

Capitolo 4

La Struttura dello Stato Solido

4.1 Teorema di Bloch

4.1.1 Introduzione

Si cominci col considerare un *solido unidimensionale*, il cui modello sarà quello di un elettrone in un potenziale periodico. L'equazione di Schrödinger indipendente dal tempo che descrive questo sistema è, come è noto:

$$\frac{d^2}{dx^2}\psi(x) + \frac{2m}{\hbar} [E - V(x)] \psi(x) = 0$$

dove il potenziale periodico con periodo a è rappresentato dalla condizione:

$$V(x) = V(x + a)$$

È possibile dimostrare (teorema di Floquet) che le funzioni d'onda soluzioni di questa equazione di Schrödinger (autofunzioni dell'energia) sono tali da rimanere uguali a meno di un fattore moltiplicativo Q , detto *fattore di Floquet*, per spostamento di un periodo:

$$\psi(x + a) = Q\psi(x)$$

La dimostrazione è basata sul fatto che a causa della periodicità del potenziale, l'equazione di Schrödinger è invariante per traslazione di un periodo. In sostanza, l'Hamiltoniana è invariante sotto trasformazioni $x \rightarrow x + a$, perché il potenziale periodico non cambia, così come non cambia l'operatore laplaciano, che è un operatore differenziale e non risente quindi della costante a .

Il teorema di Bloch che ci si accinge a studiare parte da questo assunto e si occupa di classificare le soluzioni dell'equazione di Schrödinger con potenziale periodico cercando i possibili valori che può assumere il fattore moltiplicativo di Floquet Q .

Come anticipazione si osservi che in particolare una funzione del tipo:

$$\psi(x) = e^{\pm ikx} u_k(x)$$

con $u_k(x)$ funzione periodica di periodo a , cioè tale che $u_k(x) = u_k(x+a)$, è proprio del tipo richiesto. La verifica è del tutto banale:

$$\psi(x + a) = e^{\pm ik(x+a)} u_k(x + a) = e^{\pm ika} e^{\pm ikx} u_k(x + a) = e^{\pm ika} \psi(x)$$

In questa forma particolare dunque il fattore di Floquet $Q = e^{\pm ika}$ è costituito da una fase. Questo è compatibile con il fatto che spostandosi di a la fisica del sistema non cambia:¹ infatti un fattore di fase non cambia la descrizione fisica data dalla funzione d'onda. Si noti che il k che compare in questa particolare forma del fattore di Floquet può variare e pertanto descrive una famiglia di possibili valori.

¹Perché nell'ipotesi di potenziale periodico, naturalmente.

4.1.2 Teorema di Bloch

Enunciato del teorema

Una volta stabilito tramite il teorema di Floquet che le funzioni d'onda soluzioni dell'equazione di Schrödinger di un sistema con potenziale periodico sono tali da rimanere uguali per spostamento di un periodo, a meno di un fattore moltiplicativo denominato *fattore di Floquet* Q :

$$\psi(x + a) = Q\psi(x)$$

si vuole dimostrare ora che queste funzioni si possono scrivere proprio nella forma:

$$\psi(x) = e^{\pm ikx} u_k(x) = b(x; k) \quad \text{Funzioni di Bloch}$$

con $u_k(x)$ funzione periodica di periodo a .

In altri termini, si vuole dimostrare che il fattore di Floquet Q può assumere solo i valori $Q = e^{\pm ika}$.

Note supplementari

Il teorema di Floquet è un teorema matematico del tutto generale che riguarda le equazioni differenziali con coefficienti non costanti periodici, come nel caso dell'equazione di Schrödinger con un potenziale periodico; esso introduce in effetti solo il fattore di Floquet. Bloch ha dato un'interpretazione fisica di questo teorema, scrivendo il fattore di Floquet come un esponenziale e quindi introducendo un vettore d'onda k (in realtà a tre dimensioni, infatti \vec{k} risulta essere un vero vettore). Questa interpretazione fisica si presta, oltre che al modello quantistico dei solidi, ad altre situazioni fisiche in cui ci siano onde in un mezzo periodico.

L'equazione di Schrödinger è un'equazione differenziale e dunque se non si specificano le condizioni iniziali ha una famiglia di infinite soluzioni. Tuttavia, trattandosi di un'equazione differenziale lineare del secondo ordine, le infinite soluzioni differiscono solo per due costanti. D'altra parte, il teorema di esistenza e unicità garantisce che una volta specificate le condizioni iniziali la soluzione è effettivamente unica. Queste due condizioni insieme implicano che se si trova che una certa funzione è soluzione dell'equazione (anche senza specificare le condizioni al contorno), allora la *forma delle soluzioni* è nota e fissata, e possono variare solo per delle costanti.

Nella dimostrazione ad un certo punto si dirà "si supponga che la soluzione abbia la forma...". Questo può confondere, perché sembra un'ipotesi aggiuntiva. In realtà per quanto detto sopra, se si riesce a dimostrare che la forma ipotizzata soddisfa effettivamente l'equazione, allora per il teorema di esistenza e unicità si sa che la forma generale della soluzione cercata è realmente quella ipotizzata.

Dimostrazione del teorema

Si consideri una coppia di soluzioni indipendenti dell'equazione di Schrödinger $f(x)$ e $g(x)$. Essendo l'equazione di Schrödinger un'equazione differenziale del second'ordine, esistono sicuramente due soluzioni linearmente indipendenti e tutte le altre soluzioni risultano essere combinazioni lineari di queste.

Poiché l'equazione di Schrödinger va in sé stessa se ci si sposta di un periodo, anche $f(x + a)$ e $g(x + a)$ devono essere soluzione di tale equazione.² Siccome si tratta di un'equazione differenziale del secondo ordine, essa ha solo due soluzioni linearmente indipendenti e quindi queste altre due soluzioni devono potersi scrivere come combinazione lineare di $f(x)$ e $g(x)$:

$$\begin{aligned} f(x + a) &= \alpha_1 f(x) + \alpha_2 g(x) \\ g(x + a) &= \beta_1 f(x) + \beta_2 g(x) \end{aligned}$$

$\alpha_1, \alpha_2, \beta_1, \beta_2$ sono dei parametri che dipendono dall'energia E e dalle condizioni iniziali: fissato il valore dell'energia, queste sono quindi determinate solo dalle condizioni iniziali.³

²Per convincersene basta sostituire $x \rightarrow x + a$ nell'Hamiltoniana e notare che sia il potenziale periodico che l'operatore laplaciano (operatore differenziale) restano invariati, tranne che per un fattore moltiplicativo.

³Per convincersene basta pensare che l'equazione in questione è un'equazione differenziale lineare con un parametro.

Avendo già assodato tramite il teorema di Floquet che le soluzioni dell'equazione di Schrödinger in questione godono della proprietà:

$$\psi(x+a) = Q\psi(x)$$

occorre cercare i possibili valori del fattore di Floquet Q . La soluzione dell'equazione si deve poter scrivere come combinazione lineare delle due soluzioni linearmente indipendenti $f(x)$ e $g(x)$:

$$\psi(x) = Af(x) + Bg(x)$$

Si consideri ora la funzione $\psi(x+a)$:

$$\begin{aligned}\psi(x+a) &= Af(x+a) + Bg(x+a) = A[\alpha_1 f(x) + \alpha_2 g(x)] + B[\beta_1 f(x) + \beta_2 g(x)] = \\ &= [A\alpha_1 + B\beta_1]f(x) + [A\alpha_2 + B\beta_2]g(x)\end{aligned}$$

che sostituite nella relazione $\psi(x+a) = Q\psi(x)$ permette di scrivere:

$$[A\alpha_1 + B\beta_1]f(x) + [A\alpha_2 + B\beta_2]g(x) = Q[Af(x) + Bg(x)] = QAf(x) + QBg(x)$$

uguagliando i coefficienti di queste due combinazioni lineari di $f(x)$ e $g(x)$ si ottiene un sistema di due equazioni:

$$\begin{cases} A\alpha_1 + B\beta_1 = QA \\ A\alpha_2 + B\beta_2 = QB \end{cases} \Rightarrow \begin{cases} A(\alpha_1 - Q) + B\beta_1 = 0 \\ A\alpha_2 + B(\beta_2 - Q) = 0 \end{cases}$$

Si consideri il significato di queste quantità. Le incognite di questo sistema sono le costanti A e B , cioè la soluzione del sistema fornisce il valore di A e B per i quali è valida la relazione $y(x+a) = Qy(x)$. Si devono quindi imporre delle condizioni sui coefficienti α_1 , α_2 , β_1 , β_2 e Q affinché questo sistema lineare omogeneo nelle incognite A e B non abbia solo la soluzione banale, si deve cioè imporre che il determinante della matrice dei coefficienti sia nullo:

$$\begin{vmatrix} \alpha_1 - Q & \beta_1 \\ \alpha_2 & \beta_2 - Q \end{vmatrix} = 0 \Rightarrow Q^2 - (\alpha_1 + \beta_2)Q + \alpha_1\beta_2 - \alpha_2\beta_1 = 0$$

Questa quindi è l'equazione cui deve soddisfare il fattore di Floquet Q per avere una soluzione non banale dei coefficienti A e B .

Questa equazione può essere messa in forma più semplice. Si consideri la combinazione lineare di partenza e se ne calcoli *Wronskiano*. Il *Teorema del Wronskiano* afferma che questo è indipendente dal punto in cui è calcolato. Ora, il Wronskiano calcolato in x è dato da:

$$\mathcal{W}(x) = \begin{vmatrix} f(x) & g(x) \\ f'(x) & g'(x) \end{vmatrix} = f(x)g'(x) - g(x)f'(x)$$

mentre il Wronskiano calcolato in $x+a$ è dato da:

$$\mathcal{W}(x+a) = \begin{vmatrix} f(x+a) & g(x+a) \\ f'(x+a) & g'(x+a) \end{vmatrix} = \begin{vmatrix} \alpha_1 f(x) + \alpha_2 g(x) & \beta_1 f(x) + \beta_2 g(x) \\ \alpha_1 f'(x) + \alpha_2 g'(x) & \beta_1 f'(x) + \beta_2 g'(x) \end{vmatrix}$$

che sviluppato fornisce:

$$\begin{aligned}& [\alpha_1 f(x) + \alpha_2 g(x)] [\beta_1 f'(x) + \beta_2 g'(x)] - [\beta_1 f(x) + \beta_2 g(x)] [\alpha_1 f'(x) + \alpha_2 g'(x)] = \\ &= \alpha_1 \beta_1 f(x) f'(x) + \alpha_1 \beta_2 f(x) g'(x) + \alpha_2 \beta_1 g(x) f'(x) + \alpha_2 \beta_2 g(x) g'(x) - \\ &\quad - \beta_1 \alpha_1 f(x) f'(x) - \beta_1 \alpha_2 f(x) g'(x) - \beta_2 \alpha_1 g(x) f'(x) - \beta_2 \alpha_2 g(x) g'(x) = \\ &= (\alpha_1 \beta_2 - \beta_1 \alpha_2) f(x) g'(x) - (\beta_2 \alpha_1 - \alpha_2 \beta_1) g(x) f'(x) = \\ &= \mathcal{W}(x) (\alpha_1 \beta_2 - \alpha_2 \beta_1)\end{aligned}$$

il teorema del Wronskiano permette dunque di concludere che:

$$\alpha_1 \beta_2 - \alpha_2 \beta_1 = \begin{vmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{vmatrix} = 1$$

e quindi l'equazione che deve soddisfare il fattore di Floquet Q assume la forma:

$$Q^2 - (\alpha_1 + \beta_2)Q + 1 = 0$$

che ammette evidentemente due soluzioni:

$$Q_{1,2} = \frac{(\alpha_1 + \beta_2) \pm \sqrt{(\alpha_1 + \beta_2)^2 - 4}}{2}$$

Queste due soluzioni soddisfano la proprietà:

$$Q_1 Q_2 = \frac{(\alpha_1 + \beta_2) + \sqrt{(\alpha_1 + \beta_2)^2 - 4}}{2} \cdot \frac{(\alpha_1 + \beta_2) - \sqrt{(\alpha_1 + \beta_2)^2 - 4}}{2} = \frac{(\alpha_1 + \beta_2)^2 - (\alpha_1 + \beta_2)^2 + 4}{4} \equiv 1$$

Dunque la loro particolarità è che per qualunque scelta dei parametri α_1 e β_2 (quindi, in definitiva, per ogni scelta dell'energia) si ha $Q_1 Q_2 = 1$. Quindi è questo il vincolo su Q_1 e Q_2 , cioè in definitiva su Q , cui si giunge imponendo che $\psi(x+a) = Q\psi(x)$.

Nell'ambito di questo vincolo, si possono presentare diversi casi:

- Q reale:

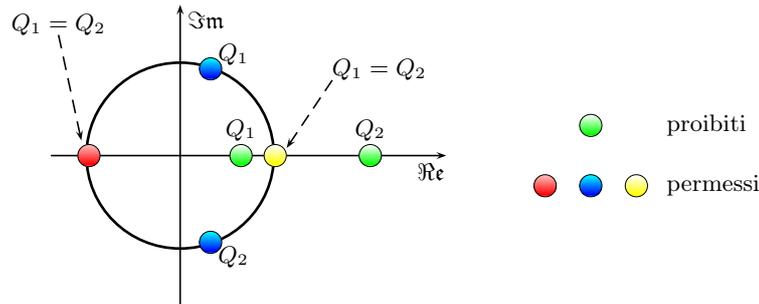
$$\begin{aligned} \rightarrow Q_1 = 1, Q_2 = 1 & \Rightarrow Q_1 Q_2 = 1 \text{ (indicato con } \bullet \text{)} \\ \rightarrow Q_1 = -1, Q_2 = -1 & \Rightarrow Q_1 Q_2 = 1 \text{ (indicato con } \bullet \text{)} \\ \rightarrow Q_1 < 1, Q_2 > 1 & \text{ tali che } Q_1 Q_2 = 1 \text{ (indicato con } \bullet \text{)} \\ \rightarrow Q_1 > 1, Q_2 < 1 & \text{ tali che } Q_1 Q_2 = 1 \text{ (indicato con } \bullet \text{)} \end{aligned}$$

- Q complesso:

$$\rightarrow Q_1, Q_2 \text{ complessi coniugati di modulo 1, in modo che } Q_1 Q_2 = 1 \text{ (indicato con } \bullet \text{)}$$

Si noti però che gli ultimi due casi di Q reale, indicati con \bullet , vanno scartati, poiché implicherebbero una funzione d'onda non più sommabile. Infatti, $Q > 1$ implica che man mano che ci si sposta di periodo in periodo con il valore di x , il valore di $\psi(x)$ viene moltiplicato sempre per un numero maggiore di 1 e quindi diverge all'infinito. Questo significa proprio che la funzione non è a quadrato sommabile e quindi non ha significato fisico.

I valori permessi e proibiti di Q possono essere rappresentati nel piano complesso:



da cui si deduce che i valori permessi giacciono tutti su un cerchio di raggio unitario.

Ricapitolando: ad ogni valore dell'energia corrispondono due autofunzioni, caratterizzate da determinati valori dei parametri α_1, α_2 e β_1, β_2 . Infatti, l'equazione di Schrödinger è del secondo ordine e dunque per ogni valore di E ci sono due soluzioni linearmente indipendenti, essendo tutte le altre combinazioni lineari di due soluzioni indipendenti scelte come base. L'energia è quindi due volte degenere. Questi valori di α_1, α_2 e β_1, β_2 a loro volta determinano una coppia di valori del fattore di Floquet. Si è visto che *solo le energie corrispondenti ai valori del fattore di Floquet sul cerchio unitario del piano complesso sono fisicamente accettabili.*

In questi casi, le funzioni d'onda del sistema possono essere rappresentate nella forma:

$$b(x; k) = e^{ikx} u(x; k)$$

con $u(x; k)$ funzione periodiche in x , di periodo a , denominate *funzioni di Bloch*. Questa forma garantisce la proprietà di periodicità, infatti:

$$\begin{aligned}\psi(x+a) &= Q\psi(x) \\ b(x+a; k) &= e^{ik(x+a)}u(x+a; k) = e^{ika}e^{ikx}u(x; k) = e^{ika}b(x; k) = Qb(x; k)\end{aligned}$$

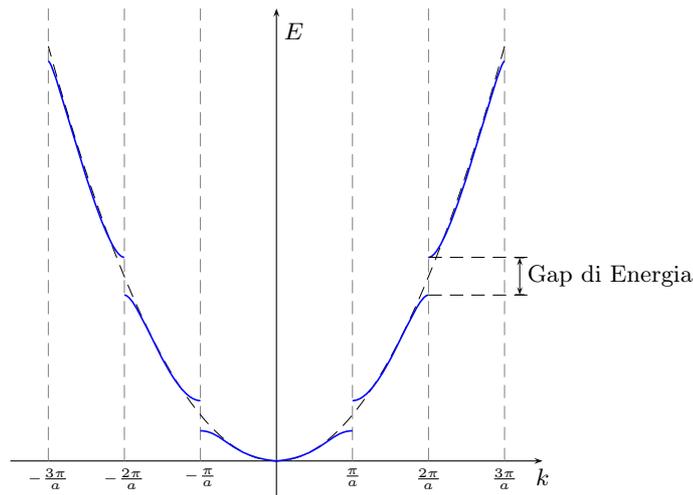
col che risulta dimostrato il teorema di Bloch.

4.1.3 Relazione di dispersione

Occorre ora indagare sulla relazione fra un determinato autovalore dell'energia E e il fattore di Floquet $Q = e^{\pm ik a}$ della corrispondente autofunzione. In altre parole, si cerca la funzione che lega l'energia E e il valore di k .

Questa dipendenza risulta essere quasi lineare. In particolare, questa funzione è monotona crescente, cioè al crescere di k (in modulo) l'energia cresce sempre, nel dominio in cui è definita questa funzione.

Il dominio di definizione deve invece essere definito accuratamente. Si è visto infatti che il fattore di Floquet può assumere dei valori reali diversi da ± 1 che, sebbene permessi da un punto di vista matematico, sono inaccettabili per motivi fisici. Per questo motivo, ci si deve aspettare che i valori $\frac{2n\pi}{a}$ di k siano dei punti di discontinuità per la funzione $E(k)$. Il grafico di questa funzione risulta in effetti del tipo:



Le zone fra due punti di discontinuità sono chiamate *zone di Brillouin*.

Alcune note sulle zone di Brillouin

Gli elettroni, oltre a comportarsi come particelle discrete, possiedono un carattere ondulatorio. Questa proprietà è utilizzata nella descrizione del modello di Sommerfeld degli elettroni liberi. Si consideri quindi un reticolo monodimensionale, con cella unitaria di dimensione a e si supponga di far crescere lentamente l'energia e corrispondentemente k . Quando k diventa abbastanza grande (λ abbastanza piccola) l'onda elettronica andrà incontro ad una riflessione di Bragg, come succede ai raggi X.

La legge di Bragg è espressa da $2d \sin \theta = n\lambda$, in particolare per il caso monodimensionale in studio d coincide con la distanza fra gli atomi a , quindi in una dimensione i riflessi di Bragg si realizzano quando:

$$k = \frac{n\pi}{a} \quad n = \pm 1, \pm 2, \dots$$

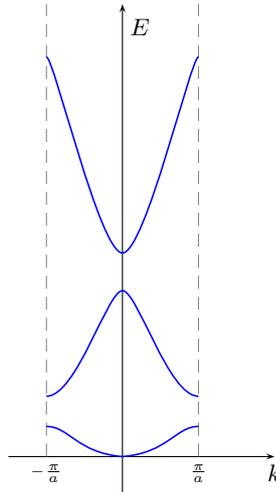
Infatti $k = 2\pi/\lambda$ e per valori $k = \frac{n\pi}{a}$ si ha:

$$2a = n\lambda$$

che rappresenta la condizione di Bragg nel caso monodimensionale. La stessa riflessione di Bragg si verifica quando k raggiunge $\frac{n\pi}{a}$ da valori superiori. *Qualunque tipo di onda propagantesi in un mezzo periodico subisce lo stesso fenomeno (Brillouin).*

Dalla periodicità delle funzioni di Bloch si può desumere che il valore di k è periodico nello spazio k -modulo $\frac{2\pi}{a}$, cioè che un valore k può essere sempre sostituito da $k + \frac{2n\pi}{a}$ (con n intero). Quindi, si possono restringere i valori di k ad un intervallo di lunghezza $\frac{2\pi}{a}$ senza la necessità di considerare k al di fuori di questo range poiché non si guadagna nessuna nuova informazione. Per convenzione l'intervallo è scelto simmetrico attorno a $k = 0$, cioè $-\frac{\pi}{a} \leq k \leq \frac{\pi}{a}$.

La procedura di restrizione dei valori di k in questo intervallo è illustrata graficamente nella figura seguente.



Si procederà ora ad uno studio più dettagliato.

Lemma sugli zeri delle funzioni di Bloch.

La funzione d'onda del sistema (funzione di Bloch):

$$b(x; k) = e^{ikx} u_k(x)$$

non ha zeri sull'asse reale.

Dimostrazione:

Si supponga che esista un $x_0 \in \mathbb{R}$: $b(x_0; k) = 0$. Si consideri la derivata prima di $b(x_0; k)$ in quel punto e la si scriva in termini di modulo e fase, cosa sicuramente possibile in quanto $b(x_0; k)$ è complessa:

$$b'(x_0; k) = \delta e^{i\gamma}$$

D'altra parte, le funzioni di Bloch sono soluzioni di un'equazione differenziale e quindi sono note a meno di una costante moltiplicativa, si può perciò moltiplicare l'intera funzione per la costante $e^{-i\gamma}$ senza cambiare la fisica. In altri termini, si sfrutta il fatto che la funzione d'onda è definita a meno di un fattore di fase per renderla reale:

$$b'(x_0; k) = \delta e^{i\gamma} e^{-i\gamma} = \delta$$

Ne consegue che la funzione $b(x; k)$ stessa può essere resa reale, quindi si possono scegliere le due relazioni:

$$b(x_0; k) = 0 \quad b'(x_0; k) = \delta$$

come condizioni iniziali dell'equazione di Schrödinger. L'equazione è a coefficienti reali, le condizioni iniziali sono reali e quindi le funzioni di Bloch devono essere reali. Ma questo è un assurdo, perché le funzioni di Bloch sono sicuramente funzioni complesse. Infatti, presentano il fattore moltiplicativo di Floquet che può assumere dei valori complessi. Poiché l'assurdo deriva dal fatto che la funzione

di Bloch abbia uno zero sull'asse reale, la conclusione è che le funzioni di Bloch non si annullano mai sull'asse reale.⁴

Proposizione sulla dipendenza di E da k

Si vuole ora dimostrare la monotonia della dipendenza di E da k . Si vedrà che tale monotonia può venir meno solo quando k è tale che Q si trova nei punti ± 1 sulla circonferenza complessa.

Si considerino due valori molto prossimi di k , ovvero k e $k + \epsilon$. Le funzioni di Bloch corrispondenti soddisfano entrambe l'equazione di Schrödinger:

$$\begin{aligned} \left[-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right] b(x; k) &= E(k)b(x; x) \\ \left[-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right] b(x; k + \epsilon) &= E(k)b(x; k + \epsilon) \end{aligned}$$

Si consideri la funzione:

$$\frac{b^2(x; k)}{b(x; k + \epsilon)}$$

questa funzione non diverge perché si è dimostrato che il denominatore $b(k + \epsilon; x)$ non si annulla mai sull'asse reale.

Il fattore di Floquet del denominatore è e^{2ikx} , mentre quello del denominatore è $e^{i(k+\epsilon)x}$, dunque il fattore complessivo è:

$$e^{2ikx} e^{-i(k+\epsilon)x} = e^{i(2k-k-\epsilon)x} = e^{i(k-\epsilon)x}$$

dunque se la funzione $b(x; k + \epsilon)$ ha numero d'onda $k + \epsilon$, la funzione $\frac{b^2(x; k)}{b(x; k + \epsilon)}$ ha numero d'onda $k - \epsilon$. Inoltre, si può provare che $\frac{b^2(x; k)}{b(x; k + \epsilon)}$ è un'autofunzione dell'energia il cui autovalore è:

$$2E(k) - E(k + \epsilon)$$

Per ϵ che tende a zero, si può approssimare questa espressione sviluppando il secondo termine in serie rispetto ad ϵ intorno allo zero:

$$2E(k) - E(k + \epsilon) \approx 2E(k) - E(k) - \epsilon \frac{dE}{dk} = E(k) - \epsilon \frac{dE}{dk}$$

e questo è vero a meno di termini quadratici rispetto a $E(k + \epsilon) - E(k)$ che da quanto appena visto, è a sua volta dell'ordine di $\frac{dE}{dk}$ e quindi si possono trascurare.

Prima di dimostrare che effettivamente che l'autofunzione di Bloch $\frac{b^2(x; k)}{b(x; k + \epsilon)}$, il cui numero d'onda si è visto essere $k - \epsilon$, ha come autovalore $2E(k) - E(k + \epsilon) \approx E(k) - \epsilon \frac{dE}{dk}$, si noti che questo significa che diminuendo k di una quantità ϵ , anche l'energia diminuisce di ϵ se la derivata $\frac{dE}{dk}$ è positiva: se k decresce, anche l'energia E decresce. Se la derivata è invece negativa, la dipendenza risulta invertita ma rimane monotona. Inoltre, come anticipato, la dipendenza è quasi lineare.

La dimostrazione che la funzione di Bloch $\frac{b^2(x; k)}{b(x; k + \epsilon)}$ è autofunzione con autovalore $2E(k) - E(k + \epsilon)$ si ottiene sostituendo direttamente la funzione nell'equazione di Schrödinger:

$$\left[-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right] \frac{b^2(x; k)}{b(x; k + \epsilon)} = [2E(k) - E(k + \epsilon)] \frac{b^2(x; k)}{b(x; k + \epsilon)}$$

Invece di seguire il calcolo, che risulta lungo e non particolarmente istruttivo, si può notare che il metodo delle separazione di variabili dice che l'autofunzione totale è il prodotto delle autofunzioni unidimensionali e che il corrispondente autovalore è la somma di quelli a una dimensione. Ricordando che l'autovalore di $b(x; k)$ è E da questo si può passare a dire che l'autovalore di $b^2(x; k)$ è $2E$ e poi dal fatto che l'autovalore di $b(x; k + \epsilon)$ è $E(k + \epsilon)$ si passa al fatto che l'autovalore di $[b(x; k + \epsilon)]^{-1}$ è $-E(k + \epsilon)$, e mettendo insieme il tutto si ottiene l'autovalore riportato.

⁴Si noti che dire che la $b(x; k)$ non ha zeri sull'asse reale significa dire in pratica che non si annulla mai. Infatti, anche se dal punto di vista matematico si possono avere degli zeri complessi, una volta considerata come funzione d'onda di un sistema fisico, la variabile x può assumere solo valori reali.

4.1.4 Alcune considerazioni sulla relazione di dispersione

La zona di Brillouin è l'intervallo dei valori di k di ampiezza pari al periodo del fattore di Floquet, infatti l'esponenziale e^{ikx} è una funzione periodica rispetto a k di periodo $2\pi/a$. La prima zona di Brillouin è quella centrata nell'origine, a seguire le altre.

Se si ha a che fare con un sistema con potenziale periodico, si è dimostrato che le funzioni d'onda sono funzioni di Bloch, cioè del tipo:

$$\psi(x) = e^{ikx}u(x; k)$$

In particolare, avranno quindi questa forma le autofunzioni dell'energia.

Si è appena dimostrato inserendo direttamente un'autofunzione con questa forma nell'equazione di Schrödinger che l'autovalore dell'energia corrispondente a questa autofunzione dipende da k in maniera quasi lineare. Più in generale le autofunzioni di un sistema con potenziale periodico hanno la forma:

$$Q\varphi(x)$$

Ora, da una parte è noto che alcune autofunzioni, con certi valori di Q , debbono essere scartate per motivi fisici, insieme con le corrispondenti energie. D'altra parte, avendo dimostrato che la dipendenza tra E e k è monotona, si conclude che il fattore di Floquet Q si comporta nel modo seguente al variare dell'energia.

Sia $Q_1 = Q_2$. All'aumentare dell'energia si ottengono i valori sulla circonferenza complessa di raggio unitario. Questi sono tutti valori fisicamente permessi corrispondenti quindi a valori permessi dell'energia e a due fattori di Floquet in forma esponenziale, corrispondenti a loro volta a due valori opposti di k , compresi tra 0 e $\pm\pi/a$.

Giunti in $k = \pm\pi/a$, quindi a $Q_1 = Q_2 = -1$, continuando a salire in energia Q assume dei valori reali fisicamente proibiti e quindi saranno proibite le corrispondenti energie. La funzione $E(k)$, per questi valori di k avrà dunque delle discontinuità.

Continuando a salire in energia, il fattore di Floquet può ritornare sulla circonferenza unitaria, dando luogo ad altre energie permesse, essendo però la $E(k)$ monotona se si continua a salire in energia si arriva ad altre energie proibite, per $k = \pm n\pi/a$.

Quello descritto qui è dunque uno spettro di energia fatto da intervalli continui di energie permesse e intervalli continui di energie proibite. Questi intervalli si denominano *bande*, e si parla di *struttura a bande dello spettro di energia*. A causa di questa struttura, una certa energia può essere individuata specificando il valore di k compreso nel suo intervallo $[-\pi/a, \pi/a]$ (prima zona di Brillouin), e da un numero intero che identifica la banda di energia permessa a cui appartiene l'energia in questione. Questo intero si definisce *indice di banda* e si indica con q .

Le autofunzioni dell'Hamiltoniana del sistema sono allora delle funzioni di Bloch distinte da un indice di banda q , le cui energie corrispondenti sono delle funzioni di k anche esse distinte dall'indice q :

$$\psi(x) = b_q(x; k) = e^{ik_q x}u_q(x) \quad E = w_q(k)$$

queste funzioni godono della proprietà di periodicità:

$$\begin{aligned} b_q(x + a; k) &= e^{ika}b_q(x; k) \\ b_q(x; k + \frac{2\pi}{a}) &= b_q(x; k) \\ w_q(k + \frac{2\pi}{a}) &= w_q(k) \end{aligned}$$

e di simmetria:

$$\begin{aligned} b_q(x; -k) &= b_q^*(x; k) \\ w_q(-k) &= w_q(k) \end{aligned}$$

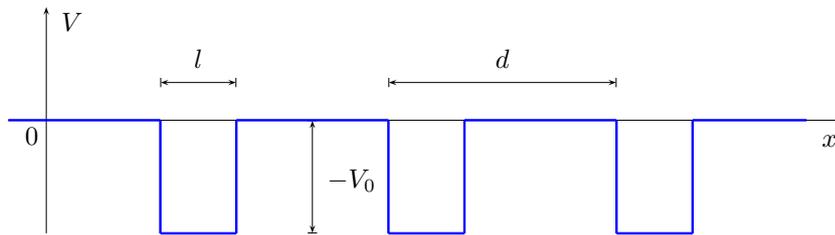
quest'ultima proprietà è legata all'invarianza per time reversal dell'equazione di Schrödinger.

4.2 Modello di Kronig–Penney

Si vuole ora vedere l'applicazione del teorema di Bloch e le sue conseguenze ad un caso specifico, ricavando esplicitamente le bande. Si introdurrà quindi prima un modello che descrive il sistema “elettrone in un solido” (metallico). Si risolverà quindi l'equazione di Schrödinger indipendente dal tempo associata, ricavando gli autovalori e le autofunzioni dell'energia. Si potrà quindi studiare la relazione di dispersione $E(k)$, mettendo in evidenza come, a causa della discontinuità nel dominio di definizione di k e grazie alla monotonia della funzione, si possano organizzare i livelli energetici in una struttura a *bande* e *gap*.

4.2.1 Introduzione al modello

Si consideri un sistema il cui potenziale periodico sia di tipo a buca rettangolare. Questo esempio non è analitico, ma ha il pregio di essere abbastanza semplice.



Si considereranno dapprima le buche finite, con l'intenzione di mandare successivamente la profondità all'infinito e la larghezza a zero, ma in modo tale che il loro prodotto sia costante: le buche saranno quindi rappresentate da funzioni delta.

Fuori dalle buche

Al di fuori delle buche l'equazione di Schrödinger è quella di una particella libera:

$$\frac{d^2}{dx^2} \psi(x) + \frac{2mE}{\hbar} \psi(x) = 0$$

Poiché nelle soluzioni compare la radice quadrata del coefficiente del secondo termine, cioè quello che in genere si chiama k ,⁽⁵⁾ si distinguono due casi a seconda del segno dell'energia. Si definiscono quindi per i due casi due quantità positive e si riscrive l'equazione di Schrödinger in maniera compatta:

$$\begin{cases} \alpha^2 \equiv \frac{2mE}{\hbar} & \rightarrow \frac{d^2}{dx^2} \psi(x) + \alpha^2 \psi(x) = 0 & ; & E > 0 \\ \beta^2 \equiv -\frac{2mE}{\hbar} & \rightarrow \frac{d^2}{dx^2} \psi(x) + \beta^2 \psi(x) = 0 & ; & E < 0 \end{cases}$$

All'interno delle buche

Si consideri preliminarmente il caso in cui il potenziale abbia una sola buca nell'origine. All'interno della buca centrata nello zero l'equazione di Schrödinger è:

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) - V_0 \psi(x) = E \psi(x) \rightarrow -\frac{d^2}{dx^2} \psi(x) = \frac{2m}{\hbar^2} (V_0 + E) \psi(x)$$

Si risolverà quindi questa equazione differenziale integrando ambo i membri intorno all'origine e mandando successivamente $l \rightarrow 0$ e $V_0 \rightarrow \infty$, in modo però che il loro prodotto resti finito. In

⁵Si faccia attenzione a non confonderlo con il k del fattore di Bloch, che è una quantità diversa

pratica si integra in un intorno dell'origine largo quanto la buca (finita), e si manda tale larghezza a zero:

$$\lim_{\substack{l \rightarrow 0 \\ V_0 \rightarrow \infty}} \int_{-l/2}^{l/2} \frac{d^2}{dx^2} \psi(x) dx = -\frac{2m}{\hbar^2} \lim_{\substack{l \rightarrow 0 \\ V_0 \rightarrow \infty}} \left[\int_{-l/2}^{l/2} V_0 \psi(x) dx + \int_{-l/2}^{l/2} E \psi(x) dx \right]$$

nell'integrare il membro di destra si tiene conto che nell'intervallo di integrazione la $\psi(x)$ è quasi costantemente uguale a $\psi(0)$, e si può quindi portare fuori dall'integrale come fosse una vera costante:

$$\lim_{\substack{l \rightarrow 0 \\ V_0 \rightarrow \infty}} \left[\frac{d}{dx} \psi(x) \Big|_{l/2} - \frac{d}{dx} \psi(x) \Big|_{-l/2} \right] = -\frac{2m}{\hbar^2} \lim_{\substack{l \rightarrow 0 \\ V_0 \rightarrow \infty}} \psi(0) [V_0 l - El]$$

svolgendo ora i limiti, nel membro di destra il termine $V_0 l$ si mantiene costante e il termine El va a zero, mentre il primo membro non dipende da V_0 :

$$\frac{d}{dx} \psi(x) \Big|_+ - \frac{d}{dx} \psi(x) \Big|_- = -\frac{2mV_0 l}{\hbar^2} \psi(0) \equiv -2\eta \psi(0) \quad \left(\eta \equiv \frac{mV_0 l}{\hbar^2} \right)$$

Questa equazione rappresenta una condizione sul salto della derivata di $\psi(x)$ in corrispondenza della buca a delta, cioè la differenza di valore che la derivata di $\psi(x)$ deve assumere subito prima e subito dopo la buca. In particolare, questo salto è legato al valore della funzione in quel punto tramite il fattore η che può essere definito *forza della buca*.

A questo punto può essere utilizzato il teorema di Bloch, il quale afferma che:

$$\psi(0) = e^{-ikd} \psi(d)$$

Siccome il fattore di Floquet non dipende da x , questa relazione deve valere anche per le derivate e il fattore può essere portato fuori della derivata:

$$\frac{d}{dx} \psi(x) \Big|_{x=0} = e^{-ikd} \frac{d}{dx} \psi(x) \Big|_{x=d}$$

che sostituita nella condizione sul salto della derivata fornisce:

$$\frac{d}{dx} \psi(x) \Big|_{x=0} - \frac{d}{dx} \psi(x) \Big|_{x=d} e^{-ikd} = -\frac{2mV_0 l}{\hbar^2} \psi(0)$$

Si è quindi utilizzato il teorema di Bloch per sostituire la derivata sinistra con la derivata calcolata in $x = d$, passando quindi ad una condizione sui bordi (di un periodo) della buca.

4.2.2 Stati legati della buca a delta

Si cercheranno ora gli stati legati del sistema costituito da una sola buca a delta messa nell'origine. Si scoprirà che ne esiste uno solo: questo argomento è importante soprattutto per il modello dell'elettrone fortemente legato.

Poiché la larghezza della buca a delta è per definizione infinitesima (al limite, nulla), il potenziale è quasi ovunque nullo⁶ e quindi l'equazione di Schrödinger sarà un'equazione di particella libera. Si distinguono due casi a seconda se si considerano energie negative o positive.

Energia negativa

In questo caso l'equazione da risolvere è:

$$\frac{d^2}{dx^2} \psi(x) + \beta^2 \psi(x) = 0 \quad \left(\beta^2 \equiv -\frac{2mE}{\hbar^2} \right)$$

si tratta di risolvere questa equazione a destra e a sinistra della buca a delta centrata in 0 e poi raccordare queste due equazioni tramite la condizione trovata sul salto della derivata. In pratica, è

⁶ *Quasi ovunque nullo* è da intendersi in senso matematico.

la tecnica standard: si risolve l'equazione nelle regioni con potenziale costante e poi si impongono opportune condizioni di raccordo.

L'equazione algebrica associata è:

$$y^2 = -\beta^2 \rightarrow y = \pm i\sqrt{\beta^2} = \pm i\sqrt{-\frac{2mE}{\hbar}} = \mp\sqrt{\frac{2mE}{\hbar}} = \mp\beta$$

e quindi la generica autofunzione è:

$$Ae^{-\beta x} + Be^{\beta x}$$

Nella regione a sinistra della buca (per $x < 0$) si deve porre $A = 0$, perché l'esponenziale negativo esplose per $x \rightarrow -\infty$ (quindi non è a quadrato integrabile) e pertanto non rappresenta uno stato fisico. Analogamente, a destra si deve imporre $B = 0$.

L'altra costante ha il ruolo di costante di normalizzazione, che per ora può essere posta per comodità ad 1, rinunciando quindi a normalizzare. Riassumendo, le soluzioni trovate sono:

$$\begin{cases} e^{\beta x} & \text{per } x < 0 \\ e^{-\beta x} & \text{per } x > 0 \end{cases}$$

si devono ora raccordare queste due funzioni in 0 imponendo la condizione sulle derivate (seconda condizione al contorno):

$$\begin{aligned} \frac{d}{dx} e^{-\beta x} \Big|_{+\epsilon} - \frac{d}{dx} e^{\beta x} \Big|_{-\epsilon} &= -2\eta e^{-\beta|x|} \Big|_{x=0} \rightarrow \\ -\beta e^{-\beta x} \Big|_{+\epsilon} - \beta e^{\beta x} \Big|_{-\epsilon} &= -2\eta e^{-\beta|x|} \Big|_{x=0} \end{aligned}$$

da cui evidentemente discende $\beta = \eta$. La condizione di raccordo sul salto della derivata permette quindi l'esistenza di un solo stato legato, dato da:

$$\psi(x) = Ae^{-\eta|x|} \quad \text{Stato legato per } E < 0$$

Energia positiva

Per energie positive non esistono stati legati. Infatti in questo caso l'equazione da risolvere è:

$$\frac{d^2}{dx^2} \psi(x) + \alpha^2 \psi(x) = 0 \quad \left(\alpha^2 \equiv \frac{2mE}{\hbar} \right)$$

e quindi l'equazione algebrica associata risulta:

$$y^2 = -\alpha^2 \rightarrow y = \pm\sqrt{\alpha} = \pm i\sqrt{\frac{2mE}{\hbar}} = \pm i\alpha$$

che ammette per soluzioni degli esponenziali complessi, che su intervalli aperti a infinito non sono sommabili. Riassumendo, per energie positive non esistono stati legati.

4.2.3 Spettro del modello di Kronig e Penney

Si riconsideri ora il modello completo con le buche periodiche e si concentri l'attenzione sulla prima *zona di Brillouin*. Si possono quindi cercare le soluzioni dell'equazione di Schrödinger indipendente dal tempo, usando le due condizioni al contorno:

1. $\psi(0) = e^{ikd}\psi(d)$ che viene dal teorema di Bloch;
2. $\frac{d}{dx} \psi(x) \Big|_{x=0} - \frac{d}{dx} \psi(x) \Big|_{x=d} e^{-ikd} = -2\eta\psi(0)$ che viene dall'integrazione nella buca

Come già accennato, si considera l'equazione di Schrödinger al di fuori delle buche perché la larghezza delle buche tende a zero. Inoltre ci si limita qui ad esplicitare il caso $E > 0$, poiché l'altro caso può essere ottenuto da questo semplicemente sostituendo alle funzioni trigonometriche le rispettive funzioni iperboliche.

L'equazione da risolvere è quindi:

$$\frac{d^2}{dx^2}\psi(x) + \alpha^2\psi(x) = 0 \quad \alpha^2 = \frac{2mE}{\hbar^2}$$

di cui l'algebra associata è:

$$y^2 = -\alpha^2 \quad \rightarrow \quad y = \pm i\alpha$$

che ammette soluzioni del tipo $\psi(x) = A'e^{i\alpha x} + B'e^{-i\alpha x}$ o, in forma trigonometrica, $\psi(x) = A \sin(\alpha x) + B \cos(\alpha x)$. Una generica autofunzione è quindi:

$$\psi(x) = A \sin(\alpha x) + B \sin \alpha(d - x)$$

La prima condizione al contorno diviene:

$$\psi(0) = e^{ikd}\psi(d) \quad \rightarrow \quad B \sin \alpha d = e^{-ikd} A \sin \alpha d$$

da cui la forma generica dell'autofunzione soluzione:

$$\psi(x) = B [\sin \alpha(d - x) + e^{ikd} \sin(\alpha x)]$$

Per impostare ora la seconda condizione al contorno conviene prima calcolare la derivata della generica autofunzione:

$$\frac{d\psi(x)}{dx} = B [-\alpha \cos \alpha(d - x) + e^{ikd} \alpha \cos(\alpha x)]$$

ed imponendo la seconda condizione, che è:

$$\left. \frac{d\psi(x)}{dx} \right|_0 - \left. \frac{d\psi(x)}{dx} \right|_d e^{-ikd} = -2\eta\psi(0)$$

si ricava:

$$B [-\alpha \cos(\alpha d) + e^{ikd} \alpha] - e^{-ikd} B [-\alpha + \alpha e^{ikd} \cos(\alpha d)] = -2\eta B \sin(\alpha d)$$

per semplificare le cose si può ancora una volta porre $B = 1$ rinunciando così a normalizzare la funzione.⁷ Questa condizione può essere espressa in una forma più semplice:

$$\begin{aligned} -\alpha \cos(\alpha d) + \alpha e^{ikd} - e^{-ikd} [-\alpha + \alpha e^{ikd} \cos(\alpha d)] &= -2\eta \sin(\alpha d) \\ -\alpha \cos(\alpha d) + \alpha e^{ikd} + \alpha e^{-ikd} - \alpha e^{-ikd} e^{ikd} \cos(\alpha d) &= -2\eta \sin(\alpha d) \\ -\alpha \cos(\alpha d) + \alpha (e^{ikd} + e^{-ikd}) - \alpha \cos(\alpha d) &= -2\eta \sin(\alpha d) \\ -2\alpha \cos(\alpha d) + \alpha (\cos kd + i \sin kd + \cos kd - i \sin kd) &= -2\eta \sin(\alpha d) \\ -2\alpha \cos(\alpha d) + 2\alpha \cos kd &= -2\eta \sin(\alpha d) \end{aligned}$$

ed eliminando il fattore 2α in comune:

$$\cos kd - \cos \alpha d = -\frac{\eta}{\alpha} \sin(\alpha d) \quad \rightarrow \quad \cos \alpha d - \frac{\eta}{\alpha} \sin(\alpha d) = \cos kd$$

Ricordando ora che $\alpha = \frac{\sqrt{2mE}}{\hbar}$, si ottiene in definitiva (*relazione di dispersione per energie positive*):⁸

$$\boxed{\cos \left(\frac{\sqrt{2mE}}{\hbar} d \right) - \frac{\eta \hbar}{\sqrt{2mE}} \sin \left(\frac{\sqrt{2mE}}{\hbar} d \right) = \cos kd}$$

⁷ B ha infatti così il ruolo della costante di normalizzazione.

⁸Si ricordi per inciso che η risulta legato alla forza della buca e vale $\frac{mV_0 l}{\hbar^2}$.

Si può dire quindi che l'imporre la seconda condizione al contorno si traduce in una relazione tra l'energia e il parametro k , nella forma:

$$\mathcal{F}(E) = \cos kd$$

In questa funzione è contenuta la struttura a bande.

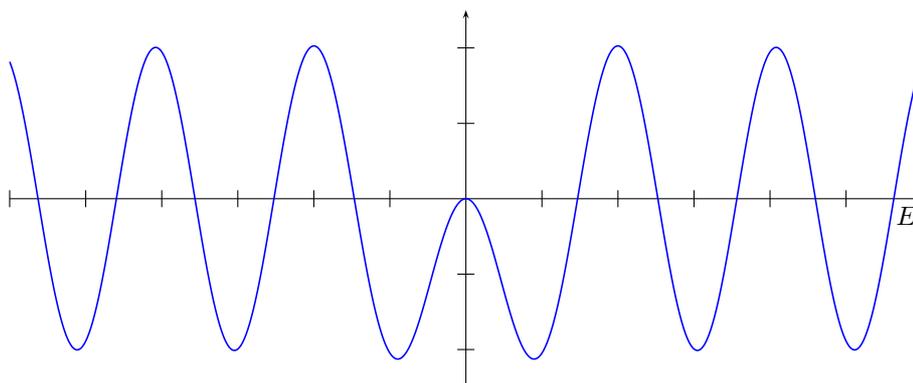
Questo è il caso in cui $E > 0$. Nel caso invece di $E < 0$, la relazione tra E e k cui conduce la seconda condizione al contorno è (basta passare alle funzioni iperboliche):

$$\cosh \beta d - \frac{\eta}{\beta} \sinh(\beta d) = \cos kd$$

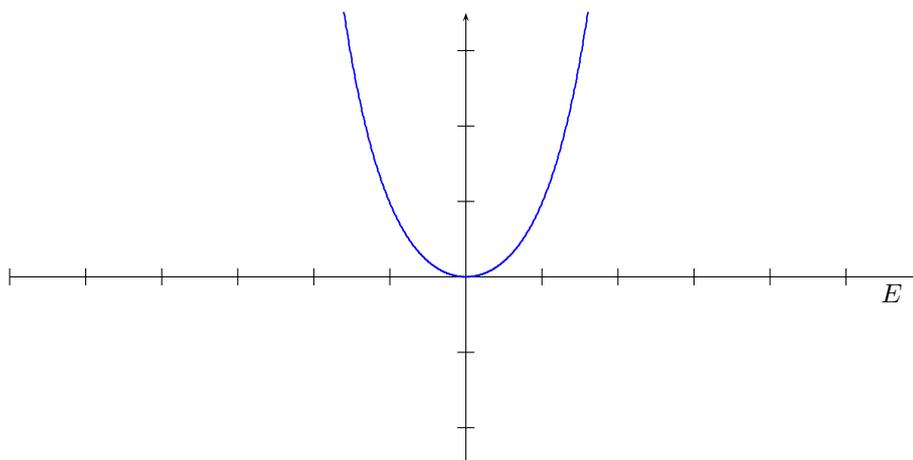
dove, si ricordi, vale $\beta = i\frac{\sqrt{2mE}}{\hbar}$. La relazione di dispersione per energie negative è dunque:

$$\cosh \left(i\frac{\sqrt{2mE}}{\hbar} d \right) - \frac{\eta\hbar}{\sqrt{2mE}} \sinh \left(i\frac{\sqrt{2mE}}{\hbar} d \right) = \cos kd$$

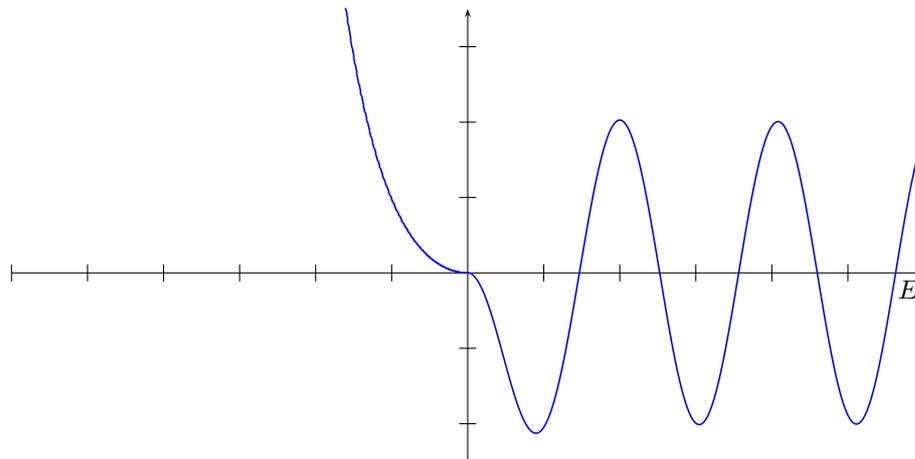
Siccome sono funzioni alquanto complicate, conviene dare dapprima solo una descrizione grafica qualitativa di queste due equazioni. In particolare, ecco un grafico del primo membro della prima equazione in funzione dell'energia.



e del primo membro della seconda equazione:



dunque l'andamento corretto totale è dato dalla parte destra del primo grafico e dalla parte sinistra del secondo:



Poiché le due equazioni consistono nell'eguagliare ad un coseno questi primi membri, si può concludere senz'altro che esistono soluzioni solo quando questi primi membri, nelle loro rispettive zone di competenza, siano compresi tra -1 e 1.

4.2.4 Studio dello spettro (bande e gap)

Nel paragrafo precedente, si sono ottenute finalmente le soluzioni dell'equazione di Schrödinger, cioè le autofunzioni dell'Hamiltoniana per il sistema "particella in potenziale periodico formato da buche a delta".

Questa equazione differenziale del secondo ordine ha due condizioni al contorno. Una ha permesso di determinare una costante arbitraria, mentre l'altra si è tradotta in due relazioni – a seconda che l'energia sia positiva o negativa – tra l'energia e il parametro k . Si noti che questo parametro deriva direttamente dal teorema di Bloch (§4.1) e che si può considerare come il numero d'onda delle autofunzioni del sistema.

Come visto, queste relazioni sono:

$$\begin{array}{l} \cos(\alpha d) - \frac{\eta}{\alpha} \sin(\alpha d) = \cos kd \quad E > 0 \\ \cosh(\beta d) - \frac{\eta}{\beta} \sinh(\beta d) = \cos kd \quad E < 0 \end{array}$$

dove:

$$\alpha = \frac{\sqrt{2mE}}{\hbar} \quad \beta = i \frac{\sqrt{2mE}}{\hbar}$$

Le autofunzioni che soddisfano queste condizioni sono le uniche che possono essere prese in considerazione, perché derivano dall'imposizione delle opportune condizioni di raccordo. In particolare, preso un certo valore dell'energia, ad esso sarà associato una determinata autofunzione caratterizzata a sua volta da un certo valore del parametro k . Solo se questi due valori di E e k soddisfano l'equazione (la prima o la seconda a seconda del segno di E) la funzione può essere accettata come un'autostato fisicamente permesso.

Queste due equazioni sono trascendenti e quindi di difficile soluzione. Si adotterà dunque un approccio grafico per uno studio qualitativo.

Mettendo in un grafico in funzione dell'energia i due membri di sinistra, e ricordando che il primo grafico va considerato solo per la parte destra ($E > 0$) e il secondo solo per quella sinistra ($E < 0$), si vede che dovendo eguagliare un coseno gli unici valore permessi di k sono quelli per cui i membri di sinistra sono compresi tra -1 e 1.

Si vede che esistono degli intervalli di energia in cui il grafico va fuori dell'intervallo in ordinata $[-1, 1]$,⁹ alternati ad intervalli in cui è compreso. Questo in definitiva significa che ci sono intervalli di energie permesse detti *bande* e intervalli di energie proibite detti *gap*.

⁹Anche se nella figura seguente non è molto evidente, le prime punte della parte destra del grafico sono superiori a 1, quindi generano dei *gap* di energia. Questo risulta più evidente nel grafico riportato più avanti, per $\eta = 2.5$.

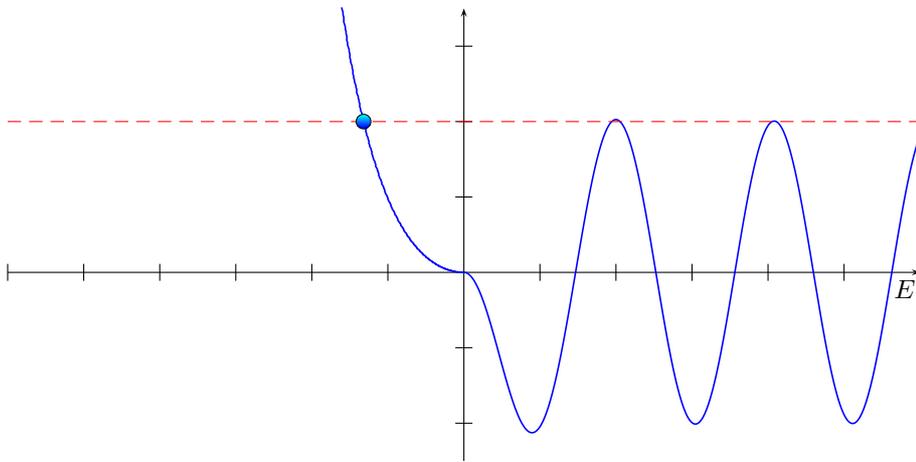
Per energie molto alte la quantità η/α tende a zero, il membro di sinistra tende a ridursi al solo coseno e non ci saranno quindi energie proibite. In realtà questo è un comportamento limite, infatti le gap sono sempre presenti anche ad alte energie, anche se vanno restringendosi sempre più.

Connessione tra le bande e gli stati legati del sistema con potenziale periodico la cui ripetizione dà origine al reticolo.

Si è visto come nel caso di una sola buca a delta esista un solo stato legato. È facile istituire una relazione tra la prima banda e questo stato legato della delta. Questo lo si può vedere facendo tendere d all'infinito.

Per fare questo, si individuano i valori degli estremi della banda, cioè la minima e la massima energia permesse per la prima banda. Risulta più comodo ricavare i valori di α e β che contengono a loro volta l'energia.

Si comincerà con il ricavare l'estremo inferiore β_L della prima banda di energia. Lo si denota con β perché cade nella regione ad energie negative. Osservando il grafico del membro di destra:



il minimo valore permesso per β si ha quando tale membro vale 1. La condizione risulta dunque:

$$\cosh(\beta d) - \frac{\eta}{\beta} \sinh(\beta d) = 1 \rightarrow \cosh(\beta d) - 1 = \frac{\eta}{\beta} \sinh(\beta d) \rightarrow \frac{\beta}{\eta_L} = \frac{\sinh(\beta d)}{\cosh(\beta d) - 1}$$

dalle formule di bisezione delle funzioni iperboliche,¹⁰ si nota che il membro di destra è una cotangente iperbolica. Moltiplicando inoltre numeratore e denominatore del membro di destra per d (in modo da avere la stessa quantità che compare come argomento della cotangente iperbolica) si ottiene infine la seguente equazione trascendente per il limite inferiore:

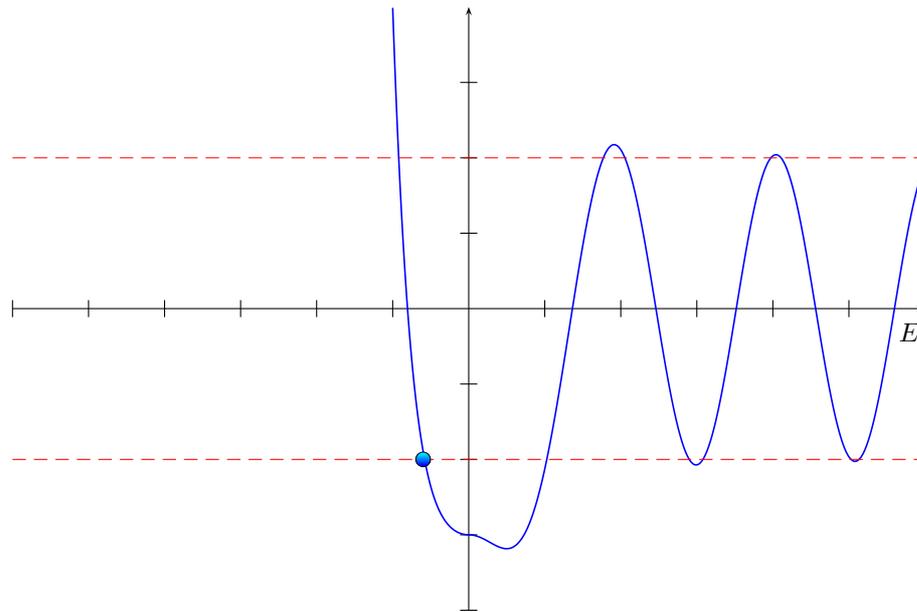
$$\boxed{\coth \frac{1}{2} \beta_L d = \frac{\beta_L d}{\eta d}}$$

Si ricaverà ora l'estremo superiore della prima banda, che sarà denotato α_U o β_U a seconda che ricada nella zona ad energie positive o negative. Occorre infatti rilevare che il valore di η influenza pesantemente la struttura del grafico. I grafici mostrati in precedenza sono in effetti a $\eta = 1$, ecco invece un esempio con $\eta = 2.5$:

¹⁰La formula di bisezione della tangente iperbolica è:

$$\tanh x = \frac{\cosh(2x) - 1}{\sinh(2x)}$$

da cui la cotangente iperbolica ($\coth x = 1/\tanh x$).



e quindi in questo caso (si veda punto blu) il limite superiore della prima banda di energia si trova ad energie negative. Si devono quindi distinguere due casi a seconda che questo limite superiore cada ad energie negative o positive. In ogni caso la condizione da imporre sarà che il primo membro dell'equazione sia uguale a -1.

Ad energie negative, la condizione da imporre è:

$$\cosh(\beta_U d) - \frac{\eta}{\beta_U} \sinh(\beta_U d) = -1 \rightarrow \cosh(\beta_U d) + 1 = \frac{\eta}{\beta_U} \sinh(\beta_U d) \rightarrow \frac{\beta_U}{\eta_L} = \frac{\sinh(\beta_U d)}{\cosh(\beta_U d) + 1}$$

e sempre dalle formule di bisezione per le funzioni iperboliche si ha la seguente equazione trascendente per trovare il limite superiore:

$$\tanh \frac{1}{2} \beta_U d = \frac{\beta_U d}{\eta d}$$

Nel caso invece in cui il limite cada ad energie positive la condizione da imporre è:

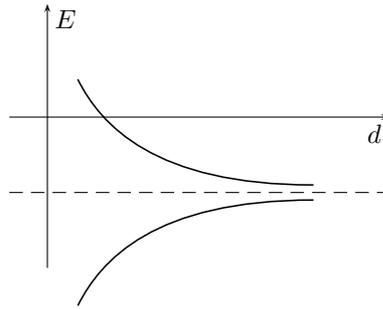
$$\cos(\alpha_U d) - \frac{\eta}{\alpha_U} \sin(\alpha_U d) = -1 \rightarrow \frac{\alpha_U d}{\eta d} = \frac{\sin(\alpha_U d)}{\cos(\alpha_U d) + 1}$$

e quindi in questo caso, ancora utilizzando le formule di bisezione, l'equazione trascendente da risolvere è:

$$\tan \frac{1}{2} \alpha_U d = \frac{\alpha_U d}{\eta d}$$

Da notare che β deve essere negativo perché possa esistere uno stato legato, perché la tangente iperbolica – o la cotangente iperbolica – per $d \rightarrow \infty$ vanno a 1. Quindi significa che per d tendente a infinito risulta $\eta = \beta$. Ma questo restituisce proprio lo stato legato trovato per il potenziale con singola buca a delta.

Ora, per trovare i due valori estremi di questa prima banda di energia si dovrebbero risolvere altrettante equazioni trascendenti. In questo contesto si farà invece affidamento sul risultato di Wannier, che dice che il loro grafico è del tipo:



Questo significa che la banda di energia permessa va restringendosi man mano che il reticolo diventa meno fitto. La cosa importante da sottolineare è che nel limite per $d \rightarrow \infty$ i due estremi della banda tendono a coincidere asintoticamente tra loro e con l'unico stato legato del modello con un'unica buca a delta. Questo fa pensare che anche per le altre bande ci sia questo comportamento, cioè che esista un legame tra bande di energia di un reticolo e stati legati del sistema base che costituisce il reticolo stesso (ad esempio gli atomi). Si ricordi che per le molecole accadeva qualcosa di simile, si veda a questo proposito la discussione sullo stato legante e stato antilegante della molecola di idrogeno in §3.3.

Si è quindi concluso che, nel modello di Kronig e Penney, allo stato legato del sistema a buca singola corrisponde la prima banda del sistema con buche multiple. Si può anzi immaginare la cosa pensando il sistema con singola buca come un sistema con buche periodiche a distanza infinita tra di esse. Tale sistema ha un solo stato legato, che, facendo avvicinare le buche tra loro, diventa una banda che man mano si allarga. Potrebbe allora nascere il dubbio che esista una sola banda di energia permessa nel sistema con potenziale periodico, visto che il sistema con singola buca ha un solo stato legato.

Si vuole invece dimostrare che per il sistema con potenziale periodico ci sono infinite bande, cioè che man mano che si sale in energia esiste sempre un susseguirsi di bande permesse e gap proibite. Quello che succede è che le gap si restringono sempre più.

Volendo, si può dire che nel processo di “avvicinamento delle buche” le prime bande si originano dagli stati legati del sistema singolo, ma ci sono bande che si originano dalla parte di spettro continuo. In altri termini, un potenziale periodico fa sentire i suoi effetti (rendendo proibiti alcuni intervalli di energia) a qualunque zona di energia, anche molto alta.

Dimostrazione

Per dimostrare la presenza delle bande a qualunque energia si ricorre ad un modello perturbativo. Si parta da una particella libera, imponendo però che le sue autofunzioni siano periodiche con lo stesso periodo del potenziale periodico, in modo da soddisfare il teorema di Bloch. Questo sistema si definisce *reticolo vuoto*. A questo sistema si applicherà poi il potenziale periodico come una perturbazione. Un modo comodo di imporre questa condizione di periodicità di Bloch sulle autofunzioni è imporre le cosiddette *condizioni al contorno di Bohr–von Kármán*, che consistono nel dire che la catena periodica di atomi è chiusa ad anello.¹¹ Questa condizione impone automaticamente i valori di k .

A proposito dei valori di k , si noti che quando si è studiata la relazione tra l'energia e k , si era discusso solo della monotonia e quasi linearità di questa relazione, ma non si è indagato sui possibili valori di k . D'altra parte k ha il ruolo di un numero quantico e quindi, almeno per gli stati legati, discreti, è plausibile che assuma solo alcuni valori discreti. La questione è che non ci si può pronunciare sui valori possibili di k fino a quando non si introducono delle condizioni al contorno, come appunto quelle di Bohr–von Kármán.

Posto che la catena sia formata da N atomi, con passo a , queste condizioni al contorno consistono nell'imporre che il valore della funzione d'onda del sistema assuma lo stesso valore nella posizione Na e nell'origine:

$$\psi(Na) = \psi(0)$$

¹¹Ovvero, che la funzione d'onda (e gli spostamenti del reticolo) del primo atomo della catena sono coincidenti con quelli dell'ultimo.

Il teorema di Bloch prevede che $\psi(Na) = e^{ikNa}\psi(0)$ e mettendo insieme le due relazioni si ha che:

$$e^{ikNa}\psi(0) = \psi(0) \quad \Rightarrow \quad e^{ikNa} = 1$$

La condizione sull'esponenziale, una volta scritta in termini di funzioni trigonometriche fornisce la relazione:

$$\cos(kNa) + i \sin(kNa) = 1$$

ovvero:

$$k = \frac{2\pi q}{Na} \quad 0 \leq q \leq N-1$$

Questo significa appunto che *il numero di valori possibili per k è pari al numero di atomi nel reticolo*. Se si considera la zona di Brillouin da $-\pi/a$ a π/a , allora q è un intero che va da $-N/2$ a $N/2$, estremi compresi o esclusi a seconda se N sia pari o dispari. Infatti se q deve essere intero e N è dispari, q non può allora uguagliare $N/2$. Nell'ipotesi che il cristallo sia finito, lavorando cioè con un numero finito di atomi, si può normalizzare a 1 su tutto il cristallo.¹²

Si supponga di avere la seguente funzione di Bloch:

$$b(x; k) = e^{ikx}u_k(x)$$

e si supponga di aver normalizzato la $u_k(x)$ su un passo reticolare:

$$\int_0^a u_k^2(x) dx = 1$$

In tal caso la funzione completa risulta normalizzata se si moltiplica per la costante di normalizzazione $1/\sqrt{Na}$. Infatti:

$$\|e^{ikx}u_k(x)\| = \int_0^{Na} (e^{ikx}u_k(x))^* e^{ikx}u_k(x) dx = \int_0^{Na} e^{-ikx}e^{ikx}u_k^2(x) dx = \int_0^{Na} u_k^2(x) dx = Na$$

La funzione normalizzata è quindi:

$$b(x; k) = \frac{1}{\sqrt{Na}} e^{ikx}u_k(x)$$

Le autofunzioni del reticolo vuoto sono pertanto le onde piane cui però si deve imporre la proprietà di periodicità. Le onde piane "pure" sono le funzioni d'onda:

$$\psi_k(x) = b(x; k) = e^{ikx}$$

con $k \in \mathbb{R}$. Bisogna imporre su queste la condizione su k che provengono dalle condizioni al contorno di Bohr-von Kármán, ovvero:

$$k = q \frac{2\pi}{Na} \quad q \in \mathbb{N}_0$$

Inoltre è comodo lavorare con k appartenente solo alla prima zona di Brillouin: si impone quindi che q sia compreso tra $-N/2$ e $N/2$ e si somma a k un numero intero di volte la costante $2\pi/a$, ottenendo quindi la seguente forma per le autofunzioni del reticolo vuoto:

$$\psi_k(x) = e^{i(k+n\frac{2\pi}{a})x}$$

Si consideri la relazione di dispersione del sistema: trattandosi di una particella libera, la relazione di dispersione è parabolica. Si suddivida l'asse dei k in intervalli di π/a e si spostino le parti di parabola fuori della prima zona di Brillouin di $n\frac{2\pi}{a}$, in modo da far rientrare tutto nella prima zona di Brillouin (si veda figura a pag. 204).

Trovata l'autobase del sistema imperturbato, si può introdurre la perturbazione periodica:

$$V(x) = \sum_{s=-\infty}^{\infty} V_s e^{is\frac{2\pi}{a}x}$$

¹²Avendo a che fare con un cristallo infinito si deve normalizzare invece su tutto l'asse reale, e quindi bisogna normalizzare a delta perché si tratta di stati del continuo.

4.3 Vibrazioni elastiche dei reticoli

In questa trattazione si introdurranno dapprima i modi normali in un contesto classico. In seguito, attraverso la quantizzazione dei modi normali, si introdurrà il concetto di *fononi*, eccitazioni collettive dei solidi in termini dei quali si può calcolare il calore specifico degli isolanti (oggetto del paragrafo §5.3).

In prima istanza, si può utilizzare per i solidi lo stesso approccio usato per le molecole. Si supponga dunque di aver trovato le energie elettroniche nell'approssimazione a nuclei fissi: si può allora applicare il metodo di approssimazione di Born-Oppenheimer, cioè considerare quest'energia degli elettroni come il campo medio in cui si muovono i nuclei.

Il potenziale periodico di cui si è discusso nel modello di Kronig e Penney (§4.2) è una possibile schematizzazione di questo potenziale medio che include gli effetti degli altri elettroni e l'effetto dei nuclei. Si può quindi utilizzare questa approssimazione per studiare lo spettro di energia totale degli elettroni come funzione delle posizioni di equilibrio dei nuclei. Successivamente, si può sviluppare in serie di potenze attorno alle posizioni di equilibrio dei nuclei, ottenendo dei termini quadratici.

Le vibrazioni dei reticoli cristallini possono quindi essere trattate in approssimazione armonica, cioè supponendo che l'energia elettronica, corrispondente in questo modello all'energia potenziale che descrive il moto dei nuclei, sia sviluppabile al secondo ordine e che si possa quindi descrivere come un'interazione di tipo elastico. Si studiano quindi i moti degli ioni (o nuclei) del solido come piccole oscillazioni attorno alle posizioni di equilibrio (*siti del reticolo*).

4.3.1 Modello Classico

Si consideri dunque per semplicità il caso unidimensionale e si prenda il modello in cui un singolo elettrone si muove nel potenziale generato dai nuclei fissi e dagli altri elettroni. Questo potenziale sarà approssimato con un potenziale armonico, questo corrisponde a prendere una Lagrangiana o una Hamiltoniana che descrivano le forze elastiche di oscillatore armonico.

Il modello per descrivere le vibrazioni di una catena infinita di atomi è quello della seguente lagrangiana:

$$L = \frac{M}{2} \sum_{s=1}^{\infty} \dot{u}_s^2 - \frac{1}{2} \sum_{s=1}^{+\infty} \sum_{p=-\infty}^{+\infty} c_p (u_{s+p} - u_s)^2$$

Si potrebbe certamente considerare una catena finita con N siti e, utilizzando delle condizioni al contorno periodiche, imporre che l'atomo in prima posizione abbia lo stesso spostamento dell'atomo nella posizione N (le già viste condizioni al contorno di Born-Von Kármán), ma questo è un dettaglio.

Le u_s rappresentano gli spostamenti dalla posizione di equilibrio degli atomi della catena. Il primo termine è l'energia cinetica associata a questi spostamenti, il secondo termine descrive invece le interazioni elastiche tra un atomo in un sito e un atomo che si trova in un altro sito. Notare che ci possono essere dei termini di accoppiamento delle forze elastiche fra un atomo di un sito e un atomo di un qualsiasi altro sito.

Nel seguito si farà l'approssimazione semplificatrice di ammettere che le uniche forze elastiche esercitate riguardino i primi vicini, mentre in linea di principio possono riguardare qualsiasi atomo della catena. Il secondo termine, dunque, rappresenta l'interazione elastica tra gli atomi del sito s e del sito $s + p$ e con costante elastica c_p , dipendente solo da p , ovvero dal numero di siti tra i due atomi in questione.

Questa Lagrangiana è quindi invariante per traslazione: se il sistema compie una traslazione pari a un multiplo del passo del reticolo, questa resta evidentemente immutata.

Prima di procedere ulteriormente si noti che i coefficienti c_p soddisfano le condizioni:

$$\begin{aligned} c_{-p} &= c_p && \text{terzo principio della dinamica} \\ c_0 &= 0 && \text{non interazione con sé stesso} \end{aligned}$$

A partire da questa lagrangiana, si ricavano le equazioni del moto:¹³

$$\left\{ \begin{array}{l} M \frac{d^2 u_0}{dt^2} = \sum_{p=-\infty}^{+\infty} c_p (u_p - u_0) \\ \vdots \\ M \frac{d^2 u_s}{dt^2} = \sum_{p=-\infty}^{+\infty} c_p (u_{s+p} - u_s) \\ \vdots \end{array} \right.$$

Si tratta di un sistema di infinite equazioni accoppiate, una per ogni sito del reticolo individuato dall'indice s . Le equazioni sono accoppiate perché, naturalmente, nel potenziale compaiono oltre a u_s anche tutte le altre u_{s+p} . Si tratta quindi di tante equazioni che accoppiano lo spostamento u_s dell'atomo nel sito s con gli spostamenti degli altri atomi della catena (u_{s+p}).

Si vogliono quindi ricavare quelle soluzioni che descrivono moti di oscillazione con una determinata frequenza (i *modi normali*), si suppone quindi che le soluzioni abbiano questa forma armonica:

$$u_s = a_s e^{-i\omega t}$$

Questa condizione trasforma le equazioni differenziali in equazioni armoniche.

Si noti che il moto di un atomo, pur supponendo che si tratti sempre di moto armonico, può avere ampiezza (e in principio, frequenza) diversa da quella di tutti gli altri. Per questo compare un pedice all'ampiezza a_s : è come dire che l'ampiezza dipende dalla posizione, anche se questa variabile è discreta essendo in pratica la posizione lungo la catena.

In questo senso, si può fare l'ipotesi che valga il teorema di Bloch per le ampiezze a_s , ovvero che spostandosi di un certo numero di siti¹⁴ l'ampiezza dell'oscillazione cambi solo per un fattore esponenziale immaginario, dipendente da un certo parametro k e dallo spostamento $s \cdot a$:

$$a_s = a_0 e^{iksa}$$

Imporre questa condizione permette di risolvere in maniera immediata il sistema di equazioni accoppiate. Questa ipotesi fa infatti scomparire la dipendenza da s e quindi le equazioni di Newton diventano la stessa per ogni atomo del reticolo. Infatti sostituendo si ha:

$$\begin{aligned} M \frac{d^2}{dt^2} (a_0 e^{iksa} e^{i\omega t}) &= \sum_{p=-\infty}^{+\infty} c_p (a_0 e^{ik(s+p)a} e^{i\omega t} - a_0 e^{iksa} e^{i\omega t}) \\ -\omega^2 M a_0 e^{iksa} e^{i\omega t} &= \sum_{p=-\infty}^{+\infty} c_p (a_0 e^{iksa} e^{ikpa} e^{i\omega t} - a_0 e^{iksa} e^{i\omega t}) \\ -\omega^2 M a_0 e^{iksa} e^{i\omega t} &= a_0 e^{iksa} e^{i\omega t} \sum_{p=-\infty}^{+\infty} c_p (e^{ikpa} - 1) \end{aligned}$$

e evidentemente scompare la dipendenza da s .

Si è dunque dimostrato che una legge del moto che soddisfi contemporaneamente l'ipotesi di oscillazione armonica e il teorema di Bloch riguardo alle ampiezze, soddisfa anche le equazioni di Newton. Trattandosi inoltre di equazioni differenziali, il teorema di esistenza e unicità garantisce che la soluzione sia unica e che quindi questa soluzione in particolare rappresenti effettivamente i moti degli atomi.

Come si può vedere, le equazioni di Newton assumono la forma di relazioni tra k e ω . Dividendo per M e cambiando di segno si ottiene:

$$\omega^2 = \frac{1}{M} \sum_{p=-\infty}^{+\infty} c_p (e^{ikpa} - 1)$$

¹³Tramite applicazione dell'equazione di Eulero-Lagrange, naturalmente.

¹⁴Ovvero di un tratto $s \cdot a$, con s un intero e a passo del reticolo.

È da notare che l'ipotesi di validità del teorema di Bloch per le ampiezze è fondamentale, in quanto è questa che fa sì che le equazioni di Newton si disaccoppino, in altri termini *l'ipotesi di validità del teorema di Bloch permette di diagonalizzare la matrice dei coefficienti*.

4.3.2 Calcolo della relazione di dispersione

Il fatto che gli spostamenti siano delle oscillazioni armoniche da una parte e che valga il teorema di Bloch per le ampiezze dall'altra, comporta che la forma degli spostamenti u_s sia:

$$u_s(t) = a_0 e^{iks_a} e^{-i\omega t} = a_0 e^{i(ks_a - \omega t)}$$

Questo fa pensare che le vibrazioni degli atomi del reticolo possono essere viste complessivamente come un'onda elastica che si propaga nel reticolo stesso con numero d'onda k e pulsazione ω .

Nella teoria ondulatoria si studia la relazione di dispersione, che è proprio la relazione $\omega(k)$ tra la frequenza e il numero d'onda. In particolare la derivata della relazione di dispersione fornisce la velocità di propagazione (velocità di gruppo) dell'onda:¹⁵

$$v_g = \frac{d\omega}{dk}$$

In meccanica quantistica, a causa delle analogie con la teoria ondulatoria, la relazione di dispersione è la relazione tra il numero d'onda e il corrispondente autovalore dell'energia di una certa autofunzione d'onda dell'Hamiltoniana. Sebbene si stia utilizzando un formalismo classico, è possibile fare un paragone con il caso quantistico. Siccome la frequenza ω è legata all'energia del sistema, una relazione tra ω e k è in effetti analoga alla relazione tra E e k quantistica. Si tratta appunto di una *relazione di dispersione*.

Si procederà quindi ora alla ricerca di questa relazione.

Per semplificare, la somma si può spezzare in tre parti:

$$\omega^2 = \frac{1}{M} \left[\sum_{p<0} c_p (1 - e^{ikpa}) + c_0 (1 - e^{ik0a}) + \sum_{p>0} c_p (1 - e^{ikpa}) \right]$$

utilizzando ora il fatto che $c_0 = 0$ ⁽¹⁶⁾ e che $c_{-p} = c_p$:

$$\begin{aligned} \omega^2 &= \frac{1}{M} \left[\sum_{p>0} c_p (1 - e^{-ikpa}) + \sum_{p>0} c_p (1 - e^{ikpa}) \right] = \\ &= \frac{1}{M} \left[\sum_{p>0} (c_p - c_p e^{-ikpa}) + \sum_{p>0} (c_p - c_p e^{ikpa}) \right] = \\ &= \frac{1}{M} \sum_{p>0} (2c_p - c_p e^{-ikpa} - c_p e^{ikpa}) = \\ &= \frac{1}{M} \sum_{p>0} c_p (2 - e^{-ikpa} - e^{ikpa}) = \\ &= \frac{1}{M} \sum_{p>0} c_p (2 - 2 \cos(kpa)) = \\ &= \frac{2}{M} \sum_{p>0} c_p (1 - \cos(kpa)) \end{aligned}$$

Si impone ora l'ipotesi che k si trovi solo nella prima zona di Brillouin, ovvero che $-\pi/a \leq k \leq \pi/a$. Questo è ragionevole, infatti due valori di k che differiscono di $2\pi/a$ forniscono la stessa frequenza.

Si è dunque ricavata la banda per le frequenze di vibrazione del reticolo, analoga alla banda di energie per degli stati elettronici. Notare che qui esiste una sola banda: se k varia tra $-\pi/a$ e π/a si

¹⁵Questa affermazione si dimostra tramite metodo della fase stazionaria.

¹⁶In ogni caso il termine corrispondente è zero, perché il termine fra parentesi fornisce $(1 - 1) = 0$.

ottiene un'unica funzione periodica di periodo appunto $2\pi/a$, a differenza di quello che accade agli stati elettronici, dove il numero delle bande è infinitamente grande.¹⁷

Per semplificare i calcoli per ottenere la relazione $\omega(k)$ cercata, un'approssimazione abituale consiste nel supporre che ci sia interazione solo tra i primi vicini, che è un'approssimazione dello stesso tipo adoperato per le molecole.

Quindi:

$$\omega^2 = \frac{2}{M} c_1 (1 - \cos(ka))$$

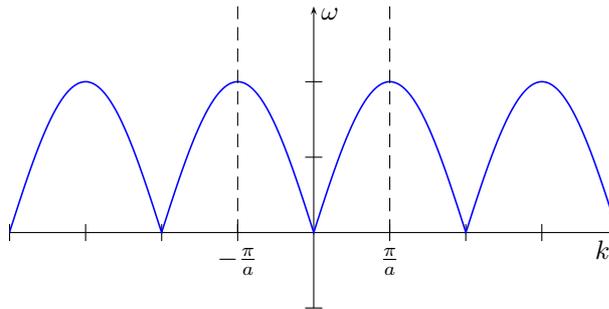
e dalle formule di bisezione:¹⁸

$$\omega^2 = \frac{2}{M} c_1 2 \sin^2\left(\frac{ka}{2}\right) = \frac{4c_1}{M} \sin^2\left(\frac{ka}{2}\right)$$

ed estraendo la radice quadrata, si trova la forma definitiva per la relazione di dispersione:

$$\omega = 2\sqrt{\frac{c_1}{M}} \left| \sin\left(\frac{ka}{2}\right) \right|$$

il modulo deriva dal fatto che le frequenze sono per definizione positive. Il grafico di questa funzione ha l'andamento mostrato in figura.



Come è noto, la velocità di un'onda è data dalla sua velocità di gruppo. Si calcolerà quindi questa velocità ai bordi ed al centro della zona di Brillouin.

Caso $k \rightarrow \pm \frac{\pi}{a}$

Per questi valori di k la relazione di dispersione $\omega(k)$ ha un massimo e quindi risulta evidentemente:

$$\left. \frac{d\omega}{dk} \right|_{k=\pm \frac{\pi}{a}} = 2 \frac{c_1}{M} \left| \cos\left(\frac{ka}{2}\right) \right| \Big|_{k=\pm \frac{\pi}{a}} = 0$$

se si interpreta quindi la forma della $u_s(t) = a_0 e^{i(ksa - \omega t)}$ come un'onda elastica propagantesi nel solido, questa relazione dice che la velocità di gruppo di quest'onda è zero: dunque *per* $k = \pm \frac{\pi}{a}$ l'onda è un'onda stazionaria.

Teoria di Bragg

Alla stessa conclusione si può arrivare anche col seguente ragionamento. Si ammetta che vaga la condizione di Bragg:

$$2d \sin \theta = n\lambda$$

¹⁷La banda è unica perché c'è solo un intervallo di valori di ω . Infatti, anche se si considerassero altri valori di k al di fuori dell'intervallo $]-\pi/a, \pi/a[$, il coseno ripeterebbe gli stessi valori, ripetendo quindi i valori di ω . La differenza con lo spettro degli elettroni è il fatto che la funzione $\omega(k)$ in questo caso non è monotona, ma limitata per via del coseno.

¹⁸In particolare:

$$\sin x = \sqrt{\frac{1 - \cos 2x}{2}}$$

Si immagini di avere una oscillazione in tanti piani perpendicolari alla direzione di propagazione. In altri termini le masse che oscillano sono considerate come dei “piani di atomi” che oscillano. In un reticolo tridimensionale, le onde del tipo considerato qui sono tali che gli atomi di un piano si muovono in fase fra loro. Queste onde possono essere longitudinali o trasversali a seconda che gli spostamenti avvengono nella direzione di propagazione o in direzione perpendicolare a questa.

Nella condizione di Bragg, d è la distanza tra due atomi che in questo caso è a , mentre θ vale $\pi/2$. La condizione di Bragg diventa allora:

$$2a = \frac{2n\pi}{k}$$

che per $n = 1$ fornisce proprio $k = 2\pi/a$.

Dunque la prima riflessione di Bragg avviene esattamente al bordo zona, in questo punto si hanno cioè i valori del vettore d'onda per i quali avviene la prima riflessione di Bragg. Questa riflessione che avviene al bordo della zona di Brillouin è come se impedisse la propagazione ulteriore delle onde, dando origine ad un'onda stazionaria all'interno della zona.

Questo è vero in generale e dunque lo si vedrà anche per gli elettroni. Infatti, anche alle autofunzioni di Bloch per gli elettroni può essere attribuita una velocità (di propagazione) e ai bordi della zona di Brillouin, cioè agli estremi di banda, si ottiene $\frac{dE}{dk} = 0$ e quindi in corrispondenza di questi estremi la velocità di gruppo degli elettroni è zero.

Le funzioni di spostamento $u_s(t) = a_0 e^{i(ksa - \omega t)}$ negli estremi di banda sono:

$$u_s(t) = a_0 e^{i(k s \pi - \omega t)} = a_0 e^{i k s \pi} e^{-i \omega t}$$

ma $e^{i k s \pi}$ vale ± 1 a seconda che s sia pari o dispari rispettivamente. Dunque, *ogni atomo oscilla con la stessa pulsazione, ma in opposizione di fase rispetto ai suoi vicini*: e questo descrive proprio un'onda stazionaria.

Caso $k \rightarrow 0$

Valori molto piccoli di k corrispondono a valori molto grandi della lunghezza d'onda. In particolare, la lunghezza d'onda è molto più grande del passo del reticolo e le oscillazioni si propagano come un'onda in un mezzo continuo. Formalmente, per valori piccoli di k vale $\sin x \simeq x$ e si può confondere il seno con l'angolo. La relazione $\omega(k)$ diventa allora lineare, proprio come per le onde in un continuo:

$$\omega = 2\sqrt{\frac{c_1}{M}} \left| \sin\left(\frac{ka}{2}\right) \right| \simeq \sqrt{\frac{c_1}{M}} |ka|$$

e dunque la velocità di propagazione vale:

$$v_g = \frac{d\omega}{dk} = \sqrt{\frac{c_1}{M}} a$$

A questa conclusione si giunge anche considerando l'equazione (di Newton) cui soddisfano le funzioni degli spostamenti nell'ipotesi di interazione con i primi vicini:

$$M \frac{d^2 u_s}{dt^2} = c_1 (u_{s+1} - u_s + u_{s-1} - u_s)$$

si noti come nel termine di destra compaiano solo le distanze tra i primi vicini e la costante elastica è la stessa.

Il termine di destra si può riscrivere moltiplicando e dividendo per a^2 :

$$\frac{d^2 u_s}{dt^2} = \frac{c_1}{M} a^2 \left[\frac{1}{a} \left(\frac{u_{s+1} - u_s}{a} + \frac{u_{s-1} - u_s}{a} \right) \right] = \frac{c_1}{M} a^2 \left[\frac{1}{a} \left(\frac{u_{s+1} - u_s}{a} - \frac{u_s - u_{s-1}}{a} \right) \right]$$

Ora, mandando a a zero, ovvero nel limite della lunghezza d'onda infinitamente grande, a destra compare un rapporto incrementale del rapporto incrementale di una $u(x)$, ovvero una derivata seconda in x :

$$\frac{d^2 u_s}{dt^2} = \frac{c_1}{M} a^2 \frac{d^2 u_s}{dx^2}$$

che è l'equazione delle onde in un continuo (equazione della corda), con velocità di propagazione data da:

$$v = \sqrt{\frac{c_1}{M}} a$$

Dunque in un certo senso, per $k \rightarrow 0$ l'ipotesi di discretizzazione scompare e si ritrova un sistema continuo.

Conclusioni

Si è studiato il moto degli ioni di un reticolo cristallino.

Innanzitutto si è fatta l'ipotesi che ogni ione risenta di un potenziale, dovuto agli elettroni di valenza e agli altri ioni, che si può sviluppare in serie attorno alle posizioni di equilibrio degli ioni, ovvero intorno ai siti del reticolo. Si è quindi arrestato lo sviluppo al secondo ordine, ottenendo un potenziale quadratico rispetto all'elongazione: in sostanza, questo significa approssimare la forza di cui risente ogni ione con una forza elastica. Sono state quindi scritte le equazioni di Newton del sistema completo, descrivendo le forze elastiche che intercorrono tra gli ioni. Si sono scritte le soluzioni di queste equazioni in forma di moti armonici.

Il passo rilevante è stato l'uso del teorema di Bloch sull'ampiezza a_s di questi moti armonici. Questo porta infatti a disaccoppiare le equazioni di Newton e a trovare la relazione di dispersione. L'ipotesi di interazione solo con i primi vicini semplifica già i calcoli: in questo caso infatti la matrice dei coefficienti del sistema di equazioni accoppiate di Newton diventa "quasi diagonale", nel senso che gli elementi non nulli sono solo quelli sulla diagonale principale e sulle due diagonali subito sopra e subito sotto la diagonale principale. L'utilizzo del teorema di Bloch invece permette di diagonalizzare perfettamente la matrice.

A differenza dello spettro degli elettroni, si è trovata una sola banda per le frequenze. Questo fatto è coerente con quello che si vedrà studiando i modi normali e in particolare col fatto che il numero di modi normali, cioè di frequenze permesse, è finito e pari al numero degli atomi nel reticolo. Quindi deve essere finito il numero delle possibili frequenze di vibrazione.

In realtà anche nel caso in studio i valori di k permessi sono pari al numero di atomi nel reticolo, come si vede una volta imposte le condizioni al contorno. Se per esempio si impongono condizioni al contorno periodiche (lo spostamento dell'atomo N -esimo uguale allo spostamento dell'atomo in zero) si trova che sono permessi solo alcuni valori di k , e quindi di ω , all'interno della banda. Infatti con queste condizioni al contorno il teorema di Bloch implica che:

$$u_N = u_0 \rightarrow a_0 e^{i(kNa - \omega t)} = u_0 \rightarrow a_0 e^{i(k0a - \omega t)} \rightarrow e^{ikNa} = e^{ik0a} = 1$$

da cui:

$$kNa = 2n\pi \rightarrow k = \frac{2n\pi}{Na}$$

e questo implica che i valori possibili di k sono N , se si ricorda che k varia da $-\pi/a$ a π/a .

4.3.3 Modi Normali

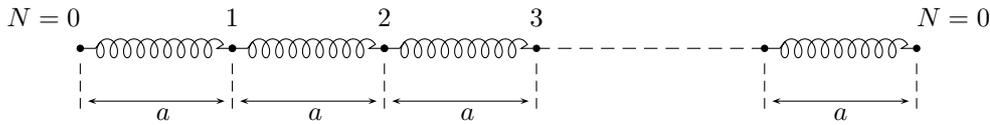
Si riaffronta ora lo studio delle vibrazioni in un reticolo in un modo più rigoroso (e, almeno in un secondo momento, quantistico) in modo da determinare i modi normali.

In particolare si studierà, ancora con un formalismo classico Newtoniano, un reticolo unidimensionale, quindi una catena di atomi, supponendo che le interazioni tra loro siano di tipo elastico. Le condizioni al contorno imposte saranno ancora quelle periodiche, e quindi il moto dell'ultimo atomo coinciderà con quello del primo, in altri termini la catena è ancora un anello.

Si ricaverà quindi una trasformazione per le posizioni degli atomi del reticolo, espressa attraverso esponenziali del tipo e^{ik} . Con queste trasformazioni, e con opportune condizioni al contorno, le equazioni del moto (equazioni di Newton) si riducono ad equazioni di oscillatore armonico. Si definiscono quindi delle coordinate normali che come si vedrà rappresentano proprio i modi normali di oscillazione del reticolo.¹⁹

Sia una catena lineare:

¹⁹La quantizzazione delle vibrazioni reticolari sarà utilizzata in seguito per dare conto del comportamento del calore specifico degli isolanti.



Si utilizzeranno condizioni al contorno periodiche, che si esprimono dicendo che lo spostamento del primo atomo è uguale a quello dell'ultimo, formalmente:²⁰

$$u_0 = u_N$$

Si impone poi che per le funzioni u_s valga il teorema di Bloch per gli spostamenti lungo il reticolo:

$$u_s = u_0 e^{iks a}$$

La condizione di periodicità implica che:

$$u_N = e^{ikNa} u_0 = u_0$$

e queste condizioni determinano il valore di k :

$$e^{ikNa} u_0 = u_0 \Rightarrow e^{ikNa} = 1 \Rightarrow kNa = 2p\pi \Rightarrow k_p = \frac{2p\pi}{Na}$$

con $0 \leq p \leq N - 1$ intero. Alternativamente, si può scegliere p intero e compreso tra $-N/2$ e $N/2$, l'estremo incluso se p è pari. Questa seconda scelta è più comoda perché rende il problema simmetrico rispetto a k . Nei calcoli seguenti si userà però la prima convenzione, che ha come vantaggio il fatto che i domini di variabilità dei due indici n e p sono uguali.²¹

Il fattore di Bloch dell' n -simo atomo della catena è e^{ikna} , e considerando che le condizioni al contorno impongono che k assuma solo determinati valori si ha:

$$e^{ikna} = e^{ina \frac{2p\pi}{Na}} = e^{i2\pi \frac{np}{N}}$$

dove gli indici n e p scorrono entrambi da 0 a $N - 1$. Si ricordi che n è l'indice associato alle posizioni degli atomi nella catena e p è l'indice associato ai valori permessi del numero d'onda k . È possibile dunque costruire una matrice $N \times N$ con i fattori di Bloch appena visti.

In queste considerazioni preliminari è stato citato il teorema di Bloch solo come richiamo per mettere in evidenza i legami con esso, tuttavia nell'approccio che si sta per sviluppare non si fa ricorso ad esso. Come si vedrà, questo nuovo approccio consiste nella ricerca delle trasformazioni a coordinate normali.

È tuttavia possibile mostrare la piena coerenza con quanto fatto nell'approccio precedente, che utilizzava appunto il teorema di Bloch. In particolare, si mostrerà come le soluzioni del moto ottenute con questo secondo approccio soddisfino il teorema di Bloch.

L'approccio più formale alla teoria delle vibrazioni dei reticoli cristallini seguito qui consiste nel cercare la trasformazione che permette di passare dalle coordinate u_n alle *coordinate normali* del sistema, cioè le coordinate in cui le equazioni del moto del sistema si disaccoppiano (metodo di separazione delle variabili).

Proposizione sulla trasformazione a coordinate normali:

La trasformazione a coordinate normali è rappresentata da una matrice i cui elementi sono:

$$\left(e^{i2\pi \frac{np}{N}} \right)_{np}$$

Si scriva la trasformazione:

$$x_p = \sum_{n=0}^{N-1} e^{i2\pi \frac{np}{N}} u_n$$

²⁰È una periodicità di periodo N rispetto alla variabile discreta di spostamento lungo il reticolo.

²¹D'altra parte, per passare da una scelta all'altra basta sommare $N/2$.

è possibile rendere la trasformazione unitaria aggiungendo un opportuno fattore di normalizzazione:

$$x_p = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} e^{i2\pi \frac{np}{N}} u_n \quad \text{trasformazione a coordinate normali}$$

Si vuole dimostrare che in queste coordinate le equazioni del moto si disaccoppiano e si trasformano in N equazioni di oscillatore armonico.

Le equazioni del sistema sono le seguenti:

$$M \frac{d^2 u_n}{dt^2} = \sum_{d=0}^{+\infty} c_d (u_{n+d} - u_n)$$

introducendo l'ipotesi di interazione con i soli vicini, queste equazioni si semplificano notevolmente:

$$M \ddot{u}_s = c_1 (u_{n+1} - 2u_n + u_{n-1})$$

e applicando la trasformazione a coordinate normali a queste equazioni di Newton si ottiene:

$$\begin{aligned} \frac{M}{\sqrt{N}} \sum_{n=0}^{N-1} e^{i2\pi \frac{np}{N}} \ddot{u}_n &= \frac{c_1}{\sqrt{N}} e^{i2\pi \frac{np}{N}} (u_{n+1} - 2u_n + u_{n-1}) \rightarrow \\ M \frac{\partial^2}{\partial t^2} \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} e^{i2\pi \frac{np}{N}} u_n &= \frac{c_1}{\sqrt{N}} \left(\sum_{n=0}^{N-1} e^{i2\pi \frac{np}{N}} u_{n+1} - 2 \sum_{n=0}^{N-1} e^{i2\pi \frac{np}{N}} u_n + \sum_{n=0}^{N-1} e^{i2\pi \frac{np}{N}} u_{n-1} \right) \end{aligned}$$

Gli esponenziali possono essere riarrangiati per semplificare le formule:

$$\begin{aligned} M \ddot{x}_p &= \frac{c_1}{\sqrt{N}} \left(\sum_{n=0}^{N-1} e^{i2\pi \frac{p}{N}} e^{-2\pi i \frac{p}{N}} e^{-2\pi i \frac{np}{N}} e^{i2\pi \frac{np}{N}} u_{n+1} - 2 \sum_{n=0}^{N-1} e^{i2\pi \frac{np}{N}} u_n + \right. \\ &\quad \left. + \sum_{n=0}^{N-1} e^{i2\pi \frac{p}{N}} e^{-2\pi i \frac{p}{N}} e^{-2\pi i \frac{np}{N}} e^{i2\pi \frac{np}{N}} u_{n-1} \right) \rightarrow \\ M \ddot{x}_p &= \frac{c_1}{\sqrt{N}} \left(e^{-2\pi i \frac{p}{N}} \sum_{n=0}^{N-1} e^{i2\pi \frac{(n+1)p}{N}} u_{n+1} - 2 \sum_{n=0}^{N-1} e^{i2\pi \frac{np}{N}} u_n + e^{i2\pi \frac{p}{N}} \sum_{n=0}^{N-1} e^{i2\pi \frac{(n-1)p}{N}} u_{n-1} \right) \rightarrow \\ M \ddot{x}_p &= \frac{c_1}{\sqrt{N}} \left(e^{-2\pi i \frac{p}{N}} \sum_{n=0}^N e^{i2\pi \frac{np}{N}} u_n - 2 \sum_{n=0}^{N-1} e^{i2\pi \frac{np}{N}} u_n + e^{i2\pi \frac{p}{N}} \sum_{n=0}^{N-2} e^{i2\pi \frac{np}{N}} u_n \right) \end{aligned}$$

Ora si vede che nelle sommatorie del primo e del terzo termine al secondo membro, il termine per $n = N$ è uguale al termine per $n = 0$ e il termine per $n = -1$ è uguale al termine per $n = N - 1$, cioè:

$$e^{i2\pi \frac{Np}{N}} u_N = e^{i2\pi \frac{0 \cdot p}{N}} u_0$$

e

$$e^{i2\pi \frac{-1 \cdot p}{N}} u_{-1} = e^{i2\pi \frac{(N-1)p}{N}} u_{N-1}$$

Infatti per le condizioni al contorno periodiche $u_N = u_0$ e $u_{N-1} = u_{-1}$, mentre per gli esponenziali, ricordando che p è un intero:

$$e^{i2\pi 0 \cdot p} = e^0 = 1 = e^{i2\pi p}$$

in definitiva:

$$\begin{aligned} M \ddot{x}_p &= \frac{c_1}{\sqrt{N}} \left(e^{-2\pi i \frac{p}{N}} \sum_{n=0}^{N-1} e^{i2\pi \frac{np}{N}} u_n - 2 \sum_{n=0}^{N-1} e^{i2\pi \frac{np}{N}} u_n + e^{i2\pi \frac{p}{N}} \sum_{n=0}^{N-1} e^{i2\pi \frac{np}{N}} u_n \right) \rightarrow \\ M \ddot{x}_p &= \frac{c_1}{\sqrt{N}} \left(e^{-2\pi i \frac{p}{N}} - 2 + e^{i2\pi \frac{p}{N}} \right) x_p \rightarrow \\ M \ddot{x}_p &= \frac{2c_1}{\sqrt{N}} \left[\cos \left(\frac{2\pi p}{N} \right) - 1 \right] x_p \rightarrow \\ \ddot{x}_p &= \frac{2c_1}{M\sqrt{N}} \left[\cos \left(\frac{2\pi p}{N} \right) - 1 \right] x_p \end{aligned}$$

definendo ora, coerentemente con le considerazioni iniziali, la quantità:

$$k_p \equiv 2\pi \frac{p}{aN}$$

l'equazione di sopra prende la forma:

$$\ddot{x}_p = \frac{2c_1}{M\sqrt{N}} (\cos k_p a - 1) x_p = \frac{2c_1}{M\sqrt{N}} \left(2 \sin^2 \frac{k_p a}{2} \right) x_p = \frac{4c_1}{M\sqrt{N}} \sin^2 \frac{k_p a}{2} x_p$$

che è proprio l'equazione di un oscillatore armonico $\ddot{x}_p = \omega_p^2 x_p$, con frequenza data da:

$$\omega_p = 2 \sqrt{\frac{4c_1}{M\sqrt{N}}} \left| \sin\left(\frac{k_p a}{2}\right) \right| = 2 \sqrt{\frac{4c_1}{M\sqrt{N}}} \left| \sin\left(\frac{\pi p}{N}\right) \right|$$

che è la stessa relazione di dispersione ottenuta in precedenza seguendo un approccio più fisico.

A causa della funzione seno, che è limitata, esiste quindi un solo intervallo finito di frequenze. All'interno di questa banda esiste un numero di frequenze possibili pari al numero di atomi N e individuate dall'indice p .

Queste frequenze sono i cosiddetti *modi normali di oscillazione* e le frequenze di oscillazione sono le stesse ottenute in precedenza con l'altro approccio, che partiva dal teorema di Bloch.

Una proprietà notevole della trasformazione a coordinate normali trovate è che per essa è facile definire la trasformazione inversa.

Proposizione

La trasformazione inversa è data dalla matrice i cui elementi sono:

$$u_n = \frac{1}{\sqrt{N}} \sum_{p=0}^{N-1} e^{-i2\pi \frac{np}{N}} x_p$$

ovvero i complessi coniugati degli esponenziali degli elementi della matrice della trasformazione a coordinate normali.

Dimostrazione

La dimostrazione si effettua mostrando che il prodotto riga per colonna delle due matrici che rappresentano le due trasformazioni restituisce la matrice identica.

$$\frac{1}{N} \sum_{p=0}^{N-1} e^{i2\pi \frac{mp}{N}} e^{-i2\pi \frac{np}{N}} = \frac{1}{N} \sum_{p=0}^{N-1} e^{i2\pi \frac{(m-n)p}{N}}$$

ponendo ora:

$$z = e^{i2\pi \frac{(m-n)}{N}}$$

il termine si può riscrivere come:

$$\frac{1}{N} \sum_{p=0}^{N-1} z^p = \frac{1}{N} \left(\sum_{p=0}^{+\infty} z^p - \sum_{p=N}^{+\infty} z^p \right) = \frac{1}{N} \left(\sum_{p=0}^{+\infty} z^p - \sum_{p=0}^{+\infty} z^{(p+N)} \right) = \frac{1}{N} \left(\sum_{p=0}^{+\infty} z^p - z^N \sum_{p=0}^{+\infty} z^p \right)$$

e sommando la serie geometrica:

$$\sum_{p=-\infty}^{+\infty} z^p = \frac{1}{1-z}$$

si ricava:

$$\frac{1}{N} \left(\sum_{p=0}^{+\infty} z^p - z^N \sum_{p=0}^{+\infty} z^p \right) = \frac{1}{N} \left(\frac{1}{1-z} - z^N \frac{1}{1-z} \right) = \frac{1}{N} \frac{1-z^N}{1-z} = \frac{1}{N} \frac{1 - e^{i2\pi(m-n)}}{1 - e^{i2\pi \frac{(m-n)}{N}}}$$

Tenendo presente che m e n sono due interi, nel caso $m \neq n$ questa espressione diventa:

$$\frac{1}{N} \frac{1 - \frac{e^{i2\pi m}}{e^{i2\pi n}}}{1 - e^{i2\pi \frac{(m-n)}{N}}} = \frac{1}{N} \frac{1 - 1}{1 - e^{i2\pi \frac{(m-n)}{N}}}$$

che vale evidentemente 0. Nel caso risulti $m = n$ si ha una forma indeterminata $0/0$ che può essere risolta riscrivendo il termine come segue:

$$\begin{aligned} \frac{1}{N} \frac{e^{i\pi(m-n)} (e^{-i\pi(m-n)} - e^{i\pi(m-n)})}{e^{i\pi \frac{(m-n)}{N}} (e^{-i\pi \frac{(m-n)}{N}} - e^{i\pi \frac{(m-n)}{N}})} &= \frac{1}{N} \frac{-2e^{i\pi(m-n)} \sin \pi(m-n)}{-2e^{i\pi \frac{(m-n)}{N}} \sin \pi \frac{(m-n)}{N}} = \\ &= \frac{1}{N} \frac{e^0 \sin \pi(m-n)}{e^0 \sin \pi \frac{(m-n)}{N}} = \\ &= \frac{1}{N} \frac{\sin \pi(m-n)}{\sin \pi \frac{(m-n)}{N}} \end{aligned}$$

poiché questo è il caso in cui l'argomento del seno si annulla, si può sostituire l'argomento al seno:²²

$$\frac{1}{N} \frac{(m-n)}{\frac{(m-n)}{N}} = \frac{N}{N} = 1$$

In definitiva quindi:

$$\frac{1}{N} \sum_{p=0}^{N-1} e^{i2\pi \frac{mp}{N}} e^{-i2\pi \frac{np}{N}} = \frac{1}{N} \sum_{p=0}^{N-1} e^{i2\pi \frac{(m-n)p}{N}} = \delta_{mn}$$

ovvero il prodotto riga per colonna delle due matrici coincide effettivamente con la matrice identica e quindi le due trasformazioni sono l'una l'inversa dell'altra. Inoltre la trasformazione è unitaria, perché la trasformazione inversa è anche la trasposta coniugata e per definizione una trasformazione è unitaria quando moltiplicata per la sua trasposta²³ fornisce la matrice identica.

La trasformazione inversa:

$$u_n = \frac{1}{\sqrt{N}} \sum_{p=0}^{N-1} e^{-i2\pi \frac{np}{N}} x_p$$

permette quindi di passare dalle coordinate normali x_p (che descrivono i modo normali, appunto) alle coordinate degli spostamenti u_n . Essa permette quindi di sapere quali sono gli spostamenti, cioè le leggi orarie di ogni atomo, in corrispondenza di ogni modo normale di oscillazione. Basta infatti scrivere la trasformazione da modi normali a spostamenti applicata ad un vettore di modi normali in cui tutti i modi sono nulli tranne quello di interesse.

Per fare questo ci si deve procurare l'espressione dell'evoluzione nel tempo della coordinata normale che rappresenta il modo normale in questione. Ma questo è presto fatto se si ricorda che le coordinate normali soddisfano le equazioni di un oscillatore armonico $x_q(t) = x_q(0)e^{i\omega t}$, con $x_q(0)$ valore della posizione di equilibrio (ovvero di riposo)

Le oscillazioni corrispondenti a questo modo normale di oscillazione sono allora (moltiplicando e dividendo per a per far comparire k_q):

$$u_n(t) = \frac{x_q(0)}{\sqrt{N}} e^{i(\omega_q t - 2\pi \frac{nq}{N})} = \frac{x_q(0)}{\sqrt{N}} e^{i(\omega_q t - k_q n a)}$$

Si verifica che queste oscillazioni soddisfano la proprietà di Bloch:

$$u_n = e^{ik_q a} u_{n-1} \quad \text{e più in generale} \quad u_n = e^{ik_q d a} u_{n-d}$$

²²O, più formalmente, considerare il limite fondamentale

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1$$

²³Coniugata, se complessa.

Si noti come il fattore di Bloch dipenda dal modo di oscillazione normale in cui si trova il sistema.

Ciascuna coordinata normale corrisponde ad un modo normale del reticolo che è un'onda di momento k_q e di frequenza ω_q che si propaga attraverso il reticolo. A due valori di p che differiscono di N corrisponde lo stesso modo normale.

Per essere precisi, le x_p non sono tuttavia ancora i veri modi normali perché gli spostamenti corrispondenti sono complessi.

Poiché infatti le u_n rappresentano spostamenti fisici, queste devono essere delle quantità reali e dunque, per come sono definite, le x_p sono delle quantità complesse.²⁴ Per ottenere i veri modi normali si osservi che la forma della generica coordinata normale:

$$x_p = \sum_{n=0}^{N-1} e^{i2\pi \frac{np}{N}} u_n$$

implica che, dovendo le u_n essere reali, qualora si inverta il segno dell'indice della coordinata normale deve risultare $x_{-p} = x_p^*$. Queste due quantità complesse possono essere quindi esplicitate in termini di una parte reale A_p e di una parte immaginaria B_p :

$$\begin{aligned} x_p &= \frac{1}{\sqrt{2}} (A_p + iB_p) \\ x_{-p} &= \frac{1}{\sqrt{2}} (A_p - iB_p) \end{aligned} \quad (4.1)$$

Si sceglie ora la rappresentazione in cui p varia da $-N/2$ a $N/2$, invece di quella in cui p va da 0 a $N-1$. Si stabilisce inoltre di considerare solo i valori positivi di questo indice $0 \leq p \leq N/2$, dove il segno uguale vale solo per N pari, con la convenzione che ad ognuno di questi valori corrisponderanno due coordinate normali, l'una complessa coniugata dell'altra.

Ma, e questo è il nocciolo della questione, si possono considerare le due quantità reali A_p e B_p , parte reale e parte immaginaria delle due coordinate normali x_p e x_{-p} . Le coordinate A_p e B_p possono essere naturalmente espresse in termini delle x_p e x_{-p} :

$$\begin{aligned} A_p &= \frac{2}{\sqrt{2}} (x_p + x_{-p}) \\ B_p &= \frac{2i}{\sqrt{2}} (x_p - x_{-p}) \end{aligned}$$

Le vere coordinate normali risultano allora essere A_p e B_p . Si ricercherà quindi la forma delle u_n in termini di A_p e B_p a partire dall'espressione delle u_n in termini delle x_p . Si diano le equazioni:

$$\begin{aligned} u_n(t) &= \frac{x_p(t_0)}{\sqrt{N}} e^{i\omega_p(t-t_0)} e^{-i2\pi \frac{np}{N}} \\ u_n(t) &= \frac{x_{-p}(t_0)}{\sqrt{N}} e^{i\omega_{-p}(t-t_0)} e^{i2\pi \frac{np}{N}} \end{aligned}$$

che rappresentano il moto del generico atomo del reticolo (o piuttosto la distribuzione dei moti), quando il sistema si trova nel modo di oscillazione p o $-p$ rispettivamente.

Questo significa che la distribuzione dei moti per le vere coordinate normali del sistema è una combinazione lineare dei due modi x_p e x_{-p} . Si considerino quindi le due combinazioni lineari che danno luogo a A_p e B_p :

$$\begin{cases} u_n(t) = \frac{1}{\sqrt{N}} e^{-i2\pi \frac{np}{N}} \frac{2}{\sqrt{2}} [x_p(t) + x_{-p}(t)] \\ u_n(t) = \frac{1}{\sqrt{N}} e^{-i2\pi \frac{np}{N}} \frac{2}{\sqrt{2}} [x_p(t) - x_{-p}(t)] \end{cases}$$

²⁴Questo a causa dell'esponentiale presente nella trasformazione.

tenendo ora presente la (4.1) e la $x_q(t) = x_q(0)e^{i\omega t}$, risulta:

$$\begin{cases} x_p(t) &= \frac{1}{\sqrt{2}} [A_p(t_0) \cos \omega_p(t - t_0) + iB_p(t_0) \sin \omega_p(t - t_0)] \\ x_{-p}(t) &= \frac{1}{\sqrt{2}} [A_p(t_0) \cos \omega_p(t - t_0) - iB_p(t_0) \sin \omega_p(t - t_0)] \end{cases}$$

e quindi:

$$\begin{cases} u_n(t) &= \sqrt{\frac{2}{N}} \left(\cos 2\pi \frac{np}{N} - i \sin 2\pi \frac{np}{N} \right) 2A_p(t_0) \cos \omega_0(t - t_0) \\ u_n(t) &= \sqrt{\frac{2}{N}} \left(\cos 2\pi \frac{np}{N} - i \sin 2\pi \frac{np}{N} \right) 2iB_p(t_0) \sin \omega_0(t - t_0) \end{cases}$$

da cui:

$$\begin{cases} u_n(t) &= \sqrt{\frac{2}{N}} A_p(t_0) \cos 2\pi \frac{np}{N} \cos \omega_0(t - t_0) \\ u_n(t) &= \sqrt{\frac{2}{N}} B_p(t_0) \sin 2\pi \frac{np}{N} \sin \omega_0(t - t_0) \end{cases}$$

e moltiplicando e dividendo per a per far comparire k_a , come al solito:

$$\begin{cases} u_n(t) &= \sqrt{\frac{2}{N}} A_p(t_0) \cos(k_p n a) \cos[\omega_0(t - t_0)] \\ u_n(t) &= \sqrt{\frac{2}{N}} B_p(t_0) \sin(k_p n a) \sin[\omega_0(t - t_0)] \end{cases}$$

4.3.4 Formalismo Hamiltoniano

Si verificherà ora la separazione delle variabili utilizzando il formalismo Hamiltoniano invece del formalismo newtoniano, come fatto fin qui.

L'energia cinetica è espressa da:

$$T = \frac{M}{2} \sum_{n=0}^{N-1} \dot{x}_n^2$$

poiché vale:

$$u_n = \frac{1}{\sqrt{N}} \sum_{\mu=0}^{N-1} e^{-i2\pi \frac{\mu n}{N}} \dot{x}_\mu \quad u_n = \frac{1}{\sqrt{N}} \sum_{\nu=0}^{N-1} e^{i2\pi \frac{\nu n}{N}} \dot{x}_{-\nu}$$

si ha:

$$T = \frac{M}{2} \sum_{n=0}^{N-1} \left(\frac{1}{\sqrt{N}} \sum_{\mu=0}^{N-1} e^{-i2\pi \frac{\mu n}{N}} \dot{x}_\mu \right) \left(\frac{1}{\sqrt{N}} \sum_{\nu=0}^{N-1} e^{i2\pi \frac{\nu n}{N}} \dot{x}_{-\nu} \right) = \frac{M}{2} \sum_{n=0}^{N-1} \left(\frac{1}{N} \sum_{\mu=0}^{N-1} e^{-i2\pi \frac{\mu n}{N}} e^{i2\pi \frac{\mu n}{N}} \dot{x}_\mu \dot{x}_{-\mu} \right) \blacksquare$$

ora si può utilizzare l'unitarietà della trasformazione dimostrata in precedenza e poiché in entrambi gli esponenziali compare n , il risultato della sommatoria nella parentesi è $\delta_{n\mu}$ e quindi:

$$T = \frac{M}{2} \sum_{n=0}^{N-1} \dot{x}_n \dot{x}_{-n}$$

Si analizzino ora i termini della somma. Siccome vale:

$$\begin{cases} x_p &= \sum_{n=0}^{N-1} e^{i2\pi \frac{np}{N}} u_n \\ x_{-p} &= \sum_{n=0}^{N-1} e^{-i2\pi \frac{np}{N}} u_n \end{cases}$$

risulta evidente che per $n = 0$ e $N/2$, x_n e x_{-n} sono reali in quanto gli esponenziali valgono 1. Da notare che il termine $N/2$ esiste solo per N pari, perché n è un intero.

Negli altri casi, x_n e x_{-n} sono complessi coniugati, quindi il loro prodotto è dato dalla somma dei quadrati della parte reale e della parte immaginaria. In definitiva quindi l'energia cinetica si scrive come:²⁵

$$T = \frac{M}{2} \dot{x}_0^2 + \underbrace{\frac{M}{2} \dot{x}_{N/2}^2}_{\text{solo per } N \text{ pari}} + \frac{M}{2} \sum_{n=0}^{\lfloor N/2 \rfloor} (\dot{A}_n^2 + \dot{B}_n^2)$$

che è una quantità reale.

L'energia potenziale è invece espressa dalla forma:

$$U = \frac{M}{2} \sum_{n=0}^{N-1} \omega_n^2 u_n^2$$

che con passaggi perfettamente analoghi a quelli utilizzati per ridurre la forma dell'energia cinetica, permette di giungere alla forma reale:

$$U = \underbrace{\frac{M}{2} \omega_{N/2}^2 x_{N/2}^2}_{\text{solo per } N \text{ pari}} + \frac{M}{2} \sum_{n=0}^{\lfloor N/2 \rfloor} \omega_n^2 (A_n^2 + B_n^2)$$

Si noti che in questo caso il termine in ω_0 non compare perché è la frequenza corrispondente a k_0 che, per come è fatta la relazione di dispersione, è nulla. *Si vede dunque che il termine x_0 compare solo nell'energia cinetica e rappresenta il moto del centro di massa, che è un moto a frequenza nulla.* Infatti, l'Hamiltoniana del sistema in termini di A_n e B_n è data da:

$$\begin{aligned} H &= \frac{M}{2} \dot{x}_0^2 + \underbrace{\frac{M}{2} \dot{x}_{N/2}^2}_{\text{solo per } N \text{ pari}} + \sum_{n=0}^{\lfloor N/2 \rfloor} (\dot{A}_n^2 + \dot{B}_n^2) + \underbrace{\frac{M}{2} \omega_{N/2}^2 x_{N/2}^2}_{\text{solo per } N \text{ pari}} + \sum_{n=0}^{\lfloor N/2 \rfloor} \omega_n^2 (A_n^2 + B_n^2) \\ &= \frac{M}{2} (\dot{A}_0^2 + \dot{B}_0^2) + \underbrace{\frac{M}{2} (\dot{A}_{N/2}^2 + \dot{B}_{N/2}^2)}_{\text{solo per } N \text{ pari}} + (A_{N/2}^2 + B_{N/2}^2) + \\ &\quad + \frac{M}{2} \sum_{n=0}^{\lfloor N/2 \rfloor} (\dot{A}_n^2 + \dot{B}_n^2) + \frac{M}{2} \sum_{n=0}^{\lfloor N/2 \rfloor} \omega_n^2 (A_n^2 + B_n^2) \end{aligned}$$

che si può separare in tante Hamiltoniane della forma:

$$H_n = \frac{M}{2} (\dot{A}_n^2 + \dot{B}_n^2) + \frac{M}{2} \omega_n^2 (A_n^2 + B_n^2)$$

più un termine per $n = 0$:

$$H_0 = \frac{M}{2} \dot{x}_0^2 = \frac{1}{N} \sum_{n=0}^{N-1} e^{i4\pi \frac{n \cdot 0}{N}} \dot{u}_n^2$$

che è una Hamiltoniana di particella libera la cui posizione è la somma di tutte le posizioni degli atomi fratto il numero degli atomi e dunque la posizione del centro di massa per definizione. Gli altri termini sono tutte delle Hamiltoniane di oscillatore armonico.

Ovviamente, è possibile quantizzare queste Hamiltoniane. Nel formalismo quantistico si possono quindi risolvere le equazioni di Schrödinger corrispondenti a queste Hamiltoniane disaccoppiate e mettendo insieme le relative soluzioni (autofunzioni) si giunge alle autofunzioni dell'Hamiltoniana complessiva, che saranno descritte nella rappresentazione delle coordinate A_n e B_n . Queste autofunzioni sono dei quanti di vibrazione individuati dallo specifico numero quantico per ciascuno dei valori di k (ne esiste uno per ogni sito del reticolo). *Questi quanti sono chiamati **fononi** e descrivono le eccitazioni quantistiche del reticolo unidimensionale (catena).*

²⁵Si ricordi che con la notazione $\lfloor x \rfloor$ si indica la parte intera di x .

4.3.5 Cenni sui reticoli tridimensionali

Si daranno ora alcuni cenni su come passare dal reticolo unidimensionale a reticoli tridimensionali.

Si considerino tre vettori di traslazione elementari, che riproducono il reticolo geometrico. Accanto al reticolo diretto se ne definisce un altro, detto *reticolo reciproco*, che permette di definire dei vettori \vec{k} . Anche nel reticolo reciproco ci sono tre traslazioni elementari, due che si ottengono dal reticolo diretto più una terza. Quest'ultima si ottiene considerando la direzione perpendicolare alle prime due, ed in questa direzione si usa un vettore di traslazione elementare che ha per modulo 2π fratto il modulo dell'omologo vettore di traslazione del reticolo geometrico: si definiscono così tre traslazioni elementari che generalizzano il vettore \vec{k} .

In questo modo quella che in una dimensione era la zona di Brillouin diventa ora una *cella di Brillouin*, un poliedro intersezione di piani. Essa si costruisce considerando un atomo e i segmenti che lo congiungono a tutti i suoi primi vicini. Considerando poi i piani perpendicolari a questi segmenti e passanti per i punti medi, il minimo volume definito dall'intersezione di questi piani costituisce la cella di Brillouin. Il caso più semplice è il reticolo cubico, il cui reciproco è ancora cubico, ma con i lati più piccoli perché dati da 2π diviso i lati del reticolo diretto.

Tutte le proprietà di k enunciate sono relative al reticolo reciproco.

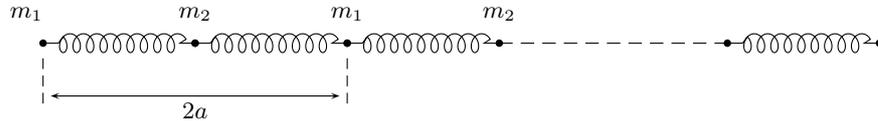
Il teorema di Bloch si definisce con i valori di k che appartengono al poliedro della prima zona di Brillouin. Inoltre, questi la riempiono in maniera uniforme se si impongono condizioni al contorno periodiche.

Le u_n sono periodiche sia nel reticolo diretto che in quello reciproco.

La struttura a bande è periodica nel reticolo reciproco.

Case preliminare: reticolo lineare con due atomi per cella

Si consideri un sistema unidimensionale di tanti oscillatori che si ripetono indefinitamente e la cui cella elementare, di lunghezza $2a$, è costituita da due oscillatori di massa diversa. Per semplicità si supponga che la distanza sia sempre a e che anche le costanti elastiche siano eguali a c



Le due equazioni del moto, una per gli atomi di posto pari e l'altra per quelli di posto dispari, usando l'ipotesi di interazione limitata ai primi vicini, è data da:

$$\begin{cases} m_1 u_{2s} = c(u_{2s+1} - 2u_{2s} + u_{2s-1}) \\ m_2 u_{2s+1} = c(u_{2s+2} - 2u_{2s+1} + u_{2s}) \end{cases}$$

dove si sono numerati i posti dando al primo atomo l'indice zero, al secondo l'indice 1 e così via. In questo modo gli atomi di posto pari hanno massa m_1 , mentre gli atomi di posto dispari hanno massa m_2 .

Questo sistema di N equazioni può essere affrontato utilizzando il teorema di Bloch, che permette di ricavare:

$$\begin{aligned} u_{2s} &= u_1 e^{i(k2sa - \omega t)} \\ u_{2s+1} &= u_2 e^{i[k2(s+1)a - \omega t]} \end{aligned}$$

Si noti che si è utilizzato direttamente il teorema di Bloch senza passare per la trasformazione a coordinate normali. In effetti, si è dimostrato che le due cose sono equivalenti e dunque le soluzioni scritte in questa forma sono già i modi normali.

Nel caso di reticolo tridimensionale con più di due atomi per cella, si vede che le bande di vibrazione sono tante quanto i gradi di libertà dell'oggetto la cui ripetizione nello spazio riproduce il reticolo. Nel caso dello spettro elettronico le bande sono infinite.

Nel caso invece delle bande vibrazionali non può essere così, perché il numero dei modi normali di vibrazione deve essere pari al numero di gradi di libertà del sistema.

Dunque, in un reticolo tridimensionale costituito da N atomi, il numero di gradi di libertà è $3N$ nel caso di un solo atomo per cella: si ottengono così tre distinte bande di frequenze di vibrazione permesse, come ad esempio accade per le onde elastiche. Queste possono essere sia longitudinali che trasversali: quelle longitudinali hanno una sola relazione di dispersione con un'unica velocità di propagazione (velocità di gruppo), mentre quelle trasversali possono avere in linea di principio due velocità di propagazione diverse (la velocità di propagazione dipende infatti dalle costanti elastiche).²⁶

Quindi, per un generico reticolo tridimensionale con un atomo per cella, le bande vibrazionali sono tre.²⁷ Le tre bande vibrazionali tendono, per $k \rightarrow 0$, alle tre onde elastiche, due trasverse e una longitudinale, con tre velocità di propagazione diverse.

Per vedere un esempio concreto, si consideri un reticolo cubico, nel quale basta vedere cosa accade nel segmento, perché quello che accade nelle tre direzioni è uguale. Se esiste più di un atomo per cella, i gradi di libertà del sistema devono aumentare e così i modi di vibrazione.

Se si considera una cella in cui ci sono R atomi, il numero di gradi di libertà complessivo è $3NR$ e dunque tante devono essere complessivamente le frequenze dei modi di vibrazione. Rispetto al caso di un atomo per cella, in cui i gradi di libertà sono $3N$, i gradi di libertà in più sono dati da:

$$3NR = 3N + x \rightarrow x = 3NR - 3N = 3N(R - 1)$$

Dunque, poiché ciascuna banda vibrazionale ha al suo interno un numero di frequenze di vibrazione (e quindi di stati) pari al numero di oggetti che costituiscono il reticolo, e cioè N , è naturale aspettarsi che compaiano $3(R - 1)$ bande in più, qualora si passi da un atomo per cella a R atomi per cella. Le tre bande originarie vanno tutte e tre a zero per $k \rightarrow 0$, le altre $3(R - 1)$ restano al di sopra.

In conclusione, a differenza dello spettro elettronico le bande dello spettro vibrazionale sono in numero finito. Inoltre il numero di frequenze permesse dal teorema di Bloch è pari al numero totale di gradi di libertà del reticolo, e quindi $3NR$. Questo risultato servirà per lo studio dei calori specifici e renderà ragione della legge di Dulong-Petit.

4.4 Elettrone fortemente legato

Si vuole ora studiare un'altra approssimazione per ottenere le bande elettroniche, cioè quella cosiddetta del metodo dell'*elettrone fortemente legato*. Si tratta di ragionare in termini degli stati legati degli oggetti i quali avvicinati insieme costituiscono il reticolo cristallino.

Si tratta dunque di un sistema fisico che ha tanti stati localizzati su tanti nuclei (o ioni) che all'inizio si trovano a distanza infinita l'uno dall'altro. Si vuole studiare cosa accade quando si avvicinano questi nuclei, si tratta quindi di un'approssimazione che costruisce le bande a partire dagli stati legati. Con l'altro metodo visto (Kronig e Penney, §4.2) si è usato invece un modello in cui gli elettroni erano liberi ma su cui agiva un potenziale periodico visto come una perturbazione, e si era giunti al risultato che a qualunque energia si susseguivano bande permesse e gap.

In un certo senso si possono reinterpretare le prime bande che compaiono: invece di essere relative all'effetto della perturbazione sugli stati di elettrone libero (che hanno quindi un'energia molto più grande dei coefficienti dello sviluppo di Fourier del potenziale), si possono interpretare come "stati atomici". Si è visto infatti che mandando il periodo del potenziale all'infinito, la prima banda si riduce allo stato legato della delta.

²⁶Le onde trasversali, d'altra parte, possono essere polarizzate in due direzioni e quindi essere legate a due costanti elastiche diverse.

²⁷Esistono cioè tre diverse relazioni di dispersione che individuano tre intervalli (sovrapposti) di frequenze permesse, all'interno di ognuno dei quali le condizioni al contorno periodiche individuano N frequenze pari al numero di atomi del reticolo.

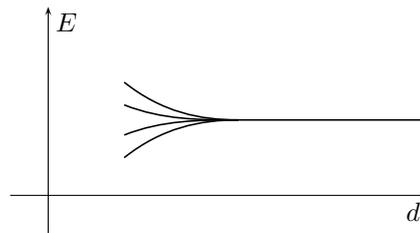
Si vuole ora generalizzare questo approccio, costruendo almeno le prime bande elettroniche a partire dagli stati legati dei vari atomi che compongono il reticolo. Sebbene questo modo di procedere renda ragione solo delle bande più basse, esso è istruttivo e indicativo perché *lega la formazione delle bande alla formazione del legame tra gli atomi*. Si ricordi che si è già vista la formazione del legame tra due atomi di idrogeno partendo dai due atomi isolati ed avvicinandoli. Mano a mano che si avvicinano, usando lo stato legante e lo stato antilegante e facendone una combinazione lineare, si possono realizzare stati in cui l'elettrone è localizzato solo sull'uno o sull'altro atomo. Questo diventa ancora più istruttivo nella molecola ione idrogeno, quando un solo elettrone è presente su due protoni. Questo anzi è un paragone più calzante con i calcoli mostrati qui, perché la struttura a bande è calcolata in approssimazione di singolo elettrone, si sta cioè supponendo che gli elettroni si muovano in modo indipendente in un potenziale unico, avendo incluso in questo potenziale periodico anche l'effetto medio (repulsivo) degli altri elettroni.

Dallo studio della molecola ione idrogeno si erano ottenuti due stati, uno legante e uno antilegante. Nel limite in cui i due protoni vanno all'infinito, questi due stati si riducono ai due stati fondamentali dell'atomo di idrogeno, che quindi sono degeneri in energia (hanno infatti la stessa energia dello stato fondamentale dell'atomo di idrogeno). In questa configurazione si ha un unico stato due volte degenero: un elettrone sul primo protone e l'altro all'infinito e viceversa.

Quando la distanza diminuisce la degenerazione viene rimossa e si ottengono due stati, uno con energia maggiore ed uno con energia minore. Mano a mano che diminuisce la distanza, la differenza in energia tra i due stati aumenta.

Si immagini di fare questo non con due soli protoni, ma con un numero N di protoni. Inizialmente ci sono N protoni a distanza infinita ed un elettrone. Ci sono N stati degeneri, che descrivono l'elettrone nello stato fondamentale di atomo di idrogeno centrato volta per volta su ognuno degli N atomi.

Avvicinando tra loro questi protoni, c'è una probabilità finita che l'elettrone passi per effetto tunnel da un protone all'altro e quindi abbassa la sua energia delocalizzandosi. Questo rimuove la degenerazione e si ottengono N stati distinti con diverse energie. Inoltre, al diminuire della distanza tra i protoni, gli stati si allargano (cioè aumenta la loro differenza in energia), dando origine ad una banda. Questa, se si vuole, è la ragione per cui una banda contiene tanti stati quanti sono gli atomi del cristallo che la costituisce: si può pensare che ogni stato è dovuto alla rimozione della degenerazione dei livelli atomici relativi a ciascun atomo. Un grafico dell'energia dei livelli in funzione della distanza tra i protoni è chiarificatore:



Come detto, questo approccio rende ragione solo delle bande a più bassa energia, tuttavia è interessante perché per queste bande si può istituire una corrispondenza tra bande e livelli atomici.

Caso unidimensionale

Nel caso unidimensionale le cose sono semplici, perché a ciascuno stato legato corrisponde una banda. Nel caso tridimensionale le cose sono più complesse, perché le bande provenienti dai singoli livelli atomici si possono mischiare.

In altri termini, la proprietà valida nel caso unidimensionale, per la quale ogni banda è separata da gap e la relazione di dispersione $E(k)$ è monotona, viene meno e dunque non si possono più individuare le bande con un indice. In realtà questo è ancora possibile perché nello spazio delle k (reticolo reciproco) gli insiemi degeneri sono di misura nulla e quindi si può utilizzare un prolungamento analitico.

Caso tridimensionale

Si consideri un sistema costituito da N atomi e si supponga per semplicità che abbiano tutti soltanto uno stato legato. Si supponga che gli stati legati di tutti gli atomi siano degeneri in energia, cioè che abbiano lo stesso autovalore dell'energia. Si supponga inoltre che attorno a questi atomi ci sia un unico elettrone. Se gli atomi si trovano a distanza infinita tra loro, esistono N stati possibili, nei quali l'elettrone è localizzato su ognuno degli N atomi. Si studierà ora cosa accade se si avvicinano gli atomi tra loro e per fare questo si utilizzerà un metodo simile al metodo L.C.A.O.

La funzione di prova utilizzata è la seguente:

$$\varphi_n(x) = \sum_{n=1}^N c_n v_n(x)$$

dove i c_n sono dei coefficienti e le v_n sono le funzioni d'onda dell'unico stato legato di ogni atomo. Si stanno evidentemente trascurando gli altri stati atomici (legati + continuo).

Si deve quindi calcolare l'elemento di matrice dell'Hamiltoniana su questa funzione di prova. Nella situazione in cui gli atomi sono infinitamente lontani, le funzioni d'onda degli stati legati di ogni atomo non si sovrappongono, quindi sono ortogonali. Quindi, almeno in un sottospazio queste costituiscono una base ortogonale e l'Hamiltoniana (di cui sono autofunzioni) è ortogonale. Inoltre, poiché si è supposto che tutti gli stati legati atomici sono degeneri in energia, sulla diagonale compare lo stesso elemento, che sarà denotato con E_0 . Man mano che gli atomi si avvicinano, le funzioni d'onda cominciano a sovrapporsi e dunque in linea di principio tutti gli elementi dell'Hamiltoniana sono non nulli.

Si può adottare per semplicità un'approssimazione di interazione ai primi vicini, che permette di supporre che la matrice dell'Hamiltoniana è bidiagonale:

$$H = \begin{pmatrix} E_0 & -A & 0 & 0 & \cdots \\ -A & E_0 & -A & 0 & \cdots \\ 0 & -A & E_0 & -A & \cdots \\ 0 & 0 & -A & E_0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

ovvero:

$$\begin{aligned} \langle v_n | H | v_n \rangle &= E_0 \\ \langle v_n | H | v_{n\pm 1} \rangle &= -A \\ \langle v_n | H | v_{n\pm p} \rangle &= 0 \quad p \neq 1 \end{aligned}$$

Inoltre, poiché il sistema deve essere invariante per traslazione, deve risultare:

$$v_n(x) = v_n(x - na)$$

che rappresenta il fatto che ogni funzione d'onda degli stati legati è uguale alle altre, ma è centrata su un diverso sito del reticolo di passo a . Il significato di A è lo stesso del cosiddetto integrale di sovrapposizione incontrato studiando le molecole biatomiche, e che era fatto come l'elemento di matrice del potenziale prodotto da un protone calcolato sulla funzione d'onda centrata sull'altro protone.

Si utilizza quindi una funzione di prova definita nel sottospazio generato dagli stati legati di singolo atomo centrati sui vari atomi e si cerca una soluzione al problema variazionale all'interno di questo sottospazio.

Ecco dunque l'equazione di Schrödinger, ovvero l'equazione agli autovalori dell'Hamiltoniana, in questo sottospazio:

$$H\varphi(x) = E\varphi(x) \quad \rightarrow \quad H \left| \sum_{n=1}^N c_n v(x) \right\rangle = E \left| \sum_{n=1}^N c_n v(x) \right\rangle$$

che proiettata a sinistra sullo stato $\varphi(x) = \sum_{n=1}^N c_n v(x)$:

$$\left\langle \sum_{n=1}^N c_n v(x) \middle| H \middle| \sum_{n=1}^N c_n v(x) \right\rangle = \left\langle \sum_{n=1}^N c_n v(x) \middle| E \middle| \sum_{n=1}^N c_n v(x) \right\rangle$$

fornisce il sistema di infinite equazioni in c_n :

$$-Ac_{n-1} + (E_0 - E)c_n - Ac_{n+1} = 0$$

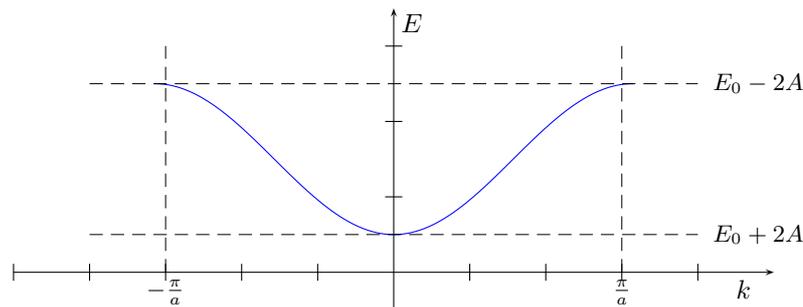
Ora, questo è un sistema di infinite equazioni se non si impongono condizioni al contorno. Imponendo invece una condizione di periodicità (o di Born–Von Kármán), questo si riduce ad un sistema di sole N equazioni:

$$\begin{cases} (E_0 - E)c_0 - Ac_1 & = 0 \\ -Ac_0 + (E_0 - E)c_1 - Ac_2 & = 0 \\ \dots & \\ -Ac_{N-1} + (E_0 - E)c_0 & = 0 \end{cases}$$

avendo utilizzato il fatto che $c_N = c_0$. Si tratta di equazioni accoppiate che però l'ipotesi di Bloch permette di disaccoppiare, ponendo in pratica $c_n = c_0 e^{ikna}$. Con questa forma, infatti, si semplifica la dipendenza da n e tutte le N equazioni diventano uguali:

$$\begin{aligned} -Ac_{n-1} + (E_0 - E)c_n - Ac_{n+1} &= 0 \rightarrow \\ -Ac_0 e^{ik(n-1)a} + (E_0 - E)c_0 e^{ikna} - Ac_0 e^{ik(n+1)a} &= 0 \rightarrow \\ -Ac_0 \cancel{e^{ikna}} e^{-ika} + (E_0 - E)c_0 \cancel{e^{ikna}} - Ac_0 \cancel{e^{ikna}} e^{ika} &= 0 \rightarrow \\ E(k) &= E_0 - 2A \cos(ka) \end{aligned}$$

Anche in questo caso, la relazione di dispersione definisce una sola banda, in quanto espressa da una funzione limitata:



dove A è una quantità negativa pari all'elemento di matrice del potenziale d'interazione $\hat{W}(x)$ tra due atomi contigui²⁸ calcolato sulle due funzioni d'onda dei due stati legati dei due atomi. Utilizzando il fatto che la $v_n(x)$ è la stessa funzione per ogni atomo, ma centrata in punti diversi, si ricava:

$$A = \int v(x) \hat{W}(x) v(x+a) dx$$

All'interno di questa banda, se si impongono le condizioni al contorno periodiche, ci sono N frequenze permesse. Infatti, dalla condizione $c_N = c_0$ si ottiene:

$$c_N = c_0 e^{ikNa} = c_0 \Rightarrow e^{ikNa} = 1 \Rightarrow kNa = 2n\pi \Rightarrow k = 2\pi \frac{n}{Na} \quad ; \quad 0 \leq n \leq N-1$$

4.5 Proprietà di conducibilità dei solidi

Si vuole ora mostrare come le proprietà di occupazione delle bande di energia permessa da parte degli elettroni diano origine alle proprietà di conducibilità dei solidi, senza tuttavia entrare in profondità nella questione del “trasporto”.

Il punto di partenza è la constatazione del fatto che gli stati di banda, cioè le funzioni d'onda di Bloch, consentono, al pari delle funzioni d'onda della particella libera, di definire un concetto di velocità, di attribuire quindi anche allo stato di banda una sorta di “velocità”.

²⁸Attrattiva, quindi negativa.

4.5.1 Evoluzione della particella libera

Le autofunzioni dell'Hamiltoniana della particella libera (per semplicità unidimensionale):

$$H = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2}$$

sono date da:

$$\varphi(x) = e^{ikx}$$

con autovalori:

$$E = \frac{\hbar^2 k^2}{2m} \quad \text{per la posizione } k = \sqrt{\frac{2mE}{\hbar}}$$

Da queste autofunzioni è possibile costruire un pacchetto d'onda pesandole tramite una funzione reale $g(k)$ e un fattore di fase $\phi(k)$:

$$\psi(x) = \int_0^{+\infty} g(k) e^{i[kx + \phi(x)]} dk$$

è possibile quindi tramite un'opportuna scelta di $g(k)$ e $\phi(k)$ costruire un pacchetto localizzato e seguirne il moto. Per fare questo si userà il metodo della fase stazionaria.

Prima di procedere è utile scrivere l'evoluzione nel tempo della funzione d'onda:

$$\psi(x, t) = \int_0^{+\infty} g(k) e^{i[kx + \phi(x)]} e^{-i\frac{E}{\hbar}t} dk = \int_0^{+\infty} g(k) e^{i\left[kx + \phi(x) - \frac{\hbar^2 k^2}{2m\hbar}t\right]} dk$$

Per sommi capi, il metodo di fase stazionaria è basato sul fatto che al variare di k varia l'argomento dell'esponenziale e quindi questo fattore oscilla rapidamente. Questa oscillazione fa sì che si abbiano contributi uguali e opposti all'integrale e che quindi danno mediamente contributo nullo. Secondo questo ragionamento, i valori di k che contribuiscono effettivamente alla funzione d'onda sono quelli per cui l'argomento dell'esponenziale è costante (stazionario) e quindi quelli per cui la sua derivata rispetto a k si annulla:

$$x - \frac{\hbar k}{m}t + \frac{d\phi(k)}{dk} = 0$$

che fornisce una relazione fra x e k . Per ogni fissata coppia "posizione x " e "tempo t ", il valore di k che soddisfa questa relazione è quello che rende la fase stazionaria e quindi quello che dà il massimo contributo all'integrale.

Il discorso si può invertire: fissato un certo valore di k , le coppie posizione-tempo che soddisfano questa relazione sono quelle che danno il maggior contributo all'integrale e dunque sono quelle per cui la $\psi(x, t)$, funzione d'onda che rappresenta il pacchetto, è per lo più non nulla. Di conseguenza, fissato k , la relazione:

$$x(t, k) = \frac{\hbar k}{m}t - \frac{d\phi(k)}{dk}$$

può essere considerata come la *legge oraria* del pacchetto. Questa legge descrive un moto uniforme con velocità $v = \frac{\hbar k}{m}$ e posizione iniziale $\frac{d\phi(k)}{dk}$. D'altra parte, per il concetto stesso di pacchetto, la $\psi(x, t)$ è non nulla nelle regioni in cui $g[k(x, t)]$ è non nulla. Se si costruisce quindi il pacchetto con una $g(k)$ molto localizzata intorno ad un valore k_0 ,²⁹ si può concludere che il pacchetto descrive una particella che si muove di moto uniforme con velocità $v_g = \frac{\hbar k_0}{m}$ (velocità di gruppo della particella libera).

²⁹Questo significa in pratica costruire il pacchetto utilizzando solo poche autofunzioni raggruppate attorno ad una determinata energia o addirittura con un'unica autofunzione.

4.5.2 Evoluzione della funzione di Bloch

Quanto visto finora riguarda una particella libera. Si trasporterà ora quanto visto al caso di un elettrone che si muove in un potenziale periodico (elettrone in un solido). Si è visto che la funzione d'onda di un tale elettrone è una funzione di Bloch,³⁰ che nel caso unidimensionale ha la forma $u(x; k)e^{ikx}$. In essa compare la funzione $u(x; k)$ periodica in x , con periodicità pari al passo del reticolo, e che dipende anche da k . Il valore di k , tramite la relazione di dispersione, determina l'autovalore dell'energia corrispondente $E \equiv E(k)$.

Si può poi costruire un pacchetto di funzioni di Bloch, la cui evoluzione nel tempo sarà data da:

$$\psi(x, t) = \int g(k)u(x; k)e^{i[kx + \phi(k) - \frac{E(k)}{\hbar}t]} dk$$

Procedendo allo stesso modo che con la particella libera, col metodo della fase stazionaria, si ottiene la legge oraria:

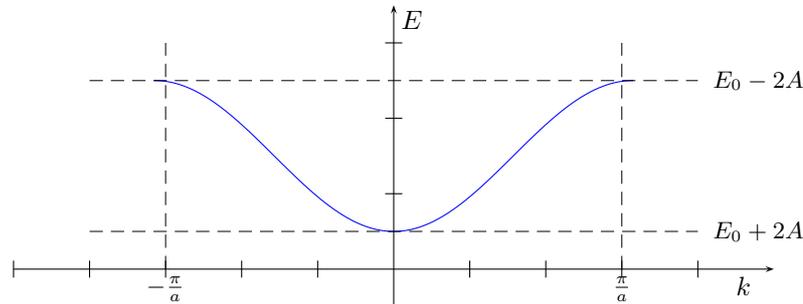
$$x(t, k) = \frac{dE(k)}{dk} \frac{1}{\hbar} t + \frac{d\phi(k)}{dk}$$

con velocità di gruppo data da $v_g = \frac{1}{\hbar} \frac{dE(k)}{dk}$. Si vede che la differenza col caso della particella libera è tutta contenuta nella relazione di dispersione $E(k)$, che è poi anche quella che definisce la banda.

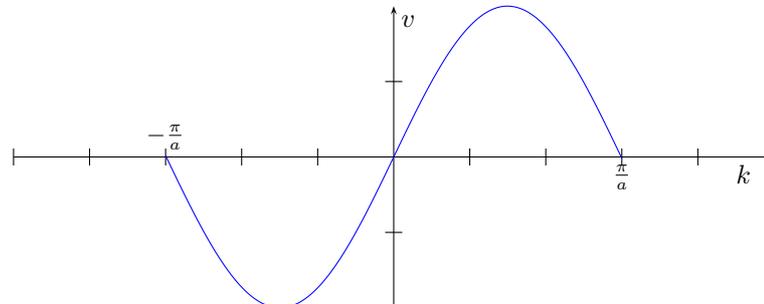
Se allora si costruisce il pacchetto con una $g(k)$ molto localizzata attorno ad un valore k_0 , scelto tra quelli permessi dalle condizioni al contorno, si può concludere che la velocità dello stato di banda corrispondente è data da:

$$v_g = \frac{1}{\hbar} \left. \frac{dE(k)}{dk} \right|_{k=k_0}$$

Per fissare le idee con un esempio, si può considerare la $E(k)$ trovata con l'approssimazione dell'elettrone fortemente legato (§4.4):



e dunque la velocità in funzione di k , cioè per ciascuno stato di banda con un certo valore di k ⁽³¹⁾ è $v(k) = 2A \sin ka$:



Dunque anche la velocità è una funzione periodica di k .

³⁰Si ricordi che questo significa che la funzione di Bloch è un'autofunzione dell'Hamiltoniana dell'elettrone in un reticolo.

³¹Scelto naturalmente tra quelli permessi dalle condizioni al contorno.

Questo concetto di velocità diventa particolarmente utile per quei valori di k per i quali la relazione di dispersione $E(k)$ è all'incirca a tangente orizzontale.³² Questo accade nei tre punti $k = 0, \pm\pi/a$.

Se si sviluppa la relazione di dispersione in serie intorno a questi punti e si tronca lo sviluppo al secondo ordine, si ottiene una relazione di dispersione quadratica in k , in quanto il termine del primo ordine, in cui compare la derivata prima, evidentemente si annulla:³³

$$E(k) = E_0 + \frac{1}{2} \left. \frac{d^2 E}{dk^2} \right|_{k=k_0} (k - k_0)^2 \quad k = 0, \pm\pi/a$$

Dunque per questi valori e nei loro pressi, la relazione di dispersione ha la stessa dipendenza da k di quella di particella libera:

$$E = \frac{\hbar^2 k^2}{2m}$$

Per portare fino in fondo questa analogia, si può assegnare all'elettrone una *massa efficace* m^* definita da:

$$m^* = \frac{1}{\hbar^2} \left. \frac{d^2 E}{dk^2} \right|_{k=k_0}$$

Nel modello dell'elettrone fortemente legato, la relazione di dispersione è tale che l'approssimazione è valida ai due bordi e al centro della prima zona di Brillouin. Al centro la massa efficace è positiva, ai bordi è negativa. Allontanandosi da queste zone, il concetto di massa efficace inizia a perdere di significato.

Se si passa in tre dimensioni, non è detto che gli estremi della banda si abbiano per valori al bordo della zona (cella) di Brillouin.³⁴ Nei sistemi tridimensionali può succedere anche che i coefficienti dello sviluppo di $E(k)$ siano diversi per le tre direzioni. In tal caso si deve parlare piuttosto di *tensore di massa efficace* e quindi si dovranno usare masse efficaci diverse in direzioni diverse.

Una volta introdotto – nelle zone in cui è possibile – il concetto di massa efficace, si può usare un'approssimazione semi-classica e descrivere l'elettrone effettivamente come una particella classica. Sotto certi aspetti, si può supporre che l'Hamiltoniana (classica) di questa particella sia la funzione $E(k)$, funzione che esprime la forma di banda:

$$\bar{H} = E(k) = E_0 + \frac{1}{2} \left. \frac{d^2 E}{dk^2} \right|_{k=k_0} (k - k_0)^2 \quad k = 0, \pm\pi/a$$

Si segue dunque l'analogia con la particella libera usando un formalismo classico. Se questa è vista come una Hamiltoniana, il secondo termine è un'energia cinetica a patto di considerare la quantità $\hbar k$ come un momento cinetico. In fisica dello stato solido questo viene chiamato *quasi-momento*.³⁵

4.5.3 Comportamento in presenza di campo elettrico

Se si tratta l'elettrone come una particella classica con momento cinetico $\hbar k$ e la cui Hamiltoniana è la forma di banda, si può capire in maniera semplice cosa accade in presenza di un campo elettrico esterno e quindi ci si può occupare del trasporto di carica all'interno dei solidi.

La presenza di un campo elettrico esterno \mathcal{E} si può rappresentare aggiungendo un termine all'Hamiltoniana del sistema che diventa:

$$\bar{H} = E(\hbar k) + e\mathcal{E}x$$

³²Per i quali, fra l'altro, la funzione è circa parabolica rispetto a k .

³³Questo vale per la banda del modello dell'elettrone fortemente legato in una dimensione, per il quale c'è appunto una sola banda. Se si prende il modello di Kronig–Penney c'è un susseguirsi di bande ed in questi punti k_0 c'è un susseguirsi di massimi e minimi. Se poi si passa in tre dimensioni, le cose si complicano ulteriormente, ma il discorso è sostanzialmente lo stesso.

³⁴Per esempio, per il silicio si hanno 6 minimi in 6 punti particolari dello spazio k .

³⁵In effetti, quella scritta è la prima equazione di Hamilton–Jacobi del sistema, perché fornisce la $\frac{dx}{dt}$ del sistema, ovvero la velocità.

questa è la prima equazione di Hamilton–Jacobi, che fornisce $\frac{dx}{dt}$. La seconda equazione di Hamilton–Jacobi, che fornisce $\frac{dp}{dt}$, ovvero in questo caso $\frac{dk}{dt}$, è:

$$\hbar \frac{dk}{dt} = -e\mathcal{E}$$

Quest’ultima è un’equazione differenziale in $k(t)$, che si risolve banalmente integrando i due membri, fornendo:

$$k(t) = -\frac{e\mathcal{E}}{\hbar}t$$

Dunque se si applica un campo elettrico al cristallo e vale l’approssimazione semi-classica, k cresce linearmente nel tempo. Siccome il valore di $k = \pi/a$ è del tutto equivalente al valore $k = -\pi/a$, quando k crescendo (o decrescendo, a seconda del segno della carica) arriva al bordo della (prima) zona di Brillouin, effettua una riflessione alla Bragg riapparendo dall’altra parte della zona di Brillouin e ricominciando a spazzarla.

Si noti però che man mano che il valore di k cambia, questo assume valori per i quali perde di significato la massa efficace. Inoltre si ha un comportamento strano dell’Hamiltoniana, che non è più quadratica, ma contiene anche un termine lineare in p .

Tuttavia, per una sorta di continuità, si può continuare ad adoperare l’analogia con una particella libera semi-classica. Si noti anche che la definizione di velocità di questa particella libera non dà problemi, perché rimane invece definita per ogni valore di k . In questi casi si può quindi parlare di pacchetto d’onda che si muove con una certa velocità.

4.5.4 Riempimento delle bande

Occorre ora far entrare in gioco il Principio di Pauli. Uno stato di banda può infatti essere occupato al più da due elettroni con spin opposto: quindi ogni valore di k è due volte degenero.

D’altra parte il numero di stati di banda, all’interno di ogni banda, è pari al numero di atomi che formano il reticolo³⁶. Allora, se ogni atomo contribuisce con un elettrone, la banda è semipiena perché ognuno dei possibili stati di Bloch è occupato da un solo elettrone. Se invece ogni atomo contribuisce con due elettroni, la banda è completamente piena.

Proposizione sul trasporto di corrente

Una banda completamente piena non può trasportare corrente e quindi *il trasporto di corrente può avvenire solo quando la banda è semipiena.*

Dimostrazione

Si è visto che in uno schema semi-classico l’applicazione di un campo elettrico ha l’effetto di far cambiare i valori di k di ogni stato di banda. Se si suppone che la catena di atomi sia indefinita e quindi ci siano infiniti stati di banda e se si applica un campo elettrico, accade che gli stati “saltano” l’uno nell’altro.

Si cercherà ora la velocità media degli stati nei tre casi di: 1) assenza di campo elettrico esterno, 2) campo elettrico con banda piena e 3) campo elettrico con banda semipiena.

Assenza di campo esterno

Se non c’è nessun campo elettrico esterno, a causa della simmetria della relazione di dispersione rispetto a k , indipendentemente dal fatto che la banda sia piena o semipiena, per ogni stato con un certo valore di k , ne esiste un altro col valore opposto di k .³⁷ Dunque, *in assenza di campo la velocità media è nulla.*

Banda piena

Se si applica il campo elettrico esterno ad un solido la cui banda è piena (cioè dove tutti gli stati permessi sono occupati), succede che ogni elettrone passa da uno stato al successivo ma la situazione

³⁶Lo si è visto in particolare nel modello dell’elettrone fortemente legato, §4.4.

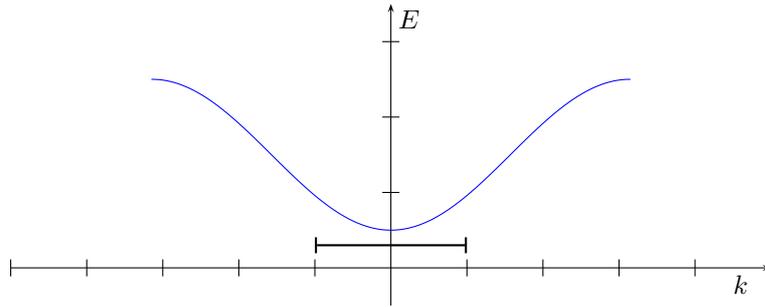
³⁷Bisogna aggiungere il fatto che gli stati di banda vengono occupati uniformemente partendo dalle energie più basse e salendo progressivamente.

rimane praticamente immutata e la velocità media rimane nulla (si ricordi che al bordo della banda l'elettrone riappare all'estremo opposto a causa della riflessione di Bragg). Si suppone naturalmente che il campo sia abbastanza debole da rendere impossibile il salto da uno stato di una banda ad uno stato di un'altra banda.

Dunque, *anche in caso di banda piena, la velocità media è nulla.*

Banda semipiena

In questo caso si ha un insieme di stati occupati, corrispondente ad un segmento della zona di Brillouin, e non a tutta la zona:



infatti, quando all'inizio non c'è ancora il campo elettrico, gli stati si dispongono al centro occupandoli a partire dalle energie più basse via via verso le più alte. Quando si accende il campo elettrico esterno, secondo il modello sviluppato qui il valore di k di ogni stato di banda si sposta e quindi questo segmento trasla fino al bordo della zona di Brillouin. Arrivato al bordo si "riflette" a causa della periodicità della relazione di dispersione (riflessione di Bragg).

Dunque *nel caso di banda semipiena e campo elettrico esterno, la velocità media non è nulla, ma oscilla perché la distribuzione dei valori di k occupati non è simmetrica.*

Questo modello in definitiva prevede che applicando un campo elettrico ad un reticolo cristallino con banda semipiena (conduttore) si ottiene una corrente alternata.

Si procederà quindi a stimare secondo il modello l'ordine di grandezza della frequenza di oscillazione di questa corrente alternata. Dall'equivalente della seconda equazione di Hamilton-Jacobi si è calcolata la $k(t)$, che è data da:

$$k(t) = -\frac{e\mathcal{E}}{\hbar}t$$

e quindi l'ordine di grandezza della frequenza della corrente alternata è dato da:

$$\nu \equiv \frac{1}{t} = -\frac{e\mathcal{E}}{\hbar}k$$

L'ordine di grandezza di \hbar è 10^{-15} eV·s, mentre quello di k è quello di $2\pi/a$ e quindi quello dell'inverso del raggio di Bohr, 10^8 cm $^{-1}$. Per $e\mathcal{E}$ si può considerare come ordine di grandezza 1. Si ottiene quindi una frequenza dell'ordine di 10^{-7} s, ovvero 10 MHz.

Questo non si osserva mai sperimentalmente, quello che si osserva invece è la cosiddetta legge di Ohm. Quindi il modello deve essere ritoccato.

4.5.5 Un modello più accurato

Il problema è che il modello studiato non sta tenendo conto delle deformazioni del reticolo, considerandolo al contrario come indeformabile, ovvero tale che la distanza tra gli atomi a sia costante lungo tutto il reticolo. In realtà il reticolo si deforma, e queste deformazioni sono di due tipi.

Deformazioni di primo tipo: sono deformazioni dovute ad imperfezioni del reticolo, sono quindi deformazioni statiche, ovvero costanti nel tempo.

Deformazioni di secondo tipo: sono deformazioni dovute alle vibrazioni termiche del reticolo, che sono quindi deformazioni variabili nel tempo.

Sugli elettroni l'effetto di entrambe queste deformazioni è lo stesso, se si ricorda che vale l'approssimazione adiabatica: questa approssimazione prevede infatti che gli elettroni seguano fedelmente i movimenti dei nuclei (ioni) attorno ai quali sono dislocati.

Una vibrazione del reticolo può inoltre sempre essere descritta con un modo normale di oscillazione, e quindi in definitiva come un fonone che viaggia lungo il reticolo (§4.3).

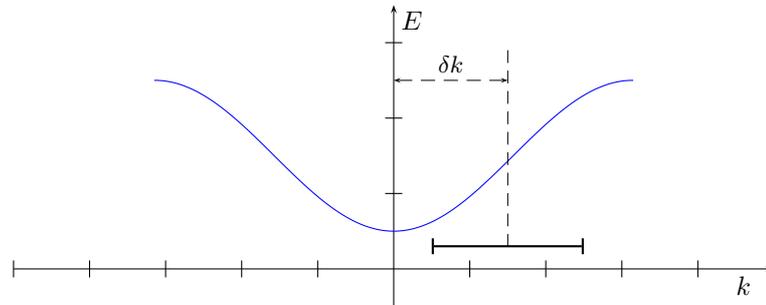
In entrambi i casi quando l'elettrone,³⁸ spostandosi nel reticolo, incontra una deformazione, avviene una sorta di "urto".

Se l'elettrone incontra una deformazione del reticolo dovuta alla vibrazione (termica o altro) si può descrivere la situazione come un urto tra lo stato di banda e un fonone. Si tratta in questo caso di una interazione in cui viene scambiata dell'energia, e lo stato di banda perde velocità (un fonone, alla stregua di un elettrone, trasporta e scambia energia). Dunque, è schematizzabile come un urto perfettamente elastico.

La deformazione statica dovuta ad una imperfezione del reticolo (casuale o ottenuta artificialmente con un *drogaggio*) è invece come un ostacolo rigido, contro cui l'elettrone ha un urto anelastico, che non fa scambiare energia ma momento cinetico.

A causa di questi fenomeni dissipativi, dopo un tempo caratteristico τ (molto breve) il segmento di valori di k popolati, sull'asse k del grafico della relazione di dispersione, si ferma e quindi la velocità media degli elettroni raggiunge un valore finale, che manterrà nel tempo.

Per calcolare la dislocazione δk del segmento:



si ricorre ancora all'espressione della velocità di spostamento di k :

$$\delta k = -\frac{e\mathcal{E}}{\hbar}\tau$$

In tre dimensioni invece di un segmento sull'asse k si ha una sfera nello spazio \vec{k} , detta *sfera di Fermi*. Per collisioni con impurezze si ha una variazione del solo k .

Ogni elettrone a regime, cioè dopo il transiente τ , ha variato il suo quasi-momento di $\hbar\delta k$. Per ottenere la velocità basta dividere questa quantità per la massa efficace. Poiché all'istante iniziale senza campo esterno la velocità (in media) era zero, questa rappresenta anche la velocità costante finale.

Se n è il numero di elettroni per unità di volume, la corrente è dunque:

$$j = n \frac{\hbar\delta k}{m^*} e$$

e sostituendo l'espressione di δk si ottiene:

$$j = n \frac{\hbar e \mathcal{E} \tau}{m^*} e = \frac{n \hbar e^2 \tau}{m^*} \mathcal{E} \equiv \sigma \mathcal{E}$$

che esprime piena coerenza con la legge di Ohm: proporzionalità tra corrente e campo applicato, tramite il fattore $\sigma = \frac{n \hbar e^2 \tau}{m^*}$.

³⁸In astratto, lo stato di banda.