

Chapitre 9

Electrons dans un cristal

9.1 Préliminaires

Avant d'en venir à la description quantitative du mouvement d'une particule (un électron) dans un champ de force périodique dérivant d'un potentiel, il convient de donner quelques éléments justifiant la pertinence de cette problématique pour l'étude des solides cristallisés.

Un solide, ordonné ou non, est un système fort complexe, même si l'on s'en tient aux interactions élémentaires les plus simples, celle de Coulomb par exemple. En effet, un solide contient des noyaux et des électrons, une telle distinction n'étant pas d'ailleurs la plus naturelle : on a vu que, selon les cas et la nature physique du solide, il était plus judicieux de le considérer comme un assemblage d'ions (solides ioniques ou métalliques) ou de molécules (solides moléculaires ou covalents).

En tout état de cause, la toute première approximation est évidemment celle de Born et Oppenheimer. La séparation entre noyaux (ou cœurs) et électrons étant faite, on en vient en outre à distinguer, le cas échéant, plusieurs sortes d'électrons : d'une part ceux qui sont dans les couches atomiques profondes, quasiment insensibles à ce qui se passe lors de la formation du solide (électrons de cœur) et ceux de la couche externe (complète ou incomplète) qui sont au contraire aux premières loges pour encaisser les altérations survenant lors de la formation du solide (électrons de valence¹). On peut considérer que cette distinction repose essentiellement sur un simple critère énergétique : les électrons de valence ont une énergie de liaison comparable, par nature, à celle d'une liaison chimique (1 eV ou une fraction d'eV), alors que les électrons internes sont très fortement liés (pour un élément de masse atomique même modeste, l'énergie de la couche (1s) est très vite² de l'ordre du keV).

À l'issue de ce dégrossissage, et en se plaçant dans la première étape de l'approximation de Born et Oppenheimer³, on reste *a priori* avec des ions et les électrons de valence, ces derniers étant soumis au potentiel plus ou moins fort des premiers. En plus du potentiel périodique du réseau, les électrons sont évidemment en interaction mutuelle, au moins par leur répulsion électrostatique.

On se doute – il suffit de se souvenir du cas de l'atome d'hélium – qu'il s'agit, tel quel, d'un problème redoutable, généralement appelé Problème à N -corps, que l'on ne sait résoudre qu'approximativement, à l'exception

¹En outre, pour un métal, une partie des électrons de valence à l'état atomique se retrouvent presque libres dans le métal : ceux-ci sont alors appelés électrons de conduction, en opposition avec ceux qui restent liés aux ions, que l'on continue à appeler électrons de valence.

²Pour un métal un peu lourd, comme le tungstène ou le molybdène, l'éjection d'un électron 1s donne lieu à réarrangement interne produisant les raies X notées K_α , K_β , etc.

³La deuxième étape consiste à étudier le mouvement des cœurs, qui donne lieu aux vibrations du solide et, dans l'hypothèse des petites oscillations, conduit à des oscillateurs harmoniques dont la quantification introduit naturellement les phonons du réseau – voir chapitre 10.

de quelques modèles simples se bornant à singer la répulsion électronique⁴. Dans une approche élémentaire, suivie ici, on décide d'oublier purement et simplement la répulsion de Coulomb entre électrons, ce qui, en se souvenant à nouveau du cas de l'hélium, peut paraître très grossier. En réalité, et de façon assez surprenante, nombreux sont les solides où l'approximation subséquente permet d'expliquer un très grand nombre de propriétés (pas toutes, évidemment) et, en particulier, d'énoncer les critères permettant de savoir pourquoi il existe des isolants et des conducteurs. La justification de ce petit miracle n'est pas aisée, mais il est possible d'imaginer que, en raison de la forte densité d'électrons, des phénomènes d'écrantage viennent réduire considérablement la portée et l'importance de la répulsion de Coulomb⁵ entre des électrons qui, de surcroît, disposent d'un vaste espace (le solide) et sont, du fait de cette délocalisation, relativement éloignés les uns des autres. Pour ces raisons, l'interaction *effective* est loin de ressembler à la répulsion de Coulomb nue.

Tout ceci étant admis, on reste – en ce qui concerne la première étape de Born - Oppenheimer – avec quelques électrons presque indépendants par atome, chacun d'entre eux étant soumis à un potentiel statique périodique, celui des ions immobiles situés aux nœuds d'un réseau de Bravais. Les électrons étant déclarés indépendants, une étape préliminaire consiste à étudier le mouvement de l'un d'entre eux. L'état (mécanique ou thermodynamique) d'une assemblée de N électrons s'en déduira par simple produit tensoriel des espaces d'états à une particule.

En définitive, dans le cadre ainsi précisé, l'équation aux valeurs propres à résoudre est celle associée au Hamiltonien d'un unique électron dans un potentiel périodique :

$$H = \frac{\vec{p}^2}{2m} + V(\vec{r}) , \quad V(\vec{r} + \vec{R}) = V(\vec{r}) \quad \forall \vec{R} \in \mathcal{B} . \quad (9.1)$$

On se doute que ceci n'est que le point de départ de nouvelles approximations, puisque la classe de potentiels $V(\vec{r})$ exactement solubles est fort réduite⁶. Avant d'en venir à quelques schémas approximatifs de portée universelle (voir plus loin), il est important d'énoncer des résultats exacts, ne reposant que sur l'existence de la symétrie de translation discrète du réseau, traduite par l'invariance de l'énergie potentielle V dans toute translation d'un vecteur $\vec{R} \in \mathcal{B}$.

9.2 Théorème de Bloch

La symétrie de translation caractéristique de V entraîne que le Hamiltonien (9.1) commute avec tous les opérateurs de translation discrète $\mathcal{T}(\vec{R})$ associés chacun à une translation géométrique $T(\vec{R})$. De ce seul fait résulte le théorème de Bloch, dont l'énoncé est le suivant⁷ :

Pour un Hamiltonien périodique du type (9.1), toute fonction (état) propre $\psi(\vec{r})$ peut être mise sous la forme d'une onde plane modulée en amplitude par une fonction $u(\vec{r})$ ayant la symétrie de périodicité du réseau :

$$\forall \psi(\vec{r}), H\psi(\vec{r}) = \varepsilon\psi(\vec{r}), \quad \exists u(\vec{r}), \psi(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} u(\vec{r}), \quad u(\vec{r}) = u(\vec{r} + \vec{R}) \quad \forall \vec{R} \in \mathcal{B} . \quad (9.2)$$

⁴Le plus célèbre d'entre eux est sans doute le modèle de Hubbard, dont la version sur réseau unidimensionnel (rigide) a été résolue en 1968 par Lieb et Wu. Si l'énoncé du problème est simple, sa résolution ne l'est pas.

⁵On a vu que la densité électronique est globalement caractérisée par la longueur r_s , définissant le rayon de la sphère occupée en moyenne par un électron. Cette longueur est de quelques unités atomiques pour un métal. Au contraire, quand $r_s \gg a_0$, les électrons sont quasiment immobiles, au sens où l'énergie cinétique est très petite devant l'énergie potentielle ; Wigner avait d'ailleurs prévu théoriquement que, dans cette situation, les électrons forment un réseau triangulaire (solide de Wigner). Ceci a été observé à la fin des années '80, par dépôt d'électrons à la surface de l'hélium liquide et par mesure de la modification des ondes capillaires.

⁶S'agissant d'exemple de potentiel périodique, celui de Kronig - Penney (un peigne de Dirac) est un grand classique.

⁷Des arguments similaires peuvent se transposer pour des Hamiltoniens périodiques en temps ; le même théorème est alors connu en Mathématiques sous le nom de *Théorème de Floquet*.

Par ailleurs, si les paramètres de maille du réseau tendent vers zéro (potentiel constant !), (9.2) dit que les fonctions propres d'une particule libre sont des ondes planes $e^{i\vec{k}\cdot\vec{r}}$: c'est l'invariance galiléenne.

Bien sûr, chaque fonction propre et chaque énergie propre doivent être indexées par un (des) nombre(s) quantique(s), collectivement notés n , associés aux observables qui, avec H , forment un ECOG. En outre, l'apparition du vecteur \vec{k} – paramètre homogène à un vecteur d'onde dont la signification sera donnée plus loin –, conduit, quand il le faut, à écrire (9.2) plus précisément (notations traditionnelles) :

$$H\psi_{n\vec{k}}(\vec{r}) = \varepsilon_n(\vec{k}) \psi_{n\vec{k}}(\vec{r}) , \quad \psi_{n\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} u_{n\vec{k}}(\vec{r}) , \quad u_{n\vec{k}}(\vec{r}) = u_{n\vec{k}}(\vec{r} + \vec{R}) . \quad (9.3)$$

L'affirmation contenue dans le théorème de Bloch peut d'ailleurs être énoncée autrement. En effet, on a la suite d'égalités :

$$\psi_{n\vec{k}}(\vec{r} + \vec{R}) = e^{i\vec{k}\cdot(\vec{r} + \vec{R})} u_{n\vec{k}}(\vec{r} + \vec{R}) = e^{i\vec{k}\cdot\vec{R}} e^{i\vec{k}\cdot\vec{r}} u_{n\vec{k}}(\vec{r}) , \quad (9.4)$$

soit finalement :

$$\psi_{n\vec{k}}(\vec{r} + \vec{R}) = e^{i\vec{k}\cdot\vec{R}} \psi_{n\vec{k}}(\vec{r}) . \quad (9.5)$$

On utilisera selon les cas la forme (9.3) ou la forme (9.5).

Ce théorème peut être établi de plusieurs façons ; deux démonstrations sont données dans la suite, chacune d'entre elles ayant ses vertus et son utilité pour la référence ultérieure. Auparavant, il convient de noter que les deux relations (9.3) et (9.5) traduisent en fait une propriété de symétrie, ici de translation : le réseau se superpose à lui-même dans n'importe quelle translation de vecteur \vec{R} , de sorte que le module carré de la fonction d'onde doit être invariant :

$$|\psi_{n\vec{k}}(\vec{r} + \vec{R})|^2 = |\psi_{n\vec{k}}(\vec{r})|^2 . \quad (9.6)$$

Bien sûr – tout comme dans le cas de l'invariance de jauge –, la fonction elle-même n'est pas invariante et acquiert au plus un facteur de phase sous l'opération de symétrie⁸. C'est ce que dit, et précise, l'équation (9.5).

1^{ère} démonstration

Elle introduit explicitement les opérateurs $\mathcal{T}(\vec{R})$, définis précisément comme suit⁹ :

$$\forall \phi(\vec{r}), \mathcal{T}(\vec{R})\phi(\vec{r}) = \phi(\vec{r} - \vec{R}) , \quad (9.8)$$

où l'on reconnaît la définition usuelle de l'opérateur agissant dans l'espace des états, associé à une transformation géométrique effectuée dans l'espace réel \mathbb{R}^3 , ici une translation. Comme il s'agit d'opérations de symétrie, tous les $\mathcal{T}(\vec{R})$ commutent avec H donné par (9.1) :

$$\forall \vec{R} \in \mathcal{B} : [H, \mathcal{T}(\vec{R})] = 0 ; \quad (9.9)$$

le sens physique de cette relation est clair : elle signifie qu'un état propre $\psi(\vec{r})$ de H étant donné, son transformé $\mathcal{T}(\vec{R})\psi(\vec{r})$ est aussi état propre avec la *même* énergie¹⁰ (la moindre des choses). En effet, (9.9) dit que, pour toute fonction quelconque ϕ , on a $H\mathcal{T}(\vec{R})\phi = \mathcal{T}(\vec{R})H\phi$; choisissons pour ϕ un état propre de H , ψ , satisfaisant $H\psi = \varepsilon\psi$. Il vient alors¹¹ :

$$H(\mathcal{T}\psi) = \mathcal{T}H\psi = \mathcal{T}\varepsilon\psi = \varepsilon(\mathcal{T}\psi) ; \quad (9.10)$$

⁸Ici la phase est constante (elle est indépendante de la position \vec{r}) – dans le cas d'une transformation de jauge, au contraire, la phase est variable dans l'espace. Elle garde toutefois toute sa pertinence : $\psi_{\vec{k}}$ et $\psi_{\vec{k}'}$ représentent deux états physiques distincts, tant que la différence $\vec{k} - \vec{k}'$ n'est pas un vecteur du réseau réciproque (voir plus loin)

⁹La définition (9.8) de la fonction transformée par l'opérateur de translation est la plus naturelle, – c'est celle qui a été utilisée dans le cours de Mécanique quantique lors de l'étude de la symétrie – et c'est cette définition qui est retenue pour la suite. On trouve parfois dans la littérature la définition "duale" alternative :

$$\mathcal{T}(\vec{R})\phi(\vec{r}) = \phi(\vec{r} + \vec{R}) . \quad (9.7)$$

¹⁰On retiendra qu'un sous-espace dégénéré est *globalement* invariant dans les opérations de symétrie.

Pour un groupe de symétrie non abélien, les états propres de H ne sont pas forcément propres des opérateurs de symétrie : penser au groupe des rotations dans \mathbb{R}^3 dans le cas d'un champ central ; les harmoniques sphériques, propres du moment cinétique orbital, ne sont pas propres des opérateurs de rotation (L_{\pm} échange les Y_{lm} les unes dans les autres, et les L_u sont les générateurs des rotations). Toutefois, L_u ne fait pas changer la valeur du nombre quantique l .

¹¹les parenthèses sont redondantes algébriquement, mais sont utiles pour faire comprendre les choses.

ceci montre que la fonction $\mathcal{T}\psi$ est propre avec la valeur propre ε , tout comme l'est ψ .

De toute évidence, et en s'appuyant sur la géométrie élémentaire, l'ensemble des $\mathcal{T}(\vec{R})$ peut être muni d'une structure de groupe, au moyen de la relation de définition qui s'impose :

$$\mathcal{T}(\vec{R})\mathcal{T}(\vec{R}') = \mathcal{T}(\vec{R} + \vec{R}') . \quad (9.11)$$

L'inverse est défini comme :

$$\mathcal{T}^{-1}(\vec{R}) = \mathcal{T}(-\vec{R}) . \quad (9.12)$$

Ce groupe a ici la puissance de \mathbb{N} (translations *discrètes*) ; il est commutatif (abélien), puisque la somme $\vec{R} + \vec{R}'$ est commutative :

$$\mathcal{T}(\vec{R})\mathcal{T}(\vec{R}') = \mathcal{T}(\vec{R} + \vec{R}') = \mathcal{T}(\vec{R}' + \vec{R}) = \mathcal{T}(\vec{R}')\mathcal{T}(\vec{R}) \iff [\mathcal{T}(\vec{R}), \mathcal{T}(\vec{R}')] = 0 \quad (9.13)$$

Il en résulte que tous les \mathcal{T} ont des vecteurs propres communs. Par ailleurs, les $\mathcal{T}(\vec{R})$ sont des opérateurs unitaires, puisqu'ils forment un sous-ensemble discret des translations continues ; parmi ces dernières existent les translations infinitésimales, qui ne sauraient être anti-unitaires¹². En conséquence :

$$\mathcal{T}^\dagger(\vec{R}) = \mathcal{T}^{-1}(\vec{R}) = \mathcal{T}(-\vec{R}) . \quad (9.14)$$

Comme H et $\mathcal{T}(\vec{R})$ commutent, ils ont des vecteurs propres en commun. Soit donc une fonction propre¹³ $\psi(\vec{r})$ telle que les deux équations suivantes sont satisfaites :

$$H\psi(\vec{r}) = \varepsilon\psi(\vec{r}) , \quad (9.15)$$

$$\mathcal{T}(\vec{R})\psi(\vec{r}) = \tau(\vec{R})\psi(\vec{r}) \quad \forall \vec{R} \in \mathcal{B} , \quad (9.16)$$

où, tout naturellement, la valeur propre τ de $\mathcal{T}(\vec{R})$ dépend de \vec{R} . L'équation (9.15) n'est pas explicitement utile pour la suite : elle rappelle seulement qu'une fonction telle que ψ , commune à H et aux $\mathcal{T}(\vec{R})$, *existe*.

En conséquence de (9.16) et (9.11), on a :

$$\mathcal{T}(\vec{R})\mathcal{T}(\vec{R}')\psi(\vec{r}) = \tau(\vec{R})\tau(\vec{R}')\psi(\vec{r}) \quad (9.17)$$

et :

$$\mathcal{T}(\vec{R} + \vec{R}')\psi(\vec{r}) = \tau(\vec{R} + \vec{R}')\psi(\vec{r}) . \quad (9.18)$$

Les équations (9.17) et (9.18) sont vraies $\forall \psi$ également propre de H , donc pour une base complète. Il en résulte :

$$\tau(\vec{R} + \vec{R}') = \tau(\vec{R})\tau(\vec{R}') . \quad (9.19)$$

τ est donc une fonction exponentielle – d'ailleurs, en vertu de l'unitarité des \mathcal{T} toutes leurs valeurs propres sont de la forme $e^{i \times \text{phase réelle}}$, ce que l'on choisit d'écrire plus précisément, pour la commodité :

$$\tau(\vec{R}) = e^{-i\Phi(\vec{R})} , \quad (9.20)$$

où Φ est une fonction à valeurs réelles ; (9.19) s'écrit alors :

$$\Phi(\vec{R} + \vec{R}') = \Phi(\vec{R}) + \Phi(\vec{R}') ; \quad (9.21)$$

cette relation dit que Φ est une forme linéaire de \vec{R} . $\forall \vec{R}$, une telle forme linéaire peut toujours s'écrire en produit scalaire¹⁴ ; $\Phi(\vec{R})$ est donc de la forme $\vec{k} \cdot \vec{R}$, où \vec{k} est quelconque et, en définitive :

$$\tau(\vec{R}) = e^{-i\vec{k} \cdot \vec{R}} . \quad (9.22)$$

Compte tenu de ce résultat, la relation (9.16) s'écrit maintenant, $\forall \psi$ propre de H et des \mathcal{T} :

$$\mathcal{T}(\vec{R})\psi(\vec{r}) = e^{-i\vec{k} \cdot \vec{R}}\psi(\vec{r}) \iff \psi(\vec{r} - \vec{R}) = e^{-i\vec{k} \cdot \vec{R}}\psi(\vec{r}) . \quad (9.23)$$

Ceci n'est autre que la formulation (9.5) du théorème de Bloch, avec le vecteur $-\vec{R}$.

¹²Un théorème dû à Wigner affirme que l'on peut associer à toute opération de symétrie un opérateur qui est soit linéaire unitaire (produit scalaire invariant), soit anti-linéaire unitaire (produit scalaire changé en son complexe conjugué).

En outre, les générateurs des translations sont proportionnels à l'impulsion \vec{p} , qui est hermitique. Pour une translation finie de vecteur \vec{r}_0 , on a $\mathcal{T}(\vec{r}_0) = e^{(1/i\hbar)\vec{p} \cdot \vec{r}_0}$. Le développement de l'exponentielle, avec $\vec{p} = -i\hbar \vec{\nabla}$ montre que pour toute fonction entière $f(\vec{r})$ on a $\mathcal{T}(\vec{r}_0)f(\vec{r}) = f(\vec{r} - \vec{r}_0)$, en conformité avec (9.8).

¹³Pour la simplicité de l'écriture, on omet les indices ($n\vec{k}$) apparaissant dans les énoncés (9.3) et (9.5) du théorème de Bloch.

¹⁴C'est ainsi, par exemple, que l'on peut définir l'espace dual d'un espace vectoriel.

2^{ème} démonstration

La démonstration précédente est en un sens cantonnée à l'espace réel. Au contraire, celle qui suit se place d'emblée dans l'espace réciproque, exploite les propriétés de la transformée de Fourier du potentiel périodique $V(\vec{r})$ et détermine en fait la fonction propre $\psi(\vec{r})$ par sa transformée de Fourier $\phi(\vec{q})$.

Ici arrive une petite difficulté technique : la fonction ψ est étendue sur tout le cristal (en principe infini) et donc n'a pas de transformée de Fourier au sens usuel. Pour contourner cette difficulté, on introduit provisoirement un cristal de taille finie, construit pour la commodité sur un (très grand) nombre de mailles primitives. L'espace direct étant dès lors *fini*, la transformée de Fourier de ψ est une intégrale sur un domaine fini (mais "grand"), de sorte que la question de son existence ne se pose pas. Tout naturellement, on a envie de prendre des conditions aux limites franches, en disant que ψ s'annule strictement en-dehors de la grande boîte (l'électron ne peut sortir de la boîte) ; en fait, ces conditions se révèlent peu commodes, techniquement parlant¹⁵. Pour cette raison, on choisit des conditions *cycliques*, dites de Born - von Karman, partant de l'idée que, le système physique étant par hypothèse (implicite) bien conditionné, le détail de ce qui se passe à la frontière de la grande boîte est sans importance physique (les éventuelles différences entre un jeu de conditions aux limites et un autre se traduisent par des termes s'annulant dans la limite d'un système infiniment étendu – l'analogie ici de la "limite thermodynamique").

Un jeu de vecteurs primitifs de \mathcal{B} , \vec{a}_i ($i = 1, 2, 3$), étant effectivement choisi, ces conditions identifient le point¹⁶ \vec{r} et le point $\vec{r} + N_i \vec{a}_i$ pour chaque direction de \mathcal{B} ($i = 1, 2, 3, N_i \gg 1$). Ces deux points étant confondus (ce sont donc finalement les mêmes, l'espace étant replié sur lui-même), il en résulte :

$$f(\vec{r} + N_i \vec{a}_i) = f(\vec{r}) \quad \forall \text{ fonction } f \quad (i = 1, 2, 3; N_i \gg 1) . \quad (9.25)$$

Ceci étant admis, on voit que le vecteur \vec{k} apparaissant dans (9.5) est maintenant quantifié à toute petite échelle. En effet, en prenant pour f une onde plane, on a :

$$e^{i\vec{k} \cdot (\vec{r} + N_i \vec{a}_i)} = e^{i\vec{k} \cdot \vec{r}} , \quad (9.26)$$

soit :

$$e^{iN_i \vec{k} \cdot \vec{a}_i} = 1 . \quad (9.27)$$

Cette relation contraint les valeurs de \vec{k} . En développant \vec{k} sur des vecteurs primitifs \vec{b}_j du réseau réciproque, $\vec{k} = \kappa_1 \vec{b}_1 + \kappa_2 \vec{b}_2 + \kappa_3 \vec{b}_3$, et avec $\vec{a}_i \cdot \vec{b}_j = 2\pi \delta_{ij}$, on voit que les composantes κ_j satisfont $e^{2\pi i N_j \kappa_j} = 1$, ce qui impose :

$$\kappa_j = \text{entier relatif} \times \frac{1}{N_j} . \quad (9.28)$$

Le vecteur \vec{k} admet donc¹⁷ la décomposition :

$$\vec{k} = \sum_{j=1}^3 \frac{m_j}{N_j} \vec{b}_j \quad (m_j \in \mathbb{Z}) . \quad (9.29)$$

Comme toujours, l'application de conditions aux limites produit la quantification (les composantes de \vec{k} sont des multiples entiers de quantités finies bien définies) ; le vecteur d'onde \vec{k} est maintenant quantifié suivant une

¹⁵Pour une particule enfermée dans une grande boîte mais libre par ailleurs, les fonctions propres sont des sinus (penser au puits infini à une dimension centré sur l'origine). Quand on replie l'espace à la Born - von Karman, les fonctions propres sont des ondes planes, beaucoup plus agréables à manipuler.

À une dimension d'espace, le repliement transforme un segment ouvert de grande longueur en cercle de grand rayon.

¹⁶Noter que la cyclicité impose aussi :

$$\mathcal{T}(\vec{R} = N_i \vec{a}_i) = \mathbf{1} . \quad (9.24)$$

¹⁷À ce stade, \vec{k} décrit (presque) tout \mathbb{R}^3 (c'est-à-dire \mathbb{R}^3 discrétisé) – il n'y a pas d'autre restriction pour les k_j , donc pas de restriction sur les m_j relativement aux N_j . On verra par la suite que deux états caractérisés par \vec{k} et $\vec{k} + \vec{K}$ ($\vec{K} \in \mathcal{B}$) représentent en fait le même état physique, ce qui permettra d'épuiser tous les états physiques distincts en cantonnant \vec{k} à une maille primitive de \mathcal{B} , soit en se bornant (par exemple) à $1 \leq m_j \leq N_j$.

grille de dimensions linéaires $\frac{1}{N_j}$; chaque volume élémentaire est une image en miniature d'une maille primitive du réseau réciproque¹⁸ et a pour volume $\delta^3 k$:

$$\delta^3 k = \frac{1}{N_1 N_2 N_3} |\vec{b}_1 \cdot (\vec{b}_2 \times \vec{b}_3)| \equiv \frac{1}{N} |(\vec{b}_1, \vec{b}_2, \vec{b}_3)| = \frac{1}{N} \frac{(2\pi)^3}{v}, \quad (9.30)$$

v désignant toujours le volume d'une maille primitive de \mathcal{B} . En désignant¹⁹ par V le volume du cristal :

$$\delta^3 k = \frac{(2\pi)^3}{V}. \quad (9.31)$$

Le nombre des “gros points” ainsi définis situés à l'intérieur d'une maille primitive du réseau réciproque est égal à N ($= N_1 N_2 N_3$).

Maintenant, toute fonction $\psi(\vec{r})$ satisfaisant ces conditions cycliques peut être développée sur des ondes planes satisfaisant elles-mêmes ces conditions. On peut donc écrire *a priori* :

$$\psi(\vec{r}) = \sum_{\vec{q}} \phi(\vec{q}) e^{i\vec{q} \cdot \vec{r}}, \quad (9.32)$$

où $\phi(\vec{q})$ est la composante de Fourier de $\psi(\vec{r})$ pour la valeur \vec{q} et où la sommation (discrète) court sur tous les \vec{q} de la forme :

$$\vec{q} = \sum_{j=1}^3 \frac{m_j}{N_j} \vec{b}_j \quad (m_j \in \mathbb{Z}) \iff e^{i N_i \vec{q} \cdot \vec{a}_i} = 1. \quad (9.33)$$

m_j étant dans tout \mathbb{Z} , \vec{q} n'est pas borné et décrit “tout” \mathbb{R}^3 (et pas seulement une maille de $\tilde{\mathcal{B}}$), cet espace étant simplement ponctué par les petits volumes résultant des conditions cycliques ; en quelque sorte, la somme $\sum_{\vec{q}}$ dans (9.32) est un précurseur de l'intégrale (c'est presque une intégrale²⁰). Il n'y a pas de doute que l'expression (9.32) fournit une fonction satisfaisant les conditions cycliques²¹ ; avec (9.33), on peut écrire en effet :

$$\psi(\vec{r} + N_i \vec{a}_i) = \sum_{\vec{q}} \phi(\vec{q}) e^{i\vec{q} \cdot (\vec{r} + N_i \vec{a}_i)} = \sum_{\vec{q}} \phi(\vec{q}) e^{i\vec{q} \cdot \vec{r}} = \psi(\vec{r}). \quad (9.34)$$

L'objectif final est de trouver $\psi(\vec{r})$: selon (9.32), il est équivalent de trouver ses composantes de Fourier $\phi(\vec{q})$. Pour cela, exploitons maintenant la périodicité de $V(\vec{r})$. De ce fait (et pour un potentiel $V(\vec{r})$ suffisamment régulier), le développement suivant existe :

$$V(\vec{r}) = \sum_{\vec{K} \in \tilde{\mathcal{B}}} \mathcal{V}(\vec{K}) e^{-i\vec{K} \cdot \vec{r}}, \quad (9.35)$$

où la somme court maintenant sur le réseau réciproque. En effet, comme $e^{i\vec{K} \cdot (\vec{r} + \vec{R})} = e^{i\vec{K} \cdot \vec{r}}$, le second membre est bien l'expression d'une fonction périodique dans l'espace réel – ce n'est d'ailleurs rien d'autre qu'une série de Fourier spatiale. Il existe évidemment une relation inverse :

$$\mathcal{V}(\vec{K}) = \frac{1}{v} \int_{\text{maille élémentaire}} V(\vec{r}) e^{i\vec{K} \cdot \vec{r}} d^3 r, \quad (9.36)$$

l'intégration étant restreinte à une maille élémentaire de \mathcal{B} . En effet, on a :

$$\int_{\text{maille élémentaire}} V(\vec{r}) e^{i\vec{K} \cdot \vec{r}} d^3 r = \sum_{\vec{K}' \in \tilde{\mathcal{B}}} \mathcal{V}(\vec{K}') \int_{\text{maille élémentaire}} e^{i(\vec{K} - \vec{K}') \cdot \vec{r}} d^3 r. \quad (9.37)$$

¹⁸On peut dire que c'est un “gros point”.

¹⁹Télescopage de notations : ne pas confondre le volume du cristal avec le potentiel $V(\vec{r})$!

²⁰On peut même la considérer comme une somme de Darboux.

²¹Il est moins évident que toute fonction satisfaisant ces conditions peut ainsi être développée.

Par ailleurs $\vec{K} - \vec{K}'$ est de la forme $\sum_j m_j \vec{b}_j$ et $\vec{r} = \sum_i x_i \vec{a}_i$ avec $x_i \in [0, 1]$ puisque \vec{r} est restreint à une maille élémentaire ; l'intégrale dans (9.37) est donc :

$$\int_{\text{maille élémentaire}} e^{2\pi i \sum_j m_j x_j} d^3 r \quad (9.38)$$

et vaut 0 si dès qu'un seul m_j est différent de zéro, et v si tous les m_j sont nuls ; en définitive :

$$\int_{\text{maille élémentaire}} e^{i(\vec{K}-\vec{K}')\cdot\vec{r}} d^3 r = v \delta_{\vec{K} \vec{K}'} , \quad (9.39)$$

où $\delta_{\vec{K} \vec{K}'}$ est le symbole de Kronecker. Transposées aux variables espace - vecteur d'onde, les relations (9.35), (9.36) et (9.39) ne sont rien d'autre que ($n, n' \in \mathbb{Z}$) :

$$f(t) = \sum_{n \in \mathbb{Z}} f_n e^{-i n \omega t} , \quad f_n = \frac{\omega}{2\pi} \int_{\text{période}} f_n e^{i n \omega t} dt , \quad \int_{\text{période}} e^{i(n-n')\omega t} dt = 2\pi \delta_{nn'} , \quad (9.40)$$

pour le couple énergie - pulsation dans le cas d'une fonction de période $T = \frac{2\pi}{\omega}$. Notons qu'avec la définition (9.36), V et \mathcal{V} ont la même dimension physique, celle d'une énergie.

Avant de continuer, quelques remarques donnant lieu à des simplifications dans la suite :

- On peut toujours choisir

$$\mathcal{V}(\vec{K} = 0) = 0 , \quad (9.41)$$

puisque $V(\vec{r})$ est défini à une constante additive près. Adopter cette relation, c'est fixer à zéro la moyenne spatiale sur une maille primitive du potentiel périodique (cette convention sera utile pour l'approximation des électrons presque libres ; au stade actuel, elle n'est pas nécessaire).

- $V(\vec{r}) \in \mathbb{R} \iff \mathcal{V}(\vec{K}) = \mathcal{V}(-\vec{K})^*$
- si le cristal a la symétrie d'inversion d'espace ($V(-\vec{r}) = V(\vec{r})$), alors $\mathcal{V}(\vec{K})$ est aussi réel puisque l'intégrale $\int d^3 r V(\vec{r}) \sin(\vec{K}\cdot\vec{r})$, possédant un intégrand impair, est nulle. En pareil cas :

$$\mathcal{V}(-\vec{K}) = \mathcal{V}(\vec{K}) = \mathcal{V}(\vec{K})^* . \quad (9.42)$$

En reportant les développements (9.32) et (9.35) dans l'équation aux valeurs propres écrite en représentation- q :

$$\left[\frac{\vec{p}^2}{2m} + V(\vec{r}) \right] \psi(\vec{r}) = \varepsilon \psi(\vec{r}) , \quad (9.43)$$

on obtient deux types de termes ($\vec{p} = -i\hbar\vec{\nabla}_{\vec{r}}$) :

$$\frac{\vec{p}^2}{2m} \sum_{\vec{q}} \phi(\vec{q}) e^{i\vec{q}\cdot\vec{r}} = -\frac{\hbar^2}{2m} \sum_{\vec{q}} (i\vec{q})^2 \phi(\vec{q}) e^{i\vec{q}\cdot\vec{r}} , \quad (9.44)$$

et :

$$V(\vec{r}) \sum_{\vec{q}} \phi(\vec{q}) e^{i\vec{q}\cdot\vec{r}} = \sum_{\vec{q}} \sum_{\vec{K}} \mathcal{V}(\vec{K}) \phi(\vec{q}) e^{i\vec{q}\cdot\vec{r}} e^{-i\vec{K}\cdot\vec{r}} = \sum_{\vec{q}'} \sum_{\vec{K} \in \vec{\mathcal{B}}} \mathcal{V}(\vec{K}) \phi(\vec{q}' + \vec{K}) e^{i\vec{q}'\cdot\vec{r}} . \quad (9.45)$$

En identifiant maintenant dans les deux membres de l'équation aux valeurs propres (9.43) le coefficient de $e^{i\vec{q}\cdot\vec{r}}$, on obtient :

$$+\frac{\hbar^2 \vec{q}^2}{2m} \phi(\vec{q}) + \sum_{\vec{K} \in \vec{\mathcal{B}}} \mathcal{V}(\vec{K}) \phi(\vec{q} + \vec{K}) = \varepsilon \phi(\vec{q}) \iff [\varepsilon - \varepsilon^{(0)}(\vec{q})] \phi(\vec{q}) = \sum_{\vec{K} \in \vec{\mathcal{B}}} \mathcal{V}(\vec{K}) \phi(\vec{q} + \vec{K}) . \quad (9.46)$$

où on a posé :

$$\varepsilon^{(0)}(\vec{q}) = \frac{\hbar^2 \vec{q}^2}{2m} . \quad (9.47)$$

Par inspection, on voit que le système linéaire²² (9.46) couple exclusivement entre elles les composantes $\phi(\vec{q})$ ne différant que d'un vecteur \vec{K} appartenant à $\tilde{\mathcal{B}}$. Autrement dit, \vec{q} étant confiné dans une maille primitive de $\tilde{\mathcal{B}}$, seuls les vecteurs de la forme $\vec{q} + \vec{K}$, \vec{K} quelconque²³ dans $\tilde{\mathcal{B}}$, apparaissent dans la sommation de l'expression (9.32). En désignant par \vec{k} une valeur particulière pour \vec{q} , on obtient une fonction ψ paramétrée par \vec{k} , que l'on note $\psi_{\vec{k}}(\vec{r})$ et dont le développement est de la forme :

$$\psi_{\vec{k}}(\vec{r}) = \sum_{\vec{K} \in \tilde{\mathcal{B}}} \phi(\vec{k} + \vec{K}) e^{i(\vec{k} + \vec{K}) \cdot \vec{r}} = e^{i\vec{k} \cdot \vec{r}} \sum_{\vec{K} \in \tilde{\mathcal{B}}} \phi(\vec{k} + \vec{K}) e^{i\vec{K} \cdot \vec{r}} . \quad (9.48)$$

À la réflexion, il n'est pas surprenant que la somme sur \vec{q} ($= \vec{k}$) donnant le développement de Fourier de $\psi_{\vec{k}}$ ne contienne que des vecteurs d'onde équivalents par translation de $\tilde{\mathcal{B}}$. \vec{k} est un bon nombre quantique et H , ou l'un de ses avatars, ne peut avoir d'éléments de matrice entre deux états associés à deux \vec{k} inéquivalents, donnant deux valeurs propres distinctes pour un même (et quelconque) opérateur de translation.

Maintenant, la somme située le plus à droite dans (9.48) est visiblement une fonction de \vec{r} ayant la symétrie de translation du réseau direct, puisque $e^{i\vec{K} \cdot (\vec{R} + \vec{r})} = e^{i\vec{K} \cdot \vec{r}}$: elle n'est rien d'autre que la fonction $u_{\vec{k}}$ introduite dans la formulation (9.2) du théorème de Bloch ; (9.48) peut ainsi se récrire :

$$\psi_{\vec{k}}(\vec{r}) = e^{i\vec{k} \cdot \vec{r}} u_{\vec{k}}(\vec{r}) , \quad u_{\vec{k}}(\vec{r}) = \sum_{\vec{K} \in \tilde{\mathcal{B}}} \phi(\vec{k} + \vec{K}) e^{i\vec{K} \cdot \vec{r}} = u_{\vec{k}}(\vec{r} + \vec{R}) \quad \forall \vec{R} \in \mathcal{B} ; \quad (9.49)$$

ainsi s'achève la deuxième démonstration de ce théorème. Par ailleurs, l'écriture de droite montre que :

$$u_{\vec{k} + \vec{K}_0}(\vec{r}) = \sum_{\vec{K} \in \tilde{\mathcal{B}}} \phi(\vec{k} + \vec{K}_0 + \vec{K}) e^{i\vec{K} \cdot \vec{r}} = \sum_{\vec{K}' \in \tilde{\mathcal{B}}} \phi(\vec{k} + \vec{K}') e^{i(\vec{K}' - \vec{K}_0) \cdot \vec{r}} = e^{-i\vec{K}_0 \cdot \vec{r}} u_{\vec{k}}(\vec{r}) \quad \forall \vec{K}_0 \in \tilde{\mathcal{B}} . \quad (9.50)$$

9.3 Premières conséquences du théorème de Bloch

9.3.1 Première zone de Brillouin

Quand on adopte les conditions cycliques de Born - von Karman, les valeurs de \vec{k} sont de la forme (9.33) :

$$\vec{k} = \sum_{j=1}^3 \frac{n_j}{N_j} \vec{b}_j \quad (n_j \text{ entier}) . \quad (9.51)$$

Le cristal est donc de taille finie, boîte parallélépipédique construite sur les trois vecteurs $N_i \vec{a}_i$ ($N_i \gg 1$). Ceci étant, il n'y a que $N = N_1 N_2 N_3$ valeurs distinctes possibles pour \vec{k} . En effet, si n_j est plus grand que N_j , ceci veut dire qu'il existe deux entiers p_j et r_j tels que :

$$n_j = p_j N_j + r_j \quad (r_j < N_j) . \quad (9.52)$$

Reporté dans (9.51), ceci montre que \vec{k} est de la forme :

$$\vec{k} = \sum_{j=1}^3 \frac{p_j N_j + r_j}{N_j} \vec{b}_j = \sum_{j=1}^3 \left(p_j + \frac{r_j}{N_j} \right) \vec{b}_j \equiv \vec{K} + \vec{k}_1 , \quad (9.53)$$

\vec{K} désignant toujours un vecteur du réseau réciproque. Comparons maintenant $\psi_{\vec{k}}$ et $\psi_{\vec{k}_1}$. Comme $e^{i\vec{K} \cdot \vec{R}} = 1$, on a :

$$\psi_{\vec{k}}(\vec{r} + \vec{R}) = e^{i\vec{k} \cdot \vec{R}} \psi_{\vec{k}}(\vec{r}) = e^{i(\vec{k}_1 + \vec{K}) \cdot \vec{R}} \psi_{\vec{k}}(\vec{R}) = e^{i\vec{k}_1 \cdot \vec{R}} \psi_{\vec{k}}(\vec{r}) . \quad (9.54)$$

²²Il s'agit bien d'un *système* : il existe une telle équation pour chaque valeur de \vec{q} choisie sur la petite grille plongée dans \mathbb{R}^3 définie plus haut.

²³à l'exception de $\vec{K} = 0$ en vertu de (9.41).

Par ailleurs, on a toujours :

$$\psi_{\vec{k}_1}(\vec{r} + \vec{R}) = e^{i\vec{k}_1 \cdot \vec{R}} \psi_{\vec{k}_1}(\vec{r}) , \quad (9.55)$$

de sorte que $\psi_{\vec{k}}$ et $\psi_{\vec{k}_1}$ satisfont la même relation de déphasage donnée par le théorème de Bloch : c'est ceci qui permet de les identifier. D'ailleurs, le développement (9.48) montre clairement que $\psi_{\vec{k}}$ et $\psi_{\vec{k} + \vec{K}_0}$ sont identiques ; il suffit de en effet faire un changement de variable muette de sommation pour montrer que les deux expressions de $\psi_{\vec{k}}$ et $\psi_{\vec{k} + \vec{K}_0}$ coïncident strictement :

$$\psi_{\vec{k} + \vec{K}_0}(\vec{r}) = \sum_{\vec{K} \in \tilde{\mathcal{B}}} \phi(\vec{k} + \vec{K}_0 + \vec{K}) e^{i(\vec{k} + \vec{K}_0 + \vec{K}) \cdot \vec{r}} = \sum_{\vec{K}' \in \tilde{\mathcal{B}}} \phi(\vec{k} + \vec{K}') e^{i(\vec{k} + \vec{K}') \cdot \vec{r}} \equiv \psi_{\vec{k}}(\vec{r}) . \quad (9.56)$$

Ceci se voit encore comme suit ; en vertu de (9.50), on a encore :

$$\psi_{\vec{k} + \vec{K}_0}(\vec{r}) = e^{i(\vec{k} + \vec{K}_0) \cdot \vec{r}} u_{\vec{k} + \vec{K}_0}(\vec{r}) = e^{i(\vec{k} + \vec{K}_0) \cdot \vec{r}} e^{-i\vec{K}_0 \cdot \vec{r}} u_{\vec{k}}(\vec{r}) = e^{i\vec{k} \cdot \vec{r}} u_{\vec{k}}(\vec{r}) \equiv \psi_{\vec{k}}(\vec{r}) . \quad (9.57)$$

Ainsi, $\psi_{\vec{k}}$ et $\psi_{\vec{k} + \vec{K}}$, $\vec{K} \in \tilde{\mathcal{B}}$ représentent un seul et même état physique : il n'y a pas lieu de distinguer ces deux fonctions²⁴. Cette propriété de périodicité de $\psi_{\vec{k}}$ se reporte évidemment sur la valeur propre associée $\varepsilon(\vec{k})$. On retiendra donc :

$$\psi_{\vec{k} + \vec{K}}(\vec{r}) = \psi_{\vec{k}}(\vec{r}) , \quad \varepsilon(\vec{k} + \vec{K}) = \varepsilon(\vec{k}) \quad \forall \vec{K} \in \tilde{\mathcal{B}} . \quad (9.58)$$

D'où le résultat important : le vecteur d'onde \vec{k} apparaissant dans les fonctions de Bloch peut être toujours choisi à l'intérieur d'une maille primitive du réseau réciproque. L'usage est de choisir la maille de Wigner - Seitz de $\tilde{\mathcal{B}}$, traditionnellement alors appelée 1^{ère} zone de Brillouin, en abrégé BZ1.

9.3.2 Équation pour la fonction $u_{\vec{k}}$

Toute fonction propre d'un Hamiltonien périodique dans l'espace peut donc être cherchée sous la forme d'une fonction de Bloch :

$$\psi_{\vec{k}}(\vec{r}) = e^{i\vec{k} \cdot \vec{r}} u_{\vec{k}}(\vec{r}) , \quad (9.59)$$

avec :

$$u_{\vec{k}}(\vec{r} + \vec{R}) = u_{\vec{k}}(\vec{r}) . \quad (9.60)$$

Tout revient donc à trouver la fonction $u_{\vec{k}}(\vec{r})$; il est facile d'écrire l'équation qu'elle doit satisfaire. Partant de :

$$\left[-\frac{\hbar^2}{2m} \Delta + V(\vec{r}) \right] \psi_{\vec{k}}(\vec{r}) = \varepsilon(\vec{k}) \psi_{\vec{k}}(\vec{r}) , \quad (9.61)$$

et injectant la forme (9.59), on trouve :

$$-\frac{\hbar^2}{2m} \left[(i\vec{k})^2 u_{\vec{k}} + 2i\vec{k} \cdot \vec{\nabla} u_{\vec{k}} + \Delta u_{\vec{k}} \right] + V(\vec{r}) u_{\vec{k}}(\vec{r}) = \varepsilon(\vec{k}) u_{\vec{k}}(\vec{r}) , \quad (9.62)$$

soit :

$$\frac{\hbar^2}{2m} (-i\vec{\nabla} + \vec{k})^2 u_{\vec{k}}(\vec{r}) + V(\vec{r}) u_{\vec{k}}(\vec{r}) = \varepsilon(\vec{k}) u_{\vec{k}}(\vec{r}) , \quad (9.63)$$

ou encore, puisque $-i\vec{\nabla} = \hbar^{-1}\vec{p}$:

$$\frac{1}{2m} (\vec{p} + \hbar\vec{k})^2 u_{\vec{k}}(\vec{r}) + V(\vec{r}) u_{\vec{k}}(\vec{r}) = \varepsilon(\vec{k}) u_{\vec{k}}(\vec{r}) , \quad (9.64)$$

En vertu de la périodicité de $u_{\vec{k}}$, (9.60), il suffit de résoudre (9.63) à l'intérieur d'une maille primitive de \mathcal{B} . Comme toujours, il convient de préciser des conditions aux limites sur la frontière du domaine de résolution, faute de quoi, l'équation aux dérivées partielles possédant une infinité de solutions, le problème serait mal posé physiquement. On sait que toute fonction propre doit être continue et à dérivée continue tant que le potentiel

²⁴Tout comme pour une fonction 2π -périodique $f(\theta)$, les deux fonctions $f(\theta)$ et $f(\theta + n2\pi)$ ($n \in \mathbb{Z}$) sont en fait une seule et même fonction.

ne présente pas de saut d'amplitude infinie, ce que l'on suppose. Il en résulte qu'ici les conditions aux limites sont :

$$u_{\vec{k}} \text{ et } \vec{\nabla} u_{\vec{k}} \text{ continues sur la frontière de la maille primitive .} \quad (9.65)$$

Comme toujours, l'application de conditions aux limites produit la quantification des valeurs propres ε , désormais affublées de l'indice discret n , tout comme les fonctions propres. Une notation précise est donc :

$$H_u(\vec{r}, \vec{p}; \vec{k}) u_{n\vec{k}}(\vec{r}) = \varepsilon_n(\vec{k}) u_{n\vec{k}}(\vec{r}) , \quad (9.66)$$

avec, d'après (9.64) :

$$H_u(\vec{r}, \vec{p}; \vec{k}) = \frac{1}{2m} (\vec{p} + \hbar\vec{k})^2 + V(\vec{r}) . \quad (9.67)$$

Les valeurs propres $\varepsilon_n(\vec{k})$ dépendent évidemment de \vec{k} ; pour chaque valeur de n , l'ensemble des valeurs de \vec{k} formant un ensemble quasi-continu, $\varepsilon_n(\vec{k})$ varie quasi-continûment et définit une *bande d'énergie* s'étendant de $\inf_{\vec{k}} \varepsilon_n$ à $\sup_{\vec{k}} \varepsilon_n$.

La caractéristique essentielle de cette bande est sa loi de dispersion, définie par la variation de ε_n avec \vec{k} . L'existence de ces bandes d'énergie est clairement une *conséquence* directe de la symétrie de translation. Noter un fait majeur : la fonction $\varepsilon_n(\vec{k})$ est ce qu'elle est et, *a priori*, peut être fort différente de la fonction associée à l'électron libre $\varepsilon^{(0)}(\vec{k}) = \frac{\hbar^2 \vec{k}^2}{2m}$. On verra que, même quand le potentiel de réseau $V(\vec{r})$ est petit, la fonction $\varepsilon_n(\vec{k})$ possède des propriétés qui fait la démarque *qualitativement* de $\varepsilon^{(0)}(\vec{k})$ (en un sens précisé ultérieurement, elle a des "trous").

Par ailleurs, comme on a été conduit à identifier deux états \vec{k} et $\vec{k} + \vec{K}$ ne différant que par un vecteur de $\tilde{\mathcal{B}}$, les valeurs de $\varepsilon_n(\vec{k})$ et de $\varepsilon_n(\vec{k} + \vec{K})$ coïncident ; $\psi_{\vec{k}}$ et $\psi_{\vec{k} + \vec{K}}$ sont les mêmes, les valeurs propres sont donc aussi les mêmes. Les fonctions $\varepsilon_n(\vec{k})$ sont donc invariantes dans toute translation du réseau réciproque :

$$\varepsilon_n(\vec{k} + \vec{K}) = \varepsilon_n(\vec{k}) \quad \forall \vec{K} \in \tilde{\mathcal{B}} . \quad (9.68)$$

9.3.3 Impulsion de l'électron dans le cristal

Le vecteur $\hbar\vec{k}$ apparaissant dans le théorème de Bloch ressemble indéniablement à une impulsion au sens usuel. Il ne s'agit cependant pas de la vraie impulsion de l'électron, quoique la terminologie traditionnelle l'appelle "impulsion de l'électron dans le cristal". En effet, la véritable impulsion, notée \vec{p} comme d'habitude, est représentée par l'opérateur $-i\hbar\vec{\nabla}$; ce n'est pas une constante du mouvement ($[H, \vec{p}] \neq 0$), puisque l'électron est soumis à une force donnée par l'opposé du gradient du potentiel périodique. Au contraire, \vec{k} est un bon "nombre" quantique, qui, en tant que tel, étiquette correctement les états propres d'un électron.

Plus précisément, l'équation de Heisenberg pour \vec{p} est²⁵ :

$$i\hbar \frac{d}{dt} \vec{p} = [\vec{p}, H] = -i\hbar \vec{\nabla} V \quad \iff \quad \frac{d}{dt} \vec{p} = [\vec{p}, H] = -\vec{\nabla} V . \quad (9.70)$$

D'ailleurs, $\psi_{n\vec{k}}$ n'est pas propre de \vec{p} :

$$\vec{p} \psi_{n\vec{k}} = -i\hbar \vec{\nabla} e^{i\vec{k}\cdot\vec{r}} u_{n\vec{k}}(\vec{r}) = -i\hbar [i\vec{k} u_{n\vec{k}} + (\vec{\nabla} u_{n\vec{k}})] e^{i\vec{k}\cdot\vec{r}} = \hbar\vec{k} \psi_{n\vec{k}} - i\hbar e^{i\vec{k}\cdot\vec{r}} (\vec{\nabla} e^{-i\vec{k}\cdot\vec{r}} \psi_{n\vec{k}}) . \quad (9.71)$$

Le second terme n'a aucune raison d'être proportionnel à $\psi_{n\vec{k}}$ – il ne disparaît que si $u_{n\vec{k}}$ est constante, c'est-à-dire pour un électron *libre*. Enfin, comme le vecteur \vec{k} est finalement défini à un vecteur \vec{K} près (puisque, par exemple, $\varepsilon_n(\vec{k})$ a la symétrie de translation décrite par (9.68)), il existe un arbitraire pour \vec{k} ; à lui seul, cet arbitraire démontre le fait que $\hbar\vec{k}$ ne peut être une impulsion physique au sens usuel.

²⁵En laissant tomber l'indice H pour simplifier les notations ; avec les notations précises :

$$\frac{d}{dt} \vec{p}_H = -(\vec{\nabla} V)_H . \quad (9.69)$$

9.3.4 Vitesse d'un électron de Bloch

On appelle désormais électron de Bloch un électron décrit par une fonction d'onde de Bloch, soit $\psi_{n\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} u_{n\vec{k}}(\vec{r})$. On va maintenant établir la relation suivante, remarquable et très simple, entre la vitesse moyenne d'un tel électron et l'énergie propre associée à $\psi_{n\vec{k}}$:

$$\langle \vec{v} \rangle_{n\vec{k}} \equiv \langle \psi_{n\vec{k}} | \frac{\vec{p}}{m} | \psi_{n\vec{k}} \rangle = \hbar^{-1} \vec{\nabla}_{\vec{k}} \varepsilon_n \equiv \vec{\nabla}_{\hbar\vec{k}} \varepsilon_n . \quad (9.72)$$

Pour établir ce résultat, on utilise le théorème de Hellmann - Feynman, qui stipule que pour un Hamiltonien quelconque $H(\lambda)$ dépendant d'un paramètre λ et avec :

$$H(\lambda) \psi(\lambda) = E(\lambda) \psi(\lambda) \quad \langle \psi(\lambda) | \psi(\lambda) \rangle = 1 , \quad (9.73)$$

on a :

$$\frac{dE}{d\lambda} = \langle \psi(\lambda) | \frac{\partial H}{\partial \lambda} | \psi(\lambda) \rangle . \quad (9.74)$$

En effet :

$$\begin{aligned} \frac{d}{d\lambda} \langle \psi(\lambda) | H(\lambda) | \psi(\lambda) \rangle &= \langle \frac{\partial \psi}{\partial \lambda} | H | \psi(\lambda) \rangle + \langle \psi(\lambda) | \frac{\partial H}{\partial \lambda} | \psi(\lambda) \rangle + \langle \psi(\lambda) | H(\lambda) | \frac{\partial \psi}{\partial \lambda} \rangle \\ &= E(\lambda) \left(\langle \frac{\partial \psi}{\partial \lambda} | \psi(\lambda) \rangle + \langle \psi(\lambda) | \frac{\partial \psi}{\partial \lambda} \rangle \right) + \langle \psi(\lambda) | \frac{\partial H}{\partial \lambda} | \psi(\lambda) \rangle . \end{aligned} \quad (9.75)$$

La parenthèse est nulle en vertu de la normalisation de ψ , et (9.74) en résulte.

Ici, le Hamiltonien pour la fonction $u_{\vec{k}}$ est H_u donné par l'expression (9.67) ; il dépend du vecteur \vec{k} . Le théorème de Hellmann - Feynman s'écrit ici :

$$\vec{\nabla}_{\vec{k}} \varepsilon_n(\vec{k}) = \langle u_{n\vec{k}} | \vec{\nabla}_{\vec{k}} H_u | u_{n\vec{k}} \rangle = \langle u_{n\vec{k}} | \frac{\hbar}{m} (\vec{p} + \hbar\vec{k}) | u_{n\vec{k}} \rangle = \frac{\hbar^2}{m} \langle u_{n\vec{k}} | -i\vec{\nabla}_{\vec{r}} + \vec{k} | u_{n\vec{k}} \rangle . \quad (9.76)$$

On a :

$$\begin{aligned} (-i\vec{\nabla}_{\vec{r}} + \vec{k}) u_{n\vec{k}} &= (-i\vec{\nabla}_{\vec{r}} + \vec{k}) e^{-i\vec{k}\cdot\vec{r}} \psi_{n\vec{k}}(\vec{r}) = (-i)(-i\vec{k}) e^{-i\vec{k}\cdot\vec{r}} \psi_{n\vec{k}} - i e^{-i\vec{k}\cdot\vec{r}} \vec{\nabla}_{\vec{r}} \psi_{n\vec{k}} + \vec{k} e^{-i\vec{k}\cdot\vec{r}} \psi_{n\vec{k}} \\ &= -i e^{-i\vec{k}\cdot\vec{r}} \vec{\nabla}_{\vec{r}} \psi_{n\vec{k}} . \end{aligned} \quad (9.77)$$

D'où :

$$\vec{\nabla}_{\vec{k}} \varepsilon_n(\vec{k}) = \frac{\hbar^2}{m} (-i) \langle u_{n\vec{k}} | e^{-i\vec{k}\cdot\vec{r}} \vec{\nabla}_{\vec{r}} | \psi_{n\vec{k}} \rangle = -i \frac{\hbar^2}{m} \langle \psi_{n\vec{k}} | \vec{\nabla}_{\vec{r}} | \psi_{n\vec{k}} \rangle = \frac{\hbar}{m} \langle \psi_{n\vec{k}} | \vec{p} | \psi_{n\vec{k}} \rangle , \quad (9.78)$$

soit la relation annoncée :

$$\langle \psi_{n\vec{k}} | \frac{\vec{p}}{m} | \psi_{n\vec{k}} \rangle = \frac{1}{\hbar} \vec{\nabla}_{\vec{k}} \varepsilon_n(\vec{k}) . \quad (9.79)$$

On remarque au passage l'analogie avec une équation de Hamilton :

$$\langle \vec{v} \rangle = \frac{\partial \varepsilon_n}{\partial(\hbar\vec{k})} \quad \longleftrightarrow \quad \dot{q} = \frac{\partial H}{\partial p} , \quad (9.80)$$

la pseudo-impulsion $\hbar\vec{k}$ jouant le rôle d'un moment conjugué et $\varepsilon_n(\vec{k})$ celui d'un Hamiltonien. Comme dans tout état propre – qui produit un état stationnaire par simple adjonction du facteur de phase $e^{-i\hbar^{-1}\varepsilon_n(\vec{k})t}$ –, la vitesse d'un électron de Bloch est constante dans le temps. L'hypothèse d'un réseau d'ions fixes est évidemment essentielle pour ce point : on devine que si l'on introduit les vibrations du réseau, les états de Bloch ne sont plus stationnaires et, alors, la vitesse moyenne d'un électron peut de fait varier – et d'ailleurs se dégrader – par collisions (création - annihilation de phonons).

9.3.5 Symétrie de la fonction $\varepsilon(\vec{k})$

Pour simplifier, on omet ci-dessous l'indice n et on considère exclusivement une bande d'énergie $\varepsilon(\vec{k})$. Il s'agit de montrer que l'on a :

$$\varepsilon(-\vec{k}) = \varepsilon(\vec{k}) , \quad \vec{\nabla}\varepsilon\left(\frac{\vec{K}}{2}\right) = 0 \quad \forall \vec{K} \in \vec{B} . \quad (9.81)$$

dans l'hypothèse où la fonction $\varepsilon(\vec{k})$ est partout deux fois dérivable.

La fonction de Bloch satisfait :

$$H \psi_{\vec{k}} = \varepsilon(\vec{k}) \psi_{\vec{k}} . \quad (9.82)$$

On sait que, lorsque l'on ne prend pas en compte les degrés de spin, l'opération de renversement du temps se traduit par la conjugaison complexe de la fonction d'onde. Par ailleurs, lorsque le Hamiltonien est invariant par renversement du temps, une fonction propre ψ et sa transformée ψ^* correspondent à la même valeur propre E . Dans le contexte actuel, ces résultats généraux s'expriment comme suit.

H étant supposé invariant par renversement du temps²⁶, une fonction de Bloch $\psi_{\vec{k}}(\vec{r})$ et sa complexe conjuguée $\psi_{\vec{k}}^*(\vec{r})$ ont la même énergie $\varepsilon(\vec{k})$. Par ailleurs, on a :

$$\psi_{\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} u_{\vec{k}}(\vec{r}) \quad \psi_{\vec{k}}^*(\vec{r}) = e^{-i\vec{k}\cdot\vec{r}} u_{\vec{k}}^*(\vec{r}) . \quad (9.83)$$

$\psi_{\vec{k}}^*(\vec{r})$ est donc une fonction de Bloch caractérisée par le préfacteur $e^{-i\vec{k}\cdot\vec{r}} \equiv e^{i(-\vec{k})\cdot\vec{r}}$, elle est donc du genre $u_{-\vec{k}}(\vec{r})e^{i(-\vec{k})\cdot\vec{r}}$. La comparaison avec (9.83) permet l'identification :

$$u_{-\vec{k}}(\vec{r}) = u_{\vec{k}}^*(\vec{r}) . \quad (9.84)$$

Il en résulte que :

$$\psi_{\vec{k}}^*(\vec{r}) = \psi_{-\vec{k}}(\vec{r}) . \quad (9.85)$$

Ces deux fonctions, transformées l'une de l'autre par renversement du temps, ont la même énergie ; mais $\psi_{-\vec{k}}(\vec{r})$ satisfait :

$$H \psi_{-\vec{k}} = \varepsilon(-\vec{k}) \psi_{-\vec{k}} . \quad (9.86)$$

d'où il résulte :

$$\varepsilon(-\vec{k}) = \varepsilon(\vec{k}) . \quad (9.87)$$

$\varepsilon_{\vec{k}}$ étant une fonction paire, sa dérivée est une fonction impaire :

$$\vec{\nabla}\varepsilon(-\vec{k}) = -\vec{\nabla}\varepsilon(\vec{k}) . \quad (9.88)$$

En outre, $\varepsilon(\vec{k})$ est une fonction périodique en \vec{K} , tout comme son gradient ; en particulier :

$$\varepsilon(\vec{k} - \vec{K}) = \varepsilon(\vec{k}) \quad \text{et} \quad \vec{\nabla}\varepsilon(\vec{k} - \vec{K}) = \vec{\nabla}\varepsilon(\vec{k}) . \quad (9.89)$$

Pour les vecteurs d'onde tels que $\vec{k} - \vec{K} = -\vec{k}$, soit $\vec{k} = \frac{\vec{K}}{2}$ (ce sont certains points remarquables de la frontière de la maille de Wigner - Seitz), les équations (9.88) et (9.89) donnent simultanément²⁷ :

$$\vec{\nabla}\varepsilon\left(-\frac{\vec{K}}{2}\right) = -\vec{\nabla}\varepsilon\left(\frac{\vec{K}}{2}\right) \quad \text{et} \quad \vec{\nabla}\varepsilon\left(-\frac{\vec{K}}{2}\right) = \vec{\nabla}\varepsilon\left(\frac{\vec{K}}{2}\right) \quad \iff \quad \vec{\nabla}\varepsilon\left(\frac{\vec{K}}{2}\right) = 0 . \quad (9.90)$$

$\varepsilon(\vec{k})$ est *extremum* en ces points. De plus, $\varepsilon(\vec{k})$ est évidemment stationnaire en $\vec{k} = 0$, par parité.

²⁶C'est toujours le cas quand $H = \frac{\vec{p}^2}{2m} + V(\vec{r})$, une expression qui exclut la présence d'un champ magnétique. L'invariance de H par renversement du temps vient du fait que H est *quadratique* par rapport à la vitesse. Au contraire, un champ magnétique implique forcément une boucle de courant (en termes plus savants : un moment cinétique, de spin ou orbital), qui s'inverse dans le changement $t \rightarrow -t$.

²⁷La discussion en cours exclut les cas où $\varepsilon(\vec{k})$ présente des points anguleux.

9.3.6 Surface de Fermi

La surface de Fermi est définie et existe dans tout modèle à électrons indépendants – son observation expérimentale dans un métal est d'ailleurs usuellement considérée comme un test certifiant la pertinence d'une telle approximation²⁸ –, quel que soit le potentiel à un corps ressenti par un électron dans le modèle effectif ou renormalisé considéré. Dans ces conditions, la construction de l'état fondamental d'un ensemble de N électrons s'effectue facilement, en conformité avec principe de Pauli : on prend les états électroniques un par un, par ordre d'énergie croissante, et on y place deux électrons de spins opposés, jusqu'à épuisement du stock d'électrons. L'énergie maximum ainsi atteinte s'appelle énergie de Fermi, ε_F . Si les niveaux sont repérés par un vecteur d'onde \vec{k} , les niveaux occupés de plus haute énergie sont tels que :

$$\varepsilon(\vec{k}) = \varepsilon_F . \quad (9.91)$$

Cette relation définit une surface dans l'espace \vec{k} , appelée surface de Fermi.

La situation la plus simple est celle d'électrons libres. Alors, la loi de dispersion est :

$$\varepsilon^{(0)}(\vec{k}) = \frac{\hbar^2 \vec{k}^2}{2m} \quad (9.92)$$

de sorte que la surface de Fermi a pour équation :

$$\frac{\hbar^2 \vec{k}^2}{2m} = \varepsilon_F . \quad (9.93)$$

C'est donc une sphère, de rayon noté k_F . La valeur de k_F s'obtient comme suit. On admet des conditions cycliques, associant un petit cube de côté $2\pi/L$ à tout vecteur \vec{k} . En tenant compte du spin, et s'agissant de caser N électrons, on peut écrire :

$$2 \frac{(4\pi/3) k_F^3}{(2\pi/L)^3} = N \iff k_F = \left(3\pi^2 \frac{N}{L^3} \right)^{1/3} \equiv (3\pi^2 \rho)^{1/3} . \quad (9.94)$$

Pour un électron soumis à un potentiel périodique, la surface de Fermi est définie par considération des plus hautes bandes occupées. Dans le cas le plus simple, il n'y a qu'une seule plus haute bande occupée²⁹, alors la surface de Fermi est encore définie par l'équation (9.91), mais $\varepsilon(\vec{k})$ est maintenant la loi de dispersion relative à cette bande. Toutefois, comme il ne faut compter qu'une fois chaque état et que deux états ne différant que par un vecteur \vec{K} de $\tilde{\mathcal{B}}$ sont identiques, \vec{k} doit être restreint à une maille primitive de $\tilde{\mathcal{B}}$, la première zone de Brillouin par exemple. Il est donc utile de préciser :

$$\varepsilon(\vec{k}) = \varepsilon_F \quad \forall \vec{k} \in \text{BZ1} . \quad (9.95)$$

La valeur de ε_F est déterminée en intégrant la densité d'états $\mathcal{D}(E)$ (définie plus loin) jusqu'à ε_F et en écrivant que cette intégrale est égale au nombre d'électrons à caser. Précisément, pour un cristal à d dimensions, on écrit :

$$\int_0^{\varepsilon_F} dE 2 \left(\frac{L}{2\pi} \right)^d \int_{\mathbb{R}^d} d^d k \delta(\varepsilon(\vec{k}) - E) = N_e , \quad (9.96)$$

où N_e est le nombre d'électrons à répartir dans la bande ayant la dispersion $\varepsilon(\vec{k})$; de façon équivalente :

$$\int_0^{\varepsilon_F} dE \int_{\mathbb{R}^d} d^d k \delta(\varepsilon(\vec{k}) - E) = 2^{d-1} \pi^d \rho_e , \quad (9.97)$$

où ρ_e est la densité électronique.

Notons enfin que, compte tenu de la périodicité de $\varepsilon(\vec{k})$ dans $\tilde{\mathcal{B}}$, la surface de Fermi complète s'obtient par les translations de $\tilde{\mathcal{B}}$ de l'objet défini par (9.95) ; cette construction est utile pour l'analyse des problèmes de transport et permet, notamment, de distinguer le caractère ouvert ou fermé des orbites (trajectoires) semi-classiques.

²⁸ A tort ou à raison, l'image de particules indépendantes fait penser à un fluide ; un ensemble de tels fermions est appelé liquide de Fermi. On dit souvent que l'observation d'une surface de Fermi est la signature d'un liquide de Fermi.

²⁹ Quand il y a recouvrement en énergie entre les bandes, il peut y avoir plusieurs bandes incomplètement remplies chacune jusqu'à la même énergie ε_F (penser aux vases communicants).

9.4 Applications choisies du théorème de Bloch

Le théorème de Bloch énonce un résultat exact sur la nature des modes propres d'une particule dans un potentiel périodique, et montre notamment que les valeurs propres $\varepsilon_n(\vec{k})$ sont des fonctions (périodiques dans \vec{B}) du vecteur \vec{k} : ce sont les lois de dispersion de la particule dans le cristal, donnant la variation de l'énergie en fonction de l'impulsion, plus précisément de ce qui tient lieu d'impulsion – voir les mises en garde terminologiques du chapitre précédent. Il s'agit maintenant de présenter deux applications importantes de ce théorème, qui définissent deux cadres d'approximation adaptés à deux situations physiques extrêmes, identifiées par les caractères de ces lois de dispersion :

1. l'approximation des électrons presque libres, correspondant au cas où, dans le solide, les électrons sont seulement faiblement perturbés par le potentiel de réseau par comparaison à leur état libre. Il en résulte que la loi de dispersion ressemble à la parabole $\varepsilon^{(0)}(\vec{k}) = \hbar^2 \frac{\vec{k}^2}{2m}$; elle est donc fortement dispersée et on verra qu'elle est constituée d'arcs séparés par des petits *gaps*. Ceci correspond à des bandes larges et à une forte délocalisation spatiale des électrons (onde plane faiblement modulée). Cette situation prévaut dans les métaux et les semi-conducteurs
2. à l'opposé, l'approximation des liaisons fortes décrit le cas où chaque électron ressent, à l'ordre zéro, le potentiel localisé d'un atome (ou ion) du réseau. À cet ordre, le spectre d'énergie est celui d'un atome : des niveaux discrets, bien séparés. En fait, chaque électron peut se déplacer d'un atome à l'autre, ce qui donne une petite dispersion aux niveaux issus des niveaux atomiques. Il en résulte que les bandes d'énergie sont étroites et séparées par des grands *gaps*. Cette situation – lorsque la plus haute bande est pleine – est typique des isolants.

9.4.1 L'approximation des électrons presque libres

On adopte donc ici l'hypothèse suivant laquelle le potentiel de réseau $V(\vec{r})$ constitue une petite perturbation sur les états propres libres :

$$\psi_{\vec{k}}^{(0)}(\vec{r}) = C e^{i\vec{k}\cdot\vec{r}} \quad \varepsilon^{(0)}(\vec{k}) = \frac{\hbar^2 \vec{k}^2}{2m}, \quad (9.98)$$

ce qui autorise d'emblée à envisager un traitement de type perturbatif, avec les précautions et les hypothèses usuelles.

Ces états propres à l'ordre zéro sont qualitativement différents des vrais états propres de l'électron dans le cristal, puisque ceux-ci doivent avoir la symétrie exigée par le théorème de Bloch et ses conséquences en raison de la symétrie de translation discrète – ce n'est pas le cas des états libres (9.98), qui sont caractéristiques d'une symétrie de translation continue, celle d'une particule libre (une onde plane telle que (9.98) est formellement une fonction de Bloch – $u_{\vec{k}}$ étant une constante (ce qui est bien une fonction périodique. . .) –, mais \vec{k} et $\vec{k} + \vec{K}$ ne représentent pas le même état physique, au contraire des “vrais” états de Bloch) ; d'ailleurs, dans (9.98), $\hbar\vec{k}$ est visiblement la *vraie* impulsion de l'électron.

Pour tout dire, on rencontre ici un bon exemple du fait que la symétrie fonctionne par tout ou rien et que des considérations *quantitatives* – du genre de celles que l'on énonce pour mettre en évidence un petit paramètre par exemple – appartiennent à un autre plan. Que le potentiel de réseau soit faible ou intense ne change rien au fait que la symétrie de translation devient discrète même si $V(\vec{r})$ est infiniment petit³⁰. Il s'agit bien d'une *brisure* de symétrie, puisque la symétrie est devenue discrète sous l'effet du réseau de pas fini. Le groupe des translations discrètes est visiblement un sous-groupe des translations continues ; on retrouve ce dernier en faisant tendre vers zéro le pas du réseau.

³⁰On a toujours intérêt à faire jouer la symétrie le plus tôt possible. Notamment, l'introduction explicite de vecteurs ayant la symétrie de la perturbation produit des matrices diagonales par blocs. Il arrive même, quand la brisure de symétrie est assez forte, qu'à l'intérieur d'un sous-espace dégénéré, la projection de l'opérateur de perturbation dans ce sous-espace est représentée par une matrice complètement diagonale ; la simple lecture de ces éléments diagonaux donne alors immédiatement, sans aucun calcul supplémentaire, les corrections d'énergie au 1^{er} ordre.

Comme on l'a vu précédemment, l'existence du potentiel de réseau entraîne que les fonctions propres et les valeurs propres sont périodiques dans $\tilde{\mathcal{B}}$:

$$\psi_{n\vec{k}-\vec{K}}(\vec{r}) = \psi_{n\vec{k}}(\vec{r}) , \quad \varepsilon_n(\vec{k}-\vec{K}) = \varepsilon_n(\vec{k}) \quad \forall \vec{K} \in \tilde{\mathcal{B}} . \quad (9.99)$$

Ceci signifie qu'il suffit de les connaître à l'intérieur d'une maille primitive de $\tilde{\mathcal{B}}$ et, dans la suite, on pourra toujours, si besoin est, choisir le vecteur \vec{k} dans la première zone de Brillouin³¹. Compte tenu de l'hypothèse sur la petitesse de V , on va utiliser une méthode de perturbation. Il convient donc de préciser complètement le problème, tel qu'il se pose à l'ordre zéro, sans hésiter d'ailleurs à le reformuler dans la perspective du but recherché. Par ailleurs, on sait que la technique de perturbation, avant d'être mise en œuvre, exige l'analyse préalable de la dégénérescence du problème non-perturbé.

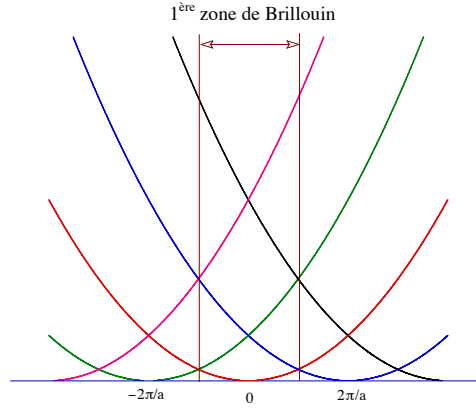


Figure 9.1: Schéma en zones réduites pour un réseau \mathcal{B} unidimensionnel de pas a . Les différentes paraboles représentent les $\varepsilon_{\vec{K}}^{(0)}(\vec{k})$ (voir (9.101)). La première zone de Brillouin s'étend entre $\pm \frac{\pi}{a}$.

Les solutions exactes sont périodiques dans $\tilde{\mathcal{B}}$, ainsi que l'exprime (9.99). D'un autre côté, comme V reste en toute hypothèse petit, on s'attend nécessairement à ce que l'on ait toujours :

$$\varepsilon(\vec{k}) \approx \varepsilon^{(0)}(\vec{k}) . \quad (9.100)$$

La bonne façon de formuler le problème à l'ordre zéro est donc de prolonger par périodisation, conformément à (9.99), tous les arcs de la parabole libre, faisant dès lors de la loi de dispersion libre une fonction elle aussi infiniment multivaluée. Ceci se fait en translatant tous les arcs de la parabole libre qui sont en-dehors de cette zone, et on obtient alors le schéma dit en zones réduites (voir figure 9.1). Dans cette représentation, $\forall \vec{k} \in \text{BZ1}$, l'énergie d'un électron en fonction de \vec{k} est infiniment multivaluée³². Avec la même notation que précédemment, $\varepsilon^{(0)}(\vec{k})$, le prolongement périodique revient donc à définir une infinité de lois paraboliques $\varepsilon_{\vec{K}}^{(0)}$ telles que :

$$\varepsilon_{\vec{K}}^{(0)}(\vec{k}) = \varepsilon^{(0)}(\vec{k}-\vec{K}) \equiv \frac{\hbar^2}{2m}(\vec{k}-\vec{K})^2 \quad \forall \vec{K} \in \tilde{\mathcal{B}} . \quad (9.101)$$

Il y aura dégénérescence stricte à chaque fois que ces paraboles se croisent, c'est-à-dire lorsque :

$$\varepsilon_{\vec{K}_1}^{(0)}(\vec{k}) = \varepsilon_{\vec{K}_2}^{(0)}(\vec{k}) \iff \varepsilon^{(0)}(\vec{k}-\vec{K}_1) = \varepsilon^{(0)}(\vec{k}) , \quad (9.102)$$

où $\vec{K} = \vec{K}_2 - \vec{K}_1$ est encore un vecteur de $\tilde{\mathcal{B}}$. Comme $\varepsilon^{(0)}(\vec{k})$ ne dépend que du module des vecteurs d'onde, cette égalité (9.102) est satisfaite à chaque fois que³³ :

$$\|\vec{k}\| = \|\vec{k}-\vec{K}\| = \|\vec{K}-\vec{k}\| . \quad (9.103)$$

³¹Dès qu'il existe un potentiel périodique, \vec{k} n'est défini physiquement que *modulo* $\vec{K} \in \tilde{\mathcal{B}}$, tout comme un angle est défini à 2π près ; on peut dire que l'équivalent de BZ1 est alors le segment $[-\pi, +\pi]$ pour une fonction 2π -périodique.

³²On peut se demander ce que devient la périodisation de $\varepsilon_{\vec{K}}^{(0)}$ quand on en revient au cas strictement libre. De fait, la limite de l'électron libre se retrouve comme suit : il suffit d'imaginer que l'on fait tendre vers zéro la maille primitive du réseau, ce qui à la limite produit un potentiel $V(\vec{r})$ prenant la même valeur *partout*. Dans cette limite, la longueur de tout vecteur $\vec{K} \in \tilde{\mathcal{B}}$ tend vers l'infini et on récupère *in fine* l'unique parabole qui est la limite de l'arc de plus basse énergie dans le schéma en zones réduites.

³³En l'absence de champ magnétique, $\varepsilon^{(0)}$ est aussi une fonction paire de \vec{k} ; toutefois, il n'y a pas lieu de considérer la dégénérescence $\vec{k} \rightarrow -\vec{k}$ correspondante : \vec{k} est un bon nombre quantique et les éléments de matrice $\langle \psi_{\vec{k}} | H | \psi_{-\vec{k}} \rangle$ sont nuls.

On peut dire les choses autrement. Pour un électron libre, la condition de dégénérescence est :

$$\varepsilon^{(0)}(\vec{k}) = \varepsilon^{(0)}(\vec{k}') . \quad (9.104)$$

Cette égalité est vraie si \vec{k} et \vec{k}' ont même module, en particulier si $\vec{k} = \vec{k}'$. Mais, dès qu'un potentiel périodique est présent, l'égalité des vecteurs d'onde – ou l'égalité des modules – doit être comprise *modulo* \vec{K} . Dans cette optique, (9.104) est équivalente à (9.103).

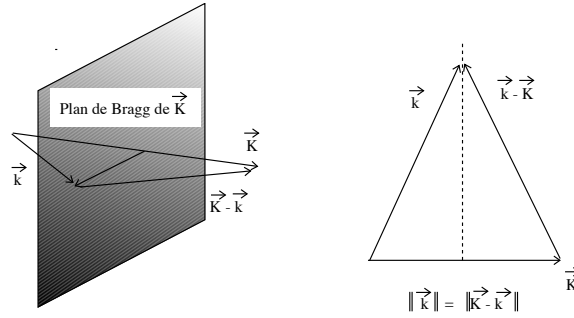


Figure 9.2: Schéma de définition du plan de Bragg d'un vecteur $\vec{K} \in \tilde{\mathcal{B}}$.

L'équation (9.103) signifie que, \vec{K} étant fixé, l'extrémité de \vec{k} est dans le plan médiateur de \vec{K} , autrement dit le plan de Bragg associé à \vec{K} . Ainsi, il apparaît une dégénérescence stricte apparaît lorsque \vec{k} appartient à un plan de Bragg³⁴ ; en pareil cas, on pourra toujours poser :

$$\vec{k} = \frac{\vec{K}}{2} + \vec{q} , \quad \vec{q} \perp \vec{K} . \quad (9.105)$$

Soit alors un vecteur primitif \vec{b}_0 définissant une certaine direction de $\tilde{\mathcal{B}}$, on peut distinguer deux cas :

- ou bien $\vec{K} = \text{entier pair} \times \vec{b}_0 \equiv 2n_0\vec{b}_0$. Dans ces conditions $\frac{\vec{K}}{2} = \vec{K}_0 \in \tilde{\mathcal{B}}$ et, après la bonne translation, on se retrouve à l'origine de $\tilde{\mathcal{B}}$. On parle alors de dégénérescence en centre de zone
- ou bien $\vec{K} = \text{entier impair} \times \vec{b}_0 \equiv (2n_0 + 1)\vec{b}_0$. Alors, $\frac{\vec{K}}{2} = \vec{K}_0 + \frac{1}{2}\vec{b}_0 \notin \tilde{\mathcal{B}}$ et la dégénérescence stricte se produit en un certain point de la frontière de la maille de Wigner - Seitz³⁵. On parle alors de dégénérescence en bord de zone.

Comme le paramètre \vec{k} varie continûment³⁶, il ne serait pas satisfaisant d'utiliser presque partout dans la zone la théorie de perturbation relative au cas non-dégénéré, sauf en certains points isolés où on appliquerait le formalisme du cas dégénéré ; une telle procédure conduirait d'ailleurs à des difficultés techniques puisque les dénominateurs d'énergie peuvent devenir graduellement de plus en plus petits (voir aussi (9.106)). Il faut donc construire un schéma de calcul plus subtil que celui énoncé par la théorie des perturbations "académique"³⁷.

On sait que le critère convenable pour qu'une théorie de perturbation fonctionne bien est de rassembler – en les traitant sur un même pied d'égalité – tous les états d'ordre zéro $\psi_n^{(0)}$ dont les énergies $E_n^{(0)}$ ont des différences au plus comparables aux éléments de matrice de la perturbation entre ces états. Ici, le potentiel de réseau (la perturbation) est supposé faible, c'est-à-dire que ses éléments de matrice sont réputés petits par

³⁴sans préjuger d'autres coïncidences par ailleurs, résultant d'une haute symétrie ponctuelle.

³⁵définie comme le polyèdre dont les faces sont les plans de Bragg de tous les vecteurs réciproques joignant l'origine de $\tilde{\mathcal{B}}$ aux nœuds premiers voisins et communément appelée 1^{ère} zone de Brillouin. L'intérieur de la 1^{ère} zone de Brillouin est l'ensemble des points de $\tilde{\mathcal{B}}$ qui sont plus près du nœud origine $\vec{K} = 0$ que de tous les autres nœuds (c'est la maille de Wigner - Seitz du réseau réciproque).

³⁶À une petite quantification près imposée par les conditions cycliques.

³⁷On rencontre une situation analogue pour un atome soumis à un champ magnétique \mathcal{B} ; alors les valeurs propres dépendent continûment de \mathcal{B} et le traitement de perturbation ordinaire doit être redéfini en présence d'un croisement de niveaux : on ne peut pas s'en tenir à la stricte recette énoncée par la théorie de perturbations dans sa version élémentaire.

rapport à une énergie typique ; compte tenu des paramètres physiques disponibles, cette énergie ne peut être que³⁸ $\frac{\hbar^2}{ma^2}$. Au total, le bon critère s'énonce par une condition du type :

$$|\varepsilon_{\vec{K}}^{(0)}(\vec{k}) - \varepsilon_{\vec{K}'}^{(0)}(\vec{k})| \lesssim |V| \ll \frac{\hbar^2}{ma^2} , \quad (9.106)$$

et le traitement de perturbation doit considérer *simultanément* tous les états d'ordre zéro satisfaisant (9.106).

V étant suffisamment faible, les arguments précédents permettent de délimiter deux situations, \vec{k} étant dans BZ1 :

1. \vec{k} est tel que toutes les branches $\varepsilon_{\vec{K}}^{(0)}(\vec{k})$ sont bien isolées les unes des autres. Alors, une théorie de perturbation non-dégénérée est acceptable
2. \vec{k} est tel que plusieurs branches $\varepsilon_{\vec{K}_i}^{(0)}(\vec{k})$ se trouvent au voisinage les unes des autres, l'étendue du voisinage se mesurant par l'ordre de grandeur du potentiel périodique conformément à (9.106). Notamment, ceci se produit, comme on vient de le voir, lorsque \vec{k} est près du centre ou au voisinage du plan de Bragg de $\vec{K} = \vec{K}_i - \vec{K}_j$ (par exemple sur la frontière de BZ1, quand $\vec{K}_1 = 0$, et \vec{K}_2 joint l'origine à un nœud premier voisin dans \vec{B} , par exemple). S'il y a g telles branches, le moins que l'on puisse faire est alors de considérer *simultanément* les g fonctions $\psi_{\vec{k}-\vec{K}_i}^{(0)}$, ou plus précisément le sous-espace $g \times g$ qu'elles engendrent. S'y prendre ainsi c'est finalement étendre la recette de perturbation dégénérée dans le *voisinage* de la dégénérescence $\|\vec{k} - \vec{K}_i\| = \|\vec{k} - \vec{K}_j\|$ et non plus strictement au point de dégénérescence.

Dans le cas le plus simple, on a $g = 2$; c'est la situation lorsque \vec{k} est non seulement dans le plan de $\vec{K} = \vec{K}_i - \vec{K}_j$ mais encore pas trop éloigné de la direction du vecteur \vec{K} lui-même. Si l'écart grandit, \vec{k} peut se rapprocher de la droite d'intersection de deux plans de Bragg et alors visiblement, $g > 2$.

Examinons maintenant comment les deux cas se présentent naturellement, en revenant au problème exact et en le représentant sur une base d'ondes planes (les fonctions à l'ordre zéro) – ceci revient à se placer dans l'espace de Fourier. Une fonction d'onde de Bloch s'écrit :

$$\psi_{\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} u_{\vec{k}}(\vec{r}) , \quad (9.107)$$

et comme la fonction $u_{\vec{k}}(\vec{r})$ a la périodicité du réseau, on peut la décomposer en série de Fourier sur le réseau réciproque :

$$u_{\vec{k}}(\vec{r}) = \sum_{\vec{K} \in \vec{B}} U_{\vec{k}}(\vec{K}) e^{-i\vec{K}\cdot\vec{r}} . \quad (9.108)$$

Il en résulte que l'on peut écrire :

$$\psi_{\vec{k}}(\vec{r}) = \sum_{\vec{K} \in \vec{B}} U_{\vec{k}}(\vec{K}) e^{i(\vec{k}-\vec{K})\cdot\vec{r}} . \quad (9.109)$$

On sait que deux états de Bloch \vec{k} et $\vec{k} + \vec{K}_0$, $\vec{K}_0 \in \vec{B}$, sont en fait un seul et même état, d'où il résulte que la fonction $U_{\vec{k}}(\vec{K})$ ne dépend que de la différence $\vec{k} - \vec{K}$. En effet, (9.109) où on remplace \vec{k} par $\vec{k} + \vec{K}_0$ donne :

$$\psi_{\vec{k}+\vec{K}_0}(\vec{r}) = \sum_{\vec{K} \in \vec{B}} U_{\vec{k}+\vec{K}_0}(\vec{K}) e^{i(\vec{k}+\vec{K}_0-\vec{K})\cdot\vec{r}} ; \quad (9.110)$$

on fait le changement de variable muette $\vec{K} - \vec{K}_0 = \vec{K}'$ dans la sommation, d'où :

$$\psi_{\vec{k}+\vec{K}_0}(\vec{r}) = \sum_{\vec{K}' \in \vec{B}} U_{\vec{k}+\vec{K}_0}(\vec{K}' + \vec{K}_0) e^{i(\vec{k}-\vec{K}')\cdot\vec{r}} \equiv \sum_{\vec{K} \in \vec{B}} U_{\vec{k}+\vec{K}_0}(\vec{K} + \vec{K}_0) e^{i(\vec{k}-\vec{K})\cdot\vec{r}} ; \quad (9.111)$$

comme $\psi_{\vec{k}} = \psi_{\vec{k}+\vec{K}_0}$, l'identification terme à terme de (9.109) et (9.111) donne :

$$U_{\vec{k}+\vec{K}_0}(\vec{K} + \vec{K}_0) = U_{\vec{k}}(\vec{K}) . \quad (9.112)$$

³⁸ a est l'ordre de grandeur de la dimension linéaire d'une maille primitive.

Ceci signifie que la fonction $U_{\vec{k}}(\vec{K})$ est invariante dans toute translation de $\vec{\mathcal{B}}$ et ne dépend donc que de $\vec{k} - \vec{K}$. En posant désormais $U_{\vec{k}}(\vec{K}) = c_{\vec{k}-\vec{K}}$, (9.109) prend la forme :

$$\psi_{\vec{k}}(\vec{r}) = \sum_{\vec{K} \in \vec{\mathcal{B}}} c_{\vec{k}-\vec{K}} e^{i(\vec{k}-\vec{K}) \cdot \vec{r}} . \quad (9.113)$$

$\psi_{\vec{k}}(\vec{r})$ est ainsi un développement en ondes planes, chacune de ces dernières satisfaisant le théorème de Bloch – une fonction constante dans l'espace est évidemment périodique en \mathcal{B} – et ses conséquences (9.99). En termes de *kets*, on peut écrire :

$$|\psi_{\vec{k}}\rangle = \sum_{\vec{K} \in \vec{\mathcal{B}}} c_{\vec{k}-\vec{K}} |f_{\vec{k}-\vec{K}}\rangle , \quad (9.114)$$

où :

$$\langle \vec{r} | f_{\vec{k}-\vec{K}} \rangle = e^{i(\vec{k}-\vec{K}) \cdot \vec{r}} \equiv \psi_{\vec{k}-\vec{K}}^{(0)}(\vec{r}) \quad (9.115)$$

sont les ondes planes non-perturbées, solutions du problème à l'ordre zéro. Les $c_{\vec{k}-\vec{K}}$ sont donc les coefficients cherchés, que l'on imagine – dans un traitement perturbatif – être des séries entières en V .

Par ailleurs, la fonction $\psi_{\vec{k}}$ satisfait :

$$-\frac{\hbar^2}{2m} \Delta \psi_{\vec{k}}(\vec{r}) + V(\vec{r}) \psi_{\vec{k}}(\vec{r}) = \varepsilon(\vec{k}) \psi_{\vec{k}}(\vec{r}) . \quad (9.116)$$

D'autre part, $\mathcal{V}(\vec{K})$ désignant toujours les composantes de Fourier³⁹ du potentiel périodique $V(\vec{r})$:

$$V(\vec{r}) = \sum_{\vec{K} \in \vec{\mathcal{B}}} \mathcal{V}(\vec{K}) e^{-i\vec{K} \cdot \vec{r}} \iff \mathcal{V}(\vec{K}) = \frac{1}{v} \int_{\text{maille élémentaire de } \mathcal{B}} d^3r V(\vec{r}) e^{i\vec{K} \cdot \vec{r}} . \quad (9.117)$$

En reportant les développements (9.113) et (9.117) dans (9.116), on trouve :

$$\begin{aligned} - \sum_{\vec{K} \in \vec{\mathcal{B}}} \frac{\hbar^2}{2m} [i(\vec{k} - \vec{K})]^2 c_{\vec{k}-\vec{K}} e^{i(\vec{k}-\vec{K}) \cdot \vec{r}} &+ \sum_{\vec{K}' \in \vec{\mathcal{B}}, \vec{K}'' \in \vec{\mathcal{B}}} \mathcal{V}(\vec{K}') c_{\vec{k}-\vec{K}''} e^{-i\vec{K}' \cdot \vec{r}} e^{i(\vec{k}-\vec{K}'') \cdot \vec{r}} \\ &= \sum_{\vec{K} \in \vec{\mathcal{B}}} \varepsilon(\vec{k}) c_{\vec{k}-\vec{K}} e^{i(\vec{k}-\vec{K}) \cdot \vec{r}} , \end{aligned} \quad (9.118)$$

soit, en posant $\vec{K}' + \vec{K}'' = \vec{K}$ dans la somme double :

$$\sum_{\vec{K} \in \vec{\mathcal{B}}} [\varepsilon^{(0)}(\vec{k} - \vec{K}) - \varepsilon(\vec{k})] c_{\vec{k}-\vec{K}} e^{i(\vec{k}-\vec{K}) \cdot \vec{r}} + \sum_{\vec{K}, \vec{K}' \in \vec{\mathcal{B}}^2} \mathcal{V}(\vec{K}') c_{\vec{k}-\vec{K}+\vec{K}'} e^{i(\vec{k}-\vec{K}) \cdot \vec{r}} = 0 . \quad (9.119)$$

Par identification des termes en $e^{-i\vec{K} \cdot \vec{r}}$ et après simplification par $e^{i\vec{k} \cdot \vec{r}}$, on en déduit le *système* d'équations pour les coefficients de Fourier de $\psi_{\vec{k}}$:

$$[\varepsilon^{(0)}(\vec{k} - \vec{K}) - \varepsilon(\vec{k})] c_{\vec{k}-\vec{K}} + \sum_{\vec{K}' \in \vec{\mathcal{B}}} \mathcal{V}(\vec{K}') c_{\vec{k}-\vec{K}+\vec{K}'} = 0 , \quad \forall \vec{K} \in \vec{\mathcal{B}} , \quad (9.120)$$

ou, de façon équivalente :

$$[\varepsilon^{(0)}(\vec{k} - \vec{K}) - \varepsilon(\vec{k})] c_{\vec{k}-\vec{K}} + \sum_{\vec{K}'' \in \vec{\mathcal{B}}} \mathcal{V}(\vec{K} - \vec{K}'') c_{\vec{k}-\vec{K}''} = 0 , \quad \forall \vec{K} \in \vec{\mathcal{B}} . \quad (9.121)$$

En choisissant convenablement le zéro d'énergie potentielle – à la valeur moyenne de $V(\vec{r})$ sur une maille primitive –, la composante uniforme $\mathcal{V}(\vec{K}) = 0$ est nulle ; dans ces conditions, la somme dans (9.121) court explicitement sur les valeurs de $\vec{K}'' \neq \vec{K}$. Par ailleurs, ce choix entraîne qu'il n'y a pas de correction d'énergie

³⁹ $\mathcal{V}(\vec{K})$ est homogène à une énergie.

au premier ordre en l'absence de dégénérescence. En effet, en pareil cas, la correction au premier ordre est simplement la valeur moyenne de la perturbation prise sur un état d'ordre zéro, soit ici :

$$\int_{\mathbb{R}^3} d^3r e^{-i(\vec{k}-\vec{K})\cdot\vec{r}} V(\vec{r}) e^{i(\vec{k}-\vec{K})\cdot\vec{r}} = \int_{\mathbb{R}^3} d^3r V(\vec{r}) = 0 . \quad (9.122)$$

Jusqu'à présent, aucune approximation n'a été faite : le système (9.121) est exact, il est tout juste une reformulation de l'équation aux valeurs propres (9.116) dans l'espace de Fourier. Inexploitable en tant que tel dans le cas général, on en effectue maintenant l'analyse perturbative.

Le système (9.121) s'écrit aussi :

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K})] c_{\vec{k}-\vec{K}} = \sum_{\vec{K}' \in \tilde{\mathcal{B}}, \vec{K}' \neq \vec{K}} \mathcal{V}(\vec{K} - \vec{K}') c_{\vec{k}-\vec{K}'} , \quad \forall \vec{K} \in \tilde{\mathcal{B}} . \quad (9.123)$$

Si $V \equiv 0$, il se réduit à :

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K})] c_{\vec{k}-\vec{K}} = 0 , \quad \forall \vec{K} \in \tilde{\mathcal{B}} . \quad (9.124)$$

L'équation (9.124) dit que $c_{\vec{k}-\vec{K}}$ est forcément nul tant que $\varepsilon(\vec{k}) \neq \varepsilon^{(0)}(\vec{k} - \vec{K})$ et inversement : c'est seulement lorsque $\varepsilon(\vec{k})$ est égal à $\varepsilon^{(0)}(\vec{k} - \vec{K})$, que $c_{\vec{k}-\vec{K}}$ n'est pas nul et est, à ce stade, indéterminé ; c'est une constante non nulle faute de quoi la solution serait identiquement nulle, condition rencontrée à propos de toute équation aux fonctions propres. En tant que constante à la limite $V \rightarrow 0$, c'est donc un terme d'ordre $\mathcal{O}(V^0)$.

Dans le cas 1., par définition, toutes les énergies $\varepsilon^{(0)}(\vec{k} - \vec{K})$ sont, pour le \vec{k} choisi, bien isolées les unes des autres. Considérons-en une, celle ayant $\vec{K} = \vec{K}_1$ et récrivons le système (9.123) en exhibant à part la ligne $\vec{K} = \vec{K}_1$:

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K}_1)] c_{\vec{k}-\vec{K}_1} = \sum_{\vec{K}' \in \tilde{\mathcal{B}}, \vec{K}' \neq \vec{K}_1} \mathcal{V}(\vec{K}_1 - \vec{K}') c_{\vec{k}-\vec{K}'} , \quad (9.125)$$

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K})] c_{\vec{k}-\vec{K}} = \sum_{\vec{K}' \in \tilde{\mathcal{B}}, \vec{K}' \neq \vec{K}} \mathcal{V}(\vec{K} - \vec{K}') c_{\vec{k}-\vec{K}'} , \quad \forall \vec{K} \in \tilde{\mathcal{B}}, \vec{K} \neq \vec{K}_1 . \quad (9.126)$$

Quand $V \rightarrow 0$, un seul crochet $[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K})]$ figurant aux premiers membres du système (9.123) s'annule, celui de (9.125), et la valeur propre perturbée $\varepsilon(\vec{k})$ tend vers l'unique valeur isolée $\varepsilon^{(0)}(\vec{k} - \vec{K}_1)$ considérée. Bien sûr, tous les autres premiers membres de (9.126) s'annulent aussi, mais c'est alors parce que, le crochet n'étant pas nul à la limite, le coefficient $c_{\vec{k}-\vec{K}}$ lui-même tend vers zéro $\forall \vec{K} \neq \vec{K}_1$. Dans ces conditions, chaque $c_{\vec{k}-\vec{K}}, \vec{K} \neq \vec{K}_1$ est au moins d'ordre 1 en V ; il en résulte que la somme au second membre de (9.125) – qui ne contient que les $c_{\vec{k}-\vec{K}'}, \vec{K}' \neq \vec{K}_1$ – est au moins du second ordre⁴⁰. En définitive, la différence d'énergie $\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K}_1)$ est du second ordre⁴¹ et la correction résultant du petit potentiel périodique est très petite. La fonction d'onde diffère de la fonction d'onde d'ordre zéro par des termes du premier ordre : l'ordre zéro est la série (9.108) réduite alors à un seul terme $e^{i(\vec{k}-\vec{K})\cdot\vec{r}}$, c'est bien une onde plane. Au total, il n'y a pas de changement qualitatif notable.

Dans le cas 2., toujours avec \vec{k} fixé, il existe par hypothèse un certain nombre, g , d'énergies quasi dégénérées associées à des vecteurs \vec{K}_i distincts :

$$|\varepsilon^{(0)}(\vec{k} - \vec{K}_i) - \varepsilon^{(0)}(\vec{k} - \vec{K}_j)| \lesssim |V| \quad (i, j = 1, 2, \dots, g) . \quad (9.127)$$

Mettons à part toutes les lignes où ces \vec{K}_i sont au premier membre de (9.123) :

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K}_i)] c_{\vec{k}-\vec{K}_i} = \sum_{\vec{K}' \in \tilde{\mathcal{B}}, \vec{K}' \neq \vec{K}_i} \mathcal{V}(\vec{K}_i - \vec{K}') c_{\vec{k}-\vec{K}'} , \quad \vec{K}_i = \vec{K}_1, \vec{K}_2, \dots, \vec{K}_g , \quad (9.128)$$

⁴⁰ Comme toujours quand il s'agit de perturbation ordinaire (non-singulière), on fait l'hypothèse suivant laquelle tous les développements sont en puissances entières de la perturbation, ici V .

⁴¹ Bien sûr, l'absence de correction d'énergie au 1^{er} ordre est une conséquence du calage de $V(\vec{r})$ assurant que $\mathcal{V}(\vec{K}) = 0$.

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K})] c_{\vec{k}-\vec{K}} = \sum_{\vec{K}' \in \tilde{\mathcal{B}}, \vec{K}' \neq \vec{K}} \mathcal{V}(\vec{K} - \vec{K}') c_{\vec{k}-\vec{K}'}, \quad \forall \vec{K} \in \tilde{\mathcal{B}}, \quad \vec{K} \neq \vec{K}_i, \quad i = 1, 2, \dots, g. \quad (9.129)$$

Le même argument que celui développé pour le cas 1. montre alors que les g crochets $[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K}_i)]$ des premiers membres de (9.128) s'annulent à la limite $V = 0$, donc les g coefficients $c_{\vec{k}-\vec{K}_i}$ sont tous d'ordre $\mathcal{O}(V^0)$. Tous les autres sont au moins d'ordre $\mathcal{O}(V)$: tout comme dans le cas 1., ils apparaissent au premier membre de (9.129), multipliés par le crochet de la différence des énergies, qui ne s'annule pas, cependant que le second membre contient des termes d'ordre $\mathcal{O}(V)$ (les g termes du genre $\mathcal{V}c_{\vec{k}-\vec{K}_i}$ - et aussi des termes d'ordre supérieur) et donc s'annule à la limite.

Alors, en extrayant du système (9.123) les seules équations où ces \vec{K}_i figurent au premier membre (il y en a g), il vient :

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K}_i)] c_{\vec{k}-\vec{K}_i} = \sum_{\vec{K}' \in \tilde{\mathcal{B}}, \vec{K}' \neq \vec{K}_i} \mathcal{V}(\vec{K}_i - \vec{K}') c_{\vec{k}-\vec{K}'}, \quad \forall \vec{K} \in \{\vec{K}_i\}. \quad (9.130)$$

Au second membre de (9.130), figurent $g-1$ coefficients d'ordre zéro, $c_{\vec{k}-\vec{K}_j}$, $j \neq i$, tous les autres étant d'ordre $\mathcal{O}(V)$. Le système (9.130) tronqué aux g lignes pertinentes peut donc s'écrire :

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K}_i)] c_{\vec{k}-\vec{K}_i} = \sum_{\vec{K}_j \in \{\vec{K}_i\}, j \neq i} \mathcal{V}(\vec{K}_i - \vec{K}_j) c_{\vec{k}-\vec{K}_j} + \mathcal{O}(V^2) \quad (i = 1, 2, \dots, g), \quad (9.131)$$

où tous les coefficients $c_{\vec{k}-\vec{K}_j}$ explicitement écrits sont $\mathcal{O}(V^0)$. Ceci étant, (9.131) montre alors que la correction d'énergie est une somme finie de termes d'ordre $\mathcal{O}(V)$, elle est donc aussi d'ordre $\mathcal{O}(V)$. Si on décide de rester à l'ordre le plus bas non trivial, il suffit de considérer seulement la somme explicitée au second membre de (9.131), et on est réduit à devoir résoudre :

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K}_i)] c_{\vec{k}-\vec{K}_i} = \sum_{\vec{K}_j \in \{\vec{K}_i\}, j \neq i} \mathcal{V}(\vec{K}_i - \vec{K}_j) c_{\vec{k}-\vec{K}_j} \quad (i = 1, 2, \dots, g). \quad (9.132)$$

La somme au second membre contient au total $g-1$ termes ; l'omission des autres termes est le pendant, au voisinage d'une dégénérescence, de ce que prescrit la recette de perturbation dégénérée : elle précise de ne considérer que la *projection* de l'opérateur de perturbation à l'intérieur de chaque sous-espace dégénéré, en omettant tous les couplages non-diagonaux entre deux sous-espaces dégénérés distincts. En définitive, on obtient un système linéaire $g \times g$ pour les g coefficients $c_{\vec{k}-\vec{K}_i}$ relatifs aux g énergies quasi-dégénérées : c'est bien une version plus habile de la méthode de perturbations standard.

Afin de fixer les idées dans ce cas 2., supposons qu'il n'existe que deux valeurs de \vec{K} , \vec{K}_1 et \vec{K}_2 , telles que l'égalité $\varepsilon^{(0)}(\vec{k} - \vec{K}_1) = \varepsilon^{(0)}(\vec{k} - \vec{K}_2)$ peut être satisfaite ; comme tout est périodique dans le réseau réciproque, on peut toujours supposer que l'un des deux \vec{K}_i est nul, par exemple \vec{K}_1 , et poser alors $\vec{K}_2 = \vec{K}$ pour simplifier les écritures. Dès lors, on se trouve dans le cas où \vec{k} est au voisinage d'un seul plan de Bragg, précisément celui de \vec{K} . Le système simplifié (9.132), se réduit alors au couple d'équations :

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k})] c_{\vec{k}} - \mathcal{V}(-\vec{K}) c_{\vec{k}-\vec{K}} = 0, \quad (9.133)$$

$$-\mathcal{V}(\vec{K}) c_{\vec{k}} + [\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K})] c_{\vec{k}-\vec{K}} = 0. \quad (9.134)$$

Ce système linéaire homogène n'a de solution non triviale que si son déterminant est nul, ce qui fournit les deux valeurs propres approchées :

$$\varepsilon_{\pm}(\vec{k}) = \Sigma\varepsilon \pm \sqrt{[\Delta\varepsilon]^2 + |\mathcal{V}(\vec{K})|^2} \equiv \Sigma\varepsilon \pm h(\vec{k}), \quad (9.135)$$

où on a posé :

$$\Sigma\varepsilon = \frac{1}{2}[\varepsilon^{(0)}(\vec{k}) + \varepsilon^{(0)}(\vec{k} - \vec{K})] \quad \Delta\varepsilon = \frac{1}{2}[\varepsilon^{(0)}(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K})]. \quad (9.136)$$

Cette expression donne donc la levée de dégénérescence dans le voisinage d'un point de l'espace réciproque où $\varepsilon^{(0)}(\vec{k}) = \varepsilon^{(0)}(\vec{k} - \vec{K})$; cette égalité est certainement vérifiée si :

$$\|\vec{k}\| = \|\vec{k} - \vec{K}\| \quad (9.137)$$

exprimant que l'extrémité du vecteur \vec{k} se trouve dans le plan bissecteur de \vec{K} , qui est par définition le plan de Bragg de \vec{K} , noté $\mathcal{P}_{\frac{\vec{K}}{2}}$.

Il y a donc maintenant deux solutions distinctes $\varepsilon_{\pm}(\vec{k})$, traduisant la levée de dégénérescence, laquelle est stricte quand \vec{k} est dans ce plan ; dans cette configuration, $\Delta\varepsilon$ s'annule et (9.135) donne :

$$\varepsilon_{\pm}(\vec{k} \in \mathcal{P}_{\frac{\vec{K}}{2}}) = \varepsilon^{(0)}(\vec{k}) \pm |\mathcal{V}(\vec{K})| . \quad (9.138)$$

Ainsi, les deux énergies qui coïncidaient à l'ordre zéro quand l'extrémité de \vec{k} est dans un plan de Bragg sont maintenant distinctes : le potentiel périodique a ouvert un *gap* d'énergie, la valeur propre $\varepsilon(\vec{k})$ différant nettement de $\varepsilon^{(0)}(\vec{k})$. Par continuité, il en va de même tant que \vec{k} reste au voisinage d'un tel plan. Quand l'extrémité de \vec{k} s'éloigne de tout plan de Bragg, les énergies perturbées rejoignent gentiment les $\varepsilon^{(0)}(\vec{k})$, à des termes $\propto(V^2)$ près.

En particulier, le plus petit vecteur \vec{k} de cette sorte est $\vec{k} = \frac{\vec{K}}{2}$, parallèle à une direction de $\tilde{\mathcal{B}}$ et on a alors plus précisément :

$$\varepsilon_{\pm}(\vec{k} = \frac{\vec{K}}{2}) = \varepsilon^{(0)}(\vec{k} = \frac{\vec{K}}{2}) \pm |\mathcal{V}(\vec{K})| . \quad (9.139)$$

Il existe un vecteur primitif \vec{b}_0 de $\tilde{\mathcal{B}}$ tel que $\vec{K} = n\vec{b}_0$ où $n \in \mathbb{Z}$. Si n est impair, le *gap* apparaît en bord de zone ; au contraire, si n est pair, celui-ci se forme en centre de zone. Ainsi par exemple, si on part de $\vec{k} = 0$, le premier *gap* rencontré dans la direction \vec{b}_0 est situé en $\vec{k} = 1 \times \frac{\vec{b}_0}{2}$, il est au bord de BZ1 ; le suivant est en $\vec{k} = 2 \times \frac{\vec{b}_0}{2} = \vec{b}_0$. Dans le schéma en zones réduites, il est au centre de BZ1, et ainsi de suite.

Le schéma de $\varepsilon^{(0)}(\vec{k})$ en zones réduites montre bien l'apparition d'un *gap* à chaque fois que deux courbes représentant un arc translaté $\varepsilon^{(0)}(\vec{k} - \vec{K})$ se croisent. À l'inverse, si l'on revient à la parabole libre non périodisée, on voit qu'elle se transforme en une suite d'arcs disjoints, les (petites) discontinuités survenant pour $\vec{k} = \pm\frac{\vec{K}}{2}, \pm\vec{K}, \pm 3\frac{\vec{K}}{2}, \dots$. L'amplitude des *gaps* décroît quand l'énergie augmente, partant de l'idée de $\mathcal{V}(\vec{q})$ est une fonction décroissante de $\|\vec{q}\|$ (pour $V(\vec{r})$ donné, plus \vec{q} a un grand module, plus l'intégrale de Fourier est oscillante). Par ailleurs, évidemment, les *gaps* décroissent lorsque le potentiel V décroît et, à la limite $V = 0$, on retrouve la parabole libre, dénuée de discontinuités.

Il est utile de regarder d'un peu plus près le voisinage d'un *gap*. Revenons à (9.135) en choisissant $\vec{K}_1 = 0$ (alors $\vec{K} = \vec{K}_2$) et, pour la commodité, mesurons \vec{k} relativement à $\frac{\vec{K}}{2}$:

$$\vec{k} = \frac{1}{2}\vec{K} + \vec{q} , \quad (9.140)$$

où \vec{q} est orienté de façon *quelconque* par rapport à \vec{K} (\vec{k} n'est dans le plan de Bragg que si $\vec{q} \perp \vec{K}$). Alors $\vec{k} - \vec{K} = -\frac{1}{2}\vec{K} + \vec{q}$ et :

$$\Sigma\varepsilon = \frac{1}{2} \frac{\hbar^2}{2m} \left[\left(\vec{q} + \frac{1}{2}\vec{K} \right)^2 + \left(\vec{q} - \frac{1}{2}\vec{K} \right)^2 \right] = \varepsilon^{(0)}\left(\frac{\vec{K}}{2}\right) + \varepsilon^{(0)}(\vec{q}) , \quad \Delta\varepsilon = \frac{\hbar^2}{m} \vec{q} \cdot \vec{K} . \quad (9.141)$$

Les valeurs propres approchées données par (9.135) s'écrivent maintenant :

$$\varepsilon_{\pm}(\vec{k}) = \varepsilon^{(0)}\left(\frac{\vec{K}}{2}\right) + \varepsilon^{(0)}(\vec{q}) \pm \sqrt{\left(\frac{\hbar^2}{2m}\right)^2 (\vec{K} \cdot \vec{q})^2 + |\mathcal{V}(\vec{K})|^2} . \quad (9.142)$$

Quand de surcroît l'extrémité de \vec{k} est dans le plan de Bragg (ce que l'on note abusivement $\vec{k} \in \mathcal{P}_{\frac{\vec{K}}{2}}$), alors $\vec{q} \perp \vec{K}$ et il reste :

$$\varepsilon_{\pm}(\vec{k} \in \mathcal{P}_{\frac{\vec{K}}{2}}) = \varepsilon^{(0)}\left(\frac{\vec{K}}{2}\right) + \varepsilon^{(0)}(\vec{q}) \pm |\mathcal{V}(\vec{K})| = \varepsilon^{(0)}\left(\frac{\vec{K}}{2}\right) + \frac{\hbar^2}{2m}(\vec{k}^2 - \vec{k} \cdot \vec{K} + \frac{\vec{K}^2}{2}) \pm |\mathcal{V}(\vec{K})| , \quad (9.143)$$

en agrément avec (9.138). De ceci il résulte :

$$\left(\vec{\nabla}_{\vec{k}} \varepsilon_{\pm}\right)_{\vec{k} \in \mathcal{P}_{\frac{\vec{K}}{2}}} = \frac{\hbar^2}{m} \left(\vec{k} - \frac{\vec{K}}{2}\right) , \quad (9.144)$$

quantité qui s'annule bien en $\vec{k} = \frac{\vec{K}}{2}$ (voir (9.90)). Le gradient de ε_{\pm} donne la direction de la normale à la surface $\varepsilon_{\pm}(\vec{k}) = C^{\text{ste}}$. On voit géométriquement que, l'extrémité de \vec{k} étant dans $\mathcal{P}_{\frac{\vec{K}}{2}}$, $\vec{k} - \frac{\vec{K}}{2}$ est dans ce plan. En d'autres termes, toute surface d'énergie constante intersecte perpendiculairement tout plan de Bragg.

Par ailleurs, quand \vec{k} est perpendiculaire à $\mathcal{P}_{\frac{\vec{K}}{2}}$ (c'est-à-dire $\vec{k} \parallel \vec{K}$ et $\vec{q} \parallel \vec{K}$) :

$$\varepsilon_{\pm}(\vec{k} \parallel \vec{K}) = \varepsilon^{(0)}\left(\frac{\vec{K}}{2}\right) + \varepsilon^{(0)}(\vec{q}) \pm \sqrt{\left(\frac{\hbar^2 K q}{2m}\right)^2 + |\mathcal{V}(\vec{K})|^2} . \quad (9.145)$$

Cette expression permet de visualiser la variation des énergies de bande quand \vec{k} traverse un plan de Bragg tout en restant parallèle à \vec{K} ; quand \vec{k} est près de $\frac{\vec{K}}{2}$, \vec{q} est très petit et on peut faire un développement limité en q :

$$\begin{aligned} \varepsilon_{\pm}(\vec{k} \parallel \vec{K}, \vec{k} \simeq \frac{\vec{K}}{2}) &\simeq \varepsilon^{(0)}\left(\frac{\vec{K}}{2}\right) + \frac{\hbar^2 q^2}{2m} \pm |\mathcal{V}(\vec{K})| \left[1 + \frac{[\hbar^2 K q / (2m)]^2}{2|\mathcal{V}(\vec{K})|^2}\right] \\ &= \varepsilon^{(0)}\left(\frac{\vec{K}}{2}\right) \pm |\mathcal{V}(\vec{K})| + \frac{\hbar^2 q^2}{2m} \left(1 \pm \frac{\hbar^2 K^2}{4m|\mathcal{V}(\vec{K})|}\right) . \end{aligned} \quad (9.146)$$

Depuis le début, on fait l'hypothèse que le couplage de réseau est faible devant les autres énergies pertinentes ; se donnant l'ordre de grandeur a du pas du réseau, il faut donc notamment que :

$$|\mathcal{V}(\vec{K})| \ll \frac{\hbar^2}{ma^2} . \quad (9.147)$$

Il en résulte que le signe du coefficient de q^2 dans (9.146) est positif pour ε_+ , négatif pour ε_- . Ceci montre que, comme attendu, les deux bandes se repoussent et ressemblent à deux paraboles de concavités opposées, séparées, là où elles sont extrémales, par l'énergie $2|\mathcal{V}(\vec{K})|$, qui est l'énergie de *gap*. Au voisinage du *gap*, tout se passe comme si l'électron situé dans la bande ε_- avait une masse *négative*. Il est d'ailleurs tentant de définir une masse effective m_{eff} comme :

$$m_{\text{eff}} = \hbar^{-2} \left[\frac{\partial^2 \varepsilon}{\partial q^2} \right]^{-1} = \frac{m}{1 \pm \frac{\hbar^2 K^2}{4m} |\mathcal{V}(\vec{K})|} , \quad (9.148)$$

où m est la masse "nue" de l'électron. Ceci vaut lorsque \vec{k} est parallèle à \vec{K} . Plus généralement, on définit un tenseur de masse effective, d'éléments :

$$m_{\text{eff } uv} = \hbar^{-2} \left[\frac{\partial^2 \varepsilon}{\partial k_u \partial k_v} \right]^{-1} \quad (u, v = x, y, z) , \quad (9.149)$$

qui joue un rôle essentiel dans la description du transport électronique dans un métal.

Une fois les valeurs propres (approchées) obtenues, il est possible de trouver les fonctions propres. Pour simplifier les notations, on suppose dans la suite que le cristal possède la symétrie d'inversion⁴² :

$$V(-\vec{r}) = V(\vec{r}) , \quad (9.150)$$

⁴²C'est toujours vrai pour une structure cristalline sans motif, avec des objets à symétrie sphérique.

ce qui, allié au fait que V est réel, assure que $\mathcal{V}(\vec{K}) \in \mathbb{R}$. Les deux équations (9.133) et (9.134) peuvent être considérées comme la forme matricielle de l'équation aux vecteurs propres :

$$H_{\text{eff}} |\psi_{\vec{k}}\rangle = \varepsilon(\vec{k}) |\psi_{\vec{k}}\rangle \quad H_{\text{eff}} = \Sigma \varepsilon \mathbf{1} + \Delta \varepsilon \sigma_z + \mathcal{V}(\vec{K}) \sigma_x, \quad (9.151)$$

définie dans le sous-espace $\{|f_{\vec{k}}\rangle, |f_{\vec{k}-\vec{K}}\rangle\}$, avec $\langle \vec{r} | f_{\vec{q}} \rangle \equiv f_{\vec{q}}(\vec{r}) = e^{i\vec{q} \cdot \vec{r}}$. On a :

$$H_{\text{eff}} - \Sigma \varepsilon \mathbf{1} = h(\vec{k}) (\cos 2\theta \sigma_z + \sin 2\theta \sigma_x) = h(\vec{k}) e^{-i\theta \sigma_y} \sigma_z e^{+i\theta \sigma_y}. \quad (9.152)$$

où $h(\vec{k}) = \sqrt{[\Delta \varepsilon]^2 + |\mathcal{V}(\vec{K})|^2}$ (voir (9.135)) et avec :

$$\tan 2\theta = \frac{\mathcal{V}(\vec{K})}{\Delta \varepsilon}. \quad (9.153)$$

Les deux vecteurs propres $|\psi_{\vec{k}\pm}\rangle$ sont donc :

$$|\psi_{\vec{k}+}\rangle = e^{-i\theta \sigma_y} |f_{\vec{k}}\rangle \quad |\psi_{\vec{k}-}\rangle = e^{-i\theta \sigma_y} |f_{\vec{k}-\vec{K}}\rangle. \quad (9.154)$$

En effet, on a par exemple :

$$H_{\text{eff}} |\psi_{\vec{k}+}\rangle = h(\vec{k}) e^{-i\theta \sigma_y} \sigma_z e^{+i\theta \sigma_y} e^{-i\theta \sigma_y} |f_{\vec{k}}\rangle = h(\vec{k}) e^{-i\theta \sigma_y} \sigma_z |f_{\vec{k}}\rangle = +h(\vec{k}) e^{-i\theta \sigma_y} |f_{\vec{k}}\rangle \equiv +h(\vec{k}) |\psi_{\vec{k}+}\rangle, \quad (9.155)$$

et de même :

$$H_{\text{eff}} |\psi_{\vec{k}-}\rangle = h(\vec{k}) e^{-i\theta \sigma_y} \sigma_z e^{+i\theta \sigma_y} e^{-i\theta \sigma_y} |f_{\vec{k}-\vec{K}}\rangle = h(\vec{k}) e^{-i\theta \sigma_y} \sigma_z |f_{\vec{k}-\vec{K}}\rangle = -h(\vec{k}) e^{-i\theta \sigma_y} |f_{\vec{k}-\vec{K}}\rangle \equiv -h(\vec{k}) |\psi_{\vec{k}-}\rangle. \quad (9.156)$$

De (9.154) on déduit la forme explicite des fonctions de Bloch approchées $\psi_{\vec{k}\pm}$:

$$\psi_{\vec{k}+}(\vec{r}) = \cos \theta f_{\vec{k}}(\vec{r}) + \sin \theta f_{\vec{k}-\vec{K}}(\vec{r}) = e^{i\vec{k} \cdot \vec{r}} (\cos \theta + \sin \theta e^{-i\vec{K} \cdot \vec{r}}) \equiv e^{i\vec{k} \cdot \vec{r}} u_{\vec{k}+}(\vec{r}), \quad (9.157)$$

$$\psi_{\vec{k}-}(\vec{r}) = -\sin \theta f_{\vec{k}}(\vec{r}) + \cos \theta f_{\vec{k}-\vec{K}}(\vec{r}) = e^{i\vec{k} \cdot \vec{r}} (-\sin \theta + \cos \theta e^{-i\vec{K} \cdot \vec{r}}) \equiv e^{i\vec{k} \cdot \vec{r}} u_{\vec{k}-}(\vec{r}). \quad (9.158)$$

Chacune de ces fonctions redonne une simple onde plane dès que $\theta \simeq 0$: c'est ce qui arrive dès que $|\vec{K} \cdot (\vec{k} - \frac{\vec{K}}{2})|$ est supérieur ou de l'ordre de $m\hbar^{-2} |\mathcal{V}(\vec{K})| \ll a^{-2}$: dès que l'extrémité de \vec{k} s'éloigne un tout petit peu du plan de Bragg, les fonctions perturbées coïncident pratiquement avec les ondes planes, propres à l'ordre zéro. D'une façon générale, on trouve que $|\psi_{\vec{k}\pm}(\vec{r})|^2 = 1 \pm \sin 2\theta \cos \vec{K} \cdot \vec{r}$.

Les expressions (9.157) et (9.158) se simplifient lorsque l'extrémité de \vec{k} est précisément dans un plan de Bragg. Alors, $\Delta \varepsilon = 0$, $\theta = +\frac{\pi}{4}$ si $\mathcal{V}(\vec{K})$ est positif et :

$$\psi_{\vec{k}+}(\vec{r}) = \sqrt{2} e^{i(\vec{k} - \frac{\vec{K}}{2}) \cdot \vec{r}} \cos \frac{\vec{K} \cdot \vec{r}}{2}, \quad (9.159)$$

$$\psi_{\vec{k}-}(\vec{r}) = -i\sqrt{2} e^{i(\vec{k} - \frac{\vec{K}}{2}) \cdot \vec{r}} \sin \frac{\vec{K} \cdot \vec{r}}{2}. \quad (9.160)$$

Maintenant, les fonctions propres ne sont plus des ondes planes, mais des ondes stationnaires, combinaisons linéaires à poids égaux des deux exponentielles $e^{\pm i \frac{\vec{K}}{2} \cdot \vec{r}}$. C'est pourquoi il est usuel de dire que, en pareil cas, l'électron subit une réflexion de Bragg.

Enfin, il est visible que les $\psi_{\vec{k}\pm}(\vec{r})$, (9.157) et (9.158), satisfont le théorème de Bloch (les $u_{\vec{k}\pm}$ sont périodiques dans \mathcal{B}).

9.4.2 L'approximation des liaisons fortes

Il s'agit de la situation extrême opposée à celle que l'on vient de décrire : on va maintenant supposer que le potentiel ressenti par un électron au voisinage d'un cœur est essentiellement celui de ce dernier, le cristal pouvant alors être vu, en première approximation, comme une simple juxtaposition d'atomes, objets neutres en interaction faible. Il en résulte que les "bonnes" fonctions d'ordre zéro sont maintenant les fonctions atomiques centrées sur un atome donné, notées $\psi_{\nu \vec{R}}(\vec{r})$ (orbitales atomiques) ; ν désigne un ensemble de nombres quantiques. En gros, l'approximation consiste à supposer que, quand on s'éloigne d'un cœur, le potentiel de réseau $V(\vec{r})$, commence à différer du pur potentiel atomique quand on atteint une distance au moins égale à la portée des fonctions atomiques $\psi_{\nu \vec{R}}(\vec{r})$.

Toutes les fonctions $\psi_{\nu \vec{R}}$, ν fixé, sont en fait les répliques les unes des autres et se transforment les unes dans les autres par les translations du réseau de Bravais. Si on désigne simplement par $\psi_{\nu \vec{R}=\vec{0}}(\vec{r}) \equiv \psi_{\nu}(\vec{r})$ les fonctions centrées sur le nœud origine, on a :

$$\psi_{\nu \vec{R}}(\vec{r}) = \mathcal{T}(\vec{R}) \psi_{\nu}(\vec{r}) = \psi_{\nu}(\vec{r} - \vec{R}) . \quad (9.161)$$

À ce stade, on dispose *a priori* d'un jeu constitué par tous les états propres d'un Hamiltonien atomique (celui de l'atome situé au nœud origine $\vec{R} = \vec{0}$, H_0) et de toutes leurs répliques par le jeu des translations du réseau de Bravais. Cet ensemble est visiblement surcomplet, au sens où ses éléments ne sont pas tous linéairement indépendants les uns des autres. En effet, H_0 à lui seul offre une base complète ; en conséquence, toute fonction peut être développée sur cette base, en particulier n'importe quelle fonction du type $\psi_{\nu \vec{R} \neq \vec{0}}(\vec{r})$. Il existe donc des développements du genre :

$$\psi_{\nu \vec{R} \neq \vec{0}}(\vec{r}) = \sum_{\nu'} C_{\nu \nu'}(\vec{R}) \psi_{\nu'}(\vec{r}) . \quad (9.162)$$

avec leurs coefficients non tous identiquement nuls. Il en résulte que l'on ne doit pas retenir pour $\{\psi_{\nu}\}_{\nu}$ tous les états propres de H_0 , mais seulement un sous-ensemble. En pratique, d'ailleurs, on ne peut jamais introduire (pour des raisons techniques) qu'un petit nombre de fonctions de ce genre. En outre, l'idée physique sous-jacente à l'approximation en cours repose sur la notion d'état localisé : de ce point de vue, il serait stupide d'inclure parmi les ψ_{ν} les états non liés de H_0 – qui sont pourtant nécessaires pour assurer la complétude de la base propre de cet opérateur⁴³. Evacuer d'emblée ces états incongrus élimine de fait la difficulté liée à la surcomplétude.

D'ailleurs, finalement, rien n'oblige à choisir comme états localisés certains états propres de H_0 . Le caractère essentiel à préserver est la localisation autour de $\vec{R} = 0$, ce qui reste acquis quand on fait des combinaisons linéaires d'états $\psi_{\nu}(\vec{r})$: une combinaison linéaire en nombre fini d'états localisés donne clairement une fonction qui l'est encore. En définitive, l'ingrédient de base dans la suite est une (en fait plusieurs) fonction(s) notée(s) $\phi(\vec{r})$, localisée(s) dans le voisinage de $\vec{R} = \vec{0}$ et (mentalement) développable(s) sur un *petit* jeu de g fonctions atomiques ψ_{ν} localisées sur ce site :

$$\phi(\vec{r}) = \sum_{\nu=1}^g c_{\nu} \psi_{\nu}(\vec{r}) , \quad (9.163)$$

avec :

$$H_0 \psi_{\nu}(\vec{r}) = E_{\nu} \psi_{\nu}(\vec{r}) . \quad (9.164)$$

Les coefficients c_{ν} dans (9.163) sont déterminés suivant une procédure qu'il n'est pas utile de préciser ici et que l'on considèrera comme donnés. Celle-ci étant achevée, on dispose de g fonctions ϕ localisées autour du site $\vec{R} = 0$. Un label permet de les distinguer les unes des autres, le cas échéant, mais est omis dans la suite sauf nécessité. Propres du Hamiltonien atomique relatif à ce site, les énergies propres correspondantes sont bien séparées, leurs différences étant de l'ordre de l'eV.

Dans l'hypothèse des liaisons fortes, l'image physique à l'ordre zéro est celle d'un électron confiné près du site \vec{R} et placé dans une orbitale $\psi_{\nu \vec{R}}$ (ou une fonction du type ϕ ci-dessus), fortement localisée. \vec{R} variant,

⁴³L'inclusion de tels états délocalisés (étendus) ferait repartir clairement en direction de l'approximation des électrons presque libres.

toutes ces fonctions sont, à cet ordre, strictement équivalentes : on peut placer l'électron n'importe où, son énergie (à cet ordre) sera toujours la même⁴⁴. En réalité, l'électron a la possibilité de sauter d'un site à l'autre, ce qui se traduit par le fait que les éléments de matrice $\langle \psi_{\nu, \vec{R}} | H | \psi_{\nu', \vec{R}'} \rangle$ sont différents de zéro – quoique petits pour que le schéma de liaisons fortes ait un sens. L'existence de ce mouvement par sauts donne à l'électron une énergie cinétique, qui s'exprime par une loi de dispersion $\varepsilon(\vec{k})$. Comme l'amplitude des sauts est réputée petite, la variation de ε en \vec{k} sera faible : au total, la bande sera peu dispersée.

Outre $\phi(\vec{r})$, il convient donc d'introduire toutes les fonctions translatées :

$$\phi_{\vec{R}}(\vec{r}) = \mathcal{T}(\vec{R}) \phi(\vec{r}) = \phi(\vec{r} - \vec{R}) ; \quad (9.165)$$

notons dès maintenant que rien n'impose aux $\phi_{\vec{R}}$ d'être orthogonales entre elles ; en général :

$$\langle \phi_{\vec{R}} | \phi_{\vec{R}'} \rangle \neq 0 \quad \forall \vec{R}, \vec{R}' . \quad (9.166)$$

Quelle que soit l'image en arrière-plan (suggérant une approximation ou une autre), il s'agit toujours de trouver les états d'un électron dans un potentiel périodique $V(\vec{r})$, qui, de ce fait, satisfont le théorème de Bloch. À partir des $\phi_{\vec{R}}(\vec{r})$, il convient donc de trouver des états $\psi_{\vec{k}}(\vec{r})$ satisfaisant :

$$\mathcal{T}(\vec{R}) \psi_{\vec{k}}(\vec{r}) = e^{-i\vec{k} \cdot \vec{R}} \psi_{\vec{k}}(\vec{r}) , \quad (9.167)$$

qui est équivalent à :

$$\psi_{\vec{k}}(\vec{r} + \vec{R}) = e^{+i\vec{k} \cdot \vec{R}} \psi_{\vec{k}}(\vec{r}) . \quad (9.168)$$

Il existe des combinaisons linéaires des $\phi_{\vec{R}}(\vec{r})$ satisfaisant la relation (9.168). En effet, posons :

$$\psi_{\vec{k}}(\vec{r}) = \sum_{\vec{R}' \in \mathcal{B}} A_{\vec{k}}(\vec{R}') \phi_{\vec{R}'}(\vec{r}) = \sum_{\vec{R}' \in \mathcal{B}} A_{\vec{k}}(\vec{R}') \phi(\vec{r} - \vec{R}') . \quad (9.169)$$

En reportant dans (9.168), il vient :

$$\sum_{\vec{R}' \in \mathcal{B}} A_{\vec{k}}(\vec{R}') \phi(\vec{r} + \vec{R} - \vec{R}') = e^{i\vec{k} \cdot \vec{R}} \sum_{\vec{R}' \in \mathcal{B}} A_{\vec{k}}(\vec{R}') \phi(\vec{r} - \vec{R}') . \quad (9.170)$$

En posant $\vec{R}'' = \vec{R}' - \vec{R}$ au premier membre :

$$\sum_{\vec{R}'' \in \mathcal{B}} A_{\vec{k}}(\vec{R} + \vec{R}'') \phi(\vec{r} - \vec{R}'') = e^{i\vec{k} \cdot \vec{R}} \sum_{\vec{R}' \in \mathcal{B}} A_{\vec{k}}(\vec{R}') \phi(\vec{r} - \vec{R}') . \quad (9.171)$$

Les variables \vec{R}' et \vec{R}'' sont muettes, d'où :

$$A_{\vec{k}}(\vec{R} + \vec{R}') = e^{i\vec{k} \cdot \vec{R}} A_{\vec{k}}(\vec{R}') , \quad (9.172)$$

ce qui fournit la solution :

$$A_{\vec{k}}(\vec{R}) = C e^{i\vec{k} \cdot \vec{R}} \quad (C = \text{constante arbitraire}) . \quad (9.173)$$

En conclusion, étant donné un jeu de fonctions $\phi_{\vec{R}}$ déduites les unes des autres par les translations⁴⁵,

⁴⁴Le problème est donc très fortement dégénéré.

⁴⁵Noter toutefois que ce résultat ne dépend visiblement pas de l'aspect localisé éventuel des $\phi_{\vec{R}}(\vec{r})$: le point essentiel est que toutes ces fonctions se déduisent les unes des autres par translation.

toute fonction du type⁴⁶ :

$$\psi_{\vec{k}}(\vec{r}) = C \sum_{\vec{R} \in \mathcal{B}} e^{i\vec{k} \cdot \vec{R}} \phi_{\vec{R}}(\vec{r}) \quad (9.175)$$

satisfait le théorème de Bloch⁴⁷, quelle que soit la constante arbitraire C , choisie par normalisation⁴⁸. Elle peut se calculer en prenant des conditions cycliques dans une grande boîte contenant N sites au total ; en remarquant que les produits scalaires $\langle \phi_{\vec{R}} | \phi_{\vec{R}'} \rangle$ ne sont pas nuls, on trouve facilement⁴⁹ :

$$N |C|^2 \sum_{\vec{R}} e^{i\vec{k} \cdot \vec{R}} \langle \phi_{\vec{R}=\vec{0}} | \phi_{\vec{R}} \rangle = 1 . \quad (9.177)$$

Si l'on dispose de g fonctions localisées de type ϕ , $\{\phi_n\}$, et donc aussi de toutes leurs translatées $\{\phi_{n\vec{R}}\}$, on pourra construire g fonctions de Bloch à \vec{k} donné :

$$\psi_{n\vec{k}}(\vec{r}) = C \sum_{\vec{R} \in \mathcal{B}} e^{i\vec{k} \cdot \vec{R}} \phi_{n\vec{R}}(\vec{r}) \quad (9.178)$$

L'usage des fonctions $\psi_{n\vec{k}}$ simplifie considérablement le problème sur un plan technique. En effet, elles prédiagonalisent H , par le fait que, satisfaisant le théorème de Bloch, elles sont fonctions propres des opérateurs de translation :

$$\mathcal{T}(\vec{R}) \psi_{n\vec{k}}(\vec{r}) = e^{-i\vec{k} \cdot \vec{R}} \psi_{n\vec{k}}(\vec{r}) . \quad (9.179)$$

En effet, H commute avec tous les opérateurs \mathcal{T} :

$$[H, \mathcal{T}(\vec{R})] = 0 \iff H\mathcal{T}(\vec{R}) = \mathcal{T}(\vec{R})H . \quad (9.180)$$

Prenons l'élément de matrice de (9.180) entre deux fonctions $\psi_{n\vec{k}}$ et $\psi_{n'\vec{k}'}$:

$$\langle \psi_{n\vec{k}} | H\mathcal{T}(\vec{R}) | \psi_{n'\vec{k}'} \rangle = \langle \psi_{n\vec{k}} | \mathcal{T}(\vec{R})H | \psi_{n'\vec{k}'} \rangle . \quad (9.181)$$

En vertu de (9.179), le premier membre est :

$$e^{-i\vec{k}' \cdot \vec{R}} \langle \psi_{n\vec{k}} | H | \psi_{n'\vec{k}'} \rangle . \quad (9.182)$$

Par définition de l'adjoint d'un opérateur, le second membre de (9.181) est :

$$(|\psi_{n\vec{k}}\rangle, \mathcal{T}(\vec{R})H|\psi_{n'\vec{k}'}\rangle) = ([\mathcal{T}(\vec{R})]^\dagger |\psi_{n\vec{k}}\rangle, H|\psi_{n'\vec{k}'}\rangle) . \quad (9.183)$$

Comme les \mathcal{T} sont unitaires et que $[\mathcal{T}(\vec{R})]^{-1} = \mathcal{T}(-\vec{R})$, ceci vaut encore :

$$(\mathcal{T}(-\vec{R})|\psi_{n\vec{k}}\rangle, H|\psi_{n'\vec{k}'}\rangle) = (e^{-i\vec{k} \cdot (-\vec{R})} |\psi_{n\vec{k}}\rangle, H|\psi_{n'\vec{k}'}\rangle) = e^{-i\vec{k} \cdot \vec{R}} (|\psi_{n\vec{k}}\rangle, H|\psi_{n'\vec{k}'}\rangle) , \quad (9.184)$$

⁴⁶Il est intéressant de remarquer que, finalement, la fonction de bonne symétrie (9.175) peut être obtenue directement en faisant agir l'opérateur \mathbf{T} :

$$\mathbf{T} = \sum_{\vec{R} \in \mathcal{B}} e^{i\vec{k} \cdot \vec{R}} \mathcal{T}(\vec{R}) \quad (9.174)$$

sur une fonction $\phi_{\vec{R}_0}$ quelconque vis-à-vis de la symétrie de translation. Ce fait est général : pour toute symétrie (géométrique, de permutations, etc.), la théorie des groupes permet de définir les bons projecteurs construisant automatiquement une fonction de symétrie requise. Ces opérateurs sont des combinaisons linéaires des opérateurs du groupe de symétrie, les coefficients étant (ici) les valeurs propres de ces opérateurs (plus généralement, ce sont les caractères des opérateurs, définis relativement à un type de symétrie donné). Pour un ensemble de particules identiques, la construction de la fonction ayant la bonne symétrie de permutation relève du même traitement : on applique un opérateur $Y_1 = N!^{-1} \sum_{\lambda} \varepsilon_{\lambda} P_{\lambda}$ à une fonction de symétrie *quelconque* pour obtenir la fonction de symétrie voulue.

⁴⁷ $\psi_{\vec{k}}$ satisfait le théorème de Bloch, mais n'est pas pour autant un état de Bloch exact, au sens où la fonction $u_{\vec{k}}$ doit satisfaire sa propre équation, ce qui n'est pas assuré en général avec le choix limité des fonctions "atomiques" $\phi_{\nu\vec{R}}$.

⁴⁸Si on dispose de g fonctions sur chaque site, on peut définir g fonctions ϕ_n linéairement indépendantes, donc autant de fonctions $\psi_{n\vec{k}}$ satisfaisant le théorème de Bloch.

⁴⁹Au total, la fonction de Bloch normalisée est, pour une seule dimension d'espace :

$$\psi_{\vec{k}}(x) = \frac{1}{\sqrt{N}} \sum_{n \in \mathbb{Z}} e^{in\vec{k}a} \phi_n(x) . \quad (9.176)$$

où l'antilinearité dans le bra a été utilisée pour écrire la dernière égalité. La quantité à droite n'est autre que $e^{-i\vec{k}\cdot\vec{R}} \langle \psi_{n\vec{k}} | H | \psi_{n'\vec{k}'} \rangle$. En rapprochant maintenant (9.182) et (9.184), il vient :

$$(e^{-i\vec{k}'\cdot\vec{R}} - e^{-i\vec{k}\cdot\vec{R}}) \langle \psi_{n\vec{k}} | H | \psi_{n'\vec{k}'} \rangle = 0 , \quad (9.185)$$

d'où il résulte que si $\vec{k}' \neq \vec{k}$, l'élément de matrice de H est nul : les $\psi_{n\vec{k}}$ assurent bien une prédiagonalisation de H . Toutefois, H possède *a priori* des éléments de matrice non-nuls entre deux états de même k mais de n différents :

$$\langle \psi_{n\vec{k}} | H | \psi_{n'\vec{k}} \rangle \neq 0 . \quad (9.186)$$

À ce stade, le problème est complètement posé dans les meilleures conditions et on est réduit à résoudre une équation aux valeurs propres dans un espace de dimension g , pour lequel la procédure numérique est la plus efficace.

Afin d'extraire les idées essentielles, supposons qu'en fait *une seule* orbitale atomique est introduite sur chaque site ($g = 1$), $\psi_{1\vec{R}}$, réplique par translation de l'orbitale ψ_1 définie sur le site à l'origine. Les définitions précédentes restent valables, avec partout $g = 1$; l'indice ν ne prend désormais qu'une seule valeur $\nu = 1$ mais est maintenu pour des raisons de clarté dans les notations. Notons aussi que la fonction ϕ définie en (9.163) se réduit à ψ_1 :

$$\phi(\vec{r}) = \psi_1(\vec{r}) , \quad H_0\psi_1 = E_1\psi_1 . \quad (9.187)$$

On sent bien que cette simplification sera justifiée si les niveaux atomiques sont bien séparés en énergie et si la modification résultant de la présence du réseau n'introduit que des éléments de matrice petits devant ces différences d'énergie. Cette simplification étant admise, et en conséquence de (9.185), H se trouve représenté par une matrice complètement diagonale sur les N fonctions $\{\psi_{\vec{k}}\}_{\vec{k}}$. Il reste dès lors à calculer l'élément diagonal $\langle \psi_{\vec{k}} | H | \psi_{\vec{k}} \rangle$, qui n'est autre, dans cette approximation, que la valeur propre $\varepsilon(\vec{k})$ cherchée. Pour calculer cet élément de matrice, on s'y prend comme suit.

On part de :

$$\psi_{\vec{k}} = C \sum_{\vec{R} \in \mathcal{B}} e^{i\vec{k}\cdot\vec{R}} \psi_1(\vec{r} - \vec{R}) , \quad (9.188)$$

d'où :

$$\varepsilon(\vec{k}) = |C|^2 \sum_{\vec{R}} \sum_{\vec{R}'} e^{-i\vec{k}\cdot\vec{R}'} e^{i\vec{k}\cdot\vec{R}} \langle \psi_{1\vec{R}'} | H | \psi_{1\vec{R}} \rangle . \quad (9.189)$$

L'élément de matrice se transforme en jouant avec l'invariance par translation de H :

$$\begin{aligned} \langle \psi_{1\vec{R}'} | H | \psi_{1\vec{R}} \rangle &= \int d^3r \psi_1^*(\vec{r} - \vec{R}') H \psi_1(\vec{r} - \vec{R}) = \int d^3r \psi_1^*(\vec{r}) H \psi_1(\vec{r} - \vec{R} + \vec{R}') \\ &\equiv \langle \psi_1 | H | \psi_{1\vec{R}-\vec{R}'} \rangle , \end{aligned} \quad (9.190)$$

d'où :

$$\varepsilon(\vec{k}) = |C|^2 \sum_{\vec{R}} \sum_{\vec{R}'} e^{i\vec{k}\cdot(\vec{R}-\vec{R}')} \langle \psi_1 | H | \psi_{1\vec{R}-\vec{R}'} \rangle = |C|^2 \sum_{\vec{R}} \sum_{\vec{R}''} e^{i\vec{k}\cdot\vec{R}''} \langle \psi_1 | H | \psi_{1\vec{R}''} \rangle . \quad (9.191)$$

La sommation sur \vec{R} fournit N fois le même terme, de sorte que :

$$\varepsilon(\vec{k}) = N|C|^2 \sum_{\vec{R}''} e^{i\vec{k}\cdot\vec{R}''} \langle \psi_1 | H | \psi_{1\vec{R}''} \rangle . \quad (9.192)$$

Pour achever le calcul, on effectue la partition suivante du Hamiltonien. Le Hamiltonien d'un électron dans le réseau est :

$$H = \frac{\vec{p}^2}{2m} + \sum_{\vec{R} \in \mathcal{B}} v(\vec{r} - \vec{R}) , \quad (9.193)$$

où $v(\vec{r})$ est l'énergie potentielle d'un électron soumis à l'attraction du seul nœud situé à l'origine⁵⁰ ; en isolant le terme $\vec{R} = 0$:

$$H = \left[\frac{\vec{p}^2}{2m} + v(\vec{r}) \right] + \sum_{\vec{R} \in \mathcal{B}^*} v(\vec{r} - \vec{R}) \equiv H_0 + \Delta V , \quad (9.194)$$

⁵⁰On obtient bien le potentiel du réseau en effectuant sur $v(\vec{r})$ toutes les translations du réseau de Bravais.

\mathcal{B}^* désignant le réseau privé du site à l'origine. Maintenant, l'élément de matrice du Hamiltonien apparaissant dans (9.192) est :

$$\langle \psi_1 | H_0 + \Delta V | \psi_{1\vec{R}} \rangle = E_1 \langle \psi_1 | \psi_{1\vec{R}} \rangle + \langle \psi_1 | \Delta V | \psi_{1\vec{R}} \rangle , \quad (9.195)$$

d'où la nouvelle forme de (9.191) compte tenu de (9.177) :

$$\varepsilon(\vec{k}) = \frac{\sum_{\vec{R}} e^{i\vec{k}\cdot\vec{R}} (E_1 \langle \psi_1 | \psi_{1\vec{R}} \rangle + \langle \psi_1 | \Delta V | \psi_{1\vec{R}} \rangle)}{\sum_{\vec{R}} e^{i\vec{k}\cdot\vec{R}} \langle \psi_1 | \psi_{1\vec{R}} \rangle} = E_1 + \frac{\sum_{\vec{R}} e^{i\vec{k}\cdot\vec{R}} \langle \psi_1 | \Delta V | \psi_{1\vec{R}} \rangle}{\sum_{\vec{R}} e^{i\vec{k}\cdot\vec{R}} \langle \psi_1 | \psi_{1\vec{R}} \rangle} . \quad (9.196)$$

En raison du caractère localisé des $\psi_{1\vec{R}}(\vec{r})$, les éléments de matrice ont une dépendance grossièrement exponentielle par rapport à la distance $\|\vec{R}\|$ à l'origine. Le moins que l'on puisse faire est donc de retenir les seuls couples correspondant à des sites au plus premiers voisins entre eux. On obtient ainsi dans un premier temps :

$$\varepsilon(\vec{k}) = E_1 + \frac{\langle \psi_1 | \Delta V | \psi_1 \rangle + \sum_{\langle \vec{0}, \vec{R} \rangle} e^{i\vec{k}\cdot\vec{R}} \langle \psi_1 | \Delta V | \psi_{1\vec{R}} \rangle}{1 + \sum_{\langle \vec{0}, \vec{R} \rangle} e^{i\vec{k}\cdot\vec{R}} \langle \psi_1 | \psi_{1\vec{R}} \rangle} . \quad (9.197)$$

où les crochets indiquent que la sommation se fait sur les seuls premiers voisins du site origine $\vec{R} = \vec{0}$. Les produits scalaires de la somme du dénominateur sont usuellement très petits, et sont négligés le plus souvent. Dans ces conditions, on se retrouve avec :

$$\varepsilon(\vec{k}) = E_1 + \langle \psi_1 | \Delta V | \psi_1 \rangle + \sum_{\langle \vec{0}, \vec{R} \rangle} e^{i\vec{k}\cdot\vec{R}} \langle \psi_1 | \Delta V | \psi_{1\vec{R}} \rangle . \quad (9.198)$$

Le second terme peut être absorbé dans une redéfinition⁵¹ de E_1 , notée ε_0 ; avec ces diverses approximations, le résultat final retenu est :

$$\varepsilon(\vec{k}) = \varepsilon_0 + \sum_{\langle \vec{0}, \vec{R} \rangle} e^{i\vec{k}\cdot\vec{R}} \langle \psi_1 | \Delta V | \psi_{1\vec{R}} \rangle \quad (9.199)$$

c'est la dépendance en \vec{k} de la somme qui introduit la dispersion de la bande, laquelle reste faible puisque les éléments de matrice de ΔV restent petits par rapport aux énergies fixant l'ordre de grandeur de ε_0 .

À titre d'exemple, soit un réseau cubique simple ; le nœud origine a six premiers voisins situés à la distance a et tous les éléments de matrice sont égaux entre eux. En notant $w \in \mathbb{R}$ cette valeur commune, l'expression (9.199) donne :

$$\varepsilon(\vec{k}) = \varepsilon_0 + 2w (\cos k_x a + \cos k_y a + \cos k_z a) . \quad (9.200)$$

Cette expression est visiblement périodique ($\varepsilon(\vec{k} + \vec{K}) = \varepsilon(\vec{k})$). Elle définit une bande d'énergie s'étendant entre les deux valeurs⁵² $\varepsilon_0 \pm 6|w|$; mesurée à la bonne aune, la dispersion est effectivement faible (bande étroite).

Remarque

D'une façon générale (et inversement), à partir de tout état de Bloch exact⁵³ $\psi_{\vec{k}}$, on peut indépendamment définir sa transformée de Fourier par :

$$\psi_{\vec{k}}(\vec{r}) = \frac{1}{\sqrt{N}} \sum_{\vec{R} \in \mathcal{B}} e^{i\vec{k}\cdot\vec{R}} \phi_{\vec{R}}^{(W)}(\vec{r}) \iff \phi_{\vec{R}}^{(W)}(\vec{r}) = \frac{1}{\sqrt{N}} \sum_{\vec{k} \in BZ1} e^{-i\vec{k}\cdot\vec{R}} \psi_{\vec{k}}(\vec{r}) . \quad (9.201)$$

⁵¹ou, pour faire pédant, une "renormalisation".

⁵²Le signe de w est sans incidence physique, comme il se doit puisqu'il s'agit d'un élément de matrice non-diagonal. Pour y voir plus clair, il suffit de regarder en détails le cas unidimensionnel (voir note 49) ; le changement de signe de w peut s'obtenir par un rephasage $(-1)^n$ de chaque ϕ_n , ce qui remplace simplement k par $k - \frac{\pi}{a} (\frac{2n}{a})$: rien ne change (physiquement), que l'on se place dans la première zone de Brillouin (k dans l'intervalle $[-\frac{\pi}{a}, \frac{\pi}{a}]$) ou dans une zone décalée de $\frac{\pi}{a}$ ($k \in [0, \frac{2\pi}{a}]$).

Très souvent, on choisit $w < 0$ pour faire commodément les petits dessins dans la première zone de Brillouin : on remplit alors la cuvette par le centre, au lieu de remplir l'"anti-cuvette" par les bords...

⁵³c'est-à-dire obtenu en résolvant exactement l'équation pour $u_{\vec{k}}$.

La fonction $\phi_{\vec{R}}^{(W)}$ ainsi définie porte le nom de fonction de Wannier. Contrairement aux fonctions atomiques $\psi_{\nu \vec{R}}$ ou $\phi_{\vec{R}}$ introduites ci-dessus, deux fonctions de Wannier relatives à deux sites distincts sont orthogonales :

$$\langle \phi_{\vec{R}}^{(W)} | \phi_{\vec{R}'}^{(W)} \rangle = \delta_{\vec{R} \vec{R}'} , \quad (9.202)$$

et constituent une autre base possible de représentation ; l'orthogonalité des fonctions de Wannier résulte du fait que la transformation (9.201) est unitaire, les $\psi_{\vec{k}}(\vec{r})$ étant mutuellement orthogonales. En tant que transformée de Fourier d'une fonction étendue, $\phi_{\vec{R}}^{(W)}(\vec{r})$ est une fonction présentant une forte localisation, un caractère qu'elle partage avec les orbitales atomiques $\psi_{\nu \vec{R}}$ ou les fonctions $\phi_{\vec{R}}$ introduites en (9.165).

9.5 Conducteurs, semi-conducteurs, isolants

Les deux situations extrêmes qui viennent d'être étudiées mettent en évidence un phénomène indissolublement lié à l'existence du potentiel périodique : l'ouverture de petits *gaps* en énergie par comparaison avec la distribution continue en énergie, de 0 à $+\infty$, pour un électron libre faiblement perturbé par le potentiel périodique, ou l'apparition d'une (faible) dispersion des états atomiques pour un électron fortement lié. En raison de la périodicité et du théorème de Bloch qui en résulte, en parcourant l'axe des énergies on rencontre des zones où il existe des états (bandes permises) et d'autres où il n'y en a pas (bandes interdites).

Cette propriété peut être utilement précisée à l'aide de la densité d'états, $\mathcal{D}(E)$; par définition, on a :

$$\text{Nombre d'états d'énergie comprise entre } E_0 \text{ et } E_0 + \delta E = \mathcal{D}(E_0) \delta E . \quad (9.203)$$

$\mathcal{D}(E_0) \delta E$ est une mesure de la dégénérescence et donne le nombre d'états de "même" énergie E_0 se distinguant les uns des autres par les autres nombres quantiques nécessaires pour spécifier complètement l'état quantique. Un électron dans un potentiel périodique est caractérisé par sa pseudo-impulsion $\hbar \vec{k}$, et son spin. En l'absence de champ magnétique, ou de tout effet spin-orbite, l'énergie d'un électron dans une bande donnée, $\varepsilon_n(\vec{k})$, ne dépend pas du spin. Le nombre d'états d'énergie E_0 à δE près s'obtient en sommant sur toutes les valeurs de \vec{k} donnant $\varepsilon_n(\vec{k}) = E_0$, et sur le spin. Ce dernier fournit un simple facteur 2, d'où ($\delta^3 k = (\frac{2\pi}{L})^3$) :

$$\sum_{\text{états}} \equiv 2 \sum_{\vec{k}} \rightarrow 2 \left(\frac{L}{2\pi} \right)^3 \int d^3 k . \quad (9.204)$$

En conséquence, la densité d'états de la bande d'énergie $\varepsilon_n(\vec{k})$ est donnée par⁵⁴ :

$$\mathcal{D}_n(E_0) = 2L^3 \int_{\mathbb{R}^3} \frac{d^3 k}{(2\pi)^3} \delta(\varepsilon_n(\vec{k}) - E_0) \quad ([\mathcal{D}] = E^{-1}) . \quad (9.205)$$

Autrement dit, $\mathcal{D}_n(E)$ résulte de la sommation sur tous les états possibles filtrée à l'aide de la fonction de Dirac. On sait que pour une fonction monotone⁵⁵ :

$$\delta(f(x) - f_0) = \frac{1}{|f'(x_i)|} \delta(x - x_i) . \quad (9.207)$$

À trois dimensions, cette relation se généralise en⁵⁶ :

$$\delta(f(\vec{x}) - f_0) = \frac{1}{\|\vec{\nabla} f(\vec{x}_0)\|} \delta((\vec{x} - \vec{x}_0) \cdot \vec{n}) , \quad f(\vec{x}_0) = f_0 , \quad (9.211)$$

⁵⁴La densité d'états est parfois définie par unité de volume, ce qui revient à poser $L = 1$ dans la définition (9.205).

⁵⁵Si la fonction n'est pas monotone, l'égalité $f(x) = f_0$ peut être satisfaite pour plusieurs valeurs x_i de x ; alors :

$$\delta(f(x) - f_0) = \sum_i \frac{1}{|f'(x_i)|} \delta(x - x_i) \quad f(x_i) = f_0 \quad \forall i . \quad (9.206)$$

⁵⁶En effet, le développement de Taylor au voisinage de $\vec{x}_0 \in S_{f_0}$ s'écrit :

$$f(\vec{x}) = f_0 + (\vec{x} - \vec{x}_0) \cdot \vec{\nabla} f(\vec{x}_0) + \mathcal{O}((\vec{x} - \vec{x}_0)^2) . \quad (9.208)$$

où \vec{n} est le vecteur unitaire de la normale à la surface S_{f_0} d'équation $f(\vec{r}) = f_0$. Avec ces résultats, la densité (9.205) s'écrit :

$$\mathcal{D}_n(E_0) = L^3 \int_{\mathbb{R}^3} \frac{d^3k}{4\pi^3} \frac{1}{\|\vec{\nabla}_{\vec{k}} \varepsilon_n(\vec{k}_0)\|} \delta((\vec{k} - \vec{k}_0) \cdot \vec{n}) , \quad \varepsilon_n(\vec{k}_0) = E_0 . \quad (9.212)$$

La présence de la fonction de Dirac dit que seuls les points où $(\vec{k} - \vec{k}_0) \cdot \vec{n}$ est nul contribuent à l'intégrale, c'est-à-dire tous les points dont l'écart à \vec{k}_0 compté le long de la normale s'annule : il s'agit donc de tous les points dont le vecteur \vec{k} est *sur* la surface. Une expression équivalente à (9.212) pour la fonction $\mathcal{D}_n(E)$ est donc :

$$\mathcal{D}_n(E) = L^3 \int_{S_E} \frac{d^2k}{4\pi^3} \frac{1}{\|\vec{\nabla}_{\vec{k}} \varepsilon_n(\vec{k})\|} , \quad (9.213)$$

où S_E est la surface d'équation $\varepsilon_n(\vec{k}) = E$. Ces définitions se généralisent en toute dimension. Par exemple, pour un réseau plan, la densité \mathcal{D} s'obtient en effectuant un intégrale de *ligne* et non de surface, le long de la ligne d'énergie constante $\varepsilon_n(\vec{k}) = E$, \vec{k} étant maintenant un vecteur à deux composantes k_x, k_y .

À l'intérieur d'une maille primitive de $\tilde{\mathcal{B}}$, $\varepsilon_n(\vec{k})$ présente au moins un minimum et un maximum, en tant que fonction continue différentiable et périodique. En ces *extrema*, $\vec{\nabla}_{\vec{k}} \varepsilon_n$ s'annule et l'intégrand diverge. À $d = 3$, ces singularités sont intégrables et $\mathcal{D}_n(E)$ reste une fonction continue, tout en présentant des points anguleux appelés singularités de van Hove. Au contraire, à $d = 1$, quand $\vec{\nabla}_{\vec{k}} \varepsilon_n(k)$ s'annule, \mathcal{D}_n diverge logarithmiquement.

Pour une énergie donnée E , la densité totale s'obtient en faisant simplement la somme des densités des bandes :

$$\mathcal{D}(E) = \sum_n \mathcal{D}_n(E) , \quad (9.214)$$

de sorte que l'allure générale de $\mathcal{D}(E)$ est celle indiquée sur la figure 9.3, sans le souci de faire figurer les détails (et en particulier les singularités de van Hove). Bien sûr, il peut arriver – c'est fréquent – que deux bandes se recouvrent en énergie (voir fig. 9.3, à droite). En tout état de cause, le spectre des énergies à une particule d'un solide périodique n'est ni simplement connexe, ni purement ponctuel. En quelque sorte, c'est un hybride entre celui d'un électron libre et celui d'un électron lié dans un atome.

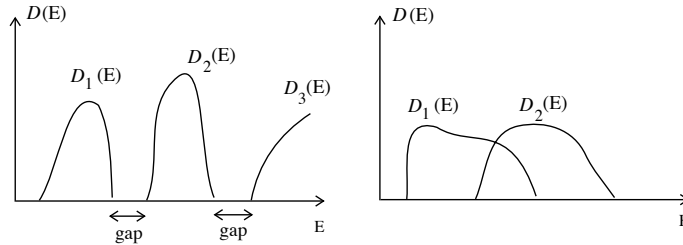


Figure 9.3: Illustration schématique de la densité d'états. À droite, l'exemple de deux bandes qui se recouvrent.

$\mathcal{D}(E)$ est tout particulièrement utile pour calculer les valeurs moyennes de grandeurs ne dépendant *que* de l'énergie ; en effet, $A(E)$ désignant une telle grandeur, sa moyenne contient des sommes du genre :

$$\mathcal{A} = \sum_{\text{états}} A(E) \quad \rightarrow \quad 2 \left(\frac{L}{2\pi} \right)^3 \int d^3k A(E) , \quad (9.215)$$

Par définition, $\vec{\nabla} f$ est perpendiculaire à la surface S_{f_0} , donc dirigé suivant la normale \vec{n} ; il en résulte que (9.208) se réécrit comme :

$$f(\vec{x}) = f_0 + \|\vec{x} - \vec{x}_0\| \|\vec{\nabla} f(\vec{x}_0)\| + \mathcal{O}((\vec{x} - \vec{x}_0)^2) , \quad (9.209)$$

ce qui donne :

$$\delta(f(\vec{x}) - f_0) = \delta(\|\vec{\nabla} f(\vec{x}_0)\|) (\vec{x} - \vec{x}_0) \cdot \vec{n} , \quad (9.210)$$

et compte tenu de (9.207), on trouve (9.211). On vérifie sans peine l'homogénéité de (9.211), sachant que $[\delta(\vec{x} \cdot \vec{n})] = L^{-1}$ (alors que $[\delta(\vec{x})] = L^{-3}$).

où le facteur 2 prend en compte le spin (ce qui suppose en outre que $A(E)$ ne dépend pas non plus du spin). L'intégration peut s'effectuer en deux temps : pour une bande de loi de dispersion $\varepsilon(\vec{k})$, on fixe la valeur de l'énergie et on intègre sur \vec{k} à E fixée :

$$\mathcal{A} = 2 \left(\frac{L}{2\pi} \right)^3 \int d^3k \delta(\varepsilon(\vec{k}) - E) A(E) . \quad (9.216)$$

Avec les règles habituelles concernant la fonction de Dirac rappelées plus haut, (9.216) devient :

$$2 \left(\frac{L}{2\pi} \right)^3 \int dE \int_{S_E} d^2k \frac{1}{\|\vec{\nabla}\varepsilon(\vec{k})\|} A(E) , \quad (9.217)$$

Au total, on a donc :

$$\mathcal{A} = \int dE \mathcal{D}(E) A(E) \equiv \int dE \tilde{A}(E) , \quad (9.218)$$

Les distinctions quantitatives des deux situations extrêmes étudiées plus haut peuvent s'illustrer comme indiqué sur la figure 9.4 : pour les électrons presque libres, les bandes sont larges et séparées par des petits *gaps* ; la situation est inverse pour un électron en liaisons fortes.

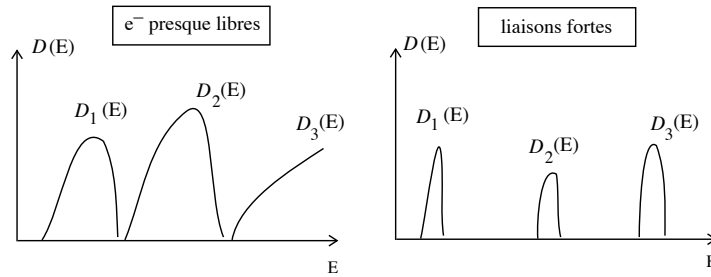


Figure 9.4: Illustration schématique des différences de la densité d'états pour les approximations des électrons presque libres (à gauche) et des liaisons fortes (à droite).

On est maintenant en mesure de comprendre – en raisonnant à température nulle⁵⁷ – l'existence de deux grandes classes de matériaux, les conducteurs et les isolants. La discussion repose crucialement sur l'hypothèse de la validité d'une théorie à une particule (ou d'une théorie de champ moyen), ce qui permet de s'appuyer sur l'existence d'états propres bien définis pour un seul électron (nu ou habillé), essentiellement corrélé aux autres par le Principe de Pauli.

Il s'agit de construire l'état du cristal – l'état fondamental –, étant donné un certain nombre d'électrons à caser, N_e . Ces derniers proviennent usuellement de la couche de valence des atomes qui s'assemblent pour former le cristal. Chaque atome fournit un petit nombre d'électrons susceptibles de participer au transport électrique, n_e ($n_e = 1$ pour un alcalin, $n_e = 8$ pour Fe, dont la configuration électronique est $[\text{Ar}] (3d)^6(4s)^2$) ; s'il y a N mailles primitives et un atome par maille (cristal monoatomique), il y a donc $N_e = n_e N$ électrons à placer dans les bandes à une particule, $\varepsilon_n(\vec{k})$, chacune d'entre elles pouvant accueillir $2N$ électrons.

Construire l'état fondamental, c'est donc remplir les états de Bloch les uns après les autres, en se déplaçant évidemment dans le sens des énergies croissantes. Pour simplifier, on supposera dans un premier temps que les bandes ne se recouvrent pas en énergie. Avec cette hypothèse, il existe une et une seule plus haute bande en énergie contenant des électrons, appelée bande de conduction⁵⁸. Deux cas sont alors possibles :

1. cette plus haute bande est pleine et contient donc $2N$ électrons. Alors, tout électron est coincé en énergie, sauf à gagner une énergie au moins égale à E_g , qu'aucun champ extérieur ne peut lui fournir. À température

⁵⁷Pour les métaux usuels, ε_F est de l'ordre de 10^4 eV ou plus, et on a toujours $k_B T \ll \varepsilon_F$; la distribution de Fermi reste donc toujours très voisine de la fonction de Heaviside $Y(\varepsilon_F - E)$.

⁵⁸L'avant-dernière bande complètement remplie s'appelle bande de valence.

nulle, le solide est donc un isolant. À température finie, les fluctuations thermiques ne peuvent rien changer tant que $k_B T \ll E_g$; comme $(k_B T)_{\text{ambiante}} \simeq 25 \text{ meV}$, le changement de comportement conductif induit par la température finie dépend de la valeur numérique de E_g :

- si E_g est de l'ordre de quelques eV (par exemple, pour le diamant $E_g = 5.5 \text{ eV}$), les probabilités d'excitation varient en gros comme $e^{-40 \times 5.5}$: autant dire qu'elles sont nulles. Le solide est isolant à toute température où il existe en tant que tel
- si E_g est nettement plus petit, une fraction d'eV pour fixer les idées (exemple : le germanium pour lequel $E_g \simeq 0.67 \text{ eV}$ à l'ambiante⁵⁹), les probabilités sont de l'ordre de e^{-30} , ce qui est encore très petit. Cependant, compte tenu de la relativement haute densité de porteurs potentiels, la conductivité devient notable, quoique très petite par rapport à un bon conducteur comme le cuivre. Le solide est alors dit semi-conducteur (intrinsèque⁶⁰). La résistivité d'un semi-conducteur *décroît* quand la température augmente, ce qui est le comportement inverse de celui observé dans les métaux.

- la plus haute bande est incomplètement remplie, ce qui situe le niveau de Fermi ε_F à l'intérieur de cette bande⁶¹ à la hauteur du dernier niveau occupé (niveau d'eau de la mer de Fermi). Les électrons qui se trouvent dans le voisinage inférieur de ε_F se voient offerts des niveaux d'énergie infiniment proches : dès lors, un champ électrique, aussi petit soit-il, peut de fait les mettre en mouvement et leur donner une énergie "orientée" – qui reste toutefois minuscule par rapport à l'énergie thermique – donnant lieu à un vecteur courant, donc au passage d'un courant électrique.

À l'intérieur d'une maille primitive de $\tilde{\mathcal{B}}$, il y a autant d'états \vec{k} distincts qu'il y a de mailles primitives dans \mathcal{B} , et chacun peut recevoir deux électrons compte tenu du spin. Si \mathcal{B} n'a pas de base (cristal monoatomique) et si chaque atome fournit un seul électron ($n_e = 1$), alors la dernière bande est à moitié remplie⁶² – toujours dans l'hypothèse de bandes qui ne se recouvrent pas – et le cristal est conducteur ; c'est encore vrai si n_e est impair. La situation est inversée si n_e est pair, et le solide est isolant⁶³

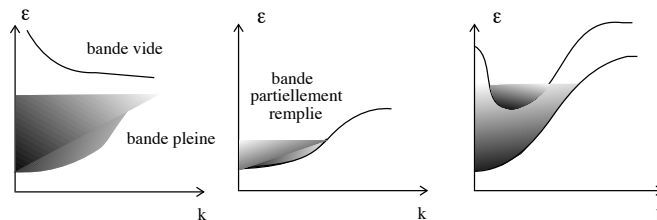


Figure 9.5: Illustration schématique selon la théorie des bandes (théorie à *une* particule) de la différence entre isolant (à gauche), métal (au milieu) et semi-métal (à droite).

L'existence de recouvrement⁶⁴ entre les bandes modifie sérieusement ces dernières conclusions. En pareil cas, si n_e est impair, les dernières bandes se partagent les électrons et on n'obtient évidemment pas un demi-remplissage ; toutefois, rien n'est changé fondamentalement en ce qui concerne les propriétés de conduction : le solide reste conducteur. Au contraire, quand n_e est pair, un solide qui est isolant en l'absence de recouvrement devient conducteur puisque les deux bandes qui se recouvrent sont inévitablement maintenant incomplètes.

⁵⁹La dilatation du réseau avec la température donne une dépendance en température aux énergies des *gaps*.

⁶⁰Cette précision est là pour le distinguer d'un matériau où l'on a délibérément introduit des impuretés (dopage). Cette opération produit des états à l'intérieur du *gap*, qui sont autant de barreaux d'échelle permettant le franchissement par étapes de ce dernier.

⁶¹ ε_F est la limite du potentiel chimique $\mu(T)$ lorsque $T \rightarrow 0$. Pour un isolant ou un semi-conducteur, cette limite se situe dans le *gap* ; sa position précise dépend des courbures des bandes de conduction et de valence. Si celles-ci ont la même courbure, $\varepsilon_F \equiv \mu(0+)$ est au milieu du *gap*.

⁶²On parle alors de *demi-remplissage*.

⁶³Usuellement, pour un solide, on appelle *bande de valence* une bande pleine, et *bande de conduction* une bande partiellement remplie. Dans tous les cas, ce sont les électrons de valence *atomiques* qui les occupent.

⁶⁴Un cas limite est celui où deux bandes, l'une pleine, l'autre vide, entrent en contact sans toutefois se dépasser l'une l'autre en un nombre fini de points de BZ1. En pareil cas, le solide n'est pas conducteur (toujours à température nulle) puisque le canal dans l'espace de \vec{k} permettant de passer de la bande pleine à la bande vide est "infiniment" étroit (dit autrement, la frontière à l'énergie de Fermi ε_F est de mesure nulle).

Il importe de réaliser le rôle crucial joué par la *dispersion* des bandes électroniques, c'est-à-dire le fait que leur énergie ε dépend de \vec{k} : un électron dans une bande plate ne peut pas gagner d'énergie sous l'effet d'un champ, même si la bande est quasi-vide. Quand la bande est partiellement remplie, les électrons peuvent certes passer d'un état \vec{k} à un état \vec{k}' sous l'effet de collisions (entre eux, sur les ions du réseau, sur des impuretés, . . .), mais leur énergie ne change pas puisqu'elle est précisément indépendante de la pseudo-impulsion. La "platitude" de l'énergie d'une bande se reporte immédiatement sur la vitesse (moyenne) d'un électron de Bloch, en vertu de la relation $\vec{v} = \hbar^{-1} \vec{\nabla}_{\vec{k}} \varepsilon$: quand la bande est plate, la vitesse est nulle.

Dans les cas moins extrêmes, les grandeurs relatives des composantes de $\vec{\nabla}_{\vec{k}} \varepsilon$ sont l'une des causes premières des anisotropies de conductivité (d'autres facteurs doivent être pris en compte, notamment la densité d'états au niveau de Fermi). Pour un solide qui se clive facilement (le carbone graphite, par exemple), cette propriété de clivage provient principalement du fait que la distance entre deux plans de réseau en nid d'abeille est plusieurs fois supérieure⁶⁵ à la plus petite distance entre deux atomes de carbone d'un même plan. Comme la dispersion des bandes est pilotée par les amplitudes de saut – lesquelles ont en gros une dépendance exponentielle en distance –, il en résulte que la dispersion des bandes dans la direction perpendiculaire aux plans est beaucoup plus faible que parallèlement aux plans, d'où une forte anisotropie de conductivité. Notons aussi qu'une bande plate donne *ipso facto* une masse effective infinie, une autre façon de garder à l'esprit l'importance fondamentale de la dispersion.

On retiendra également le lien semi-quantitatif entre dispersion et délocalisation des états. Pour les électrons presque libres, la dispersion est très grande, cependant que les états sont essentiellement des ondes planes, faiblement modulées en amplitude par la fonction $u_{\vec{k}}$; il en résulte que la densité électronique est très diffuse⁶⁶ et que, si on prépare un paquet d'ondes localisé (sur un seul site, par exemple), il s'étale très vite. Au contraire, pour des électrons en liaisons fortes, la dispersion est faible ; les états propres sont toujours étendus, puisqu'il s'agit d'états de Bloch, mais ils présentent une forte structure spatiale, traduisant le fait que l'électron, en quelque sorte, passe par effet-tunnel d'un site à l'autre. Ici encore, un paquet d'ondes localisé finit par se diluer sur le réseau, mais l'évolution se fait sur une échelle de temps beaucoup plus longue ($\sim \frac{\hbar}{t}$, t désignant toujours l'amplitude de saut élémentaire). À la limite d'une amplitude de saut infiniment petite, l'énergie ne dépend pas de \vec{k} , la dégénérescence est $\sim N$ et les fonctions atomiques localisées sont, tout autant que les états de Bloch ou n'importe quelle autre combinaison linéaire, des fonctions propres du Hamiltonien d'un électron. En définitive, la dispersion ou la non-dispersion des bandes est un indice commode de la nature, localisée ou étendue, des états électroniques.

Remarques

1. Comme rappelé au début de la discussion, cette explication élémentaire de la distinction entre conducteurs et isolants repose crucialement sur l'hypothèse de la validité d'une théorie à une particule, où les interactions entre électrons sont oubliées – ou en tout cas lorsqu'elles peuvent être traitées par une méthode de type champ moyen qui, par nature, restitue formellement *in fine* le cadre d'un modèle à particules indépendantes. Lorsque tel n'est pas le cas, la question de la distinction isolant *vs* conducteur doit être entièrement reconsidérée. Il est facile de comprendre pourquoi en raisonnant physiquement par référence à un cas-limite.

Pour simplifier la discussion, supposons que chaque atome apporte un seul électron (demi-remplissage), auquel cas, selon la théorie des bandes, on est en présence d'un conducteur. Introduisons maintenant l'interaction entre les électrons (on sort du schéma des bandes) et admettons que leur répulsion mutuelle (traditionnellement notée U) est si forte que leur énergie cinétique puisse être négligée devant l'énergie potentielle d'interaction. Il est bien clair, alors, que chaque électron a envie de rester sur un atome : dès qu'il veut en bouger, les autres le repoussent fortement. En définitive, dans cette image extrême, les électrons eux-mêmes ont tendance à former un réseau statique⁶⁷, coïncidant pour cette situation de demi-remplissage, au réseau de Bravais sous-jacent du solide⁶⁸. Il est physiquement évident qu'un tel solide ne

⁶⁵environ 2.4 fois plus grande.

⁶⁶On parle alors usuellement d'états *étendus*.

⁶⁷Dans le régime des faibles densités ($r_s \gg a_0$), c'est le réseau de Wigner.

⁶⁸Quand on n'est pas à demi-remplissage, la densité d'électrons n'est pas commensurable avec la densité d'atomes, et la physique devient beaucoup plus riche (possibilité d'ondes de densité de charge, appelées CDW, acronyme de *charge density waves*).

peut être conducteur – tous les porteurs de charge sont bloqués. Il est non moins évident qu’il n’y a pas de contradiction : le schéma explicatif en terme de bande est invalide, puisqu’il résulte d’une théorie à une particule. Pour cette raison, il s’agit d’une classe particulière d’isolant : on l’appelle *isolant de Mott*.

Dans un cas moins extrême, la discussion repose sur les valeurs comparées de la répulsion électronique – d’ordre U quand deux électrons sont sur le même site – et de l’énergie cinétique, dont l’ordre de grandeur est noté (traditionnellement) t . Tout (ou presque) dépend du rapport $\alpha = \frac{t}{U}$. Conceptuellement, on peut imaginer une variation graduelle de ce rapport. Quand il est très grand, le solide considéré ci-dessus est sûrement conducteur ; quand il est très petit, on a un isolant. Le passage d’une phase à l’autre porte le nom de transition de Mott, où l’on voit la conductivité chuter brutalement à zéro quand α décroît au voisinage d’une certaine valeur critique α_C , comportement reminiscent d’une transition de phase survenant à une certaine température critique T_C .

Physiquement, on peut imaginer que cette variation graduelle se fait en dilatant progressivement le réseau d’un alcalin. Pour les valeurs usuelles de la maille, le solide est conducteur ; distendre le réseau fait décroître α . Quand le paramètre a de la maille est très grand, les “ions” se trouvent très loin les uns des autres et la structure est celle d’un antiferromagnétique, les atomes portant chacun un électron dont le spin alterne d’un site à l’autre (ceci pour l’état singulet de spin total nul). Quand a décroît, t augmente puisque le passage d’un site à l’autre est facilité, le rapport $\frac{t}{U}$ croît et la transition de Mott peut se produire, le cas échéant. En pratique, la variation de a peut s’obtenir en comprimant le cristal : sous très haute pression, un isolant de Mott peut basculer dans l’état conducteur.

2. Dans un tout autre ordre d’idées, l’explication de la distinction isolant - conducteur donnée plus haut repose également sur une autre hypothèse importante, celle d’un réseau *rigide*. Dans certains cas, autoriser une déformation de réseau peut donner lieu à une modification spectaculaire de l’image fournie par la théorie des bandes avec un réseau fixé. L’exemple le plus connu est l’instabilité dite de Peierls.

L’idée est facile à comprendre dans le cas très simple d’un réseau unidimensionnel à un électron par maille. Si le réseau est rigide, la première zone de Brillouin s’étend entre $\pm\frac{\pi}{a}$ et est à moitié remplie : le solide est conducteur. Imaginons maintenant que l’atome initialement en na est déplacé d’une petite quantité $(-1)^n\delta a$. Autrement dit, on “dimérise” le réseau, les atomes $2n$ et $2n + 1$ étant maintenant distants de $a - 2\delta a$. Dès lors, le réseau de Bravais a une maille élémentaire de dimension $2a$ et la nouvelle première zone de Brillouin s’étend entre $\pm\frac{\pi}{2a}$. En repliant la bande d’énergie considérée⁶⁹, on voit qu’il y a une stricte dégénérescence en bord de zone ($k = \pm\frac{\pi}{2a}$) ; c’est donc à ces endroits que les *gaps* vont s’ouvrir quand le réseau se déforme. Maintenant, la bande inférieure est pleine, la bande supérieure est vide, un *gap* existe et le solide est isolant !

Il s’agit bien d’une instabilité au sens où l’énergie du gaz d’électrons a spontanément tendance à décroître dès que la dimérisation est rendue possible⁷⁰. Cet abaissement d’énergie entre en compétition avec l’énergie élastique de déformation de réseau, puisque déformer le réseau coûte de l’énergie. En faisant le bilan, on trouve que l’énergie totale présente un minimum pour une certaine valeur δa_0 , qui fixe le taux de dimérisation du réseau, la conclusion restant inchangée : dans ce cas précis, autoriser le réseau à se déformer transforme un conducteur en isolant

⁶⁹Après tout, rien n’interdit de représenter le réseau *non déformé* comme un réseau de maille double avec deux atomes par maille. On retrouvera cette façon de faire lors de l’étude des vibrations de réseau.

⁷⁰Visiblement, l’énergie de la bande pleine avec les *gaps* est plus basse que celle de la bande non déformée.