parte e’ basato sui libri di Kittel,²² Ibach e Lüth,²³ e Ashcroft e Mermin.²⁴

$$g = \frac{2\pi}{a} \quad (635)$$

i vettori del reticolo reciproco hanno la forma:

$$G = pg \quad p \in \mathbb{Z} \quad (636)$$

Riprendiamo il sistema che costituisce l'equazione centrale. A p. 159 avevamo ripartito tale sistema in N sottosistemi indipendenti, dove N e' il numero di celle primitive che costituisce il campione macroscopico.

Riporto per vostra comodita' il j -esimo sottosistema:

$$\begin{aligned} \left(\frac{\hbar^2 k_j^2}{2m} - \epsilon \right) C_{\vec{k}_j} + U_{\vec{G}_1} C_{\vec{k}_j - \vec{G}_1} + U_{\vec{G}_2} C_{\vec{k}_j - \vec{G}_2} + \dots &= 0 \\ \left(\frac{\hbar^2 |\vec{k}_j - \vec{G}_1|^2}{2m} - \epsilon \right) C_{\vec{k}_j - \vec{G}_1} + U_{\vec{G}_1} C_{\vec{k}_j - \vec{G}_1 - \vec{G}_1} + U_{\vec{G}_2} C_{\vec{k}_j - \vec{G}_1 - \vec{G}_2} + \dots &= 0 \\ \left(\frac{\hbar^2 |\vec{k}_j - \vec{G}_2|^2}{2m} - \epsilon \right) C_{\vec{k}_j - \vec{G}_2} + U_{\vec{G}_1} C_{\vec{k}_j - \vec{G}_2 - \vec{G}_1} + U_{\vec{G}_2} C_{\vec{k}_j - \vec{G}_2 - \vec{G}_2} + \dots &= 0 \\ \dots & \end{aligned}$$

\vec{k}_j e' il j -esimo punto interno alla prima zona di Brillouin (\vec{k}_j e' anche, in generale, un punto \vec{k} del reticolo reciproco macroscopico).

Per il seguito ci conviene riscrivere questo sottosistema in un modo piu' funzionale a cio' che diremo. Tralasciamo l'indice j , riscriviamo i vettori \vec{k} e \vec{G} come degli scalari (cioe' quello che in realta' essi sono per il caso monodimensionale che stiamo considerando) e scriviamo i G_i sulla base della 636.

$$\begin{aligned}
& \left(\frac{\hbar^2 k^2}{2m} - \epsilon \right) C_k + \sum_{i=-\infty}^{\infty} U_{ig} C_{k-ig} = 0 \\
& \left(\frac{\hbar^2 (k-g)^2}{2m} - \epsilon \right) C_{k-g} + \sum_{i=-\infty}^{\infty} U_{ig} C_{k-g-ig} = 0 \\
& \left(\frac{\hbar^2 (k-2g)^2}{2m} - \epsilon \right) C_{k-2g} + \sum_{i=-\infty}^{\infty} U_{ig} C_{k-2g-ig} = 0 \\
& \dots \\
& \left(\frac{\hbar^2 (k+g)^2}{2m} - \epsilon \right) C_{k+g} + \sum_{i=-\infty}^{\infty} U_{ig} C_{k+g-ig} = 0 \\
& \left(\frac{\hbar^2 (k+2g)^2}{2m} - \epsilon \right) C_{k+2g} + \sum_{i=-\infty}^{\infty} U_{ig} C_{k+2g-ig} = 0 \\
& \dots
\end{aligned}$$

Notate che, essendo k e g degli scalari, la notazione del modulo quadrato e' diventata un semplice elevamento al quadrato. Notate anche che, siccome il sistema e' monodimensionale, i "vettori" G possono avere solo due versi: quello positivo (cioe' nel verso positivo dell'asse k) e quello negativo (cioe' nel verso negativo dell'asse k); questo e' stato messo in evidenza nella riscrittura del sottosistema, dove ora compaiono esplicitamente i multipli di g sottratti o sommati a k (ora g e' uno scalare positivo, eq. 635). Per lo stesso motivo, le sommatorie corrono sull'indice intero i che varia da $-\infty$ a ∞ .

Ricordate che questo sottosistema ha infinite soluzioni (infiniti sets di coefficienti C , ciascun set con un suo autovalore) indicizzate dall'indice di banda n :

$$\epsilon_{kn}, \quad \Psi_{kn}(r) = \sum_{j=-\infty}^{\infty} C_{(k-jg)n} \exp(i(k-jg)r)$$

Nel sottosistema non ho indicato n sia per alleggerire la notazione e sia perche', come vedrete nel seguito, dovremo considerare non una sola banda, ma due bande successive.

238. La generica equazione del sottosistema puo' essere scritta nel modo seguente:

$$\left(\frac{\hbar^2 (k-jg)^2}{2m} - \epsilon \right) C_{k-jg} + \sum_{i=-\infty}^{\infty} U_{ig} C_{k-jg-ig} = 0, \quad j \in \mathbb{Z} \quad (637)$$

E' importante osservare che, essendo per ipotesi il potenziale periodico molto piccolo, gli U_{ig} saranno molto piccoli e quindi le sommatorie:

$$\sum_{i=-\infty}^{\infty} U_{ig} C_{k-jg-ig}$$

saranno anch'esse molto piccole.

239. Come prima cosa dimostreremo che, per tutti i punti della prima zona di Brillouin lontani⁴⁵ da $k = 0$ e $k = \pm\pi/a = \pm g/2$ la dipendenza dell'energia da k e', con buona approssimazione⁴⁶, quella parabolica del gas di Fermi. I punti $k = 0$ e $k = \pm\pi/a = \pm g/2$ sono speciali e vedremo che proprio quando k e' vicino a uno di essi la presenza di un potenziale periodico non nullo determina la creazione di un band gap.

Riprendiamo la generica equazione del sottosistema corrispondente al punto k interno alla prima zona di Brillouin (eq. 637) e riscriviamola cosi':

$$C_{k-jg} = \frac{\sum_{i=-\infty}^{\infty} U_{ig} C_{k-jg-ig}}{\left(\epsilon - \frac{\hbar^2(k-jg)^2}{2m}\right)} \quad (638)$$

Questa e' un'espressione per il generico coefficiente C_{k-jg} del sottosistema. Ovviamente c'e' un problema perche' al secondo membro ci sono delle incognite: tutti i coefficienti $C_{k-jg-ig}$ e l'energia ϵ . Tuttavia, grazie alle ipotesi che abbiamo fatto sul potenziale periodico (elettroni quasi liberi, cioe' potenziale periodico molto piccolo), possiamo ottenere una stima del secondo membro e quindi del coefficiente C_{k-jg} .

Come gia' accennato al punto 238, essendo il potenziale periodico molto piccolo, i coefficienti del suo sviluppo di Fourier saranno "piccoli" e quindi tutta la sommatoria al numeratore del secondo membro della 638 sara' "piccola" (diciamo $\rightarrow 0$).

Inoltre, come gia' sottolineato (punto 236), un secondo aspetto del modello degli elettroni quasi liberi e' che gli autovalori dell'equazione di Schrödinger differiscono di poco da quelli del gas di Fermi.

Allora, limitandoci per il momento alla banda di energia piu' bassa (quella per $n = 1$), possiamo scrivere il valore di ϵ come quello per il gas di Fermi piu' un discostamento α tendente a zero:

$$\epsilon = \frac{\hbar^2 k^2}{2m} + \alpha \quad \text{con: } |\alpha| \rightarrow 0 \quad (639)$$

(non perdetevi: k e' il punto interno alla prima zona di Brillouin che individua un intero sottosistema dell'equazione centrale 591; la 638 rappresenta la j -esima equazione di tale sottosistema)

Quindi possiamo riscrivere la 638 come:

$$C_{k-jg} = \frac{\sum_{i=-\infty}^{\infty} U_{ig} C_{k-jg-ig}}{\left(\frac{\hbar^2 k^2}{2m} + \alpha - \frac{\hbar^2(k-jg)^2}{2m}\right)} \quad (640)$$

A questo punto si ha:⁴⁷

⁴⁵Mi rendo conto che "lontano" in matematiche non significa nulla, ma non stiamo facendo matematica; il significato di "lontani" sara' chiaro dal seguito.

⁴⁶Anche queste parole si potrebbero chiarire in modo piu' preciso; questo lo fanno Ashcroft e Mermin nel loro libro, p. 152-156.

⁴⁷Quello che segue e' molto poco formale. Si potrebbe essere piu' rigorosi usando la teoria delle perturbazioni, ma mi sembra un'inutile complicazione; il senso di quello che voglio dire dovrebbe essere chiaro comunque.

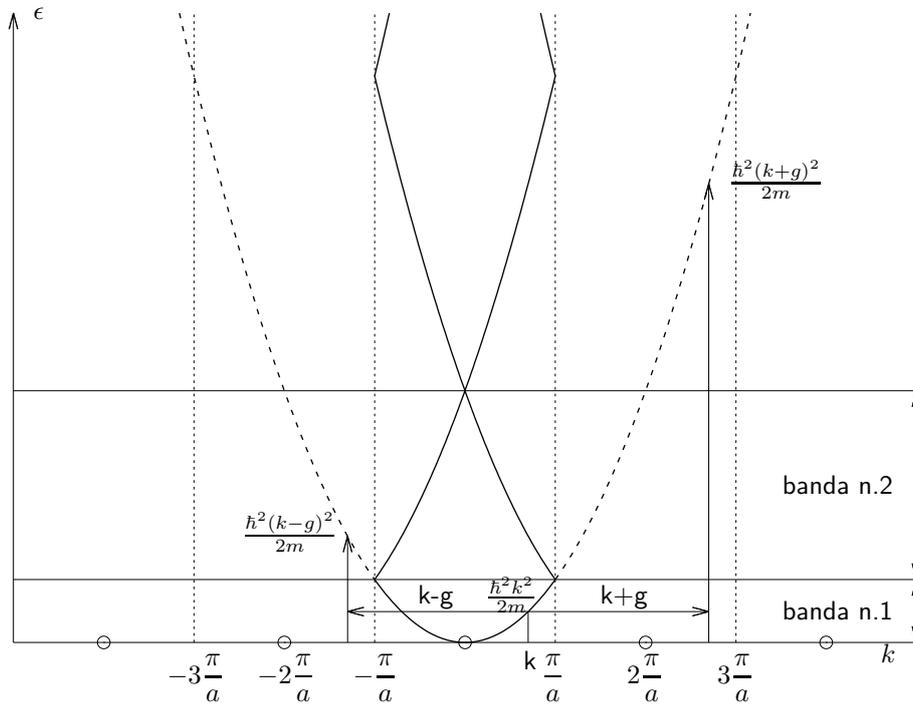


Figura 61:

$\forall j \neq 0$ il denominatore e' finito (cioe' non piccolo, perche' la differenza: $(\hbar^2 k^2/2m - \hbar^2 (k - jg)^2/2m)$ e' finita e $|\alpha| \rightarrow 0$); siccome il numeratore e' piccolo ($\rightarrow 0$) siamo in una situazione del tipo:⁴⁸

$$\frac{[\text{qualcosa} \rightarrow 0]}{[\text{qualcosa finito}]} = [\text{qualcosa} \rightarrow 0]$$

$j = 0$ la frazione diventa:

$$C_k = \frac{\sum_{i=-\infty}^{\infty} U_{ig} C_{k-jg-ig}}{\alpha} \quad (641)$$

e quindi siamo in una situazione del tipo:

$$\frac{[\text{qualcosa} \rightarrow 0]}{[\text{qualcosa} \rightarrow 0]} = [\text{qualcosa finito}]$$

Ho rappresentato la situazione nella figura 61. Sono mostrate le bande n.1 e 2 e la parabola che da' la relazione di dispersione per il gas di Fermi. Nella prima zona di Brillouin ($[-\pi/a, \pi/a]$) e' preso un punto k ; l'autovalore piu' basso ($n = 1$) corrispondente a tale punto k e' uguale (con buona approssimazione) a quello che si puo' individuare sulla parabola valida per il gas di Fermi: $\hbar^2 k^2/2m$. Nella figura sono indicate anche le immagini del punto k che si ottengono sommando e

⁴⁸Passatemi la notazione poco ortodossa: spero aiuti a chiarire.

sottraendo ad esso il vettore piu' corto del reticolo reciproco ($j = \pm 1$ nella 638), ($k + g$) e ($k - g$): i corrispondenti valori dell'energia calcolati sulla parabola del gas di Fermi differiscono per una quantita' finita da quello che si ha per k . Dalla figura potete anche comprendere facilmente che la differenza:

$$\left(\frac{\hbar^2 k^2}{2m} - \frac{\hbar^2 (k - jg)^2}{2m} \right) \quad (642)$$

diventa sempre piu' grande (in valore assoluto) all'aumentare di j .

240. In definitiva, abbiamo trovato che per ogni punto della prima zona di Brillouin lontano da $k = 0$ e $k = \pm\pi/a = \pm g/2$ tutti i coefficienti dello sviluppo in serie di Fourier della soluzione dell'equazione di Schrödinger (per $n = 1$) tendono a zero, *tranne* il coefficiente C_k (cioe' C_{k-jg} con $j = 0$). Inoltre, l'autovalore corrispondente e' uguale (a meno di un errore tendente a zero) a quello previsto dal modello del gas di Fermi. Cioe':

$$\Psi_{k1}(r) = C_{k1} \exp(ikr) \quad (643)$$

$$\epsilon_1(k) = \frac{\hbar^2 k^2}{2m} \quad (644)$$

Questo e' esattamente il risultato che avevamo gia' trovato per il gas di Fermi (il coefficiente C_k va comunque rideterminato per la condizione di normalizzazione); tutto cio' e' assolutamente ragionevole: stiamo assumendo che il potenziale periodico sia molto piccolo e quindi ci aspettiamo che tutto funzioni in modo molto simile al caso in cui il potenziale periodico e' rigorosamente nullo (gas di Fermi).

241. Per le bande successive alla prima ($n > 1$) vale un discorso analogo; l'unica differenza e' che l'autovalore corrispondente al punto k interno alla prima zona di Brillouin e' ottenuto, in questo caso, calcolando la parabola del gas di Fermi per un valore di k esterno alla prima zona di Brillouin; a parte cio', i passaggi restano identici a quelli svolti al punto 239.

La situazione per la banda n.2 e' rappresentata nella figura 62.

L'autovalore della seconda banda corrispondente a k si trova sul ramo di parabola A che e' il ramo A' traslato di g verso destra ed e' quindi ottenuto calcolando la parabola del gas di Fermi in ($k - g$) (guardate il "percorso" indicato dalle frecce). Analogamente al caso della banda n.1 si avra':

$$C_{k-jg} = \frac{\sum_{i=-\infty}^{\infty} U_{ig} C_{k-jg-ig}}{\left(\frac{\hbar^2 (k-g)^2}{2m} + \alpha - \frac{\hbar^2 (k-jg)^2}{2m} \right)} \quad (645)$$

e quindi nello sviluppo in serie di Fourier solo il coefficiente C_{k-g} (cioe' C_{k-jg} con $j = 1$) avra' un valore apprezzabilmente diverso da zero (mentre tutti gli altri saranno praticamente nulli). Per la seconda banda il risultato sara' pertanto:

$$\Psi_{k2}(r) = C_{(k-g)2} \exp(i(k-g)r) \quad (646)$$

$$\epsilon_2(k) = \frac{\hbar^2 (k-g)^2}{2m} \quad (647)$$

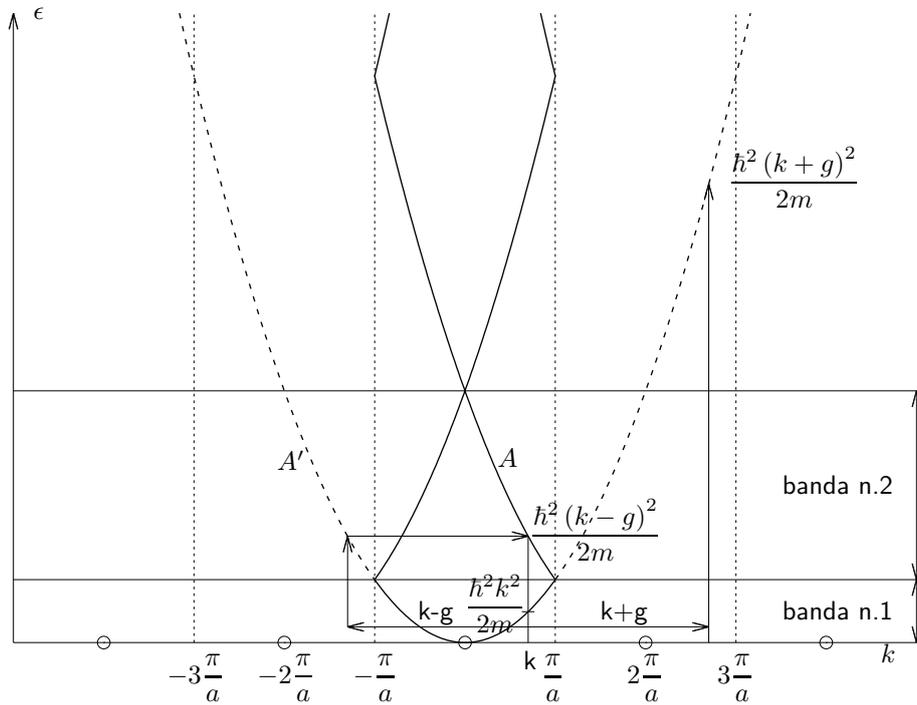


Figura 62:

Di nuovo, questo e' esattamente il risultato che si ottiene per il gas di Fermi.

242. In generale, per k lontano da 0 e $\pm g/2$ e per ogni banda n , l'eq. 638 assume la forma:

$$C_{k-jg} = \frac{\sum_{i=-\infty}^{\infty} U_{ig} C_{k-jg-ig}}{\left(\frac{\hbar^2(k-pg)^2}{2m} + \alpha - \frac{\hbar^2(k-jg)^2}{2m} \right)} \quad (648)$$

dove $p \in \mathbb{Z}$ e' un opportuno numero intero che, fissato k nella prima zona di Brillouin, determina l'autovalore n -esimo tramite:

$$\epsilon_{k,n} = \frac{\hbar^2(k-pg)^2}{2m} + \alpha \quad (649)$$

Notate che p e' funzione sia di k che dell'indice di banda n . Ad esempio, con riferimento alla figura 61, per il punto k mostrato e la terza banda si avrebbe:

$$p(k, 3) = -1 \quad (650)$$

cioe':

$$\epsilon_{k,3} = \frac{\hbar^2(k+g)^2}{2m} + \alpha \quad (651)$$

mentre se considerassimo il punto k' simmetrico rispetto all'origine di quello mostrato, si avrebbe:

$$p(k', 3) = 1 \quad (652)$$

cioe':

$$\epsilon_{k',3} = \frac{\hbar^2 (k-g)^2}{2m} + \alpha \quad (653)$$

La cosa importante in questo contesto e' che, in ogni caso, il denominatore dell'eq. 638 e' sempre finito tranne *nel singolo caso* in cui $j = p$ e quindi:

$$\begin{aligned} j \neq p \quad C_{k-jg} &= \frac{\sum_{i=-\infty}^{\infty} U_{ig} C_{k-jg-ig}}{\left(\frac{\hbar^2 (k-pg)^2}{2m} + \alpha - \frac{\hbar^2 (k-jg)^2}{2m} \right)} \rightarrow 0 \\ j = p \quad C_{k-pg} &= \frac{\sum_{i=-\infty}^{\infty} U_{ig} C_{k-pg-ig}}{\alpha} \neq 0 \text{ (finito)} \end{aligned} \quad (654)$$

Cioe': per k lontano da 0 e $\pm g/2$ e per ogni banda n , *un solo coefficiente* dello sviluppo in serie di Fourier della $\Psi_{kn}(\vec{r})$ "sopravvive" e ritroviamo la soluzione del modello "empty lattice":

$$\Psi_{kn}(r) = C_{(k-pg)n} \exp(i(k-pg)r) \quad (655)$$

$$\epsilon_n(k) = \frac{\hbar^2 (k-pg)^2}{2m} \quad (656)$$

243. Fino a questo punto sembra non ci sia nulla di nuovo rispetto a quanto gia' sapevamo dal modello "empty lattice".

Vediamo ora cosa succede alle bande n. 1 e 2 quando il punto k che consideriamo si trova in un piccolo intorno sinistro del confine della prima zona di Brillouin: $k \in [g/2 - \delta, g/2]$.

Innanzitutto osserviamo che per $k = g/2 = \pi/a$ e nel modello "empty lattice" le bande n.1 e 2 (cosi' come le bande n. 3 e 4 e cosi' via) "si toccano" (guardate ad esempio la figura 62); cioe': nel modello "empty lattice" l'energia corrispondente al confine della prima zona di Brillouin e' doppiamente degenerata rispetto all'indice di banda n :

$$k = \frac{g}{2} \Rightarrow \epsilon_1(k) = \epsilon_2(k) \quad (657)$$

Ora vedremo che se il potenziale periodico viene "leggermente alzato" rispetto al valore nullo (transizione dal modello "empty lattice" al modello degli elettroni quasi liberi), questa degenerazione viene rimossa.

Come nei casi precedenti, il punto di partenza e' ancora l'eq. 638:

$$C_{k-jg} = \frac{\sum_{i=-\infty}^{\infty} U_{ig} C_{k-jg-ig}}{\left(\epsilon - \frac{\hbar^2 (k-jg)^2}{2m} \right)}$$

Ci concentriamo sul piccolo intervallo di energie:

$$\frac{\hbar^2}{2m} \left(\frac{g}{2} \right)^2 - \gamma_1 \leq \epsilon \leq \frac{\hbar^2}{2m} \left(\frac{g}{2} \right)^2 + \gamma_2 \quad (658)$$

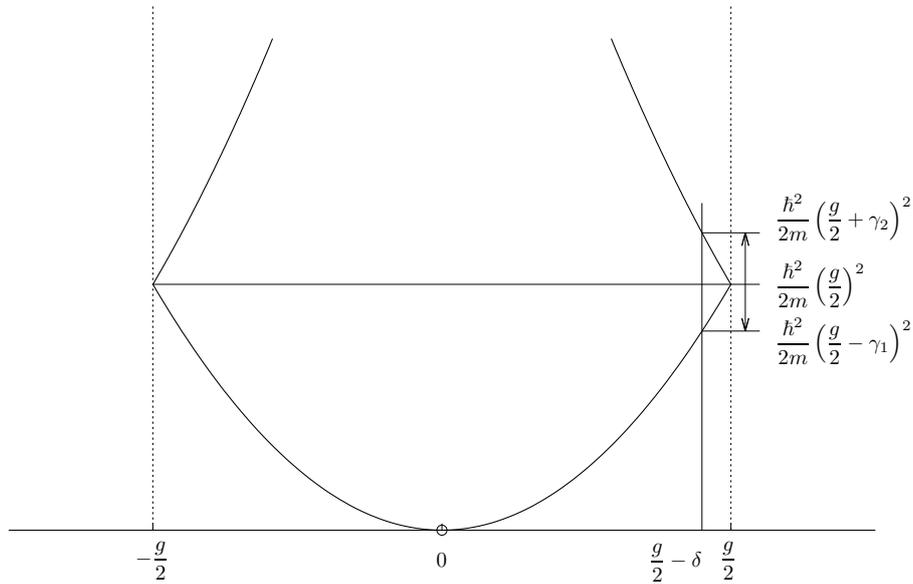


Figura 63:

determinato dall'intervallo $[g/2 - \delta, g/2]$ sull'asse k (guardate la figura 63⁴⁹). Siccome ora siamo interessati non a un singolo valore di ϵ ma ai valori di ϵ contenuti in un intervallo, lasciamo ϵ indicato come variabile, sapendo che il suo valore e' uguale, a meno di un errore infinitesimo α , a quello del gas di Fermi:

$$\epsilon = \frac{\hbar^2 \left(\frac{g}{2}\right)^2}{2m} + \alpha \quad |\alpha| \rightarrow 0 \quad (659)$$

Per lo stesso motivo, lasciamo anche k indicato come variabile, sapendo che il suo valore e' molto vicino (o identicamente uguale) a $g/2$.

Nei casi precedenti avevamo visto che il denominatore della 638 diventava molto piccolo solamente *per un singolo* coefficiente e cio' faceva si che quel singolo coefficiente fosse in pratica l'unico presente nello sviluppo in serie di Fourier dell'autofunzione.

Il caso presente si differenzia proprio in questo.

Essendo:

$$k \approx \frac{g}{2} \quad (660)$$

$$\epsilon = \frac{\hbar^2 \left(\frac{g}{2}\right)^2}{2m} + \alpha \quad (661)$$

la 638 assume la forma:

⁴⁹Nonostante possa sembrare dalla figura, l'intervallo sull'asse ϵ non e' simmetrico rispetto al valore comune alle bande n. 1 e 2 (e cioe': $\frac{\hbar^2}{2m} \left(\frac{g}{2}\right)^2$): riuscite a vedere perche'?

$$C_{k-jg} = \frac{\sum_{i=-\infty}^{\infty} U_{ig} C_{k-jg-ig}}{\left(\frac{\hbar^2 \left(\frac{g}{2}\right)^2}{2m} + \alpha - \frac{\hbar^2 \left(\frac{g}{2} - jg\right)^2}{2m} \right)}$$

da cui si vede che ora, diversamente dal caso precedente in cui k era "lontano" da $g/2$, ci sono *due* valori di j per cui il denominatore della 638 diventa molto piccolo. Si ha infatti:

$$\frac{\hbar^2 \left(\frac{g}{2}\right)^2}{2m} + \alpha - \frac{\hbar^2 \left(\frac{g}{2} - jg\right)^2}{2m} = \alpha \rightarrow 0 \iff j = 0 \text{ OPPURE } j = 1 \quad (662)$$

(Come vedremo fra poco al punto 250, una cosa simile avviene anche quando k e' interno a un piccolo intervallo centrato a $k = 0$)

Per qualsiasi altro valore di j la differenza al denominatore della 638 e' sempre finita e quindi i corrispondenti coefficienti di Fourier diventano trascurabili (con un ragionamento identico a quello gia' fatto al punto 239).

In definitiva, abbiamo scoperto che, nelle condizioni che stiamo considerando e in particolare quando k e' vicino al confine della prima zona di Brillouin, tutti i coefficienti dello sviluppo in serie di Fourier delle soluzioni dell'equazione di Schrödinger corrispondenti ad autovalori compresi nell'intervallo mostrato nella figura 63 (quindi appartenenti alle bande n. 1 e 2) sono trascurabili, tranne:

$$C_k \text{ e } C_{k-g}$$

(cioe': C_{k-jg} per $j = 0$ e $j = 1$)

Vediamo ora che forma approssimata assume il sottosistema 637 in queste condizioni.

Innanzitutto, siccome solo i coefficienti C_k e C_{k-g} hanno un valore apprezzabile, la sommatoria della 637 puo' essere approssimata ai soli due termini contenenti i due coefficienti. Notate che questi due termini della sommatoria su i non sono gli stessi in tutte le equazioni, ma dipendono da j (j indicizza le equazioni del sottosistema):

$$\sum_{i=-\infty}^{\infty} U_{ig} C_{k-jg-ig} \quad (663)$$

$$C_{k-jg-ig} = C_k \iff i = -j \quad (664)$$

$$C_{k-jg-ig} = C_{k-g} \iff i = 1 - j \quad (665)$$

e quindi:

$$\sum_{i=-\infty}^{\infty} U_{ig} C_{k-jg-ig} \approx U_{-jg} C_k + U_{(1-j)g} C_{k-g} \quad (666)$$

Il sottosistema 637 diventa:

$$\left(\frac{\hbar^2 (k-g)^2}{2m} - \epsilon \right) C_{k-g} + U_{-g} C_k + U_0 C_{k-g} = 0 \quad (j=1) \quad (667)$$

$$\left(\frac{\hbar^2 (k)^2}{2m} - \epsilon \right) C_k + U_0 C_k + U_g C_{k-g} = 0 \quad (j=0) \quad (668)$$

$$U_{-jg} C_k + U_{(1-j)g} C_{k-g} = 0 \quad (j \neq 0, j \neq 1) \quad (669)$$

Nelle equazioni per $(j \neq 0, j \neq 1)$ i termini:

$$\left(\frac{\hbar^2 (k-jg)^2}{2m} - \epsilon \right) C_{k-jg}$$

sono spariti perché abbiamo visto che i coefficienti C_{k-jg} sono tutti trascurabili. Per ridurre il sottosistema a una forma trattabile, assumiamo ancora che in tutte le equazioni per $(j \neq 0, j \neq 1)$ i coefficienti U dello sviluppo in serie del potenziale siano sufficientemente piccoli da rendere il primo membro ≈ 0 (nonostante, come abbiamo visto, C_k e C_{k-g} non siano piccoli). Questa approssimazione è permessa dal modello degli elettroni quasi liberi (cioè il potenziale $U(r)$ è molto piccolo).

Se accettiamo questa approssimazione, tutte le equazioni per $(j \neq 0, j \neq 1)$ diventano delle identità ($0=0$) e quindi spariscono. La dimensione del sistema si riduce a 2:

$$\begin{aligned} \left(\frac{\hbar^2 (k-g)^2}{2m} - \epsilon \right) C_{k-g} + U_0 C_{k-g} + U_{-g} C_k &= 0 \\ U_g C_{k-g} + U_0 C_k + \left(\frac{\hbar^2 (k)^2}{2m} - \epsilon \right) C_k &= 0 \end{aligned} \quad (670)$$

244. Per semplificare ulteriormente il problema facciamo altre due ipotesi connesse al potenziale periodico $U(r)$.

1. È sempre possibile introdurre nell'hamiltoniano una costante additiva senza cambiare la soluzione:

$$\left(-\frac{\hbar^2}{2m} \nabla^2 + U(\vec{r}) \right) \Psi(\vec{r}) = \epsilon \Psi(\vec{r}) \quad (671)$$

$$\left(-\frac{\hbar^2}{2m} \nabla^2 + U(\vec{r}) + C \right) \Psi(\vec{r}) = \epsilon \Psi(\vec{r}) \quad (672)$$

$$\left(-\frac{\hbar^2}{2m} \nabla^2 + U(\vec{r}) \right) \Psi(\vec{r}) + C \Psi(\vec{r}) = \epsilon \Psi(\vec{r}) \quad (673)$$

$$\left(-\frac{\hbar^2}{2m} \nabla^2 + U(\vec{r}) \right) \Psi(\vec{r}) = (\epsilon - C) \Psi(\vec{r}) \quad (674)$$

L'unico effetto e' lo shift dell'energia di tutti gli stati della stessa quantita': ma questo non e' mai un problema visto che l'energia degli stati e' definita, in generale, a meno di una costante.

Sulla base di cio', possiamo introdurre nell'hamiltoniano la costante uguale a $-U_0$, dove U_0 e' il coefficiente dello sviluppo in serie di Fourier del potenziale per $G = 0$ (cioe' per $i = 0$ nella sommatoria della 637).

Questo accorgimento ci consente di assumere che nello sviluppo in serie di Fourier del potenziale periodico il coefficiente U_0 sia nullo:

$$\begin{aligned}
& -\frac{\hbar^2}{2m}\nabla^2 - U_0 + U(\vec{r}) \\
&= -\frac{\hbar^2}{2m}\nabla^2 - U_0 + \sum_{\vec{G}} U_{\vec{G}} \exp(i\vec{G} \cdot \vec{r}) \\
&= -\frac{\hbar^2}{2m}\nabla^2 - U_0 + U_0 + \sum_{\vec{G} \neq 0} U_{\vec{G}} \exp(i\vec{G} \cdot \vec{r}) \\
&= -\frac{\hbar^2}{2m}\nabla^2 + \sum_{\vec{G} \neq 0} U_{\vec{G}} \exp(i\vec{G} \cdot \vec{r}) \tag{675}
\end{aligned}$$

Con cio' i termini contenenti U_0 spariscono dal sistema 670:

$$\begin{aligned}
& \left(\frac{\hbar^2 (k-g)^2}{2m} - \epsilon \right) C_{k-g} + U_{-g} C_k = 0 \\
& U_g C_{k-g} + \left(\frac{\hbar^2 (k)^2}{2m} - \epsilon \right) C_k = 0 \tag{676}
\end{aligned}$$

2. Assumiamo che il potenziale $U(r)$ sia una funzione reale e pari, cioe':

$$U(r) = U(-r) \tag{677}$$

(nel caso tridimensionale questo equivarrebbe a supporre che la struttura cristallina abbia un centro di simmetria e quindi: $U(\vec{r}) = U(-\vec{r})$)

Sotto questa ipotesi, per la proprieta' generale dei coefficienti di Fourier di una funzione reale e pari vista al punto 55, vale:

$$U_{ig} = U_{-ig} \tag{678}$$

(ovvero, nel caso tridimensionale, $U_{\vec{G}} = U_{-\vec{G}}$)

e quindi, in particolare:

$$U_g = U_{-g} \tag{679}$$

In definitiva, possiamo riscrivere il sistema come:

$$\begin{aligned} \left(\frac{\hbar^2 (k-g)^2}{2m} - \epsilon \right) C_{k-g} + U_g C_k &= 0 \\ U_g C_{k-g} + \left(\frac{\hbar^2 (k)^2}{2m} - \epsilon \right) C_k &= 0 \end{aligned} \quad (680)$$

245. Come vedete, quando k e' vicino al confine della prima zona di Brillouin, riusciamo a ridurre il sottosistema 637 a un sistema omogeneo di due sole equazioni!

Le incognite in questo sistema sono i due coefficienti C_k e C_{k-g} e l'energia ϵ ; quest'ultima e' approssimativamente data dal valore del gas di Fermi (ricordate che siamo nel modello degli elettroni quasi liberi), tuttavia ora vedremo che il sistema ottenuto ci consente di determinare l'energia con un livello di precisione maggiore.⁵⁰

Il sistema 680 ammette una soluzione non banale se e solo se il determinante dei coefficienti e' nullo, cioe':

$$\begin{vmatrix} \left(\frac{\hbar^2 (k-g)^2}{2m} - \epsilon \right) & U_g \\ U_g & \left(\frac{\hbar^2 (k)^2}{2m} - \epsilon \right) \end{vmatrix} = 0 \quad (681)$$

Per alleggerire, poniamo:

$$\lambda_{k-g} = \frac{\hbar^2 (k-g)^2}{2m} \quad (682)$$

$$\lambda_k = \frac{\hbar^2 (k)^2}{2m} \quad (683)$$

Si ha allora:

⁵⁰Se riguardate il procedimento del punto 239 in cui abbiamo ricavato l'energia e le autofunzioni per punti k lontani dal confine della prima zona di Brillouin, vi dovrete rendere conto che, in pratica, il punto chiave e' stato quello di assumere che la sommatoria:

$$\sum_{i=-\infty}^{\infty} U_{ig} C_{k-jg-ig}$$

sia trascurabile. In questo caso, invece, non trascuriamo *del tutto* la sommatoria perche' ne consideriamo i termini contenenti U_g e U_{-g} ; e' per questo che il livello di precisione e' maggiore.

$$\begin{aligned}
(\lambda_{k-g} - \epsilon)(\lambda_k - \epsilon) - U_g^2 &= 0 \\
\epsilon^2 - \epsilon(\lambda_k + \lambda_{k-g}) + \lambda_k \lambda_{k-g} - U_g^2 &= 0 \\
\epsilon &= \frac{\lambda_k + \lambda_{k-g} \pm \sqrt{(\lambda_k + \lambda_{k-g})^2 - 4(\lambda_k \lambda_{k-g} - U_g^2)}}{2} \\
\epsilon &= \frac{1}{2}(\lambda_k + \lambda_{k-g}) \pm \left\{ \frac{1}{4}(\lambda_k - \lambda_{k-g})^2 + U_g^2 \right\}^{1/2} \quad (684)
\end{aligned}$$

L'espressione appena ottenuta fornisce la dipendenza dell'energia degli stati elettronici dal vettore d'onda k in un piccolo intorno sinistro di $k = g/2$ (dove sono valide le ipotesi fatte per ottenerla).

Siccome λ_k e λ_{k-g} dipendono in modo quadratico da k , lo stesso vale per ϵ ; cioè: in un piccolo intorno sinistro di $k = g/2$ la dipendenza di ϵ da k e' di tipo quadratico.

246. Notate che per $U_g \rightarrow 0$ (modello "empty lattice") le due soluzioni dell'equazione di secondo grado diventano:

$$\epsilon \approx \frac{1}{2}(\lambda_k + \lambda_{k-g}) \pm \frac{1}{2}(\lambda_k - \lambda_{k-g}) = \begin{cases} \lambda_{k-g} = \frac{\hbar^2 (k-g)^2}{2m} \\ \lambda_k = \frac{\hbar^2 (k)^2}{2m} \end{cases} \quad (685)$$

che sono le espressioni per la bande n. 1 e 2 nel modello "empty lattice" (eq. 607 e 608).

247. Come vi avevo preannunciato, la degenerazione dell'energia per $k = g/2$ e' stata rimossa, perche' ora ci sono due valori distinti di ϵ . Quando k e' esattamente uguale a $g/2$, si ha:

$$\lambda_{\frac{g}{2}} = \lambda_{-\frac{g}{2}} \quad (686)$$

$$\epsilon = \lambda_{\frac{g}{2}} \pm |U_g| = \overbrace{\frac{\hbar^2 \left(\frac{\pi}{a}\right)^2}{2m}}^{\text{free electron}} \pm |U_g| \quad (687)$$

Il termine evidenziato e' il valore dell'energia per il gas di Fermi (per $k = g/2 = \pi/a$): quindi l'accensione del potenziale periodico splitta l'energia al confine della prima zona di Brillouin in due valori che stanno al disopra e al di sotto del valore per il gas di Fermi di una quantita' uguale al valore assoluto del primo coefficiente U_g dello sviluppo in serie di Fourier del potenziale. La differenza fra i due valori di energia trovati e' pari a $2|U_g|$:

$$\epsilon_+ - \epsilon_- = \lambda_{\frac{g}{2}} + |U_g| - (\lambda_{\frac{g}{2}} - |U_g|) = 2|U_g| \quad (688)$$

Ora la prima e la seconda banda non sono piu' contigue, ma separate da un'intervallo di energia $[\epsilon_-, \epsilon_+]$; un elettrone del cristallo *non puo' possedere* un'energia che appartenga a tale intervallo, che costituisce quindi quello che abbiamo chiamato band gap.

248. L'autovalore ϵ_- e' il massimo della banda n. 1, mentre l'autovalore ϵ_+ e' il minimo della banda n. 2; le due bande sono ora discoste.

La modificazione degli autovalori rispetto al valore per gli elettroni liberi rappresentata dalla 684 non vale solo quando k e' esattamente uguale a $g/2$, ma anche quando k e' sufficientemente vicino a $g/2$ da fare in modo che gli argomenti utilizzati per ottenerla siano ancora validi (fondamentalmente, se riguardate il procedimento, le ipotesi principali sono:

$$\begin{aligned} k &\approx \frac{g}{2} \\ \epsilon &\approx \frac{\hbar^2 \left(\frac{g}{2}\right)^2}{2m} = \frac{\hbar^2 \left(-\frac{g}{2}\right)^2}{2m} \end{aligned} \quad (689)$$

)
Quindi ci aspettiamo che i due rami di parabola che descrivono la prima e la seconda banda nel caso degli elettroni liberi si modifichino nelle immediate vicinanze del confine della prima zona di Brillouin "respingendosi" e che questo effetto sia massimo quando $k = g/2$.

Notate anche che i due rami di parabola sono perpendicolari alla retta verticale passante per il confine della prima zona di Brillouin. Detto in altri termini, la derivata dei due rami di parabola e' nulla per $k = g/2$. Questo si vede facilmente derivando rispetto a k la 684:

$$\begin{aligned} \frac{d\epsilon}{dk} &= \frac{d}{dk} \left(\frac{1}{2} (\lambda_k + \lambda_{k-g}) \pm \left\{ \frac{1}{4} (\lambda_k - \lambda_{k-g})^2 + U_g^2 \right\}^{1/2} \right) \\ &= \frac{1}{2} \left(\frac{d\lambda_k}{dk} + \frac{d\lambda_{k-g}}{dk} \right) \pm \frac{1}{2} \left\{ \frac{1}{4} (\lambda_k - \lambda_{k-g})^2 + U_g^2 \right\}^{-1/2} \\ &\times \left(\frac{1}{2} (\lambda_k - \lambda_{k-g}) \right) \left(\frac{d\lambda_k}{dk} - \frac{d\lambda_{k-g}}{dk} \right) \end{aligned} \quad (690)$$

Siccome:

$$k = \frac{g}{2} \implies \frac{d\lambda_k}{dk} + \frac{d\lambda_{k-g}}{dk} = \lambda_k - \lambda_{k-g} = 0 \quad (691)$$

ne segue che:

$$\left(\frac{d\epsilon}{dk} \right)_{k=g/2} = 0 \quad (692)$$

La situazione e' rappresentata nella figura 64, dove sono mostrate la prima e la seconda banda e l'apertura del band gap (le linee tratteggiate mostrano l'andamento dell'energia per gli elettroni liberi).

249. Potete notare che la figura 64 e' simmetrica, nel senso che le modifiche che hanno luogo in un intorno sinistro di $k = g/2$ si verificano anche in un intorno destro di $k = -g/2$.⁵¹

⁵¹Notate che questa **non** e' una conseguenza della periodicita' della funzione $\epsilon_n(k)$ nello spazio reciproco (punto 223); semmai, tale periodicita' replica l'intorno sinistro di $k = g/2$ nell'intorno *sinistro* di $k = -g/2$; torneremo su questo aspetto fra poco.

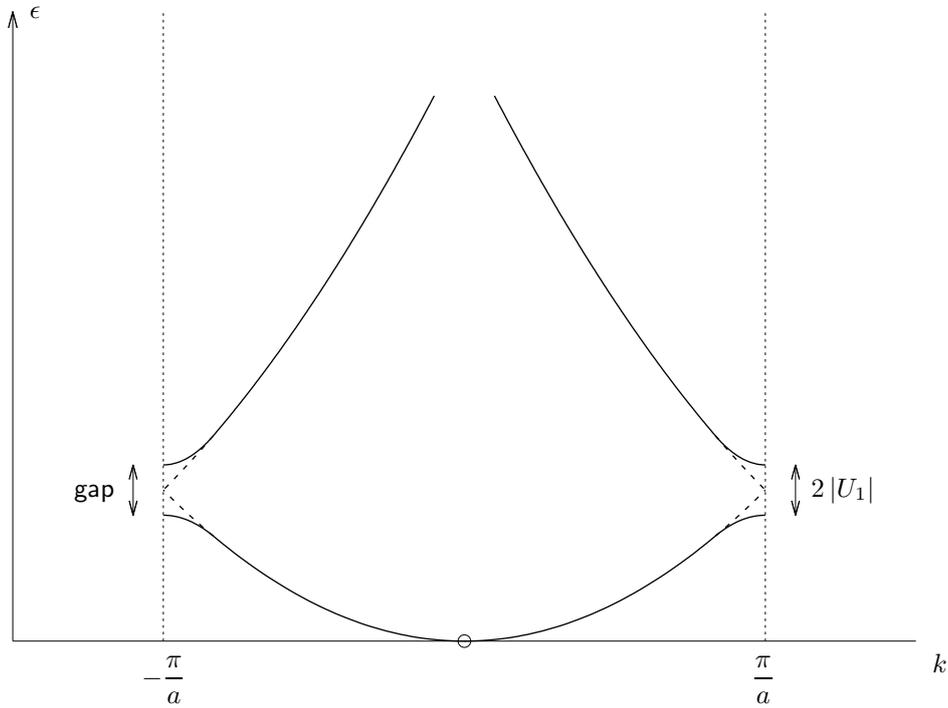


Figura 64:

A questo punto dovrebbe essere semplice rendersi conto che le cose devono stare effettivamente così'.

Se ripetete il ragionamento sviluppato dal punto 243 in poi per $k \in [-g/2, -g/2 + \delta]$ (cioè per k appartenente a un piccolo intorno destro di $k = -g/2$), troverete che in questo caso ci sono ancora due soli coefficienti dello sviluppo in serie dell'autofunzione che assumono un valore apprezzabile:

$$C_k \text{ e } C_{k+g} \quad (693)$$

(cioè: C_{k-jg} , per $j = 0$ e $j = -1$)

I termini della sommatoria nella 637 che sopravvivono sono ottenuti con:

$$\sum_{i=-\infty}^{\infty} U_{ig} C_{k-jg-ig} \quad (694)$$

$$C_{k-jg-ig} = C_k \iff i = -j \quad (695)$$

$$C_{k-jg-ig} = C_{k+g} \iff i = -1 - j \quad (696)$$

e quindi la sommatoria può essere approssimata a:

$$\sum_{i=-\infty}^{\infty} U_{ig} C_{k-jg-ig} \approx U_{-jg} C_k + U_{(-1-j)g} C_{k+g} \quad (697)$$

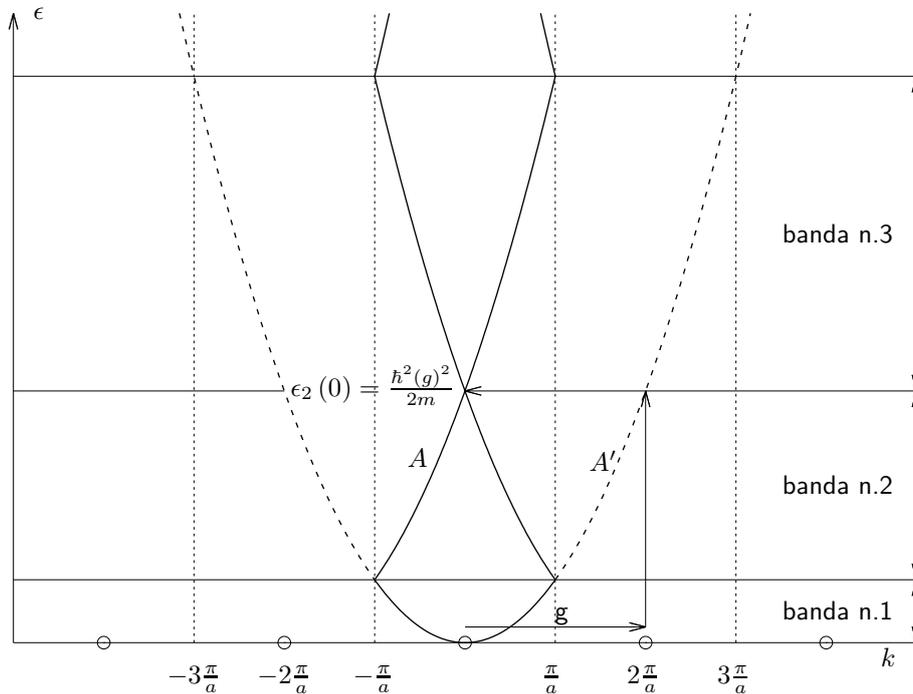


Figura 65: Il tratto A interno alla prima zona di Brillouin e' ottenuto traslando il tratto A' di $-g$. $\epsilon_2(0)$ e' ottenuto calcolando la parabola valida per gli elettroni liberi in $k = g$.

Cio' conduce ad un sistema identico al sistema 680, con $k + g$ al posto di $k - g$. Tutte le osservazioni fatte al punto 245 e successivi valgono invariate (a parte ovvi adattamenti): verificatelo.

250. Nei punti precedenti abbiamo discusso la formazione di un band gap fra la prima e la seconda banda di un cristallo (monodimensionale) costituito da elettroni quasi liberi. Questo fenomeno non e' limitato solo a queste due bande: l'accensione di un potenziale non nullo (anche se opportunamente piccolo) determina l'apertura di band gaps per tutte le bande del sistema.

Sulla base di quanto abbiamo gia' visto, non e' difficile spiegarsi la comparsa di un band gap fra la seconda e la terza banda.

A questo scopo consideriamo i punti k appartenenti a un piccolo intervallo centrato attorno a $k = 0$. Nel modello degli elettroni quasi liberi, per tali punti si ha:

$$\epsilon_2(k) = \frac{\hbar^2(g)^2}{2m} + \alpha \quad \text{con } |\alpha| \rightarrow 0 \quad (698)$$

Notate: fissato k , $\epsilon_2(k)$ e' il *secondo* autovalore, cioe' l'energia appartenente alla seconda banda. Il valore di $\epsilon_2(0)$ e' ricavato dalla parabola valida per gli elettroni liberi calcolata per $k = g$ (figura 65).

Notate anche che, analogamente al caso $k = g/2$ già analizzato (punto 243), anche per $k = 0$ le bande 2 e 3 hanno un punto in comune, cioè l'energia ha una doppia degenerazione rispetto all'indice di banda n :

$$\epsilon_2(0) = \epsilon_3(0) \quad (699)$$

Ripercorrendo i passaggi visti al punto 243 e successivi, potete verificare che due soli coefficienti hanno un valore apprezzabilmente diverso da 0:

$$C_{k-g} \text{ e } C_{k+g} \quad (700)$$

(cioè: C_{k-jg} per $j = 1$ e $j = -1$)

Allora, tenendo conto che:

$$C_{k-jg-ig} = C_{k-g} \iff i = 1 - j \quad (701)$$

$$C_{k-jg-ig} = C_{k+g} \iff i = -1 - j \quad (702)$$

la sommatoria della 637 può essere approssimata a:

$$\sum_{i=-\infty}^{\infty} U_{ig} C_{k-jg-ig} \approx U_{(1-j)g} C_{k-g} + U_{(-1-j)g} C_{k+g} \quad (703)$$

Con i passaggi già visti si arriva al seguente sistema:

$$\left(\frac{\hbar^2 (k-g)^2}{2m} - \epsilon \right) C_{k-g} + U_{-2g} C_{k+g} = 0 \quad (704)$$

$$U_{2g} C_{k-g} + \left(\frac{\hbar^2 (k+g)^2}{2m} - \epsilon \right) C_{k+g} = 0 \quad (705)$$

Da questo punto in poi l'algebra è identica a quella vista al punto 245. Il risultato finale è l'apertura di un band gap fra le bande 2 e 3. L'ampiezza di tale band gap è $2|U_{2g}|$, cioè due volte il valore assoluto del secondo coefficiente dello sviluppo in serie di Fourier del potenziale periodico presente nel cristallo. Nella figura 66 ho rappresentato le bande n. 1, 2 e 3 (di quest'ultima solo la parte inferiore) e i corrispondenti band gaps determinati dal potenziale periodico del cristallo.

251. È possibile scrivere una forma generale del sistema di due equazioni da cui si ottiene il band gap fra le bande n e $n+1$ nel modello monodimensionale di un cristallo contenente elettroni quasi liberi.

Ricapitoliamo ed estendiamo quanto visto.

Band gap bande 1-2 (punto 243)

⇒ Il band gap si apre in un intorno del punto: $k = g/2$ (e del punto: $k = -g/2$)

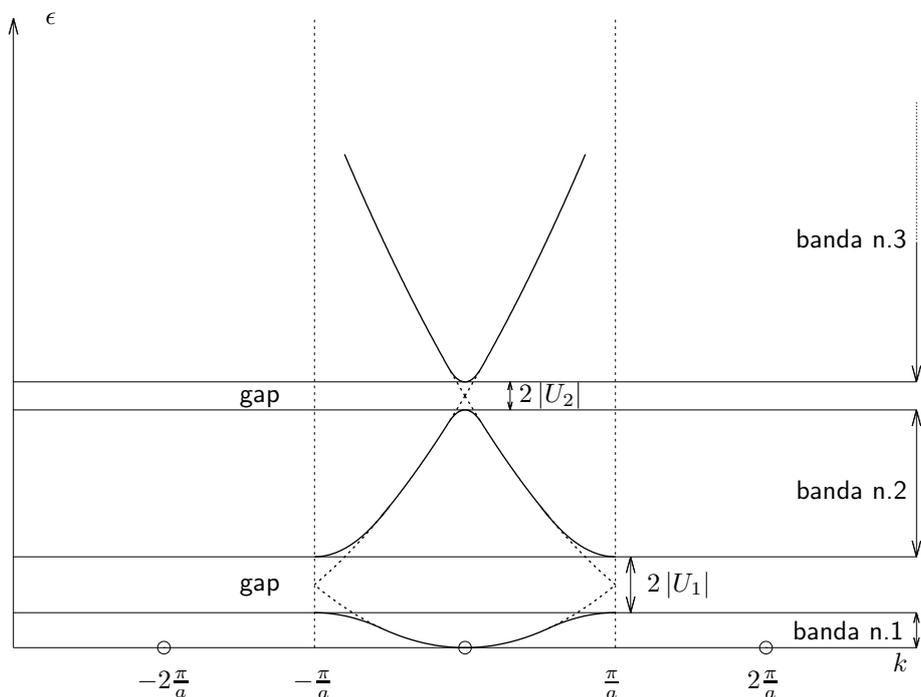


Figura 66: L'apertura di un band gap fra le bande n.1 e 2 e fra le bande n.2 e 3. Le linee tratteggiate mostrano l'andamento dell'energia per gli elettroni liberi (i.e. potenziale esattamente nullo e quindi bande contigue). La scala di questa figura e' diversa da quella della figura 65. Notate che la parte inferiore di questa figura, cioe' quella relativa al band gap fra le bande 1 e 2, e' identica (a parte la scala) alla figura 64.

⇒ L'energia degli elettroni liberi e':

$$\epsilon \approx \frac{\hbar^2 \left(\frac{g}{2}\right)^2}{2m}$$

⇒ I valori di j per cui:

$$\frac{\hbar^2 (k - jg)^2}{2m} \approx \epsilon$$

sono: $j = 0, 1$

⇒ Gli unici due coefficienti C_{k-jg} apprezzabilmente diversi da zero sono:

$$C_{k-g}, C_k$$

⇒ La sommatoria:

$$\sum_{i=-\infty}^{\infty} U_{ig} C_{k-jg-ig}$$

viene approssimata nel modo seguente.

$$C_{k-jg-ig} = C_{k-g} \iff i = 1-j$$

$$C_{k-jg-ig} = C_k \iff i = -j$$

$$\sum_{i=-\infty}^{\infty} U_{ig} C_{k-jg-ig} \approx U_{(1-j)g} C_{k-g} + U_{-jg} C_k$$

⇒ Il sistema che si ottiene e':

$$\left(\frac{\hbar^2 (k-g)^2}{2m} - \epsilon \right) C_{k-g} + U_{-g} C_k = 0$$

$$U_g C_{k-g} + \left(\frac{\hbar^2 (k)^2}{2m} - \epsilon \right) C_k = 0$$

Band gap bande 2-3 (punto 250)

⇒ Il band gap si apre in un intorno del punto: $k = 0$

⇒ L'energia degli elettroni liberi e':

$$\epsilon \approx \frac{\hbar^2 (g)^2}{2m}$$

⇒ I valori di j per cui:

$$\frac{\hbar^2 (k-jg)^2}{2m} \approx \epsilon$$

sono: $j = -1, 1$

⇒ Gli unici due coefficienti C_{k-jg} apprezzabilmente diversi da zero sono:

$$C_{k-g}, C_{k+g}$$

⇒ La sommatoria:

$$\sum_{i=-\infty}^{\infty} U_{ig} C_{k-jg-ig}$$

viene approssimata nel modo seguente.

$$C_{k-jg-ig} = C_{k-g} \iff i = 1-j$$

$$C_{k-jg-ig} = C_{k+g} \iff i = -1-j$$

$$\sum_{i=-\infty}^{\infty} U_{ig} C_{k-jg-ig} \approx U_{(1-j)g} C_{k-g} + U_{(-1-j)g} C_{k+g}$$

⇒ Il sistema che si ottiene e':

$$\begin{aligned} \left(\frac{\hbar^2 (k-g)^2}{2m} - \epsilon \right) C_{k-g} + U_{-2g} C_{k+g} &= 0 \\ U_{2g} C_{k-g} + \left(\frac{\hbar^2 (k+g)^2}{2m} - \epsilon \right) C_{k+g} &= 0 \end{aligned}$$

Band gap bande 3-4 (verificate da soli)

⇒ Il band gap si apre in un intorno del punto: $k = g/2$ (e del punto: $k = -g/2$)

⇒ L'energia degli elettroni liberi e':

$$\epsilon \approx \frac{\hbar^2 \left(\frac{3}{2}g \right)^2}{2m}$$

⇒ I valori di j per cui:

$$\frac{\hbar^2 (k-jg)^2}{2m} \approx \epsilon$$

sono: $j = -1, 2$

⇒ Gli unici due coefficienti C_{k-jg} apprezzabilmente diversi da zero sono:

$$C_{k-2g}, C_{k+g}$$

⇒ La sommatoria:

$$\sum_{i=-\infty}^{\infty} U_{ig} C_{k-jg-ig}$$

viene approssimata nel modo seguente.

$$\begin{aligned} C_{k-jg-ig} = C_{k-2g} &\iff i = 2-j \\ C_{k-jg-ig} = C_{k+g} &\iff i = -1-j \end{aligned}$$

$$\sum_{i=-\infty}^{\infty} U_{ig} C_{k-jg-ig} \approx U_{(2-j)g} C_{k-2g} + U_{(-1-j)g} C_{k+g}$$

⇒ Il sistema che si ottiene e':

$$\begin{aligned} \left(\frac{\hbar^2 (k-2g)^2}{2m} - \epsilon \right) C_{k-2g} + U_{-3g} C_{k+g} &= 0 \\ U_{3g} C_{k-2g} + \left(\frac{\hbar^2 (k+g)^2}{2m} - \epsilon \right) C_{k+g} &= 0 \end{aligned}$$

Band gap bande 4-5 (verificate da soli)

- ⇒ Il band gap si apre in un intorno del punto: $k = 0$
 ⇒ L'energia degli elettroni liberi e':

$$\epsilon \approx \frac{\hbar^2 (2g)^2}{2m}$$

- ⇒ I valori di j per cui:

$$\frac{\hbar^2 (k - jg)^2}{2m} \approx \epsilon$$

sono: $j = -2, 2$

- ⇒ Gli unici due coefficienti C_{k-jg} apprezzabilmente diversi da zero sono:

$$C_{k-2g}, C_{k+2g}$$

- ⇒ La sommatoria:

$$\sum_{i=-\infty}^{\infty} U_{ig} C_{k-jg-ig}$$

viene approssimata nel modo seguente.

$$\begin{aligned} C_{k-jg-ig} = C_{k-2g} &\iff i = 2 - j \\ C_{k-jg-ig} = C_{k+2g} &\iff i = -2 - j \end{aligned}$$

$$\sum_{i=-\infty}^{\infty} U_{ig} C_{k-jg-ig} \approx U_{(2-j)g} C_{k-2g} + U_{(-2-j)g} C_{k+2g}$$

- ⇒ Il sistema che si ottiene e':

$$\begin{aligned} \left(\frac{\hbar^2 (k - 2g)^2}{2m} - \epsilon \right) C_{k-2g} + U_{-4g} C_{k+2g} &= 0 \\ U_{4g} C_{k-2g} + \left(\frac{\hbar^2 (k + 2g)^2}{2m} - \epsilon \right) C_{k+2g} &= 0 \end{aligned}$$

A questo punto il pattern dovrebbe essere emerso.

Se n e' dispari (cioe' band gap fra le bande $n.1$ e $2, 3$ e 4 etc), allora, posto $s = (n + 1) / 2$ si ha:

$$\begin{aligned} \left(\frac{\hbar^2 (k - sg)^2}{2m} - \epsilon \right) C_{k-sg} + U_{-ng} C_{k+(s-1)g} &= 0 \\ U_{ng} C_{k-sg} + \left(\frac{\hbar^2 (k + (s-1)g)^2}{2m} - \epsilon \right) C_{k+(s-1)g} &= 0 \end{aligned}$$

Se n e' pari (cioe' band gap fra le bande $n.2$ e $3, 4$ e 5 etc), allora, posto $s = n/2$ si ha:

$$\begin{aligned} \left(\frac{\hbar^2 (k - sg)^2}{2m} - \epsilon \right) C_{k-sg} + U_{-ng} C_{k+sg} &= 0 \\ U_{ng} C_{k-sg} + \left(\frac{\hbar^2 (k + sg)^2}{2m} - \epsilon \right) C_{k+sg} &= 0 \end{aligned}$$

Ho riportato questo solo per farvi notare che l'ampiezza del band gap fra la banda n e la banda $n+1$ e' uguale a $2|U_n| = 2|U_{-n}|$. E siccome il valore assoluto dei coefficienti dello sviluppo in serie di Fourier del potenziale periodico diminuisce all'aumentare di n , se ne deduce che per questo modello il band gap diventa sempre piu' piccolo all'aumentare dell'indice di banda (cioe' dell'energia).

Per certi versi, questo e' cio' che ci si aspetterebbe per un qualsiasi sistema quantistico confinato: la spaziatura dei livelli energetici tende a zero all'aumentare dell'energia (pensate ad esempio ai livelli energetici atomici).

FIXME (start)

figure band gap repeated zone and extended zone

FIXME (end)

252. Una giustificazione fisica del band gap.

Possiamo dare un'interpretazione fisica del risultato che abbiamo ottenuto col modello monodimensionale degli elettroni quasi liberi.

Riprendiamo il sistema di due equazioni (sistema 680) valido in un piccolo intorno (sinistro) di $k = g/2$ che porta all'apertura del band gap fra le bande $n. 1$ e 2 . Quando k e' *esattamente* uguale a $g/2$ il sistema diventa:

$$(\lambda - \epsilon) C_{-\frac{g}{2}} + U_g C_{\frac{g}{2}} = 0 \quad (706)$$

$$U_g C_{-\frac{g}{2}} + (\lambda - \epsilon) C_{\frac{g}{2}} = 0 \quad (707)$$

con:

$$\lambda = \lambda_{\frac{g}{2}} = \lambda_{-\frac{g}{2}} = \frac{\hbar^2 (g/2)^2}{2m} = \frac{\hbar^2 (-g/2)^2}{2m} \quad (708)$$

Le due soluzioni per l'energia diventano (eq. 684 con $k = g/2$ e quindi $\lambda_k = \lambda_{k-g} = \lambda = \hbar^2 (g/2)^2 / (2m)$):

$$\epsilon = \lambda \pm |U_g| \quad (709)$$

Sostituendo nella prima (o nella seconda) equazione del sistema troviamo:

$$\frac{C_{\frac{g}{2}}}{C_{-\frac{g}{2}}} = \pm 1 \quad (710)$$

Questo significa che i due coefficienti $C_{\frac{g}{2}}$ e $C_{-\frac{g}{2}}$ hanno lo stesso valore assoluto, ma in corrispondenza a uno dei due valori di ϵ hanno lo stesso segno, mentre in corrispondenza all'altro hanno segno opposto.

Ricordate che $C_{\frac{g}{2}}$ e $C_{-\frac{g}{2}}$ sono gli unici due coefficienti "importanti" nello sviluppo in serie della funzione di stato. Quindi le due funzioni di stato corrispondenti ai due autovalori trovati avranno la forma (riguardate ad es. l'eq. 595):

$$\Psi_+(r) \propto \exp\left(i\frac{g}{2}r\right) + \exp\left(-i\frac{g}{2}r\right) = 2\cos\left(\frac{g}{2}r\right) = 2\cos\left(\frac{\pi}{a}r\right) \quad (711)$$

$$\Psi_-(r) \propto \exp\left(i\frac{g}{2}r\right) - \exp\left(-i\frac{g}{2}r\right) = 2i\sin\left(\frac{g}{2}r\right) = 2i\sin\left(\frac{\pi}{a}r\right) \quad (712)$$

(Ho usato il simbolo \propto e non $=$ perché non ho specificato il fattore che moltiplica le due espressioni, determinato dalla condizione di normalizzazione, che non ha alcuna importanza in questo contesto)

A questo punto osserviamo che le due funzioni $\Psi_{\pm}(r)$ sono onde *stazionarie*, cioè i punti dell'asse r dove si annullano *non* cambiano nel tempo. Ricordate che la soluzione completa dell'equazione di Schrödinger è ottenuta moltiplicando la parte spaziale per l'esponenziale:

$$\exp(-i\omega t), \quad \omega = \frac{\epsilon}{\hbar} \quad (713)$$

(eq. 310)

Se la parte spaziale *non* è di tipo esponenziale, come in questo caso, spazio e tempo "rimangono separati" e il comportamento della soluzione è quello di un'onda stazionaria (punto 68).

La densità di probabilità ρ_{\pm} (che consente di calcolare la probabilità di trovare un elettrone in un certo intervallo dell'asse r) per le due funzioni è proporzionale al rispettivo modulo quadrato, cioè:

$$\rho_+ \propto |\Psi_+(r)|^2 \propto \cos^2\left(\frac{\pi}{a}r\right) \quad (714)$$

$$\rho_- \propto |\Psi_-(r)|^2 \propto \sin^2\left(\frac{\pi}{a}r\right) \quad (715)$$

Osservate subito che ρ_{\pm} *non* è costante, diversamente da quanto avviene per la densità di probabilità per un elettrone libero (FE="free electron"):

$$\rho_{FE} \propto |\Psi_{FE}(r)|^2 \propto |\exp(ikr)|^2 = 1 \quad (716)$$

Osservate ancora che ρ_+ e ρ_- sono periodiche di periodo a , cioè hanno la stessa periodicità del cristallo monodimensionale che stiamo discutendo:

$$\begin{aligned} \cos^2\left(\frac{\pi}{a}(r+a)\right) &= \cos^2\left(\frac{\pi}{a}r + \pi\right) = \left(-\cos\left(\frac{\pi}{a}r\right)\right)^2 = \cos^2\left(\frac{\pi}{a}r\right) \\ \sin^2\left(\frac{\pi}{a}(r+a)\right) &= \sin^2\left(\frac{\pi}{a}r + \pi\right) = \left(-\sin\left(\frac{\pi}{a}r\right)\right)^2 = \sin^2\left(\frac{\pi}{a}r\right) \end{aligned}$$

Inoltre ρ_+ e ρ_- sono sfasate di $a/2$ (mezza lunghezza d'onda) e quindi una assume il suo valore massimo dove l'altra vale 0.

Allora guardate la figura 67.⁵² In essa e' rappresentato il cristallo monodimensionale con un potenziale periodico "ragionevole" (cioe' energia potenziale degli elettroni circa zero lontano dai nuclei e rapidamente decrescente ($\rightarrow -\infty$) nelle loro vicinanze). La figura mostra inoltre l'andamento di ρ_+ , ρ_- e ρ_{FE} .

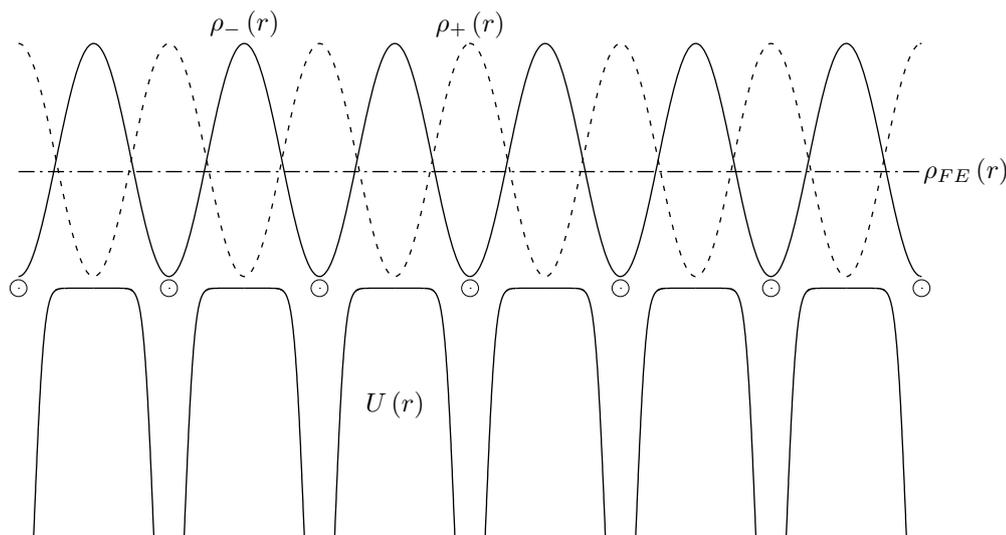


Figura 67: Andamenti del potenziale periodico $U(r)$, delle densita' di probabilita' ρ_{\pm} e della densita' di probabilita' ρ_{FE} per un cristallo monodimensionale; i cerchietti indicano la posizione dei nuclei.

ρ_{FE} e' costante e quindi la probabilita' di trovare un elettrone descritto da un'onda piana e' la stessa in ogni regione dell'asse r ; le due densita' di probabilita' ρ_{\pm} non sono costanti e tendono ad "accumulare" l'elettrone (cioe' danno valori di probabilita' maggiori) in regioni diverse dell'asse r : ρ_+ tende a localizzare l'elettrone con maggior probabilita' vicino ai nuclei; viceversa, ρ_- tende a localizzarlo con maggior probabilita' lontano da essi, cioe' nelle regioni internucleari.

Ma allora l'energia di un elettrone descritto da $\Psi_+(r)$ (che trascorre piu' tempo vicino ai nuclei) deve essere minore dell'energia di un elettrone descritto da $\Psi_-(r)$ (che trascorre piu' tempo lontano da essi).

Abbiamo svolto questa analisi per $k = g/2$, ma si puo' facilmente verificare che essa rimane valida per tutti i valori di k corrispondenti al confine delle zone di Brillouin:

$$k = j \frac{g}{2} \quad j = 0, \pm 1, \pm 2, \dots \quad (717)$$

In ultima analisi, dunque, l'apertura del band gap e' dovuta al fatto che elettroni aventi vettori d'onda corrispondenti ai confini delle zone di Brillouin sono

⁵²Trovate questa figura nel libro di Kittel²⁵ oppure nel libro di Ibach e Lüth.²⁶

descritti da funzioni di stato che sono onde stazionarie e che li possono localizzare mediamente piu' vicino o piu' lontano dai nuclei atomici; e questo fa si che la loro energia (*che sarebbe esattamente la stessa in assenza dei nuclei*) sia invece differente.

253. Band gap e diffrazione.

Abbiamo visto al punto precedente che quando $k = g/2$ ci sono due soluzioni possibili dell'equazione di Schrödinger, aventi entrambe la forma di onde stazionarie.

In questo punto voglio mostrarvi che l'origine fisica di cio' e' la diffrazione.

L'elettrone e' un'onda e come tale puo' subire diffrazione in un reticolo, esattamente come i raggi X (p. 79).

La condizione affinche' un elettrone venga diffratto dal reticolo degli ioni che costituisce il solido e' la stessa che abbiamo visto in generale (eq. 294):

$$\vec{k} \cdot \vec{n}_G = \frac{1}{2}G \quad (718)$$

In una dimensione, l'angolo fra \vec{k} ed \vec{n}_G puo' puo' essere solamente 0 oppure π e quindi il prodotto scalare e':

$$\vec{k} \cdot \vec{n}_G = \pm k \quad (719)$$

dove $k \geq 0$ indica il modulo del vettore \vec{k} .

In definitiva:

$$\pm k = \frac{1}{2}G \Rightarrow k = \pm \frac{1}{2}G = j \frac{g}{2} \quad j = 0, \pm 1, \pm 2, \dots \quad (720)$$

Quindi: quando il vettore d'onda dell'elettrone e' lontano dai valori $kg/2$, la sua interazione con il reticolo degli ioni e' trascurabile (nel modello degli elettroni quasi liberi) ed esso si comporta in pratica come un elettrone libero (punto 240); quando pero' il suo vettore d'onda e' uguale (o molto vicino) al valore che soddisfa la condizione per la diffrazione, allora esso viene diffratto.

In una dimensione, la diffrazione e' equivalente a una pura riflessione; cioe' quando su un nucleo incide un'onda (in questo caso l'elettrone), il nucleo si comporta come una sorgente che riemette la radiazione che l'ha colpito (con la stessa ampiezza, frequenza e fase) in *tutte le direzioni*: ma in un cristallo monodimensionale *tutte le direzioni* si riducono a due soltanto: la stessa direzione dell'onda incidente, oppure la direzione opposta.

La riemissione dell'onda incidente nella stessa direzione non e' una diffrazione: semplicemente l'onda procede indisturbata. Invece, la riemissione nella direzione opposta e' una diffrazione.

Siccome l'elettrone incidente e quello diffratto sono sempre "lo stesso elettrone", la funzione d'onda che ne descrive il comportamento deve essere la somma dell'onda incidente e di quella diffratta, per il principio di sovrapposizione (punto 93).

Un elettrone descritto dalla funzione d'onda:

$$\Psi(r) \propto \exp\left(i \frac{g}{2} r\right) \quad (721)$$

ha una lunghezza d'onda pari a $2a$ (ricordate che $g = 2\pi/a$), cioe' il doppio della costante di cella del reticolo diretto monodimensionale. Siccome le onde

diffratte sono originate dai nuclei, che hanno un periodo di ripetizione pari ad a , ne segue che verranno originate *due* onde diffratte, aventi una differenza di fase pari ad a (guardate la figura 68):

$$\Psi(r) \propto \exp\left(-i\frac{g}{2}r\right) \quad (722)$$

$$\Psi(r) \propto \exp\left(-i\frac{g}{2}(r+a)\right) \quad (723)$$

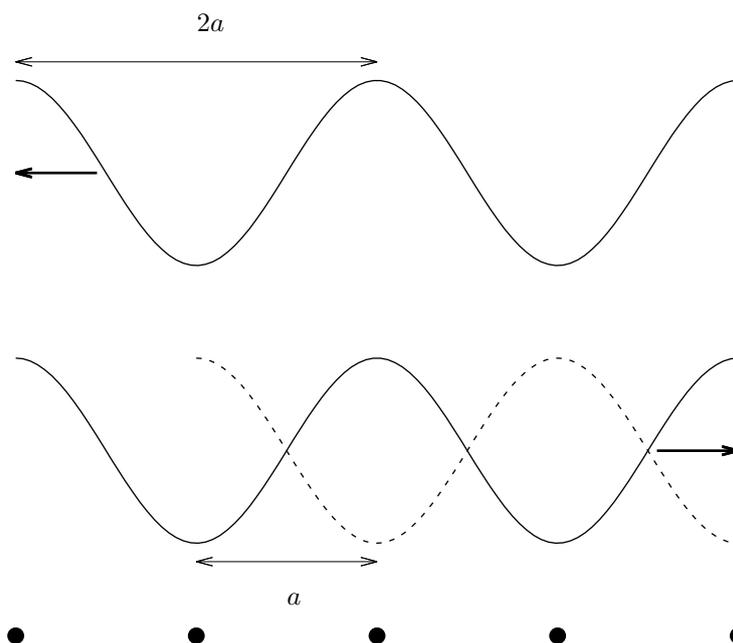


Figura 68: Un elettrone con vettore d'onda π/a si muove verso sinistra (in alto); siccome la condizione per la diffrazione e' soddisfatta, gli ioni (cerchietti pieni) generano onde (elettroniche) diffratte verso destra (in basso); la condizione per l'interferenza costruttiva (e cioe' che il cammino percorso da due onde diffratte deve essere uguale a un numero intero di lunghezze d'onda) viene soddisfatta separatamente dagli ioni di posto pari e da quelli di posto dispari; si generano cosi' due onde diffratte sfasate di a . Entrambe le onde diffratte sovrappongono con l'onda elettronica incidente dando due onde stazionarie distinte.

Queste sono onde progressive che si muovono in direzione *opposta* a quella dell'onda (cioe' l'elettrone) incidente.

La sovrapposizione dell'onda incidente con ciascuna delle due onde diffratte descrive uno di due possibili stati dell'elettrone prodotti dalla diffrazione.

Si avra':

$$\begin{aligned}
\Psi_+(r) &\propto \exp\left(i\frac{g}{2}r\right) + \exp\left(-i\frac{g}{2}r\right) = 2\cos\left(\frac{g}{2}r\right) = 2\cos\left(\frac{\pi}{a}r\right) \quad (724) \\
\Psi_-(r) &\propto \exp\left(i\frac{g}{2}r\right) + \exp\left(-i\frac{g}{2}(r+a)\right) \\
&= \cos\left(\frac{\pi}{a}r\right) + i\sin\left(\frac{\pi}{a}r\right) + \cos\left(\frac{\pi}{a}(r+a)\right) - i\sin\left(\frac{\pi}{a}(r+a)\right) \\
&= \cos\left(\frac{\pi}{a}r\right) + i\sin\left(\frac{\pi}{a}r\right) + \cos\left(\frac{\pi}{a}r + \pi\right) - i\sin\left(\frac{\pi}{a}r + \pi\right) \\
&= 2i\sin\left(\frac{\pi}{a}r\right) \quad (725)
\end{aligned}$$

Ma queste sono esattamente le due onde stazionarie che avevamo trovato al punto 252 (eq. 711 e 712) e quindi da questo punto in poi si applica lo stesso ragionamento e si raggiunge la stessa conclusione.

254. A questo punto possiamo riassumere quanto trovato per il modello “nearly free electron”.

In presenza di un potenziale periodico piccolo, ma non trascurabile (come era nel modello “empty lattice”), il comportamento degli elettroni e' descritto con buona approssimazione quasi ovunque nello spazio \vec{k} dagli autovalori e autofunzioni degli elettroni liberi (punto 239 e seguenti).

Tuttavia, quando il vettore d'onda di un elettrone e' uguale o molto vicino al valore che soddisfa la condizione per la sua diffrazione da parte del reticolo su cui si basa la struttura cristallina del solido (e che e' responsabile della periodicit  del potenziale che compare nell'equazione di Schr dinger), la degenerazione degli autovalori (che si avrebbe in assenza del potenziale) viene rimossa e si ha la formazione di un intervallo di energia proibito (band gap) che separa due bande consecutive (punto 243 e seguenti).

255. Abbiamo analizzato il meccanismo di formazione del band gap in una sola dimensione, perche' cio' rende tutta la matematica connessa estremamente piu' semplice.

L'estensione a due e tre dimensioni non comporta nulla di nuovo da un punto di vista concettuale; e' solo piu' complicata.

Anche in piu' di una dimensione si aprono dei band gaps in corrispondenza ai vettori d'onda che soddisfano (perfettamente, o quasi perfettamente) la condizione per la diffrazione:

$$\vec{k} \cdot \vec{n}_G = \frac{1}{2}G$$

Nel caso piu' generale in cui lo spazio \vec{k} e' uno spazio tridimensionale, tale condizione e' verificata da tutti i vettori d'onda che si appoggiano sul piano perpendicolare al vettore reciproco \vec{G} e che lo divide in due segmenti di uguale lunghezza (figura 38). Tale piano costituisce proprio la faccia di una zona di Brillouin 68.

Quindi anche nel caso 2D o 3D i band gaps si aprono in corrispondenza ai confini delle zone di Brillouin.

La rappresentazione della struttura a bande quando la dimensione del sistema e' maggiore di 1 e' ottenuta in modo identico a quanto abbiamo gia' visto per il modello “empty lattice” (punti 232 e 233).

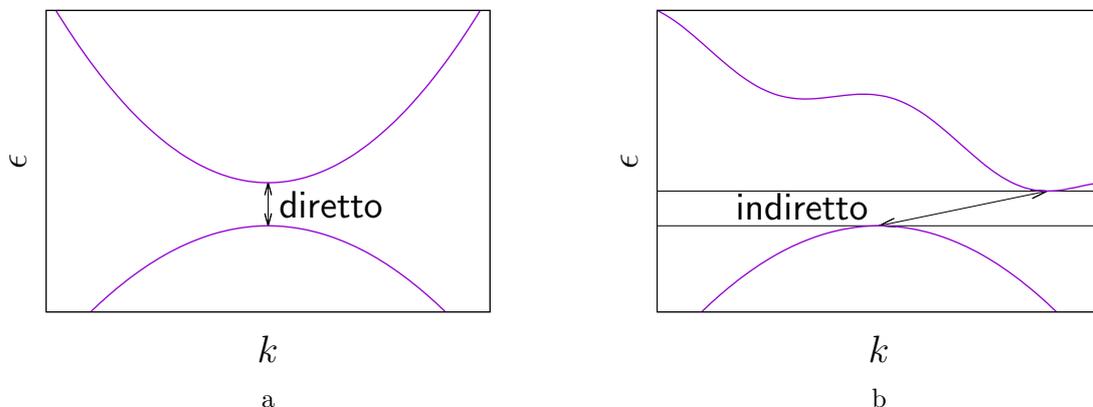


Figura 69:

256. Un modo è quello di rappresentare un certo numero di bande per i vettori d'onda corrispondenti ai punti di un segmento interno (o di una traiettoria costituita da segmenti interni) alla prima zona di Brillouin. Ovviamente, nel caso del modello "nearly free electron" (e nei casi reali) le bande non sono in generale contigue, ma compaiono dei band gaps.

Il band gap fra due bande consecutive può essere *diretto* oppure *indiretto*.

Se l'autovalore di minima energia della banda "superiore" e quello di massima energia della banda "inferiore" hanno lo stesso vettore d'onda, allora si parla di band gap diretto (figura 69a). Se invece i vettori d'onda corrispondenti ai due autovalori sono diversi, allora si parla di band gap indiretto (figura 69b).

Come già visto per il modello "empty lattice", anche nel caso del modello "nearly free electron" si può avere sovrapposizione di bande. Se due bande sono sovrapposte non può esserci un band gap fra esse.

257. Anche la superficie di Fermi si complica nel caso 2 o 3D del modello "nearly free electron". Purtroppo non ci sono calcoli semplici (anche solo approssimati) a cui si possa ricorrere per ricavare un confronto fra la superficie di Fermi in assenza e in presenza di un band gap per un sistema a dimensionalità > 1 . Quello che possiamo fare è una previsione qualitativa di come la superficie di Fermi viene modificata in presenza di un non trascurabile (anche se piccolo) potenziale periodico.

Riprendiamo la figura 60 che si applica ad un sistema 2D. Se l'energia di Fermi è tale per cui la superficie di Fermi è completamente contenuta nella prima zona di Brillouin (come in figura 60a), allora non c'è differenza fra il modello "empty lattice" e quello "nearly free electron". Se però la superficie di Fermi interseca il confine della prima zona di Brillouin, allora si creano delle discontinuità in corrispondenza alle intersezioni, dove il vettore d'onda soddisfa la condizione per la diffrazione (punto 253).

Possiamo raffigurare in modo qualitativo queste discontinuità tenendo presente quanto detto al punto 248, e cioè che le bande si "deformano" in corrispondenza ai confini delle zone di Brillouin intersecandoli perpendicolarmente.

La figura 60b mostra la rappresentazione "reduced zone" della superficie di Fermi per il sistema bidimensionale "empty lattice" (punto 229) nel caso in cui

tale superficie si estenda sia sulla prima che sulla seconda banda. Lo stesso scenario per il modello “nearly free electron” e’ mostrato nella figura 70 (dove ho riprodotto anche le corrispondenti figure per il modello “empty lattice”, per maggiore comodita’). Notate come si sono formate delle discontinuita’ in corrispondenza al confine fra la prima e la seconda zona di Brillouin; notate anche come il profilo della superficie incide perpendicolarmente sul confine fra le due zone. Quando i rami della seconda zona vengono riportati entro la prima (rappresentazione “reduced zone”), si ottiene una figura molto simile alla 60b, salvo per il fatto che il perimetro della superficie di Fermi presenta ora delle interruzioni in corrispondenza ai gaps generati dal potenziale periodico.

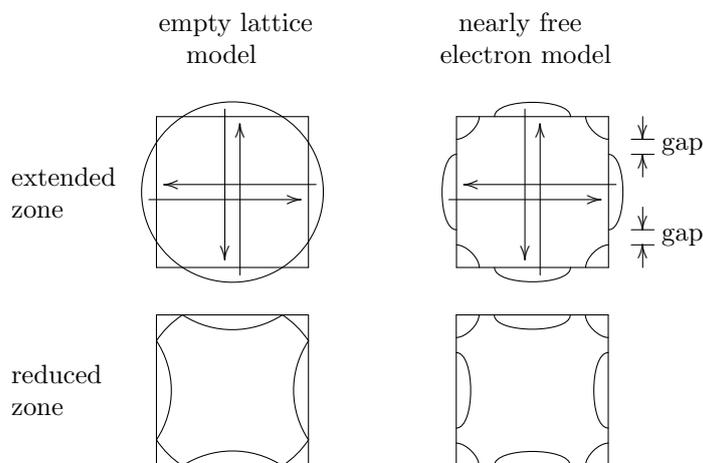


Figura 70: Confronto della superficie di Fermi per i modelli bidimensionali “empty lattice” e “nearly free electron” nel caso in cui la superficie sia distribuita sia sulla prima che sulla seconda zona di Brillouin. Le frecce indicano le parti “esterne” della superficie di Fermi che vengono riportate nella prima zona di Brillouin per ottenere la rappresentazione “reduced zone” da quella “extended zone”; ho anche indicato due dei band gaps che si aprono nel perimetro della superficie di fermi a causa del potenziale periodico.

In generale, l’effetto del potenziale periodico apprezzabile (“nearly free electron model”) e’ quello di “smussare” tutte le cuspidi che sono presenti quando esso e’ trascurabile (“empty lattice model”). Questo si vede bene nel confronto “empty lattice” “nearly free electron” con la rappresentazione “repeated zone”, mostrato nella figura 71.

Ovviamente, se il potenziale periodico non e’ piccolo (cioe’ non si puo’ applicare il modello “nearly free electron”), le cose si complicano ulteriormente e si deve necessariamente ricorrere al calcolo.

Il modello “tight binding”^{53 54}

⁵³Per questa parte ho seguito la derivazione di Ibach e Lüth.²⁷ Una derivazione simile e’ presentata anche da Kittel,²⁸ mentre Ashcroft e Mermin presentano un trattamento piu’ generale, ma anche piu’ complesso.²⁹

⁵⁴Anche in questo caso usero’ l’inglese. “Tight” vuol dire stretto, ben fissato: la traduzione sarebbe “modello degli elettroni strettamente (o saldamente) legati”.

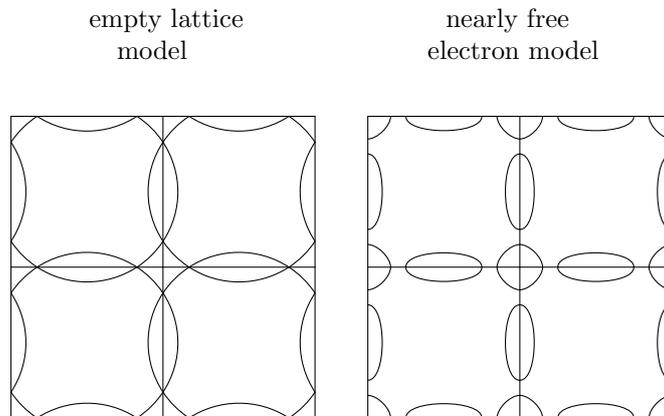


Figura 71: Confronto della superficie di Fermi per i modelli bidimensionali “empty lattice” e “nearly free electron” nella rappresentazione “repeated zone” (ho ripetuto quattro celle). Oltre alla presenza di gaps nella superficie di Fermi per il modello “nearly free electron”, dovrebbe essere evidente l’effetto di “smussamento” menzionato nel testo.

258. Nella parte precedente abbiamo visto due dei piu’ semplici modelli per l’analisi (non numerica) della struttura elettronica dei solidi (cristallini): il modello degli elettroni liberi e quello degli elettroni “quasi” liberi (“nearly free electron”). In entrambi i casi l’elemento fondamentale del modello e’ che in un solido gli elettroni (di valenza) degli atomi che lo costituiscono non risentano (“free electron”) o risentano molto poco (“nearly free electron”) del potenziale generato dagli ioni fissi nelle posizioni della struttura solida (cristallina).

Come suggerisce il suo nome, il modello “tight binding” affronta il problema partendo dall’estremo opposto, e cioe’ considerando gli elettroni di un solido fortemente legati agli atomi costituenti.

Sia i modelli gia’ visti che il modello “tight binding” che vedremo ora sono semplicistici e lo studio della struttura elettronica dei solidi reali viene condotto interamente per via numerica, utilizzando tecniche di calcolo molto sofisticate. Tuttavia, (come dovrete gia’ esservi resi conto) i modelli semplici hanno il pregio di consentire un trattamento analitico che aiuta a comprendere *la fisica* di questi sistemi.

259. Immaginate di espandere in modo uniforme un cristallo di sodio (reticolo cubico a corpo centrato) fino a far diventare la costante di cella dell’ordine di 1 cm. E’ chiaro che in tali condizioni ciascun atomo di sodio (con i suoi elettroni) ignora totalmente la presenza degli altri: lo stato degli elettroni e’ descritto in modo esatto dalle funzioni di stato dell’atomo isolato.

Se ora avviciniamo gli atomi di questo “cristallo” riducendo progressivamente la costante di cella, si dovra’ necessariamente arrivare al punto in cui gli elettroni su ciascun atomo cominciano a sentire la presenza del resto del cristallo. Questo punto dipende dallo stato degli elettroni e in particolare da quanto la funzione di stato atomica di ciascun elettrone si estende verso l’esterno dal nucleo. Possiamo aspettarci che l’ordine in cui gli elettroni risentono dell’influenza del resto del cristallo intorno a se’ sia $3s$ $2p$ $2s$ $1s$, perche’ questo e’ l’ordine con cui tali

funzioni di stato si estendono allontanandosi dal nucleo.

Nella figura 72 sono mostrate le funzioni $3s$ $2p$ $2s$ $1s$ centrate su due atomi di sodio a una distanza pari a quella del sodio cristallino. E' evidente che l'overlap delle funzioni $1s$ $2s$ e $2p$ e' trascurabile, mentre quello delle funzioni $3s$ (che descrivono l'elettrone di valenza del sodio) e' apprezzabile.

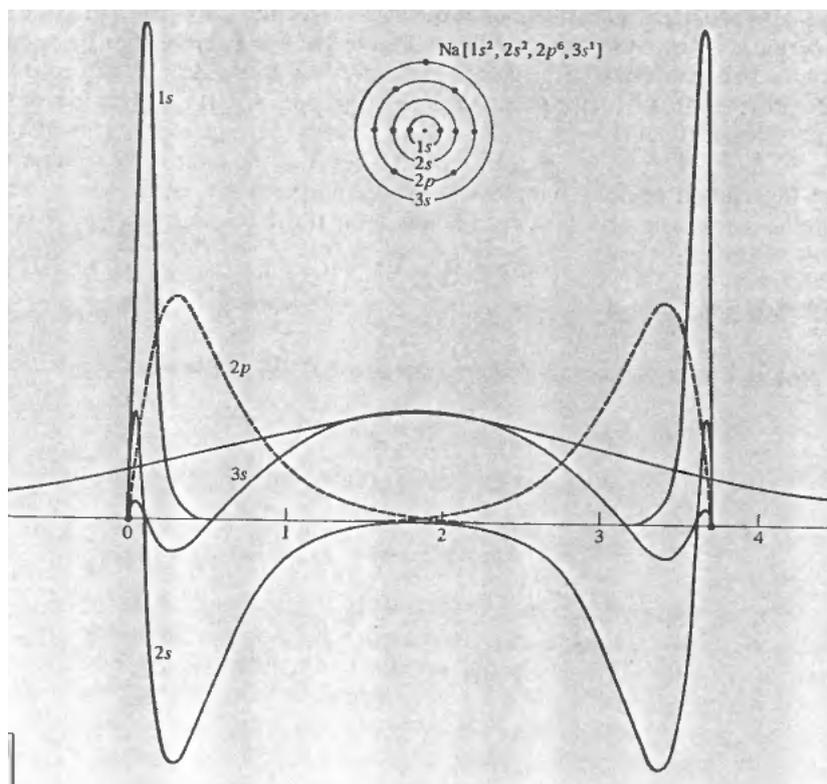


Figura 72:

Il modello "tight binding" assume che l'overlap delle funzioni di stato elettroniche centrate sugli atomi che costituiscono il solido sia sufficientemente grande da richiedere una loro correzione rispetto agli atomi isolati, ma sufficientemente piccolo da poter considerare le funzioni degli atomi isolati come una buona approssimazione.

260. L'equazione d'onda monoelettronica indipendente dal tempo per un atomo isolato nel sito \vec{r}_n del cristallo e'⁵⁵:

$$\left(-\frac{\hbar^2}{2m} \nabla^2 + V_{at}(\vec{r} - \vec{r}_n) \right) \varphi_j(\vec{r} - \vec{r}_n) = E_j \varphi_j(\vec{r} - \vec{r}_n) \quad (726)$$

dove V_{at} e' il potenziale dell'atomo isolato ed E_j indica il suo j -esimo autovalore. Per sviluppare il modello "tight binding" assumiamo che le autofunzioni φ_j dell'atomo isolato siano note e molto localizzate, cioe' l'overlap fra due funzioni

⁵⁵ $V_{at}(\vec{r} - \vec{r}_n)$ e $\varphi_j(\vec{r} - \vec{r}_n)$ sono, rispettivamente, il potenziale dell'atomo isolato e l'autofunzione corrispondente all'autovalore E_j traslati dall'origine al punto \vec{r}_n

centrate su atomi in posizione relativa di primi vicini sia piccolo. Supponiamo inoltre che, in vicinanza di ciascun atomo (la cui posizione \vec{r}_n coincide con un punto del reticolo di Bravais⁵⁶) l'hamiltoniano completo H del cristallo possa essere approssimato dall'operatore hamiltoniano H_{at} dell'atomo isolato piu' una correzione $v(\vec{r} - \vec{r}_n)$ che decade rapidamente allontanandosi dall'atomo e che per semplicita' supporremo possieda simmetria sferica:

$$H = H_{at} + v = -\frac{\hbar^2}{2m}\nabla^2 + V_{at}(\vec{r} - \vec{r}_n) + v(\vec{r} - \vec{r}_n)$$

Il termine $v(\vec{r} - \vec{r}_n)$ tiene conto del fatto che gli atomi nel cristallo non sono *esattamente* isolati e che il potenziale totale del cristallo deve essere periodico. Il nostro intento e' quello di trovare un'espressione per l'energia degli stati elettronici soluzioni dell'equazione di Schrödinger:

$$H\Psi_{\vec{k}}(\vec{r}) = \epsilon(\vec{k})\Psi_{\vec{k}}(\vec{r}) \quad (727)$$

dove $\Psi_{\vec{k}}(\vec{r})$ e' una funzione avente la forma prescritta dal teorema di Bloch (punto 187 e successivi) e quindi dipende dal vettore d'onda \vec{k} , i cui valori permessi sono determinati dalle condizioni al contorno periodiche di Born e von Karman (eq. 540).

Moltiplicando ambo i membri della 727 per $\Psi_{\vec{k}}^*(\vec{r})$ e integrando su tutto lo spazio si ottiene:

$$\epsilon(\vec{k}) = \frac{\langle \Psi_{\vec{k}} | H | \Psi_{\vec{k}} \rangle}{\langle \Psi_{\vec{k}} | \Psi_{\vec{k}} \rangle} \quad (728)$$

dove ho usato la notazione "bra-ket" per indicare gli integrali, cioe', ad esempio:

$$\langle \Psi_{\vec{k}} | H | \Psi_{\vec{k}} \rangle = \int \Psi_{\vec{k}}^*(\vec{r}) H \Psi_{\vec{k}}(\vec{r}) d\vec{r}$$

(e' sottinteso che l'integrale e' esteso a tutto lo spazio)

L'espressione 728 per $\epsilon(\vec{k})$ e' del tutto generale; il modello "tight binding" viene sviluppato assumendo che l'autofunzione $\Psi_{\vec{k}}$ possa essere espressa come una opportuna combinazione lineare delle autofunzioni atomiche corrispondenti a un certo autovalore atomico E_j di tutti gli atomi presenti nel cristallo:

$$\Psi_{\vec{k}}(\vec{r}) = \sum_n a_n \varphi_j(\vec{r} - \vec{r}_n) \quad (729)$$

L'idea e' che le autofunzioni dell'atomo isolato siano una buona approssimazione alla funzione che descrive lo stato di un elettrone con energia vicina a E_j nell'intero cristallo. In pratica, l'eq. 729 replica l'autofunzione atomica φ_j su ogni sito reticolare e definisce questa "estensione" dell'autofunzione atomica come la funzione che descrive lo stato di un elettrone in tutto il cristallo; in termini equivalenti, possiamo dire che quando l'elettrone si trova in prossimita' del nucleo situato a \vec{r}_n , viene descritto dall'autofunzione φ_j centrata su tale nucleo.

⁵⁶Cioe': assumiamo per semplicita' che il cristallo sia costituito da un solo tipo di atomi e che la base associata a ciascun punto reticolare sia costituita da un solo atomo.

261. Come abbiamo detto, la funzione $\Psi_{\vec{k}}$ deve avere la forma prescritta dal teorema di Bloch. Cio' si puo' ottenere se i coefficienti della combinazione lineare 729 vengono espressi come:

$$a_n = \exp\left(i\vec{k} \cdot \vec{r}_n\right) \quad (730)$$

E' semplice verificare che se gli a_n hanno la forma 730 allora la $\Psi_{\vec{k}}$ soddisfa il teorema di Bloch.

Per un generico vettore \vec{R} del reticolo diretto si ha infatti:

$$\begin{aligned} \Psi_{\vec{k}}(\vec{r} + \vec{R}) &= \sum_n \exp\left(i\vec{k} \cdot \vec{r}_n\right) \varphi_j(\vec{r} - \vec{r}_n + \vec{R}) \\ &= \sum_n \exp\left(i\vec{k} \cdot (\vec{r}_n - \vec{R} + \vec{R})\right) \varphi_j(\vec{r} - \vec{r}_n + \vec{R}) \\ &= \sum_n \exp\left(i\vec{k} \cdot (\vec{r}_n - \vec{R})\right) \exp\left(i\vec{k} \cdot \vec{R}\right) \varphi_j(\vec{r} - (\vec{r}_n - \vec{R})) \\ &= \exp\left(i\vec{k} \cdot \vec{R}\right) \sum_n \exp\left(i\vec{k} \cdot (\vec{r}_n - \vec{R})\right) \varphi_j(\vec{r} - (\vec{r}_n - \vec{R})) \end{aligned}$$

Sia \vec{r}_n che \vec{R} sono vettori del reticolo di Bravais e quindi anche $\vec{r}'_n = \vec{r}_n - \vec{R}$ e' un vettore del reticolo di Bravais⁵⁷; e allora:

$$\begin{aligned} &= \exp\left(i\vec{k} \cdot \vec{R}\right) \sum_n \exp\left(i\vec{k} \cdot \vec{r}'_n\right) \varphi_j(\vec{r} - \vec{r}'_n) \\ &= \exp\left(i\vec{k} \cdot \vec{R}\right) \Psi_{\vec{k}}(\vec{r}) \end{aligned}$$

che e' la condizione 493 per essere della forma prescritta dal teorema di Bloch.

262. Ora sostituiamo la 729 con gli a_n dati dalla 730 nella 728:

$$\epsilon(\vec{k}) = \frac{\int \left(\sum_m \exp\left(i\vec{k} \cdot \vec{r}_m\right) \varphi_j(\vec{r} - \vec{r}_m)\right)^* H \sum_n \exp\left(i\vec{k} \cdot \vec{r}_n\right) \varphi_j(\vec{r} - \vec{r}_n) d\vec{r}}{\int \left(\sum_m \exp\left(i\vec{k} \cdot \vec{r}_m\right) \varphi_j(\vec{r} - \vec{r}_m)\right)^* \sum_n \exp\left(i\vec{k} \cdot \vec{r}_n\right) \varphi_j(\vec{r} - \vec{r}_n) d\vec{r}} \quad (731)$$

263. Consideriamo prima il denominatore della 731, che puo' essere riscritto come:

$$\begin{aligned} &\sum_{m,n} \int \left(\exp\left(i\vec{k} \cdot \vec{r}_m\right) \varphi_j(\vec{r} - \vec{r}_m)\right)^* \exp\left(i\vec{k} \cdot \vec{r}_n\right) \varphi_j(\vec{r} - \vec{r}_n) d\vec{r} \\ &= \sum_{m,n} \int \exp\left(i\vec{k} \cdot (\vec{r}_n - \vec{r}_m)\right) \varphi_j^*(\vec{r} - \vec{r}_m) \varphi_j(\vec{r} - \vec{r}_n) d\vec{r} \end{aligned}$$

⁵⁷La somma 729 e' estesa a tutti i siti del cristallo; se il generico vettore \vec{R} e' tale per cui $\vec{r}'_n = \vec{r}_n - \vec{R}$ "esce" dal cristallo, e' sottinteso che la funzione $\varphi_j(\vec{r} - \vec{r}'_n)$ e' l'immagine periodica tramite le condizioni di Born e van Karman della corrispondente funzione all'interno del cristallo.

Grazie all'ipotesi che le funzioni atomiche siano localizzate (punto 260) possiamo assumere (in prima approssimazione) che solo gli integrali con $n = m$ siano apprezzabilmente diversi da zero; quindi si ha (il termine esponenziale e' uguale a 1 quando $n = m$):

$$= \sum_n \int \varphi_j^* (\vec{r} - \vec{r}_n) \varphi_j (\vec{r} - \vec{r}_n) d\vec{r} = \sum_n 1 = N \quad (732)$$

dove N e' il numero di atomi presenti nel cristallo. La penultima uguaglianza deriva dal fatto che, senza perdere in generalita', si puo' assumere che le funzioni atomiche φ_j siano normalizzate e quindi l'integrale e' uguale a 1.

264. Ora consideriamo il numeratore della 731.

Per l'applicazione dell'hamiltoniano H sulla somma alla sua destra si ha:

$$\begin{aligned} & H \sum_n \exp \left(i\vec{k} \cdot \vec{r}_n \right) \varphi_j (\vec{r} - \vec{r}_n) \\ &= (H_{at} + v) \sum_n \exp \left(i\vec{k} \cdot \vec{r}_n \right) \varphi_j (\vec{r} - \vec{r}_n) \end{aligned}$$

L'hamiltoniano e' lineare e quindi possiamo scambiarlo di posto con la somma; inoltre, il termine esponenziale non dipende da \vec{r} e quindi puo' essere spostato a sinistra dell'hamiltoniano; in definitiva:

$$\begin{aligned} &= \sum_n \exp \left(i\vec{k} \cdot \vec{r}_n \right) (H_{at} + v (\vec{r} - \vec{r}_n)) \varphi_j (\vec{r} - \vec{r}_n) \\ &= \sum_n \exp \left(i\vec{k} \cdot \vec{r}_n \right) \left(\underbrace{H_{at} \varphi_j (\vec{r} - \vec{r}_n)}_{= E_j \varphi_j (\vec{r} - \vec{r}_n)} + v (\vec{r} - \vec{r}_n) \varphi_j (\vec{r} - \vec{r}_n) \right) \end{aligned}$$

Ora sfruttiamo il fatto che $\varphi_j (\vec{r} - \vec{r}_n)$ e' autofunzione con autovalore E_j dell'hamiltoniano H_{at} :

$$\begin{aligned} &= \sum_n \exp \left(i\vec{k} \cdot \vec{r}_n \right) (E_j \varphi_j (\vec{r} - \vec{r}_n) + v (\vec{r} - \vec{r}_n) \varphi_j (\vec{r} - \vec{r}_n)) \\ &= \sum_n \exp \left(i\vec{k} \cdot \vec{r}_n \right) (E_j + v (\vec{r} - \vec{r}_n)) \varphi_j (\vec{r} - \vec{r}_n) \end{aligned}$$

Ora possiamo sostituire l'espressione ottenuta nel numeratore della 731:

$$\begin{aligned} & \int \left(\sum_m \exp \left(i\vec{k} \cdot \vec{r}_m \right) \varphi_j (\vec{r} - \vec{r}_m) \right)^* \sum_n \exp \left(i\vec{k} \cdot \vec{r}_n \right) (E_j + v (\vec{r} - \vec{r}_n)) \varphi_j (\vec{r} - \vec{r}_n) d\vec{r} \\ &= \int \sum_{m,n} \exp \left(-i\vec{k} \cdot \vec{r}_m \right) \varphi_j^* (\vec{r} - \vec{r}_m) \exp \left(i\vec{k} \cdot \vec{r}_n \right) (E_j + v (\vec{r} - \vec{r}_n)) \varphi_j (\vec{r} - \vec{r}_n) d\vec{r} \\ &= \sum_{m,n} \exp \left(i\vec{k} \cdot (\vec{r}_n - \vec{r}_m) \right) \int \varphi_j^* (\vec{r} - \vec{r}_m) (E_j + v (\vec{r} - \vec{r}_n)) \varphi_j (\vec{r} - \vec{r}_n) d\vec{r} \end{aligned}$$

Nell'ultimo passaggio si sfrutta il fatto che $\exp\left(i\vec{k} \cdot (\vec{r}_n - \vec{r}_m)\right)$ non dipende da \vec{r} e quindi si puo' portare fuori dall'integrale.

Procediamo separando l'integrale nei due addendi che contiene e facendo la stessa cosa sulla risultante sommatoria:

$$\begin{aligned}
&= \sum_{m,n} \exp\left(i\vec{k} \cdot (\vec{r}_n - \vec{r}_m)\right) \left(E_j \int \varphi_j^*(\vec{r} - \vec{r}_m) \varphi_j(\vec{r} - \vec{r}_n) d\vec{r} \right. \\
&\quad \left. + \int \varphi_j^*(\vec{r} - \vec{r}_m) v(\vec{r} - \vec{r}_n) \varphi_j(\vec{r} - \vec{r}_n) d\vec{r} \right) \\
&= \sum_{m,n} \exp\left(i\vec{k} \cdot (\vec{r}_n - \vec{r}_m)\right) E_j \int \varphi_j^*(\vec{r} - \vec{r}_m) \varphi_j(\vec{r} - \vec{r}_n) d\vec{r} \\
&\quad + \sum_{m,n} \exp\left(i\vec{k} \cdot (\vec{r}_n - \vec{r}_m)\right) \int \varphi_j^*(\vec{r} - \vec{r}_m) v(\vec{r} - \vec{r}_n) \varphi_j(\vec{r} - \vec{r}_n) d\vec{r}
\end{aligned}$$

Il primo termine dell'ultima uguaglianza e' identico, a parte il fattore costante E_j che si puo' portare a sinistra del segno di sommatoria, a quello visto al punto 263 e quindi, assumendo anche in questo caso che solo gli integrali con $n = m$ siano apprezzabilmente diversi da zero, esso e' uguale a NE_j :

$$= NE_j + \sum_{m,n} \exp\left(i\vec{k} \cdot (\vec{r}_n - \vec{r}_m)\right) \int \varphi_j^*(\vec{r} - \vec{r}_m) v(\vec{r} - \vec{r}_n) \varphi_j(\vec{r} - \vec{r}_n) d\vec{r}$$

265. Ora ci occupiamo dell'ultimo termine. Cominciamo con lo scrivere la somma su m ed n come due somme separate:

$$= NE_j + \sum_n \left(\sum_m \exp\left(i\vec{k} \cdot (\vec{r}_n - \vec{r}_m)\right) \int \varphi_j^*(\vec{r} - \vec{r}_m) v(\vec{r} - \vec{r}_n) \varphi_j(\vec{r} - \vec{r}_n) d\vec{r} \right)$$

Poi evidenziamo il termine con $m = n$ nella somma su m :

$$\begin{aligned}
&= NE_j + \sum_n \left(\int \varphi_j^*(\vec{r} - \vec{r}_n) v(\vec{r} - \vec{r}_n) \varphi_j(\vec{r} - \vec{r}_n) d\vec{r} \right. \\
&\quad \left. + \sum_{m \neq n} \exp\left(i\vec{k} \cdot (\vec{r}_n - \vec{r}_m)\right) \int \varphi_j^*(\vec{r} - \vec{r}_m) v(\vec{r} - \vec{r}_n) \varphi_j(\vec{r} - \vec{r}_n) d\vec{r} \right)
\end{aligned}$$

e spezziamo la somma su n nei due addendi che la costituiscono:

$$\begin{aligned}
&= NE_j + \sum_n \int \varphi_j^*(\vec{r} - \vec{r}_n) v(\vec{r} - \vec{r}_n) \varphi_j(\vec{r} - \vec{r}_n) d\vec{r} \\
&\quad + \sum_n \left(\sum_{m \neq n} \exp\left(i\vec{k} \cdot (\vec{r}_n - \vec{r}_m)\right) \int \varphi_j^*(\vec{r} - \vec{r}_m) v(\vec{r} - \vec{r}_n) \varphi_j(\vec{r} - \vec{r}_n) d\vec{r} \right)
\end{aligned} \tag{733}$$

Gli integrali della prima somma su n sono tutti uguali per la simmetria del reticolo di Bravais, perche' contengono la *stessa* funzione φ_j e lo *stesso* potenziale v , solamente shiftati su tutti i punti reticolari del cristallo.

Poniamo allora:

$$\int \varphi_j^* (\vec{r} - \vec{r}_n) v (\vec{r} - \vec{r}_n) \varphi_j (\vec{r} - \vec{r}_n) d\vec{r} = -A \quad (734)$$

e quindi:

$$\sum_n \int \varphi_j^* (\vec{r} - \vec{r}_n) v (\vec{r} - \vec{r}_n) \varphi_j (\vec{r} - \vec{r}_n) d\vec{r} = - \sum_n A = -A \sum_n 1 = -NA \quad (735)$$

Per poter scrivere la seconda somma della 733 in modo piu' semplice facciamo ora due ipotesi addizionali:

1. limitiamo il calcolo degli integrali ai soli primi vicini del sito n ; in altre parole supponiamo che l'indice m (nella somma su $m \neq n$) corra sui soli primi vicini del sito n . Questa assunzione e' ragionevole dal momento che siamo nell'ambito del modello "tight binding", che prevede proprio il "quasi isolamento" degli atomi presenti nel cristallo
2. assumiamo che la funzione atomica φ_j abbia simmetria sferica; questo in pratica significa assumere che sia una funzione di tipo s (le funzioni atomiche monoelettroniche di tipo p d etc *non* hanno simmetria sferica)

Queste due ipotesi non sono essenziali, ma ci consentono di semplificare molto il trattamento seguente.

Allora: grazie all'ipotesi 1 possiamo dire che:

$$|\vec{r}_n - \vec{r}_m| = \text{costante}$$

perche' tutti i primi vicini di un sito sono a distanza costante da esso (notate che il modulo di $\vec{r}_n - \vec{r}_m$ e' costante, ma non la sua direzione).

Ma allora, grazie all'ipotesi 2, l'integrale nella somma:

$$\left(\sum_{m \neq n} \exp \left(i\vec{k} \cdot (\vec{r}_n - \vec{r}_m) \right) \int \varphi_j^* (\vec{r} - \vec{r}_m) v (\vec{r} - \vec{r}_m) \varphi_j (\vec{r} - \vec{r}_m) d\vec{r} \right)$$

della 733 ha sempre lo stesso valore, indipendente sia da m che da n . E' indipendente da m perche', limitandoci ai soli primi vicini, la funzione $\varphi_j^* (\vec{r} - \vec{r}_m)$ e' centrata sempre alla stessa distanza dalla $\varphi_j (\vec{r} - \vec{r}_n)$ (punto 1 sopra) e le due funzioni (e il potenziale $v (\vec{r} - \vec{r}_n)$) hanno simmetria sferica (punto 2); e' indipendente da n per la simmetria traslazionale del reticolo di Bravais, come gia' notato per il primo termine della 733. Allora possiamo porre:

$$\int \varphi_j^* (\vec{r} - \vec{r}_m) v (\vec{r} - \vec{r}_m) \varphi_j (\vec{r} - \vec{r}_m) d\vec{r} = -B \quad (736)$$

e quindi:

$$\begin{aligned} & \sum_n \left(\sum_{m \neq n} \exp \left(i\vec{k} \cdot (\vec{r}_n - \vec{r}_m) \right) \int \varphi_j^* (\vec{r} - \vec{r}_m) v (\vec{r} - \vec{r}_n) \varphi_j (\vec{r} - \vec{r}_n) d\vec{r} \right) \\ &= -B \sum_n \sum'_m \exp \left(i\vec{k} \cdot (\vec{r}_n - \vec{r}_m) \right) \end{aligned}$$

dove l'apice ci ricorda che m corre sui soli primi vicini del sito n .

Di nuovo, per la simmetria del reticolo di Bravais, la somma su m non dipende da n e quindi arriviamo a:

$$= -B \sum'_m \exp \left(i\vec{k} \cdot (\vec{r}_n - \vec{r}_m) \right) \sum_n 1 = -NB \sum'_m \exp \left(i\vec{k} \cdot (\vec{r}_n - \vec{r}_m) \right) \quad (737)$$

266. Sostituiamo a ritroso la 735 e la 737 nella 733 e quest'ultima insieme con la 732 nella 731:

$$\epsilon \left(\vec{k} \right) = E_j - A - B \sum'_m \exp \left(i\vec{k} \cdot (\vec{r}_n - \vec{r}_m) \right) \quad (738)$$

Questo e' il risultato finale.

L'equazione 738 stabilisce una relazione fra un certo set di autovalori dell'equazione d'onda per il cristallo e un particolare autovalore dell'atomo isolato, E_j .

E' facile vedere che $\epsilon \left(\vec{k} \right)$ e' periodica nello spazio \vec{k} . Per un generico vettore reciproco \vec{G} :

$$\begin{aligned} \epsilon \left(\vec{k} + \vec{G} \right) &= E_j - A - B \sum'_m \exp \left(i \left(\vec{k} + \vec{G} \right) \cdot (\vec{r}_n - \vec{r}_m) \right) \\ &= E_j - A - B \sum'_m \exp \left(i\vec{k} \cdot (\vec{r}_n - \vec{r}_m) \right) \exp \left(i\vec{G} \cdot (\vec{r}_n - \vec{r}_m) \right) \\ &= E_j - A - B \sum'_m \exp \left(i\vec{k} \cdot (\vec{r}_n - \vec{r}_m) \right) \end{aligned}$$

perche' il prodotto scalare di un vettore reciproco (\vec{G}) per un vettore diretto ($\vec{r}_n - \vec{r}_m$) e' uguale a un multiplo intero di 2π (eq. 126).

Quindi il set di autovalori del cristallo trovato e' individuato dai punti \vec{k} interni alla prima zona di Brillouin e costituisce quella che nei due modelli precedenti (elettroni liberi e quasi liberi) abbiamo chiamato banda.

Diversamente dai modelli precedenti, nel modello "tight binding" ciascuna banda e' legata a un livello elettronico dell'atomo isolato. Si parla quindi di bande s , p , d etc.

L'ampiezza della banda e' data dalla differenza fra il valore massimo e quello minimo del termine che dipende da \vec{k} ed e' scalata dal fattore B . Quest'ultimo (eq. 736) e' un integrale di sovrapposizione fra la funzione φ_j centrata sul sito n e la stessa funzione, ma centrata su un sito m suo primo vicino.

Da cio' segue che bande corrispondenti a livelli elettronici dell'atomo isolato piu' bassi (cioe' corrispondenti a orbitali atomici interni) saranno piu' strette delle bande corrispondenti a livelli elettronici piu' alti, caratterizzati da orbitali piu' espansi (orbitali di valenza), che danno maggiore sovrapposizione.

Anche il termine A che compare nella equazione 738 e' un integrale di sovrapposizione (eq. 734).

Se immaginiamo di espandere progressivamente il cristallo, la distanza interatomica aumenta e quindi sia A che B tendono a zero: l'ampiezza della banda tende a zero ed $\epsilon(\vec{k}) \rightarrow E_j$, cioe' si ritrova giustamente il risultato valido per l'atomo isolato.

Visto dalla parte opposta, se immaginiamo di costruire un cristallo avvicinando gli atomi che lo costituiscono, l'eq. 738 dice che ogni livello elettronico dell'atomo isolato si "espande" progressivamente in una banda la cui ampiezza e' tanto maggiore quanto maggiore e' la sovrapposizione fra gli orbitali corrispondenti e il cui "baricentro" si sposta di A rispetto all'autovalore dell'atomo isolato.

267. Possiamo farci un'idea piu' concreta del risultato ottenuto se ricaviamo esplicitamente la dipendenza $\epsilon(\vec{k})$ per il caso specifico di un cristallo avente un reticolo di Bravais cubico semplice con costante di cella a .

In questo caso ci sono 6 primi vicini di un atomo preso come origine, a una distanza pari ad a . I vettori differenza ($\vec{r}_n - \vec{r}_m$) saranno quindi:

$$(\pm a, 0, 0), (0, \pm a, 0), (0, 0, \pm a)$$

Posto $\vec{k} = (k_x, k_y, k_z)$, la somma nell'equazione 738 diventa:

$$\begin{aligned} \sum'_m \exp(i\vec{k} \cdot (\vec{r}_n - \vec{r}_m)) &= \exp(ik_x a) + \exp(-ik_x a) \\ &+ \exp(ik_y a) + \exp(-ik_y a) \\ &+ \exp(ik_z a) + \exp(-ik_z a) \\ &= 2(\cos(k_x a) + \cos(k_y a) + \cos(k_z a)) \end{aligned}$$

e quindi:

$$\epsilon(\vec{k}) = E_j - A - 2B(\cos(k_x a) + \cos(k_y a) + \cos(k_z a))$$

La somma dei tre coseni e' compresa nell'intervallo $[-3, 3]$ per cui l'ampiezza della banda e' $2B \times (3 - (-3)) = 12B$.

La trasformazione dei livelli elettronici dell'atomo isolato nelle corrispondenti bande man mano che gli atomi si avvicinano per formare il cristallo e' mostrata qualitativamente nella figura 73.

I grafici di $\epsilon(\vec{k})$ lungo la direzione reciproca $[111]$ si ottengono osservando che, lungo tale direzione, si ha:

$$k_x = k_y = k_z = k$$

e quindi:

$$\epsilon(\vec{k}) = E_j - A - 6B \cos(ka)$$

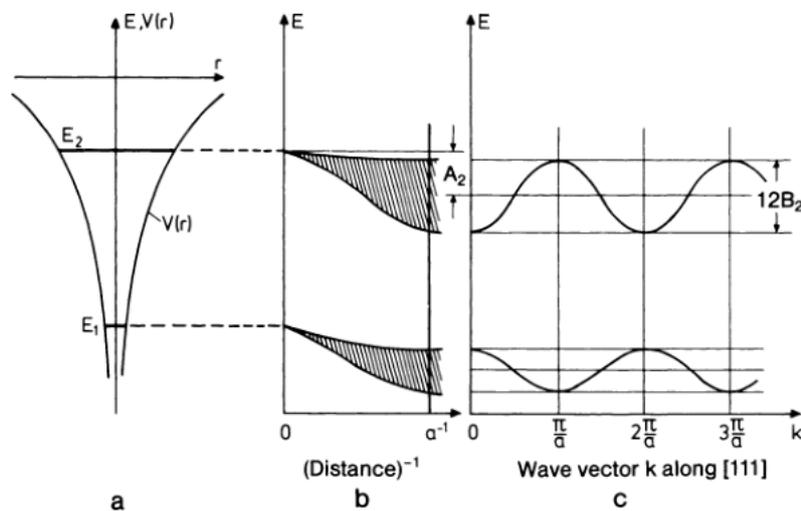


Figura 73: a: due livelli elettronici nel potenziale dell'atomo isolato; b: man mano che gli atomi si avvicinano (la distanza inversa cresce), i livelli si allargano in corrispondenti bande; c: l'andamento di $\epsilon(\vec{k})$ per le due bande nella direzione reciproca [111].

Conduttori, semiconduttori e isolanti

268. Tutto cio' che abbiamo visto finora costituisce solo una piccola parte della teoria che tratta la struttura elettronica dei solidi. Non abbiamo visto, ne' vedremo per mancanza di spazio, alcuna applicazione di questa teoria.

Col poco che sappiamo possiamo pero' comprendere da un punto di vista fondamentale il motivo per cui alcuni elementi o composti sono dei conduttori elettrici, altri sono degli isolanti e altri ancora hanno caratteristiche intermedie (e per questo vengono detti semiconduttori).

269. Innanzitutto ricordiamo che ogni banda contiene N vettori \vec{k} permessi, dove N e' il numero di celle primitive presenti nel cristallo (punto 204). N e' il numero di punti \vec{k} permessi nella prima zona di Brillouin e sia gli autovettori che i corrispondenti autovalori per un dato indice di banda sono funzioni periodiche di \vec{k} (punto 223): quindi ogni banda contiene N stati *indipendenti* (per ogni vettore del reticolo reciproco \vec{G} e un indice di banda fissato, lo stato $\Psi_{\vec{k}+\vec{G}}(\vec{r})$ e' un'immagine dello stato $\Psi_{\vec{k}}(\vec{r})$, con \vec{k} interno alla prima zona di Brillouin).

Ad ogni punto \vec{k} permesso corrispondono 2 possibili stati elettronici che differiscono per lo spin; quindi ogni banda puo' "ospitare" $2N$ elettroni.

270. Limitandoci per semplicita' a cristalli costituiti da un unico elemento, il "riempimento" delle bande dipendera' dal numero atomico e dal numero di atomi contenuti nella cella primitiva del cristallo.

In un cristallo costituito da un elemento con numero atomico Z e avente una cella primitiva che contiene N_A atomi sono presenti NN_AZ elettroni. A 0K tali elettroni occuperanno le bande del cristallo a partire da quelle a energia piu' bassa.

Sulla base di quanto visto per il modello “tight binding”, possiamo dire che gli elettroni di core occuperanno le bande a energia minore mentre gli elettroni di valenza andranno ad occupare quelle a energia maggiore.

Tenendo presente quanto detto sulla distribuzione di Fermi–Dirac (punto 123 e successivi) e’ chiaro che, ai fini delle proprieta’ di conduzione della corrente elettrica (e di molte altre proprieta’ dei solidi), ha senso considerare solo gli elettroni a maggior energia (cioe’ quelli vicini al livello di Fermi); infatti, gli elettroni a piu’ bassa energia non hanno livelli energetici “vuoti” nei quali possano essere eccitati.

La distribuzione degli elettroni di valenza nelle bande a energia maggiore dipende dal loro numero.

Per cristalli con un numero *dispari* di elettroni di valenza per cella primitiva ($N_A Z$ dispari), le prime $N(N_A Z - 1)/2N = (N_A Z - 1)/2$ bande sono completamente occupate mentre la banda a energia maggiore sara’ *semioccupata*. Per cristalli con un numero di elettroni di valenza per cella primitiva *pari* ($N_A Z$ pari) tutte le prime $NN_A Z/2N = N_A Z/2$ bande saranno *completamente occupate*.

271. Se la banda a energia maggiore (detta anche *banda di valenza*) e’ semioccupata (o in generale parzialmente occupata) gli elettroni con energia vicina all’energia di Fermi hanno livelli energetici vuoti immediatamente accessibili.⁵⁸ Se il solido viene sottoposto a una differenza di potenziale, gli elettroni possono venire accelerati (aumentando la loro energia e cambiando il loro vettore d’onda) e il solido avra’ un comportamento da conduttore elettrico.

Se invece la banda a energia maggiore e’ completamente occupata ed e’ separata dalla banda successiva (che viene detta *banda di conduzione*) da un band gap, allora l’accensione di una differenza di potenziale a due estremi opposti del solido non causera’ passaggio di corrente perche’ gli elettroni non possono accedere (facilmente) a stati con energia maggiore: in questo caso il comportamento del solido sara’ quello di un isolante.

272. L’alluminio ha 3 elettroni di valenza e un singolo atomo per cella primitiva: esso e’ un ottimo conduttore elettrico. Lo stesso vale per tutti i metalli alcalini, che hanno un solo elettrone di valenza.

Il carbonio ha 4 elettroni di valenza e (nella struttura del diamante) 2 atomi per cella primitiva: in tali condizioni esso e’ un isolante.

273. I metalli alcalino terrosi sono dei conduttori, pur avendo 2 elettroni di valenza. Il motivo di cio’ e’ che in questi sistemi la banda di valenza completamente occupata e’ sovrapposta alla banda successiva e quindi gli elettroni di valenza possono acquistare facilmente energia da un campo elettrico.

274. Non e’ detto che se la banda di valenza e’ completamente occupata il solido sia necessariamente un isolante. Se il gap energetico fra la banda di valenza e la banda di conduzione e’ sufficientemente piccolo (dell’ordine di 1 – 2 eV) allora a temperatura ambiente una frazione degli elettroni che a 0 K sarebbero tutti nella banda di valenza, puo’ trovarsi nella banda di conduzione. In questo caso gli elettroni nella banda di conduzione (e anche le vacanze elettroniche lasciate

⁵⁸La spaziatura dei livelli appartenenti alla stessa banda e’ dell’ordine di:

$$\frac{\hbar^2 \left(\sqrt[3]{\frac{8\pi^3}{V}} \right)^2}{2m}$$

dove V e’ il volume del solido ed m la massa dell’elettrone; questo numero e’ cosi’ piccolo che i livelli di una stessa banda sono a tutti gli effetti continui.

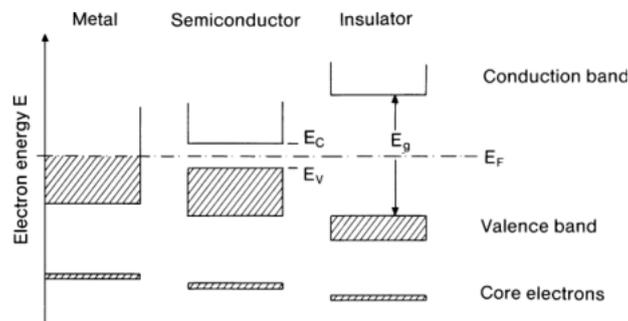


Figura 74:

nella banda di valenza) possono determinare una apprezzabile conducibilità elettrica. I sistemi di questo tipo vengono definiti semiconduttori. La figura 74 mostra molto schematicamente la relazione fra la posizione reciproca delle bande di valenza e conduzione, la loro occupazione e il comportamento da conduttore, semiconduttore o isolante.

Riferimenti bibliografici

- [1] Wikipedia contributors, Euler's formula — Wikipedia, The Free Encyclopedia. 2019; https://en.wikipedia.org/w/index.php?title=Euler%27s_formula&oldid=904652298, [Online; accessed 24-August-2019].
- [2] Andre' Nicolas, Integral of periodic function over the length of the period is the same everywhere. 2012; <https://math.stackexchange.com/questions/98409/integral-of-periodic-function-over-the-length-of-the-period-is-the-same-everywhere>, [Online; accessed 24-August-2019].
- [3] Kittel, C. Introduction To Solid State Physics, 7th ed.; John Wiley & Sons: New York, 1996.
- [4] Ref. 3, ch. 2.
- [5] Harris, J. W.; Stocker, H. The Handbook of Mathematics and Computational Science; Springer-Verlag: Berlin, Heidelberg, 1998.
- [6] Ron Horgan, Brillouin zones pictures. 2012; <http://www.damtp.cam.ac.uk/user/rrh/notes/bzpics.pdf>.
- [7] Ashcroft, N. W.; Mermin, N. D. Solid State Physics; Harcourt Brace College Publishers, 1976.
- [8] Ref. 7, p. 40-42.
- [9] Schroeder, D. V. An Introduction To Thermal Physics; Addison Welsey Longman, 2000.

- [10] Ref. 9, p.247.
- [11] Ref. 7, p. 144.
- [12] Dirac, P. The Principles of Quantum Mechanics; Comparative Pathobiology - Studies in the Postmodern Theory of Education; Clarendon Press, 1981.
- [13] Ref. 7, ch. 8.
- [14] Ref. 3, ch. 7.
- [15] Ibach, H.; Lüth, H. Solid-State Physics: An Introduction to Principles of Materials Science, 3rd ed.; Springer: Berlin, 2003.
- [16] Ref. 15, ch. 7.
- [17] Ref. 7, app. D.
- [18] Ref. 3, ch. 7.
- [19] Ref. 3, ch. 7, fig. 8.
- [20] Ref. 15, fig. 7.3.
- [21] Ref. 3, ch. 9, figg. 6, 7, 8 e 9.
- [22] Ref. 3, pp. 189-192.
- [23] Ref. 15, sec. 7.2.
- [24] Ref. 7, pp. 156-159.
- [25] Ref. 3, ch. 7, fig. 3a.
- [26] Ref. 15, fig. 7.4.
- [27] Ref. 15, sec. 7.3.
- [28] Ref. 3, p. 245-248.
- [29] Ref. 7, ch. 10.